



REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE



UNIVERSITE ABOU-BEKR BELKAID - TLEMCCEN

THÈSE

Présentée à :

FACULTE DES SCIENCES – DEPARTEMENT D'INFORMATIQUE

Pour l'obtention du diplôme de :

DOCTORAT EN SCIENCES

Spécialité : Informatique

Par :

Mr Mohammed M'HAMEDI

Sur le thème

L'intelligence Artificielle au Service de l'E-santé : Applications Pour le Diagnostic des Maladies de la Peau

Soutenue publiquement le 27/02/2025 à Tlemcen devant le jury composé de :

BENAMAR Abdelkrim	Pr	Université de Tlemcen	Président
Mr HADJILA Mourad	MCA	Université de Tlemcen	Directeur de thèse
Mr MERZOUG Mohammed	MCA	Université de Tlemcen	Co-Directeur de thèse
Mr HADJILA Fethallah	MCA	Université de Tlemcen	Examineur
Mr HACHEMI Mohammed Hicham	MCA	Université des Sciences et de la Technologie d'Oran	Examineur
Mr MENEZLA Fayssal	MCA	Centre universitaire El Bayadh	Examineur

*Laboratoire de recherche en informatique de Tlemcen LRIT
BP 119, 13000 Tlemcen - Algérie*

Remerciements

Louange à Allah qui m'a donné patience et courage pour mener à bien ce travail de thèse malgré les difficultés rencontrées. Je tiens à remercier en premier lieu Mr Merzoug Mohammed et Mr Hadjila Mourad, mes directeurs de thèse pour leurs précieux conseils.

Je remercie également les membres du jury qui ont accepté l'évaluation de notre travail, et je cite en l'occurrence Pr Benamar Abdelkrim, Dr Hadjila Fethallah, Dr Hachemi Mohammed Hicham, et Dr Menezla Fayssal.

Je remercie aussi ma famille pour son soutien, son écoute et ses encouragements tout au long de cette thèse. Je tiens à remercier Dr. Imane Nedjar pour son encouragement et son soutien.

Je dédie ce modeste travail
A mes très chers parents, qu'Allah tout puissant les protège
A ma très chère femme, mon enfant, mon frère et mes soeurs
A mes collègues

Résumé

Les affections cutanées, dont les cancers de la peau, représentent un défi important pour la santé publique. L'apparence souvent trompeuse des lésions souligne l'importance d'un diagnostic précis par un dermatologue. Parmi les cancers de la peau, le mélanome malin se distingue par sa gravité et sa capacité à se métastaser rapidement. Bien qu'il soit moins fréquent que d'autres types de cancer cutané, comme le carcinome basocellulaire et le carcinome spinocellulaire, il nécessite une attention particulière en raison de son pronostic. La dermoscopie est un outil essentiel pour les dermatologues car elle permet de détecter les mélanomes à un stade précoce, ce qui est crucial pour un traitement efficace. La combinaison de la dermoscopie et de l'intelligence artificielle offre de nouvelles perspectives pour un diagnostic plus précis et plus rapide du mélanome. Les algorithmes d'apprentissage profond, entraînés sur des datasets d'images de haute qualité, peuvent aider les dermatologues à détecter les mélanomes à un stade précoce, améliorant ainsi les chances de guérison des patients. Cette thèse présente le développement d'un système d'aide au diagnostic du cancer de la peau basé sur l'apprentissage profond, visant à classifier précisément le mélanome malin. L'étude explore diverses architectures de réseaux neuronaux convolutifs (CNN), en intégrant l'apprentissage par transfert pour tirer parti des modèles pré-entraînés, l'augmentation de données pour enrichir les ensembles d'entraînement, et des architectures hybrides CNN-LSTM pour améliorer les performances du modèle. L'objectif est d'améliorer la détection précoce du mélanome grâce à des méthodes computationnelles avancées, ce qui pourrait potentiellement augmenter les taux de survie des patients.

Mots clés : Maladies de peau, Dermatologie, Cancer de la peau, Mélanome malin, CNN, Apprentissage par transfert, Augmentation des données, , Architecture hybride, LSTM.

Abstract

Skin diseases, including skin cancers, represent a major public health challenge. The often misleading appearance of lesions underlines the importance of accurate diagnosis by a dermatologist. Among skin cancers, malignant melanoma stands out for its severity and ability to metastasise rapidly. Although it is less common than other types of skin cancer, such as basal cell carcinoma and squamous cell carcinoma, it requires special attention because of its prognosis. Dermoscopy is an essential tool for dermatologists as it enables melanomas to be detected at an early stage, which is crucial for effective treatment. The combination of dermoscopy and artificial intelligence offers new prospects for more accurate and faster diagnosis of melanoma. Deep learning algorithms trained on high-quality image datasets can help dermatologists detect melanoma at an early stage, improving patients' chances of recovery. This thesis presents the development of a skin cancer diagnostic aid system based on deep learning, aimed at accurately classifying malignant melanoma. The study explores various convolutional neural network (CNN) architectures, integrating transfer learning to take advantage of pre-trained models, data augmentation to enrich the training sets, and data mining to improve the accuracy of the system. The aim is to improve early detection of melanoma using advanced computational methods, which could potentially increase patient survival rates.

Keywords: Skin diseases, Dermatology, Skin cancer, Malignant Melanoma, CNN, Transfer learning, Data augmentation, Hybrid architecture, LSTM.

الملخص:

الأمراض الجلدية، بما في ذلك سرطان الجلد، تمثل تحديًا كبيرًا للصحة العامة. غالبًا ما يؤدي المظهر الخادع للأورام إلى التأكيد على أهمية التشخيص الدقيق من قبل طبيب الأمراض الجلدية. من بين أنواع سرطان الجلد، يبرز الميلانوما الخبيث بشدة وقدرته على الانتشار بسرعة. لى الرغم من أنه أقل شيوعًا من أنواع أخرى من سرطان الجلد، مثل سرطان الخلايا القاعدية وسرطان الخلايا الحرشفية، إلا أنه يتطلب اهتمامًا خاصًا بسبب صعوبة تشخيصه تعتبر الفحص بالمجهر البصري (dermoscopy) أداة أساسية لأطباء الأمراض الجلدية، حيث تمكنهم من اكتشاف الميلانوما في مرحلة مبكرة، وهو أمر حاسم للعلاج الفعال. يمثل الجمع بين الفحص بالمجهر البصري والذكاء الاصطناعي آفاقًا جديدة لتشخيص الميلانوما بشكل أسرع وأكثر دقة. يمكن أن تساعد خوارزميات التعلم العميق المدربة على مجموعات بيانات عالية الجودة من الصور أطباء الأمراض الجلدية في اكتشاف الميلانوما في مرحلة مبكرة، مما يحسن فرص شفاء المرضى. تقدم هذه الرسالة تطوير نظام مساعد لتشخيص سرطان الجلد يستند إلى التعلم العميق، بهدف تصنيف الميلانوما الخبيث بدقة. تستكشف هذه الدراسة مختلف هياكل الشبكات العصبونية التلافيفية (CNN) ، مع دمج التعلم الانتقالي للاستفادة من النماذج المدربة مسبقًا، وزيادة البيانات لتخصيص مجموعات التدريب، واستخراج البيانات لتحسين دقة النظام. يهدف هذا البحث إلى تحسين الكشف المبكر عن الميلانوما باستخدام طرق الحوسبة المتقدمة، مما يمكن أن يزيد من معدلات بقاء المرضى.

الكلمات المفتاحية: الأمراض الجلدية، الأمراض الجلدية (الجلدية)، سرطان الجلد، الورم الميلانيني الخبيث، الشبكة العصبية الالتلافيفية (CNN) ، التعلم بالنقل، تعزيز البيانات، العمارة الهجينة، الشبكة العصبية التكرارية طويلة المدى (LSTM).

Table des matières

Remerciements	i
Résumé	iii
Sommaire	vi
Table des figures	ix
Liste des tableaux	xi
Introduction générale	1
1 Intelligence artificielle, Machine Learning, et Deep Learning	6
1.1 Introduction	7
1.2 Histoire de l'intelligence artificielle	7
1.3 Sous-domaines de l'intelligence artificielle	8
1.4 Apprentissage automatique	9
1.5 Apprentissage profond (Deep Learning)	9
1.5.1 Avantages de l'apprentissage profond	10
1.5.2 Réseaux de neurones artificiels (ANN)	11
1.5.2.1 Les composants des Réseaux Neuronaux Artificiels	13
1.5.2.2 Types de réseaux de neurones artificiels	17
1.6 Réseau de neurones convolutif (CNN)	20
1.6.1 Architecture de réseau CNN	20
1.6.1.1 L'image d'entrée	21
1.6.1.2 La couche de convolution	21
1.6.1.3 La couche de Pooling	22
1.6.1.4 La couche Flattening	23
1.6.2 Exemple d'implémentation d'un CNN	23
1.6.3 L'apprentissage par transfert et modèles pré-entraînés	25
1.6.3.1 Définition du modèle pré-entraîné	26
1.6.3.2 L'extraction de caractéristiques	26
1.6.3.3 L'ajustement fin	28
1.6.3.4 Modèles pré-entraînés populaires	29
1.7 Préparation des données	30
1.7.1 Chargement et prétraitement des données	30
1.7.2 Normalisation des données	31

1.7.3	Augmentation des données	31
1.7.4	Division des données	32
1.8	Métriques d'évaluation	33
1.9	Implémentation d'un réseau CNN	35
1.9.1	Identifier une image par le modèle pré-entraîné VGG	35
1.9.2	L'ajustement fin sur le modèle VGG16	37
1.9.3	L'ajustement fin sur VGG avec des couches ANN personnalisées	41
1.10	Conclusion	43
2	Background et généralités sur le cancer de la peau : méthodes de diagnostic et travaux existants sur le diagnostic du mélanome par Deep Learning	44
2.1	Introduction	45
2.2	Présentation des lésions cutanées	45
2.3	Cancer de la peau	47
2.3.1	Cancer de la peau non- mélanome	47
2.3.1.1	Facteurs de risque du cancer de la peau non- mélanome	48
2.3.1.2	Diagnostic des cancers de la peau non- mélanome	48
2.3.2	Cancer de la peau mélanome malin	50
2.3.2.1	Types de mélanome malin cutané	50
2.3.2.1.1	Mélanome à propagation superficielle	50
2.3.2.1.2	Mélanome nodulaire	50
2.3.2.1.3	Mélanome lentigo maligne	52
2.3.2.1.4	Mélanome acral-lentigineux	52
2.3.2.2	Etiologie (étude des causes)	53
2.3.2.3	Les symptômes	53
2.3.2.4	Prévention du mélanome malin	54
2.3.2.5	Méthodes de diagnostic du mélanome malin	54
2.3.2.5.1	La règle ABCDE	55
2.3.2.5.2	La liste de contrôle en 3 points	55
2.3.2.5.3	La liste de contrôle en 7 points	55
2.3.2.5.4	La méthode de Menzies	56
2.3.2.5.5	L'analyse des formes	56
2.3.3	Techniques d'imagerie du cancer de la peau	56
2.3.3.1	La microscopie confocale par réflectance	57
2.3.3.2	La Tomographie en Cohérence Optique	57
2.3.3.3	L'échographie	57
2.3.3.4	La dermoscopie	57
2.4	Utilisation de l'intelligence artificielle pour le diagnostic du mélanome malin	58
2.4.1	Les ensembles de données (datasets)	59
2.4.2	Les travaux existants : Méthodes de Deep Learning appliquées au diagnostic du mélanome	60
2.5	Conclusion	64
3	Identification du mélanome malin dans les images de lésions cutanées à l'aide de l'apprentissage par transfert	65
3.1	Introduction	65
3.2	Présentation du dataset utilisé	66

3.2.1	Description des métadonnées	66
3.2.2	Exploration des caractéristiques	68
3.2.3	Caractéristiques de la cible (sortie)	68
3.2.4	Caractéristiques du sexe	69
3.2.4.1	Caractéristiques de l'âge approximatif	70
3.2.5	Caractéristiques de la position anatomique	71
3.2.6	Caractéristiques du diagnostic	72
3.2.7	Caractéristiques du nom d'image et l'identificateur du patient	73
3.3	Prétraitement des données	74
3.3.1	Élimination des noms d'images redondantes	74
3.3.2	Division du dataset	75
3.4	Architecture du modèle de classification utilisant l'apprentissage par transfert	76
3.5	Résultats et discussion	78
3.5.1	Résultats d'entraînement et de validation	78
3.5.2	Test et évaluation	80
3.6	Conclusion	82
4	Amélioration de la classification du mélanome dans les images de lésions cutanées	83
4.1	Introduction	83
4.2	Approche d'augmentation des données	84
4.2.1	Prétraitement des données	86
4.2.1.1	Génération de nouvelles données par la technique d'augmentation des données	86
4.2.1.2	Division du dataset	90
4.2.2	Résultats et discussion	91
4.2.2.1	Résultats d'entraînement et de validation	91
4.2.2.2	Test et évaluation	93
4.3	Approche hybride CNN et LSTM basée sur l'apprentissage par transfert	95
4.3.1	Architecture du réseau hybride CNN et LSTM	95
4.3.2	Résultats et discussion	97
4.3.2.1	Résultats d'entraînement et de validation	97
4.3.2.2	Test et évaluation	100
4.4	Évaluation comparative des performances : État de l'art vs. Nos approches proposées	101
4.5	Conclusion	102
	Conclusion générale	104
	Liste des publications	107
	Bibliographie	108

Table des figures

1	v
1.1	Vue d'ensemble : ML vs DL vs AI.	10
1.2	Impact de la quantité des données sur la performance.	10
1.3	Neurone biologique.	11
1.4	Neurone artificiel.	12
1.5	Réseaux de couches.	12
1.6	CNN, exemple de classification de véhicules.	20
1.7	Architecture et couches du réseau de neurones convolutif CNN.	21
1.8	Matrice pour image en lettre L.	21
1.9	Détecteur de caractéristiques et la carte de caractéristiques.	22
1.10	Détecteurs de caractéristiques et de carte multiples.	22
1.11	Max pooling.	23
1.12	Procédure de Flattening.	25
1.13	Modèles pré-entraînés : l'extraction de caractéristiques et l'ajustement fin.	27
1.14	Architecture divisée CNN – base convolutive et classificateur.	27
1.15	Couche de base convolutive réutilisable.	28
1.16	Division des données training set, validation set, et test set.	33
1.17	Image de pierre à identifier par VGG16.	36
1.18	Affichage du résumé du modèle VGG16 après son lancement (les trois dernières couches).	38
1.19	Résumé du modèle créé après l'ajout de la nouvelle couche.	39
1.20	Accuracy et perte de l'entraînement et validation du modèle VGG16.	40
1.21	Résumé du modèle après l'ajout des couches ANN personnalisées.	42
1.22	Accuracy et perte de l'entraînement et validation du modèle VGG16 personnalisé.	43
2.1	Type de lésion cutanée (a) Les lésions primaires, et (b) Les lésions secondaires.	46
2.2	Facteurs influençant les conséquences d'une maladie.	49
2.3	Mélanome malin à propagation superficielle.	51
2.4	Mélanome malin nodulaire.	51
2.5	Mélanome lentigo maligne.	52
2.6	Mélanome acral-lentigineux.	53
2.7	Images dermoscopiques : formes et couleurs irrégulières.	58
3.1	Visualisation des caractéristiques de la cible.	68
3.2	Visualisation des caractéristiques du sexe.	69

3.3	Distribution de sexe par rapport au, (a) : maligne, (b) : bénigne. . . .	70
3.4	Distribution approximative de l'âge moyen par sexe.	70
3.5	Distribution approximative de l'âge moyen par cible (sortie).	71
3.6	Nombre d'images par rapport à la position anatomique.	72
3.7	Nombre d'images par rapport à diagnostic.	73
3.8	Distribution de l'image_name par patient_id (uniquement les nombre > 75).	73
3.9	Préparation des données : suppression des images en double.	75
3.10	Modèle proposé pour la prédiction du mélanome.	76
3.11	Description de l'architecture du réseau (a) : modèle basé sur VGG-19, (b) : Modèle basé sur MobilerNetV2 et (c) : Les couches ajoutées. . .	77
3.12	Accuracy/Perte de modèle basé sur VGG-19 vs. époque.	79
3.13	Accuracy/Perte de modèle basé sur MobileNetV2 vs. époque.	79
3.14	Matrice de confusion du, (a) : modèle basé sur VGG-19 et (b) : du modèle basé sur MobileNetV2.	81
4.1	Préparation des données : Approche d'augmentation des données. . .	87
4.2	Processus d'augmentation d'images.	89
4.3	Images du générateur de données a) image originale. (b) 9 images aug- mentées.	90
4.4	Accuracy/Perte de modèle basé sur VGG-19 vs. époque.	91
4.5	Accuracy/Perte de modèle basé sur MobileNetV2 vs. époque.	92
4.6	: Métriques de performance pour les deux modèles : (a) l'entraînement et (b) validation.	93
4.7	Matrice de confusion du modèle basé sur, (a) : VGG-19 et (b) : Mobi- leNetV2.	94
4.8	Architecture du réseau hybride CNN pré-entraîné et LSTM.	96
4.9	Accuracy/Perte de modèle basé sur VGG-19-LSTM Vs. époque. . . .	98
4.10	Accuracy/Perte de modèle basé sur MobilNetV2-LSTM Vs. époque. .	98
4.11	Métriques de performance pour les deux modèles : (a) l'entraînement et (b) validation.	99
4.12	Matrice de confusion du modèle basé sur, (a) : VGG19-LSTM et (b) : MobileNetV2-LSTM.	101

Liste des tableaux

1.1	Les fonctions d'activation	14
1.2	Fonctions de coût ou de perte	15
1.3	Fonctions d'optimisation	16
1.4	Types de réseaux de neurones, leurs architectures, algorithmes, applications et illustrations	19
1.5	Composants d'un Réseau de Neurones Convolutif (CNN) et leurs descriptions	24
1.6	Les modèles pré-entraînés en CNN	30
1.7	Les principales métriques de classification [34], [35]	34
1.8	Matrice de confusion.	35
2.1	Liste de contrôle en sept points pour le mélanome malin.	54
2.2	Critères de la règle ABCDE pour le calcul du score dermoscopique total (TDS).	55
2.3	Interprétation de score dermoscopique total (TDS).	56
2.4	Résumé des travaux existants.	63
3.1	Statistiques sur les mélanomes et non-mélanomes	67
3.2	Division de l'ensemble de donnée par classe.	76
3.3	Performance des modèles basés sur VGG19 et MobileNetV2.	81
4.1	Description des paramètres de transformations.	85
4.2	Paramètres et valeurs utilisés pour l'augmentation des données.	86
4.3	Division de l'ensemble de donnée par classe.	90
4.4	Performance des modèles basés sur VGG19, et MobileNetV2.	94
4.5	Performance des modèles basé sur VGG19-LSTM, et MobileNetV2-LSTM.	101
4.6	Performance des approches proposées par rapport aux techniques de références.	102

Introduction générale

Les maladies de la peau représentent une part importante de l'ensemble des maladies, affectant des millions de personnes dans le monde. La dermatologie est la spécialité médicale responsable de l'étude de plus de 4 000 maladies de la peau et des lésions cutanées, représentant 15 à 30 % des soins médicaux ambulatoires dans les systèmes de santé, incorporant un large ensemble de ressources diagnostiques, thérapeutiques et esthétiques [1]. Les maladies de la peau sont connues de l'humanité depuis son origine, la composante essentiellement visuelle de ces affections ayant permis leur reconnaissance précoce. Les premières classifications cutanées remontent à une époque très ancienne, où elles ont été décrites par les grandes civilisations qui ont façonné la médecine occidentale. L'étude des maladies de la peau est liée à la médecine générale depuis des siècles. Ce n'est qu'au 18ème siècle, sous l'impulsion des progrès de la science des domaines de la recherche, qu'apparaissent les premiers ouvrages consacrés spécifiquement à l'étude des maladies de la peau. Tout au long des 19ème et 20ème siècles, la révolution scientifique et les innovations technologiques ont transformé la dermatologie, en améliorant les techniques de diagnostic et en apportant de nouvelles ressources thérapeutiques. Au 21ème siècle, les maladies de la peau sont devenues un domaine d'intérêt majeur pour la recherche médicale et la santé publique en raison de leur prévalence croissante et de l'impact des facteurs liés à l'environnement et au mode de vie. L'augmentation de maladies telles que l'eczéma, le psoriasis et le cancer de la peau a entraîné des progrès dans les traitements et les technologies dermatologiques, y compris l'utilisation de l'intelligence artificielle pour la détection précoce et la médecine personnalisée [2].

Les types de lésions cutanées les plus fréquents dans la pratique des maladies infectieuses sont les infections de la peau et des tissus mous causées par des bactéries, des virus, des champignons et des parasites, ainsi que différents types d'exanthèmes accompagnés de fièvre. Les infections de la peau et des tissus mous comptent parmi les types d'infections les plus répandus et comprennent la cellulite, l'érysipèle, l'impétigo, l'ecthyma, la folliculite, les furoncles, les escarres, les abcès, les infections liées aux traumatismes, les ulcères décubitaux, la fasciite nécrosante, la gangrène de Fournier et les infections dues à des morsures humaines ou animales. Leur présentation, leur étiologie et leur gravité sont variables.

Dans certains cas, les lésions cutanées sont un symptôme important d'une maladie systémique ou localisée et permettent de diagnostiquer certaines infections ou affections non infectieuses, mais diverses maladies infectieuses et non infectieuses peuvent produire des lésions cutanées similaires, et une même affection peut présenter différents types de lésions cutanées en raison d'un nombre limité de manifestations

cutanées [3]. Certaines lésions cutanées, telles que l'herpès zoster, le psoriasis et la dermatite atopique, peuvent être correctement reconnues même sans grande expérience, mais beaucoup d'autres sont plus exigeantes.

Les maladies de peau sont multiples et variées. Si certaines sont bénignes et facilement traitables, d'autres peuvent être plus complexes et nécessitent un suivi médical régulier. Le cancer est loin d'être la seule manifestation des maladies de la peau que les gens doivent connaître. L'institut national de l'arthrite et des maladies musculo-squelettiques et cutanées (National Institute of Arthritis and Musculoskeletal and Skin Diseases) identifie 13 troubles cutanés bénins importants [4]. Bien qu'elles ne soient pas intrinsèquement cancéreuses, ces affections majeures ont un impact négatif sur la qualité de vie des individus et peuvent même se transformer en affections potentiellement mortelles si elles sont infectées ou négligées.

Parmi les affections cutanées les plus fréquentes, on retrouve l'acné juvénile, l'eczéma atopique, le psoriasis en plaques, la rosacée, les verrues vulgaires, l'herpès labial récurrent et les différents types de cancers de la peau. De nombreuses affections cutanées répondent bien aux traitements et peuvent être contrôlées par des adaptations du mode de vie. Une bonne hygiène de vie, la connaissance des facteurs déclenchants et des consultations régulières chez un dermatologue sont essentiels pour prévenir les récurrences et améliorer le confort au quotidien. Parmi les maladies de la peau, les cancers, cette catégorie englobe des formes graves notamment le carcinome basocellulaire, le carcinome spinocellulaire, et le mélanome. Bien qu'ils soient moins fréquents que les autres affections cutanées mentionnées, ces cancers peuvent entraîner des complications sérieuses s'ils ne sont pas dépistés et traités de manière précoce. Le carcinome se développe souvent après l'âge de cinquante ans et est facilement guérissable dans la majorité des cas. Les carcinomes basocellulaires ne donnent jamais de métastases, mais les carcinomes spinocellulaires en donnent, principalement dans les ganglions lymphatiques autour de la tumeur. Les cancers passent souvent par différents stades. Les carcinomes spinocellulaires peuvent commencer par une lésion localisée dans l'épiderme. Une croûte (kératose actinique) ou un eczéma (maladie de Bowen) apparaît à la surface de la peau. L'infiltration plus profonde du derme caractérise le stade du carcinome invasif [5].

Le mélanome malin, un type de cancer de la peau, est l'une des maladies de la peau les plus dangereuses. Bien que moins fréquent que d'autres cancers de la peau tels que le carcinome basocellulaire et le carcinome spinocellulaire, le mélanome est particulièrement dangereux en raison de sa capacité à se propager (métastases) rapidement à d'autres parties du corps. Lorsque le mélanome se propage au-delà de la peau et atteint des organes tels que les poumons, le foie ou le cerveau, il devient beaucoup plus difficile à traiter et peut mettre la vie en danger [6].

Les mélanomes, le type de cancer de la peau le plus mortel, se développent à partir des cellules mélanocytaires, qui produisent la mélanine, responsable de la coloration brune ou rouge de la peau. Il existe deux types de pigmentation de la peau : les pigments bruns, qui procurent un bronzage et une certaine protection contre les UV, et les pigments rouges (peau rouge), qui ne le font pas. Les sujets qui génèrent principalement des pigments rouges ne bronzent pas et ont donc un risque plus élevé de développer un cancer de la peau. Les personnes qui vivent dans les régions montagneuses sont plus susceptibles de développer un cancer de la peau en raison de leur exposition au soleil et de la nature de la couleur de leur peau.

Cette thèse explore la maladie de la peau la plus dangereuse, qui est le mélanome malin, une forme très agressive de cancer de la peau. Diagnostiquer le mélanome

malin à un stade précoce peut être difficile, car ses caractéristiques cliniques sont souvent similaires à celles d'un nævus atypique. S'il est identifié suffisamment tôt, le mélanome peut être traité avec succès par une simple intervention chirurgicale, ce qui améliore considérablement les chances de survie. Avec le traitement, les patients à ce stade ont un taux de survie global à cinq et dix ans de plus de 90% [7].

Un dermatologue peut confirmer ou infirmer un diagnostic de cancer de la peau précisément le mélanome. L'aspect de la lésion est souvent suffisant, mais dans toutes les situations douteuses, un examen microscopique sera nécessaire. Le schéma clinique ABCD (asymétrie, irrégularité des bords, variation de couleur et diamètre supérieur) est l'une des techniques les plus fréquemment utilisées [8]. Malheureusement, certaines formes de mélanome malin peuvent ne pas être détectées par la règle clinique ABCD, car le problème de la détection des mélanomes est que leurs caractéristiques se confondent avec celles des grains de beauté bénins. C'est pourquoi la dermoscopie est une méthode plus efficace pour identifier les mélanomes malins. En effet, la dermoscopie est un outil précieux pour l'évaluation des lésions cutanées, offrant une précision accrue qui améliore considérablement le diagnostic du mélanome et est donc largement privilégiée par les dermatologues [9].

Dans la science avancée, c'est un péché d'utiliser les méthodes ou approches traditionnelles pour guérir ou atténuer la gravité du cancer. Grâce aux progrès de la technologie clinique dans le monde médical, les médecins et les experts médicaux utilisent divers équipements médicaux pour diagnostiquer rapidement le cancer de la peau et en arrêter la gravité.

Une autre technique susceptible d'améliorer l'efficacité du diagnostic du mélanome malin est l'utilisation de l'intelligence artificielle (IA). Le développement de l'AI est motivé par l'objectif d'améliorer la santé publique, d'améliorer l'accessibilité aux soins de santé, de réduire les coûts des soins de santé et de faire progresser les connaissances médicales, tout en utilisant la puissance de la technologie et de l'innovation. Une étape importante qui a marqué l'ère de l'intelligence artificielle moderne en dermatologie fut la démonstration des capacités de classification des cancers de la peau par les réseaux neuronaux convolutionnels (CNN) d'apprentissage profond (DL), qui étaient comparables aux performances des dermatologues certifiés.

Ces dernières années, l'apprentissage profond (DL) s'est imposé comme un outil puissant dans le domaine de la détection des cancers. Les modèles de DL, tels que les réseaux neuronaux convolutifs (CNN) et les réseaux neuronaux récurrents (RNN), sont capables d'apprendre des modèles complexes à partir de grandes quantités de données et d'identifier automatiquement les tumeurs potentielles, sans avoir besoin de règles explicites. Bien que les CNN peuvent diagnostiquer des maladies le cancer de la peau avec une précision similaire à celle des dermatologues, mais ils nécessitent un grand nombre de photographies cliniques pour leur entraînement. Par conséquent, des ensembles de données (dataset) tels que HAM10000, BCN20000, ISCI2018, et ISIC 2020 qui contiennent différentes images dermoscopiques du cancer de la peau ont été créés. Ces datasets sont en effet des références incontournables pour la recherche en classification de lésions cutanées, notamment les mélanomes.

Dans cette thèse, nous proposons un système d'aide au diagnostic pour les cancers de la peau basé sur l'apprentissage profond. Notre objectif est de développer un modèle robuste et généralisable capable de classer de manière précise et différenciée le mélanome malin. Pour ce faire, nous explorerons différentes architectures de CNN, en mettant l'accent sur les techniques d'apprentissage par transfert et d'augmentation de données et les architectures hybrides CNN-LSTM, afin d'améliorer les performances

globales et de détecter le cancer à un stade précoce. Ce projet de recherche pourrait avoir un impact significatif sur la détection précoce du mélanome et, par conséquent, sur le taux de survie des patients.

L'entraînement du modèle CNN est principalement basé sur l'ensemble de données à partir duquel il apprend. La qualité, la diversité et la taille de l'ensemble de données resteront un facteur majeur de la performance et de la réalisabilité des modèles d'apprentissage automatique. Les limites de ces modèles sont souvent dues à la qualité des données utilisées pour leur entraînement. Il est essentiel que les classes soient bien équilibrées et que l'ensemble de données soit suffisamment diversifié pour permettre au modèle d'apprendre efficacement ces caractéristiques. Ces limites du déséquilibre des classes sont de plus en plus mises en évidence dans le domaine de l'apprentissage statistique pour le domaine de la santé, et il est nécessaire de proposer de bons classificateurs face à ces problèmes. Le déséquilibre des classes a été la principale cause des anomalies dans l'ensemble de données sur le cancer de la peau en particulier le mélanome, car les patients infectés par un mélanome avaient un nombre d'images considérablement inférieur à celui des patients présentant des lésions cutanées non cancéreuses.

La solution que nous proposons pour résoudre ce problème consiste à augmenter l'ensemble de données avec de nouvelles images dérivées de l'original. Nous utilisons une technique d'augmentation d'image avec ces paramètres de transformation tels que, inversions, rotations, zooms et cisaillements. Nos principales contributions peuvent se résumer ainsi :

- Nous proposons deux modèles CNN basés sur l'apprentissage par transfert, utilisant les architectures pré-entraînées VGG-19 et MobileNetV2.
- Le problème des classes déséquilibrées est traité par la technique d'augmentation des données.
- Nous optimisons nos résultats en combinant les forces des réseaux de neurones convolutionnels (CNN) et récurrents (LSTM) au sein d'un modèle hybride.
- Les deux modèles CNN sont évalués à l'aide de l'ensemble de données ISIC 2020.

Le manuscrit de la thèse est structuré en quatre chapitres. Dans le premier chapitre intitulé Introduction à l'intelligence artificielle (IA) et aux modèles CNN pré-entraînés, les concepts fondamentaux de l'intelligence artificielle et de l'apprentissage automatique sont explorés pour poser les bases de la compréhension de ces technologies. Ce chapitre introduit ensuite les réseaux de neurones convolutifs (CNN) et leur rôle essentiel en vision par ordinateur, une sous-discipline de l'IA dédiée à l'analyse et à l'interprétation des données visuelles. L'accent est mis sur les techniques d'apprentissage par transfert, qui permettent de réutiliser des modèles CNN pré-entraînés pour accomplir des tâches spécifiques, réduisant ainsi le besoin de données massives pour entraîner des modèles efficaces. Enfin, une revue des principaux modèles pré-entraînés, tels que VGG19 et MobileNetV2, est présentée, soulignant leur pertinence particulière pour la classification d'images médicales, où la précision et la fiabilité sont cruciales.

Dans le second chapitre intitulé Maladies cutanées et classification du mélanome, une description approfondie des maladies cutanées courantes et des troubles dermatologiques est proposée pour contextualiser l'importance de la détection et du diagnostic précoces. Ce chapitre se concentre ensuite sur le cancer de la peau, avec une attention particulière au mélanome, en examinant ses caractéristiques cliniques

et pathologiques qui le distinguent des autres affections cutanées. Enfin, une revue des recherches antérieures sur la classification du mélanome à l'aide des techniques d'intelligence artificielle (IA) et de réseaux de neurones convolutifs (CNN) est effectuée, soulignant les avancées significatives et les défis persistants dans l'application de ces technologies pour améliorer le diagnostic et le traitement de cette forme de cancer.

Le troisième chapitre intitulé Classification et détection du mélanome avec CNN et apprentissage par transfert explore en détail la méthodologie employée pour appliquer les réseaux de neurones convolutifs (CNN) à la classification du mélanome. Ce chapitre présente les approches spécifiques basées sur les architectures VGG-19 et MobileNetV2, mettant en lumière leur utilisation pour la détection du mélanome. Une évaluation rigoureuse des performances de ces modèles est réalisée, en se concentrant sur des métriques clés telles que l'accuracy, ainsi que d'autres indicateurs de performance. Enfin, une discussion approfondie des résultats obtenus est menée, abordant les limites et les défis rencontrés lors de l'utilisation de ces architectures pour la classification du mélanome, tout en suggérant des pistes d'amélioration.

Dans le quatrième chapitre intitulé Amélioration des performances par augmentation de données et réseaux hybrides CNN-LSTM, les techniques d'augmentation des données sont présentées comme un moyen essentiel pour améliorer la robustesse et la généralisation des modèles de classification du mélanome. Le chapitre introduit ensuite les architectures hybrides CNN-LSTM, combinant les capacités de traitement spatial des CNN avec la gestion temporelle des LSTM (Long Short-Term Memory), en mettant en œuvre des modèles tels que VGG-19-LSTM et MobileNetV2-LSTM pour la détection du mélanome. Une évaluation comparative des performances de ces modèles hybrides est réalisée par rapport aux modèles CNN classiques et aux méthodes de pointe actuelles, offrant une analyse des gains en termes de précision et de fiabilité, ainsi que des perspectives sur l'amélioration continue des techniques de classification du mélanome.

Finalement, nous couronnons ce manuscrit par une conclusion générale et des perspectives.

Intelligence artificielle, Machine Learning, et Deep Learning

Sommaire

1.1	Introduction	7
1.2	Histoire de l'intelligence artificielle	7
1.3	Sous-domaines de l'intelligence artificielle	8
1.4	Apprentissage automatique	9
1.5	Apprentissage profond (Deep Learning)	9
1.5.1	Avantages de l'apprentissage profond	10
1.5.2	Réseaux de neurones artificiels (ANN)	11
1.5.2.1	Les composants des Réseaux Neuronaux Artificiels	13
1.5.2.2	Types de réseaux de neurones artificiels	17
1.6	Réseau de neurones convolutif (CNN)	20
1.6.1	Architecture de réseau CNN	20
1.6.1.1	L'image d'entrée	21
1.6.1.2	La couche de convolution	21
1.6.1.3	La couche de Pooling	22
1.6.1.4	La couche Flattening	23
1.6.2	Exemple d'implémentation d'un CNN	23
1.6.3	L'apprentissage par transfert et modèles pré-entraînés	25
1.6.3.1	Définition du modèle pré-entraîné	26
1.6.3.2	L'extraction de caractéristiques	26
1.6.3.3	L'ajustement fin	28
1.6.3.4	Modèles pré-entraînés populaires	29
1.7	Préparation des données	30
1.7.1	Chargement et prétraitement des données	30
1.7.2	Normalisation des données	31
1.7.3	Augmentation des données	31
1.7.4	Division des données	32
1.8	Métriques d'évaluation	33
1.9	Implémentation d'un réseau CNN	35
1.9.1	Identifier une image par le modèle pré-entraîné VGG	35
1.9.2	L'ajustement fin sur le modèle VGG16	37
1.9.3	L'ajustement fin sur VGG avec des couches ANN personnalisées	41
1.10	Conclusion	43

1.1 Introduction

L'intelligence artificielle (IA) est un domaine de l'informatique qui se concentre sur la création de systèmes capables de réaliser des tâches qui nécessitent normalement l'intelligence humaine. L'émergence de l'Intelligence Artificielle remonte aux années 1950, lorsque les premiers programmes ont été créés pour permettre aux ordinateurs d'exécuter des tâches intelligentes. Depuis lors, cette technologie a été continuellement explorée et repoussée par des chercheurs et des innovateurs passionnés. De nos jours, l'IA est largement reconnue comme l'une des avancées majeures de la quatrième révolution industrielle et est perçue comme le pilier de l'avenir de notre monde [10]. L'impact de l'IA est profond et touche de nombreux aspects de notre vie quotidienne. Grâce à ses capacités, elle a considérablement simplifié nos activités et processus, en offrant des solutions efficaces dans une multitude de domaines. Afin de rendre l'intelligence artificielle plus accessible et performante, une combinaison de techniques avancées est utilisée, dont l'apprentissage automatique (Machine Learning) et l'apprentissage profond (Deep Learning). L'objectif ultime de l'intelligence artificielle est de reproduire et d'égaliser les capacités de raisonnement et de logique humaines. Cependant, il reste encore des défis à relever dans cette quête, et les chercheurs continuent de travailler sur de nouvelles avancées pour améliorer les performances et les capacités de l'IA [10]. Les réseaux neuronaux étant couramment utilisés dans les méthodes d'apprentissage profond, l'expression « Deep Learning » s'est imposée. L'un des réseaux neuronaux profonds les plus couramment utilisés est le réseau neuronal convolutif (CNN). Contrairement aux CNN, les techniques traditionnelles d'extraction de caractéristiques, telles que SIFT et LBP, nécessitent une intervention humaine pour leur extraction. Les CNN ont extrait les caractéristiques directement à partir d'un ensemble de données ou d'images brutes (dataset). Dans ce chapitre, des concepts de base importants pour la compréhension de cette thèse seront expliqués. Dans un premier temps, nous allons présenter les aspects fondamentaux de l'intelligence artificielle. Ensuite, des termes plus généraux relatifs à l'apprentissage automatique seront introduits, suivis de concepts et de définitions spécifiques au sous-ensemble de défis d'apprentissage automatique rencontrés dans le cadre de cette étude, nous approfondirons sur l'apprentissage profond, dans lequel nous détaillerons particulièrement les réseaux CNN.

1.2 Histoire de l'intelligence artificielle

L'intelligence artificielle existe depuis longtemps et au fil des années, elle a connu une évolution constante, réalisant des avancées qui semblaient auparavant inimaginables. L'histoire de l'IA remonte à 1943 lorsque Warren McCulloch et Walter Pitts ont créé le premier modèle mathématique pour un réseau de neurones.

En 1952, Arthur Samuel a développé un logiciel capable d'apprendre à jouer aux échecs de manière autonome. La première conférence consacrée à l'intelligence artificielle a eu lieu en 1956, dirigée par John McCarthy.

En 1959, Arthur Samuel a introduit le terme "Machine Learning" (ML) pour décrire la capacité des machines à apprendre. Au cours des années 1970, le gouvernement britannique a publié le rapport "Lighthill", qui exprimait les déceptions de la recherche en IA. Cela a été suivi d'une période de doute qui a persisté jusqu'en 1980.

Dans les Années 1990, le machine Learning monte en puissance, avec des algo-

algorithmes comme les machines à vecteurs de support (SVM) et les avancées en reconnaissance de la parole et de l'image.

En 1997, l'IA a marqué l'histoire en battant l'homme lors d'une partie d'échecs, avec la victoire de DeepBlue contre le champion du monde Garry Kasparov.

En 2006 Geoffrey Hinton et son équipe redécouvrent le potentiel des réseaux de neurones profonds, menant à l'explosion du deep learning. En 2010, l'IA connaît une renaissance grâce à la disponibilité de grandes quantités de données (Big Data) et à la puissance de calcul des GPU, permettant des avancées majeures en reconnaissance d'image, traitement du langage naturel, et jeux (ex. AlphaGo de DeepMind en 2016).

En 2016, l'IA a remporté une nouvelle victoire contre l'homme, avec la victoire du système AlphaGo de Google DeepMind au jeu de go [11].

Aujourd'hui, les technologies de l'IA y compris l'apprentissage automatique (Machine Learning) et l'apprentissage profond (Deep Learning) sont intégrées dans de nombreux domaines tels que la santé, les finances, les transports, et le divertissement [12].

1.3 Sous-domaines de l'intelligence artificielle

Les sous-domaines de l'intelligence artificielle offrent des avantages et des perspectives uniques, chacun se concentrant sur des aspects différents du développement et du déploiement de l'IA. En se spécialisant dans ces domaines, les chercheurs et les praticiens peuvent améliorer les capacités des systèmes d'IA et résoudre des problèmes complexes plus efficacement. Parmi ces sous-domaines figurent : l'apprentissage automatique (ML), l'apprentissage profond (DL), le traitement du langage naturel (NLP), la vision par ordinateur, et la robotique.

- Apprentissage automatique (ML) : Le ML est une branche de l'intelligence artificielle et de l'informatique qui se concentre sur l'utilisation de données et d'algorithmes pour permettre à l'IA d'imiter la façon dont les humains apprennent, en améliorant progressivement sa précision.
- Apprentissage profond (DL) : Le DL est une méthode d'intelligence artificielle qui apprend aux ordinateurs à traiter les données en s'inspirant du cerveau humain.
- Traitement du langage naturel (NLP) : Le NLP vise à permettre aux ordinateurs de comprendre, d'interpréter et de générer du langage humain. Il comprend des tâches telles que la reconnaissance vocale, la traduction linguistique, l'analyse de texte, l'analyse des sentiments et la réponse aux questions.
- Vision par ordinateur : La Vision par ordinateur vise à permettre aux ordinateurs d'identifier et de comprendre les objets et les personnes dans les images et les vidéos. Elle comprend des tâches telles que la reconnaissance d'objets, la segmentation d'images, la reconnaissance faciale et la compréhension de scènes.
- Robotique : La robotique combine l'IA avec des systèmes physiques pour créer des machines intelligentes capables d'interagir avec le monde physique et de le manipuler. Elle comprend des domaines tels que la perception robotique, la planification des mouvements, les systèmes de contrôle et l'interaction homme-robot.

1.4 Apprentissage automatique

L'apprentissage automatique (ML) permet aux ordinateurs d'apprendre à partir de données sans être explicitement programmés. Il implique l'utilisation de modèles et de techniques statistiques qui permettent à l'ordinateur d'apprendre des motifs à partir des données, puis d'appliquer cet apprentissage pour faire des prédictions, en utilisant de grandes quantités de données pour apprendre et s'adapter. L'apprentissage automatique est un domaine diversifié et en évolution rapide qui emploie une variété de méthodes différentes pour atteindre ses objectifs. Cela se fait à l'aide d'algorithmes qui peuvent identifier des modèles et faire des prédictions sur de nouvelles données. Le ML est utilisé dans une vaste gamme d'applications. Les principales catégories et techniques de l'apprentissage automatique sont [13] :

- Apprentissage supervisé : Les modèles sont entraînés sur un jeu de données étiqueté, c'est-à-dire que chaque entrée d'apprentissage est associée à une sortie correcte. Ce type d'apprentissage utilise des algorithmes comme la régression linéaire, les arbres de décision, les machines à vecteurs de support (SVM), et les réseaux de neurones.
- Apprentissage non supervisé : Les modèles tentent de trouver des structures ou des modèles cachés dans des données non étiquetées. Ce type d'apprentissage utilise des algorithmes comme, k-means, algorithme des k-plus proches voisins (k-NN), analyse en composantes principales (PCA), autoencodeurs.
- Apprentissage par renforcement : Dans l'apprentissage par renforcement, les agents apprennent à prendre des décisions en interagissant avec un environnement et en recevant soit des récompenses soit des pénalités. Dans cette technique, on peut utiliser des algorithmes comme le Q-learning, les politiques de Monte Carlo, le Deep Q-Network (DQN), les méthodes de politique d'acteur-critiques.
- Apprentissage semi-supervisé : Cette technique est une combinaison d'apprentissage supervisé et non supervisé, où une petite quantité de données étiquetées est utilisée avec une grande quantité de données non étiquetées, des algorithmes comme Transductive SVM, et graphes de consistance peuvent être employés.

1.5 Apprentissage profond (Deep Learning)

L'apprentissage profond (DL) est un type du ML (voir figure 1.1) qui utilise des réseaux neuronaux artificiels, qui sont inspirés du cerveau humain. Les réseaux neuronaux peuvent apprendre des modèles complexes à partir de grandes quantités de données et sont utilisés dans des applications de pointe telles que la conduite autonome et la reconnaissance faciale. L'apprentissage profond est basé sur les réseaux neuronaux artificiels, qui sont modélisés d'après la structure et la fonction du cerveau humain.

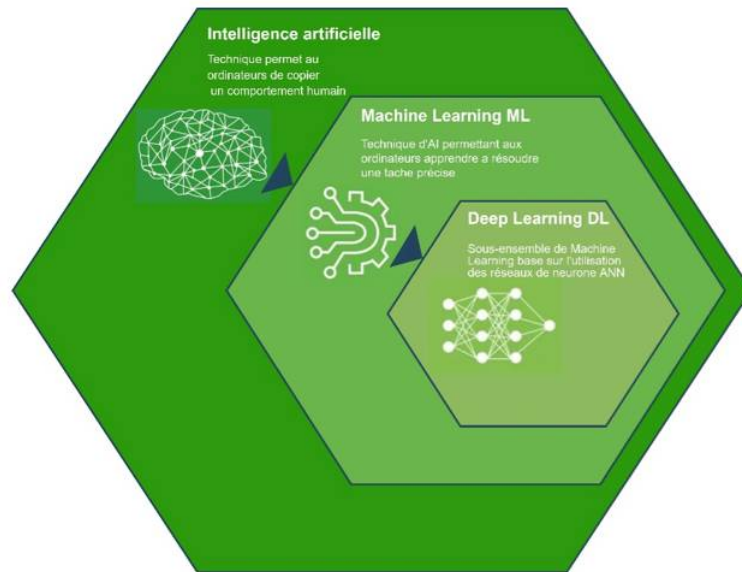


Figure 1.1 – Vue d'ensemble : ML vs DL vs AI.

1.5.1 Avantages de l'apprentissage profond

Le deep learning présente plusieurs avantages par rapport aux algorithmes traditionnels d'apprentissage automatique. Les algorithmes de Deep Learning apprennent automatiquement à extraire des caractéristiques de haut niveau à partir de données brutes en utilisant plusieurs couches de réseaux neuronaux artificiels (ANN). Cela permet aux modèles de Deep Learning d'apprendre à partir de vastes quantités de données et d'exécuter des tâches telles que la reconnaissance d'images et le traitement du langage naturel avec une grande précision. Les algorithmes de Machine Learning peuvent effectuer des tâches telles que la classification et la régression, mais ils peuvent ne pas être aussi efficaces que les algorithmes de Deep Learning pour traiter des données complexes et non structurées (voir Figure 1.2) [14].

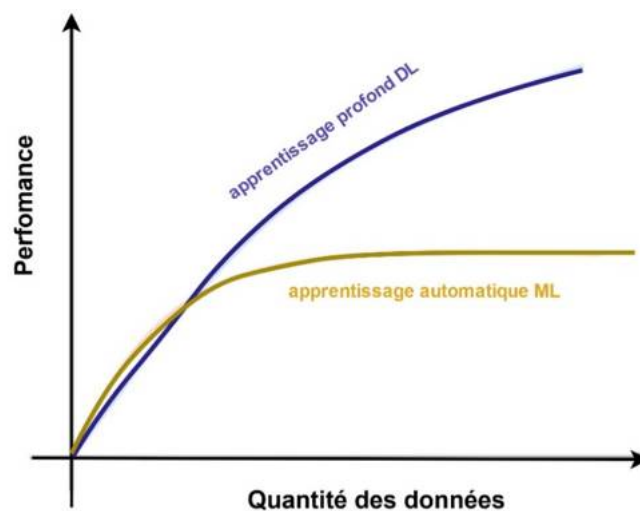


Figure 1.2 – Impact de la quantité des données sur la performance.

1.5.2 Réseaux de neurones artificiels (ANN)

Les réseaux neuronaux artificiels (ANN) simulent le mécanisme d'apprentissage pour les organismes biologiques. Le système nerveux humain contient des cellules appelées neurones. Les neurones sont connectés entre eux par des axones et des dendrites, et les régions de connexion entre les axones et les dendrites sont appelées synapses. Ces connexions sont illustrées dans la Figure 1.3.

Les dendrites reçoivent des signaux entrants d'autres neurones via les synapses. Ces signaux sont ensuite intégrés dans le corps cellulaire, et si la somme des signaux dépasse un certain seuil, le neurone déclenche un potentiel d'action le long de son axone. Ce signal électrique parcourt l'axone jusqu'à ses terminaisons, où il est transmis à d'autres neurones dans le système nerveux. Un neurone est généralement structuré en trois parties. La première partie est le corps cellulaire qui reçoit un signal et le traite. Si le signal est pertinent, le corps cellulaire s'excite et se charge en électricité. Si l'électricité atteint un certain seuil, le corps cellulaire permet à son excitation d'être transférée à la cellule suivante.

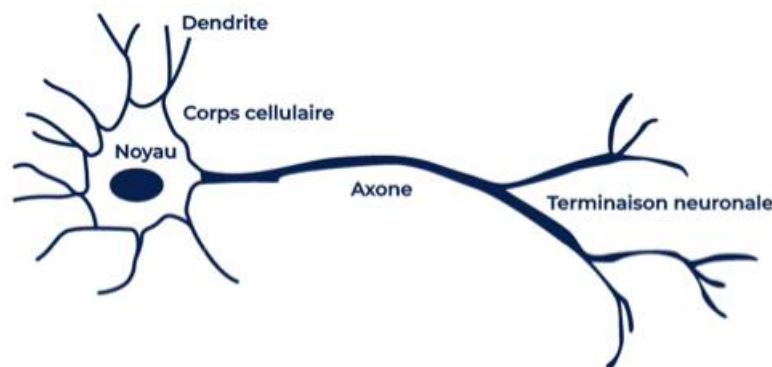


Figure 1.3 – *Neurone biologique.*

Les neurones artificiels s'inspirent de la biologie et reproduisent le fonctionnement de notre cerveau. Notre cerveau compte des milliards de cellules neuronales qui nous aident à interpréter ce que l'on appelle des signaux. Aujourd'hui, les neurones artificiels permettent de transférer ce principe dans un ordinateur en reproduisant un neurone biologique sous forme de programme.

Un neurone réel peut être excité ou non par un signal électrique. Un neurone artificiel possède également cette capacité, on peut imiter les propriétés naturelles d'un neurone réel en introduisant ce que l'on appelle des poids. Un poids est multiplié par l'entrée et donne une nouvelle valeur. D'autre part un neurone biologique accumule toutes ces informations et se charge en électricité. Un neurone artificiel tente de copier ce comportement en additionnant toutes les différentes valeurs pondérées pour chaque entrée. C'est ce que nous appelons la fonction de transfert. La somme de toutes les valeurs pondérées nouvellement calculées est appelée l'entrée du réseau (voir figure 1.4). Lorsqu'un neurone est chargé, il transmet l'électricité au neurone suivant si un certain seuil est dépassé. Nous pouvons mettre en œuvre ce principe en introduisant une fonction appelée la fonction d'activation. Cette dernière reçoit l'entrée du réseau et vérifie si elle dépasse un seuil prédéfini [14], [15].

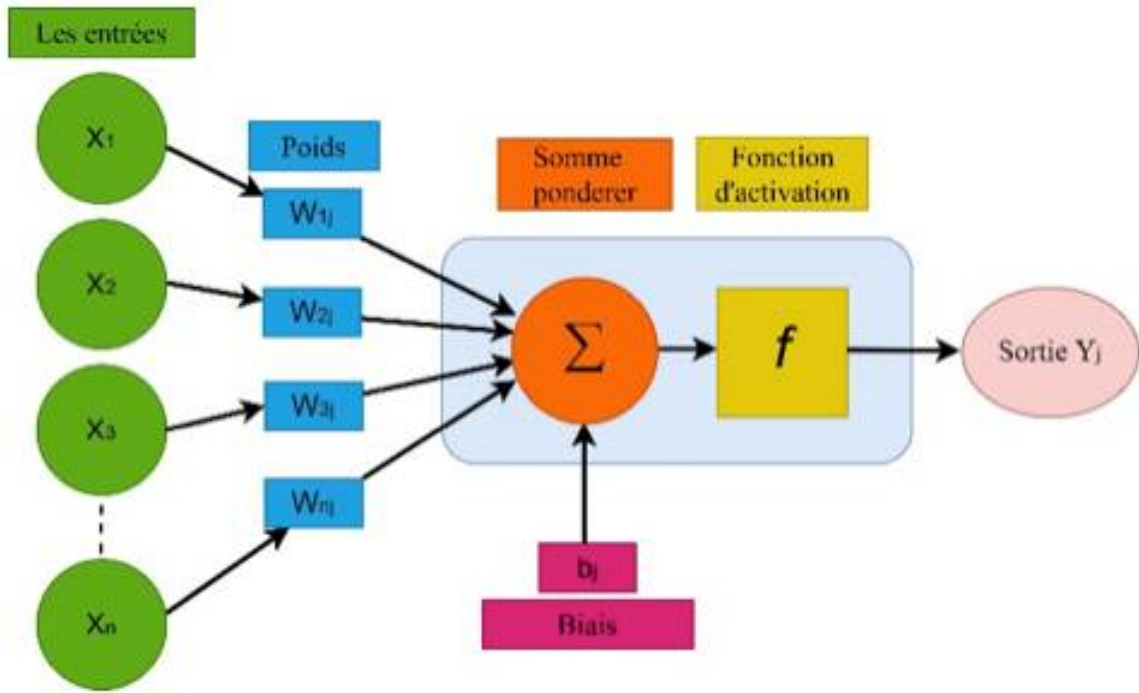


Figure 1.4 – Neurone artificiel.

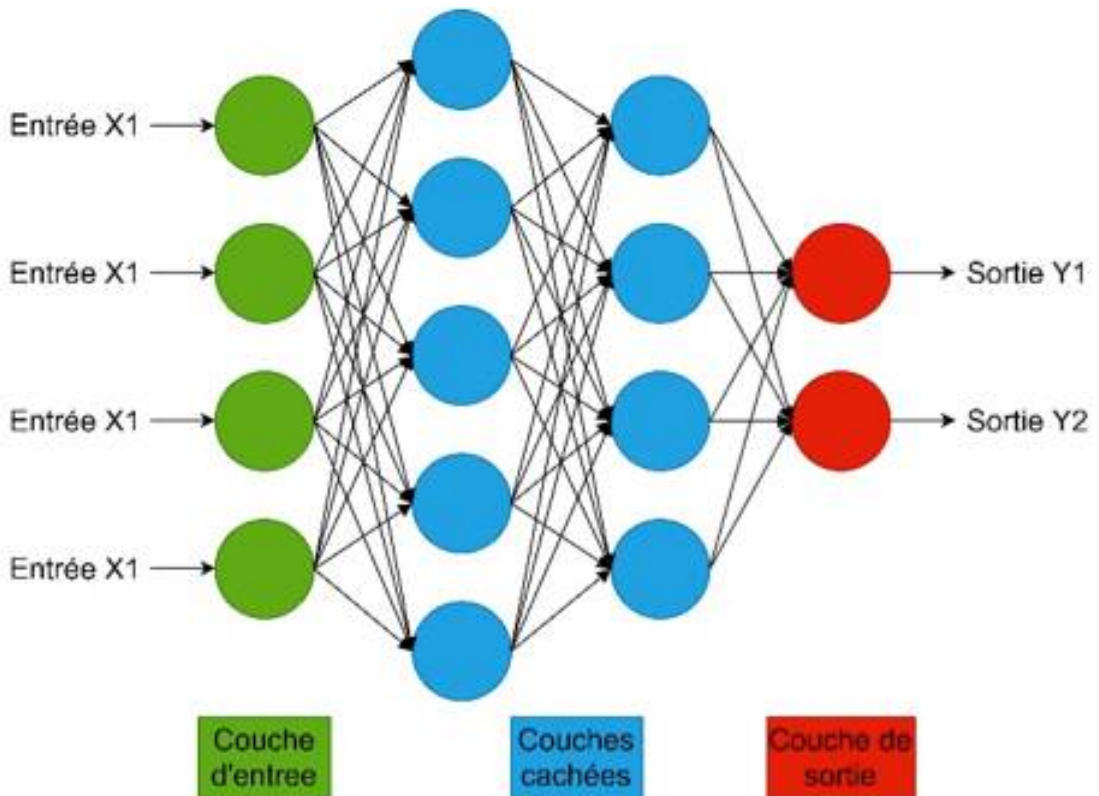


Figure 1.5 – Réseaux de couches.

1.5.2.1 Les composants des Réseaux Neuronaux Artificiels

Les réseaux de neurones artificiels (RNA) sont constitués de plusieurs éléments qui travaillent ensemble pour traiter l'information et prendre des décisions. Ces éléments incluent [16] :

1. **Neurones (unités) :** Egalement appelés des perceptrons, ce sont les unités de base d'un réseau neuronal. Ils sont responsables du traitement des informations, et chaque neurone reçoit des signaux en entrée, les traite et envoie des signaux en sortie à d'autres neurones.
2. **Couches de neurones :** Une couche de neurones est un groupe de neurones artificiels interconnectés. Les couches de neurones sont empilées les unes sur les autres pour former l'architecture du réseau, et permettent aux réseaux d'apprendre des relations complexes entre les données d'entrée et les sorties souhaitées. Il existe différents niveaux de couches de neurones, une couche d'entrée reçoit les données brutes du réseau, les couches cachées qui sont situées entre les couches d'entrée et de sortie et essentiellement effectuent le traitement principal de l'information, et la couche de sorties qui est la dernière couche du réseau neuronal qui produit les sorties finales du réseau (voir figure 1.5).
3. **Poids et biais :** Comme illustre la figure 1.4, les poids représentent la force de la connexion entre deux neurones. Ils sont initialisés à des valeurs aléatoires et sont ensuite ajustés pendant l'entraînement pour minimiser l'erreur de prédiction. Par contre, les biais se sont des termes constants et des valeurs ajoutées à la somme pondérée des entrées pour ajuster le modèle, qui permettent mieux répondre aux données d'entraînement.
4. **Fonctions d'activation :** La fonction d'activation est une fonction mathématique qui est appliquée à la sortie pondérée d'un neurone pour déterminer son activation. La fonction d'activation introduit une non-linéarité dans le réseau, ce qui lui permet d'apprendre des relations complexes entre les données d'entrée et les sorties souhaitées. Parmi les fonctions courantes, on trouve ReLU (Rectified Linear Unit), Sigmoid, Tanh (Hyperbolic Tangent), et Softmax [16]. Les fonctions d'activation seront présentées dans le tableau 1.1.
5. **Fonction de coût ou de perte :** C'est une fonction qui mesure l'écart entre les prédictions du modèle et les valeurs réelles. Elle sert aussi à quantifier l'erreur du modèle pour guider l'optimisation du modèle pendant l'entraînement. Parmi les fonctions coût, on trouve, Erreur Quadratique Moyenne (Mean Squared Error, MSE), et Entropie Croisée (Cross-Entropy) [16], [17]. Le tableau 1.2 illustre les différentes fonctions de perte.
6. **Processus d'optimisation :** C'est un algorithme utilisé pour ajuster les poids et les biais afin de minimiser la fonction de coût. Parmi les méthodes, il y a Descente de Gradient Stochastique (SGD), Adam (Adaptive Moment Estimation), RMSProp (Root Mean Square Propagation) (pour plus de détails, voir le tableau 1.3).
7. **Les données :** Les données englobent les données d'apprentissage qui sont utilisées pour entraîner un réseau neuronal, et doivent être représentatives du problème que le modèle doit résoudre, et les données de test qui sont un ensemble de données distinctes utilisées pour évaluer la performance finale du modèle après l'entraînement. Ce sont des données que le modèle n'a pas vues auparavant.

Les principaux composants sont résumés dans les tableaux 1.1, 1.2, et 1.3 :

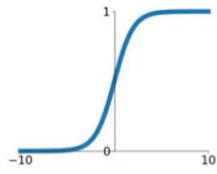
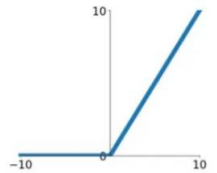
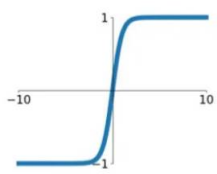
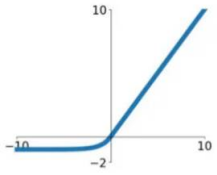
Fonction d'activation	Formule mathématique	Description/ utilisation	Forme de courbe
Sigmoid	$\sigma(x) = \frac{1}{1+e^{-z}}$	Introduit une non-linéarité, utile pour la classification binaire.	
Softmax	$softmax(x_i) = \frac{e^{x_i}}{\sum_{j=1}^n e^{x_j}}$	Utilisée dans les couches de sortie pour la classification multi-classe.	
ReLU (Rectified Linear Unit)	$ReLU(x) = \max(0, x)$	Remplace les valeurs négatives par 0, ce qui aide à résoudre le problème de gradient en déclin. Utilisée dans les couches cachées.	
Tanh (Hyperbolic Tangent)	$tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$	Normalise les valeurs entre -1 et 1. Utilisée dans les couches cachées.	
ELU	x si $x \geq 0$ $\alpha(e^x - 1)$ si $x < 0$	Réduire les biais pour les valeurs négatives tout en maintenant les non-linéarités. Utilisée dans les couches cachées.	

Tableau 1.1 – Les fonctions d'activation

Fonction de coût (ou perte)	Formule mathématique	Description	Utilisation principale
Erreur Quadratique Moyenne (MSE)	$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$	Mesure la moyenne des carrés des erreurs entre les valeurs réelles et prédites.	Régression
Entropie Croisée (Binary Cross-Entropy)	$BCE = -\frac{1}{N} \sum_{i=1}^N [(y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i))]$	Évalue la différence entre les probabilités prédites et les vraies valeurs binaires.	Classification Binaire
Entropie Croisée (Categorical Cross-Entropy)	$CategoricalCrossEntropy = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^C y_{ij} \log(\hat{y}_{ij})$	Évalue la différence entre les probabilités prédites pour chaque classe et les vraies classes.	Classification Multi-classe
Focal Loss	$FocalLoss = -\alpha(1 - p_t)^\gamma \log(p_t)$	Modifie l'entropie croisée pour se concentrer sur les exemples difficiles.	Classification
Hinge Loss	$HingeLoss = \max(0, 1 - \hat{y}_i y_i)$	Utilisé pour les SVM, mesure la distance à la marge pour la classification binaire.	Classification Binaire avec SVM
Kullback-Leibler Divergence (KL Divergence)	$KL(P Q) = \sum_i P(i) \log \frac{P(i)}{Q(i)}$	Mesure la différence entre deux distributions de probabilité.	Modélisation des distributions de probabilité

Tableau 1.2 – Fonctions de coût ou de perte

Fonction d'optimisation	Formule mathématique	Description	Avantage
Descente de Gradient	$\theta = \theta - \alpha \nabla J(\theta)$	Ajuste les paramètres dans la direction opposée au gradient de la fonction de coût.	Simple à comprendre et à implémenter.
SGD (Descente de Gradient Stochastique)	$\theta = \theta - \alpha \nabla J(\theta; x^{(i)}; y^{(i)})$	Utilise un exemple ou un petit lot d'exemples pour chaque mise à jour des paramètres.	Convergence plus rapide que la descente de gradient classique. Efficace pour les grands ensembles de données.
Adam	$m_t = \beta_1 m_{t-1} + (1 - \beta_1) \nabla J(\theta)$ $v_t = \beta_2 v_{t-1} + (1 - \beta_2) (\nabla J(\theta))^2$ $\hat{m}_t = \frac{m_t}{1 - \beta_1^t} \quad \hat{v}_t = \frac{v_t}{1 - \beta_2^t} \quad \theta = \theta - \alpha \frac{\hat{m}_t}{\sqrt{\hat{v}_t + \epsilon}}$	Utilise des moyennes mobiles des gradients et des carrés des gradients.	Convergence rapide. Adaptatif aux différents types de données.
RMSProp	$v_t = \beta v_{t-1} + (1 - \beta) (\nabla J(\theta))^2$ $\theta = \theta - \frac{\alpha}{\sqrt{v_t + \epsilon}}$	Ajuste le taux d'apprentissage pour chaque paramètre en fonction de la moyenne des carrés des gradients passés.	Convergence stable et rapide. Efficace pour les réseaux de neurones récurrents.
Adagrad	$\theta = \theta - \frac{\alpha}{\sqrt{G_t + \epsilon}}$ $G_t = \sum_{i=1}^t (\nabla J(\theta_i))^2$	Ajuste le taux d'apprentissage pour chaque paramètre en fonction de la somme cumulative des gradients carrés.	Performant pour les problèmes de traitement du langage naturel.
L-BFGS	Utilise une approximation de la matrice Hessienne : $H_k p_k = -\nabla f(x_k)$	Méthode quasi-Newton utilisant une approche par mémoire limitée pour estimer la matrice Hessienne.	Convergence rapide. Efficace pour les problèmes de grande dimension.

Tableau 1.3 – Fonctions d'optimisation

1.5.2.2 Types de réseaux de neurones artificiels

Dans cette partie, nous souhaitons approfondir les différents types de réseaux de neurones artificiels et vous aider à mieux comprendre leur impact sur la technologie moderne. En général, chaque type de réseau de neurones artificiels possède ses propres forces et cas d'utilisation distincts [16]. Nous en avons décrit quelques-uns ci-dessous :

- Le réseau de neurones à propagation dirigée (FFNNs) : Appelé aussi réseau profond à propagation dirigée est l'un des types de réseaux de neurones artificiels (ANN) les plus simples. Dans ce type d'ANN, les données passent par plusieurs nœuds, jusqu'à atteindre un nœud de sortie. Un réseau à propagation dirigée n'a pas de cycles ou de boucles, ce qui signifie que l'information ne peut pas faire demi-tour et influencer les étapes précédentes. Ces types de réseaux peuvent rencontrer des difficultés pour prédire précisément la sortie pour une entrée donnée, en particulier lorsqu'ils traitent des problèmes complexes. Dans ce cas, les poids et les biais du réseau doivent être ajustés pour réduire l'écart entre la sortie prédite et la sortie réelle. C'est là qu'intervient la rétro-propagation, qui permet au réseau d'ajuster ses poids et ses biais. La rétro-propagation aborde ce problème en calculant l'erreur entre la sortie prédite par le réseau et la sortie réelle attendue. Cette erreur est ensuite propagée en arrière à travers le réseau, couche par couche. À chaque étape, on calcule la contribution de chaque poids et biais à l'erreur globale.
- Le perceptron Multi-Couches (MLP) : Le MLP est un type spécifique de FFNN, d'où l'apprentissage dans un MLP se fait généralement par l'algorithme de rétro-propagation. Ce type d'ANN possède au moins trois couches et permet de classifier des données qui ne sont pas linéairement séparables (c'est-à-dire qui ne peuvent être distinguées par une droite). Pour parvenir à cela, il est entièrement connecté, ce qui signifie que chaque nœud (neurone artificiel) d'une couche est connecté à tous les nœuds de la couche suivante.
- Réseaux Neuronaux Récurrents (RNN) : Ce type de réseau de neurones artificiels est similaire à un réseau de neurones à propagation dirigée, mais il enregistre la sortie d'une couche spécifique et la réinjecte en tant qu'entrée. Cela lui permet de prendre en compte les informations passées et de potentiellement prédire plusieurs sorties possibles à partir d'une couche donnée. Pour plus de précisions, si la première couche reçoit la sortie et la renvoie en tant qu'entrée, les couches suivantes commenceront le processus neuronal récurrent. et la renvoie en tant qu'entrée, les couches suivantes commenceront le processus du réseau récurrent, et chaque nœud conserve l'étape précédente. Et donc, le système se souvient des mauvaises prédictions et en tire des leçons pour améliorer les suivantes. En bref, les RNN peuvent apprendre de chaque étape pour prédire le résultat de l'étape suivante.
- Les réseaux Neuronaux à Mémoire Longue et Court Terme (LSTM) : Les réseaux LSTM sont un sous-ensemble de RNN capables d'apprendre les dépendances à long terme. Comme il s'agit d'un RNN, le LSTM est également un réseau séquentiel, ce qui signifie qu'il peut se souvenir des informations du processus précédent et utiliser ces connaissances pour traiter l'entrée actuelle. Cependant, les RNN ne peuvent pas mémoriser les dépendances à long terme, ni fournir des résultats pour des entrées qui dépendent d'informations remontant très loin dans le temps. C'est là que la LSTM entre en jeu, le réseau peut choisir

de conserver certaines informations pour les utiliser à l'avenir et d'oublier les informations non adéquates.

- Réseau de neurones convolutifs (CNN) : Ce type d'ANN applique une version différente des MLP en possédant plusieurs couches qui peuvent être complètement interconnectées. L'objectif premier des CNN est de déchiffrer les caractéristiques spécifiques d'une image donnée, telle qu'un visage. Ils identifient les caractéristiques en fonction de la distance qui sépare chaque pixel (qui constitue chaque caractéristique du visage) d'un point de référence. Les réseaux de neurones convolutifs seront plus détaillés dans la section suivante.
- Réseaux Neuronaux à Graphes (GNN) : Les GNN sont conçus pour traiter des données organisées sous forme de graphes. Un graphe est une structure composée de nœuds (entités) connectés par des arêtes (relations). Les GNN permettent d'analyser et de tirer parti des relations entre les entités d'un graphe pour effectuer des tâches de prédiction ou de classification. Contrairement aux réseaux de neurones convolutifs (CNN) conçus pour les images en grille ou aux réseaux de neurones récurrents (RNN) pour les séquences, les GNN traitent des structures plus flexibles. Ils fonctionnent en propageant des messages entre les nœuds voisins dans un graphe. À chaque étape de propagation, un nœud agrège les informations de ses voisins et met à jour sa propre représentation. Ce processus permet au réseau d'apprendre des relations complexes entre les entités au sein du graphe.

Le tableau 1.4 présente des types de réseaux neuronaux artificiels avec une vue d'ensemble des caractéristiques, des exemples d'algorithmes et des applications.

Type de Réseau	Architecture	Algorithmes	Applications	Illustration
Le réseau de neurones à propagation dirigée	<ul style="list-style-type: none"> • Couche d'entrée • Couches cachées • Couche de sortie 	Perceptron, MLP	Classification, Régression, détection d'objets	Feedforward Neural Network (FFNN)
Le perceptron Multi-Couches	<ul style="list-style-type: none"> • Couche d'entrée • Couches cachées • Couche de sortie 	Descente de Gradient Stochastique (SGD), Adam (Adaptive Moment Estimation)	Prédiction de séries temporelles, traitement du langage naturel (NLP)	Multilayer Perceptron (MLP)
Réseaux Neuronaux Récurrents	<ul style="list-style-type: none"> • Couche récurrente • Couche de sortie 	RNN de Base, LSTM, GRU	Traitement du langage naturel, Analyse de séries temporelles	Recurrent Neural Networks (RNN)
Les réseaux Neuronaux à Mémoire Longue et Court Terme	Cellule LSTM	LSTM Standard, Bi-directional LSTM	Traduction automatique, Reconnaissance vocale	Long Short-Term Memory, (LSTM)
Réseau de neurones convolutifs	<ul style="list-style-type: none"> • Convolution • Pooling • Classification 	LeNet-5, AlexNet, VGGNet, ResNet...	Vision par ordinateur, Traitement d'image et de vidéo	Convolutional Neural Networks, (CNN)
Réseaux Neuronaux à Graphes	<ul style="list-style-type: none"> • Nœuds • Arêtes 	GCN (Graph Convolutional Network), GAT (Graph Attention Network)	Réseaux sociaux, Analyse de molécules	Graph Neural Networks, (GNN)

Tableau 1.4 – *Types de réseaux de neurones, leurs architectures, algorithmes, applications et illustrations*

1.6 Réseau de neurones convolutif (CNN)

Le réseau de neurones convolutif fonctionne différemment des autres types de réseaux car il traite les données de manière spatiale. Au lieu de connecter chaque neurone à tous les neurones de la couche suivante, il utilise des connexions locales et partagées, ce qui lui permet de capturer des motifs locaux dans les données. Les réseaux de neurones convolutifs (CNN) sont utilisés pour diverses tâches en vision par ordinateur, telles que l'identification d'images, le regroupement par similarité et la reconnaissance d'objets dans des scènes. Voici une explication détaillée de ces tâches, accompagnée d'exemples et de concepts clés. Pour comprendre la vision par ordinateur, discutons de la vision humaine. La vision humaine est la capacité de l'œil et du cerveau humain à voir et à reconnaître des objets. La vision par ordinateur est le processus qui consiste à donner à une machine une compréhension similaire, voire meilleure, de la vision et de l'identification des objets dans le monde réel. Le réseau de neurones artificiels (ANN) le plus couramment appliqué en vision par ordinateur est le réseau de neurones convolutif (CNN). Il y a deux raisons principales pour lesquelles les CNN sont utilisés dans les problèmes de vision par ordinateur. Premièrement, avec les ANN traditionnels, la détermination du problème de vision par ordinateur est déjà difficile pour des images de petite taille. En outre, les CNN nécessitent un prétraitement d'image relativement minimal par rapport à d'autres algorithmes de classification d'images, ce qui signifie que les CNN peuvent apprendre les filtres par eux-mêmes. La figure 1.6 montre un exemple d'un réseau CNN dédié pour la classification des véhicules.

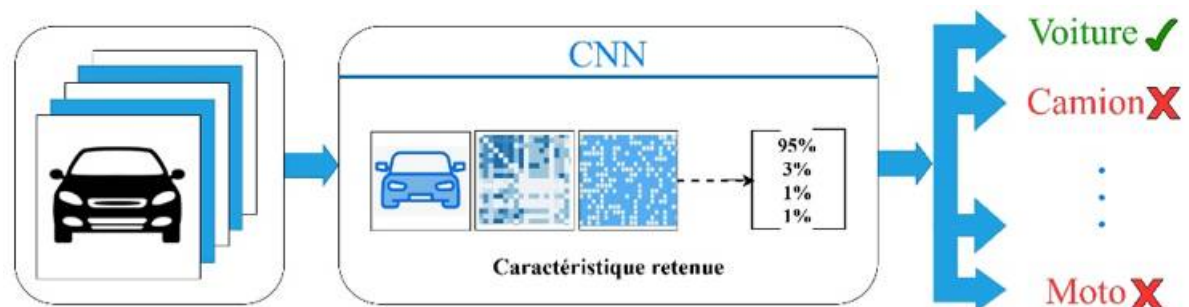


Figure 1.6 – CNN, exemple de classification de véhicules.

1.6.1 Architecture de réseau CNN

Dans les CNN, le réseau contient plusieurs couches appelées (voir figure 1.7), couche d'entrée, couche de sortie et plusieurs couches cachées [18]. Les couches cachées sont constituées par d'autres couches telles que : les couches de convolution (Convolutional Layers), les couches ReLU (Rectified Linear Unit) en tant que fonction d'activation, les couches de normalisation (Normalization Layers), les couches de pooling (Pooling Layers), et les couches entièrement connectées (FC Layers : Fully Connected Layers) [19].

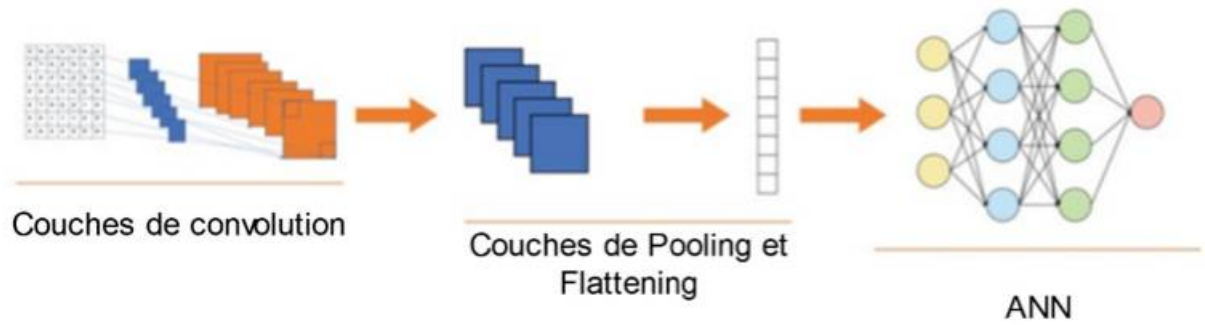


Figure 1.7 – Architecture et couches du réseau de neurones convolutif CNN.

Les principaux composants de l'architecture du CNN sont les suivants : l'image d'entrée, les couches de convolution, les couches de Pooling et la couche de Flattening.

1.6.1.1 L'image d'entrée

A l'entrée du réseau de neurones convolutif (CNN), l'image d'entrée est l'image brute qui est alimentée au réseau pour traitement. L'image peut être de n'importe quel type : humain, un animale, image médical (rayon X, scanner, IRM, etc.), un objet quelconque. La figure 1.8 explique comment un ordinateur visualise l'image de la lettre L. Tous les blocs ayant une valeur de '1' représentent les données, tandis que les '0' représentent l'espace vide.

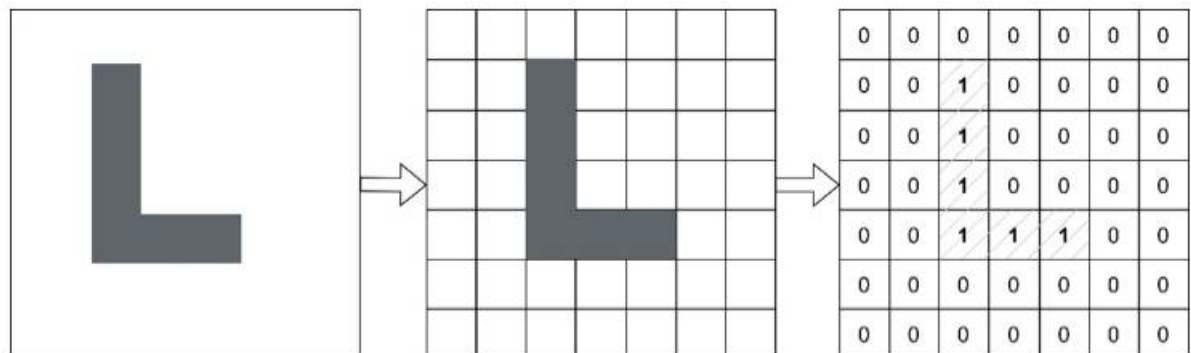


Figure 1.8 – Matrice pour image en lettre L.

1.6.1.2 La couche de convolution

La couche de convolution est le point de départ du traitement de l'image. Elle est conçue pour détecter les motifs et les caractéristiques locales dans les données d'entrée. Une couche de convolution se compose de deux parties, un détecteur de caractéristiques (ou filtre), et une carte de caractéristiques (feature map).

Un filtre est typiquement une petite matrice, par exemple de taille 3x3, 5x5, etc. Ce filtre est déplacé sur l'image d'entrée, généralement pixel par pixel, qu'il doit à chaque position faire un produit scalaire entre ces valeurs et les valeurs correspondantes de la région de l'image est calculé. Le résultat de cette opération est une valeur unique qui est stockée dans une carte de caractéristiques (feature map) [20] (voir la figure 1.9).

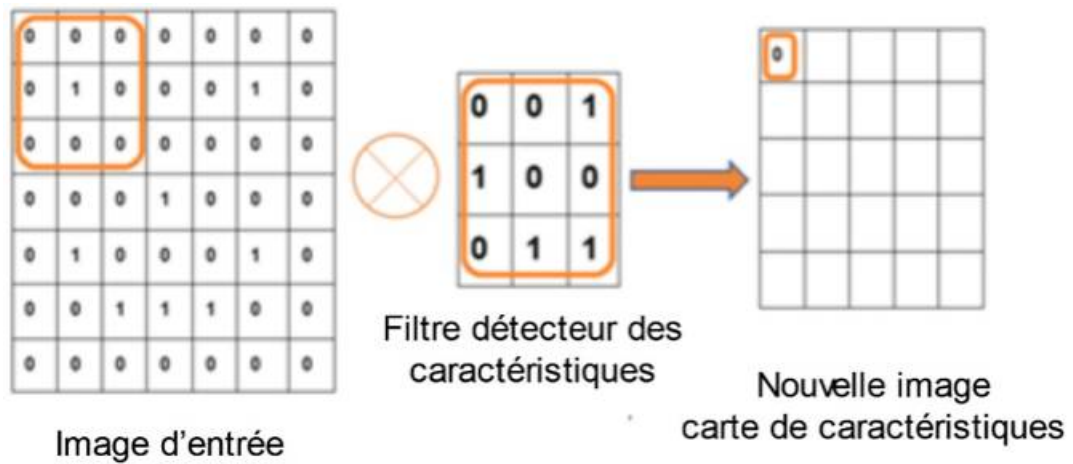


Figure 1.9 – Détecteur de caractéristiques et la carte de caractéristiques.

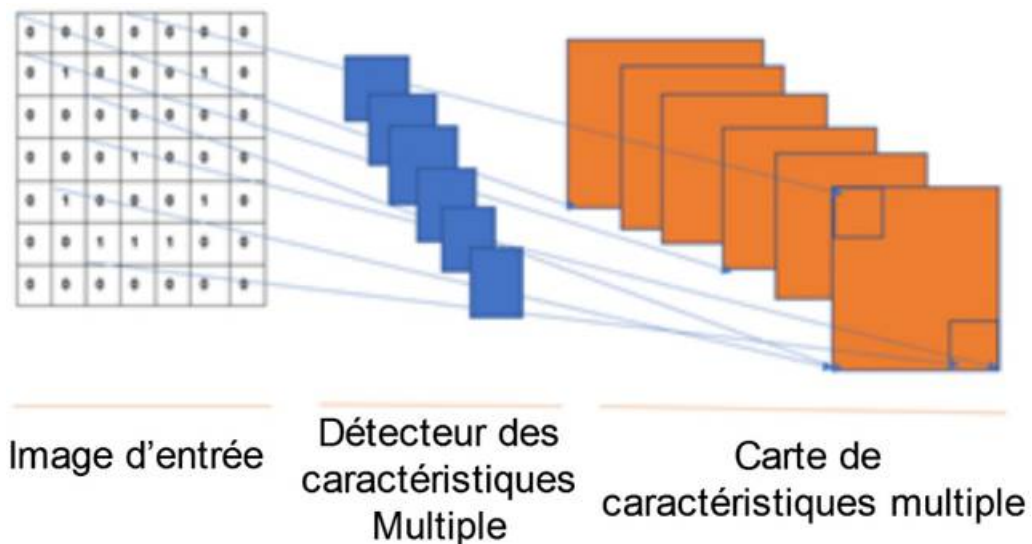


Figure 1.10 – Détecteurs de caractéristiques et de carte multiples.

Comme on peut le voir, ce détecteur de caractéristiques est placé (superposé) sur l'image originale et le calcul est effectué sur les éléments correspondants. Le calcul est effectué en multipliant les éléments correspondants (voir la figure 1.9). Ce processus est répété pour toutes les cellules. Il en résulte une nouvelle image traitée.

Dans un CNN réel, plusieurs détecteurs de caractéristiques sont utilisés pour produire de nombreuses cartes de caractéristiques (voir la figure 1.10).

1.6.1.3 La couche de Pooling

La couche de Pooling également appelée couche de sous-échantillonnage ou couche de réduction, nous permet d'ignorer les données les moins importantes de l'image et de réduire davantage l'image, tout en préservant ses caractéristiques importantes.

Par exemple pour déterminer si une image contient un chat ou non, le réseau neuronal analyse l'image. Il peut examiner la forme des oreilles, des yeux, etc. Au même temps, l'image inclut de nombreuses caractéristiques qui ne sont pas liées aux

chats. Le mécanisme de pooling aide l'algorithme à comprendre quelles parties de l'image sont pertinentes et quelles parties ne le sont pas. Pour les types de pooling [21], on trouve le "Max Pooling", le "Average Pooling", et le "Global Pooling". Le Max Pooling sélectionne la valeur maximale dans chaque fenêtre. C'est le type de pooling le plus courant. Le Average Pooling détermine la moyenne des valeurs dans chaque fenêtre.

Enfin, le Global Pooling est appliqué sur l'ensemble de la carte de caractéristiques, produisant une seule valeur (utilisée en général en fin de réseau).

La carte de caractéristiques dérivée de la couche de convolution passe par une couche de max pooling pour réduire encore l'image, tout en préservant la partie la plus pertinente de l'image. Cela signifie que nous sélectionnons une taille de matrice, par exemple 2x2, et que nous balayons la carte des caractéristiques et sélectionnons le nombre maximum de la matrice 2x2 qui s'enfonce dans le bloc [21], [22]. La figure 1.11 nous donne une idée claire du fonctionnement du max pooling.

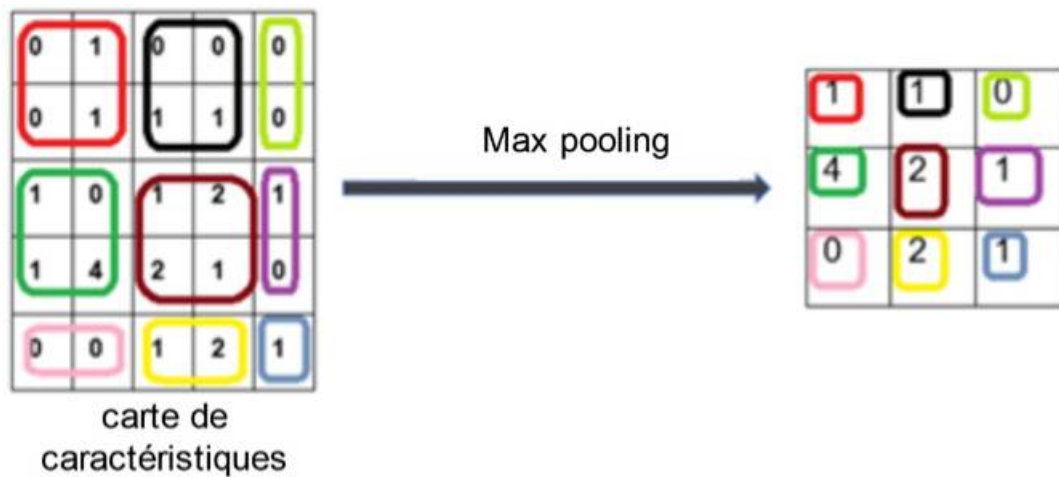


Figure 1.11 – Max pooling.

1.6.1.4 La couche Flattening

La couche Flattening est une étape courante dans les réseaux de neurones convolutifs (CNN) qui intervient après les couches de convolution et de pooling.

La fonction de la couche Flattening est la transformation de la dimensionnalité qui convertit les données de sortie des couches précédentes (souvent des matrices 3D avec hauteur, largeur et profondeur) en un vecteur unidimensionnel qui constitue l'entrée pour les couches FC layers d'où la nécessité de cette transformation.

Dans l'exemple de la figure 1.12, chaque ligne est transformée en colonne et empilée. Ici, nous avons converti une matrice 3x3 en une matrice 1xn, où n, dans ce cas, est 9 :

1.6.2 Exemple d'implémentation d'un CNN

Cette partie montre un exemple complet de la construction d'un CNN pour une tâche de classification d'images en utilisant le module Tensorflow-Keras.

Le tableau 1.5 montre les séquences d'un réseau de neurones convolutif typique.

Composant	Description	Exemple de Code en Python
Couche d'Entrée	Recevoir les données brutes, telles que des images, et les préparer pour le traitement.	<code>input_shape = (28, 28, 1)</code> <code> #(dimension d'image d'entrée)</code>
Couche Convulsive	Appliquer des filtres pour extraire des caractéristiques locales comme des bords et des textures.	<code>Conv2D(32, (3,3), activation='relu',</code> <code> input_shape=(28, 28, 1))</code> <code> #(32 filtres, taille de filtre 3x3)</code>
Couche ReLU	Ajouter une activation non linéaire pour améliorer l'apprentissage en activant les neurones positifs.	<code>Activation('relu')</code>
Couche de Normalisation	Standardiser les activations pour améliorer la performance et la stabilité du modèle.	<code>BatchNormalization()</code>
Couche de Pooling	Réduire la taille des caractéristiques pour extraire les informations les plus significatives.	<code>MaxPooling2D((2, 2))</code>
Deuxième Couche Convulsive	Extraire des caractéristiques plus complexes à partir des caractéristiques traitées précédemment.	<code>Conv2D(64, (3, 3))</code> <code> #(64 filtres, taille de filtre 3x3)</code>
Deuxième Couche ReLU	Ajouter une activation non linéaire après la deuxième convolution	<code>Activation('relu')</code>
Deuxième Couche de Normalisation	Standardiser les activations après la deuxième couche convulsive.	<code>BatchNormalization()</code>
Deuxième Couche de Pooling	Réduire davantage la taille des caractéristiques après la deuxième convolution.	<code>MaxPooling2D((2, 2))</code>
Couche de Flattening	Aplatir les données avant les couches FC layers.	<code>Flatten()</code>
Couche Entièrement Connectée (FC layers)	Combiner les caractéristiques extraites pour produire des prédictions.	<code>Dense(128, activation='relu')</code>
Couche de Sortie	Produire les résultats finaux du modèle, souvent sous forme de probabilités pour chaque classe.	<code>Dense(10, activation='softmax')</code> <code> # (10 classes)</code>

Tableau 1.5 – *Composants d'un Réseau de Neurones Convolutif (CNN) et leurs descriptions*

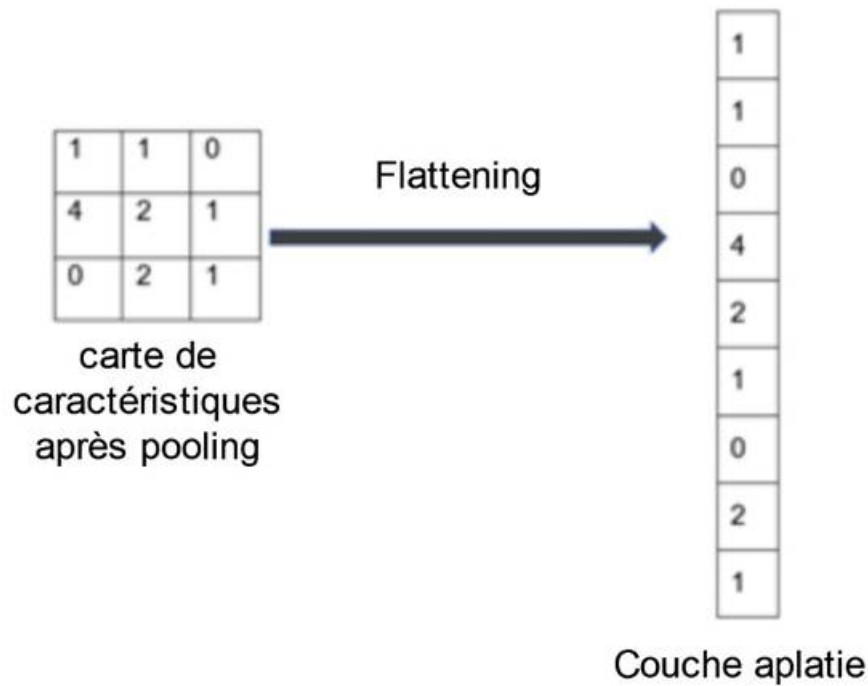


Figure 1.12 – Procédure de Flattening.

Le listing 2.1 présente le code correspond à l'exemple donné par le tableau 1.5.

Listing 1.1 – Code Python réseau CNN

```

from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Conv2D, MaxPooling2D, ReLU,
    BatchNormalization, Flatten, Dense
# Définir le modele CNN
model = Sequential([
    # Couche d'entr e et couche convolutive
    Conv2D(32, (3, 3), input_shape=(28, 28, 1)), # Couche
    convolutive
    ReLU(), # Activation ReLU
    BatchNormalization(), # Normalisation des activations
    MaxPooling2D((2, 2)), # Couche de pooling
    # Deuxi me bloc de couches
    Conv2D(64, (3, 3)), # Couche convolutive
    ReLU(), # Activation ReLU
    BatchNormalization(), # Normalisation des activations
    MaxPooling2D((2, 2)), # Couche de pooling
    # Aplatir les donn es avant les couches FC layers
    Flatten(),

    # Une seule couche entierement connectee
    Dense(128, activation='relu'), # Couche dense
    # Couche de sortie
    # 10 classes pour une t che de classification

```

1.6.3 L'apprentissage par transfert et modèles pré-entraînés

Dans cette partie, nous présentons les concepts importants de l'apprentissage par transfert et des réseaux pré-entraînés également connus sous le nom de modèles

pré-entraînés. Supposons que nous voulions utiliser deux approches pour savoir si une photo représente un chien ou un chat. La première consiste à construire un modèle d'apprentissage profond à partir de zéro, puis à transmettre les nouvelles images aux réseaux. Une autre option consiste à utiliser un modèle de réseau neuronal d'apprentissage profond pré-entraîné qui a déjà été construit en utilisant des images de chats et de chiens, au lieu de créer un réseau neuronal à partir de zéro. L'utilisation d'un modèle pré-entraîné permet d'économiser le temps de calcul et des ressources.

L'utilisation d'un réseau pré-entraîné peut présenter certains avantages imprévus. Par exemple, presque toutes les photos de chiens et de chats comportent d'autres objets, tels que des arbres, le ciel et des meubles. Nous pouvons même utiliser ce modèle pré-entraîné pour identifier des objets tels que les arbres, le ciel et les meubles.

1.6.3.1 Définition du modèle pré-entraîné

Les êtres humains apprennent par l'expérience. Nous appliquons les connaissances acquises dans une situation donnée à des situations similaires auxquelles nous serons confrontés à l'avenir. Supposons que vous souhaitiez apprendre à conduire un véhicule SUV (Sport Utility Vehicule) et vous ne savez conduire que des petites voitures avec une boîte de vitesses simple. Les dimensions du SUV sont considérablement plus grandes que celles de voiture ordinaire. Toutefois, certains systèmes de base (tels que l'embrayage, l'accélérateur et les freins) restent similaires à ceux de voiture ordinaire. Ainsi, le fait de savoir comment conduire une petite voiture et toutes les connaissances que vous avez acquises en conduisant une voiture ordinaire peuvent être utilisées lorsque vous apprenez à conduire un gros SUV. C'est précisément cela l'apprentissage par transfert. Par définition, l'apprentissage par transfert est un concept d'apprentissage automatique dans lequel nous stockons et utilisons les connaissances acquises dans une activité lors de l'apprentissage d'une autre activité similaire [23–25]. Un réseau pré-entraîné est un réseau sauvegardé, et qui a été déjà entraîné sur un très grand ensemble de données, la plupart du temps sur des problèmes de classification d'images. Pour travailler sur un réseau pré-entraîné, il est nécessaire de comprendre les concepts d'extraction de caractéristiques et d'ajustement fin.

1.6.3.2 L'extraction de caractéristiques

Pour comprendre l'extraction de caractéristiques, nous devons revoir l'architecture d'un réseau neuronal convolutif CNN (voir Figure 1.13). L'architecture complète d'un CNN, se compose des éléments suivants : une couche de convolution, les couches de pooling et flattening et un réseau neurones artificiel (ANN).

Maintenant, si on divise cette architecture en deux parties, la première partie contient les deux premiers éléments sauf l'ANN, tandis que la seconde partie ne contient que l'ANN. Dans ce cas, la première partie est appelée base convolutive ainsi que la seconde partie est appelée le classificateur (voire figure 1.14).

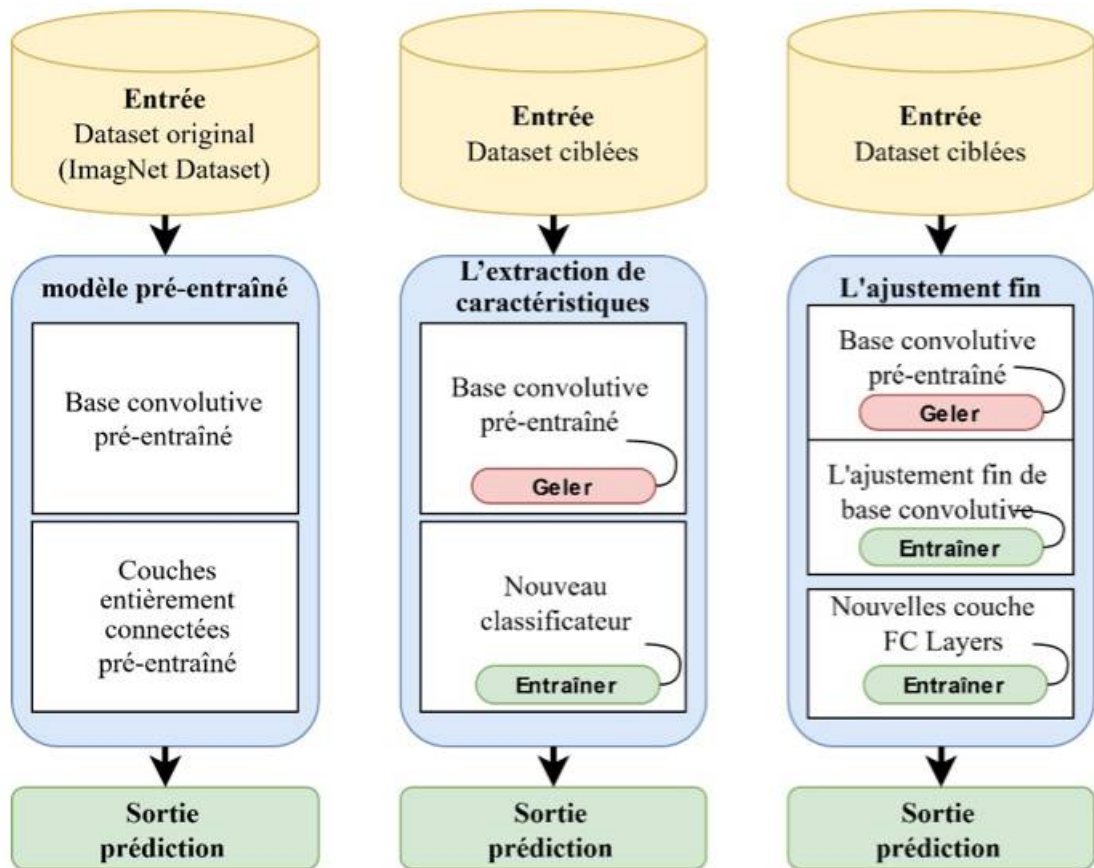


Figure 1.13 – Modèles pré-entraînés : l'extraction de caractéristiques et l'ajustement fin.

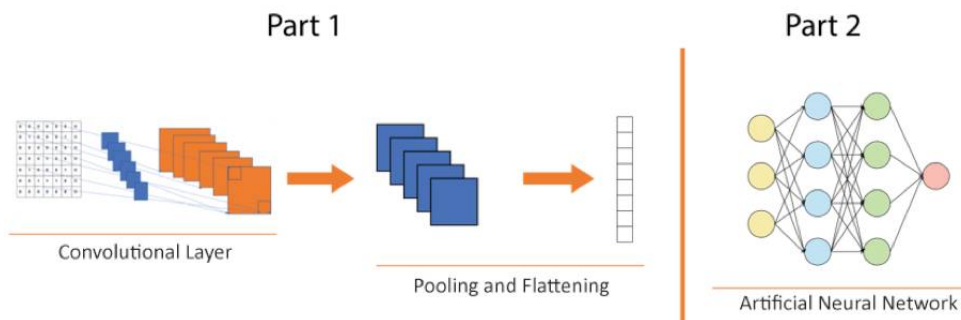


Figure 1.14 – Architecture divisée CNN – base convolutive et classificateur.

Lors de l'extraction des caractéristiques, nous préservons la base convolutive et nous modifions le classificateur [26–28]. Donc, nous continuons à réutiliser les apprentissages de la couche convolutive et nous pouvons personnaliser différents classificateurs. Un classificateur peut classifier un chien ou un chat, un vélo ou une voiture, ou même des images radiographiques médicales pour classer les tumeurs, les infections, etc.

L'apprentissage du classificateur est principalement spécifique aux classes sur lesquelles le modèle a été formé. Il est préférable de transférer uniquement la couche

convolutive de base et excluant le classificateur. L'une des caractéristiques les plus importantes de l'apprentissage pré-entraîné est de comprendre le concept de geler certaines couches. Il s'agit d'une technique qui consiste à ne pas modifier les paramètres de certaines couches d'un modèle pré-entraîné lors de son adaptation à une nouvelle tâche.

Le gel signifie essentiellement que nous arrêtons le processus de mise à jour du poids de certaines des couches inférieures qui se trouve dans la couche de base convolutive, car ces couches apprennent des représentations et connaissances abstraites à partir de grandes quantités de données d'entraînement. En gelant ces couches, on préserve ces connaissances et on évite de les perdre en les adaptant à une tâche spécifique. D'autre part, geler les couches inférieures permet de réduire le risque de sur-apprentissage [16], un phénomène où le modèle apprend trop précisément les données d'entraînement et perd sa capacité à généraliser à de nouvelles données. En fin, le gel des couches permet d'économiser du temps et des ressources de calcul, car il réduit le nombre d'opérations d'apprentissage à effectuer.

Le processus de gel de certaines couches et d'entraînement d'autres couches est appelé ajustement fin (fine-tuning) d'un réseau.

1.6.3.3 L'ajustement fin

Nous pouvons geler certaines des couches initiales du réseau afin de ne pas perdre les informations stockées dans ces couches. Nous avons procédé à un ajustement appelé ajustement fin (fine-tuning). Comme le montre la figure 1.15, l'ajustement fin consiste à ajuster notre réseau neuronal de manière à ce qu'il soit plus adapté à la tâche à accomplir.

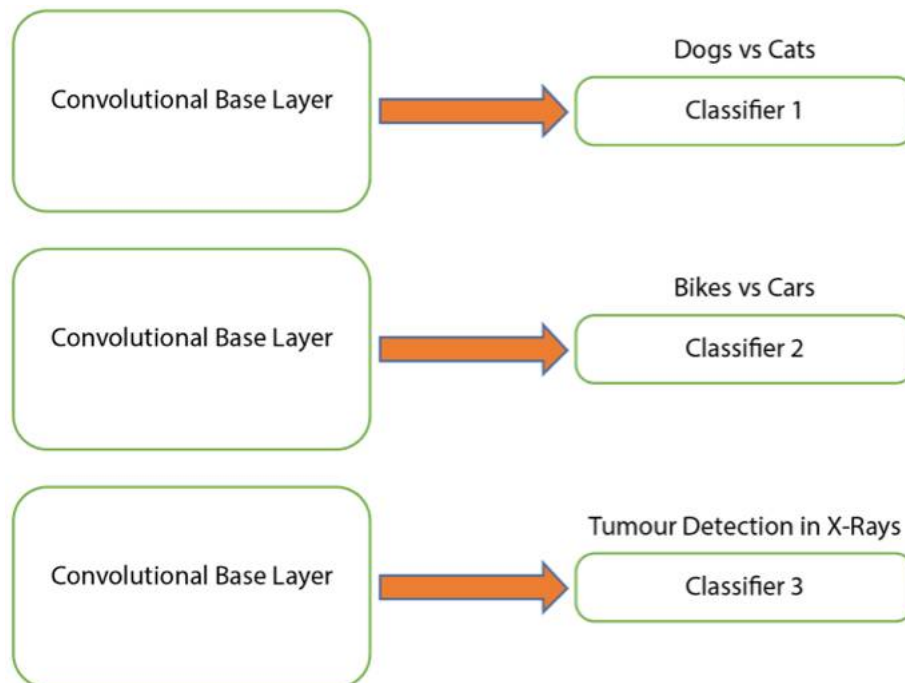


Figure 1.15 – Couche de base convolutive réutilisable.

L'ajustement fin d'un réseau pré-entraîné est une technique puissante en apprentissage profond qui permet de tirer parti de modèles ayant déjà appris des repré-

sentations générales à partir de vastes quantités de données. Supposons que nous disposions d'un réseau pré-entraîné qui identifie les animaux. Si nous voulons identifier des animaux spécifiques, comme les chiens et les chats, nous pouvons modifier légèrement les couches pour qu'elles apprennent à quoi ressemblent les chiens et les chats. Cela revient à utiliser l'ensemble du réseau pré-entraîné, puis à ajouter une nouvelle couche composée d'images de chiens et de chats. Nous ferons une activité similaire en utilisant un réseau préconstruit et en y ajoutant un classificateur, qui sera entraîné sur des images de chiens et de chats.

Voici les trois points pour réaliser l'ajustement fin [29], [30] :

- Ajouter un classificateur (ANN) à un système pré-entraîné, en ajoutant de nouvelles couches entièrement connectées (FC layers) qui seront entraînées sur la tâche spécifique.
- Geler la base convolutive (couche de convolution, les couches de pooling) qui consiste à empêcher leurs poids d'être mis à jour pendant l'entraînement.
- Compiler et entraîner conjointement les couches et le classificateur ajouté.

Pour un ajustement fin plus approfondi, vous pouvez dégeler certaines couches supérieures du modèle de base et les ré-entraîner avec un taux d'apprentissage plus faible, ceci est optionnel.

1.6.3.4 Modèles pré-entraînés populaires

Dans le cadre d'une expérience professionnelle réelle, nous n'aurons presque jamais besoin de construire un modèle convolutionnel (CNN) de base par nous-même. Nous utiliserons toujours des modèles pré-entraînés. Les modèles pré-entraînés suivants peuvent être considérés comme des couches convolutives de base. Nous utilisons ces réseaux avec l'adaptation à un classificateur (ANN). Parmi les modèles pré-entraînés, on trouve :

- - VGGNet (VGG16, VGG19) [31],
- - ResNet (ResNet50, ResNet101) [32],
- - Inception (InceptionV3, GoogleNet) [33], [34],
- - EfficientNet (EfficientNetB0, EfficientNetB7) [35],
- - MobileNet (MobileNet, MobileNet V2) [36] [37],
- - Xception [38].

Le tableau 1.6 montre une liste non exhaustive des modèles pré-entraînés.

Modèles	Architecture	Utilisations	bibliothèque
VGGNet	CNN avec des couches convolutives simples	Classification d'images, reconnaissance d'objets.	PyTorch, Keras
ResNet	CNN profond avec des connexions résiduelles	Classification d'images, détection d'objets, segmentation d'images.	PyTorch, TensorFlow
Inception	CNN avec des modules d'Inception	Classification d'images, détection d'objets, reconnaissance d'objets dans des scènes complexes.	TensorFlow, Keras
EfficientNet	CNN avec des coefficients de mise à l'échelle composés	Classification d'images, détection d'objets, segmentation d'images.	TensorFlow, PyTorch
MobileNet	CNN léger optimisé pour les appareils mobiles	Applications mobiles, détection d'objets en temps réel avec des ressources limitées.	TensorFlow, PyTorch
Xception	CNN inspiré par les réseaux Inception mais avec des convolutions séparables en profondeur	Classification d'images, segmentation sémantique, reconnaissance d'objets.	TensorFlow, Keras

Tableau 1.6 – Les modèles pré-entraînés en CNN

1.7 Préparation des données

Après la sélection d'un modèle pré-entraîné qui a été entraîné sur un ensemble de données vaste et diversifié comme ImageNet [39], la préparation des données pour l'apprentissage par transfert avec les CNNs comporte plusieurs étapes.

1.7.1 Chargement et prétraitement des données

Chargez vos données images dans votre environnement de programmation par exemple Python ensuite importez les bibliothèques courantes telles que Keras et TensorFlow qui fournissent des fonctions à cet effet. Le prétraitement consiste à nettoyer, organiser et redimensionner les images pour qu'elles correspondent à la taille d'entrée attendue par le modèle pré-entraîné. Cela implique souvent de les convertir à une taille et à un format uniformes (par exemple, 224x224 pixels, canaux RVB) [40].

La plupart des frameworks d'apprentissage profond prennent en charge les formats d'image courants tels que JPEG, PNG et BMP. Nous pouvons utiliser des bibliothèques comme OpenCV ou scikit-image pour gérer des formats supplémentaires.

Organisez les images dans des dossiers en fonction de leurs classes, ou utilisez les fichiers de format csv ou Excel. Cela permet de déduire automatiquement la découverte du chemin de différentes images par rapport aux différentes classes pendant le chargement. Le fichier Metadatas est un fichier organisé qui contient toutes les informations nécessaires telles que les noms des images, la classe de chaque image et les différentes caractéristiques. Après la lecture et le chargement des informations des images, on prétraite ces images pour les utiliser dans un modèle de réseau de neurones convolutifs (CNN) soit à partir d'un dossier en utilisant la fonction `flow_from_directory` de la bibliothèque Keras, soit à partir d'un fichier Metadata csv/xls en utilisant l'ensemble de données (DataFrame) géré par le module Pandas. La fonction utilisée dans ce cas est `flow_from_dataframe` de la même bibliothèque Keras.

1.7.2 Normalisation des données

La normalisation permet de s'assurer que les données de l'image ont une échelle cohérente. C'est un élément crucial pour l'apprentissage efficace du modèle CNN. Les techniques de normalisation courantes consistent à soustraire l'intensité moyenne des pixels de chaque canal ou à mettre à l'échelle les valeurs des pixels dans une plage spécifique (par exemple 0 à 1 ou -1 à 1).

1.7.3 Augmentation des données

L'augmentation des données est une tâche facultative consistant à élargir l'ensemble des images en créant des variations de ces dernières. Cela permet au modèle d'apprendre des caractéristiques qui résistent à de légères modifications telles que la rotation, la mise à l'échelle ou le retournement. Elle peut améliorer les performances de généralisation du modèle et réduire le sur-apprentissage (overfitting) [41].

La transformation des données peut consister à faire pivoter les images, à les décaler, à les retourner, à les agrandir, etc. En appliquant cette transformation, nous obtenons des images plus diversifiées dans les lots dans l'objectif de disposer beaucoup plus de données qu'à l'origine.

Pour l'augmentation de données type images, l'objet `ImageDataGenerator` de Tensorflow est utilisé [42] d'où les paramètres `shear_range`, `zoom_range` et `horizontal_flip`, etc. sont tous utilisés pour transformer les images. L'augmentation des données est une partie importante du traitement des images qui permet de [41] :

- Réduire le sur-apprentissage excessif en créant plusieurs versions de la même image.
- Augmenter le nombre d'images : Une seule image agit comme plusieurs images. L'ensemble de données comporte donc moins d'images, mais chaque image peut être convertie en plusieurs images grâce à l'augmentation du nombre d'images. Cette augmentation est utilisée pour équilibrer le jeu de données afin d'améliorer la performance du modèle et d'éviter qu'elle ne soit biaisée vers les classes majoritaires.

- Faciliter la prédiction de nouvelles images : Imaginons qu'une seule image soit observée sous différents angles et que chaque angle soit considéré comme une image distincte. Cela signifie que l'algorithme sera plus précis dans la prédiction de nouvelles images.

1.7.4 Division des données

La division des données pour les modèles ANN est une étape cruciale pour l'amélioration des processus d'entraînement. Elle consiste à diviser l'ensemble des données en sous-ensembles distincts.

Pour l'apprentissage et les tests de notre modèle, nous devrions avoir nos données réparties en trois ensembles de données distincts (voir figure 1.16) :

- L'ensemble d'entraînement (Training set) : Il s'agit de l'ensemble des données utilisées pour former le modèle et lui permettre d'apprendre les caractéristiques/modèles cachés dans les données. Le training set est utilisé pour l'apprentissage. À chaque itération, les mêmes données d'apprentissage (Training set) sont transmises vers le réseau neuronal de manière répétée, et le modèle continue d'apprendre les caractéristiques des données.
- L'ensemble de validation (Validation set) : C'est un ensemble de données, fractionné à partir de l'ensemble training set, qui est utilisé pour valider les performances de notre modèle pendant l'entraînement. Ce processus de validation fournit des informations qui nous aident à ajuster les hyperparamètres et les configurations du modèle. En conséquence, cela peut nous dire si l'entraînement va dans la bonne direction ou non. Le modèle est entraîné sur l'ensemble training set et, simultanément, évalué sur l'ensemble validation set après chaque itération. D'autre part, la division de l'ensemble de données en un ensemble de validation set permet d'empêcher le sur-apprentissage, c'est-à-dire que le modèle devient très bon pour classer les échantillons de l'ensemble d'apprentissage training set mais ne peut pas se généraliser et effectuer des classifications précises sur les données qu'il n'a pas vues auparavant.
- L'ensemble de test (Test set) : L'ensemble de test est un ensemble distinct de données utilisé pour tester et évaluer les performances globales du modèle, après avoir terminé l'entraînement. Il fournit les métriques des performances du modèle final en termes d'accuracy, de précision, etc.

La création de différents ensembles de données nous permet d'ajuster les performances réelles du modèle. Le pourcentage et le processus de division de l'ensemble de données dépendent du nombre d'échantillons présents dans les données et du modèle.

La pratique courante en matière d'apprentissage est d'effectuer une division 70 : 15 : 15% ou 60 : 20 : 20% pour les ensembles training set, validation set, et test set.

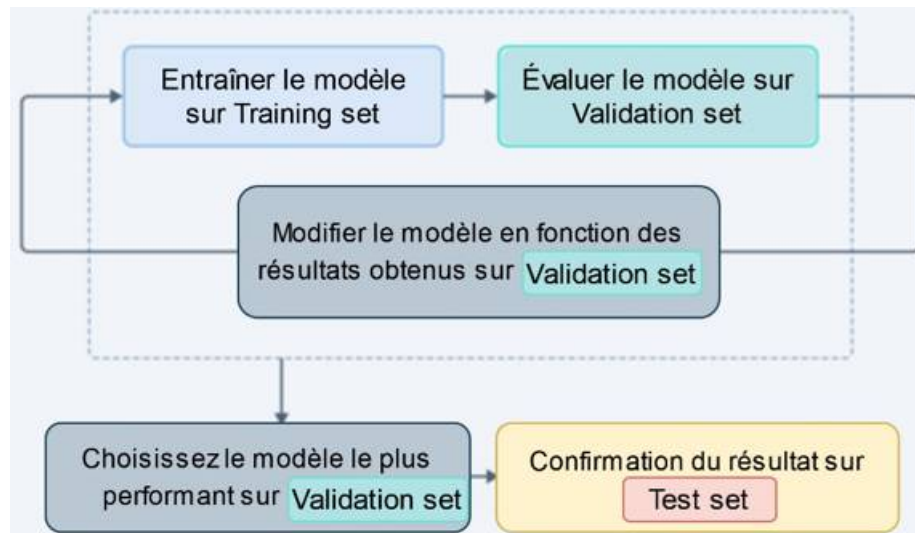


Figure 1.16 – Division des données training set, validation set, et test set.

1.8 Métriques d'évaluation

Il est important de pouvoir évaluer nos modèles efficacement, car elle ne se limite pas qu'aux performances techniques, mais prend également en compte le contexte global de notre problématique. Les métriques d'évaluation courantes pour les tâches de classification comprennent l'exactitude (Accuracy), la précision, le rappel (Recall) appelé aussi sensibilité (ou taux de vrais positifs, TPR), spécificité (taux de vrais négatifs TNR), le taux de faux positifs (FPR), le taux de faux négatif (FNR), le F1-score, la courbe caractéristique du récepteur (ROC) et l'aire sous la courbe (AUC). Toutes ces métriques sont récapitulées dans le tableau 1.7.

Métrique	Formule	Description
Exactitude/ Accuracy	$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$	C'est le nombre de prédictions correctes divisé par le nombre total de prédictions. Pourcentage de prédictions correctes parmi toutes les prédictions.
Précision	$Precision = \frac{TP}{TP+FP}$	Pourcentage de prédictions positives correctes parmi toutes les prédictions positives. Évaluer l'exactitude des prédictions positives faites par un modèle.
Recall/ Sensibilité/ Taux de vrais positifs (TPR)	$Recall = \frac{TP}{TP+FN}$	Pourcentage des instances de vrais positifs parmi tous les échantillons qui sont réellement positifs.
Spécificité /Taux de vrais négatifs (TNR)	$Specificity = \frac{TN}{TN+FP}$	Pourcentage des instances de vrais négatifs parmi tous les échantillons qui sont réellement négatifs.
Taux de faux positifs (FPR)	$FPR = \frac{FP}{TN+FP}$	Pourcentage des échantillons négatifs incorrectement classés comme positifs.
Taux de faux négatifs (FNR)	$FNR = \frac{FN}{FN+TP}$	Pourcentage des échantillons positifs incorrectement classés comme négatifs.
F1 Score	$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$	Moyenne harmonique entre la précision et le rappel, donnant une mesure équilibrée des deux métriques.
ROC	Pas de formule unique, mais la courbe est tracée avec ROC = Taux de Vrais Positifs (TPR) vs. Taux de Faux Positifs (FPR)	Courbe qui trace le TPR (Recall) contre le FPR pour différents seuils de classification.
AUC	$AUC = \int_{-\infty}^{\infty} f(x) dx$	Mesure de la capacité du modèle à distinguer entre les classes, en particulier dans les courbes ROC.

Tableau 1.7 – Les principales métriques de classification [34], [35]

D'autre part l'évaluation de la qualité des classifications est fondée sur une matrice de confusion et les quatre instances suivantes [43] :

- **Vrai Positif (TP)** : Il s'agit du nombre de résultats qui étaient initialement positifs et qui ont été prévus comme étant positifs (classe 1 par exemple).
- **Faux Positif (FP)** : Il s'agit du nombre de résultats qui étaient initialement négatifs, mais qui ont été prédits positifs. Cette erreur est également appelée une erreur de type 1.
- **Vrai Négatif (TN)** : Il s'agit du nombre de résultats initialement négatifs qui ont été correctement prédits comme négatifs (classe 0 par exemple).
- **Faux Négatif (FN)** : il s'agit du nombre de résultats qui étaient initialement positifs, mais dont la prévision était négative. Cette erreur est aussi appelée une erreur de type 2.

Le but est de maximiser les valeurs des TN et TP, et de minimiser les valeurs des FN et FP.

La matrice de confusion est un outil indispensable pour évaluer la performance d'un modèle de classification. Elle nous permet de visualiser les erreurs commises par le modèle et de calculer plus précisément les métriques. Le tableau 1.8 montre une matrice de confusion et illustre comment les prédictions du modèle se comparent aux valeurs réelles.

	Prédit Négatif	Prédit Positif
Réel Négatif	TN	FP
Réel Positif	FN	TP

Tableau 1.8 – *Matrice de confusion.*

1.9 Implémentation d'un réseau CNN

L'implémentation d'un réseau CNN (Convolutional Neural Network) peut être réalisée en suivant plusieurs étapes clés, en utilisant des bibliothèques comme Keras, TensorFlow ou PyTorch. Dans cette activité, nous avons travaillé avec TensorFlow et Keras, avec l'utilisation du réseau pré-entraîné VGG16.

1.9.1 Identifier une image par le modèle pré-entraîné VGG

Nous avons un exemple d'image d'une pierre. Nous allons utiliser le réseau pré-entraîné VGG16 [31] pour traiter et identifier l'image. Voici comment procéder étape par étape :



Figure 1.17 – Image de pierre à identifier par VGG16.

Listing 1.2 – Importer les bibliothèques

```
import numpy as np
from keras.applications.vgg16 import VGG16
from keras.preprocessing import image
from keras.applications.vgg16 import preprocess_input,
    decode_predictions
from keras.preprocessing import image
```

La dernière couche de prédictions (Dense) comporte 1 000 valeurs. Cela signifie que le VGG16 a un total de 1 000 classes et que notre image sera l'une de ces 1 000 classes.

Listing 1.3 – Charger le modèle VGG

```
vgg_model = VGG16(weights='imagenet', include_top=True, input_shape
    =(224, 224, 3))
```

Bien que VGG16 soit conçu pour des images en couleurs de taille 224x224, nous allons adapter l'image de la figure 1.17 de 200x200 en niveaux de gris à la taille requise par VGG16. Nous allons charger l'image par la fonction `image.load_img` depuis le chemin du système, ensuite transformer l'image en un vecteur par l'utilisation de la fonction `img_to_array`. Ensuite, l'image doit être en quatre dimensions pour que le VGG16 puisse la traiter. Nous développons la dimension de l'image par la fonction `np.expand_dims`. Enfin, il faut faire un prétraitement de l'image à l'aide de la fonction `preprocess_input`.

Listing 1.4 – Traitement de l'image

```
stone_image= image.load_img('.././stone.jpg',target_size=(224, 224)
    )
stone_image
stone_array = image.img_to_array(stone_image)
stone_array.shape #dimension afficher est (224,224,3)
stone_array = np.expand_dims(stone_array, axis=0)
stone_array.shape #dimension afficher est (1,224,224,3)
stone_array = preprocess_input(stone_array)
```

Lancer et créer la variable de prédiction. La dimension de la variable prédictive **pred** doit être (1,1000), car la base de données ImageNet comporte 1000 catégories d'images appelé aussi classe. La variable prédictive indique la probabilité que notre image soit l'une de ces images.

Une fois la variable prédictive **pred** est créée, on va mettre les prédictions sous une forme lisible. L'affichage de l'étiquette la plus probable se fait à partir du résultat de la fonction **decode_predictions**, le résultat le plus probable est obtenu, par exemple, la probabilité la plus élevée.

Listing 1.5 – Prédiction

```
pred = vgg_model.predict(stone_array)
pred.shape #dimension afficher est (1,1000)
label = decode_predictions(pred)
decoded_label = label[0][0]
# La classification affiche:
print('%s (%.2f%%)' % (decoded_label[1], decoded_label[2]*100 ))

# affichage: stone wall (99.73%)
```

Dans cette activité, nous avons prédit une image qui dit (avec une probabilité de 99.73 %) que l'image est un mur en pierre. Il est clair qu'une plus grande précision signifie qu'un objet relativement similaire à notre image est présent dans la base de données ImageNet, et que le modèle VGG16 a réussi à identifier l'image.

1.9.2 L'ajustement fin sur le modèle VGG16

Avant de commencer, il faut y avoir un ensemble de données d'images (dataset) dans notre répertoire de travail. Nous avons utilisé un Dataset appelé `car_flower` [44].

L'ensemble de donnée a besoin d'être divisé en trois dataset (dossiers), un dossier d'entraînement `training_set` et d'un dossier de validation `validation_set` pour valider notre modèle. Chacun de ces dossiers contiendra des images de voitures, et un dossier contenant des images de fleurs.

Nous travaillons à l'ajustement fin (**Fin_tuning**) du modèle VGG16, pour cela nous allons geler le réseau et supprimer la dernière couche de VGG16, qui contient 1000 classes. Après avoir supprimé la dernière couche, nous construirons un nouveau classificateur ANN adapté à notre dataset `car_flowers` c-à-d. pour les voitures et fleurs. Ainsi, quelle que soit l'image que nous fournissons au modèle, il la classera comme une fleur ou une voiture en fonction de sa probabilité de prédiction. En fait, nous allons remplacer la dernière couche du VGG16 par une couche définie par nous-même.

Le code pour ces étapes est le suivant :

Listing 1.6 – Importer les bibliothèques

```
import numpy as np
from keras.applications.vgg16 import VGG16
from keras.preprocessing import image
from keras.applications.vgg16 import preprocess_input,
    decode_predictions
from keras.preprocessing.image import ImageDataGenerator
from keras.preprocessing import image
import keras
from keras.layers import Dense
from tensorflow import random
```

Listing 1.7 – Lancer le modèle

```
Vgg_model = VGG16(weights='imagenet', include_top=True, input_shape
    =(224, 224, 3))
vgg_model.summary()
```

Lorsque nous exécutons l'instruction `vgg_model.summary()`, le résumé du modèle s'affichera. Le résumé montre les couches du modèle VGG16, y compris les Dense FC layers et la couche de sortie, comme montre la figure 1.18.

Model: "vgg16"

Layer (type)	Output Shape	Param #
input_layer (InputLayer)	(None, 224, 224, 3)	0
block1_conv1 (Conv2D)	(None, 224, 224, 64)	1,792
block1_conv2 (Conv2D)	(None, 224, 224, 64)	36,928
block1_pool (MaxPooling2D)	(None, 112, 112, 64)	0
block2_conv1 (Conv2D)	(None, 112, 112, 128)	73,856
block2_conv2 (Conv2D)	(None, 112, 112, 128)	147,584
block2_pool (MaxPooling2D)	(None, 56, 56, 128)	0
⋮	⋮	⋮
flatten (Flatten)	(None, 25088)	0
fc1 (Dense)	(None, 4096)	102,764,544
fc2 (Dense)	(None, 4096)	16,781,312
predictions (Dense)	(None, 1000)	4,097,000

Figure 1.18 – Affichage du résumé du modèle VGG16 après son lancement (les trois dernières couches).

Si nous voulons supprimer la dernière couche qui contient 1000 classes de prédictions dans le modèle `model_vgg`, il faut créer un nouveau modèle appelé modèle de la classe séquentielle et itérer à travers toutes les couches du modèle VGG. Ajoutez-les toutes les couche, à l'exception de la dernière couche. Après nous allons geler les couches en itérant à travers les couches et en réglant le paramètre `trainable` sur `False`. Enfin, nous ajoutons une nouvelle couche de sortie de taille 1 (classe binaire) avec une fonction d'activation sigmoïde et imprimons le résumé du modèle. La Figure 1.13 montre la sortie du code suivant :

Listing 1.8 – Nouveau modèle (classifier)

```
last_layer = str(vgg_model.layers[-1])
np.random.seed(42)
random.set_seed(42)
```

```

model= keras.Sequential()

for layer in vgg_model.layers:
if str(layer) != last_layer:
model.add(layer)
#Geler les couche
for layer in model.layers:
layer.trainable=False

#Ajouter la couche sigmoid
model.add(Dense(1, activation='sigmoid'))
model.summary()

```

Le résumé du modèle donné par la figure 1.19 montre les couches du modèle VGG16, y compris la nouvelle couche Dense avec la fonction d'activation sigmoid.

Model: "sequential"

Layer (type)	Output Shape	Param #
block1_conv1 (Conv2D)	(None, 224, 224, 64)	1,792
block1_conv2 (Conv2D)	(None, 224, 224, 64)	36,928
block1_pool (MaxPooling2D)	(None, 112, 112, 64)	0
block2_conv1 (Conv2D)	(None, 112, 112, 128)	73,856
⋮	⋮	⋮
flatten (Flatten)	(None, 25088)	0
fc1 (Dense)	(None, 4096)	102,764,544
fc2 (Dense)	(None, 4096)	16,781,312
dense (Dense)	(None, 1)	4,097

Figure 1.19 – Résumé du modèle créé après l'ajout de la nouvelle couche.

Maintenant, on doit compiler le réseau avec un optimiseur **adam** et une fonction de perte **binary cross-entropy** et calculer la métrique exactitude (**accuracy**) pendant l'entraînement.

Listing 1.9 – Compilation du modèle

```

model.compile(optimizer='adam', loss='binary_crossentropy', metrics
=[ 'accuracy' ])

```

Dans cette étape, nous allons préparer nos données Car_Flower [45], et créer des générateurs de données d'entraînement et de validation, comme nous l'avons vu précédemment. Nous normalisons les images d'entraînement et de validation de 1/255 de façon que toutes les valeurs soient comprises entre 0 et 1.

Ensuite, nous créons un jeu de donnée d'entraînement à partir du dossier « training set ». Notre modèle CNN a une taille d'image de 224x224, donc la même taille doit être

passée pour toutes les image, le paramètre `batch_size=64` est le nombre d'images dans un seul lot, et le paramètre `class_mode` est binaire puisque nous créons un classificateur binaire (deux classes voiture ou fleurs).

Listing 1.10 – Préparation des donnée training et validation set

```
generate_train_data = ImageDataGenerator(rescale = 1./255)
generate_validation_data = ImageDataGenerator(rescale =1./255)
training_dataset = generate_train_data.flow_from_directory(
    ' path /training_set',
    target_size = (224, 224),
    batch_size = 64,
    class_mode = 'binary')
validation_dataset = generate_test_data.flow_from_directory(\
    ' path /validation_set',
    target_size = (224, 224),
    batch_size = 64,
    class_mode = 'binary')

#Le nombre d images de training set est : 10764
#Le nombre d images de validation set est : 2674

nb_train_steps = tarin_img_number//batch_size
nb_val_steps = validation_img_number//batch_size
nb_epochs=25
```

Listing 1.11 – Lancer l'entraînement du modèle

```
history=model.fit(
    training_set, steps_per_epoch = nb_train_steps,
    epochs = nb_epochs,
    validation_data = validation_set,
    validation_steps = nb_val_steps)
```

Les résultats de l'apprentissage la perte (Loss) et l'exactitude (Accuracy) de notre modèle sont montrés par la figure 1.20.

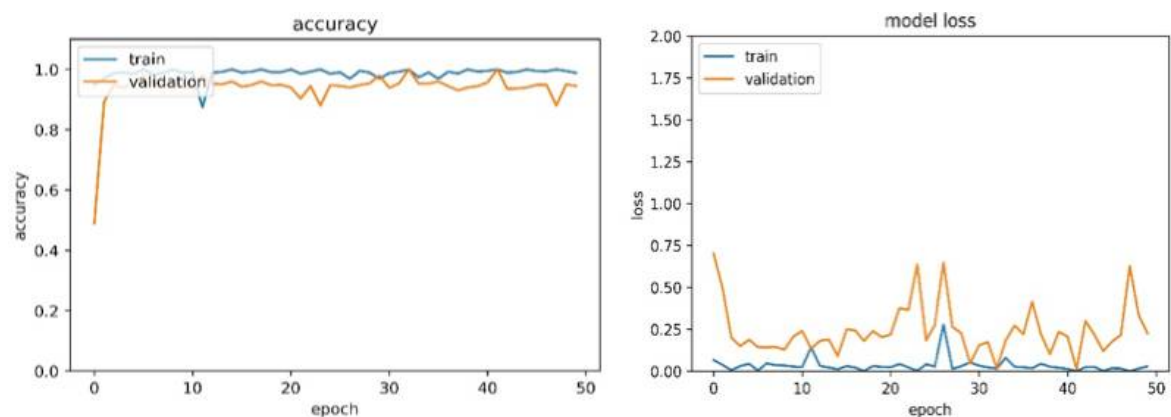


Figure 1.20 – Accuracy et perte de l'entraînement et validation du modèle VGG16.

Pour tester notre modèle et prédire une nouvelle image (le code est le même qui a été utilisé dans la section I.9.1., identifier une image par le modèle CNN). Tout d'abord, chargez l'image à partir de dossier `'... path/ Prediction/test_image.jpg'` et

définissez la taille cible sur (224, 224) puisque le modèle VGG16 accepte les images avec cette dimension.

La méthode de prédiction du classificateur et l’affichage de la sortie dans un format lisible sont donnés par le listing suivant :

Listing 1.12 – Résultat de prédiction

```
result = model.predict(test_image)
if result[0][0] == 1:
    prediction = 'It is a flower'
else:
    prediction = 'It is a car'
print(prediction)
```

1.9.3 L’ajustement fin sur VGG avec des couches ANN personnalisées

Ce processus vous permet d’ajuster finement un modèle VGG16 en ajoutant des couches ANN personnalisées pour une tâche de classification binaire. Nous adaptons les hyperparamètres et les données d’entraînement en utilisant les images du même dataset (car_flower) que dans la partie précédente.

Listing 1.13 – Charger le modèle

```
Vgg_model = VGG16(weights='imagenet', include_top=False, input_shape
    =(224, 224, 3))
vgg_model.summary()
```

Le paramètre **include_top=False** permet de supprimer toutes les couches ANN, y compris la couche de flattening, du modèle pré-entraîné VGG parce que nous utilisons une couche de convolution pour l’extraction des caractéristiques et que nous voulons former des couches entièrement connectées pour la classification de nos images.

Maintenant, on va procéder à geler les couches du modèle VGG16. Cela permet de conserver les caractéristiques apprises par le modèle VGG16.

Listing 1.14 – Geler les couches

```
for layer in vgg_model.layers:
    layer.trainable = False
```

- Ensuite nous avons ajouté des couches ANN personnalisées :
- **Flatten()** : Aplati la sortie des convolutions pour qu’elle puisse être utilisée par les couches denses.
- **Dense(256, activation='relu')** : Ajoute quatre couches dense avec 256 unités et une fonction d’activation ReLU (Rectified Linear Unit), qui aide à introduire la non-linéarité.
- **Dense(1, activation='sigmoid')** : Ajoute la couche de sortie avec une seule unité et une fonction d’activation sigmoid pour la classification binaire. La sortie sera une probabilité comprise entre 0 et 1 (voiture ou fleur).

Listing 1.15 – Ajouter des couches ANN personnalisées

```
x = vgg_model.output
x = Flatten()(x)
```

```
x = Dense(256, activation='relu')(x)
x = Dense(256, activation='relu')(x)
x = Dense(256, activation='relu')(x)
x = Dense(256, activation='relu')(x)
output = Dense(1, activation='sigmoid')(x)
```

Enfin, nous créons le modèle final comme le montre la figure 1.21.

Listing 1.16 – Création du modèle final

```
model = Model(inputs=base_model.input, outputs=output)
model.summary()
```

Model: "sequential"

Layer (type)	Output Shape	Param #
block1_conv1 (Conv2D)	(None, 224, 224, 64)	1,792
block1_conv2 (Conv2D)	(None, 224, 224, 64)	36,928
block1_pool (MaxPooling2D)	(None, 112, 112, 64)	0
block2_conv1 (Conv2D)	(None, 112, 112, 128)	73,856
⋮	⋮	⋮
flatten (Flatten)	(None, 25088)	0
dense (Dense)	(None, 256)	6,422,784
dense_1 (Dense)	(None, 256)	65,792
dense_2 (Dense)	(None, 256)	65,792
dense_3 (Dense)	(None, 256)	65,792
dense_4 (Dense)	(None, 1)	257

Figure 1.21 – Résumé du modèle après l'ajout des couches ANN personnalisées.

Nous suivons ensuite les mêmes étapes que dans la section I.9.2, concernant la préparation des données pour l'ensemble training set et validation set, la compilation et l'entraînement du modèle. On obtient ainsi la courbe de précision suivante montrée par la figure 1.22 :

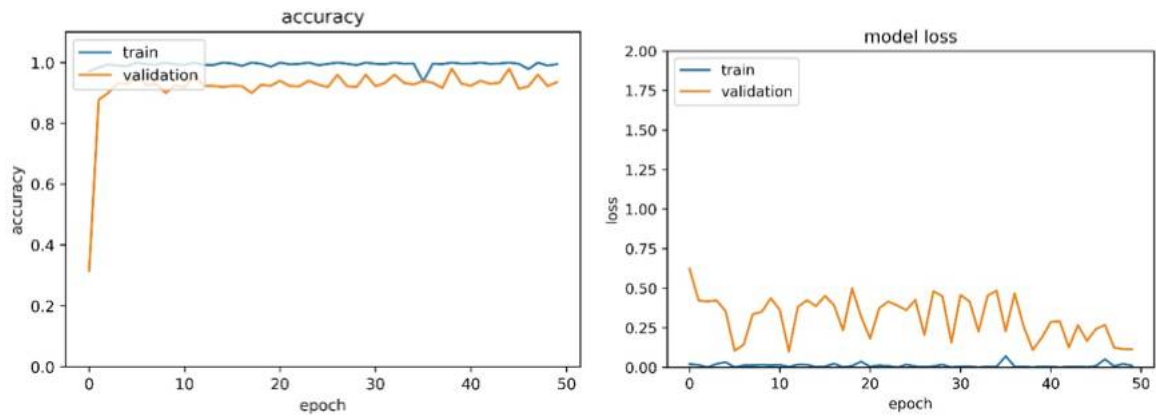


Figure 1.22 – *Accuracy et perte de l’entraînement et validation du modèle VGG16 personnalisé.*

1.10 Conclusion

L’intelligence artificielle est tout simplement un outil informatique doté d’un mécanisme permettant de simuler l’intelligence humaine. Une fois que cet outil est construit à l’aide de divers algorithmes et modules d’apprentissage automatique, il peut lire des images, du texte, de la vidéo ou de l’audio.

En outre, l’apprentissage automatique est une technique basée sur l’intelligence artificielle pour développer un système informatique intelligent. De nos jours, la technologie fonctionne grâce à l’intelligence artificielle et à l’apprentissage automatique. L’apprentissage automatique est un outil de traitement des données et d’analyse très puissant, capable de traiter des ensembles de données vastes et étendus, structurés ou non. Nous avons consacré ce chapitre à la présentation de notions de base telles que l’IA, l’apprentissage automatique et ses types et nous concluons avec l’apprentissage profond où nous avons montré une vue d’ensemble du perceptron et de son type avec son fonctionnement. En outre ce chapitre a fourni des détails sur les fondements théoriques pertinents nécessaires pour comprendre le DL, y compris les architectures existantes basées sur les CNN, ainsi que les modèles pré-entraînés.

Dans le chapitre suivant, nous allons présenter les maladies du cancer de la peau et plus particulièrement le mélanome.

Background et généralités sur le cancer de la peau : méthodes de diagnostic et travaux existants sur le diagnostic du mélanome par Deep Learning

Sommaire

2.1	Introduction	45
2.2	Présentation des lésions cutanées	45
2.3	Cancer de la peau	47
2.3.1	Cancer de la peau non- mélanome	47
2.3.1.1	Facteurs de risque du cancer de la peau non- mélanome	48
2.3.1.2	Diagnostic des cancers de la peau non- mélanome	48
2.3.2	Cancer de la peau mélanome malin	50
2.3.2.1	Types de mélanome malin cutané	50
2.3.2.1.1	Mélanome à propagation superficielle	50
2.3.2.1.2	Mélanome nodulaire	50
2.3.2.1.3	Mélanome lentigo maligne	52
2.3.2.1.4	Mélanome acral-lentigineux	52
2.3.2.2	Etiologie (étude des causes)	53
2.3.2.3	Les symptômes	53
2.3.2.4	Prévention du mélanome malin	54
2.3.2.5	Méthodes de diagnostic du mélanome malin	54
2.3.2.5.1	La règle ABCDE	55
2.3.2.5.2	La liste de contrôle en 3 points	55
2.3.2.5.3	La liste de contrôle en 7 points	55
2.3.2.5.4	La méthode de Menzies	56
2.3.2.5.5	L'analyse des formes	56
2.3.3	Techniques d'imagerie du cancer de la peau	56
2.3.3.1	La microscopie confocale par réflectance	57
2.3.3.2	La Tomographie en Cohérence Optique	57
2.3.3.3	L'échographie	57
2.3.3.4	La dermoscopie	57
2.4	Utilisation de l'intelligence artificielle pour le diagnostic du mélanome malin	58

2.4.1	Les ensembles de données (datasets)	59
2.4.2	Les travaux existants : Méthodes de Deep Learning appliquées au diagnostic du mélanome	60
2.5	Conclusion	64

2.1 Introduction

Le cancer est l'une des principales causes de décès des êtres humains. Selon les statistiques de l'organisation mondiale de la santé, le cancer devrait être la première cause de décès (13,1 millions) d'ici 2030 [46, 47]. Parmi tous les types de cancer, le cancer de la peau est la forme de cancer la plus répandue aux États-Unis [48]. Selon les prévisions, 20 % des Américains développeront un cancer de la peau au cours de leur vie [49]. Le cancer de la peau n'est pas forcément mortel. Cependant, le diagnostic à un stade précoce joue un rôle essentiel pour sauver des vies. Pour comprendre la détection et le diagnostic précoces du cancer de la peau, il est important d'examiner la peau humaine et les différents types de cancers de la peau.

Le mélanome, une des formes les plus graves de cancer de la peau, représente un type de cancer particulièrement redouté. Les facteurs de risque possibles sont : l'exposition à la lumière du soleil ou aux rayons UV, les infections virales, l'exposition à l'arsenic, la suppression immunitaire et certaines conditions génétiques. Au début de la guérison du mélanome, la chirurgie était la seule méthode disponible pour traiter le cancer, puis, au fil du temps, la biopsie du ganglion lymphatique sentinelle, l'excision large, les inhibiteurs de points de contrôle immunitaire, les médicaments de thérapie ciblée et la chimiothérapie ont été utilisés pour traiter le mélanome.

Dans la médecine avancée, c'est un péché d'utiliser les méthodes ou approches traditionnelles pour guérir ou atténuer la gravité du cancer. Grâce aux progrès de la technologie dans le monde médical, les médecins et les experts médicaux utilisent divers équipements médicaux pour diagnostiquer rapidement le cancer de la peau et en arrêter la gravité.

Les systèmes de diagnostic assisté par ordinateur (CAD) les plus récents ont été utilisés pour identifier et diagnostiquer les mélanomes malins. Ce chapitre est divisé en trois parties : la première décrit les lésions cutanées, la deuxième explique les différents types de cancers de la peau et la troisième se concentre sur les techniques de diagnostic et assisté par ordinateur et l'intelligence artificielle.

2.2 Présentation des lésions cutanées

Une lésion cutanée est une partie de la peau qui présente une croissance anormale et dont l'aspect diffère de celui de la peau à proximité, notamment la couleur, la forme, la taille et la texture. Les lésions cutanées sont très courantes et apparaissent souvent à la suite d'une atteinte localisée de la peau, comme les coups de soleil ou la dermatite de contact. Cependant, d'autres peuvent être des manifestations de troubles sous-jacents, tels que les infections, le diabète et les troubles auto-immuns ou génétiques. Bien que la plupart des lésions cutanées soient bénignes et inoffensives, certaines d'entre elles peuvent être malignes ou précancéreuses, c'est-à-dire qu'elles ont le potentiel d'évoluer vers un cancer de la peau.

Les lésions sont généralement classées comme suit :

- Les lésions primaires : Apparaissent sur une peau précédemment saine et sont directement associées à une cause spécifique. Les taches de rousseur, les grains de beauté et les cloques, entre autres, sont des exemples courants de lésions cutanées primaires, qui peuvent apparaître à la naissance ou au cours de la vie.
- Les lésions secondaires : Se développent à partir de l'évolution d'une lésion cutanée primaire, soit à la suite d'une manipulation traumatique, telle que le grattage ou le frottement, soit en raison de son traitement ou de sa progression. Les croûtes, les plaies, les ulcères et les cicatrices sont des exemples de lésions cutanées secondaires.

Les lésions cutanées peuvent se présenter sous différentes tailles, formes et aspects (voir figure 2.1). En outre, les lésions cutanées peuvent être isolées ou groupées, localisées sur une seule zone ou étendues à l'ensemble du corps. Les risques liés aux différentes catégories de lésions cutanées sont les nævus congénitaux et les nævus atypiques. Les enfants nés avec des grains de beauté sont peu fréquents, de l'ordre de 1 à 2 %. Le risque que le nævus congénital se transforme en mélanome est faible, est d'environ 5 %.

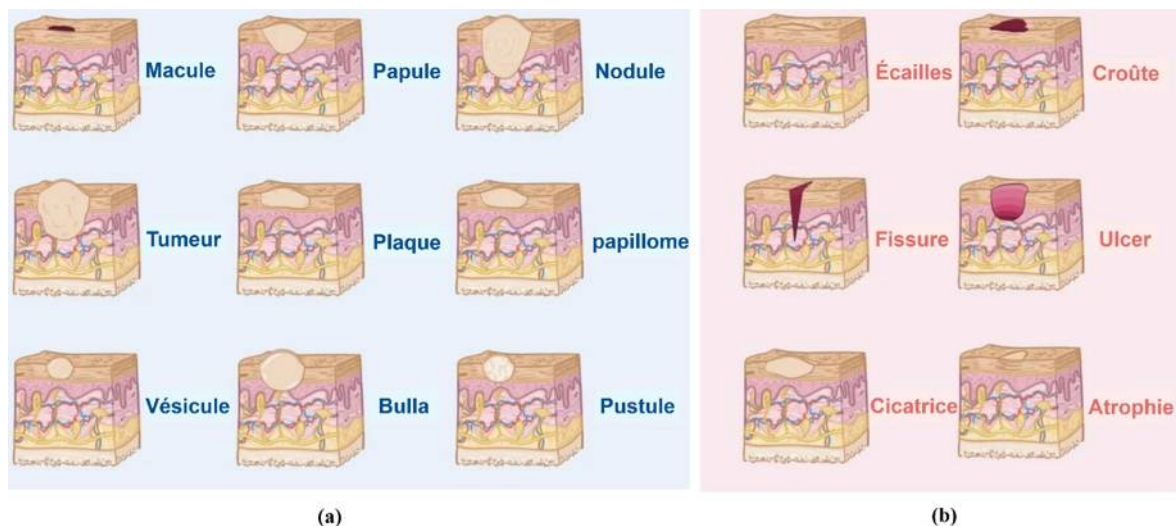


Figure 2.1 – Type de lésion cutanée (a) Les lésions primaires, et (b) Les lésions secondaires.

Les grains de beauté qui y sont associés sont de grande taille, de forme irrégulière et de couleur allant du rose au brun foncé, ce qui représente des nævus atypiques ou des nævus dysplasiques.

Une lésion cutanée bénigne est une anomalie, une excroissance ou une tumeur cutanée non cancéreuse qui peut apparaître n'importe où sur le corps. Les lésions bénignes peuvent se manifester de différentes manières, en fonction de leur cause et de leur tissu d'origine. Les lésions cutanées bénignes courantes comprennent la plupart des nævus mélanocytaires, mieux connus sous le nom de grains de beauté, les kératoses séborrhéiques, les taches cutanées, les angiomes cerise et les lipomes, entre autres. La plupart du temps, ces lésions sont inoffensives et ne nécessitent pas de traitement, à moins qu'elles ne provoquent des symptômes tels qu'une gêne ou des démangeaisons. Contrairement aux lésions malignes, les lésions cutanées bénignes sont généralement symétriques, bien circonscrites, d'aspect uniforme, stables ou évoluant lentement dans

le temps. Toutefois, dans certains cas, il peut être difficile de distinguer les lésions bénignes des lésions malignes ; dans ces cas, une biopsie ou une ablation chirurgicale de la zone affectée peut être effectuée pour écarter la possibilité d'une malignité. Une lésion cutanée maligne est, par définition, un cancer de la peau.

Les deux principaux types de cancer de la peau sont le carcinome kératinocytaire et le mélanome. Chaque type de cancer de la peau présente des caractéristiques uniques, mais les signes généraux d'un cancer de la peau peuvent inclure des lésions cutanées à croissance rapide, des changements de couleur ou de taille d'une lésion préexistante, ou une plaie croûteuse qui ne guérit pas avec le temps.

2.3 Cancer de la peau

Il existe deux principaux types de cancer de la peau classés selon les cellules où ils se développent : carcinome (non-mélanome) et mélanome.

2.3.1 Cancer de la peau non- mélanome

Le terme cancer de la peau non- mélanome est couramment utilisé pour désigner les carcinomes basocellulaires (CBC) et les carcinomes spinocellulaires (CSC). Il définit ainsi une population de patients atteints de cancer en fonction de ce qu'ils n'ont pas, ce qui empêche une étude adéquate et minimise l'importance de cette maladie. Le carcinome basocellulaire (CBC), également appelé épithélioma basocellulaire ou ulcère de rongeur, est le type de cancer de la peau le moins agressif. Il s'agit de la tumeur maligne la plus fréquente chez l'homme [50]. Ce type de cancer se développe à partir des cellules basales qui tapissent la couche la plus profonde de l'épiderme. Une croissance anormale de cette couche est ce que l'on appelle un CBC.

Il apparaît généralement dans les zones d'exposition chronique au soleil, telles que le visage, les oreilles, le cou, le cuir chevelu et les épaules, mais peut également se développer dans le dos ou sur les jambes. Le BCC a généralement une croissance lente et se métastase très rarement, bien qu'il puisse provoquer une destruction locale cliniquement significative et une défiguration s'il est négligé ou traité de manière inadéquate. Il ne met généralement pas la vie en danger. La prévision est excellente avec un traitement approprié.

En outre, le cancer de la peau non- mélanome englobe également d'autres tumeurs malignes cutanées, notamment le lymphome cutané, le carcinome à cellules de Merkel, la maladie de Paget, les angiosarcomes et les histiocytomes malins [51].

Le carcinome spinocellulaire (CSC) est la deuxième forme la plus courante de cancer de la peau et représente environ 20 % (1 sur 5) des cancers de la peau diagnostiqués [52].

Ce cancer commence dans les cellules appelées kératinocytes, qui se trouvent dans la partie supérieure de l'épiderme. Le CSC se développe le plus souvent dans les zones exposées au soleil, mais il peut également se développer dans les cicatrices, les zones de peau qui ont été brûlées dans le passé et les zones de peau qui ont été ulcérées pendant une longue période. Dans un petit nombre de cas, le CSC peut se développer dans les zones génitales. En outre, une inflammation chronique de la peau ou des conditions médicales qui suppriment le système immunitaire sur une période prolongée peuvent favoriser le développement du CSC.

Le CSC est plus agressif que le CBC car il a un taux de croissance plus rapide, des marges moins bien délimitées et un potentiel métastatique plus important [53]. Si 96 à 97 % des CSC sont localisés, le même pourcentage (96 à 97 %) des cas restants (3 à 4 %) peut se propager à d'autres parties du corps et l'issue est souvent fatale [52]. En raison du très grand nombre de cas enregistrés chaque année, les cancers de la peau autres que le mélanome constituent un problème de santé publique important.

2.3.1.1 Facteurs de risque du cancer de la peau non- mélanome

Des facteurs de risque multiples et connexes sont associés au développement des cancers de la peau non- mélanome [51]. Des causes génétiques et environnementales sont impliquées dans le processus de carcinogenèse en plusieurs étapes, y compris l'initiation de la tumeur, la promotion, la progression pré-maligne et la conversion maligne de cellules cutanées normales en lésions cutanées cancéreuses.

Parmi les facteurs de risque, nous indiquons :

- Rayonnement ultraviolet, phénotype de la peau et effet de latence et de dosage : La lumière du soleil est la cause externe la plus importante de cancers de la peau non- mélanome. Les rayons ultraviolets B (UVB), A (UVA) et C (UVC) peuvent provoquer des lésions de l'ADN dans les cellules normales de la peau. Bien que la plupart des études montrent que la principale partie du spectre solaire responsable du développement de ces cancers est l'UVB, et l'UVA est également un facteur de risque [54]. Les UVC sont absorbés par la couche d'ozone et n'atteignent pas la surface de la terre.
- Produits chimiques et tabac : L'exposition aux rayonnements ionisants, qu'elle soit iatrogène (traitement de l'acné, de la tinea capiteuse, .etc.), professionnelle (uranium dans les mines) ou accidentelle (explosion d'une bombe atomique), multiplie le risque de CBC [55]. Il a été rapporté que le tabac doublait le risque d'attendre le CSC [56]. Cependant, aucune relation claire n'a été trouvée entre le tabagisme et le risque de CBC.
- Personnes âgées : Il existe une relation claire entre la fonction du système immunitaire, le vieillissement et l'incidence des cancers de la peau non- mélanome. Cela suggère que l'immuno-sénescence affaiblit l'immunité cellulaire, inhibant la surveillance de la transformation maligne des cellules normales [57].

2.3.1.2 Diagnostic des cancers de la peau non- mélanome

Plusieurs facteurs [58] influencent l'évolution d'une maladie (voir figure 2.2). La manière la plus simple d'évaluer l'effet de la santé dans une population particulière, est d'estimer la morbidité (nombre de cas de maladies) et la mortalité associées. Il y a des indicateurs de santé plus complexes qui sont le malaise, le handicap et l'insatisfaction (qualité de vie).

Selon la figure 2.2, les facteurs diagnostiques des cancers de la peau non-mélanome peuvent être divisés en facteurs liés au cancer lui-même, au diagnostic, aux traitements, aux performances cliniques du clinicien traitant et aux facteurs liés au patient. Les facteurs liés au patient comprennent la proposition de diagnostic, l'acceptation du plan de traitement, l'observance et les performances physiques (comorbidités chroniques associées et mode de vie).

- Facteurs liés aux cancers de la peau non-mélanome : Ces facteurs incluent le type de couverture du cancer et le profil de risque de la tumeur. Pour les CSC,

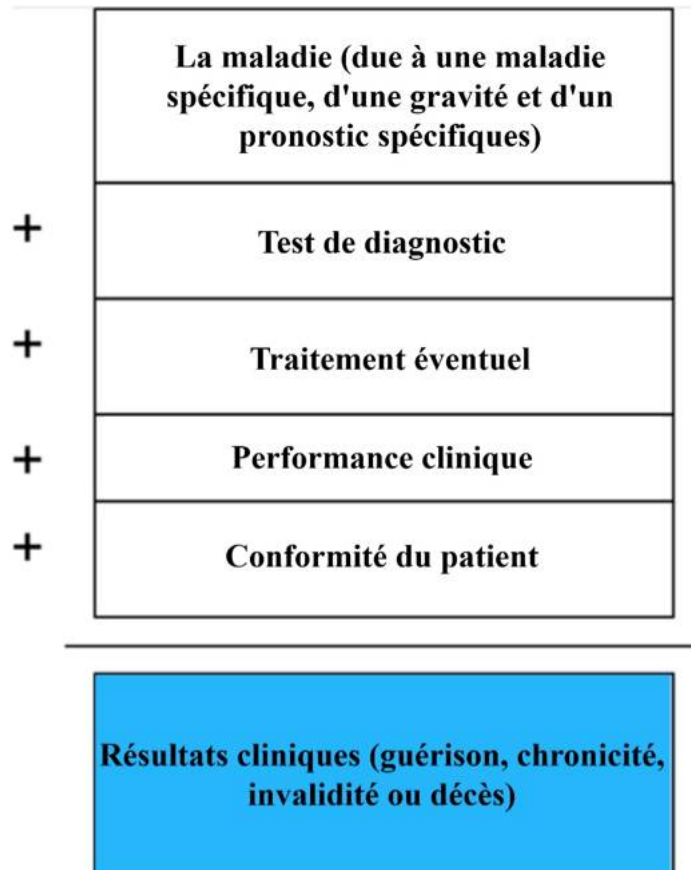


Figure 2.2 – *Facteurs influençant les conséquences d'une maladie.*

une mortalité plus élevée est observée chez les patients dont les tumeurs sont situées sur les lèvres et les oreilles, ainsi que dans les zones anogénitales. D'autres précurseurs à haut risque sont une tumeur de grande taille (supérieure à 2 cm), une caractérisation faible à modérée et une invasion profonde. Dans le cas du CBC, l'agressivité de la tumeur dépend également de la localisation, et de l'emplacement. Les tumeurs de la tête et du cou (en particulier celles situées sur l'oreille et la paupière) étant plus souvent métastatiques.

- **Facteurs de traitement** : Dans la prise en charge des patients atteints du cancer de la peau non-mélanome, il convient de prendre en considération le facteur du traitement. Il peut être modifié par les performances du clinicien traitant et par l'acceptation du patient à-propos le plan de traitement. Comme ces tumeurs sont considérées comme moins graves que d'autres en raison de leur croissance lente et de leur faible taux de métastases, un résultat cosmétique optimal est souvent plus important que l'ablation totale. Les conséquences en sont une morbidité accrue et, dans le pire des cas, la mortalité [59].
- **Facteurs liés au patient** : Parmi les facteurs liés aux patients qui influencent le diagnostic des cancers de la peau non-mélanome, sont des caractéristiques démographiques générales, telles que l'âge, le sexe et la race. La mortalité liée à ce type de cancer est plus élevée avec l'âge, le sexe masculin et dans la population blanche [59].

2.3.2 Cancer de la peau mélanome malin

Le mélanome malin est considéré comme le type dangereux de cancer de la peau qui se développe de manière incontrôlée à partir des cellules pigmentaires appelées « mélanocytes » et constitue une maladie potentiellement mortelle. Le diagnostic de mélanome malin à un stade préliminaire est difficile à établir car la plupart des caractéristiques partagées sont similaires aux caractéristiques cliniques d'un nævus atypique.

Bien que le mélanome malin se manifeste n'importe où dans le corps, y compris dans les organes internes, cette section se concentre uniquement sur les expériences de ceux qui ont développé un mélanome de la peau, souvent appelé mélanome malin cutané. Il existe quatre types fondamentaux de mélanome qui se distinguent par leur degré de gravité et leur localisation sur le corps.

2.3.2.1 Types de mélanome malin cutané

Tous les types mélanome malin présentent le même niveau de danger, en fonction de facteurs tels que la profondeur de la tumeur, la présence ou l'absence d'ulcération, le nombre de ganglions lymphatiques régionaux contenant le mélanome et l'étendue de la propagation du cancer dans les ganglions lymphatiques régionaux [60].

2.3.2.1.1 Mélanome à propagation superficielle Ce type de mélanome est le plus fréquent, environ 70 % appartiennent à ce type [61]. Comme son nom l'indique, il se développe à la surface de la couche supérieure pendant des mois, voire des années, avant de pénétrer plus profondément dans la peau. Comme l'illustre la figure 2.3, ce mélanome se présente sous la forme d'une lésion plate ou à peine surélevée, souvent avec des bords irréguliers et des variations de couleur. Les lésions apparaissent le plus souvent sur le torse, les jambes, et le haut du dos des deux sexes et elles sont plus fréquentes chez les personnes d'âge moyen [60].

Le premier signe d'un nouveau mélanome superficiel est la coloration plus foncée d'une partie d'un grain de beauté déjà existant ou l'apparition d'un nouveau grain de beauté sur une peau normale qui n'était pas touchée auparavant. Ce type de mélanome ne devient dangereux que lorsqu'il commence à se développer verticalement dans les couches plus profondes de la peau.

2.3.2.1.2 Mélanome nodulaire Le mélanome nodulaire est le type de mélanome le plus agressif et représente environ 25 % (environ 1 mélanome sur 4) de tous les mélanomes diagnostiqués [61].

Il se développe généralement plus rapidement en profondeur et en épaisseur que les autres types de mélanome et peut ne pas avoir de phase de développement radial facilement visible. Au lieu de se développer à partir d'un grain de beauté déjà présent, il peut se former à un endroit où il n'y avait aucune lésion auparavant.

Comme le montre la figure 2.4, le mélanome se présente généralement sous la forme d'un nodule bleu-noir en forme de bosse, bien que 5 % des lésions soient roses ou rouges et que 5 % soient amélanotiques (sans coloration), ce qui rend le diagnostic plus difficile. Le mélanome nodulaire est plus fréquent chez les hommes que chez les femmes, il apparaît plus couramment chez les personnes âgées de plus de 60 ans, mais il peut se développer à tout âge, et se trouve le plus souvent sur le torse ou la tête et le cou [60].



Figure 2.3 – *Mélanome malin à propagation superficielle.*



Figure 2.4 – *Mélanome malin nodulaire.*

2.3.2.1.3 Mélanome lentigo maligne Le mélanome lentigo maligne apparaît généralement sur la peau endommagée par le soleil chez les personnes d'âge moyen et les personnes âgées, particulièrement sur le visage. Dans ses premiers stades, ce type de mélanome peut être confondu avec une tache de vieillesse bénigne ou une tache de soleil, étant donné que le cancer peut être facilement confondu avec d'autres affections, il peut passer inaperçu pendant des années, ce qui peut être particulièrement dangereux. Il est souvent confondu avec le lentigo simplex, une tache brunâtre bénigne qui peut se développer chez les personnes âgées après des années d'exposition au soleil. Le mélanome lentigo maligne représente environ 10 % des mélanomes diagnostiqués au (soit environ 1 sur 10) [61]. La figure 2.5 est un exemple d'image de ce type de mélanome, qui commence par une tache étalée, plate, avec des bords irréguliers et des couleurs brunes variables.



Figure 2.5 – *Mélanome lentigo maligne.*

2.3.2.1.4 Mélanome acral-lentigineux Le mélanome acral-lentigineux est le mélanome le plus fréquent chez les Noirs et les Asiatiques. Il constitue 50 % des mélanomes chez les personnes ayant ce type de peau.

Acral vient du mot grec akron, qui signifie extrémité, et le mélanome apparaît généralement sur la paume de la main, la plante des pieds (comme dans la figure 2.6) ou sous les ongles. Il est beaucoup plus fréquent sur les pieds que sur les mains.

Les lésions sont généralement de couleur beige, brune ou noire, avec des variations de couleur et des bords irréguliers. Ce type de mélanome est souvent diagnostiqué plus tardivement que d'autres en raison de la croyance erronée selon laquelle les mélanomes ne se développent que sur les zones exposées au soleil et que les personnes noires et asiatiques sont moins susceptibles d'en être affectées [62]. Le fait de confondre les premiers signes du mélanome acral-lentigineux avec des ecchymoses ou des blessures

aux paumes, aux plantes des pieds ou aux lits des ongles peut également retarder le diagnostic.



Figure 2.6 – *Mélanome acral-lentigineux.*

2.3.2.2 Etiologie (étude des causes)

Le mélanome est une tumeur maligne dont l'étiologie (les causes) est à la fois endogène (interne) et exogène (externe). Le facteur exogène le plus important est l'exposition aux rayons ultraviolets (UV), traditionnellement assimilée à l'exposition au soleil, mais aussi, à d'autres sources artificielles, telles que les bancs solaires et les lampes à bronzer [63].

Les facteurs endogènes, dont le phototype de la peau (la quantité de pigments de mélanine dans la peau), le nombre de grains de beauté, les nævi mélanocytaires atypiques (grains de beauté inhabituels) et les antécédents familiaux de cancer de la peau, sont également des facteurs prédictifs importants du risque de mélanome [64].

2.3.2.3 Les symptômes

Les mélanomes peuvent avoir différents aspects. Ils peuvent être petits, brillants ou cireux, écailleux et rugueux, fermes et rouges, croûteux ou saignants, ou présenter d'autres caractéristiques ; c'est pourquoi la règle générale veut que tout ce qui est suspect soit examiné par un médecin.

Environ un tiers des mélanomes se développent à partir de grains de beauté existants et normaux et le reste à partir d'une peau antérieurement normale. Pour déterminer quelles lésions pigmentées nécessitent un examen plus approfondi, on peut utiliser la liste de contrôle en sept points [65] décrite ci-dessous (voir tableau 2.1).

Caractéristiques principales	Caractéristiques secondaires
Modification de la taille d'une lésion antérieure ou croissance évidente d'une nouvelle lésion.	Diamètre supérieur à 6 mm
Couleur irrégulière, diverses nuances de brun et de noir dans une lésion nouvelle ou ancienne.	Inflammation : rare dans les lésions bénignes, sauf si elles sont régulièrement traumatisées.
Forme irrégulière, asymétrie et contour irrégulier d'une lésion pigmentée nouvellement développée ou apparition de cette caractéristique dans une lésion existante.	Suintement, formation de croûtes ou saignement de la lésion.
	Modification de la sensation : généralement décrite comme une légère irritation.

Tableau 2.1 – *Liste de contrôle en sept points pour le mélanome malin.*

2.3.2.4 Prévention du mélanome malin

La prévention du cancer a été définie comme l'ensemble des mesures qui limitent la progression de la maladie à tout moment de son évolution. Les mesures de prévention peuvent intervenir à n'importe quel stade du processus de santé. Trois niveaux de prévention : primaire, secondaire et tertiaire ont été traditionnellement identifiés dans la littérature infirmière. La prévention primaire vise à éviter l'apparition du cancer, les activités de prévention secondaire se concentrent sur le diagnostic et l'intervention précoces et les activités de prévention tertiaire impliquent la réhabilitation et la restauration de niveaux optimaux de fonctionnement [66]. La prévention primaire et la prévention secondaire sont toutes deux importantes en ce qui concerne le mélanome malin.

Les mesures de la prévention secondaire comprennent des tentatives de détection précoce du mélanome chez les personnes asymptomatiques en les encourageant à adopter des pratiques telles que l'auto-examen de la peau et les examens annuels par des professionnels.

2.3.2.5 Méthodes de diagnostic du mélanome malin

L'objectif d'une détection clinique efficace du mélanome primitif reste primordial pour la réussite de la prise en charge de ce cancer.

La visualisation des lésions cutanées par l'une des techniques d'imagerie susmentionnées ne suffit pas à distinguer un mélanome malin d'un mélanome bénin. Il est nécessaire de disposer de techniques de diagnostic reproductibles qui puissent être utilisées par les cliniciens pour comprendre les types de cancer de la peau. Il existe quatre méthodes reproductibles communément acceptées pour le diagnostic des cancers de la peau, en particulier du mélanome. Il s'agit de la règle ABCD-E, la liste de contrôle en 3 points, la liste de contrôle en 7 points, la méthode de Menzies et

l'analyse des formes.

2.3.2.5.1 La règle ABCDE Cette méthode a été introduite en 1994 par [67]. ABCDE signifie Asymétrie, Bordure, Couleur, Diamètre et Elévation. Il s'agit de cinq critères d'image dermoscopique pour l'évaluation semi-quantitative des lésions cutanées. Les mélanomes sont généralement asymétriques, avec des bords irréguliers, et mesurent plus de 6 mm de diamètre, et ils ont généralement des couleurs mélangées et changent de taille, de couleur, de forme et saignent. Ces critères (à l'exception de E) ont leurs scores possibles en fonction de l'aspect de la lésion cutanée (voir tableau 2.2). Ces scores sont multipliés par les facteurs de pondération associés pour obtenir un score dermoscopique total (TDS).

Critères	Score possible	Description	poids
Asymétrie	0-8	Évaluer le contour, la couleur et les structures	1.3
Bordure	0-2	Fin brusque du schéma pigmentaire	0.1
Couleur	1-6	Présence de 6 couleurs au maximum (blanc, rouge, marron clair, marron foncé, bleu-gris, noir)	0.5
Diamètre	1-5	Présence de zones sans structure, de traces, de points et de globules	0.5

Tableau 2.2 – Critères de la règle ABCDE pour le calcul du score dermoscopique total (TDS).

Le TDS peut être calculé à l'aide de la formule ci-dessous :

$$TDS = (A \times 1.3) + (B \times 0.1) + (C \times 0.5) + (D \times 0.5) \quad (2.1)$$

Le résultat de la formule 2.1 peut être interprété selon le tableau 2.3 :

2.3.2.5.2 La liste de contrôle en 3 points Cette méthode recherche trois critères : (1) asymétrie, (2) réseau pigmentaire atypique et (3) structures bleu-blanc [68]. La présence de l'un de ces critères indique la possibilité d'un mélanome.

2.3.2.5.3 La liste de contrôle en 7 points Cette méthode utilise la technique de notation de la règle ABCDE. Il existe trois critères d'évaluation majeurs et quatre critères mineurs. Les critères majeurs sont l'existence d'un ou des tissus pigmentaires atypiques, d'une ou des zones gris-bleu et d'un ou des schémas vasculaires atypiques avec un score de (2). Par exemple, si tous ces critères sont présents dans une lésion, celle-ci est notée 6.

Il existe également des critères mineurs. Il s'agit de l'existence d'un flux radial, d'une pigmentation diffuse irrégulière, de points et de globules irréguliers et d'un modèle de régression dans une lésion cutanée avec un score de (1) pour chacun d'entre eux. Pour établir un diagnostic de mélanome, il faut obtenir un score total minimum de (3).

Score dermoscopique total (TDS)	Interprétation
<4.75	Lésion bénigne
4.8-5.54	Lésion suspecte, suivi rapproché ou ablation recommandée
>5.54	Risque élevé de mélanome
score faux positif (>5.45) parfois observés dans des cas	<ul style="list-style-type: none">• Nævus de Reed et de Spitz• Nævus de Clark avec motif globulaire• Nævus mélanocytaire congénital

Tableau 2.3 – *Interprétation de score dermoscopique total (TDS).*

2.3.2.5.4 La méthode de Menzies Cette méthode est basée sur une liste de contrôle de onze caractéristiques. Dans cette méthode, l'absence ou la présence de onze caractéristiques au total est examinée. Elle distingue les lésions bénignes des mélanomes grâce à deux ensembles de caractéristiques négatives et neuf ensembles de caractéristiques positives. L'ensemble négatif ne comprend que deux caractéristiques, à savoir la symétrie et la couleur unique, tandis que l'ensemble positif comprend neuf caractéristiques : l'existence d'un voile bleu-blanc, de multiples points bruns, de pseudopodes, d'un flux radial, d'une dépigmentation de type cicatrice, de points noirs périphériques, de multiples couleurs, de multiples points bleus/gris et d'un large réseau pigmentaire [69].

L'existence d'au moins une caractéristique de la liste des caractéristiques positives et l'absence des deux caractéristiques de la liste des caractéristiques négatives sont nécessaires pour diagnostiquer une lésion comme étant un mélanome malin.

2.3.2.5.5 L'analyse des formes L'analyse des motifs ou de forme est une autre méthode utilisée pour diagnostiquer les lésions mélanome et différencier les lésions bénignes des mélanomes malins. La méthode d'analyse des formes est utilisée pour identifier des formes spécifiques de lésions cutanées qui peuvent être globales ou locales. Certains des motifs globaux sont réticulaires, globulaires, en pavé, homogènes, en étoile, parallèles, multicomposants et non spécifiques, qui se réfèrent aux lésions bénignes. Les motifs locaux sont le réseau pigmentaire, les points/globules/moles, les stries, le voile bleu-blanc, les structures de régression, l'hypopigmentation, les taches et les structures vasculaires, qui font également référence aux lésions bénignes [69].

2.3.3 Techniques d'imagerie du cancer de la peau

S'il est diagnostiqué à un stade précoce ou peu avancé, le cancer de la peau peut être traité dans 90 % des cas, contre 50 % à un stade avancé. Avec le développement de techniques d'imagerie non invasives et à haute résolution, la précision du diagnostic des cancers de la peau ou des lésions cutanées s'est améliorée. En ce qui concerne le

mélanome en particulier, la faible précision du diagnostic est la principale raison du sur-traitement (dû à un diagnostic faussement positif) ou du traitement insuffisant (dû à un diagnostic faussement négatif).

Les faux diagnostics positifs sont les principaux responsables de l'augmentation excessive des coûts de traitement, car ils conduisent à exciser un nombre inutilement élevé de lésions bénignes en vue d'une biopsie et d'un examen pathologique. Cependant, les techniques d'imagerie à haute résolution ont un fort potentiel d'amélioration de la spécificité du diagnostic, ce qui permet de réduire les prélèvements inutiles. Les techniques d'imagerie les plus courantes actuellement utilisées pour le diagnostic des cancers de la peau sont la microscopie confocale par réflectance, la tomographie par cohérence optique, l'échographie et la dermoscopie.

2.3.3.1 La microscopie confocale par réflectance

La microscopie confocale est une méthode d'imagerie non invasive qui utilise un laser focalisé sur un point spécifique de la peau et qui permet de visualiser les détails cellulaires de la peau en temps réel. Les structures cellulaires (cellules, mélanine, hémoglobine, etc.) ayant des indices de réfraction différents. La microscopie confocale peut facilement différencier la lumière réfléchi par la peau. Cependant, elle est la plus coûteuse des autres techniques d'imagerie de la peau.

2.3.3.2 La Tomographie en Cohérence Optique

La Tomographie en Cohérence Optique peut être utilisée pour imager des structures microscopiques (quelques μm) en temps réel et peut distinguer les tissus sains des tissus cancéreux. Cependant, cette méthode n'est pas capable de visualiser les éléments subcellulaires et la membrane d'où il ne peut pas détecter la tumeur à un stade précoce. De plus, sans confirmation histologique, la Tomographie en Cohérence Optique ne peut pas déterminer complètement le diagnostic de mélanome. Elle n'est donc pas un moyen avantageux de diagnostic du mélanome.

2.3.3.3 L'échographie

L'échographie est l'une des techniques non invasives les plus courantes, car elle est polyvalente, indolore et présente peu de risques. Dans cette méthode, la morphologie de la peau peut être visualisée par les ondes ultrasonores qui reviennent du tissu. Bien que les ondes ultrasonores puissent atteindre les couches profondes de la peau et évaluer la tumeur, la faible résolution ne permet pas de distinguer les lésions cutanées d'un point de vue histomorphologique. En outre, elle ne permet pas de détecter les tumeurs à un stade initial.

2.3.3.4 La dermoscopie

La dermoscopie, également connue sous le nom de microscopie à épilumence, est une méthode non invasive et en direct très pratique pour la détection précoce du mélanome malin et d'autres lésions pigmentées. Elle permet aux utilisateurs de capturer les couleurs et les structures souterraines de la peau afin de détecter le mélanome à un stade précoce. Selon les statistiques de la littérature, l'utilisation de la dermoscopie peut augmenter la précision du diagnostic de 5 à 30 % en fonction du type de lésion cutanée [70].

Le mélanome peut être reconnu lorsque la lésion ne présente que 2 ou 3 couleurs à la dermoscopie (ou 1 ou 2 en clinique). Les mélanomes plus profonds révèlent plus de couleurs [71]. La figure 2.7 présente quelques exemples d'images dermoscopiques données par DermNet [71] avec des formes ou structures et couleurs irrégulières.

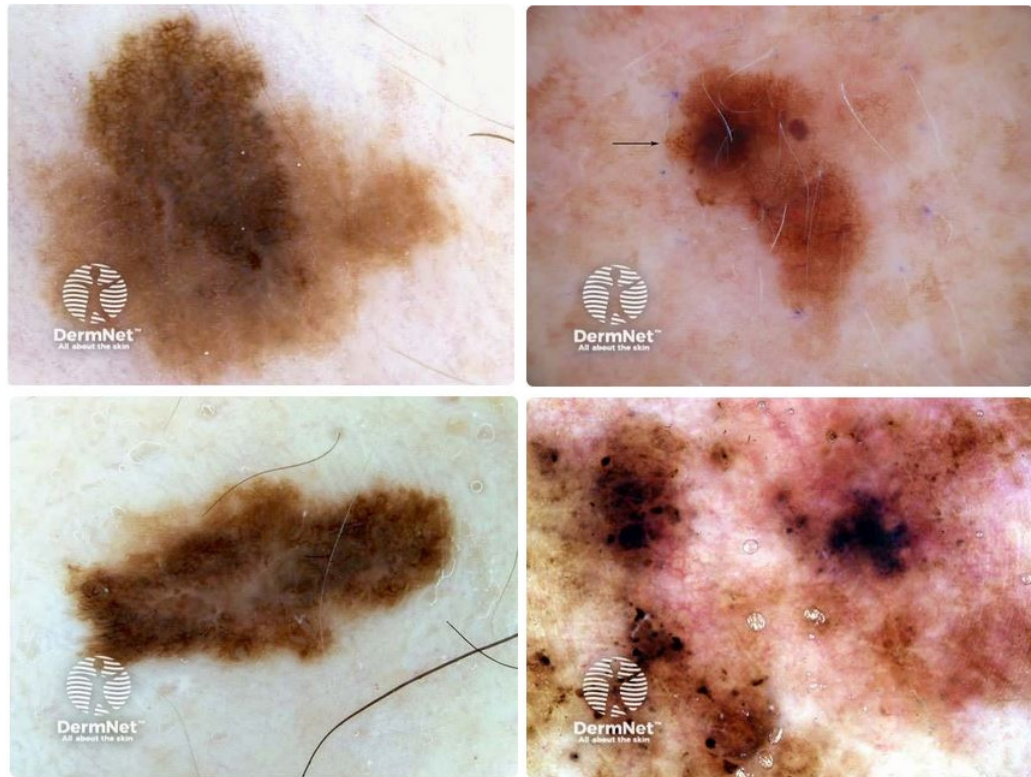


Figure 2.7 – Images dermoscopiques : formes et couleurs irrégulières.

2.4 Utilisation de l'intelligence artificielle pour le diagnostic du mélanome malin

Le diagnostic précoce du mélanome est crucial pour augmenter les chances de guérison. Les méthodes traditionnelles citées précédemment, bien que fiables, peuvent être longues et sujettes à une variabilité inter-observateurs. Le Deep Learning, un sous-domaine de l'apprentissage automatique, a émergé comme une approche prometteuse pour automatiser et améliorer la précision du diagnostic dermatologique, notamment pour le mélanome.

En général, le CNN est susceptible d'être largement appliqué à la classification des cancers de la peau et des blessures cutanées. Toutefois, les ensembles de données (dataset) existants sont insuffisants, certaines méthodes utilisent des échantillons de petite taille, il existe peu d'ensembles de données d'images cliniques et, en général, chaque méthode crée un nouvel ensemble de données pour ses besoins. En outre, les CNN sont les architectures les plus couramment utilisées pour l'analyse d'images médicales, y compris le diagnostic du mélanome. Des modèles populaires comme VGGNet, ResNet, et Inception ont été largement utilisés. Ces modèles ont démontré

une capacité remarquable à extraire des caractéristiques discriminantes des images de lésions cutanées.

2.4.1 Les ensembles de données (datasets)

Les ensembles de données jouent un rôle crucial dans le développement de modèles d'intelligence artificielle capables de diagnostiquer le mélanome malin. Ces données doivent être de haute qualité, variées et représentatives de la réalité clinique pour permettre aux algorithmes d'apprentissage automatique d'apprendre à distinguer les lésions bénignes et malignes.

Les Caractéristiques des datasets pour les cancers de la peau sont :

- Images démoscopiques de haute qualité : Les images sont claires, nettes et prises dans des conditions d'éclairage optimales.
- Annotation précise : Chaque image est associée à une annotation précise indiquant la présence ou l'absence de malignité, ainsi que le type de lésion si un mélanome est présent.
- Diversité des données : Les datasets doivent inclure des images de patients de différents âges, sexes, phototypes et origines ethniques, afin de garantir la généralisabilité des modèles.
- Équilibre des classes : Il est important d'avoir un nombre suffisant d'images de mélanomes et de lésions bénignes pour éviter les biais d'apprentissage.
- Méta-données : Des informations supplémentaires telles que l'historique médical du patient, les facteurs de risque, etc., peuvent être utiles pour améliorer la performance des modèles.

Parmi les datasets populaires dans le diagnostic des cancers de la peau sont celles de l'organisation internationale ISIC (l'International Skin Imaging Collaboration).

1. Le dataset ISIC 2017 : L'ensemble de données ISIC 2017 fait partie du défi annuel de l'International Skin Imaging Collaboration (ISIC), qui se concentre sur l'analyse et la classification des lésions cutanées. L'ensemble de données 2017 comprend une variété d'images dermatologiques annotées avec différents types de lésions cutanées, telles que le mélanome, le carcinome basocellulaire et les naevus bénins. Le dataset ISIC 2017 contient environ 2 000 images de lésions cutanées [72].
2. Le dataset HAM10000 [73] : ("Human Against Machine avec 10000 images") a été créé pour faciliter l'apprentissage des réseaux neuronaux en vue du diagnostic automatisé des lésions cutanées pigmentées. L'ensemble de données se compose de 10015 images dermatoscopiques accessibles au public dans les archives de l'ISIC. Plus précisément, il comprend 327 images de kératoses actiniques, 514 carcinomes basocellulaires, 1099 kératoses bénignes, 115 dermatofibromes, 1113 mélanomes, 6705 naevus mélanocytaires et 142 lésions cutanées vasculaires. En comparaison, l'ensemble de données PH2 comprend 40 mélanomes et 160 naevus mélanocytaires.
3. Le Dataset ISIC 2019 : Le jeu de données ISIC 2019 est une nouvelle version clé de la série de jeux de données annuels fournis par l'ISIC. Le dataset ISIC 2019 contient environ 25 000 images dermoscopiques. Ce nombre plus élevé par rapport aux années précédentes permet une plus grande diversité et une

meilleure couverture des lésions cutanées. La répartition des images par classe est comme suit [74] : le Melanome est environ 1100 à 1200 images, le Basal Cell Carcinoma (BCC) est environ 900 à 1000 images, le Squamous Cell Carcinoma (SCC) est environ 200 à 300 images, le Actinic Keratosis de 400 à 500 images, Benign Keratosis de 1000 à 1200 images, les Nevus est environ 3000 à 4000 images, et pour Vascular Lesions, le nombre d'images est de 300 à 400 images.

4. Le dataset ISIC 2020 [75] : est une autre étape importante dans l'évolution des ensembles de données pour l'analyse des lésions cutanées. Il est utilisé pour des tâches telles que la classification et la segmentation des lésions dermatologiques. Le dataset ISIC 2020 contient environ 32542 images dermoscopiques représentant un pourcentage de 98.2% pour la bénigne et 584 images avec un pourcentage de 1.8% pour les images de mélanome malin confirmé.

2.4.2 Les travaux existants : Méthodes de Deep Learning appliquées au diagnostic du mélanome

Plusieurs études ont été publiées, confirmant que l'IA peut être aussi performante, voire plus, que des professionnels qualifiés lorsqu'il s'agit de détecter des problèmes de santé, y compris des cancers de la peau. Voici quelques travaux de recherche existants utilisant des méthodes de deep learning pour le diagnostic du mélanome malin.

Dans l'article [76], Hurtado et Reales ont proposé un nouveau système de classification des cancers de la peau à partir d'images provenant d'un appareil photo standard. Ils ont étudié l'impact du bootstrapping lissé, une technique permettant d'augmenter les ensembles de données, sur les résultats de la classification. L'étude a comparé huit classificateurs, dont SVM, KNN et ANN, avec et sans augmentation des données. Les résultats ont montré que l'ANN avec l'augmentation des données a obtenu les meilleures performances et le meilleur équilibre, atteignant une AUC de 87.1

Li et al [77] ont passé en revue 45 travaux de recherche sur l'identification des maladies de la peau à l'aide de la technologie d'apprentissage profond depuis 2016. Ils évaluent ces recherches en termes de type de maladie, d'ensemble de données, de technologie de traitement des données, de technologie d'augmentation des données, de modèle d'identification des images de maladies cutanées, de cadre d'apprentissage profond, d'indicateurs d'évaluation et d'efficacité du modèle. Dans [77], les auteurs ont présenté une vue d'ensemble des méthodes d'apprentissage profond et de leurs applications dans le diagnostic des maladies de la peau. Ils ont commencé par une introduction aux problèmes de peau et aux méthodologies de collecte de photographies dermatologiques. Ils ont examiné les architectures d'apprentissage profond les plus répandues ainsi que les cadres communs pour la mise en œuvre des algorithmes d'apprentissage profond.

Srinivasu et al. [78] ont proposé un processus informatisé de classification des maladies de la peau en utilisant l'apprentissage profond basé sur la mémoire à long terme (LSTM). L'ensemble de données HAM10000 a été utilisé et la méthode proposée a surpassé les autres méthodes avec une accuracy de plus de 85

Un nouveau modèle a été construit par Goceri [79] à l'aide de MobileNet, où une nouvelle fonction de perte hybride a été utilisée. Cette étude explore l'impact de la technologie mobile et du deep learning sur le diagnostic des maladies de la peau. Elle examine comment les applications mobiles et les modèles de deep learning peuvent

être intégrés pour améliorer les diagnostics dermatologiques dans un contexte clinique et pratique. La technique proposée atteint une accuracy de 94.76

Adekanmi et al. [80] ont proposé une méthode basée sur l'apprentissage profond qui surmonte les limites de la détection et de la segmentation automatiques des lésions de mélanome. Un réseau codeur-décodeur amélioré avec des sous-réseaux codeur et décodeur reliés par une série de chemins de saut qui proposent de rapprocher le niveau sémantique des cartes de caractéristiques du codeur du niveau sémantique des cartes de caractéristiques du décodeur. Le modèle DL atteint une précision et un coefficient de dé de 95 % et 92 % sur l'ensemble de données ISIC 2017 et une précision et un coefficient de dé de 95 % et 93 % sur les ensembles de données PH2.

Hussien et al. [81] ont proposé une approche d'apprentissage profond pour la classification des mélanomes cutanés à l'aide d'un modèle CNN à 27 couches, qui donne des résultats prometteurs. Le modèle CNN est méticuleusement conçu pour extraire les caractéristiques des images de lésions cutanées et les classer dans les catégories mélanome et non mélanome. Ce modèle incorpore plusieurs couches de convolution, des couches de normalisation par lots, des couches de mise en commun maximales, des couches entièrement connectées, des couches d'exclusion et des techniques d'augmentation des données, qui contribuent toutes à la précision et à la généralisation du modèle. Les résultats expérimentaux obtenus sur des ensembles de données de référence accessibles au public pour la classification des lésions cutanées montrent que le modèle CNN proposé surpasse les approches de pointe existantes. Une bonne précision, égale à 99.99%, a été obtenue dans la détection du modèle.

Les auteurs de [82] ont proposé un système qui combine les développements récents en matière d'apprentissage profond avec des approches d'apprentissage automatique établies pour créer des ensembles de méthodes capables de segmenter les lésions cutanées, ainsi que d'analyser la zone détectée et les tissus environnants pour détecter les mélanomes. Comparé à la moyenne de huit dermatologues experts sur un sous-ensemble de 100 images de test, le système proposé produit une meilleure accuracy (76 % contre 70.5 %) et une meilleure spécificité (62 % contre 59 %) évaluées à une sensibilité (82 %).

Dans [83], Li et Shen ont proposé deux méthodes d'apprentissage profond pour traiter trois tâches principales émergeant dans le domaine du traitement des images de lésions cutanées, à savoir la segmentation des lésions (tâche 1), l'extraction des caractéristiques dermoscopiques des lésions (tâche 2) et la classification des lésions (tâche 3). Les cadres d'apprentissage profond proposés sont évalués sur l'ensemble de données ISIC 2017. Les résultats expérimentaux montrent des "accuracy" prometteuses des cadres proposés, c'est-à-dire 0.753 pour la tâche 1, 0.848 pour la tâche 2 et 0.912 pour la tâche 3.

Les auteurs de [84] ont évalué les performances de plusieurs réseaux neuronaux convolutifs de pointe. Pour améliorer la qualité des images, les auteurs ont utilisé plusieurs étapes de prétraitement. Le prétraitement et l'augmentation des données peuvent avoir un impact positif sur l'accuracy finale.

Naeem et al. [85] ont collecté l'état de la recherche en identifiant les tendances récentes de la recherche, les défis et les opportunités pour le diagnostic du mélanome et ont exploré les solutions existantes pour le diagnostic de la détection du mélanome en utilisant l'apprentissage profond, afin d'aider les chercheurs dans le domaine de la détection du mélanome, un modèle proposé, des défis et des opportunités ont été présentés.

En outre, les auteurs de [86] ont utilisé un cadre de réseau neuronal amélioré pour

réaliser un apprentissage rapide des caractéristiques et une segmentation idéale des images de mélanome. Une couche de normalisation par lots est utilisée entre la couche de convolution et la couche d'activation (telle que ReLU ou ELU) pour résoudre le problème de la disparition et de l'explosion du gradient.

Dans [87], Reis et al. ont proposé un réseau neuronal convolutif basé sur l'apprentissage profond appelé InSiNet, où la performance du modèle est basée sur plusieurs modèles pré-entraînés utilisés dans cette étude tels que GoogleNet, DenseNet-201, ResNet152V2 et EfficientNetB0. L'architecture InSiNet développée a réussi à atteindre une "accuracy" de 94.59%, 91.89 % et 90.54 % dans les ensembles de données ISIC 2018, 2019 et 2020, respectivement.

Les auteurs de [88] ont proposé un réseau d'attention longue multi-échelle (MSLANet) pour la classification des lésions cutanées, qui est composé de trois réseaux d'attention longue (LANet) où chaque LANet peut fusionner les informations contextuelles et améliorer la capacité de représentation discriminative grâce au mécanisme d'attention longue. Les résultats montrent que l'architecture MSLANet développée surpasse les méthodes de pointe, atteignant une AUC de 93.7% sur l'ensemble de données ISIC 2017 et AUC de 92.4 % sur l'ensemble de données SIIM-ISIC 2020.

Dans [89], Mijwil a sélectionné et entraîné un réseau d'apprentissage profond pour analyser plus de 24 000 images de cancer de la peau à l'aide d'un modèle de réseau neuronal convolutif (ConvNet) en appliquant trois architectures : InceptionV3, ResNet et VGG19. Afin d'identifier les meilleures architectures pour la classification de ces images et d'obtenir des résultats extrêmement satisfaisants, de nombreuses variables sont prises en considération. D'après les résultats des tests, la meilleure architecture est InceptionV3, qui a atteint une accuracy de diagnostic d'environ 86.90 %, une précision de 87.47 %, une sensibilité de 86.14 % et une spécificité de 87.66 %.

Le tableau 2.4 présente un résumé détaillé des travaux existants sur les méthodes de deep learning appliquées au diagnostic du mélanome.

Référence	Méthodes utilisées	Dataset	Résultats (%)
[76] (2021)	Machine Learning (CNN, SVM)	ISIC 2018	AUC 87.1
[90] (2020)	Etude et évaluation des travaux existant	Divers datasets	–
[77] (2021)	Analyse des méthodes de Deep Learning	Divers datasets	–
[78] (2021)	Mobilenet V2, LSTM	HAM10000	Accuracy 85.34, Spécificité 92, Sensibilité 88.24
[80] (2021)	Deep Learning, Technologie Mobile	Divers datasets	Accuracy 94.76
[80] (2020)	Deep Learning (CNN)	ISIC 2017, PH2	Accuracy 95, Spécificité 96, Sensibilité 97
[81] (2024)	Deep Learning (CNN)	Divers datasets	Precision 99
[82] (2017)	Combinant plusieurs modèles (CNN)	ISIC 2016, ISBI 2016	Accuracy 76, Spécificité 62, Sensibilité 82
[83] (2018)	Deep Learning (CNN)	ISIC 2017 et PH2	Accuracy 85.7, Spécificité 96.1, Sensibilité 49, AUC 91.2
[84] (2019)	Evaluation de plusieurs modèles, Deep learning CNN	ISIC 2018	–
[85] (2020)	Modèles (CNN) Evaluation plusieurs Dataset	ISIC, PH2, Dermofit, MED-NODE, HAM10000	–
[86] (2017)	Deep Learning (CNN) et segmentation	ISIC 2016	Accuracy 91, Précision 97.8, Spécificité 97, Sensibilité 88.6
[87] (2022)	Proposer un modèle de deep learning, InSi-Net	ISIC 2018, ISIC 2019, ISIC 2020	Accuracy 90.54, Spécificité 88.64, Sensibilité 93.33
[88] (2023)	Proposer, un réseau de neurones MSLANet	ISIC 2017, ISIC 2020	Accuracy 95.6, Spécificité 97.4, Sensibilité 59.6, AUC 92.4
[89] (2021)	Plusieurs modèles (CNN)	ISIC 2018, HAM10000	Accuracy 86.90, Précision 87.47, Spécificité 87.66, Sensibilité 86.14

Tableau 2.4 – *Résumé des travaux existants.*

2.5 Conclusion

Dans ce chapitre, nous avons présenté un aperçu sur le cancer de la peau précisément le mélanome malin, en détaillant les aspects fondamentaux et les méthodes de diagnostic traditionnelles. Nous avons examiné également les dernières avancées et travaux existant dans l'état de l'art en Deep Learning appliquées au diagnostic du cancer de la peau, en mettant en lumière les innovations et les techniques de pointe dans ce domaine. La combinaison des connaissances médicales traditionnelles avec les innovations technologiques ouvre de nouvelles perspectives pour le diagnostic précoce et la gestion du cancer de la peau, augmentant ainsi les chances de traitement réussi et d'amélioration des résultats pour les patients.

Dans le chapitre suivant, nous développons un modèle de diagnostic du mélanome basé sur des réseaux de neurones convolutifs (CNN). Nous exploitons des modèles pré-entraînés tels que VGG19 et MobileNetV2 pour accéder à l'apprentissage par transfert.

Identification du mélanome malin dans les images de lésions cutanées à l'aide de l'apprentissage par transfert

Sommaire

3.1	Introduction	65
3.2	Présentation du dataset utilisé	66
3.2.1	Description des métadonnées	66
3.2.2	Exploration des caractéristiques	68
3.2.3	Caractéristiques de la cible (sortie)	68
3.2.4	Caractéristiques du sexe	69
3.2.4.1	Caractéristiques de l'âge approximatif	70
3.2.5	Caractéristiques de la position anatomique	71
3.2.6	Caractéristiques du diagnostic	72
3.2.7	Caractéristiques du nom d'image et l'identificateur du patient	73
3.3	Prétraitement des données	74
3.3.1	Élimination des noms d'images redondantes	74
3.3.2	Division du dataset	75
3.4	Architecture du modèle de classification utilisant l'apprentissage par transfert	76
3.5	Résultats et discussion	78
3.5.1	Résultats d'entraînement et de validation	78
3.5.2	Test et évaluation	80
3.6	Conclusion	82

3.1 Introduction

Les images de dermoscopie ont prouvé leur importance pour évaluer les lésions cutanées et catégoriser les lésions normales, bénignes et malignes sous la forme d'un dermatoscope. Il s'agit de l'un des tests diagnostiques les plus utilisés en dermatologie, car il est non invasif, en temps réel, portable et peu coûteux.

L'objectif principal de ce chapitre est de présenter une approche pour le diagnostic du mélanome, en se basant sur un réseau neuronal. Cette implémentation utilise des réseaux de neurones convolutifs (CNN) et l'apprentissage par transfert avec des

modèles pré-entraînés tels que VGG19 et MobileNetV2. L'approche développée comprend quatre phases : le choix d'un dataset, le prétraitement des données, l'extraction de caractéristiques, et l'ajustement fin.

3.2 Présentation du dataset utilisé

Dans cette étude, les ensembles de données dermoscopiques ouverts disponibles sur internet sont utilisés pour l'évaluation de l'approche de classification proposée.

L'approche était basée sur un ensemble de données approuvées par le comité institutionnel de l'International Skin Imaging Collaboration (ISIC) et la Society for Imaging Informatics in Medicine (SIIM), SIIM-ISIC 2020 [75]. Les images proviennent des sources suivantes : Hospital Clinic de Barcelona, Medical University of Vienna, Memorial Sloan Kettering Cancer Center, Melanoma Institute Australia, University of Queensland et University of Athens Medical School.

Le dataset comprend 33126 images dermoscopiques, dont 32542 images représentant un pourcentage de 98.2% pour la classe 0 (bénigne) et 584 images avec un pourcentage de 1.8% pour la classe 1 maligne (mélanome confirmé). L'ensemble de données contient des images de lésions cutanées qui comprennent quelques classes telles que le diagnostic clinique, les cibles binaires et les métadonnées ou les informations sur les caractéristiques spécifiques fournies par les dermatologues.

3.2.1 Description des métadonnées

Les métadonnées présentées avec l'ensemble de données contiennent les informations supplémentaires associées aux images dermoscopiques. Ces métadonnées fournissent un contexte et des informations importantes sur l'ensemble de données et peuvent inclure divers détails décrits dans le tableau 3.1.

	Non mélanome (bénignes)	mélanome (malignes)
No. d'images	32542	584
No. d'images/sex		
Mal	16716	364
Femelle	15761	220
No. d'images/age		
<20	798	8
<40 and >20	10239	89
>40	21437	487
No. d'images/anatomical position		
Membre supérieur	4872	111
Membre inférieur	8293	124
Torse	16588	257
Paume des mains/plantes des pieds	370	5
Tête/nuque	1782	74
Orale/génitale	120	4
No. de patient	2055	428
No. de patient/sex		
Homme	977	168
Femme	1076	260
No. de patient/position anatomique		
Membre supérieur	1400	93
Membre inférieur	1623	108
Torse	1942	207
Paume des mains/plantes des pieds	195	5
Tête/nuque	816	67
Orale/génitale	77	3

Tableau 3.1 – *Statistiques sur les mélanomes et non-mélanomes*

Notre modèle doit prédire la probabilité entre 0 et 1 que la lésion de l'image soit mélanome maligne ou non. La valeur 0 indique une lésion bénigne et la valeur 1 indique une lésion maligne (mélanome malin). Le fichier des métadonnées contient les huit colonnes suivantes :

- 'image_name' : L'identifiant indique le nom de fichier de l'image correspondante.
- 'patient_id' : L'identifiant du patient.
- 'sex – sexe' : Le sexe du patient (s'il est inconnu, il n'est pas renseigné 'NAN').
- 'age_approx' : L'âge approximatif du patient au moment de l'imagerie.

- 'anatom_site_general_challenge' : cette colonne représente la position anatomique, ou la localisation de l'image dans corps.
- 'diagnosis' : Les informations détaillées sur le diagnostic.
- 'benign_malignant' : Indicateur de la malignité de la lésion.
- 'target' : La version binaire de la variable cible.

3.2.2 Exploration des caractéristiques

L'exploration des caractéristiques du dataset ISIC 2020 est une étape cruciale pour comprendre la nature des données et pour informer la stratégie de modélisation. Le dataset ISIC 2020 est largement utilisé pour la détection et la classification des lésions cutanées, en particulier pour la détection du mélanome, un type de cancer de la peau. Un aperçu général des caractéristiques de dataset ISIC 2020 est décrit comme suit :

3.2.3 Caractéristiques de la cible (sortie)

Lorsque l'on parle de "cible" dans le contexte de la modélisation, de l'apprentissage automatique ou de tout autre processus de prédiction, on fait référence à la variable que l'on cherche à prédire ou à classifier. C'est le résultat final que notre modèle est censé produire.

Le diagramme à barre de la figure 3.1 illustre la distribution de la variable cible (sortie) au sein du dataset ISIC2020. Cette variable est catégorielle et binaire : une lésion peut être classée comme "maligne" ou "bénigne".

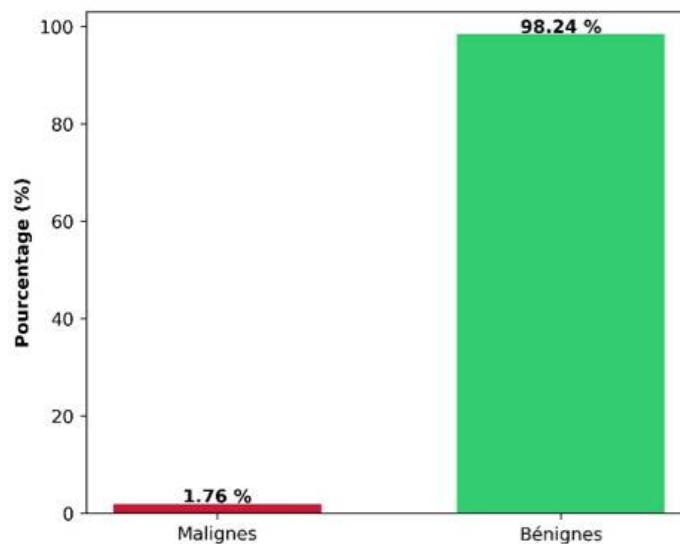


Figure 3.1 – Visualisation des caractéristiques de la cible.

Le nombre d'images de lésions bénignes est de 32542 (98.24%), tandis que le nombre d'images de lésions malignes est de 584 (1.76%). Comme indiqué ci-dessus, les images ou classe bénignes et malignes (0 et 1) sont complètement déséquilibrées. Ce déséquilibre a des conséquences importantes sur l'entraînement et l'évaluation des modèles de classification. Les modèles ont tendance à être biaisés vers la classe

majoritaire (bénigne dans ce cas), ce qui peut entraîner une mauvaise performance sur la classe minoritaire (maligne).

3.2.4 Caractéristiques du sexe

Le graphique présenté par la figure 3.2 illustre la distribution des sexes au sein du dataset ISIC2020, utilisé pour la classification des lésions cutanées. Il met en évidence une répartition relativement équilibrée entre les individus de sexe masculin et féminin. Les deux catégories (hommes et femmes) sont représentées de manière assez similaire, avec un léger avantage numérique pour les hommes (51.56% contre 48.24% pour les femmes).

Aussi, le nombre total d'individus dans le dataset semble important, avec plusieurs milliers d'observations pour chaque sexe. Cela suggère une bonne représentativité de la population et une fiabilité statistique des résultats.

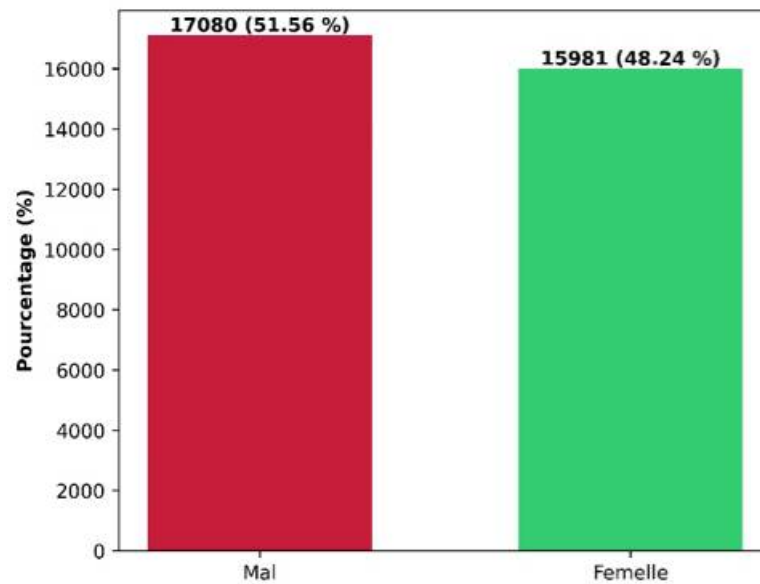


Figure 3.2 – Visualisation des caractéristiques du sexe.

La figure 3.3 présente la distribution des images de lésions cutanées maligne et bénigne en fonction du sexe. Il met en évidence une différence significative entre les hommes et les femmes. Le nombre d'images de lésions malignes chez les hommes est nettement supérieur à celui des femmes. Cette différence est visuellement frappante et suggère une prévalence plus élevée des cas de malignité chez les hommes dans cet ensemble de données. Cette différence peut-elle s'expliquer par des facteurs biologiques intrinsèques aux hommes et aux femmes, ou bien par des facteurs environnementaux (exposition au soleil, habitudes de vie, etc.).

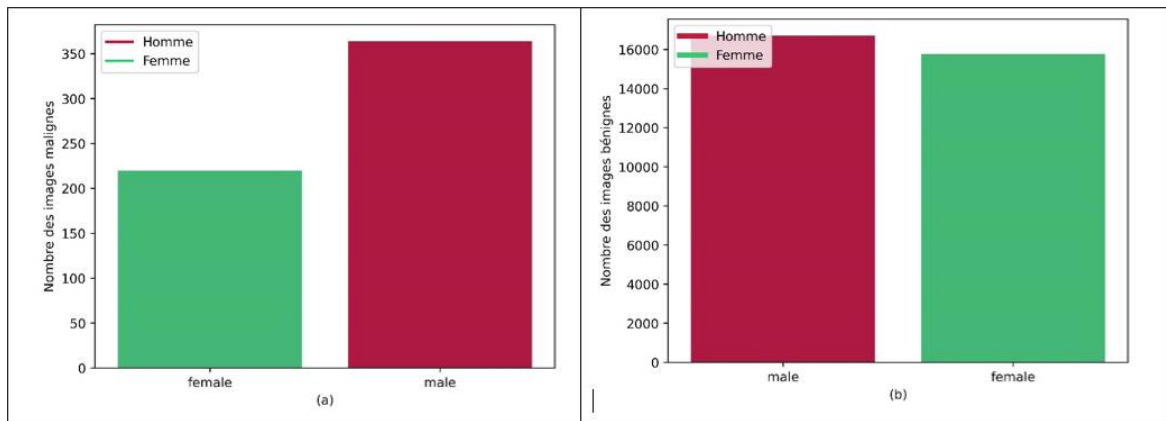


Figure 3.3 – Distribution de sexe par rapport au, (a) : maligne, (b) : bénigne.

3.2.4.1 Caractéristiques de l'âge approximatif

Tout d'abord, il faut déterminer l'âge moyen par sexe, qui est 50 ans pour les hommes, et de 47 ans pour les femmes, ce qui nous permet de conclure que les femmes sont touchées par cette maladie à un âge plus précoce que les hommes.

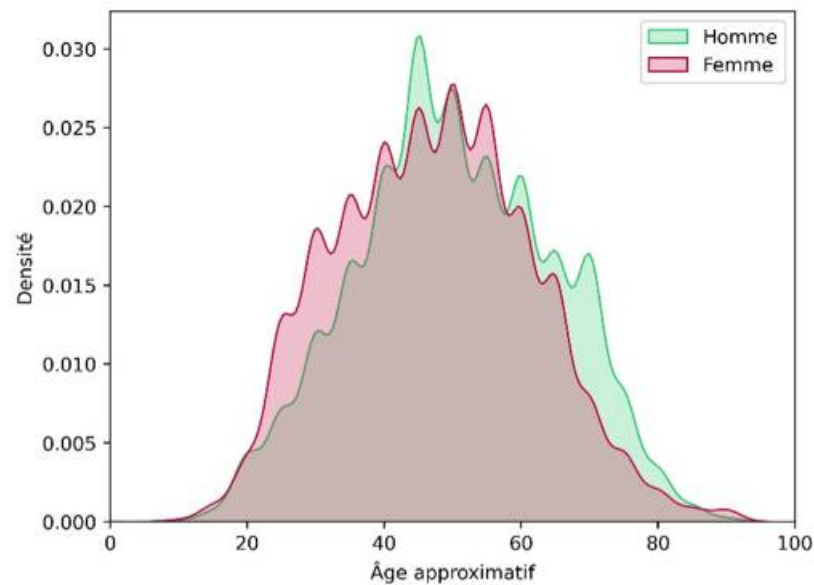


Figure 3.4 – Distribution approximative de l'âge moyen par sexe.

Maintenant, nous visualisons la caractéristique de l'âge moyen par cible (maligne ou bénigne) à l'aide d'histogramme de la figure 3.5. On constate qu'il y a plus de personnes atteintes le mélanome malin dont l'âge est compris entre 58 et 62 ans.

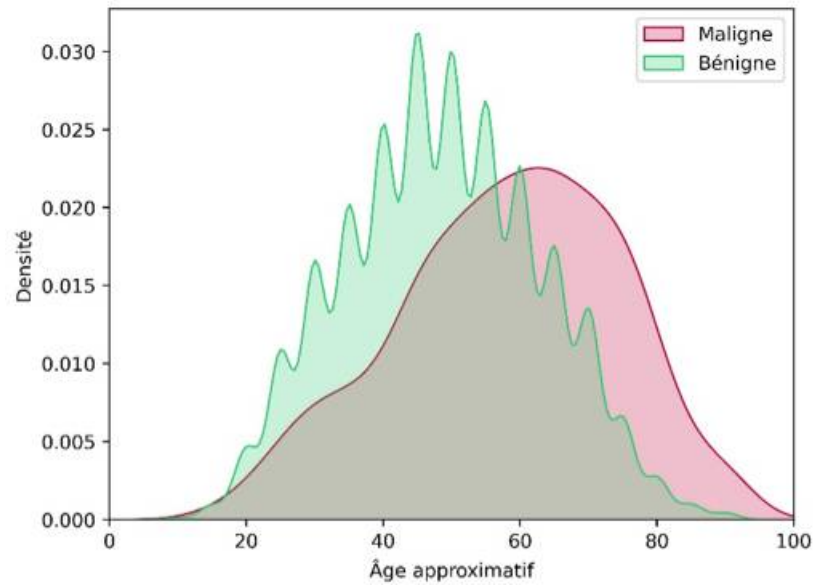


Figure 3.5 – Distribution approximative de l'âge moyen par cible (sortie).

3.2.5 Caractéristiques de la position anatomique

Cette caractéristique présente le nombre d'images des lésions cutanées en fonction de leur position anatomique sur le corps humain. Les positions anatomiques dans le dataset ISIC 2020 sont catégorisées comme suit :

- torso : La torse,
- lower extremity : Les membres inférieurs,
- upper extremity : Les membres supérieurs,
- head/neck : La tête et la nuque,
- palms/soles : Les paumes des mains et les plantes des pieds,
- oral/genital : les organes orale et génitale.

Cette distribution peut avoir des implications pour le diagnostic et le traitement des lésions cutanées. Par exemple, les dermatologues pourraient accorder une attention particulière aux zones les plus touchées lors des examens cliniques.

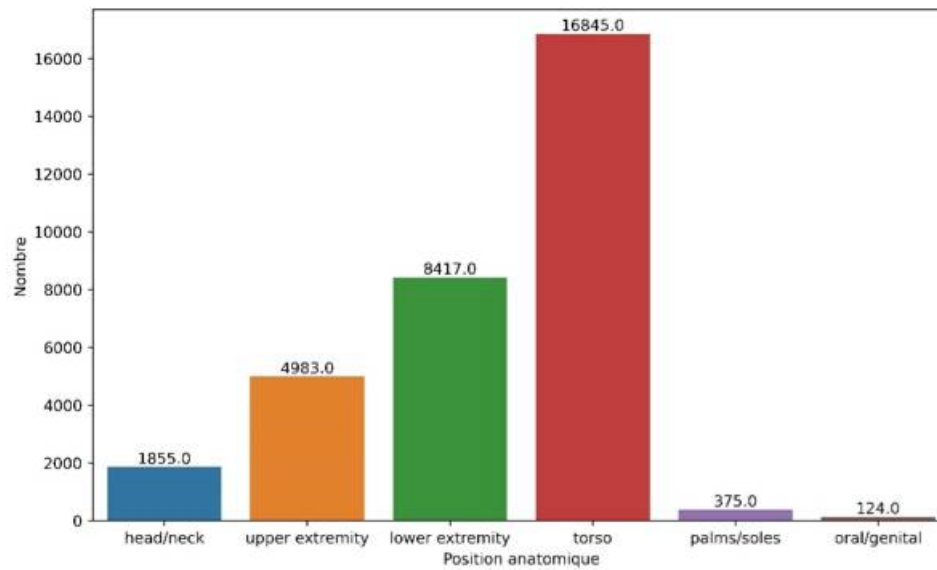


Figure 3.6 – Nombre d'images par rapport à la position anatomique.

Comme le montre le graphique en barres de la figure 3.6, la partie du corps la plus touchée est le torse (le tronc du corps humain) avec plus de 16 000 cas. Cela suggère une exposition plus importante de cette zone aux facteurs de risque liés au développement de lésions cutanées (soleil, frottements, etc.).

En deuxième position, on trouve les membres inférieurs (**lower extremity**) et supérieurs (**upper extremity**), avec respectivement environ 8400 et 4900 cas. Ces zones sont également exposées aux éléments extérieurs et peuvent être sujettes à des traumatismes. Les autres zones présentent un nombre de cas nettement inférieur, ce qui est cohérent avec une exposition moindre aux facteurs de risque et une moindre probabilité de développement de lésions cutanées.

3.2.6 Caractéristiques du diagnostic

Cette caractéristique indique des informations détaillées sur le diagnostic. Comme le montré le digramme circulaire de la figure 3.7, les cas les plus importants sont les nævus (nevus), mélanome (melanoma) et inconnu (unknown). Dans ce dataset, inconnu présente 27124 images de lésions bénignes.

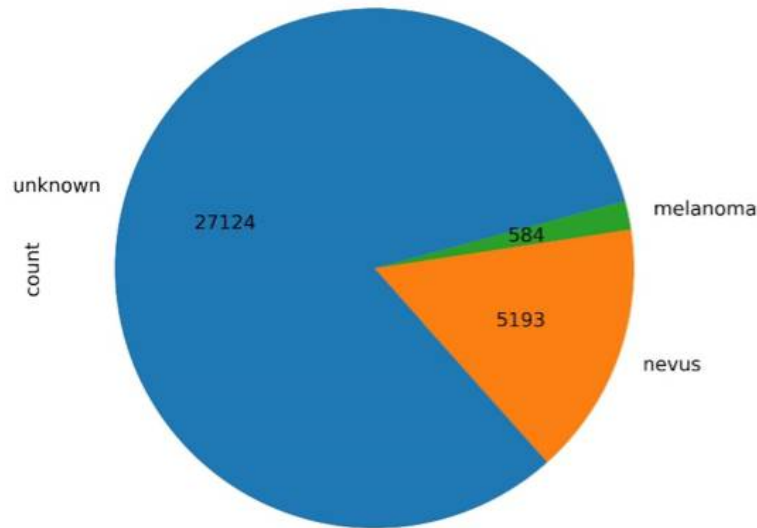


Figure 3.7 – Nombre d'images par rapport à diagnostic.

3.2.7 Caractéristiques du nom d'image et l'identificateur du patient

La répartition des noms d'images (`image_name`) selon l'identifiant du patient (`patient_id`), pour les patients ayant plus de 75 images, est illustrée par le diagramme circulaire de la figure 3.8. Chaque secteur du cercle représente un patient unique, identifié par un numéro (IP_ suivi de chiffres). La taille de chaque secteur est proportionnelle au nombre total d'images associées à ce patient. Les chiffres inscrits à l'intérieur des secteurs indiquent vraisemblablement le nombre exact d'images par patient.

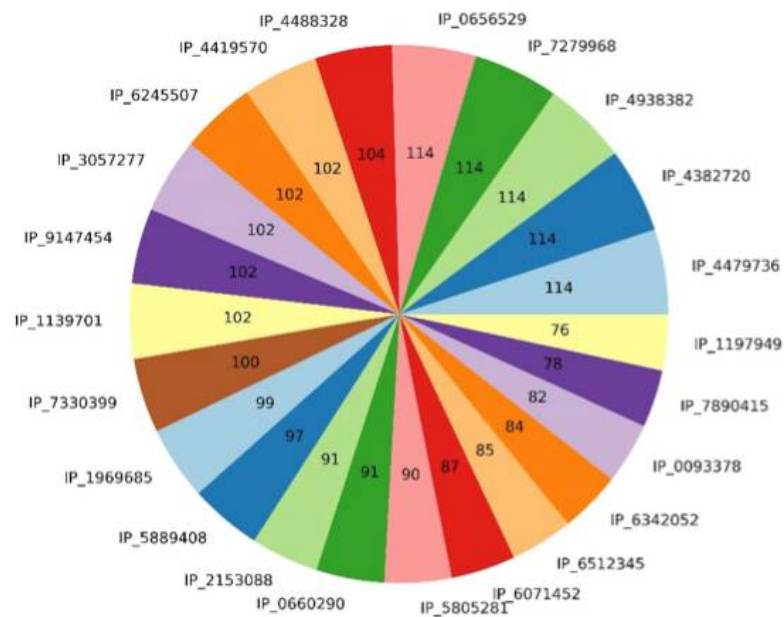


Figure 3.8 – Distribution de l' `image_name` par `patient_id` (uniquement les nombre > 75).

Selon la répartition illustrée par le diagramme de la figure 3.8, chaque patient possède une ou plusieurs images de lésions bénignes, avec un maximum de 114 images par patient, comme c'est le cas pour les ID IP_0656529 et IP_7279968, par exemple. Donc, il existe une grande dissemblance dans le nombre d'images par patient. Certains patients ont un nombre d'images nettement supérieur à la moyenne, tandis que d'autres en ont un nombre plus faible. Cette hétérogénéité peut avoir des implications pour les analyses statistiques. Par exemple, les patients avec un grand nombre d'images pourraient avoir un poids plus important dans les analyses, ce qui pourrait biaiser les résultats.

Il est donc nécessaire de procéder à un prétraitement initial pour éliminer les redondants, afin de réduire le nombre total d'images de lésions bénignes, qui est de 32542 images. Tous les détails seront expliqués dans la section suivante par l'organigramme de la figure 3.9.

3.3 Prétraitement des données

Avant l'entraînement du modèle proposé et le transfert learning, on va procéder à un nettoyage des données inutiles ce qu'on appelle un prétraitement des données qui compte deux étapes :

3.3.1 Elimination des noms d'images redondantes

D'après l'exploration des caractéristiques analysées dans la partie précédente, nous avons découvert que, suite à un problème survenu lors de l'ingestion des données dans nos archives, plusieurs centaines d'images étaient effectivement stockées en redondant sous deux noms internes (image_name). Dans les données d'apprentissage, chaque couple dupliqué appartient au même patient, c'est-à-dire que si $\text{nom_image}(x) = \text{nom_image}(y)$, cela signifie que $\text{patient_id}(x) = \text{patient_id}(y)$, ce qui signifie qu'il n'y a pas de nouvelles données sur le patient.

Après avoir éliminé les noms d'images en double sur la base du code unique du patient (patient_id) et de la position anatomique (anatom_site_general_challenge), les 32542 images initiales pour la classe bénigne ont été réduites à un maximum de 6271 images. Nous avons construit l'ensemble de données d'entrée à partir de 6271 images pour la classe 0 (classe bénigne) et 584 pour la classe 1 (classe maligne), comme le montre l'organigramme de la figure 3.9.

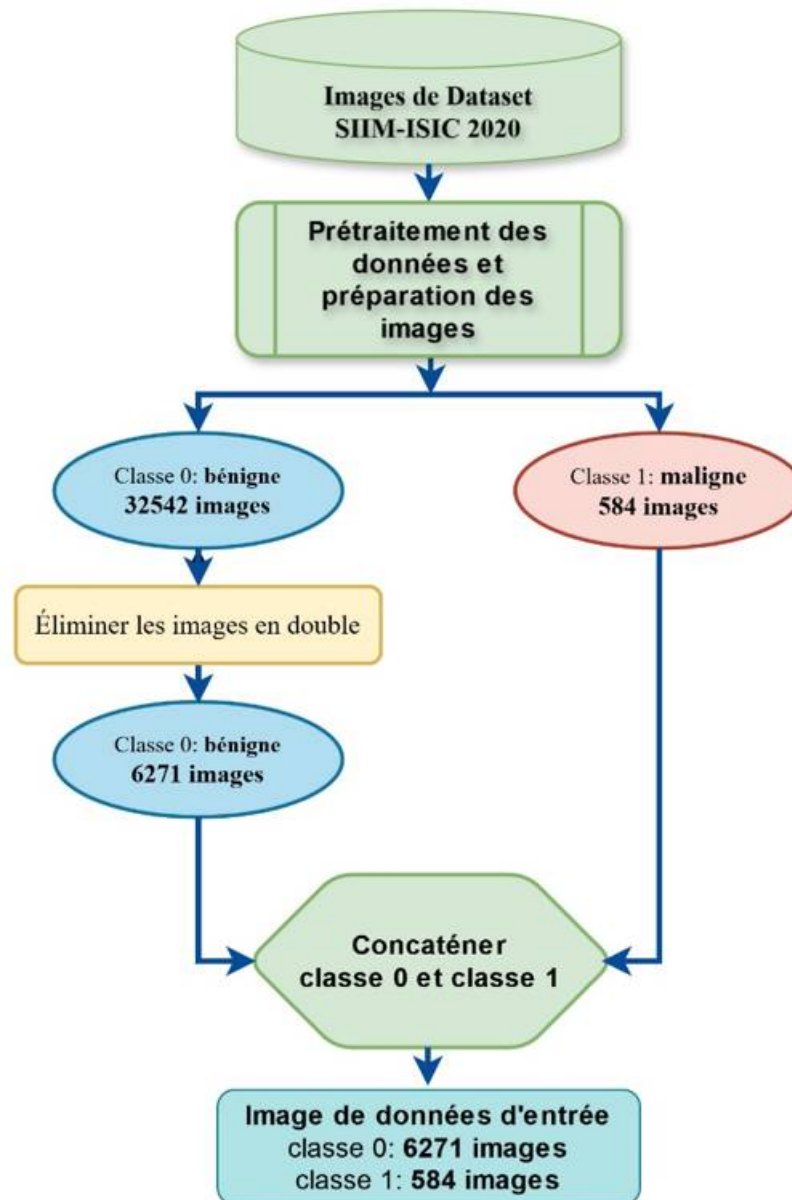


Figure 3.9 – Préparation des données : suppression des images en double.

3.3.2 Division du dataset

La structure et les étapes du modèle proposé pour la classification des mélanomes cutanés et l'apprentissage par transfert sont illustrées à la figure 3.10. Elle se compose de six éléments principaux : le prétraitement de l'image (qui génère les données d'entrée de l'image), la division de l'ensemble de données entraînement, validation et test (Training set, Validation set, et Test set), l'extraction de caractéristiques et l'ajustement fin à l'aide des modèles pré-entraînés VGG19 et MobileNetV2, l'ajout de couches de sortie denses et, enfin, la classification de l'entraînement et l'évaluation des performances.

Les données collectées doivent être divisées en trois ensembles différents : Training set, Validation set, et Test set, comme le montre la figure 3.10. De plus, la répartition des images entre les deux classes a été effectuée comme le montre le tableau 3.2.

Initialement l'ensemble de données a été divisé en 80 % de données de Training

set et 20 % de données de Test set. En plus, nous avons divisé le même ensemble de données d'entraînement en 75 % des données de l'ensemble de Training set et 25 % des données de l'ensemble de Validation set (c'est-à-dire 60 % de l'ensemble de Training set et 20 % de l'ensemble de Validation set de l'ensemble de données globale, la figure 3.10 illustre cette configuration).

Ensemble d'entraînement et de validation		Ensemble de test	
Classe (0) bénigne	Classe (1) maligne	Classe (0) bénigne	Classe (1) maligne
5017	467	1254	117

Tableau 3.2 – Division de l'ensemble de donnée par classe.

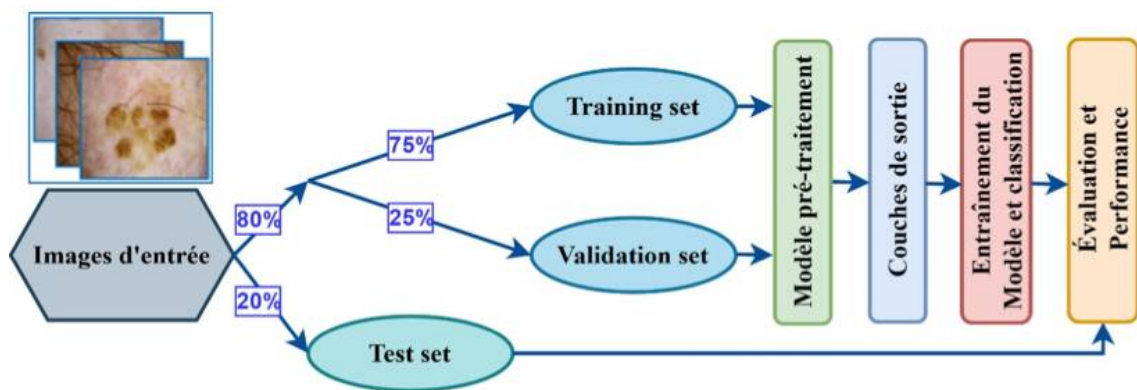


Figure 3.10 – Modèle proposé pour la prédiction du mélanome.

3.4 Architecture du modèle de classification utilisant l'apprentissage par transfert

Le CNN est un algorithme d'apprentissage profond conçu pour les images et les vidéos. Un CNN comporte plusieurs couches : une couche d'entrée, des couches cachées (extraction des caractéristiques), une couche entièrement connectée (classification) et une couche de sortie. Après le développement rapide des modèles d'apprentissage profond, dans la pratique, divers modèles pré-entraînés de classification (détaillés précédemment dans le chapitre 1) sont utilisés sur différents ensembles de données, tels que AlexNet, VGG, GoogLeNet, Tom, Inception, ResNet, et MobileNet pour la classification des maladies.

Dans notre approche, nous allons exploiter les modèles pré-entraînés VGG-19 et MobileNetV2. Les couches et l'architecture du réseau pour l'apprentissage par transfert avec l'ajustement fin du modèle VGG-19 et MobileNetV2 sont présentées dans la figure 3.11.

En ce qui concerne le modèle VGG-19, les points à noter sont qu'il a une image d'entrée de dimension (None, 224, 224, 3) et que nous avons quatre dernières couches de sortie, trois couches entièrement connectées incluant la couche de prédiction avec 1000 valeurs parce que le dataset a un nombre de classes d'images de 1000. D'autre

part, le modèle MobileNetV2 a la même dimension de l'image d'entrée que VGG-19 et possède deux couches connectées comprenant 1000 classes pour les couches de prédiction. Une couche softmax de 1000 catégories est utilisée comme couche de sortie dans les modèles VGG-19 et MobileNetV2. Nous supprimons cette couche et la remplaçons par une couche de sortie avec activation sigmoïde, adaptée à notre problème de classification binaire.

Dans un premier temps, après avoir gelé le réseau, nous supprimons les dernières couches entièrement connectées et les remplaçons par nos six couches définies, comme le montre la figure 3.11-(c). Nous avons également appliqué la régularisation dropout à chacune des couches ajoutées afin d'éviter les problèmes de sur-apprentissage (overfitting).

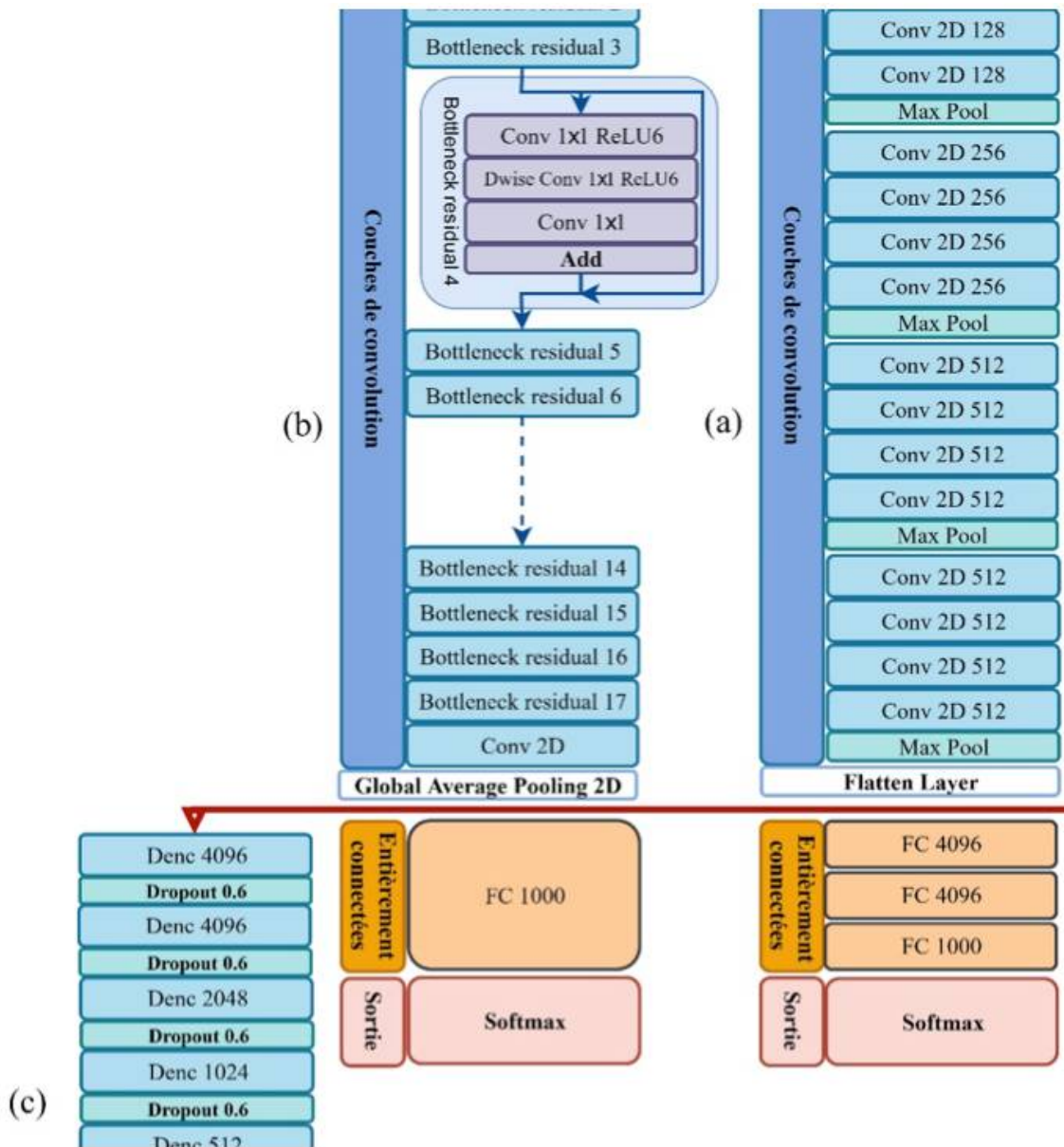


Figure 3.11 – Description de l'architecture du réseau (a) : modèle basé sur VGG-19, (b) : Modèle basé sur MobileNetV2 et (c) : Les couches ajoutées.

3.5 Résultats et discussion

L'apprentissage des réseaux CNN demande de vastes volumes de données et implique des calculs intensifs, ce qui nécessite une puissance de calcul considérable (notamment via des GPU : Graphics Processing Unit). En outre, étant donné que nous n'avons utilisé que 100 époques, en raison des limitations de temps et de puissance de calcul, nous aurions pu facilement obtenir de meilleurs scores de précision en utilisant des GPU, car ils ont tendance à être beaucoup plus rapides et à obtenir de meilleurs scores lorsqu'un plus grand nombre d'échantillons d'images sont utilisés pour l'entraînement.

De nombreux paramètres peuvent être ajustés lorsqu'il s'agit d'entraîner un modèle CNN. Les paramètres décrits ci-dessous sont utilisés dans toutes les expériences.

L'architecture du réseau utilisé pour nos expérimentations est basée sur les modèles pré entraînés VGG19 et MobilNetV2 (voir Figure 3.11). Bien que d'autres architectures puissent être utilisées pour obtenir de bons résultats, VGG19 et MobileNetV2 sont utilisés pour plusieurs raisons. Tout d'abord, il s'agit d'une architecture bien connue qui est devenue une architecture standard pour les expériences sur la classification des images. Deuxièmement, même avec des GPU puissants, l'apprentissage d'un VGG19 ou MobileNetV2 prend moins de temps, ainsi que les modèles plus complexes prennent plus de temps [91]. VGG19 et de même MobileNetV2 prennent en entrée une image RVB de 224x224.

Il reste encore un certain nombre d'hyperparamètres du réseau à définir. La plupart de ces hyperparamètres ont été choisis en expérimentant jusqu'à ce que le réseau commence à s'entraîner efficacement. Le `batch_size` de grande valeur peut entraîner une convergence lente, tandis que les `batch_size` de petite taille entraînent une convergence sous-optimale. La valeur des `batch_size` a été fixée à 64. Le taux d'apprentissage (`learning_rate`) a été fixé à 0.0001.

Le modèle a été mis en œuvre sur la plateforme TensorFlow à l'aide des modules Keras open-source et du langage de programmation python. En outre, il a utilisé l'optimiseur Adam avec la fonction de perte Crossentropie focale binaire (Binary Focal Crossentropy), définie par [92].

3.5.1 Résultats d'entraînement et de validation

La métrique la plus importante est l'accuracy. Il s'agit du nombre d'images correctement identifiées divisé par le nombre total d'images. Le taux d'erreur est également présenté, il s'agit du taux d'erreur de la classe dont l'accuracy globale est la plus faible.

Pour chaque modèle entraîné, nous avons évalué sa performance initiale en mesurant son accuracy sur un ensemble de données distinct (training set et validation set). Cette accuracy a été calculée à chaque étape de l'entraînement (époque) et est illustrée dans les graphiques de la figure 3.12 et la figure 3.13.

Un graphique montrant l'accuracy globale pour toutes les 100 époques de formation est également présenté. Il s'agit de donner au lecteur une idée de la manière dont l'apprentissage par transfert affecte la convergence d'un réseau avec apprentissage par transfert.

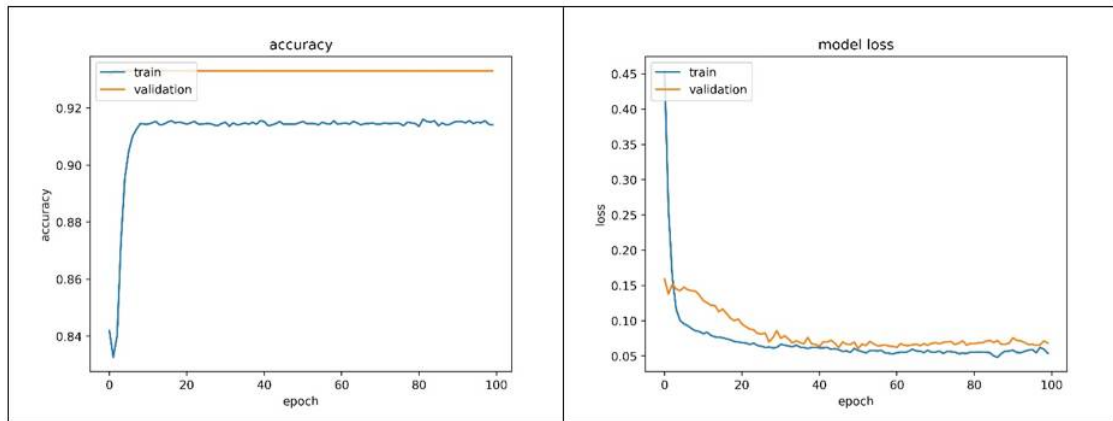


Figure 3.12 – Accuracy/Perte de modèle basé sur VGG-19 vs. époque.

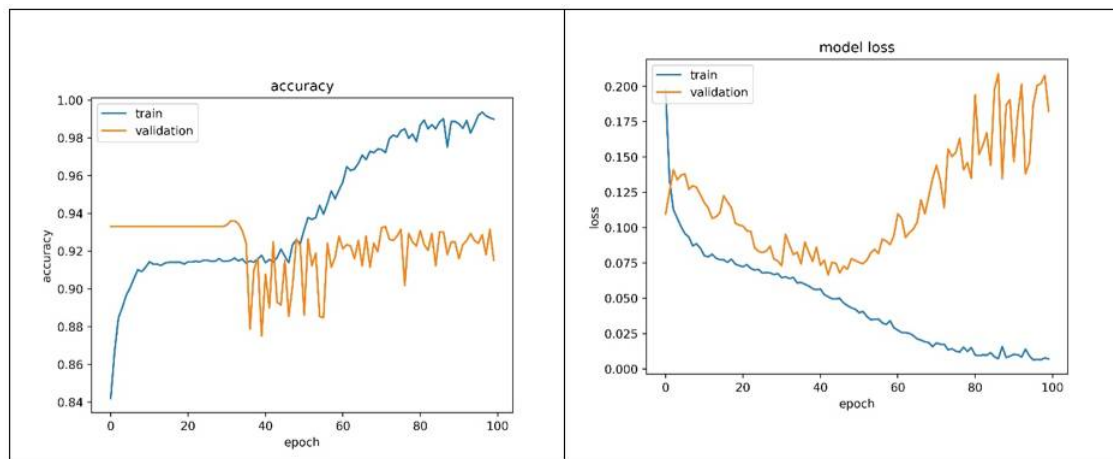


Figure 3.13 – Accuracy/Perte de modèle basé sur MobileNetV2 vs. époque.

L'accuracy et les pertes à chaque époque pendant l'entraînement et la validation du modèle basé sur VGG19 est montré dans la figure 3.12. Cette courbe montre qu'après 7 époques, l'accuracy de l'entraînement et de la validation augmentent rapidement. Après une phase d'apprentissage rapide au début, l'accuracy d'entraînement du modèle atteint un plateau à partir de la 10ème époque qui atteint la valeur 91.4%. Pour la validation, l'accuracy se stabilise et reste fixe pendant tout le temps d'apprentissage qui est autour de la valeur 93.3% pour les données de validation.

D'autre part, les pertes d'entraînement et de validation diminuent rapidement après 7 époques, et après 60 époques, les pertes d'entraînement et de validation deviennent stables, elles sont autour de 0.05 pour l'entraînement et 0.06 pour la validation.

Pour le modèle basé sur MobileNetV2, les courbes de la figure 3.13 montre qu'il y a une augmentation régulière de l'accuracy sur l'ensemble d'entraînement, atteignant environ 99% après 100 époques, ce qui montre que le modèle apprend bien à partir des données d'entraînement. Mais l'accuracy sur l'ensemble de validation commence à environ 93%, elle montre beaucoup de variations à partir de la 37ème époque, restant instable, autour de 90-92% avec de grandes oscillations. La perte sur l'ensemble

d'entraînement diminue régulièrement et atteint une valeur très basse de 0.007, tandis que, la perte sur l'ensemble de validation commence à augmenter après environ 40 époques et continue de croître avec des fluctuations importantes, ce qui est une indication claire d'un sur-apprentissage (overfitting).

Pour les deux modèles, les courbes d'entraînement montrent que les modèles apprennent de mieux en mieux avec le temps. Cependant, les courbes de validation ne suivent pas la même aptitude, indiquant un possible sur-apprentissage (overfitting). Cela se produit lorsque le modèle s'adapte trop bien aux données d'entraînement, au point de ne plus généraliser correctement aux nouvelles données.

3.5.2 Test et évaluation

Cette section présente la procédure de mise en place des tests d'évaluation du modèle ainsi que les résultats obtenus.

La matrice de confusion est une méthode ML précieuse qui détermine les métriques d'évaluation et du test tel que, l'accuracy, la spécificité, Recall (la sensibilité), l'AUC et la précision d'un modèle. La matrice de confusion présente une manière visuelle et quantitative d'évaluer les performances de deux modèles de classification basés sur VGG19 et MobileNetV2. Ces modèles sont utilisés pour distinguer des images bénignes (négatives) de celles qui sont malignes (positives). Les échantillons correctement classés par les modèles sont représentés en couleur foncée, tandis que les erreurs de classification sont indiquées en couleur claire comme le montre la figure 3.14.

D'après la matrice de confusion de la figure 3.14-a, Le modèle basé sur VGG19 semble avoir des difficultés considérables à classer correctement les images. Toutes les images classées comme malignes sont en réalité bénignes, cela indique que le taux de faux positifs est à 100%. Le modèle a tendance à sur-prédire la classe maligne, même lorsque les images sont bénignes.

Le Modèle basé sur MobileNetV2 a identifié une partie des cas positifs (malignes) correctement, mais pas tous (29%). Cela montre que le modèle est capable de détecter les positifs, bien que de manière imparfaite.

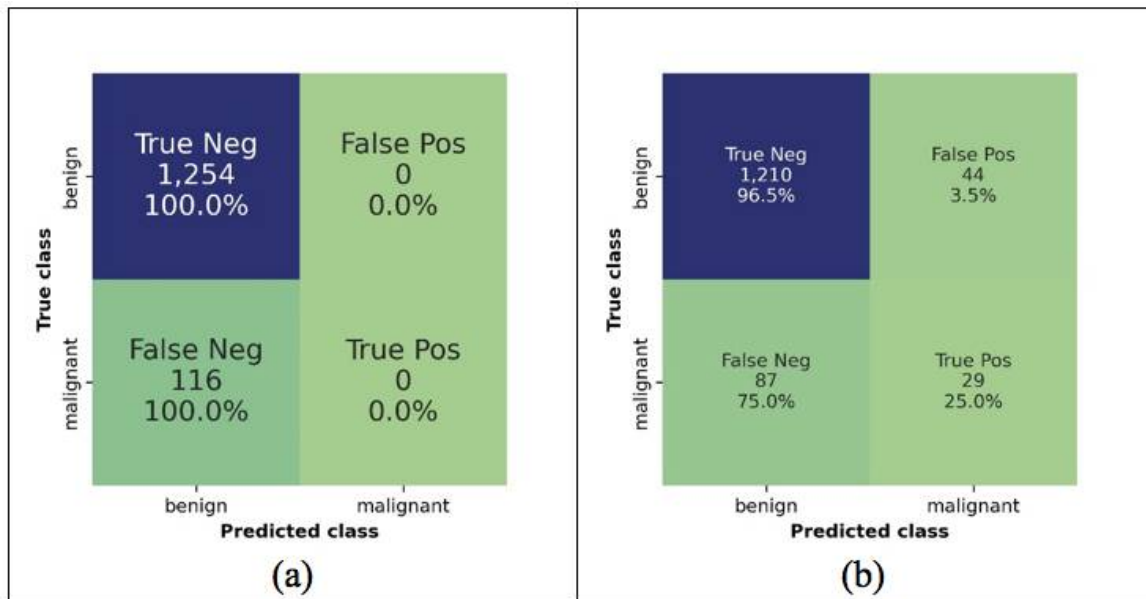


Figure 3.14 – Matrice de confusion du, (a) : modèle basé sur VGG-19 et (b) : du modèle basé sur MobileNetV2.

Les résultats des tests, en termes de métriques de performance, sont présentés dans le tableau 3.3 :

	Accuracy (%)	AUC (%)	Précision (%)	Sensibilité (%)	Spécificité (%)
Basé sur VGG19	91.53	14.1	0.00	0.00	100
Basé sur Mobil-NetV2	90.04	43.2	39.7	25	99.20

Tableau 3.3 – Performance des modèles basés sur VGG19 et MobileNetV2.

En se basant sur les métriques de performance présentées dans le tableau 3.3, nous allons discuter des résultats obtenus avec les modèles basés sur VGG19 et MobileNetV2 :

- Accuracy : Les deux modèles basés sur, VGG19 et MobileNetV2, affichent d'excellentes performances en termes d'accuracy, mais VGG19 est légèrement plus performant. Toutefois, il est important de noter que l'accuracy seule ne suffit pas toujours à évaluer un modèle, surtout lorsque les données sont déséquilibrées (c'est-à-dire qu'il y a plus d'image d'une classe par rapport à l'autre), car elle pourrait surestimer les performances du modèle sur la classe majoritaire.
- AUC : L'AUC est un indicateur clé pour mesurer la qualité d'un modèle, surtout lorsque les données sont déséquilibrées. Plus l'AUC est proche de 100%, mieux le modèle est capable de distinguer les différentes classes. Dans notre cas, MobileNetV2 surpasse nettement VGG19, ce qui signifie que le modèle basé sur MobileNetV2 est bien meilleur pour classer correctement les données.

- **Précision** : La précision reflète le pourcentage de prédictions positives (maligne) correctes parmi toutes les prédictions positives. La précision nulle de VGG19 sur la tâche de détection de mélanome, est confirmée par la matrice de confusion (Figure 3.14), qui ne présente aucune prédiction positive (maligne) correcte. MobileNetV2, avec une précision de 39.7%, démontre une capacité limitée mais non négligeable à discriminer les cas de mélanome, bien que des améliorations soient nécessaires.
- **Sensibilité** : La sensibilité mesure le pourcentage de cas positifs réels correctement identifiés par le modèle. Là encore, la sensibilité de VGG19 est de 0%, ce qui est préoccupant car cela indique que le modèle ne détecte aucun cas positif (maligne). MobileNetV2 a une sensibilité de 25%, ce qui est faible mais meilleur que VGG19, ce qui indique qu'il peut identifier certains cas positifs mais en manque beaucoup d'autres.
- **Spécificité** : La spécificité mesure le pourcentage de cas négatifs (bénin) réels correctement identifiés. Les deux modèles obtiennent de bons résultats sur cette métrique, avec un score parfait pour VGG19. Cela indique que VGG19 est bon pour identifier les cas non mélanome, mais étant donné sa mauvaise performance en précision et sensibilité, il est peut-être trop conservateur, classant trop de cas positifs comme négatifs.

Le modèle proposé basé sur MobileNetV2 montre une approche plus équilibrée par rapport à VGG-19. Il parvient à détecter certains cas de mélanome, bien qu'il manque encore de nombreuses détections. Il y a un compromis entre la capacité du modèle à identifier les malignes et les bénignes, ce qui reflète une meilleure généralisation que le modèle VGG-19, mais encore loin d'être parfait. Des améliorations pourraient impliquer une réévaluation de l'architecture du modèle, des méthodes de prétraitement des données ou de l'ajustement des seuils de décision.

3.6 Conclusion

L'objectif de cette étude était de trouver un modèle adapté à l'apprentissage par transfert, capable d'atteindre des scores d'accuracy respectables dans un court laps de temps. Nous avons réussi à trouver une architecture appropriée qui permet la classification d'images grâce à des réseaux neuronaux convolutifs et des modèles pré-entraînés VGG19 et MobileNetV2.

Comme nous l'avons montré dans la section précédente, il existe un déséquilibre observable dans l'ensemble de données qui doit être pris en considération. Nous avons constaté que la distribution des classes de lésions cutanées dans l'ensemble de données ISIC 2020 est très déséquilibrée, comme le montre la figure 3.1. Ce déséquilibre dans les données peut biaiser les résultats des modèles et doit être pris en compte.

Les résultats obtenus avec du modèle basé sur VGG19 sont très faibles. Bien que le modèle basé sur MobileNetV2 offre de meilleures performances, mais il reste perfectible, notamment en ce qui concerne la détection des cas de malignité.

Ces résultats mettent en évidence la nécessité de développer des modèles plus performants pour la détection précoce du mélanome. Si MobileNetV2 offre une première piste, il est clair que des améliorations sont indispensables pour obtenir des résultats cliniquement pertinents.

Nous aborderons dans le prochain chapitre les techniques permettant de remédier aux limitations observées.

Amélioration de la classification du mélanome dans les images de lésions cutanées

Sommaire

4.1	Introduction	83
4.2	Approche d’augmentation des données	84
4.2.1	Prétraitement des données	86
4.2.1.1	Génération de nouvelles données par la technique d’augmentation des données	86
4.2.1.2	Division du dataset	90
4.2.2	Résultats et discussion	91
4.2.2.1	Résultats d’entraînement et de validation	91
4.2.2.2	Test et évaluation	93
4.3	Approche hybride CNN et LSTM basée sur l’apprentissage par transfert	95
4.3.1	Architecture du réseau hybride CNN et LSTM	95
4.3.2	Résultats et discussion	97
4.3.2.1	Résultats d’entraînement et de validation	97
4.3.2.2	Test et évaluation	100
4.4	Évaluation comparative des performances : État de l’art vs. Nos approches proposées	101
4.5	Conclusion	102

4.1 Introduction

Le plus grand obstacle de la classification lors de l’utilisation d’algorithmes d’apprentissage profond CNN dans le domaine de la dermatologie est le manque d’ensembles de données. Les modèles CNN ont besoin de grandes quantités de données variées pour apprendre à généraliser à de nouveaux exemples. Un manque de données, surtout pour les classes minoritaires (comme les lésions malignes), peut entraîner un écart par rapport à la classe majoritaire et réduire la précision du modèle. Comme nous l’avons vu dans le chapitre 3, les données sont déséquilibrées. Le déséquilibre des classes peut entraîner un déséquilibre d’apprentissage du modèle en faveur de la classe d’échantillons majoritaires par rapport à la classe d’échantillons minoritaires,

se traduisant inévitablement par des métriques d'évaluation insatisfaisantes. Afin d'améliorer la fiabilité de nos résultats, nous avons mis en œuvre deux approches, que nous présentons dans ce chapitre.

Nous avons utilisé une approche d'augmentation des données, afin d'augmenter la classe 1 (maligne) pour qu'elle soit à peu près équivalente à la classe 0 (bénigne). Cette approche est décrite en détail dans la figure 4.1.

Dans une seconde approche, nous avons significativement amélioré les performances du modèle en intégrant un réseau hybride CNN-LSTM. Cette architecture hybride nous a permis de mieux capturer les caractéristiques spatiales et temporelles des données, conduisant ainsi à une amélioration significative des résultats. Les résultats de ces expérimentations sont présentés et discutés dans ce chapitre.

4.2 Approche d'augmentation des données

L'augmentation des données est une technique puissante pour améliorer la performance des modèles de l'apprentissage par transfert en créant des variations supplémentaires de nos données d'entraînement.

En augmentant artificiellement le nombre d'images, on réduit le risque de sur-apprentissage et on améliore la généralisation du modèle. Le modèle devient plus robuste aux variations dans les images d'entrée, ce qui est particulièrement important en dermatologie où les lésions peuvent présenter de légères différences d'une image à l'autre.

L'augmentation de données est une technique cruciale pour remédier au déséquilibre des classes en dermatologie, où les images de lésions malignes sont souvent insuffisamment représentées. Elle permet d'augmenter artificiellement le nombre d'exemples de la classe minoritaire (maligne) afin d'améliorer la performance des modèles de classification.

TensorFlow fournit la classe `ImageDataGenerator` qui permet d'appliquer facilement de nombreux paramètres de transformations à nos images, ces paramètres sont illustrés dans le tableau 4.1 :

Paramètre	Description	Type	Intervalle
rotation_range	Plage de degrés pour la rotation aléatoire.	Entier (degrés)	[0, 180]
width_shift_range	Plage pour le décalage horizontal aléatoire.	Flottant ou Tuple	[0, 1]
height_shift_range	Plage pour le décalage vertical aléatoire.	Flottant ou Tuple	[0, 1]
shear_range	Plage pour l'angle de cisaillement aléatoire.	Flottant (degrés)	[0, 45]
zoom_range	Plage pour le zoom aléatoire.	Flottant ou Tuple	[0, inf]
horizontal_flip	Retourner horizontalement de manière aléatoire.	Booléen	[True, False]
vertical_flip	Retourner verticalement de manière aléatoire.	Booléen	[True, False]
fill_mode	Stratégie pour remplir les pixels après transformation.	Chaîne de caractère	['constant', 'nearest', 'reflect', 'wrap']
Cval	Utiliser pour remplir les pixels lorsque fill_mode est 'constant'.	Flottant	[0, 255]
brightness_range	Plage pour la variation aléatoire de la luminosité.	Tuple	[0, inf] (min) et (max)
channel_shift_range	Plage pour le déplacement aléatoire des canaux de couleur.	Flottant	[0, 255]

Tableau 4.1 – Description des paramètres de transformations.

L'organigramme de la figure 4.1 présente le processus de préparation des données et d'équilibrage des classes d'images (nous utilisons la technique d'augmentation des données), étape préliminaire à l'entraînement de modèles de détection de mélanome par transfert en se basant toujours sur les deux modèles pré-entraînés VGG19 et MobileNetV2.

4.2.1 Prétraitement des données

Comme nous l'avons vu au chapitre 3, préalablement à l'entraînement du modèle et à la mise en œuvre de l'apprentissage par transfert, une étape de prétraitement des données s'impose. Celle-ci consiste notamment en la suppression des images redondantes, la partition des données, en plus, nous utilisons notre approche proposée qui est la technique d'augmentation de données.

4.2.1.1 Génération de nouvelles données par la technique d'augmentation des données

Avant de mettre en œuvre les techniques d'augmentation des données, il est essentiel de disposer d'un ensemble d'apprentissage suffisamment volumineux, ce qui est souvent nécessaire pour optimiser les performances des modèles CNN. Dans une première étape, un générateur d'images a été développé. Ce générateur applique des transformations aléatoires à chaque image en entrée, en faisant varier six paramètres de transformations différents (voir figure 4.1). Les valeurs de ces paramètres sont détaillées dans le tableau 4.2.

Paramètre	Valeur
horizontal_flip	True
vertical_flip	True
zoom_range	0.4
rotation_range	17
shear_range	0.4
height_shift_range	0.6

Tableau 4.2 – Paramètres et valeurs utilisés pour l'augmentation des données.

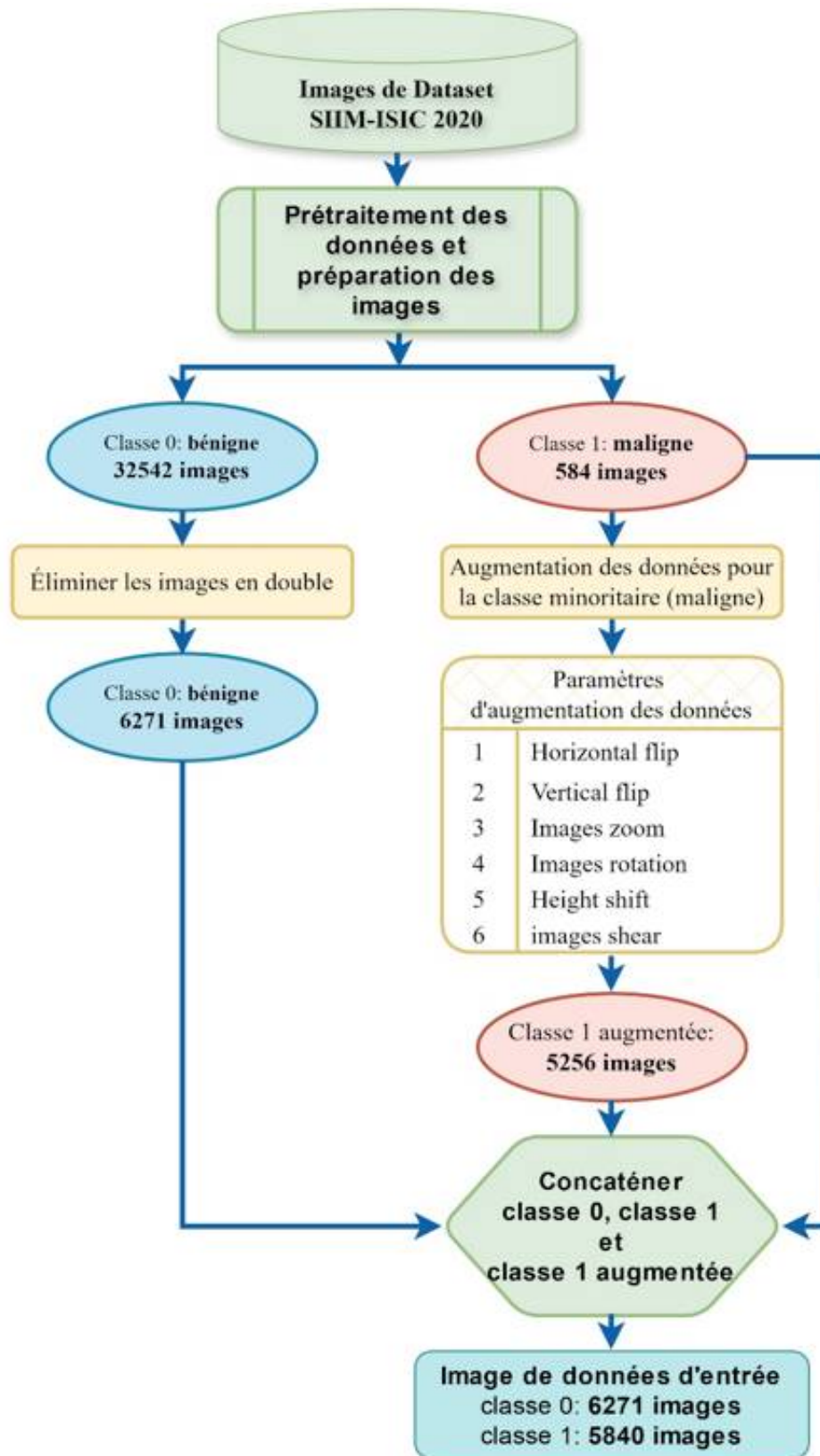


Figure 4.1 – Préparation des données : Approche d'augmentation des données.

Le jeu de données SIIM-ISIC 2020, composé d'images de lésions cutanées, a fourni les images d'entrée. Ces dernières ont subi un prétraitement pour standardiser leur format et leur qualité. Les images subissent un prétraitement comprenant une nor-

malisation, un redimensionnement et, si nécessaire, une filtration pour optimiser leur qualité. Par la suite, nous procédons à une répartition initiale des classes : 32 542 images sont classées comme bénignes (Classe 0) et 584 comme malignes (Classe 1). Afin d'assurer la fiabilité du modèle, les doublons de la classe bénigne sont éliminés par la suite, laissant un ensemble de 6 271 images uniques. Pour remédier au problème du déséquilibre des classes, nous avons mis en œuvre une technique d'augmentation de données sur la classe "maligne". Nous ajustons les paramètres de transformations aléatoires telles que `rotation_range`, `height_shift_range`, `shear_range`, `zoom_range`, `horizontal_flip`, et `vertical_flip`. Les valeurs de ces paramètres, qui ont été choisies de manière à préserver les caractéristiques pertinentes des images, sont présentées dans le tableau 4.2.

L'organigramme illustré par la figure 4.2 représente le processus d'augmentation d'image, où une boucle s'interrompt une fois qu'un nombre déterminé d'images augmentées a été généré. L'organigramme présenté décrit un processus d'augmentation d'images à partir d'un dataframe (`df_1`) contenant des images avec une classe maligne (`target=1`). Le but est de générer un certain nombre d'images augmentées à partir de ce dataframe.

Les étapes clés du processus sont :

- Filtrage des données : On sélectionne uniquement les lignes du dataframe où la cible est égale à 1 (pour la classe maligne), et création d'un dataframe appelé `df_1`, de la même manière création d'un dataframe appelé `df_0` qui contient que les images de la classe 0 (bénigne).
- Calcul du nombre d'images : On compte le nombre total d'images dans le dataframe `df_1`, (`nb_of_img_1`) en prenant la longueur du dataframe (584 images), de même le nombre total d'images dans le dataframe `df_0`, (`nb_of_img_0 = 6271` images).
- Calcul du nombre d'images à générer : Ce calcul est généré pour équilibrer les classes, la différence entre le nombre d'images des deux classes, représentant le nombre d'images à ajouter à la classe maligne (`nb_of_img = 5687`).
- Calcul du nombre d'augmentations par une seule image : On détermine combien d'augmentations seront appliquées à chaque image, c'est-à-dire le nombre d'augmentations à appliquer à chaque image de la classe maligne pour atteindre l'équilibre souhaité, en divisant `nb_of_img` par la longueur de `df_1` (`nb_of_img_1`), ici `nb_of_aug_img = 5687/584 ≈ 9`.
- Initialisation d'un générateur d'images : Un compteur `i` est initialisé à 0. Un `ImageDataGenerator` est initialisé pour générer des images augmentées avec les paramètres de transformation (nous utilisons les paramètres cités dans le tableau 4.2).
- Appel de la fonction `flow_from_dataframe` : Le générateur commence à créer des images augmentées à partir de dataframe `df_1`.
- Sauvegarde des images augmentées : Pour chaque image de la classe maligne, l'image est soumise à `ImageDataGenerator` pour générer une image augmentée. Les images générées sont sauvegardées dans le répertoire de destination.
- Incrémentation du compteur : Le compteur d'itérations `i` est incrémenté, $i = i + 1$
- Condition : $i > nb_of_aug_img?$: La boucle vérifie si `i` est supérieur à `nb_of_aug_img` (ici, 9). Si (Yes) `i` dépasse 9, la boucle s'arrête. Sinon (No), la boucle continue avec la prochaine itération pour générer plus d'images.

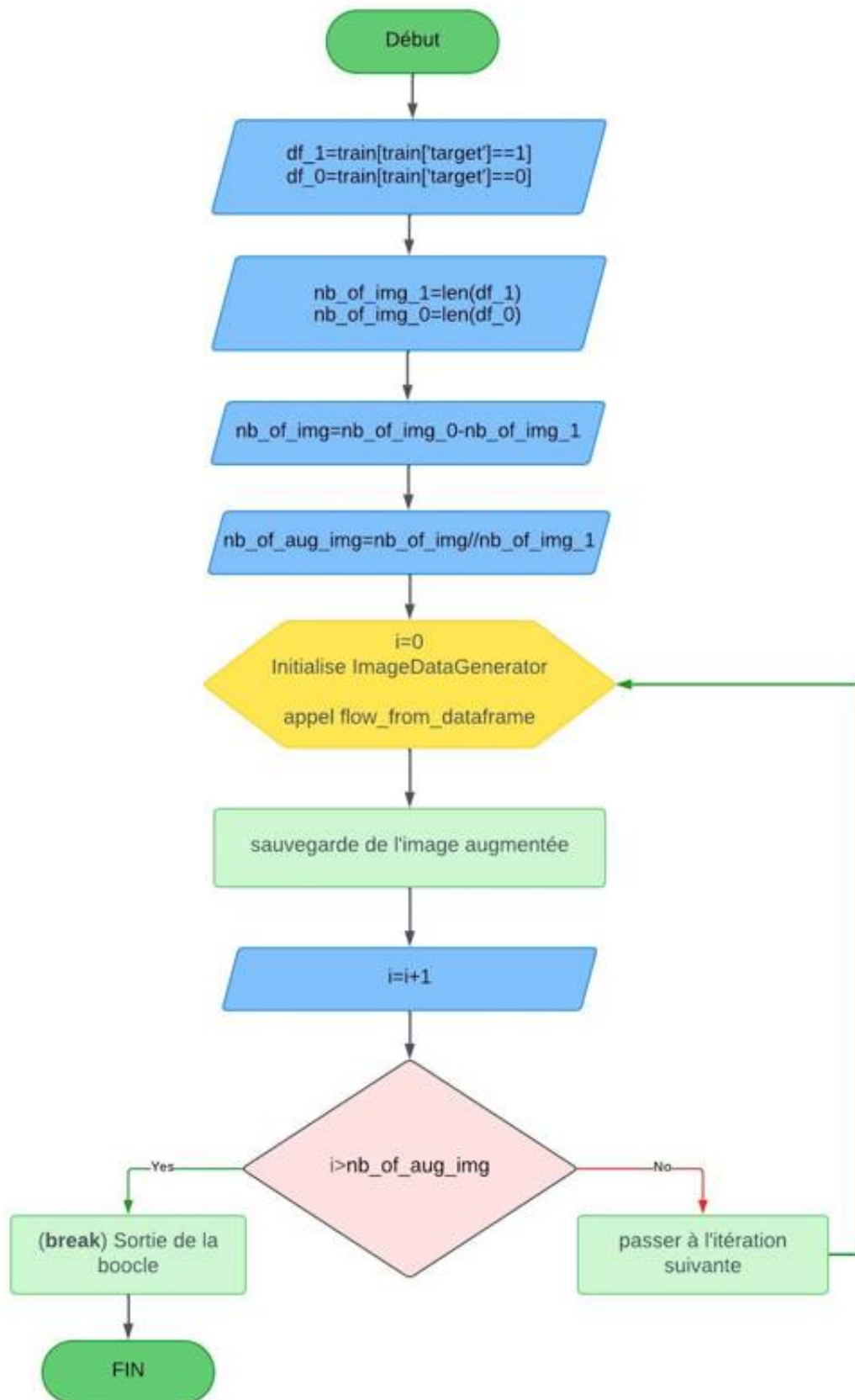


Figure 4.2 – Processus d'augmentation d'images.

Avec les images fournies, cela implique que pour chaque image maligne (classe 1), environ 9 images augmentées seront générées pour approcher le nombre d'images

bénignes (classe 0).

À partir de 584 images originales de la classe 1 (maligne), l'exécution du générateur de données d'image est répétée 9 fois, ce qui correspond au nombre de 5256 (9 fois 584) images augmentées qui sont générées. La figure 4.3-(a) présente un exemple d'image de la classe maligne. La figure 4.3-(b) illustre le résultat de l'augmentation de cette image, produisant neuf variations.

Dans ce cas, la technique de l'augmentation des données produit 5 256 images supplémentaires pour la classe maligne, rendant ainsi cette classe plus représentée dans le dataset.

Enfin, à l'issue de la phase d'augmentation de données, les images des classes bénigne et maligne sont concaténées pour former le dataset final. Ce dernier se compose de 6 271 images classées comme bénignes et de 5 840 images classées comme malignes (584 images originales + 5 256 images augmentées).

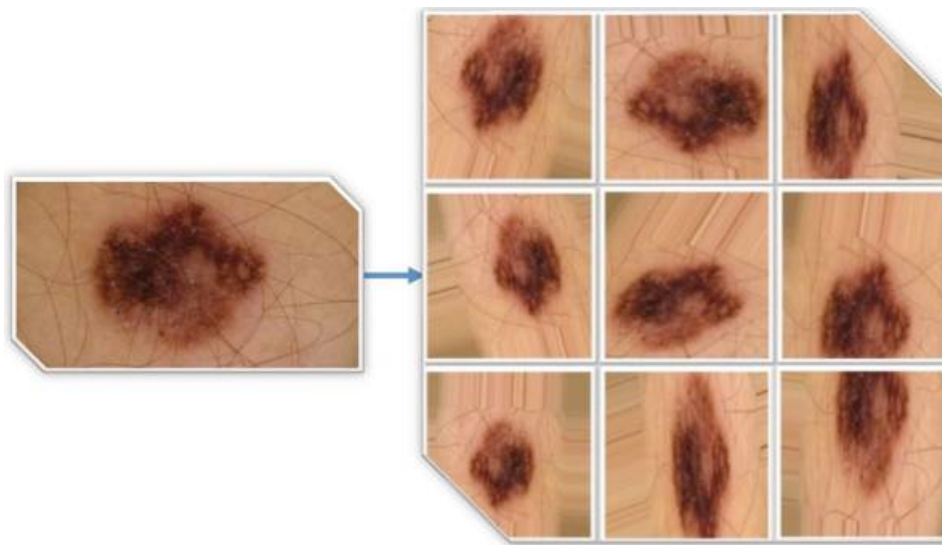


Figure 4.3 – Images du générateur de données a) image originale. (b) 9 images augmentées.

4.2.1.2 Division du dataset

Les étapes de classification des mélanomes malins et l'apprentissage par transfert mises en œuvre dans ce modèle sont identiques à celles détaillées au chapitre 3 (figure 3.10), à savoir une division de données, un prétraitement des images, un apprentissage supervisé sur un modèle pré-entraîné VGG19 ou MobileNetV2, et une évaluation des performances. De plus, la répartition des images entre la classe des mélanomes malins et celle des lésions bénignes suit la distribution présentée dans le tableau 4.3, assurant ainsi un équilibre suffisant pour l'entraînement du modèle.

Training and validation set		Test set	
Classe (0) bénigne	Classe (1) maligne	Classe (0) bénigne	Classe (1) maligne
5017	4672	1254	1168

Tableau 4.3 – Division de l'ensemble de donnée par classe.

Après avoir appliqué les techniques d’augmentation des données, nous avons divisé l’ensemble de données en trois sous-ensembles : un ensemble d’entraînement (60%), un ensemble de validation (20%) et un ensemble de test (20%). Cette répartition se traduit par 5017 images bénignes et 4672 images malignes pour l’entraînement et la validation, et 1254 images bénignes et 1168 images malignes pour les tests.

La quantité de données dans l’ensemble d’entraînement et de validation est significativement plus grande que celle dans l’ensemble de test. Cette approche est standard, car plus de données en entraînement permettent au modèle d’apprendre des représentations plus robustes, tandis que l’ensemble de test est utilisé pour évaluer de manière impartiale les performances du modèle.

4.2.2 Résultats et discussion

Dans cette approche, nous réutilisons les modèles pré-entraînés VGG-19 et MobileNetV2, précédemment exploités au chapitre 3. L’architecture du réseau ainsi que les modalités d’apprentissage par transfert, incluant l’ajustement fin, sont détaillées dans la figure 3.11 du chapitre 3.

En ce qui concerne les hyper-paramétrés du réseau, la valeur de `batch_size` est toujours fixée à 64, le taux d’apprentissage (`learning_rate`) est fixé à 0,0001. De plus, nous utilisons l’optimiseur Adam ainsi que la fonction de perte Focal Crossentropy binaire (Binary Focal Crossentropy).

4.2.2.1 Résultats d’entraînement et de validation

Les figures 4.4 et 4.5 illustrent respectivement l’accuracy et la perte de notre modèle CNN proposé, basé sur VGG-19 et MobileNetV2, sur les ensembles d’entraînement et de validation, au fil des époques de l’apprentissage. L’accuracy d’entraînement augmente régulièrement tout au long du processus d’apprentissage, tandis que l’accuracy de validation peut plafonner ou même diminuer après un certain nombre d’époques.

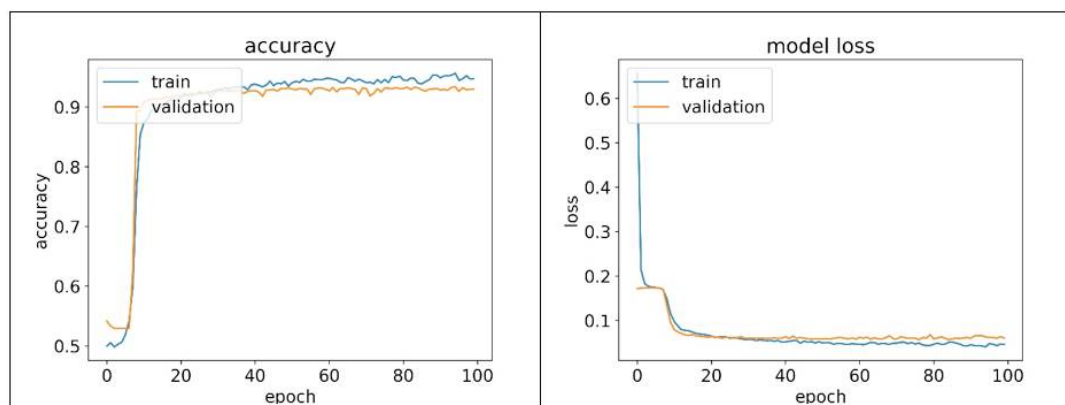


Figure 4.4 – Accuracy/Perte de modèle basé sur VGG-19 vs. époque.

Pour la courbe de l’accuracy de la figure 4.4, on observe une augmentation rapide au début, tant pour l’ensemble d’entraînement que pour l’ensemble de validation, indiquant que le modèle apprend efficacement dès les premières itérations. Vers la 20ème époque, l’accuracy se stabilise autour de 92% pour les deux ensembles, ce

qui suggère que le modèle atteint un point de convergence. La courbe relativement plate après la stabilisation montre que le modèle ne subit pas de surapprentissage significatif, car les performances sur l'ensemble de validation restent cohérentes avec celles de l'entraînement.

De la même manière, pour la courbe qui présente l'évolution de la perte du modèle en fonction des époques, nous observons que pour les deux ensembles (entraînement et validation) la perte diminue rapidement au début, ce qui est un bon signe indiquant que le modèle apprend rapidement dans les premières époques. La perte finale est très faible (environ 0.1), ce qui suggère que le modèle est performant.

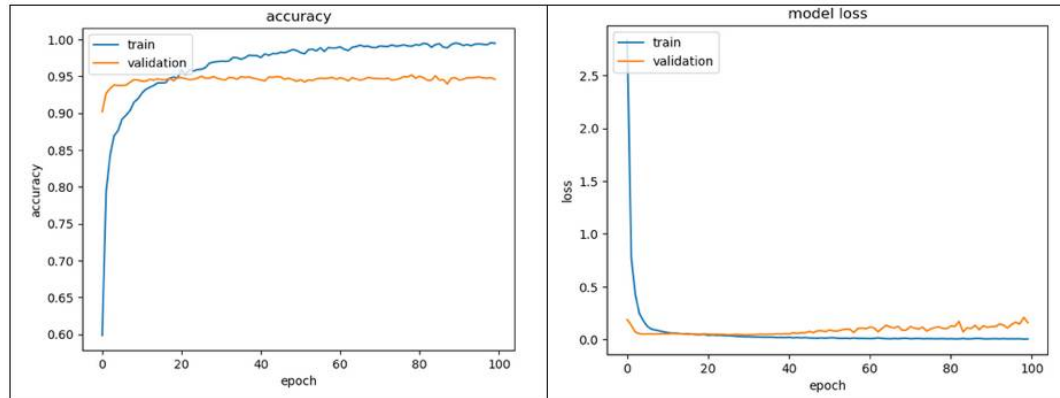


Figure 4.5 – Accuracy/Perte de modèle basé sur MobileNetV2 vs. époque.

Pour l'évolution de l'accuracy du modèle basé sur MobileNetV2 en fonction des époques pour l'ensemble d'entraînement et de validation, nous remarquons qu'au début, l'accuracy de l'entraînement augmente rapidement, atteignant environ 95% en moins de 15 époques. Cela indique que le modèle apprend rapidement les caractéristiques importantes des données. Et celle de l'ensemble de validation atteint également un seuil autour de 94% après environ 10 époques. Cela montre que le modèle a une bonne capacité de généralisation aux données non utilisées. La courbe de l'accuracy de l'entraînement continue de croître lentement, atteignant presque 99% à la fin des 100 époques. En outre, il y a un léger écart entre la valeur obtenue à partir de l'entraînement et celle de la validation (environ 0.05), ce qui est courant mais peut suggérer un début de surapprentissage (overfitting). Cependant, il est encourageant de noter que cet écart reste relativement stable.

Pour la perte, nous constatons que la perte d'entraînement tombe près de zéro, ce qui signifie que le modèle s'ajuste très bien aux données d'entraînement. Cela peut cependant indiquer un surapprentissage si la perte de validation n'est pas aussi faible. À partir de la 60ème époque environ, la courbe de perte de validation commence à montrer une légère tendance à la hausse, mais puisqu'il y a un léger écart entre les pertes d'entraînement et de validation (environ 0.01), cela suggère que le modèle n'est pas en train de sur-apprendre de manière significative.

Pour les époques ultérieures, le taux d'amélioration de l'accuracy et de la perte est ralenti et se stabilise à la fin. Selon la figure 4.6 (partie précision), nous avons obtenu un taux de validation de 93.11 % en utilisant le modèle basé sur VGG-19 et un taux de validation de 94.58 % en utilisant le modèle basé sur MobileNetV2.

La figure 4.6 montre également une comparaison détaillée entre le modèle basé sur VGG-19 et le modèle basé sur MobileNetV2. Trois mesures ont été prises en

compte : l'accuracy du modèle, l'AUC et la précision. Nous observons principalement une légère différence entre les deux modèles pour les performances en termes d'AUC et de précision. Plus précisément, lorsque nous utilisons le modèle basé sur le VGG-19, nous obtenons une AUC de validation de 94.83 % et une précision de 97.2 %. Cependant, les performances du modèle basé sur MobileNetV2 s'améliorent encore, avec une AUC de validation égale à 96.8 % et une précision égale à 98.1 %. Ce résultat confirme la supériorité du modèle basé sur MobileNetV2 sur l'autre dans le contexte du diagnostic du mélanome avec l'utilisation de la technique d'augmentation de donnée.

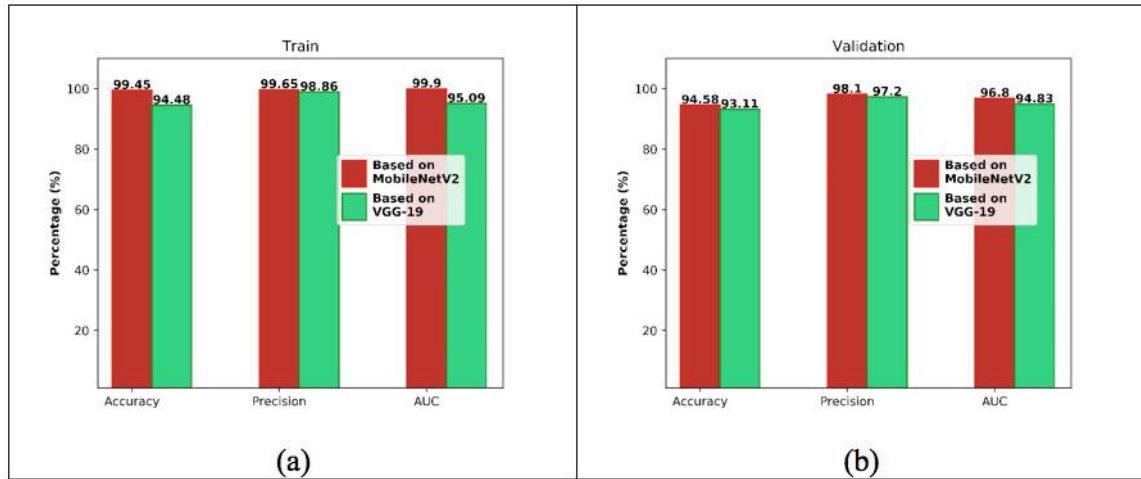


Figure 4.6 – : Métriques de performance pour les deux modèles : (a) l'entraînement et (b) validation.

4.2.2.2 Test et évaluation

Les performances des modèles proposés sont évaluées en termes d'accuracy, de précision, d'AUC, de sensibilité et de spécificité. Ces métriques ont été calculées sur un ensemble de test composé de 2422 images réparties en deux classes, 1254 pour la classe 0 et 1168 pour la classe 1. Les performances globales du modèle sont brièvement présentées dans le tableau 4.4, tandis que la matrice de confusion est présentée à la figure 4.7.

Pour le modèle basé sur VGG-19 (figure 4.7-(a)), on constate que le classificateur a une très bonne capacité à prédire les cas bénins (True Negative élevé, 97%). Cela signifie qu'il est peu susceptible de classer à tort une tumeur bénigne comme maligne. Il comporte également un taux de faux négatifs (False Negative) peu élevé (10.3%), ce qui signifie que certains échantillons malins ont été incorrectement classés comme bénins. Cela explique qu'il a tendance à manquer certains cas de tumeurs malignes, ce qui peut avoir des conséquences cliniques importantes.

Comparé au modèle VGG-19, la matrice de confusion du modèle basé sur MobileNetV2 (figure 4.7-(b)), offre une performance globale légèrement supérieure avec des taux de faux positifs et de faux négatifs plus faibles, démontrant ainsi une meilleure précision dans la classification des deux catégories. Bien que les performances soient améliorées par rapport à VGG-19, le modèle peut encore manquer quelques cas de tumeurs malignes (False Negative : 9.2% ou 107 cas).

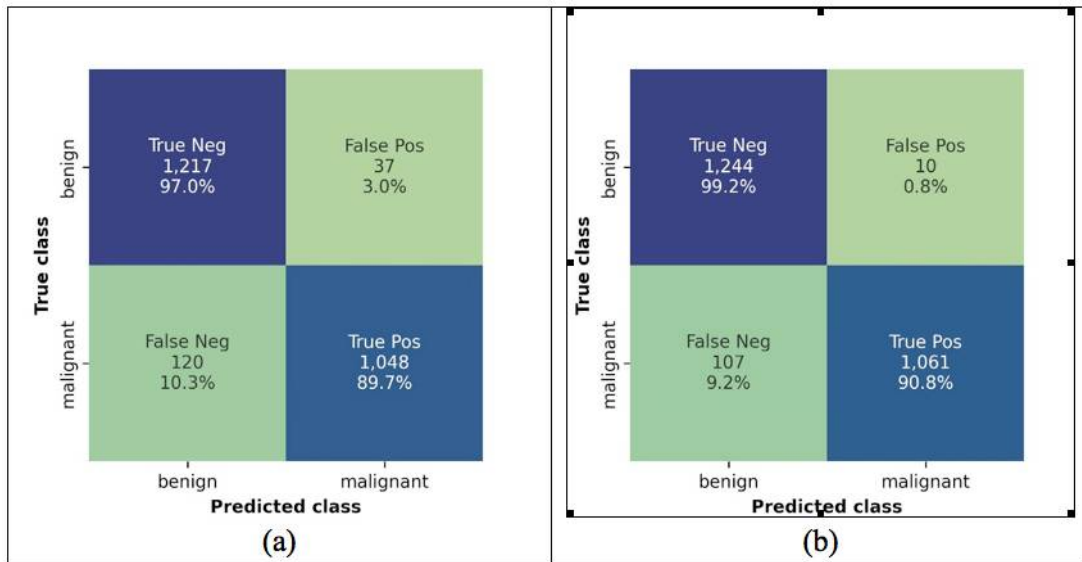


Figure 4.7 – Matrice de confusion du modèle basé sur, (a) : VGG-19 et (b) : MobileNetV2.

	Accuracy (%)	AUC (%)	Précision (%)	Sensibilité (%)	Spécificité (%)
Basé sur VGG19	93.51	95.20	96.58	89.72	97.04
Basé sur MobileNetV2	95.16	97.57	99.06	90.83	99.20

Tableau 4.4 – Performance des modèles basés sur VGG19, et MobileNetV2.

Le modèle basé sur MobileNetV2 semble offrir de meilleures performances globales que VGG19, en particulier en termes d’accuracy, de précision, et de spécificité.

Le modèle basé sur VGG19 atteint un accuracy de 93.51 %, tandis que celui basé sur MobileNetV2 affiche une valeur plus élevée de 95.16 %. Cela indique que MobileNetV2 a un taux de classification global plus précis par rapport à VGG19. Une précision plus élevée pour MobileNetV2 signifie qu’il a moins de faux positifs, ce qui est crucial dans un contexte médical pour éviter les diagnostics erronés. La sensibilité mesure la capacité du modèle à identifier correctement les échantillons malins. MobileNetV2 montre une légère amélioration par rapport à VGG19, ce qui signifie qu’il est légèrement meilleur pour détecter les cas malins. Une spécificité plus élevée pour MobileNetV2 signifie qu’il est plus efficace pour éviter de classer incorrectement les échantillons bénins comme malins.

Pour la métrique AUC, le modèle basé sur VGG19 est égal à 95.20 %, tandis que celui pour MobileNetV2 est de 97.57 %. Un AUC plus élevé signifie que MobileNetV2 est meilleur pour différencier les échantillons bénins et malins, ce qui suggère une meilleure performance globale en termes de discrimination entre les classes.

En résumé, les deux modèles basés sur VGG19 et MobileNetV2 offrent de bonnes performances pour la tâche de classification considérée. Cependant, MobileNetV2 semble être légèrement plus performant en termes de précision et de spécificité.

4.3 Approche hybride CNN et LSTM basée sur l'apprentissage par transfert

L'objectif de cette approche est de proposer une architecture hybride spécifique, CNN-LSTM, combinant les forces des réseaux de neurones convolutifs (CNN) et des réseaux de neurones récurrents (RNN), plus précisément des réseaux LSTM (Long Short-Term Memory). Cette hybridation est souvent utilisée pour des tâches nécessitant la capture de caractéristiques spatiales (images) et temporelles (séquences).

Le modèle hybride combine un CNN pré-entraîné pour extraire des caractéristiques à partir d'images, suivi d'une couche LSTM qui apprend les dépendances temporelles dans ces caractéristiques. La structure du modèle est adaptée pour capturer des informations complexes des images tout en étant régularisée pour éviter le surapprentissage. Ce type de modèle peut être utilisé dans des applications où les données sont spatiales (images) et temporelles (séquences d'images ou séries temporelles).

4.3.1 Architecture du réseau hybride CNN et LSTM

Le diagramme de la figure 4.8 illustre une architecture hybride dans laquelle un réseau neuronal convolutif (CNN) est combiné à une couche LSTM. Ce diagramme montre une architecture complexe qui combine la puissance d'extraction de caractéristiques des CNN avec la capacité des LSTM à capturer des dépendances temporelles ou séquentielles.

Cette approche tire parti de deux architectures CNN différentes, VGG19 et MobileNetV2, pour extraire des caractéristiques à partir d'images d'entrée. Ensuite, ces caractéristiques sont passées à une couche LSTM pour capturer les dépendances temporelles ou séquentielles, avant d'être traitées par des couches entièrement connectées ajoutées pour la classification finale.

Ces deux architectures sont des CNN pré-entraînés sur d'immenses bases de données d'images (ImageNet). Elles sont utilisées ici pour extraire des caractéristiques pertinentes des images d'entrée (224x224 pixels).

Dans cette deuxième approche, nous avons utilisé le même jeu de données ISIC2020, avec un prétraitement et une préparation des données reposant toujours sur la technique d'augmentation que celle décrite dans l'approche précédente (voir Figure 4.1). Nous avons également appliqué la même stratégie de partitionnement des données, comme le montre le tableau 4.3.

dèle pré-entraîné, MobileNetV2 et VGG19, respectivement. Elles sont utilisées pour convertir les caractéristiques 2D en un vecteur 1D, prêt à être utilisé par la couche suivante.

Pour la couche LSTM ajoutée, celle-ci constitue le cœur du modèle récurrent. Elle a la capacité de mémoriser des informations à long terme et de modéliser les dépendances temporelles entre les différentes entrées. Le vecteur de caractéristiques extrait par l'une des architectures CNN (selon le modèle utilisé MobileNetV2 ou VGG19) est passé à une couche LSTM avec 256 unités. La régularisation L2 est appliquée aux poids (0.01) pour éviter le surapprentissage. De même pour éviter le surapprentissage, une couche Dropout est appliquée après la LSTM, désactivant aléatoirement 50 % des unités.

Les sorties de la couche LSTM sont ensuite passées à plusieurs couches denses ajoutées (FC layers) qui effectuent la classification finale, où chaque couche est suivie d'une couche Dropout pour régulariser le modèle et éviter le surapprentissage.

La dernière couche utilise une fonction d'activation Sigmoid pour produire une sortie de probabilité binaire, indiquant la classe prédite (0 : bénigne ou 1 : maligne).

4.3.2 Résultats et discussion

Dans l'approche proposée, nous avons conservé les mêmes hyperparamètres utilisés dans l'approche précédente. La taille des lots (batch size) reste fixée à 64, et le taux d'apprentissage (learning rate) est défini à 0,0001. De plus, nous utilisons l'optimiseur Adam et la fonction de perte Focal Crossentropy binaire (Binary Focal Crossentropy).

4.3.2.1 Résultats d'entraînement et de validation

La figure 4.9 présente les courbes de l'accuracy et de perte pour un modèle basé sur VGG-19-LSTM, en fonction des époques d'entraînement. Chaque graphique comporte deux courbes, une pour les données d'entraînement et une autre pour les données de validation. Pour la courbe du jeu d'entraînement, l'accuracy progresse de manière constante et se stabilise autour de 98%, indiquant que le modèle s'ajuste correctement aux données d'entraînement. En revanche, l'accuracy du jeu de validation atteint un pic au début (autour de 95%), puis oscille entre 90% et 93%. Ces variations pourraient suggérer que le modèle commence à surapprendre après un certain nombre d'époques, ce qui se traduit par une moins bonne généralisation aux données de validation comparativement aux données d'entraînement.

La perte diminue rapidement au départ, puis continue à décroître lentement, ce qui montre que le modèle apprend à mieux minimiser l'erreur sur le jeu d'entraînement. Pour le jeu de validation, la perte diminue également au début, mais finit par se stabiliser, avec une légère tendance à augmenter ou à fluctuer. Cela suggère que le modèle commence à perdre en capacité de généralisation à partir d'un certain point. Les fluctuations et la stagnation de l'accuracy et de la perte de validation indiquent un possible surapprentissage.

De la même façon, nous allons expliquer les courbes de l'accuracy et de la perte pour un modèle basé sur MobileNetV2-LSTM, en fonction des époques d'entraînement, comme illustré par la figure 4.10. Dans ce cas, l'accuracy des données d'entraînement augmente très rapidement dès les premières époques, se stabilisant autour de 95%. Cela montre que le modèle s'adapte efficacement aux données d'entraînement. Pour les données de validation, il atteint également un niveau élevé très tôt et

fluctue légèrement autour de 94% à 95%. Cette stabilité avec de légères fluctuations pourrait indiquer que le modèle généralise relativement bien. Cela peut être interprété comme le fait que la perte sur le jeu de validation suit une tendance similaire à celle de l'entraînement : elle diminue rapidement avant de se stabiliser. La proximité des pertes d'entraînement et de validation indique que le modèle ne souffre pas de surapprentissage et conserve une bonne capacité de généralisation.

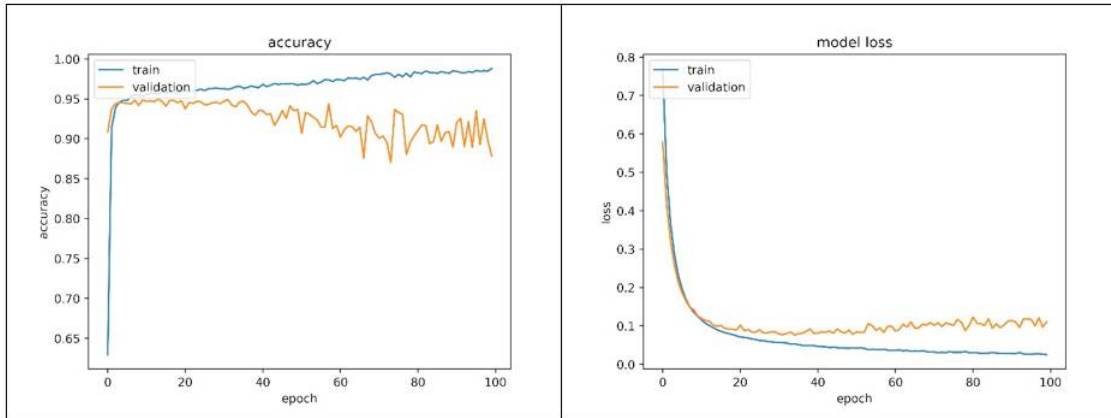


Figure 4.9 – Accuracy/Perte de modèle basé sur VGG-19-LSTM Vs. époque.

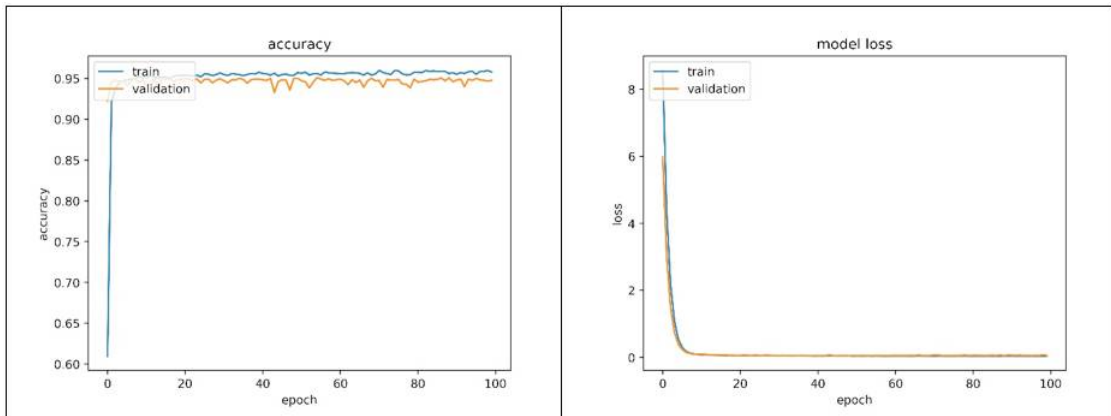


Figure 4.10 – Accuracy/Perte de modèle basé sur MobileNetV2-LSTM Vs. époque.

Afin d'effectuer une comparaison plus rigoureuse des deux modèles hybrides CNN-LSTM, nous avons non seulement suivi l'évolution de l'accuracy et de la perte au cours des données l'entraînement et de validation, mais nous avons également calculé des métriques complémentaires telles que la précision et l'AUC. La figure 4.11 compare les performances de deux modèles hybrides, l'un basé sur MobileNetV2-LSTM et l'autre sur VGG19-LSTM, pour les phases d'entraînement et de validation.

Globalement, le modèle basé sur MobileNetV2-LSTM semble performant par rapport au modèle VGG19-LSTM sur toutes les métriques et pour les deux ensembles de données (entraînement et validation). Le modèle VGG19-LSTM a un accuracy plus élevée (98.78%) que le modèle MobileNetV2-LSTM (95.79%), ce qui suggère que VGG19-LSTM a mieux appris les données d'entraînement. Mais, le mo-

dèle MobileNetV2-LSTM maintient une meilleure accuracy sur les données de validation (94.76%) par rapport à VGG19-LSTM (88.83%).

Bien que MobileNetV2-LSTM ait une meilleure précision, VGG19-LSTM présente une AUC légèrement plus élevée, ce qui indique une meilleure capacité à distinguer entre les classes sur l'ensemble d'entraînement. MobileNetV2-LSTM surpasse également VGG19-LSTM en termes de précision (98.62% contre 87.2%) et AUC (97.25% contre 96.92%). Cela indique que MobileNetV2-LSTM est plus performant et plus robuste sur les données non vues.

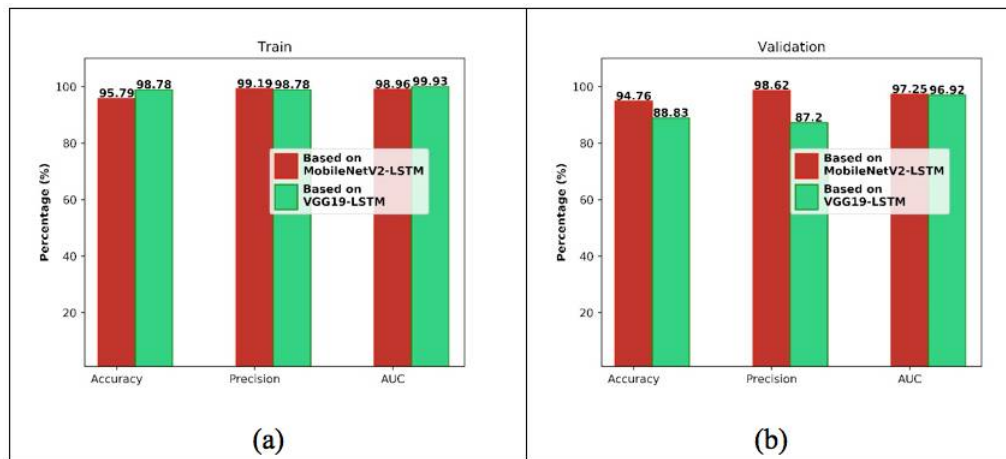


Figure 4.11 – Métriques de performance pour les deux modèles : (a) l'entraînement et (b) validation.

Ici le modèle basé sur MobileNetV2-LSTM semble bien équilibré entre l'apprentissage et la généralisation. La faible différence entre la précision et la perte pour l'entraînement et la validation indique une bonne performance du modèle sur des données non vues. Le modèle VGG-19-LSTM est capable d'apprendre les caractéristiques des données et d'atteindre une bonne performance. Le modèle semble souffrir de sur-apprentissage à partir d'un certain nombre d'époques. Cela est dû au fait que le modèle devient trop complexe et mémorise le bruit présent dans les données d'entraînement, ce qui diminue sa capacité à généraliser à de nouvelles données.

MobileNetV2-LSTM semble mieux généralisé sur les données de validation par rapport à VGG19-LSTM, bien que VGG19-LSTM ait montré de meilleures performances lors de l'entraînement. Cela pourrait indiquer que VGG19-LSTM est plus susceptible de surapprendre (overfitting) sur les données d'entraînement, tandis que MobileNetV2-LSTM trouve un meilleur équilibre entre l'apprentissage et la généralisation.

La couche LSTM, présente dans les deux modèles, joue un rôle crucial dans la capture des dépendances temporelles et contextuelles. Les résultats suggèrent que l'intégration de l'LSTM avec les deux architectures CNN améliore significativement les performances. L'analyse comparative des performances entre les modèles basés sur MobileNetV2, utilisant uniquement la technique d'augmentation des données comme dans la première approche (Figure 4.6), et le modèle hybride MobileNetV2-LSTM de la deuxième approche (Figure 4.11) révèle des différences significatives dans leur comportement pendant l'entraînement et la validation.

Bien que le modèle basé sur MobileNetV2 soit performant, le modèle hybride

MobileNetV2-LSTM semble offrir de meilleures capacités de généralisation, ce qui est essentiel pour les applications pratiques. Cela suggère que l'intégration du LSTM améliore la généralisation sur des données non vues, notamment lorsqu'il existe des dépendances temporelles ou séquentielles à capturer dans les données.

4.3.2.2 Test et évaluation

Grâce à la matrice de confusion, nous pouvons évaluer la précision du modèle en analysant le nombre de cas correctement classés comme bénins ou malins. Les valeurs sur la diagonale principale représentent les prédictions correctes, confirmant ainsi l'efficacité de notre approche proposée. Les matrices de confusion présentées à la Figure 4.12 permettent d'évaluer la précision de nos modèles dans la classification des images mélanome.

En ce qui concerne les images bénignes, MobileNetV2 a démontré une plus grande précision en classifiant correctement 1240 images sur 1254 (98.6%), contre 1210 images pour VGG19-LSTM (96.5%). Pour les images malignes, MobileNetV2 a également obtenu de meilleurs résultats avec un taux de réussite de 94.5% (1023/1168), comparé à 87.6% pour VGG19-LSTM (1023/1168).

Pour une évaluation plus approfondie, nous complétons notre analyse avec d'autres métriques comme l'AUC, la précision, la sensibilité et la spécificité, comme le montre le tableau 4.5.

Le modèle MobileNetV2-LSTM affiche une accuracy plus élevée, ce qui indique qu'il permet davantage dans la classification correcte des exemples en classes bénignes et malignes. De plus, avec un AUC supérieur, il montre une meilleure capacité à différencier les classes positives (malignes) des classes négatives (bénignes), quel que soit le seuil de classification. Par ailleurs, la précision plus élevée de MobileNetV2-LSTM (98.31%) suggère qu'il produit moins de faux positifs que le modèle VGG19-LSTM (94.17%), renforçant ainsi sa fiabilité globale. Le modèle MobileNetV2-LSTM affiche d'excellentes performances en termes de sensibilité (94.5%), ce qui signifie qu'il est capable de détecter la majorité des cas malins. De plus, sa spécificité élevée (98.6%) garantit une faible proportion de faux positifs, minimisant ainsi les erreurs de diagnostic pour les cas bénins.

En général, MobileNetV2-LSTM a significativement amélioré les performances de classification, en particulier pour les images malignes, atteignant un taux de réussite de 94.5% contre 87.6% pour VGG19-LSTM.

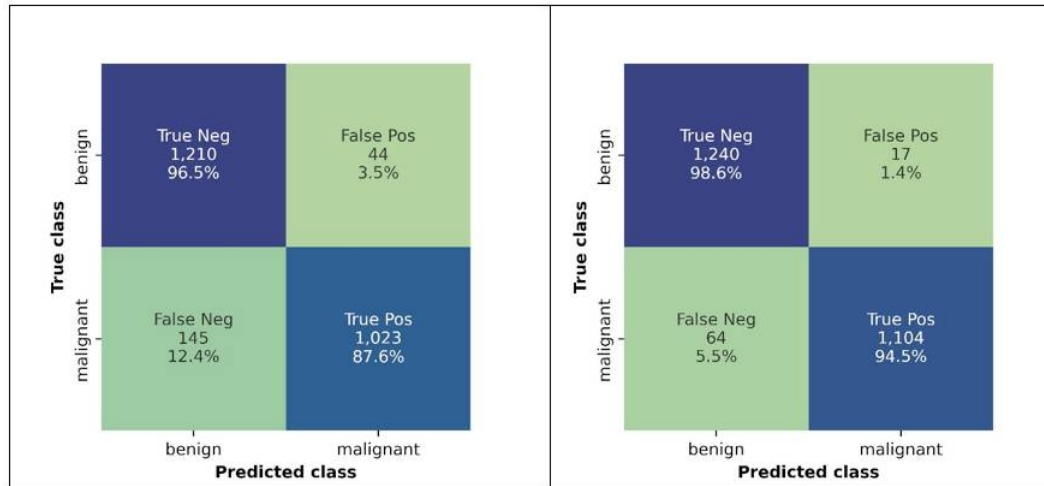


Figure 4.12 – Matrice de confusion du modèle basé sur, (a) : VGG19-LSTM et (b) : MobileNetV2-LSTM.

	Accuracy (%)	AUC (%)	Précision (%)	Sensibilité (%)	Spécificité (%)
Basé sur VGG19-LSTM	88.31	95.28	94.17	87.62	96.50
Basé sur MobileNetV2-LSTM	95.67	97.71	98.31	94.50	98.60

Tableau 4.5 – Performance des modèles basé sur VGG19-LSTM, et MobileNetV2-LSTM.

Les deux modèles montrent une bonne précision, mais le modèle MobileNetV2-LSTM présente une performance globale meilleure, avec une précision plus élevée, en particulier dans l'identification correcte des cas bénins et malins.

4.4 Évaluation comparative des performances : État de l'art vs. Nos approches proposées

Nos approches proposées sont évaluées et comparées aux méthodes de l'état de l'art vu dans le chapitre 2 (tableau 2.4), démontrant ainsi son potentiel pour améliorer le diagnostic du mélanome. Le tableau 4.6 offre une synthèse des résultats obtenus sur le dataset ISIC 2020, permettant de comparer les performances de nos approches avec celles des méthodes de l'état de l'art en deep learning pour le diagnostic du mélanome. Nous utilisons diverses métriques d'évaluation telles que l'accuracy, l'AUC, la précision, la sensibilité et la spécificité.

	Accuracy (%)	AUC (%)	Précision (%)	Sensibilité (%)	Spécificité (%)
Srinivasu et al. [78]	85.34	–	–	88.24	92
Reis et al. [87]	90.54	–	–	93.33	88.64
Wan et al. [88]	95.60	92.40	–	59.6	97.40
Mijwil [89]	86.90	–	87.47	86.14	87.66
Basé sur VGG19	93.51	95.20	96.58	89.72	97.04
Basé sur MobilNetV2	95.16	97.57	99.06	90.83	99.20
Basé sur VGG19-LSTM	88.31	95.28	94.17	87.62	96.50
Basé sur MobilNetV2-LSTM	95.67	97.71	98.31	94.50	98.60

Tableau 4.6 – Performance des approches proposées par rapport aux techniques de références.

De manière générale, les modèles des deux approches proposées, basés sur VGG19 et MobileNetV2, notamment ceux combinés avec LSTM, affichent généralement de meilleures performances que les modèles de Srinivasu et al. [78], Reis et al. [87], et Wan et al. [88]. Les modèles proposés offrent un bon équilibre entre la précision et sensibilité. Cela signifie qu'ils sont capables de correctement classifier à la fois les images de mélanomes et les images bénignes. De même, l'AUC élevé de la plupart des modèles indique une bonne capacité à discriminer entre les classes maligne et bénigne, indépendamment du seuil de classification.

Tous les modèles affichent de bonnes valeurs de spécificité, ce qui témoigne de leur capacité à identifier correctement les images bénignes. Cependant, pour la détection de la majorité des cas de mélanome, seul le modèle basé sur MobileNetV2-LSTM se distingue avec une sensibilité optimale de 94.5 %. En termes d'accuracy, d'AUC et de sensibilité, ce modèle atteint également les meilleures performances, ce qui suggère une meilleure capacité de généralisation. En effet, cette approche combine efficacement les forces de MobileNetV2 avec les LSTM, offrant ainsi des performances exceptionnelles sur toutes les métriques, particulièrement en matière de sensibilité, faisant de cette méthode une solution extrêmement robuste pour le diagnostic du mélanome.

4.5 Conclusion

Le chapitre précédent a souligné le problème du déséquilibre des classes dans notre ensemble de données ISIC 2020. Pour pallier cette limitation et obtenir des modèles plus robustes, nous avons exploré deux approches principales, l'augmentation de données et la conception d'un modèle hybride CNN-LSTM. Ces approches visent

à améliorer la robustesse et la précision de nos modèles de classification pour la détection des lésions malignes.

Nous avons utilisé la technique d'augmentation des données. Le modèle basé sur MobileNetV2 semble offrir une meilleure précision globale, en particulier pour éviter les faux positifs, ce qui est crucial dans un contexte médical où le sur-diagnostic peut avoir des conséquences importantes. En revanche, le modèle basé sur VGG-19, bien que performant, montre des faiblesses par rapport à MobileNetV2, en particulier dans la détection des cas malins.

Nous avons poussé plus loin nos investigations où nous avons exploré l'intégration de réseaux récurrents (LSTM) à ces architectures. Les résultats obtenus avec le modèle MobileNetV2-LSTM montre une performance supérieure à tous les niveaux par rapport au VGG19-LSTM. Avec des scores plus élevés en termes d'accuracy, AUC, précision, sensibilité, et spécificité, le MobileNetV2-LSTM apparaît comme un modèle plus robuste pour la classification des images bénignes et malignes. Ce modèle serait donc préféré pour des applications où la précision et la fiabilité des prédictions sont cruciales.

Comparé aux méthodes utilisées dans l'état de l'art, notre approche, en particulier l'utilisation du modèle MobileNet-LSTM, se distingue par une meilleure gestion du déséquilibre des classes, conduisant à des modèles plus robustes. Grâce à l'augmentation des données et à la combinaison des forces de MobileNet et LSTM, notre méthode surpasse les performances des techniques existantes pour la détection des lésions malignes.

Conclusion générale

Le cancer de la peau est une maladie où des cellules anormales se développent dans la peau. Il existe plusieurs types de cancer de la peau, les plus courants étant le carcinome basocellulaire, le carcinome spinocellulaire et le mélanome. Le mélanome, s'il n'est pas détecté à un stade précoce, peut rapidement évoluer vers une forme plus agressive, mettant en jeu le pronostic vital. Toutefois, certains types de mélanomes restent difficiles à classer, comme ceux des patients présentant de multiples grains de beauté atypiques et des mélanomes nodulaires. L'incidence croissante du mélanome à l'échelle mondiale souligne l'urgence de développer des outils de diagnostic précoce efficaces. Une détection rapide permet d'instaurer un traitement adapté et d'augmenter significativement les chances de guérison.

L'IA révolutionne le diagnostic des tumeurs malignes de la peau en permettant à un plus grand nombre de professionnels de santé de contribuer à une détection précoce et efficace. Le développement de modèles de classification CNN pour le cancer de la peau, et le mélanome dans ce contexte, est motivé par l'objectif d'améliorer la santé publique, d'améliorer l'accessibilité aux soins de santé, de réduire les coûts des soins de santé et de faire progresser les connaissances médicales, tout en utilisant la puissance de la technologie et de l'innovation.

Cette thèse aborde la tâche cruciale de l'identification et de la classification du mélanome malin, en utilisant l'ensemble de données SIIM-ISIC 2020 et un réseau de neurones convolutionnels profonds (CNN). Nous avons proposé de mettre en œuvre la technique d'augmentation des données pour résoudre le problème de déséquilibre des classes dans l'ensemble de données ISIC 2020, qui a été un problème courant pour les chercheurs appliquant des techniques d'apprentissage supervisé. Nous avons également développé un modèle de classification d'images spécifique à notre application, ainsi que des variantes ajustées et appliquées aux réseaux neuronaux pré-entraînés VGG-19 et MobileNetV2. Nous avons également intégré une architecture hybride CNN-LSTM utilisant les mêmes modèles pré-entraînés. Cette approche a été testée de manière approfondie en évaluant des métriques de performance clés telles que l'exactitude, la précision, la spécificité, la sensibilité et l'AUC. Par ailleurs, la technique d'augmentation de données est une technique essentielle pour améliorer la performance des modèles de classification en dermatologie, notamment lorsque les données sont déséquilibrées. En combinant des modèles différents, comme dans notre cas avec CNN et LSTM, et en évaluant soigneusement les résultats, il est possible de construire des modèles plus robustes et plus précis pour la détection des lésions malignes.

Notre première approche a consisté à combiner des techniques d'augmentation de

données avec l'apprentissage par transfert et des architectures de CNN pré-entraînées (MobileNetV2 et VGG19) pour améliorer la robustesse et la généralisation de nos modèles. Le modèle basé sur VGG19 pourrait être utilisé comme outil de dépistage initial, où une sensibilité élevée pour les cas bénins est importante pour éviter des biopsies inutiles. Cependant, il est essentiel de combiner les résultats de ce modèle avec d'autres examens pour confirmer les cas suspects de malignité. Par contre le modèle basé sur MobileNetV2 semble offrir une meilleure précision globale et pourrait être préféré pour un diagnostic plus définitif. Néanmoins, il est important de noter que même ce modèle n'est pas parfait et que les faux négatifs restent une préoccupation.

Nous avons adopté une deuxième approche qui combine des modèles CNN pré-entraînés ou l'apprentissage par transfert avec des LSTM. Cette combinaison s'avère particulièrement prometteuse pour la classification des images médicales, surtout dans un contexte où la précision et la fiabilité sont cruciales. Les résultats obtenus suggèrent que les modèles basés sur MobileNetV2-LSTM offrent les meilleures performances pour la tâche de diagnostic du mélanome sur le dataset ISIC 2020. Ces modèles combinent les avantages des architectures convolutives profondes (MobileNetV2) avec la capacité de modélisation séquentielle des réseaux récurrents (LSTM), permettant ainsi une meilleure extraction des caractéristiques pertinentes et une classification plus précise.

Les résultats montrent que les approches proposées, notamment celles basées sur MobileNetV2 et MobileNetV2-LSTM, surpassent les techniques de référence en termes d'accuracy, d'AUC, de précision, de sensibilité, et de spécificité. Ces approches offrent une meilleure capacité de diagnostic, ce qui les rend particulièrement efficaces pour la détection du mélanome dans le dataset ISIC 2020. La méthode MobileNetV2-LSTM, en particulier, se démarque par sa performance globale, ce qui en fait une approche prometteuse pour les applications futures.

Comparé aux méthodes utilisées dans l'état de l'art, nos approches offrent une meilleure gestion du déséquilibre des classes et permettent d'obtenir des modèles plus robustes. En particulier, l'utilisation de techniques telles que l'augmentation des données et la combinaison de modèles, comme le réseau CNN-LSTM, nous a permis de surpasser les performances des méthodes existantes en matière de détection des lésions malignes. Nous avons intégré cette architecture hybride avec une technique d'augmentation des données, nous avons pu surmonter les limitations des méthodes existantes, offrant ainsi une meilleure précision pour la détection des lésions malignes. Cette combinaison innovante permet de surpasser les performances des techniques actuelles, renforçant la robustesse et l'efficacité de notre modèle. Cette étude souligne le potentiel de l'apprentissage profond dans l'amélioration de la précision du diagnostic du mélanome, améliorant ainsi les soins aux patients.

Les perspectives de recherche futures devraient inclure l'exploration des approches multimodales, qui intègrent les images dermatologiques avec d'autres types de données, afin de développer des modèles de classification plus performants et robustes. Cette stratégie tirerait parti des complémentarités entre différentes sources d'information, améliorant ainsi la précision des diagnostics de mélanome. De plus, l'utilisation de l'apprentissage par renforcement, y compris l'apprentissage par renforcement profond, pourrait améliorer la prise de décision et permettre une adaptation plus efficace des modèles en fonction des retours cliniques en temps réel.

L'avenir de la classification d'images médicales s'annonce prometteur grâce aux avancées de l'intelligence artificielle et aux collaborations entre chercheurs, cliniciens et ingénieurs. En combinant ces différentes approches, il sera possible de développer

des outils de diagnostic de plus en plus précis et fiables, améliorant ainsi la prise en charge des patients.

Liste des publications

- Journaux internationaux avec comité de lecture

- [1] M'HAMEDI, Mohammed, MERZOUG, Mohammed, HADJILA, Mourad, et al. Enhancing melanoma skin cancer classification through data augmentation. TELKOMNIKA (Telecommunication Computing Electronics and Control), 2024, vol. 22, no 5, p. 1209-1220.
- [2] NEDJAR, Imane et M'HAMEDI, Mohammed. Interactive system based on artificial intelligence and robotic arm to enhance arabic sign language learning in deaf children. Education and Information Technologies, 2024, p. 1-18.
- [3] NEDJAR, Imane, M'HAMEDI, Mohammed, et BEKKAOUI, Mokhtaria. Real-Time Solid Waste Sorting Machine Based on Deep Learning. International journal of electrical and computer engineering systems, 2024, vol. 15, no 7, p. 581-589.

Bibliographie

- [1] A. Alani, M. Sadlier, A. Uddin, C. Hackett, B. Ramsay, and K. Ahmad, “An analysis of inpatient dermatologic consultations at university hospital limerick : inadequate infrastructure leads to acute skin failure,” *Irish Journal of Medical Science (1971-)*, vol. 186, pp. 305–307, 2017.
- [2] L. C. Parish, “Dermatologic therapy for the 21st century,” 2000.
- [3] E. M. Miró and N. P. Sánchez, “Cutaneous manifestations of infectious diseases,” *Atlas of dermatology in internal medicine*, pp. 77–119, 2012.
- [4] L.-R. PLAN and N. S. S. Areas, “National institute of arthritis and musculoskeletal and skin diseases (niams),” 2010.
- [5] C. Karimkhani, L. N. Boyers, R. P. Dellavalle, and M. A. Weinstock, “It’s time for “keratinocyte carcinoma” to replace the term “nonmelanoma skin cancer”,” *Journal of the American Academy of Dermatology*, vol. 72, no. 1, pp. 186–187, 2015.
- [6] M. Ernst and A. Giubellino, “The current state of treatment and future directions in cutaneous malignant melanoma,” *Biomedicines*, vol. 10, no. 4, p. 822, 2022.
- [7] C. M. Balch, A. C. Buzaid, S.-J. Soong, M. B. Atkins, N. Cascinelli, D. G. Coit, I. D. Fleming, J. E. Gershenwald, A. Houghton Jr, J. M. Kirkwood *et al.*, “Final version of the american joint committee on cancer staging system for cutaneous melanoma,” *Journal of Clinical Oncology*, vol. 19, no. 16, pp. 3635–3648, 2001.
- [8] C. Papageorgiou, D. Ioannides, Z. Apalla, E. Vakirlis, E. Lazaridou, E. Sotiriou, and A. Lallas, “Dermoscopy of difficult-to-diagnose melanomas,” *Serbian Journal of Dermatology and Venereology*, vol. 8, no. 3, pp. 121–127, 2016.
- [9] C. Longo, R. Pampena, E. Moscarella, J. Chester, M. Starace, E. Cinotti, B. M. Piraccini, G. Argenziano, K. Peris, and G. Pellacani, “Dermoscopy of melanoma according to different body sites : Head and neck, trunk, limbs, nail, mucosal and acral,” *Journal of the European Academy of Dermatology and Venereology*, vol. 37, no. 9, pp. 1718–1730, 2023.
- [10] H. Hassani, E. S. Silva, S. Unger, M. TajMazinani, and S. Mac Feely, “Artificial intelligence (ai) or intelligence augmentation (ia) : what is the future ?” *Ai*, vol. 1, no. 2, p. 8, 2020.
- [11] K. Manheim and L. Kaplan, “Artificial intelligence : Risks to privacy and democracy,” *Yale JL & Tech.*, vol. 21, p. 106, 2019.
- [12] S. J. Russell and P. Norvig, *Artificial intelligence : a modern approach*. Pearson, 2016.

- [13] E. Alpaydin, *Introduction to machine learning*. MIT press, 2020.
- [14] X.-W. Chen and X. Lin, “Big data deep learning : challenges and perspectives,” *IEEE access*, vol. 2, pp. 514–525, 2014.
- [15] Q. Zhang, H. Yu, M. Barbiero, B. Wang, and M. Gu, “Artificial neural networks enabled by nanophotonics,” *Light : Science & Applications*, vol. 8, no. 1, p. 42, 2019.
- [16] Y. Bengio, I. Goodfellow, and A. Courville, *Deep learning*. MIT press Cambridge, MA, USA, 2017, vol. 1.
- [17] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, “Focal loss for dense object detection,” in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2980–2988.
- [18] P. Le Callet, C. Viard-Gaudin, and D. Barba, “A convolutional neural network approach for objective video quality assessment,” *IEEE transactions on neural networks*, vol. 17, no. 5, pp. 1316–1327, 2006.
- [19] M. Thoma, “Analysis and optimization of convolutional neural network architectures,” *arXiv preprint arXiv :1707.09725*, 2017.
- [20] D. C. Ciresan, U. Meier, J. Masci, L. M. Gambardella, and J. Schmidhuber, “Flexible, high performance convolutional neural networks for image classification,” in *Twenty-second international joint conference on artificial intelligence*. Citeseer, 2011.
- [21] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.
- [22] D. Scherer, A. Müller, and S. Behnke, “Evaluation of pooling operations in convolutional architectures for object recognition,” in *International conference on artificial neural networks*. Springer, 2010, pp. 92–101.
- [23] S. J. Pan and Q. Yang, “A survey on transfer learning,” *IEEE Transactions on knowledge and data engineering*, vol. 22, no. 10, pp. 1345–1359, 2009.
- [24] L. Torrey and J. Shavlik, “Transfer learning,” in *Handbook of research on machine learning applications and trends : algorithms, methods, and techniques*. IGI global, 2010, pp. 242–264.
- [25] A. Sarkar, P. R. Behera, and J. Shukla, “Multi-source transfer learning for facial emotion recognition using multivariate correlation analysis,” *Scientific Reports*, vol. 13, no. 1, p. 21004, 2023.
- [26] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, “How transferable are features in deep neural networks?” *Advances in neural information processing systems*, vol. 27, 2014.
- [27] P. M. Cheng and H. S. Malhi, “Transfer learning with convolutional neural networks for classification of abdominal ultrasound images,” *Journal of digital imaging*, vol. 30, pp. 234–243, 2017.
- [28] E. C. Orenstein and O. Beijbom, “Transfer learning and deep feature extraction for planktonic image data sets,” in *2017 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, 2017, pp. 1082–1088.
- [29] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

- [30] N. Gupta, “A pre-trained vs fine-tuning methodology in transfer learning,” in *Journal of Physics : Conference Series*, vol. 1947, no. 1. IOP Publishing, 2021, p. 012028.
- [31] L. Aversano, M. L. Bernardi, M. Cimitile, M. Iammarino, and S. Rondinella, “Tomato diseases classification based on vgg and transfer learning,” in *2020 IEEE International Workshop on Metrology for Agriculture and Forestry (MetroAgri-For)*. IEEE, 2020, pp. 129–133.
- [32] Z. Wu, C. Shen, and A. Van Den Hengel, “Wider or deeper : Revisiting the resnet model for visual recognition,” *Pattern recognition*, vol. 90, pp. 119–133, 2019.
- [33] Y. Ibrahim, H. Wang, and K. Adam, “Analyzing the reliability of convolutional neural networks on gpus : Googlenet as a case study,” in *2020 International Conference on Computing and Information Technology (ICCI-1441)*. IEEE, 2020, pp. 1–6.
- [34] C. A. Ferreira, T. Melo, P. Sousa, M. I. Meyer, E. Shakibapour, P. Costa, and A. Campilho, “Classification of breast cancer histology images through transfer learning using a pre-trained inception resnet v2,” in *International conference image analysis and recognition*. Springer, 2018, pp. 763–770.
- [35] M. Tan, “Efficientnet : Rethinking model scaling for convolutional neural networks,” *arXiv preprint arXiv :1905.11946*, 2019.
- [36] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, “Mobilenets : efficient convolutional neural networks for mobile vision applications (2017),” *arXiv preprint arXiv :1704.04861*, vol. 126, 2017.
- [37] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, “Mobilenetv2 : Inverted residuals and linear bottlenecks,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 4510–4520.
- [38] F. Chollet, “Xception : Deep learning with depthwise separable convolutions,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1251–1258.
- [39] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, “Imagenet : A large-scale hierarchical image database,” in *2009 IEEE conference on computer vision and pattern recognition*. Ieee, 2009, pp. 248–255.
- [40] T. Core, “v2. 2.0,” *Just your regular densely-connected NN layer*, 2020.
- [41] K. Maharana, S. Mondal, and B. Nemade, “A review : Data pre-processing and data augmentation techniques,” *Global Transitions Proceedings*, vol. 3, no. 1, pp. 91–99, 2022.
- [42] Tensorflow, “Data augmentation : Tensorflow core,” https://www.tensorflow.org/tutorials/images/data_augmentation, 2022, accessed : 21-march-2023.
- [43] C. Robert, “Machine learning, a probabilistic perspective,” 2014.
- [44] M. A. Boudia, R. M. Hamou, A. Amine, and A. C. Lokbani, “An adaptation of a f-measure for automatic text summarization by extraction,” *Cluster Computing*, vol. 23, pp. 2389–2398, 2020.
- [45] Kaggle, “car-flowers-workshop,” <https://www.kaggle.com/datasets/mohamedmhamed/car-flowers-workshop>, 2024, accessed : 25-may-2024.

- [46] WHO, “World health organization (who),” <https://www.who.int/>, 2024, accessed : 23-May-2024.
- [47] ScienceDaily, “Cancer projected to become leading cause of death worldwide in 2010,” <https://www.sciencedaily.com/releases/2008/12/081209111516.htm>, 2008, accessed : 24-may-2024.
- [48] M. Emre Celebi, Y. Alp Aslandogan, W. V. Stoecker, H. Iyatomi, H. Oka, and X. Chen, “Unsupervised border detection in dermoscopy images,” *Skin research and technology*, vol. 13, no. 4, pp. 454–462, 2007.
- [49] R. O. Kenet and T. B. Fitzpatrick, “Reducing mortality and morbidity of cutaneous melanoma : A six year plan : B) identifying high and low risk pigmented lesions using epiluminescence microscopy,” *The Journal of Dermatology*, vol. 21, no. 11, pp. 881–884, 1994.
- [50] V. P. Chinem and H. A. Miot, “Epidemiology of basal cell carcinoma,” *Anais brasileiros de dermatologia*, vol. 86, pp. 292–305, 2011.
- [51] D. S. Rigel, R. Friedman, J. K. Robinson, M. I. Ross, C. J. Cockerell, H. Lim, E. Stockfleth, and J. M. Kirkwood, *Cancer of the Skin : Expert Consult*. Elsevier Health Sciences, 2011.
- [52] O. Website, “The skin cancer foundation,” <https://www.kaggle.com/datasets/mohamedmhamedi/car-flowers-workshop>, 2020, accessed : 08-Jul-2024.
- [53] M. Alam, “Uncertainty and variance in the management of high-risk cutaneous squamous cell carcinoma : comment on “uncertainty in the perioperative management of high-risk cutaneous squamous cell carcinoma among mohs surgeons”,” *Archives of dermatology*, vol. 146, no. 11, pp. 1231–1232, 2010.
- [54] S. Lautenschlager, H. C. Wulf, and M. R. Pittelkow, “Photoprotection,” *The Lancet*, vol. 370, no. 9586, pp. 528–537, 2007.
- [55] R. E. Shore, M. Moseson, X. Xue, Y. Tse, N. Harley, and B. S. Pasternack, “Skin cancer after x-ray treatment for scalp ringworm,” *Radiation research*, vol. 157, no. 4, pp. 410–418, 2002.
- [56] M. R. Karagas, T. A. Stukel, E. R. Greenberg, J. A. Baron, L. A. Mott, and R. S. Stern, “Risk of subsequent basal cell carcinoma and squamous cell carcinoma of the skin among patients with prior skin cancer,” *Jama*, vol. 267, no. 24, pp. 3305–3310, 1992.
- [57] W. B. Ershler and D. L. Longo, “Aging and cancer : issues of basic and clinical science,” *Journal of the National Cancer Institute*, vol. 89, no. 20, pp. 1489–1497, 1997.
- [58] D. L. Sackett, R. B. Haynes, and P. Tugwell, *Clinical epidemiology : a basic science for clinical medicine.*, 1985.
- [59] M. A. Weinstock, H. A. Bogaars, M. Ashley, V. Litle, E. Bilodeau, and S. Kimmel, “Nonmelanoma skin cancer mortality : a population-based study,” *Archives of dermatology*, vol. 127, no. 8, pp. 1194–1197, 1991.
- [60] Home, “The melanoma centre,” <https://melanomacentre.com.au/>, 2022, accessed : 08-Aug-2024.
- [61] A. C. Society, “Skin cancer : Skin cancer types,” <https://www.cancer.org/cancer/types/skin-cancer.html>, 2023, accessed : 28-Jun-2024.

- [62] J. N. Cormier, Y. Xing, M. Ding, J. E. Lee, P. F. Mansfield, J. E. Gershenwald, M. I. Ross, and X. L. Du, “Ethnic differences among patients with cutaneous melanoma,” *Archives of internal medicine*, vol. 166, no. 17, pp. 1907–1914, 2006.
- [63] BAD, “British association of dermatologists,” <http://www.bad.org.uk/>.
- [64] M. A. Tucker and A. M. Goldstein, “Melanoma etiology : where are we?” *Oncogene*, vol. 22, no. 20, pp. 3042–3052, 2003.
- [65] A. A. Marghoob, R. P. Braun, and A. W. Kopf, *Atlas of dermoscopy*. Informa healthcare London, 2012.
- [66] D. S. Rigel and J. A. Carucci, “Malignant melanoma : prevention, early detection, and treatment in the 21st century,” *CA : a cancer journal for clinicians*, vol. 50, no. 4, pp. 215–236, 2000.
- [67] F. Nachbar, W. Stolz, T. Merkle, A. B. Cagnetta, T. Vogt, M. Landthaler, P. Bilek, O. Braun-Falco, and G. Plewig, “The abcd rule of dermatoscopy : high prospective value in the diagnosis of doubtful melanocytic skin lesions,” *Journal of the American Academy of Dermatology*, vol. 30, no. 4, pp. 551–559, 1994.
- [68] H. P. Soyer, G. Argenziano, I. Zalaudek, R. Corona, F. Sera, R. Talamini, F. Barbato, A. Baroni, L. Cicale, A. Di Stefani *et al.*, “Three-point checklist of dermoscopy a new screening method for early detection of melanoma,” *Dermatology*, vol. 208, no. 1, pp. 27–31, 2004.
- [69] Dermoscopedia.org, “Menzies method,” https://dermoscopedia.org/Menzies_Method.
- [70] H. Kittler, “Diagnostic accuracy of dermoscopy/dermatoscopy,” *An Atlas of Dermoscopy*, p. 12, 2004.
- [71] DermNet®, “Dermoscopy. malignant melanoma,” <https://dermnetnz.org/cme/dermoscopy-course/dermoscopy-of-melanoma>.
- [72] M. Berseth, “Isic 2017-skin lesion analysis towards melanoma detection,” *arXiv preprint arXiv :1703.00523*, 2017.
- [73] P. Tschandl, C. Rosendahl, and H. Kittler, “The ham10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions,” *Scientific data*, vol. 5, no. 1, pp. 1–9, 2018.
- [74] I. C. 2019.
- [75] I. S. I. Collaboration *et al.*, “Siim-isic 2020 challenge dataset,” *International Skin Imaging Collaboration*, 2020.
- [76] J. Hurtado and F. Reales, “A machine learning approach for the recognition of melanoma skin cancer on macroscopic images,” *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 19, no. 4, pp. 1357–1368, 2021.
- [77] H. Li, Y. Pan, J. Zhao, and L. Zhang, “Skin disease diagnosis with deep learning : A review,” *Neurocomputing*, vol. 464, pp. 364–393, 2021.
- [78] P. N. Srinivasu, J. G. SivaSai, M. F. Ijaz, A. K. Bhoi, W. Kim, and J. J. Kang, “Classification of skin disease using deep learning neural networks with mobilenet v2 and lstm,” *Sensors*, vol. 21, no. 8, p. 2852, 2021.
- [79] E. Goceri, “Diagnosis of skin diseases in the era of deep learning and mobile technology,” *Computers in Biology and Medicine*, vol. 134, p. 104458, 2021.

- [80] A. A. Adegun and S. Viriri, “Deep learning-based system for automatic melanoma detection,” *IEEE Access*, vol. 8, pp. 7160–7172, 2019.
- [81] M. A. Hussien and A. H. H. Alasadi, “Classification of melanoma skin cancer using deep learning approach,” *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 22, no. 1, pp. 129–137, 2024.
- [82] N. C. Codella, Q.-B. Nguyen, S. Pankanti, D. A. Gutman, B. Helba, A. C. Halpern, and J. R. Smith, “Deep learning ensembles for melanoma recognition in dermoscopy images,” *IBM Journal of Research and Development*, vol. 61, no. 4/5, pp. 5–1, 2017.
- [83] Y. Li and L. Shen, “Skin lesion analysis towards melanoma detection using deep learning network,” *Sensors*, vol. 18, no. 2, p. 556, 2018.
- [84] S. H. Kassani and P. H. Kassani, “A comparative study of deep learning architectures on melanoma detection,” *Tissue and Cell*, vol. 58, pp. 76–83, 2019.
- [85] A. Naeem, M. S. Farooq, A. Khelifi, and A. Abid, “Malignant melanoma classification using deep learning : datasets, performance measurements, challenges and opportunities,” *IEEE access*, vol. 8, pp. 110 575–110 597, 2020.
- [86] X. Zhang, “Melanoma segmentation based on deep learning,” *Computer assisted surgery*, vol. 22, no. sup1, pp. 267–277, 2017.
- [87] H. C. Reis, V. Turk, K. Khoshelham, and S. Kaya, “Insinet : a deep convolutional approach to skin cancer detection and segmentation,” *Medical & Biological Engineering & Computing*, pp. 1–20, 2022.
- [88] Y. Wan, Y. Cheng, and M. Shao, “Mslanet : multi-scale long attention network for skin lesion classification,” *Applied Intelligence*, vol. 53, no. 10, pp. 12 580–12 598, 2023.
- [89] M. M. Mijwil, “Skin cancer disease images classification using deep learning solutions,” *Multimedia Tools and Applications*, vol. 80, no. 17, pp. 26 255–26 271, 2021.
- [90] L.-F. Li, X. Wang, W.-J. Hu, N. N. Xiong, Y.-X. Du, and B.-S. Li, “Deep learning in skin disease image recognition : A review,” *Ieee Access*, vol. 8, pp. 208 264–208 280, 2020.
- [91] D. Strigl, K. Kofler, and S. Podlipnig, “Performance and scalability of gpu-based convolutional neural networks,” in *2010 18th Euromicro conference on parallel, distributed and network-based processing*. IEEE, 2010, pp. 317–324.
- [92] T.-Y. Ross and G. Dollár, “Focal loss for dense object detection,” in *proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2980–2988.

Résumé

Les affections cutanées, dont les cancers de la peau, représentent un défi important pour la santé publique. L'apparence souvent trompeuse des lésions souligne l'importance d'un diagnostic précis par un dermatologue. Parmi les cancers de la peau, le mélanome malin se distingue par sa gravité et sa capacité à se métastaser rapidement. Bien qu'il soit moins fréquent que d'autres types de cancer cutané, comme le carcinome basocellulaire et le carcinome spinocellulaire, il nécessite une attention particulière en raison de son pronostic. La dermoscopie est un outil essentiel pour les dermatologues car elle permet de détecter les mélanomes à un stade précoce, ce qui est crucial pour un traitement efficace. La combinaison de la dermoscopie et de l'intelligence artificielle offre de nouvelles perspectives pour un diagnostic plus précis et plus rapide du mélanome. Les algorithmes d'apprentissage profond, entraînés sur des datasets d'images de haute qualité, peuvent aider les dermatologues à détecter les mélanomes à un stade précoce, améliorant ainsi les chances de guérison des patients. Cette thèse présente le développement d'un système d'aide au diagnostic du cancer de la peau basé sur l'apprentissage profond, visant à classifier précisément le Mélanome malin. L'étude explore diverses architectures de réseaux neuronaux convolutifs (CNN), en intégrant l'apprentissage par transfert pour tirer parti des modèles pré-entraînés, l'augmentation de données pour enrichir les ensembles d'entraînement, et des architectures hybrides CNN-LSTM pour améliorer les performances du modèle. L'évaluation expérimentale a démontré que l'architecture MobileNetV2-LSTM offre les meilleures performances en termes d'exactitude, sensibilité et spécificité, surpassant les approches existantes dans la littérature. L'objectif est d'améliorer la détection précoce du mélanome grâce à des méthodes computationnelles avancées, ce qui pourrait potentiellement augmenter les taux de survie des patients.

Mots clés : maladies de peau, dermatologie, cancer de la peau, mélanome malin, CNN, apprentissage par transfert augmentation des donnée, architecture hybride, LSTM.

Abstract

Skin diseases, including skin cancers, represent a major public health challenge. The often misleading appearance of lesions underlines the importance of accurate diagnosis by a dermatologist. Among skin cancers, malignant melanoma stands out for its severity and ability to metastasize rapidly. Although it is less common than other types of skin cancer, such as basal cell carcinoma and squamous cell carcinoma, it requires special attention because of its prognosis. Dermoscopy is an essential tool for dermatologists as it enables melanomas to be detected at an early stage, which is crucial for effective treatment. The combination of dermoscopy and artificial intelligence offers new prospects for more accurate and faster diagnosis of melanoma. Deep learning algorithms trained on high-quality image datasets can help dermatologists detect melanoma at an early stage, improving patient's chances of recovery. This thesis presents the development of a skin cancer diagnostic aid system based on deep learning, aimed at accurately classifying malignant melanoma. The study explores various convolutional neural network (CNN) architectures, integrating transfer learning to take advantage of pre-trained models, data augmentation to enrich the training sets, and hybrid CNN-LSTM architectures to improve classifier accuracy. Experimental results showed that the MobileNetV2-LSTM architecture outperformed state-of-the-art approaches, achieving the highest accuracy, sensitivity, and specificity. The aim is to improve early detection of melanoma using advanced computational methods, which could potentially increase patient survival rates.

Keywords: Skin diseases, Dermatology, Skin cancer, Malignant Melanoma, CNN, Transfer learning, Pre-trained models, Data augmentation, Hybrid architecture, LSTM.

المخلص

الأمراض الجلدية، بما في ذلك سرطان الجلد، تمثل تحديًا كبيرًا للصحة العامة. غالبًا ما يؤدي المظهر الخادع للأورام إلى التأكيد على أهمية التشخيص الدقيق من قبل طبيب الأمراض الجلدية. من بين أنواع سرطان الجلد، يبرز الميلانوما الخبيث بشدة وقدرته على الانتشار بسرعة. على الرغم من أنه أقل شيوعًا من أنواع أخرى من سرطان الجلد، مثل سرطان الخلايا القاعدية وسرطان الخلايا الحرشفية، إلا أنه يتطلب اهتمامًا خاصًا بسبب صعوبة تشخيصه. تعتبر الفحص بالمجهر البصري (dermoscopy) أداة أساسية لأطباء الأمراض الجلدية، حيث تمكنهم من اكتشاف الميلانوما في مرحلة مبكرة، وهو أمر حاسم للعلاج الفعال. يمثل الجمع بين الفحص بالمجهر البصري والذكاء الاصطناعي أفقًا جديدًا لتشخيص الميلانوما بشكل أسرع وأكثر دقة. يمكن أن تساعد خوارزميات التعلم العميق المدربة على مجموعات بيانات عالية الجودة من الصور أطباء الأمراض الجلدية في اكتشاف الميلانوما في مرحلة مبكرة، مما يحسن فرص شفاء المرضى. تقدم هذه الرسالة تطوير نظام مساعد لتشخيص سرطان الجلد يستند إلى التعلم العميق، بهدف تصنيف الميلانوما الخبيث بدقة. تستكشف هذه الدراسة مختلف هياكل الشبكات العصبونية التلافيفية (CNN)، مع دمج التعلم الانتقالي للاستفادة من النماذج المدربة مسبقًا، وزيادة البيانات لتخصيص مجموعات التدريب، واستخراج البيانات لتحسين دقة النظام. أظهرت التقييمات التجريبية أن بنية MobileNetV2-LSTM حققت أفضل النتائج من حيث الدقة والحساسية والخصوصية، متفوقةً بذلك على المناهج الحالية. يهدف هذا البحث إلى تحسين الكشف المبكر عن الميلانوما باستخدام طرق الحوسبة المتقدمة، مما يمكن أن يزيد من معدلات شفاء المرضى.

الكلمات المفتاحية: الأمراض الجلدية، الأمراض الجلدية (الجلدية)، سرطان الجلد، الورم الميلانيني الخبيث، الشبكة العصبونية التلافيفية (CNN)، التعلم بالنقل، النماذج المدربة مسبقًا، تعزيز البيانات، العمارة الهجينة، الشبكة العصبونية التكرارية طويلة المدى (LSTM).