

Projet de fin d'études C-Pred

Application de méthodes de classification et de prédiction de lésions tissulaires

Antoine Boukhelifa, Giovanni Vagliani
Encadrant : Marcel Hourani

I.	Introduction	4
1)	Contexte	4
2)	Objectif	4
3)	Les lésions de la peau	5
II.	Les bases de données	7
1)	Présentation des différentes bases	7
2)	Notre choix : HAM10000	8
3)	Exploratory Data Analysis	8
a.	Répartition des anomalies dermatologiques	8
b.	Répartition des sexes	9
c.	Répartition des localisations	10
d.	Répartition des anomalies en fonction de l'âge	10
III.	Etat de l'art	12
1)	Les métriques	12
2)	Les études	12
IV.	Deep Learning and Computer Vision	14
1)	Qu'est-ce que c'est ?	14
2)	Apprentissage	14
3)	Modèles de classification	15
4)	CNN	16
V.	Notre choix de modèle : YoLov8	18
1)	Présentation	18
2)	Méthodologie	18
a.	Equilibrage de la base de données :	18
b.	Heatmap sur les prédictions avec Eigencam	20
c.	Prétraitement des images	21
d.	Sélection de la région d'intérêt (ROI)	22
3)	Résultats	26
a.	Méthode 1	26
b.	Méthode 2	27
c.	Méthode 3	28
d.	Méthode 4	29
e.	Méthode 5	29
f.	Méthode 6	30
g.	Méthode 7	31

Antoine Boukhelifa
Giovanni Vagliani

VI. Conclusion.....	32
VII. Perspectives	34
VIII. Bibliographie	35
IX. Annexes.....	36

I. Introduction

1) Contexte

C-Pred concentre ses efforts sur la détection précoce des lésions cutanées, englobant une vaste gamme de conditions dermatologiques telles que les grains de beauté, les kératoses, les brûlures et bien d'autres. Ces lésions, bien que variées, représentent un enjeu crucial de santé publique, car une identification précoce revêt une importance capitale. En effet, ces altérations cutanées peuvent évoluer vers des complications graves si elles ne sont pas diagnostiquées et traitées à temps.

Au-delà de leur impact esthétique, les maladies dermatologiques, parmi lesquelles le mélanome occupe une place prépondérante, peuvent entraîner des complications sévères du fait de leur nature agressive et de leur capacité à se propager rapidement en l'absence de détection précoce. Ainsi, le projet C-Pred émerge comme une initiative stratégique visant à anticiper et à limiter les conséquences néfastes associées à ces lésions cutanées.

Dans cette démarche, la dermatoscopie, outil d'imagerie dermatologique, s'avère primordiale. L'utilisation de la dermatoscopie permet une analyse détaillée des lésions cutanées, offrant une vision amplifiée et des caractéristiques subtiles non visibles à l'œil nu. Cette approche renforce la précision diagnostique, offrant une efficacité accrue par rapport à l'évaluation visuelle conventionnelle.

En déployant des protocoles de classification avancés et en exploitant la dermatoscopie conjointement à des méthodes technologiques innovantes, le projet C-Pred aspire à accroître significativement l'efficacité du diagnostic.

Cette nouvelle approche, intégrant la technologie à la dermatoscopie pour une meilleure identification des lésions cutanées, promet ainsi d'optimiser les résultats thérapeutiques et d'offrir des solutions adaptées, transformant ainsi la prise en charge des maladies dermatologiques pour un plus grand nombre de personnes.

2) Objectif

L'objectif de C-Pred est de concevoir un algorithme spécialisé dans la détection et la classification des anomalies dermatologiques. Pour ce faire, notre approche se décline en deux étapes clés.

Dans un premier temps, nous approfondirons notre compréhension des lésions cutanées en exploitant les avancées de la littérature scientifique ainsi que les technologies appliquées aux maladies de la peau. En parallèle, nous sélectionnerons une base de données appropriée pour notre analyse.

Dans un second temps, notre objectif sera de mettre en œuvre des algorithmes de classification, exploitant les progrès de l'intelligence artificielle, notamment en computer vision.

L'évolution de l'intelligence artificielle, notamment dans le champ de la computer vision, a révolutionné un éventail de domaines. La capacité des systèmes d'IA à interpréter et à extraire des données à partir d'images a généré des avancées notables.

Dans le domaine médical, l'IA a ouvert de vastes horizons. La computer vision a permis la détection précoce de diverses anomalies, depuis l'identification des fractures osseuses jusqu'à la reconnaissance faciale, en passant par l'analyse d'images radiologiques pour repérer des pathologies complexes. Elle a également été employée pour surveiller les signes vitaux à partir de vidéos, détecter des anomalies cardiaques et même anticiper des complications médicales.

Notre but est donc d'appliquer ces avancées de l'intelligence artificielle aux images dermatologiques, afin de développer des outils performants de classification et de détection des lésions. Ce faisant, nous visons à instaurer des diagnostics précoces, véritable enjeu pour améliorer significativement la prise en charge des maladies dermatologiques, offrant ainsi des perspectives prometteuses pour l'amélioration des soins de santé cutanée.

3) Les lésions de la peau


Les lésions cutanées sont des altérations visibles à la surface de la peau. Elles peuvent résulter de diverses causes, notamment des infections, des traumatismes, des réactions allergiques, ou des processus pathologiques sous-jacents. Comprendre ces lésions est fondamental pour effectuer une meilleure classification et prédiction.

Parmi ces lésions, on retrouve une variété de conditions dermatologiques courantes, chacune avec ses caractéristiques distinctes. Le naevus mélanocytaire, par exemple, est une croissance bénigne de cellules pigmentaires de la peau, tandis que le mélanome qui représente sa forme maligne est potentiellement grave. Les kératoses séborrhéiques sont des lésions cutanées non cancéreuses, souvent pigmentées, tandis que le carcinome basocellulaire est une forme courante de cancer de la peau, généralement moins agressive que le mélanome.

Les kératoses actiniques, la maladie de Bowen et les lésions vasculaires de la peau présentent également des particularités distinctes. Les kératoses actiniques sont des lésions précancéreuses causées par une exposition excessive au soleil, tandis que la maladie de Bowen est une forme de carcinome in situ touchant les couches supérieures de la peau. Les lésions vasculaires impliquent les vaisseaux sanguins et peuvent prendre différentes formes telles que les angiomes ou les malformations vasculaires.

Enfin, le dermatofibrome est une lésion cutanée bénigne qui se forme généralement à la suite d'un traumatisme mineur. La reconnaissance de ces lésions et la différenciation entre les lésions bénignes et malignes revêtent une importance capitale pour le diagnostic et le traitement adéquat.

Dans le cadre du projet, un tableau synthétique (*Table 1*) a été élaboré pour récapituler les caractéristiques majeures de sept types de lésions dermatologiques. Ce tableau offre une vue d'ensemble succincte, comprenant des informations telles que la définition, les raisons de l'apparition, la dangerosité, les populations concernées et les moyens de limiter l'émergence de ces altérations cutanées.

Nom et acronyme des lésions	Définition	Raison de l'apparition	Moyens pour limiter l'apparition	Personnes concernées	Dangerosité
Naevus mélanocytaire, Nv 	Croissance bénigne de cellules pigmentaires de la peau	Pas entièrement comprise Facteurs génétiques Exposition au soleil	Protection solaire Surveillance régulière	Toutes les tranches d'âge Individus à la peau claire	Faible

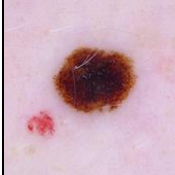
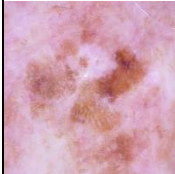
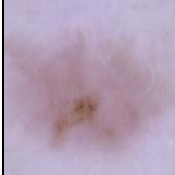
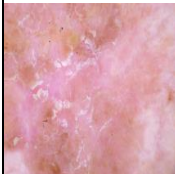
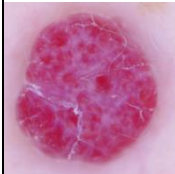
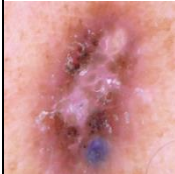
<p>Mélanome, Mel</p> 	<p>Forme maligne du naevus mélanocytaire</p>	<p>Facteurs génétiques Exposition au soleil Brûlures sévères</p>	<p>Eviter la surexposition aux UV Surveillance régulière</p>	<p>Principalement les adultes Individus à la peau claire</p>	<p>Elevée</p>
<p>Kératose séborrhéique, Bkl</p> 	<p>Lésion cutanée non cancéreuse, généralement pigmentée</p>	<p>Facteurs génétiques Vieillesse</p>	<p>Surveillance régulière</p>	<p>Personnes âgées Individus à la peau claire</p>	<p>Faible</p>
<p>Carcinome basocellulaire, Bcc</p> 	<p>Cancer de la peau, moins agressif que le mélanome</p>	<p>Facteurs génétiques Exposition cumulative au soleil très importante</p>	<p>Protection solaire Surveillance régulière</p>	<p>Principalement les adultes Plus fréquent chez les hommes</p>	<p>Modéré</p>
<p>Kératose actinique et maladie de Bowen, Akiec</p> 	<p>Lésion pré-cancéreuse due à une exposition chronique au soleil</p>	<p>Facteurs génétiques Exposition prolongée au soleil</p>	<p>Protection solaire</p>	<p>Age moyen à avancé Plus fréquent chez les hommes Individus à la peau claire</p>	<p>Modéré à élevé</p>
<p>Lésions vasculaires, Vaso</p> 	<p>Affections des vaisseaux sanguins de la peau</p>	<p>Facteurs génétiques Traumatismes cutanés Mauvais développement des vaisseaux sanguins</p>	<p>Protection solaire, même si peu efficace</p>	<p>Toutes les tranches d'âge Facteurs géographiques</p>	<p>Variable</p>
<p>Dermatofibrome, Df</p> 	<p>Lésion cutanée bénigne formée à la suite d'un traumatisme mineur</p>	<p>Inconnue Peut être liée aux réponses immunitaires</p>	<p>Aucune pour le moment</p>	<p>Principalement les adultes Plus fréquent chez les femmes</p>	<p>Faible</p>

Table 1 : Récapitulatif des lésions et de leurs caractéristiques

II. Les bases de données

1) Présentation des différentes bases

Pour identifier correctement les 7 types de lésions cutanées que nous voulons, il faut trouver une base de données adéquate pour garantir la robustesse et l'interprétation du modèle de Deep Learning. Cela implique un nombre assez conséquent pour que l'algorithme puisse s'entraîner dessus, des images de chaque lésion ainsi qu'une bonne représentation des anomalies. Il existe déjà plusieurs bases de données en libre accès, chacune avec ses avantages et inconvénients et nous allons les comparer. Il s'agit de PH2 (*Teresa, Pedro, André, & Catarina, 2015*), Atlas, Rosendahl, ISIC2017 (Bill, Connah, Andrzej, Joanna, & Moi Hoon, 2022) et HAM10000.

Le dataset PH2 est spécifiquement orienté vers la détection du mélanome, cependant, il ne comporte malheureusement pas les images requises pour l'identification exhaustive de toutes les lésions cutanées ciblées dans notre étude. Elle n'est donc pas retenue.

La base de données Atlas et Rosendahl ont le même objectif que PH2 mais sont plus complètes et proposent tous les types d'image que nous recherchons. Cependant, le nombre d'images disponibles pour chacune d'entre elles (1024 et 2259 respectivement) n'est pas suffisant pour assurer un entraînement adéquat du modèle. Par conséquent, nous les excluons de notre sélection.

Il reste dorénavant ISIC2017 et HAM10000 qui sont deux datasets possédant un bon nombre d'images (respectivement 13786 et 10015) et selon ce critère nous devrions choisir ISIC2017. Cependant, il faut prendre en compte deux nouveaux critères pour le choix. La diversité des images garantit une bonne répartition et permet d'éviter le sur ou le sous-apprentissage. La vérification pathologique signifie que les images des lésions cutanées ont été confirmées par des examens pathologiques. Cela implique qu'un professionnel de la santé spécialisé dans l'interprétation microscopique des échantillons de tissus cutanés, a examiné les lésions et a fourni une évaluation diagnostique précise. La vérification garantit donc que les diagnostics sont basés sur des analyses approfondies des caractéristiques microscopiques des tissus.

En nous aidant de ces deux nouveaux critères, HAM10000 serait plus adapté pour notre étude : il a une plus grande représentation de chaque lésion et possède un bien plus grand pourcentage d'images qui ont été vérifiées (53.3% contre 26.3% pour ISIC2017).

Voici un tableau récapitulatif des caractéristiques de chaque base de données (*Table 2*).

Dataset	PH2	Atlas	Rosendahl	ISIC2017	HAM10000
Nombre d'images	200	1024	2259	13786	10015
Vérification pathologique	20.5%	NA	100%	26.3%	53.3%
Nv	160	582	803	11861	6705
Mel	40	275	1019	342	1113
Bkl	-	70	490	575	1099
Bcc	-	42	296	33	514
Akiec	-	5	295	2	327

Vasc	-	30	3	15	142
Df	-	20	30	7	115

Table 2 : Récapitulatif des bases de données et de leurs caractéristiques

HAM10000, avec sa variété d'images dermatoscopiques, apparaît comme le meilleur choix pour créer un modèle de deep learning capable de distinguer efficacement les types de lésions.

2) Notre choix : HAM10000

HAM10000 est l'abréviation de Human Against Machine et est constitué de 10015 images dermatoscopiques appelées archives ISIC.

La base de données HAM10000, disponible sur Kaggle ([*Philipp, Cliff, & Harald, 2018*](#)), se compose de plusieurs éléments, notamment deux dossiers contenant les images et un fichier CSV nommé "HAM10000_metadata". Ce fichier répertorie sept attributs clés pour caractériser les lésions cutanées. Chaque lésion cutanée est identifiée de manière unique par un "lesion_id", ce qui permet de regrouper plusieurs images associées à la même lésion. Parallèlement, chaque image possède son propre "image_id", facilitant ainsi leur distinction. Les attributs "dx" et "dx_type" revêtent une importance particulière pour le diagnostic des lésions. "dx" donne le diagnostic clinique, comme mélanome (mel), nævus (nv), carcinome basocellulaire (bcc), et autres. En parallèle, "dx_type" spécifie la méthode diagnostique utilisée, qu'il s'agisse de l'histopathologie (histo), d'un suivi clinique (follow_up), d'un consensus d'experts (consensus), ou encore de l'utilisation de la microscopie confocale (confocal). Cette base de données renferme également des attributs démographiques essentiels pour chaque patient, tels que "age" et "sex". "Age" représente l'âge du patient au moment de la capture de l'image, tandis que "sex" indique le genre du patient, noté comme "male" ou "female". Enfin, l'attribut "localization" précise la zone anatomique où la lésion a été repérée, incluant des détails comme le dos, les bras ou le visage.

3) Exploratory Data Analysis

Avant de procéder à la construction de l'algorithme de classification, une analyse exploratoire des données a été effectuée pour mieux comprendre la nature et la répartition des données dans la base HAM10000.

a. Répartition des anomalies dermatologiques

La répartition des différents types d'anomalies dermatologiques dans la base de données est essentielle pour évaluer le déséquilibre entre les classes. Comme illustré dans la figure ci-dessous ([*Figure 1*](#)), les résultats montrent une forte disparité dans la proportion des classes, avec notamment une prédominance significative des nævus (nv) par rapport aux autres types d'anomalies en représentant environ 67% des images de la base de données. Cela s'explique notamment par leur diffusion très large chez les personnes.

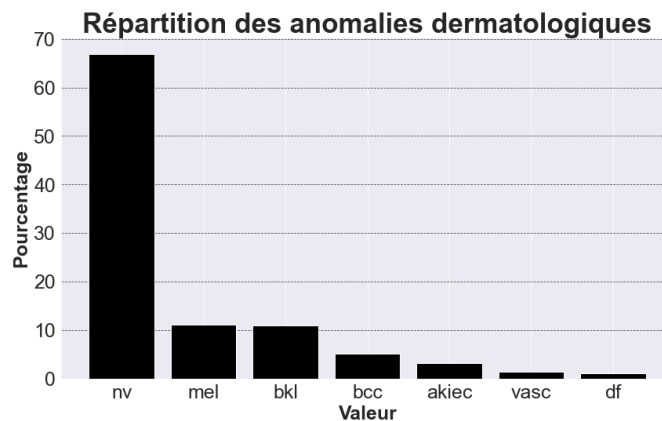


Fig. 1. Répartition des anomalies dermatologiques

Voici un tableau récapitulatif de la répartition exacte

Classe	Nombre d'images	Pourcentage
akiec	327	3.27%
bcc	514	5.13%
bkl	1099	10.97%
df	115	1.15%
mel	1113	11.11%
nv	6705	66.95%
vasc	142	1.42%

Table 3 : Récapitulatif de la répartition de HAM10000

b. Répartition des sexes

Une autre caractéristique importante à considérer est la répartition du sexe dans la base de données. Le diagramme en secteurs ci-dessous (*Figure 2*) met en évidence la distribution des patients par sexe. On observe une légère prépondérance des données provenant de patients de sexe masculin, même si cela n'a pas d'influence pour notre étude. En revanche on peut noter la présence d'images n'appartenant à aucun sexe. Après une recherche plus approfondie, ce cas ne nous posera également pas de problèmes.



Fig. 2. Répartition du genre

c. Répartition des localisations

L'exploration de la base de données HAM10000 révèle une diversité significative dans la localisation des anomalies dermatologiques sur le corps humain. Les images de lésions cutanées proviennent de différentes régions anatomiques, chacune présentant une importance variable, comme illustré par le nuage de mots pour les localisations (*Figure 3*).

Parmi ces localisations, le dos se distingue comme une zone fréquemment observée, avec une taille de mot relativement grande dans le nuage de mots. Cela suggère que le dos est l'une des régions les plus courantes pour les lésions cutanées dans la base de données. Les membres inférieurs, le tronc, les membres supérieurs et l'abdomen sont également bien représentés, avec des tailles de mots significatives, soulignant l'implication de ces régions dans les différentes anomalies dermatologiques. D'autres parties du corps, bien que moins fréquentes, sont également répertoriées dans le nuage de mots, ce qui témoigne de la variabilité spatiale des conditions cutanées dans la base de données HAM10000. Comprendre cette répartition est essentiel pour une interprétation précise des données et peut influencer les stratégies de classification et de diagnostic des lésions cutanées.



Fig. 3. Visualisation des localisations des anomalies

d. Répartition des anomalies en fonction de l'âge

Enfin, l'analyse de la distribution des anomalies en fonction de l'âge, présentée sous forme de boîtes à moustaches (*Figure 4*), permet d'identifier d'éventuelles variations d'âge en fonction du type d'anomalie dermatologique. Cette visualisation peut aider à détecter des tendances ou des particularités dans les données : par exemple, on constate que les nv sont présents chez toute la population, peu importe son âge, tandis que les anomalies telles que akiec ou bcc se manifestent surtout chez les plus de 40 ans.

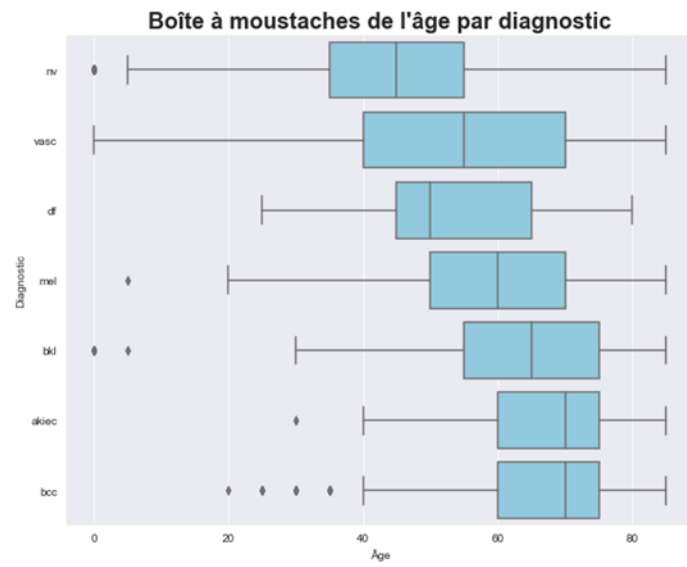


Fig. 4. Représentation des anomalies en fonction de l'âge

III. Etat de l'art

1) Les métriques

Dans l'évaluation des performances d'un modèle de classification appliqué à la base de données HAM10000, il est essentiel d'utiliser un ensemble varié de métriques pour obtenir une image complète de sa performance.

Parmi ces métriques, l'accuracy se démarque comme une mesure intuitive qui évalue le pourcentage d'échantillons correctement classés par rapport au total des échantillons. Cependant, dans les cas où les classes sont déséquilibrées, la balanced accuracy offre une vue plus équilibrée en tenant compte des déséquilibres.

$$Accuracy = \frac{TP}{TP + TN + FP + FN}$$
$$Balanced\ accuracy = \frac{1}{N} \sum_{i=1}^N \frac{TP_i}{TP_i + FN_i}$$

Le recall se concentre sur la capacité du modèle à détecter tous les échantillons positifs. Il est crucial dans les cas où la détection des faux négatifs est critique, car il mesure la sensibilité du modèle à ces échantillons.

$$Recall = \frac{TP}{TP + FN}$$

La précision, d'autre part, évalue la proportion d'échantillons positifs correctement identifiés parmi tous les échantillons identifiés comme positifs par le modèle. Elle est importante lorsque les faux positifs sont coûteux.

$$Precision = \frac{TP}{TP + FP}$$

Le F1-score, quant à lui, est une mesure harmonique de la précision et du rappel. Il offre une perspective équilibrée entre la précision et la capacité du modèle à identifier correctement tous les échantillons positifs. C'est particulièrement utile lorsque les classes sont déséquilibrées ou lorsque les faux négatifs sont coûteux, comme dans la détection de maladies.

$$F1_score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

Enfin, la matrice de confusion offre une visualisation détaillée des performances du modèle en mettant en évidence les vrais positifs, les vrais négatifs, les faux positifs et les faux négatifs. Elle permet une évaluation plus approfondie de la capacité du modèle à classer correctement les échantillons dans chaque classe.

En combinant ces différentes métriques, il est possible d'obtenir une évaluation complète et précise de la performance du modèle de classification appliqué à la base de données HAM10000, ce qui permet de prendre des décisions informées sur son utilisation et son amélioration.

2) Les études

Au fil des dernières années, la classification précise des lésions cutanées a fait l'objet de plusieurs études novatrices, principalement en utilisant la base de données HAM10000. Ces recherches ont significativement contribué aux avancées de la dermatologie numérique. Avant 2018, la base de données HAM10000 n'était pas disponible, et d'autres bases de données étaient utilisées dans les études.

Ainsi, en 2017, une recherche a exploré la classification en trois groupes - normal, anormal et mélanome - en utilisant la base de données PH2, constituée de 200 images provenant du service de dermatologie de l'Hôpital Pedro Hispano. Les résultats de cette étude, réalisée avec un réseau neuronal artificiel (ANN), ont montré des performances élevées avec une accuracy de 92,50%, une précision de 92,38% et un F1-score de 90,45%. (Ozkan & Koklu, 2017)

Cependant, il est essentiel de souligner que ces performances élevées peuvent être attribuées à la taille relativement réduite de la base de données PH2. Cette limitation a suscité une transition vers des ensembles de données plus vastes, notamment HAM10000 à partir de 2018. Cette démarche visait à assurer une généralisation plus robuste des modèles.

Maintenant, examinons de plus près les différentes études menées avec la base de données HAM1000, chacune adoptant des approches variées avec des modèles distincts :

En 2018, une étude a utilisé le modèle DenseNet afin de développer un système d'aide au diagnostic basé sur des caractéristiques morphologiques des lésions cutanées. Les résultats ont atteint une accuracy de 88,2%, une précision de 78,67% et un recall de 84,61%. (*Heller, Bussmann, Shah, Dean, & Papanikolopoulos, 2018*)

En 2019, une autre recherche s'est penchée sur la classification des images de cancer de la peau en utilisant le transfert d'apprentissage avec le modèle MobileNet. Les résultats ont montré une accuracy de 97%, une précision de 90%, et un recall de 91%. (*Younis, Bhatti, & Azeem, 2019*)

En 2020, une approche novatrice a été proposée avec un ResNet50 modifié, obtenant une accuracy de 93%, une précision de 88%, un recall de 86%, et un F1-score de 84%. (*Le, Le, Ngo, & Ngo, 2020*)

En 2021, une étude a utilisé un modèle de réseau neuronal convolutionnel profond (DCNN) avec une architecture basée sur le transfert d'apprentissage depuis ImageNet. Le modèle a atteint de très bonnes performances avec une accuracy de 91,43%, une précision de 96,57%, un recall de 93,66% et un F1-score de 93,10%. (*Ali, Miah, Haque, Rahman, & Islam, 2021*)

En 2022, une méthode appelée Bootstrapping Ensembles based Convolutional Neural Network (BE-CNN) a été introduite, affichant des performances solides avec une accuracy de 93,8%, une précision de 91,12%, un recall de 92,24%, et un F1-score de 93,13%. (*Kalaivani & Karpagaballi, 2022*)

Enfin, en 2023, une étude a combiné l'optimisation génétique et la recherche de particules (HG-PSO) avec le réseau de convolution YOLOv7 pour classifier les mélanomes cutanés. Les résultats ont montré une précision remarquable avec une accuracy de 98,86%, une précision de 99%, un recall de 98,85%, et un F1-score de 99,85%. (*Shaheen, 2023*)

Ainsi, les études ont adopté principalement des modèles de réseaux de neurones convolutionnels (CNN), démontrant la robustesse de cette approche dans la résolution de problèmes complexes de dermatologie numérique.

IV. Deep Learning and Computer Vision

1) Qu'est-ce que c'est ?

Le Deep Learning, représente une branche avancée de l'intelligence artificielle qui s'appuie sur l'utilisation de réseaux de neurones artificiels pour résoudre des problèmes complexes. Inspiré par le fonctionnement du cerveau humain, ce domaine se base sur des couches interconnectées de neurones pour apprendre des représentations hiérarchiques des données.

Cette technologie trouve un terrain d'application majeur dans le computer vision, une discipline qui vise à permettre aux machines d'interpréter et de comprendre visuellement le monde qui les entoure. Grâce au Deep Learning, les systèmes informatiques peuvent désormais interpréter, analyser et traiter des images avec une précision sans précédent. Cela se traduit par la capacité de détecter, reconnaître et classifier des objets, des formes et des structures visuelles complexes avec une fiabilité remarquable.

Les schémas de machine Learning (ML) et de Deep Learning (DL) illustrent souvent comment ces réseaux de neurones, en particulier dans le Deep Learning, sont capables de capturer des caractéristiques de plus en plus abstraites à mesure que l'information traverse les différentes couches du réseau. Cette hiérarchisation des caractéristiques permet d'obtenir des représentations plus significatives et détaillées des données, facilitant ainsi la résolution de problèmes complexes, notamment dans le domaine de la computer vision.

2) Apprentissage

Apprentissage supervisé :

Dans l'apprentissage supervisé, le modèle est entraîné sur un ensemble de données étiqueté, où chaque exemple du jeu de données est associé à une étiquette ou à une sortie souhaitée.

L'objectif principal est d'apprendre une fonction qui mappe les entrées aux sorties en se basant sur des exemples fournis dans le jeu de données d'entraînement.

Pendant la phase d'entraînement, le modèle ajuste ses paramètres en minimisant la différence entre ses prédictions et les étiquettes réelles du jeu de données d'entraînement. Une fois entraîné, le modèle peut être utilisé pour faire des prédictions sur de nouvelles données.

Voici quelques exemples d'algorithmes d'apprentissage supervisé : Régression linéaire, Régression logistique, Machine à vecteurs de support.

Apprentissage non supervisé :

A l'inverse, dans l'apprentissage non supervisé, le modèle est confronté à des données non étiquetées, et l'objectif est d'extraire des structures ou des motifs intrinsèques à partir de ces données.

Les algorithmes d'apprentissages non supervisé peuvent être utilisés pour regrouper des données en clusters ou pour réduire la dimension des données.

Voici quelques exemples d'algorithmes : K-means, Analyse en composantes principales (PCA)

3) Modèles de classification

Il existe de nombreux algorithmes d'apprentissage supervisé basés sur la classification. Voici une présentation de certains modèles :

Support Vector Machine (SVM) :

L'algorithme SVM est un classifieur linéaire qui cherche à séparer les ensembles de données à travers des lignes appelées hyperplans. Son objectif est de maximiser la distance entre la ligne séparatrice et les échantillons des différentes classes, dont les plus proches sont désignés comme vecteurs support.

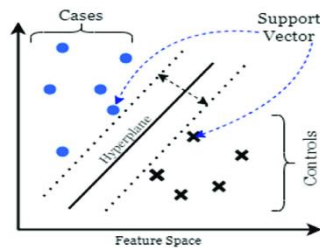


Fig. 5. Visualisation du Support Vector Machine

Arbre de décision :

Cet algorithme classe les données sous forme de branches, partant d'une racine où chaque donnée prend une direction en fonction de son comportement. Les nœuds représentent des règles de séparation des données en différentes catégories, et les feuilles contiennent les informations finales. Cet algorithme est souvent utilisé pour anticiper des variations, telles que les réactions du marché en cas de changement.

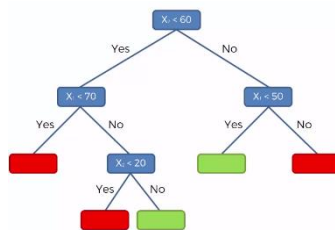


Fig. 6. Visualisation de l'Arbre de décision

K-means :

Cet algorithme trie les données en différents groupes en fonction de leurs caractéristiques. Pour cela, il établit une moyenne de référence parmi le jeu de données, ce qui permet alors de définir un profil type.

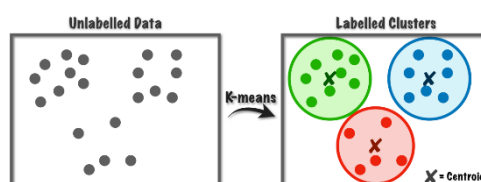


Fig. 7. Visualisation de K-means

K-nearest neighbors (KNN):

Le KNN analyse les similitudes entre les variables d'un jeu de données en calculant la distance entre les points, regroupant ceux qui sont les plus proches dans la même catégorie

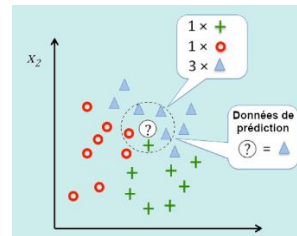


Fig. 8. Visualisation de K-nearest neighbors

4) CNN

Les réseaux de neurones convolutionnels (CNN) constituent une catégorie spécialisée d'apprentissage profond, particulièrement conçue pour la classification d'images. L'essence de leur fonctionnement repose sur la représentation des images d'entrée sous forme de matrices tridimensionnelles, où les deux premières dimensions traduisent une image en niveaux de gris, tandis que la troisième dimension encode les couleurs fondamentales (rouge, vert, bleu).

La première composante clé d'un CNN est sa partie convolutive. Cette phase est une opération mathématique qui utilise des filtres de convolution (gaussien, moyennneur...) pour extraire des caractéristiques significatives de l'image, générant ainsi des cartes de convolution. Ces cartes de convolution indiquent l'emplacement des traits et des éléments saillants dans l'image, contribuant ainsi à la représentation des caractéristiques visuelles.

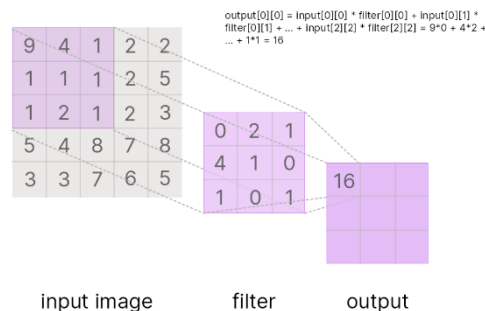


Fig. 9. Fonctionnement de la convolution

Après la convolution, intervient généralement une opération de pooling, souvent réalisée à l'aide du Max-Pooling. Ce processus réduit la dimension spatiale de la représentation, préservant les informations les plus pertinentes en sélectionnant la valeur maximale dans chaque région définie. Cette approche contribue à la robustesse du modèle tout en réduisant le risque de perte de détails importants.

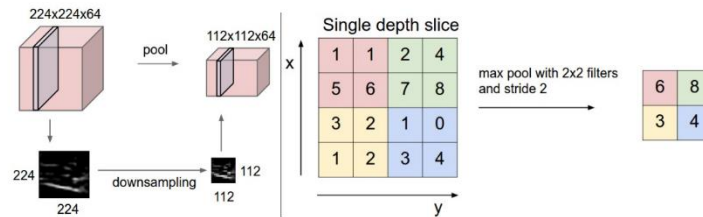


Fig. 10. Fonctionnement du Max-Pooling

Une étape cruciale dans l'architecture d'un CNN est le "Flatten". Placé entre la partie convolutive et la partie de classification, le "Flatten" convertit les données d'un format matriciel en un vecteur unidimensionnel, facilitant ainsi la transition des informations extraites par les couches convolutives vers les couches entièrement connectées du réseau.

La seconde composante majeure d'un CNN est la partie de classification, qui débute après le "Flatten". C'est à ce stade que les caractéristiques extraites sont utilisées pour attribuer des étiquettes ou des classes à l'image. Généralement, cela s'accomplit au moyen d'un modèle de perceptron multicouche (MLP). L'algorithme de rétropropagation du gradient est employé pour ajuster les poids du modèle, minimisant ainsi l'erreur de prédiction et améliorant les capacités de classification du réseau.

Un aspect fondamental dans le fonctionnement d'un CNN est l'activation. Les fonctions d'activation, telles que ReLU, sont appliquées à la sortie de chaque neurone dans une couche. Elles introduisent une non-linéarité dans le modèle, permettant au réseau d'apprendre des motifs et des relations complexes dans les données. Cette non-linéarité est cruciale pour introduire des capacités de généralisation et d'apprentissage plus complexes, améliorant ainsi les performances du modèle.

En somme, les CNN suivent généralement une architecture bien définie, combinant des couches de convolution, de pooling, d'activation ReLU, et des couches entièrement connectées. Ces réseaux ne se limitent pas à la seule classification d'images ; ils peuvent être appliqués à diverses tâches telles que la reconnaissance des formes, la détection d'anomalies, et d'autres domaines liés au traitement des données visuelles.

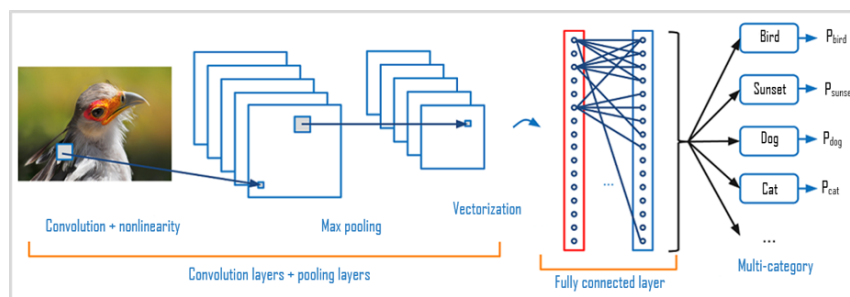


Fig. 11. Réseau de neurones convolutionnel

V. Notre choix de modèle : YoLov8

1) Présentation

YOLO (You Only Look Once) est une famille de modèles de détections d'objet en vision par ordinateur développée par Ultralytics. La série de modèles YOLO est devenue célèbre grâce à une précision excellente tout en conservant une taille de modèle réduite.

YOLOv8 est le tout dernier modèle YOLO à la pointe de la technologie qui peut être utilisé pour la détection d'objets, la classification d'images et les tâches de segmentations d'instances.

2) Méthodologie

Dans le cadre de notre étude, nous avons suivi une approche méthodique pour exploiter le dataset HAM10000, composé de 7 classes distinctes. La première étape a consisté à organiser les données en regroupant les différentes classes dans des dossiers spécifiques. Par la suite, chaque classe a été subdivisée en échantillons distincts pour l'entraînement (60%), la validation (20%), et les tests (20%).

Le modèle YOLOv8 a été sélectionné et appliqué à notre ensemble de données, suivant les étapes suivantes :

- **Chargement du Modèle YOLO préentraîné** : Nous avons débuté par charger un modèle YOLO préentraîné sur la base de données ImageNet, prêt à être adapté à nos données spécifiques. Cette étape est cruciale pour tirer parti des caractéristiques apprises sur des données similaires.
- **Entraînement du Modèle** : L'étape suivante consiste à entraîner le modèle sur nos données. Le nombre d'époch et la taille des images étaient des paramètres déjà préétabli par Ultralytics.
- **Validation du Modèle** : Après l'entraînement, nous avons validé le modèle pour évaluer sa performance. Cette étape a impliqué une vérification des paramètres, des hyperparamètres et des métriques choisies.
- **Prédiction sur l'Ensemble de Test** : Enfin, une fois le modèle validé, nous avons procédé à des prédictions sur l'ensemble de test. Cela nous a permis d'évaluer la capacité du modèle à généraliser et à fournir des prédictions précises sur des données qu'il n'avait pas rencontrées auparavant.

Une fois l'entraînement effectué sur la base de données initiale, nous avons réfléchi aux divers moyens possibles pour améliorer les performances du modèle.

a. Equilibrage de la base de données :





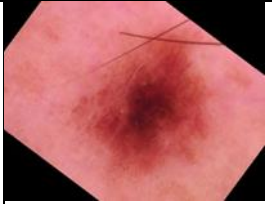
L'augmentation de données est une étape essentielle dans le processus de préparation des données pour l'apprentissage automatique, surtout lorsqu'on travaille avec des ensembles de données déséquilibrés comme celui de la base de données HAM10000. Cette méthode consiste à générer de nouvelles données à partir des données existantes, dans le but d'équilibrer les différentes classes et d'améliorer les performances du modèle.

Dans notre cas, la base de données HAM10000 présente un déséquilibre significatif en termes de nombre d'images par classe. Pour remédier à cette situation, nous avons choisi de standardiser le nombre d'images par classe à 2000. Pour les classes sous-représentées, nous avons sélectionné des images aléatoirement et appliqué des techniques d'augmentation de données, telles que la rotation, le retournement, le zoom ou le recadrage, pour créer de nouvelles images jusqu'à atteindre le nombre

désiré (upsampling). En revanche, pour les classes sur-représentées, nous avons procédé à la suppression aléatoire d'images excédentaires afin de ramener leur nombre à 2000 (downsampling). Cette démarche nous a permis de créer un ensemble de données plus équilibré, offrant ainsi une représentation plus juste de chaque classe. En standardisant le nombre d'images par classe, nous nous assurons que le modèle dispose d'une quantité suffisante de données pour apprendre efficacement les caractéristiques de chaque classe, tout en évitant le surapprentissage ou le sous-apprentissage. En conséquence, nous espérons obtenir des performances plus fiables et généralisables lors de l'évaluation du modèle sur des données du test.

Nous avons ainsi implémenté différentes transformations d'images telles que l'ajout de formes géométriques, le cisaillement, la réflexion, la rotation et l'ajustement du contraste. Chaque transformation est appliquée de manière aléatoire pour enrichir notre ensemble d'images, ce qui permet d'imiter une variété de conditions que le modèle pourrait rencontrer dans des scénarios réels. Par exemple, la rotation des images peut simuler des prises de vue sous différents angles, tandis que l'ajustement du contraste peut représenter des variations d'éclairage.

Voici dans la table suivante, une représentation visuelle de chaque méthode utilisée.

Image de référence		
Ajout d'un rectangle		
Cisaillement		
Réflexion aléatoire		
Rotation		

Ajustement du contraste	
-------------------------	--

Table 4 : Les différentes méthodes de up-sampling

b. Heatmap sur les prédictions avec Eigencam

Dans le cadre de notre analyse approfondie du modèle, nous avons mis en œuvre la méthode EigenCAM pour produire des cartes de chaleur sur des images de test. L'objectif de cette démarche était d'identifier précisément les zones d'une image qui ont le plus influencé la décision de notre modèle de machine learning, offrant ainsi une meilleure compréhension de la manière dont celui-ci perçoit et interprète les données visuelles.

EigenCAM est une technique d'explication de modèle qui cherche à visualiser les régions d'une image ayant le plus contribué à la décision d'un modèle de deep learning. Cette méthode repose sur l'identification des vecteurs propres qui maximisent la covariance entre les activations de différentes couches du réseau lorsqu'une image spécifique est traitée. En d'autres termes, EigenCAM vise à découvrir les motifs ou caractéristiques visuelles au sein de l'image ayant le plus influencé la prédiction finale du modèle.

L'analyse avec EigenCAM a révélé un aspect crucial de la préparation des données pour notre modèle : la nécessité de segmenter correctement les images. Nous avons observé que la méthode de prise de photo des anomalies pouvait inclure sur certaines images la présence de bords noirs. Ces coins ne sont pas des caractéristiques pertinentes pour la tâche de classification ou de détection, et biaisent la prédiction du modèle comme l'indique la couleur rouge dans la figure suivante.

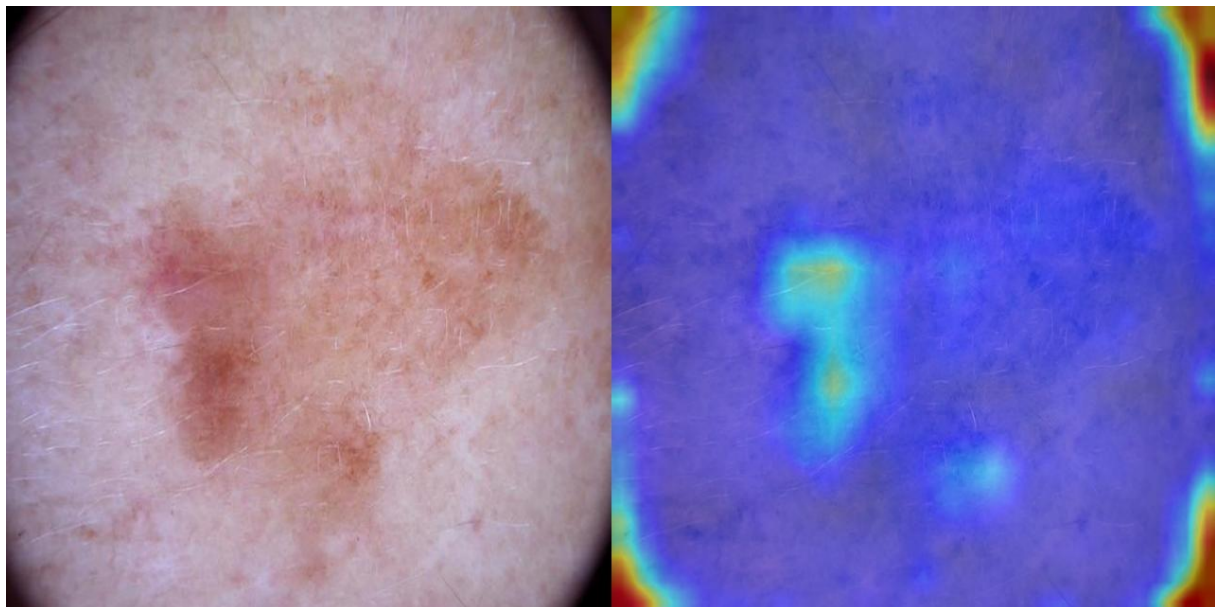


Fig. 12. Image sans et avec EigenMap

Ainsi, l'utilisation d'EigenCAM a souligné l'importance de la segmentation d'image dans notre processus de préparation des données. La segmentation d'image consiste à diviser une image en

parties ou régions, souvent pour isoler les régions d'intérêt et éliminer le fond ou les parties non pertinentes. En segmentant correctement les images, nous nous assurons que le modèle se concentre uniquement sur les caractéristiques significatives de l'image qui sont réellement indicatives de la classe ou de l'objet à détecter.

c. Prétraitement des images

La phase de traitement ou de processing que nous avons mise en œuvre intègre des techniques de filtrage avancées pour optimiser l'extraction des caractéristiques (features) à partir des images, une étape cruciale pour l'amélioration de la précision dans les applications d'analyse d'images.

Filtre Sharpen :

L'application du filtre Sharpen dans le processus de prétraitement des images pour notre modèle de deep learning est justifiée par sa capacité à améliorer la netteté et la précision des détails dans les images. Cette technique est particulièrement utile pour accentuer les contours des objets dans l'image, ce qui facilite la distinction entre les différentes classes d'anomalies dermatologiques.

Mathématiquement, le filtre Sharpen opère en modifiant les valeurs de pixel de l'image originale en fonction de leur différence avec les valeurs des pixels voisins. Plus spécifiquement, cela peut être réalisé à l'aide d'un noyau de convolution, qui est appliqué à chaque pixel de l'image. Le noyau de convolution, souvent représenté sous forme de matrice, définit les poids attribués à chaque pixel voisin lors du calcul de la nouvelle valeur du pixel cible.

Nous avons utilisé le filtre Laplacien dans notre cas, voici sa formule mathématiques :

$$\Delta^2 I(x, y) = \sum_{i=-1}^1 \sum_{j=-1}^1 \omega(i, j) \times I(x + i, y + j)$$

Où $I(x, y)$ représente la valeur d'un pixel donné et $\omega(i, j)$ les coefficients du masque de Laplace, un exemple typique étant le suivant :

$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

En appliquant ce filtre, notre objectif est d'améliorer la qualité visuelle des images dermatologiques en accentuant les contours des lésions cutanées et en augmentant leur netteté. Cela facilite non seulement l'identification des différentes classes d'anomalies dermatologiques, mais aussi l'extraction des caractéristiques pertinentes pour la classification par notre modèle de deep learning.

Filtre Enhance :

À la suite de l'application du filtre Sharpen, nous avons ajouté une étape supplémentaire en appliquant le filtre Enhance à nos images. Le filtre Enhance vise à sublimer l'image en améliorant les couleurs, la luminosité et le contraste, contribuant ainsi à rendre l'image globalement plus attrayante et esthétiquement plaisante.

Mathématiquement, le filtre Enhance fonctionne en ajustant les valeurs de pixel de l'image pour augmenter la saturation des couleurs et l'intensité des contrastes. Cela peut être réalisé en utilisant

diverses techniques telles que l'accentuation des couleurs, l'ajustement de la courbe de luminosité ou l'égalisation de l'histogramme.

En résultat, l'image obtenue après l'application du filtre Enhance est plus vivante et dynamique. Les couleurs sont plus riches, les contrastes sont plus prononcés et la luminosité est optimisée. Chaque détail de l'image ressort avec une plus grande intensité, ce qui contribue à une meilleure perception visuelle globale.

Cette étape de post-traitement permet non seulement d'améliorer la qualité visuelle des images dermatologiques, mais également de rendre les caractéristiques des lésions cutanées plus distinctes et facilement discernables par notre modèle de deep learning. Ainsi, cela contribue à améliorer la performance de notre algorithme de classification en fournissant des données d'entrée visuellement plus riches et informatives.

Filtre Clahe :

En complément du filtre Sharpen et à la place de Enhance, nous avons également appliqué le filtre CLAHE (Contrast Limited Adaptive Histogram Equalization) à nos images. Le CLAHE est une méthode avancée d'égalisation de l'histogramme adaptatif conçue pour améliorer le contraste dans les images.

Contrairement aux méthodes d'égalisation traditionnelles, le CLAHE fonctionne par régions de l'image, adaptant le contraste de chacune de manière à rendre les détails dans les zones sombres et claires plus perceptibles, sans amplification du bruit. Cette approche permet de traiter l'image par petits segments, avec une égalisation de l'histogramme pour chaque segment, avant de combiner harmonieusement l'ensemble pour obtenir un résultat final équilibré et détaillé.

En résultat, l'image obtenue après l'application du filtre CLAHE présente un contraste amélioré et des détails plus nets, ce qui permet une meilleure perception visuelle des caractéristiques des lésions cutanées. Cette étape de post-traitement contribue ainsi à enrichir les données d'entrée de notre modèle de deep learning, en fournissant des images plus détaillées et informatives pour la classification des anomalies dermatologiques.

d. Sélection de la région d'intérêt (ROI)

La segmentation d'images constitue une étape cruciale en vision par ordinateur, visant à diviser une image en régions d'intérêt distinctes. Dans notre contexte, l'objectif principal est de détecter la position de l'anomalie dermatologique sur les images de la base de données, afin de délimiter une région d'intérêt plus restreinte autour de cette anomalie. Cette approche permet de réduire la taille des images à analyser, ce qui présente plusieurs avantages stratégiques.

Tout d'abord, en diminuant la taille des images, nous réduisons la quantité de données à traiter pour notre modèle de deep learning. Cela se traduit par des temps de calcul réduits et une consommation de ressources informatiques moindre, tout en maintenant des performances élevées en termes de précision et de vitesse d'exécution.

Ensuite, la réduction de la taille des images permet de concentrer l'attention du modèle sur la région d'intérêt spécifique, à savoir l'anomalie dermatologique, en éliminant les détails parasites tels que les poils ou autres artefacts. Cette concentration accrue sur la région d'intérêt améliore la capacité du modèle à détecter et à classer correctement les anomalies cutanées, en minimisant les distractions visuelles potentielles.

Pour parvenir à cette segmentation des images, plusieurs méthodes de prétraitement ont été envisagées (détection de blobs, détection de contours et suppression des poils). Enfin, pour une

segmentation sémantique plus avancée, nous avons également envisagé l'utilisation d'un modèle de deep learning tel que U-Net.

Détection de blobs :

La méthode de détection de blobs est une technique de segmentation d'image largement utilisée en vision par ordinateur. Elle vise à détecter les régions d'intérêt en identifiant des groupes de pixels ayant des caractéristiques distinctes, telles que des variations d'intensité lumineuse ou de couleur. Ces groupes de pixels, appelés "blobs", représentent des régions homogènes de l'image qui peuvent correspondre à des objets, des contours ou des anomalies.

Dans ce processus, l'image d'entrée est d'abord convertie en niveaux de gris pour simplifier le traitement. Ensuite, une opération de flou gaussien est appliquée pour réduire le bruit de l'image. Une binarisation est ensuite effectuée pour segmenter l'image en pixels d'avant-plan et d'arrière-plan. La morphologie mathématique est ensuite utilisée pour ouvrir et fermer les contours binaires, ce qui aide à éliminer le bruit et à lisser les contours des blobs.

Enfin, la détection de blobs est effectuée à l'aide de l'algorithme de détection de blobs, où les paramètres tels que la taille des blobs, la circularité ou même l'inertie peuvent être spécifiés. Nous avons effectué plusieurs tests pour obtenir la meilleure combinaison des paramètres mais avons eu du mal à identifier correctement l'anomalie sur les images. A titre d'exemple, voici une image où cette méthode n'a pas détecté l'anomalie principale de l'image, mais une simple variation de couleur de la peau n'ayant aucun rapport.

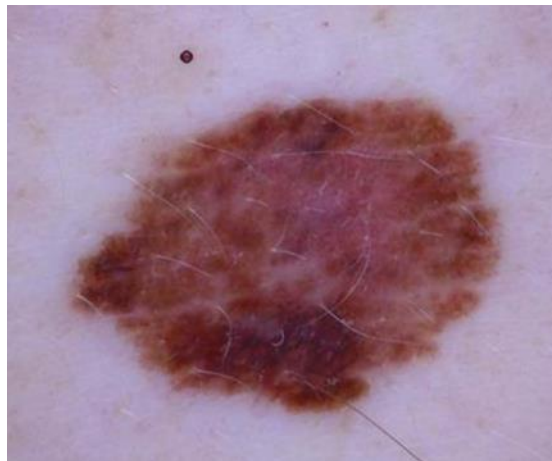


Fig. 13. Image avec l'application de Blob Detection

À la suite de nos difficultés rencontrées avec cette méthode, nous avons décidé de procéder à une autre façon.

Détection de contours :

Cette méthode repose sur l'identification des contours des objets dans l'image. Ici aussi, nous avons appliqué un filtre gaussien aux images et utilisé des formules morphologiques pour simplifier le traitement. En comparaison avec la méthode de détection de blobs, la méthode findContours diffère dans sa manière de détecter et de représenter les objets dans l'image. Alors que la détection de blobs identifie des régions d'intérêt homogènes basées sur des critères tels que la taille et la forme, la méthode findContours détecte les contours des objets en se basant sur les variations de luminosité et de couleur entre les pixels voisins. De plus, la méthode findContours permet une représentation plus

précise des contours des objets, ce qui en fait une technique adaptée pour la détection et la segmentation d'objets avec des formes complexes ou irrégulières.

La méthode s'est révélée très efficace sur nos images et a détecté l'anomalie à plusieurs reprises comme le montre la figure suivante. Les résultats avec cette méthode sont beaucoup plus concluants par-rapport à la méthode de détection de blobs et c'est donc `findContours` que l'on a utilisé pour notre partie de preprocessing.

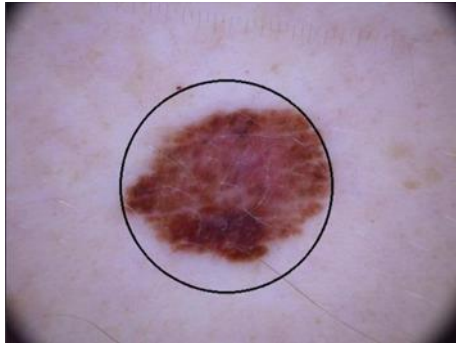


Fig. 14. Image avec l'application de `findContours`

Une fois les contours détectés, nous avons procédé à l'extraction du centre du cercle représentant l'anomalie ainsi que son rayon. Ces informations nous ont ensuite permis de définir un cadre qui englobe entièrement le cercle de l'anomalie. Cependant, pour garantir la capture de toute l'anomalie, même en présence de défauts ponctuels ou de bords flous, nous avons ajouté une marge supplémentaire de manière aléatoire autour du cadre. Cette marge supplémentaire nous assure que le cadre inclut non seulement l'anomalie visible, mais également toute variation ou étendue potentielle de celle-ci, ce qui garantit une détection complète et précise de l'anomalie dans l'image.

Ainsi, nous obtenons une deuxième image plus petite que l'originale et centrée sur l'anomalie dermatologique. On peut distinguer cette amélioration dans les 2 figures qui suivent où l'image réduite fait environ 75% de l'image initiale.

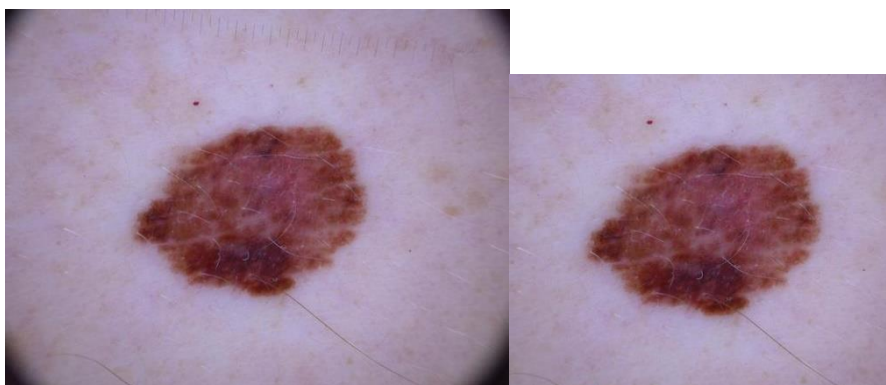


Fig. 15. Image avant/après le crop

Suppression des poils :

Une étape cruciale dans le prétraitement des images dermatologiques est la suppression des poils et autres artefacts indésirables, qui peuvent introduire du bruit et affecter la précision des analyses

ultérieures. Pour atteindre cet objectif, nous avons exploré l'utilisation d'algorithmes de traitement d'image, notamment les filtres de lissage et les techniques de détection de contours.

Nous avons commencé par appliquer le filtre de Canny, qui est une technique populaire de détection de contours basée sur la dérivation des gradients d'intensité dans l'image. Le filtre de Canny permet de détecter les bords nets et les transitions d'intensité dans l'image, ce qui est crucial pour identifier les contours des poils et autres détails indésirables.

Une fois les contours détectés à l'aide du filtre de Canny, nous avons fait appel à la fonction `HoughLinesP` de la bibliothèque OpenCV. Cette fonction est utilisée pour détecter les lignes droites dans une image, ce qui est particulièrement utile pour extraire les structures linéaires telles que les poils dans notre cas d'étude.

La méthode de Hough transforme l'espace des paramètres de l'image en un espace de paramètres de ligne, où chaque ligne dans l'image est représentée par un point. En utilisant la transformée de Hough probabiliste implémentée dans la fonction `HoughLinesP`, nous avons pu identifier les lignes représentant les poils dans l'image. Ces lignes ont ensuite été supprimées de l'image, laissant ainsi une version lissée et dépourvue d'artefacts indésirables.

Cette approche de suppression des poils à l'aide du filtre de Canny et de la fonction `HoughLinesP` s'est avérée efficace pour nettoyer les images dermatologiques et préparer les données d'entrée pour notre modèle de deep learning. Elle nous a permis de concentrer l'attention du modèle sur les régions d'intérêt pertinentes, telles que les lésions cutanées, en éliminant les distractions liées aux poils et aux autres éléments non pertinents.

En jouant sur les paramètres de la fonction de `HoughLinesP`, nous avons pu arriver à des résultats visuellement convaincants sur les images de notre base de données comme le montre la figure ci-dessous. Sur la première image, les lignes détectées correspondent aux poils de l'image comme voulu. Sur la deuxième image, ces lignes ont été lissées en remplaçant les pixels concernés par la moyenne des pixels voisins.



Fig. 16. Image avant/après la suppression des poils

Les résultats étant convaincants, nous allons utiliser cette méthode dans notre partie de preprocessing.

Segmentation sémantique avec U-Net :

Dans le cadre de notre processus de recherche de méthodes de segmentation sémantique avancée, nous avons considéré l'utilisation d'un modèle de deep learning CNN bien connu, U-Net. Ce modèle est réputé pour sa capacité à réaliser la segmentation d'images biomédicales, une tâche essentielle dans notre contexte de détection d'anomalies dermatologiques. En effet, U-Net est conçu pour identifier

automatiquement les régions d'intérêt dans une image, y compris les anomalies dermatologiques, avec une grande précision et efficacité.

L'architecture en forme de U de ce modèle lui confère une capacité unique à capturer à la fois les détails fins et les informations contextuelles dans les images, ce qui en fait un choix particulièrement prometteur pour notre application spécifique. Cependant, malgré ses avantages indéniables, nous avons été contraints de mettre de côté l'utilisation de U-Net dans notre projet.

La principale raison de cette décision réside dans l'absence d'un modèle pré-entraîné sur la base de données HAM10000, qui est la base de notre étude. En effet, l'entraînement d'un modèle U-Net à partir de zéro sur un ensemble de données aussi vaste aurait demandé des ressources computationnelles et temporelles considérables. De plus, la création manuelle des masques pour l'ensemble de la base de données aurait été une tâche extrêmement chronophage et fastidieuse, ce qui aurait retardé considérablement notre avancée dans le projet.

Face à ces contraintes, nous avons dû prendre la décision pragmatique de ne pas utiliser U-Net pour la segmentation sémantique dans notre projet. Nous sommes donc resté avec nos solutions précédents de findContours pour la sélection de la région d'intérêt, tout en gardant à l'esprit les avantages potentiels qu'offrirait l'utilisation de U-Net dans des contextes futurs où les ressources et le temps nécessaires à son déploiement seraient disponibles.

3) Résultats

a. Méthode 1

Nos premières analyses portent sur la phase de validation du modèle YoLo sur la base de données HAM10000 sans aucun traitement de notre part. Les figures suivantes représentent les courbes d'entraînement et la matrice de confusion.

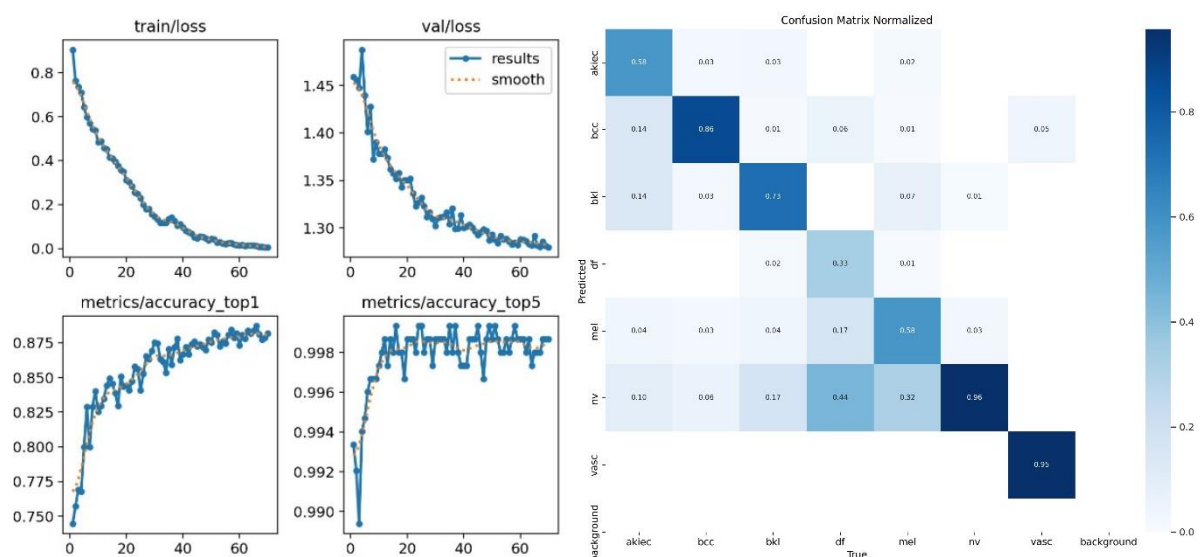


Fig. 17. Résultats

L'analyse des courbes de train loss, val loss et top1 accuracy en fonction des epochs révèle plusieurs informations importantes sur les performances et le comportement du modèle pendant l'entraînement.

Premièrement, on observe une baisse constante de la train loss tout au long des epochs, ce qui correspond à l'erreur du modèle sur les données d'entraînement. Cette baisse indique que le modèle apprend efficacement à partir des données fournies, et cette convergence vers une valeur basse suggère que le modèle parvient à capturer les motifs et les caractéristiques des données.

De même, la baisse de la val loss sans aucune hausse significative suggère l'absence de sur-apprentissage, c'est-à-dire que le modèle ne sur-ajuste pas aux données d'entraînement au détriment de la généralisation aux données de validation. Cette observation est cruciale pour garantir que le modèle est capable de généraliser ses connaissances à de nouvelles données et donc de prédire avec précision sur des données non vues.

Plus le nombre d'epochs augmente, plus les valeurs de train loss et val loss ont tendance à se stabiliser autour de valeurs spécifiques (par exemple, 0.05 pour l'entraînement et 1.3 pour la validation). Cette stabilisation suggère que le modèle a atteint son maximum de potentiel d'apprentissage et qu'il n'est pas nécessaire d'itérer sur plus d'epochs. En d'autres termes, le modèle a appris autant qu'il le peut des données disponibles.

De plus, la baisse de la loss s'accompagnant d'une hausse de l'accuracy indique que l'amélioration de la performance du modèle sur les données d'entraînement se traduit par une meilleure capacité à prédire correctement les étiquettes des données. Cette corrélation entre la perte et la précision renforce la confiance dans les performances du modèle.

Enfin, l'absence de tendance cyclique dans les courbes confirme que les hyperparamètres du modèle ont été correctement choisis, ce qui signifie que le modèle ne souffre pas de problèmes tels que l'instabilité ou le sur-apprentissage cyclique.

En analysant la matrice, nous observons que certaines classes sont bien prédites, notamment la classe nv, avec un taux de prédiction de 0,96. De même, la classe vasc présente des résultats excellents. En revanche, les autres classes ne sont pas prédites avec précision. Par exemple, la classe df est faiblement prédite (0,33) et est fortement confondue avec la classe nv, affichant un taux de 0,44.

b. Méthode 2

Notre deuxième méthode concerne la base de données initiale qui a été filtrée avec Sharpen puis Enhance. Les figures suivantes représentent les courbes d'entraînement et la matrice de confusion.

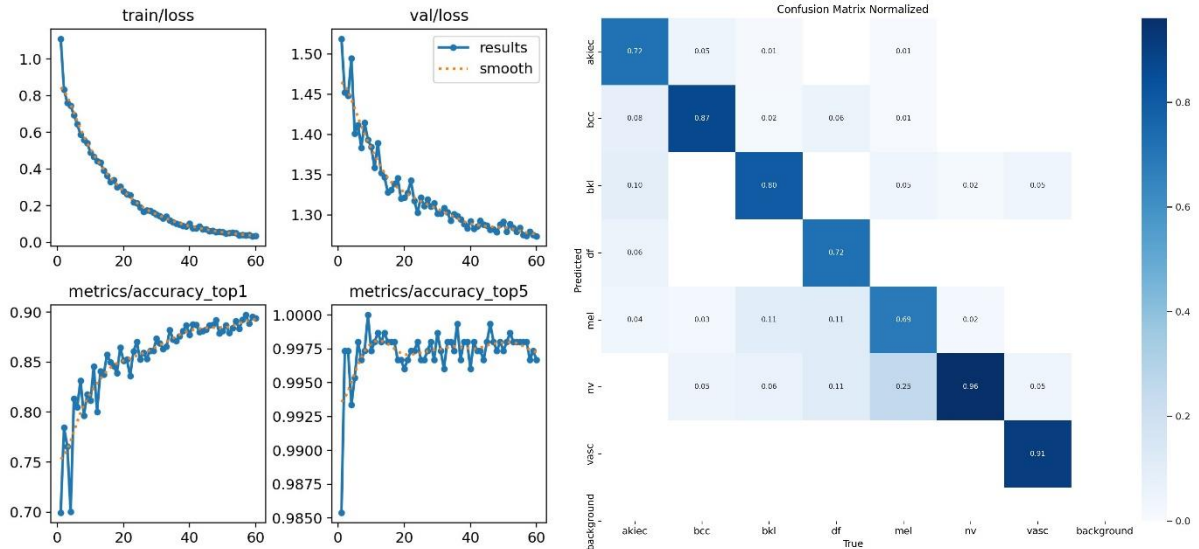


Fig. 18. Résultats

L'analyse des courbes est la même que précédemment, on ne distingue aucune anomalie. En analysant la matrice, nous observons que globalement les classes ont été mieux prédites que précédemment avec une Accuracy minimale de seulement 0,69 pour la classe mel. Voici les métriques calculées : Balanced Accuracy 85.4 %, Recall 82.9% et F1-Score 80.7%.

c. Méthode 3

Notre troisième méthode concerne la base de données initiale qui a été filtrée avec Sharpen puis Clahe. Les figures suivantes représentent les courbes d'entraînement et la matrice de confusion.

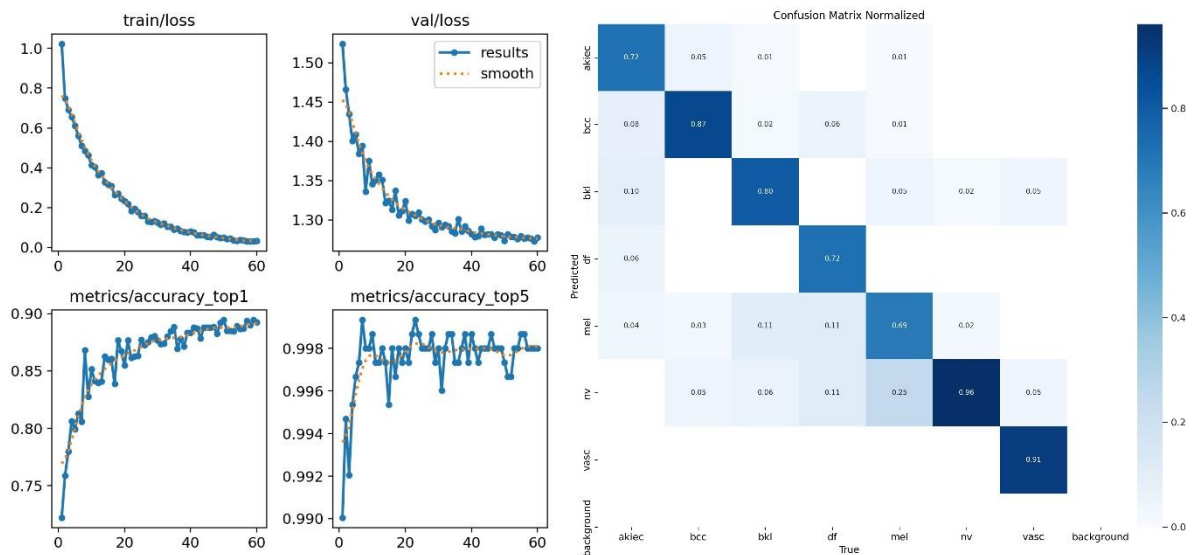


Fig. 19. Résultats

L'analyse des courbes est la même que précédemment, on ne distingue aucune anomalie.

En ce qui concerne la matrice de confusion, on retrouve les mêmes tendances qu'avec Enhance mais avec des performances très légèrement supérieures qui se ressentent aux métriques : Balanced Accuracy 86.3%, Recall 84.2% et F1-Score 81.9%.

d. Méthode 4

Notre quatrième méthode concerne la base de données initiale qui a été augmentée mais qui n'a reçu aucun autre prétraitement. Les figures suivantes représentent les courbes d'entraînement et la matrice de confusion.

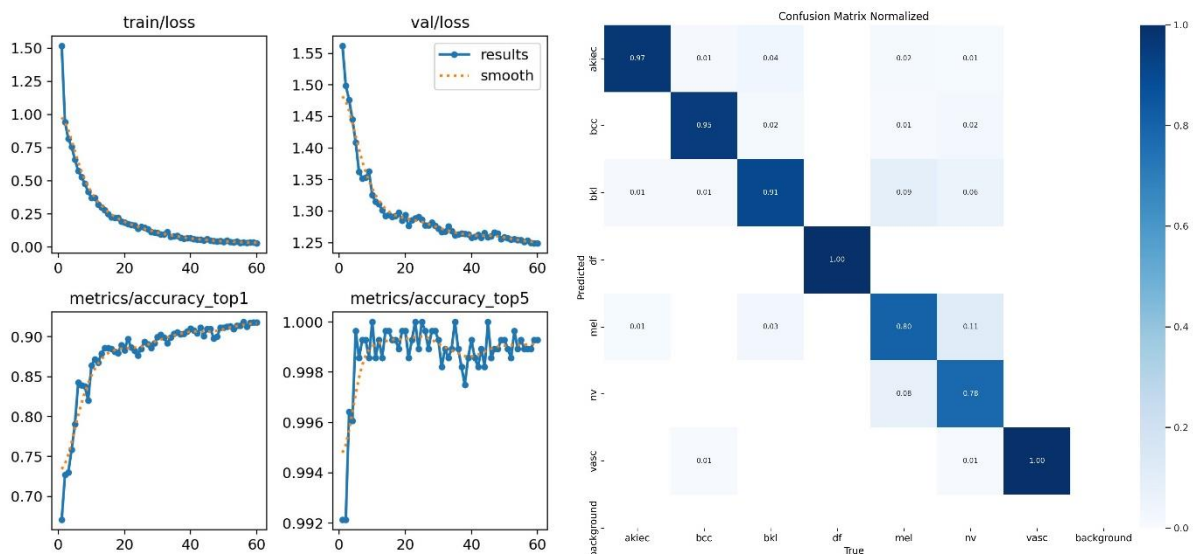


Fig. 20. Résultats

L'analyse des courbes est la même que précédemment, on ne distingue aucune anomalie.

Nous obtenons ici nos premiers résultats parfaits : la classe df et vasc ont été prédites correctement à 100%, notre but est d'obtenir ce résultat pour toutes les classes. Les classes ayant été équilibrées, nous pouvons désormais passer de la Balanced Accuracy à l'Accuracy simple : Accuracy 91.8%, Recall 91.6% et F1-Score 93.3%.

e. Méthode 5

Notre cinquième méthode concerne la base de données initiale qui a été augmentée et segmentée grâce à findContours. Les figures suivantes représentent les courbes d'entraînement et la matrice de confusion.

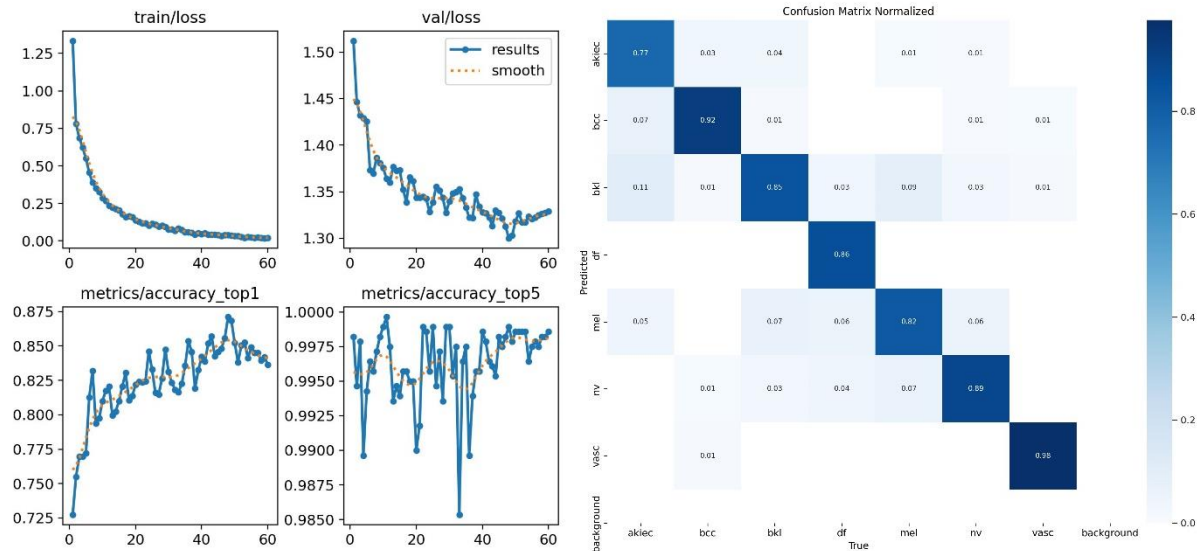


Fig. 21. Résultats

Pour la courbe de la loss du train, nous ne voyons toujours aucune anomalie. Cependant, on peut observer une remontée de la loss de la validation et une baisse de l'accuracy en fin d'époque ce qui témoigne que le modèle commençait à entrer dans une phase de sur-apprentissage. C'est donc sans étonnement que la matrice de confusion de cette méthode affiche de moins bons résultats que précédemment. Les métriques ont également chuté pour obtenir les scores suivants : Accuracy 87%, Recall 86% et F1-Score 85.1%.

f. Méthode 6

Notre sixième méthode concerne la base de données initiale qui a été augmentée et cette fois-ci privée de ses poils. Les figures suivantes représentent les courbes d'entraînement et la matrice de confusion.

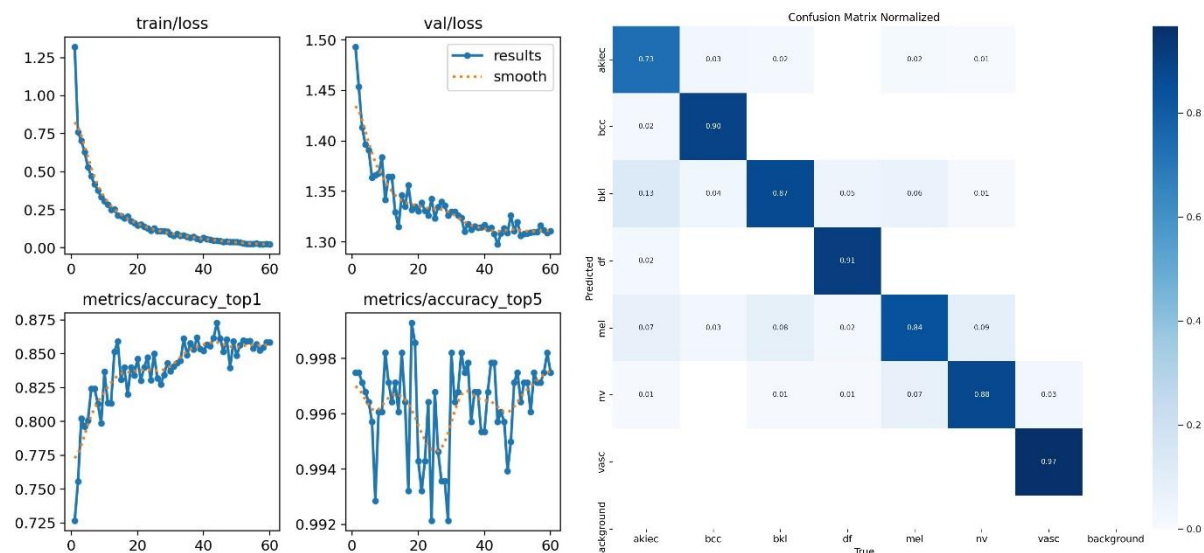


Fig. 22. Résultats

Dans ce cas-là, les courbes de la loss pour la validation et l'Accuracy ne sont pas parfaites mais ne sortent pas de la norme pour autant, elles sont correctes.

La matrice de confusion affiche des résultats légèrement supérieurs aux précédents avec une Accuracy de 87.3%, un à Recall 88.2% et un F1-Score de 86%.

g. Méthode 7

Notre dernière méthode regroupe l'ensemble des méthodes que l'on a retenues : l'augmentation des données, la sélection d'intérêt ainsi que la suppression des poils. Les figures suivantes représentent les courbes d'entraînement et la matrice de confusion.

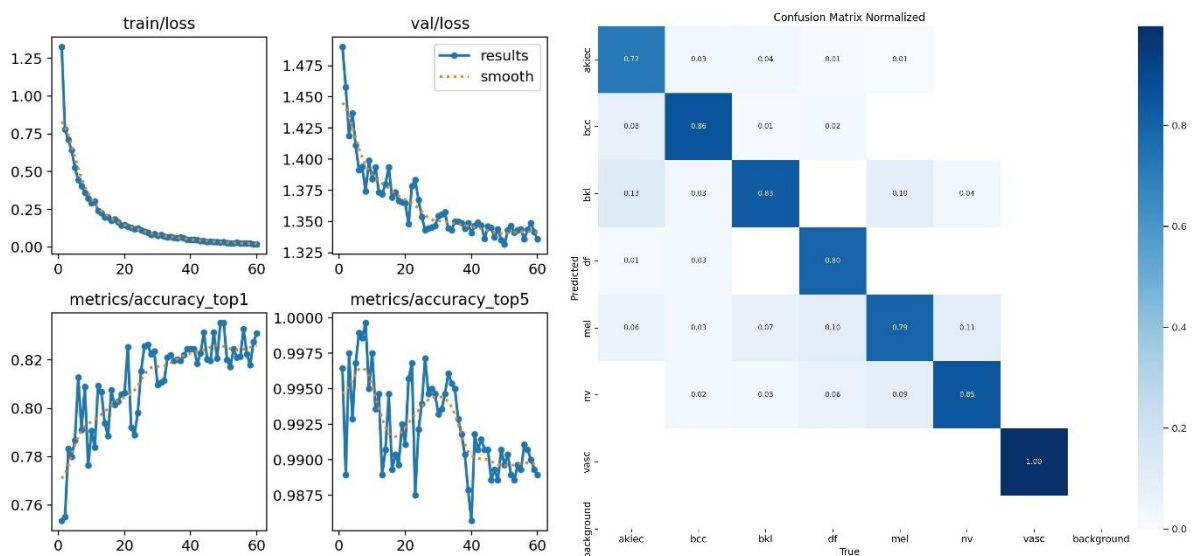


Fig. 23. Résultats

Dans cette situation, les courbes de loss du train et de la validation retrouvent une tendance normale, mais l'Accuracy a du mal à converger et ce même après 60 epoch. Le modèle n'est donc pas parfait.

La matrice de confusion affiche des résultats corrects mais pas encore suffisants, avec une Accuracy pour la classe akiec de seulement 0.72 tandis que celle de vasc est stable à 1. On obtient une Accuracy de 83.5%, un à Recall 84.5% et un F1-Score de 82%.

VI. Conclusion

Pour récapituler, voici nos résultats obtenus avec les différents changements que nous avons octroyé à la base de données.

	Enhance	Clahe	Augmentation	Segmentation	Suppression des poils	Accuracy	Balanced accuracy	Recall	F1-Score
Méthode 1						/	87.6%	86%	85.1%
Méthode 2	X					/	85.4%	82.9%	80.7%
Méthode 3		X				/	86.3%	84.2%	81.9%
Méthode 4			X			91.8%	/	91.6%	93.3%
Méthode 5			X	X		87	/	86%	85.1%
Méthode 6			X		X	87.3%	/	88.2%	86%
Méthode 7			X	X	X	83.5%	/	84.5%	82%

Table 5 : Récapitulatif des performances

Les meilleurs résultats ont été obtenus en augmentant la base de données sans y ajouter aucun autre traitement. Si on y rajoute des modifications, on voit que les performances chutent d'un coup, ce qui montre que les paramètres qui y sont utilisés ne sont pas encore optimisés pour HAM10000 (meilleure détection des contours pour findCircles, meilleurs seuils pour la détection des lignes pour la suppression des poils).

Au terme de ce projet, nous avons pu constater les progrès significatifs réalisés dans la détection et la classification des anomalies dermatologiques grâce à l'intégration de l'intelligence artificielle. Notre approche, axée sur l'utilisation de modèles de deep learning tels que YoLov8, a permis d'obtenir des résultats prometteurs, notamment avec la méthode 4 impliquant une base de données augmentée. Les performances obtenues avec cette méthode démontrent l'efficacité de l'apprentissage automatique dans la reconnaissance des lésions cutanées, offrant ainsi des perspectives nouvelles dans le domaine de la dermatologie. En exploitant les avancées de l'intelligence artificielle, nous avons pu non seulement améliorer la précision du diagnostic, mais également accroître l'efficacité et la rapidité du processus, ouvrant ainsi la voie à une prise en charge plus précoce et plus efficace des patients. L'utilité plus qu'importante de l'IA dans la classification des anomalies dermatologiques réside dans sa capacité à analyser de vastes ensembles de données, à détecter des motifs subtils et à généraliser à partir de ces informations pour des applications pratiques. En outre, l'IA permet une personnalisation accrue des traitements et des interventions, en prenant en compte les caractéristiques individuelles des patients et en adaptant les recommandations en conséquence.

Au-delà de ses avantages dans le domaine médical, l'intelligence artificielle ouvre également des perspectives dans la recherche, la prévention et l'éducation en matière de santé cutanée. En permettant une analyse approfondie des images dermatologiques, elle contribue à l'identification de nouvelles associations entre les caractéristiques des lésions et leur évolution, facilitant ainsi la découverte de nouveaux traitements et stratégies de prévention.

Antoine Boukhelifa
Giovanni Vagliani

En conclusion, ce projet met en lumière le potentiel immense de l'intelligence artificielle dans le domaine de la dermatologie, offrant des solutions innovantes pour la détection précoce, le diagnostic précis et la prise en charge efficace des anomalies cutanées. Grâce à une approche intégrée combinant l'expertise médicale et les avancées technologiques, nous sommes sur la voie de transformer radicalement la manière dont nous comprenons, diagnostiquons et traitons les maladies dermatologiques, pour le bénéfice de tous.

VII. Perspectives

Malgré les avancées significatives réalisées dans ce projet, plusieurs axes d'amélioration et de développement se profilent à l'horizon. Tout d'abord, l'exploration de YOLOv9, la version la plus récente de l'architecture YOLO, pourrait être une piste intéressante à suivre. Cette version apporte des améliorations en termes de précision et de vitesse par rapport à ses prédécesseurs, offrant ainsi de nouvelles possibilités pour la détection d'anomalies dermatologiques.

Parallèlement, l'optimisation des paramètres pour les différentes étapes de traitement constitue un domaine clé pour améliorer les performances du modèle. En ajustant finement les hyperparamètres, tels que le seuil de confiance pour la détection des objets ou les paramètres des algorithmes de prétraitement des images, il est possible d'optimiser la sensibilité et la spécificité du modèle, améliorant ainsi sa capacité à détecter et à classifier avec précision les lésions cutanées.

En outre, l'intégration de la segmentation d'image avec des modèles tels que U-Net représente une piste prometteuse pour une analyse plus approfondie des lésions cutanées. Bien que cette approche n'ait pas été explorée dans le cadre de ce projet en raison de contraintes de temps et de ressources, elle offre un potentiel considérable pour améliorer la précision de la détection en permettant une identification plus précise des contours des lésions et une meilleure séparation des tissus sains et anormaux.

En conclusion, ces perspectives ouvrent de nouvelles voies pour l'amélioration continue de notre système de détection et de classification des anomalies dermatologiques. En combinant l'exploration de nouvelles architectures de réseau, l'optimisation des paramètres et l'intégration de techniques de segmentation avancées, nous sommes en mesure de repousser les limites de la précision et de la fiabilité de notre modèle, offrant ainsi des solutions toujours plus performantes pour la prise en charge des maladies de la peau.

VIII. Bibliographie

- Ali, M. S., Miah, M. S., Haque, J., Rahman, M. M., & Islam, M. K. (2021). An enhanced technique of skin cancer classification using deep convolutional neural network with transfer learning models.
- Bill, C., Connah, K., Andrzej, B., Joanna, J.-K., & Moi Hoon, Y. (2022). Analysis of the ISIC image datasets: Usage, benchmarks and recommendations. *Elsevier*.
- Heller, N., Bussmann, E., Shah, A., Dean, J., & Papanikolopoulos, N. (2018). Computer Aided Diagnosis of Skin Lesions from Morphological Features.
- Kalaivani, & Karpagaballi. (2022). A Deep Ensemble Model for Automated Multiclass Classification Using Dermoscopy Images.
- Le, D., Le, H., Ngo, L., & Ngo, H. (2020). Transfer learning with class-weighted and focal loss function for automatic skin cancer classification.
- Mendonça, T. &. (2015). *Dermoscopy Image Analysis*. Lisbonne: CRC Press.
- Ozkan, I. A., & Koklu, M. (2017). Skin Lesion Classification using Machine Learning Algorithms.
- Philipp, T., Cliff, R., & Harald, K. (2018). The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific Data*.
- Shaheen, H. (2023). Skin lesion classification using HG-PSO and YOLOv7 based convolutional network in real time.
- Teresa, M., Pedro, F., André, R. S., & Catarina, B. (2015). PH2: A Public Database for the Analysis of Dermoscopic Images. *ResearchGate*.
- Younis, H., Bhatti, M. H., & Azeem, M. (2019). Classification of Skin Cancer Dermoscopy Images using Transfer Learning.

IX. Annexes

Etat de l'art

Etude 1 :

2017 : [714-Article_Text-2388-2-10-2019052020200107-32906-51yql9-libre.pdf \(d1wqtxts1xzle7.cloudfront.net\)](https://www.researchgate.net/publication/323882107_Skin_Lesion_Classification_using_Machine_Learning_Algorithms)

Titre : Skin Lesion Classification using Machine Learning Algorithms

Auteurs : Ilker Ali Ozkan, Murat Koklu

Date de publication : 28/12/2017

(Ozkan & Koklu, 2017)

Objectif de l'étude :

L'objectif est de pré classifier les lésions cutanées en 3 groupes (normal, anormal et mélanome) à l'aide de méthodes d'apprentissage automatique. L'étude vise également à développer un système de support à la décision qui facilitera la prise de décision pour un médecin.

Méthodologies :

Cette étude utilise la base de données PH2 constitué de 200 images provenant du service de dermatologie de l'Hôpital Pedro Hispano. Ce jeu de données se compose de 80 images normales, 80 images anormales et 40 images de mélanomes.

4 techniques de classifications sont appliquées afin de connaître le modèle le plus performant entre le Réseau de Neurones Artificiels (ANN), Machines à Vecteurs de support (SVM), K-Nearest Neighbor (KNN) et Arbre de Décision (DT). L'ANN est un système de neurones interconnectés comme le CNN, SVM est un classificateur non paramétrique, KNN est un algorithme d'apprentissage basé sur des exemples et DT est un algorithme en forme d'arbre.

Résultats :

Les résultats de performances montrent que l'ANN atteint une accuracy de 92,5%, suivi de DT à 90%, SVM à 89,5% et KNN à 82%. Ainsi l'ANN semble être plus performant que les autres algorithmes pour la classification des lésions cutanées. L'étude suggère de possibilités de progrès en explorant différentes techniques de prétraitement des données et des algorithmes de classification hybrides.

Etude 2 :

2018 : [Computer Aided Diagnosis of Skin Lesions from Morphological Features \(umn.edu\)](https://www.researchgate.net/publication/323882107_Computer_Aided_Diagnosis_of_Skin_Lesions_from_Morphological_Features)

Titre : Computer Aided Diagnosis of Skin Lesions from Morphological Features

Auteurs : Nicholas Heller, Erika Bussmann, Aneri Shah, Joshua Dean, Nikolaos Papanikolopoulos

Date de publication : 24/08/2018

(Heller, Bussmann, Shah, Dean, & Papanikolopoulos, 2018)

Objectif de l'étude :

L'objectif est de développer un système d'aide au diagnostic informatisé pour la détection des lésions cutanées basées sur leurs caractéristiques morphologiques.

Méthodologies :

La base de données utilisée est HAM10000 qui est un sous ensemble d'ISIC.

La méthode de cette étude repose sur le transfert d'apprentissage et l'augmentation des données. Comme pour l'étude précédente 5 réseaux de neurones différents ont été comparé afin d'obtenir les meilleurs résultats. Les modèles sont ResNet 50 v1, Inception v3, Xception, DenseNet 201, et InceptionResNet v2.

La procédure d'entraînement est identique pour chaque modèle. Ils sont tout d'abord préentraînés sur la base de données ImageNet, puis les données sont transférées vers une base moins complexe, ici HAM10000. Le modèle est réentraîné sur la base.

Connaissant déjà les résultats, le modèle DenseNet sera celui avec l'accuracy la plus élevée. C'est pourquoi je vais expliquer plus précisément son fonctionnement :

DenseNet est un type de réseau de neurones convolutifs (CNN) qui se caractérise par une connexion dense entre les couches. Chaque couche reçoit des entrées non seulement de la couche précédente, mais aussi de toutes les couches précédentes dans le réseau. Cela favorise le partage des caractéristiques et renforce le flux d'information à travers le réseau. Le modèle DenseNet est divisé en blocs appelés "Dense Blocks". Chaque bloc contient plusieurs couches de convolution, et les sorties de toutes ces couches sont concaténées avant d'être transmises à la couche suivante.

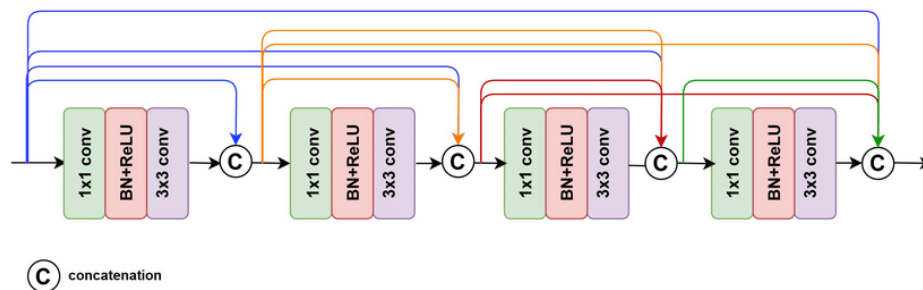


Fig. 24. Fonctionnement d'un réseau de neurones DenseNet

Résultats :

Model	Accuracy %
Inception	87.7
Xception	86.2
Inception-ResNet	87.8
ResNet	85.2
DenseNet	88.2
Ensemble	85.2

Table 6 : Comparaisons des résultats de l'étude avec des études précédentes

Etude 3 :

2019 : [Classification of Skin Cancer Dermoscopy Images using Transfer Learning | IEEE Conference Publication | IEEE Xplore](#)

Titre : Classification of Skin Cancer Dermoscopy Images using Transfer Learning

Auteurs : Haseeb Younis, Muhammad Hamza Bhatti, Muhammad Azeem

Date de publication : 02/12/2019

(Younis, Bhatti, & Azeem, 2019)

Objectif de l'étude :

Cette étude vise à développer un système plus précis pour la détection précoce du cancer de la peau, en se concentrant sur la classification d'images à l'aide d'un CNN MobileNet.

Méthodologie :

Comme l'étude précédente et toutes les autres qui vont suivre, la base de données utilisée est HAM10000.

Le modèle utilisé est MobileNet. C'est une architecture de réseau neuronal développée par Google,

L'approche de MobileNet repose sur l'utilisation de convolutions dites "depthwise separable" pour réduire le nombre de paramétrages et la charge computationnelle. Cette technique divise la convolution en 2 étapes distinctes :

- Depthwise convolution : convolution profonde de taille 3x3 qui sert à appliquer un unique filtre à chaque chanel d'entrée
- Pointwise convolution : convolution ponctuelle de taille 1x1 qui sert à combiner les sortants de la depthwise convolution.

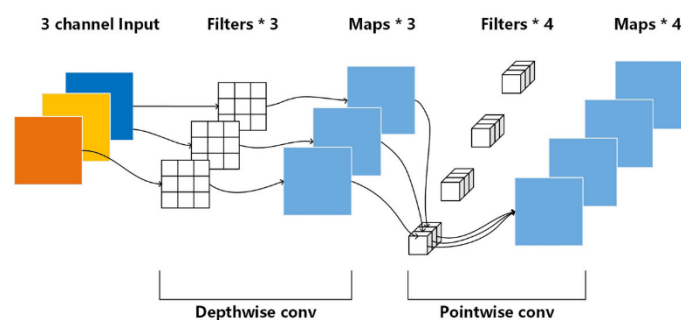


Fig. 25. Fonctionnement d'un réseau de neurones MobileNet

Résultats :

Table II: Implementation results comparison between previous research works

	Accuracy%	Recall%
Esteva[6]	72.1	96
Mohd [8]	83.33	-
Ebthial[9]	90.32	93.97
Pereira[13]	91.5	-
Ulji[7]	94.2	90.74
Anabik[12]	77.5	-
Proposed system	97.07	91.34

Table 7 : Comparaisons des résultats de l'étude avec des études précédentes

Etude 4 :

2020 : [2009.05977.pdf \(arxiv.org\)](https://arxiv.org/abs/2009.05977)

Titre : Transfer learning with class-weighted and focal loss function for automatic skin cancer classification

Auteurs : Duyen N.T. Le, Hieu X. Le, Lua T. Ngo, and Hoan T. Ngo

Date de publication : 13/09/2020

(Le, Le, Ngo, & Ngo, 2020)

Méthodologie :

L'architecture du modèle repose sur le transfert d'apprentissage, utilisant une version modifiée de l'architecture ResNet50 pré-entraînée sur la base de données ImageNet.

ResNet50 est un modèle de réseau neuronal profond qui appartient à la famille des réseaux résiduels. La caractéristique principale de ResNet50 est l'utilisation de blocs résiduels. Chaque bloc résiduel contient 2 branches :

- Une branche principale qui effectue une série de transformation linéaire (convolution, normalisation, activation) sur l'entrée
- Une branche de raccourci qui effectue une transformation linéaire légère sur l'entrée pour l'ajuster à la sortie de la branche principale.

La sortie des 2 branches est additionnée, créant une connexion résiduelle. Elle permet de faciliter l'apprentissage de fonctions complexes.

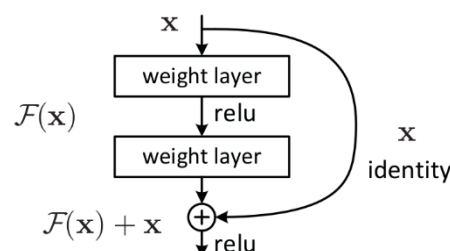


Fig. 26. Fonctionnement d'un réseau de neurones ResNet50

Dans cette étude, quelques modifications ont été effectuées incluant l'utilisation du "globale average pooling" au lieu de "l'average pooling" et le remplacement de la couche supérieure par un bloc composé d'une couche de dropout (0,5) encadrée par 2 couches entièrement connectées. Le globale average pooling est une opération qui réduit chaque features à une valeur unique en prenant la moyenne de toutes les valeurs. La couche de dropout désactive aléatoirement un pourcentage de neurones pendant l'entraînement. C'est 2 modifications servent à minimiser le surajustement.

Pour résoudre le déséquilibre des classes, une approche de poids de classe est appliquée dans la fonction de perte, attribuant des poids différents aux classes. Les poids initiaux sont basés sur les ratios entre les classes et sont ajustés pour améliorer la précision finale.

Résultats :

Le modèle a atteint une accuracy moyenne de 93%, surpassant la précision des dermatologues qui était de 84%. De plus, en comparant le modèle ResNet50 avec les modèles VGC16, MobileNet et EfficientNetB1, les résultats ont montré que le modèle est le plus performant. Ainsi, ces résultats peuvent de référence pour des études futures, aidant les scientifiques à revoir les stratégies d'entraînement et à accélérer le processus d'apprentissage.

Etude 5 :

2021 : [An enhanced technique of skin cancer classification using deep convolutional neural network with transfer learning models - ScienceDirect](#)

Titre : An enhanced technique of skin cancer classification using deep convolutional neural network with transfer learning models

Auteurs : Md Shahin Ali, Md Sipon Miah, Jahurul Haque, Md Mahbubur Rahman, Md Khairul Islam

Date de publication : 15/09/2021

(Ali, Miah, Haque, Rahman, & Islam, 2021)

Objectif de l'étude :

L'article propose un modèle de réseau neuronal convolutionnel profond (DCNN) qui classifie le cancer de la peau avec plus de précision, même si les patients sont à un stade précoce.

Méthodologie :

L'architecture de ce modèle repose également sur le transfert d'apprentissage, utilisant une version modifiée de l'architecture d'un DCNN pré-entraînée sur la base de données ImageNet.

Voici les paramètres initialisés pour ce modèle :

- batch size = 128 : cela correspond au nombre d'images traitées à chaque itération.
- Learning rate = 0,001 : il détermine le taux avec lequel débutera la procédure d'apprentissage
- Optimiseur (Adam) : il minimise la fonction de perte
- Fonction de perte (binary cross-entropy)

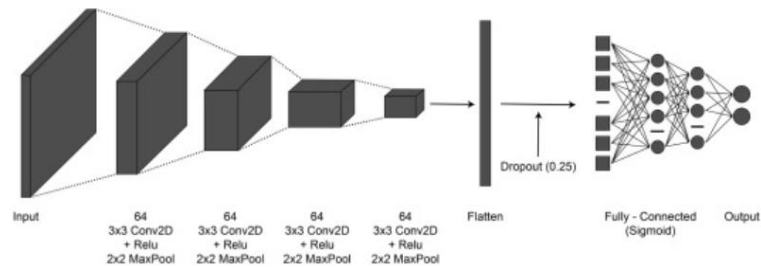


Fig. 27. Fonctionnement du réseau de neurones profond de l'étude

Résultats :

Model	Precision	Recall	F1 score	Training Acc	Testing Acc
AlexNet (Han et al., 2017)	97.88	92.01	94.85	93.82	90.86
ResNet (Targ et al., 2016)	96.00	93.63	94.80	91.25	90.93
VGG-16 (Guan et al., 2019)	1.00	86.02	92.48	86.17	86.02
DenseNet (Carcagni et al., 2019)	91.42	94.83	93.09	91.64	88.33
MobileNet (Sinha & El-Sharkawy, 2019)	92.08	94.15	93.10	90.31	88.26
Proposed DCNN model	96.57	93.66	95.09	93.16	91.43

Table 8 : Comparaisons des résultats de l'étude avec des études précédentes

Etude 6 :

2022 : [A Deep Ensemble Model for Automated Multiclass Classification Using Dermoscopy Images | IEEE Conference Publication | IEEE Xplore](#)

Titre : A Deep Ensemble Model for Automated Multiclass Classification Using Dermoscopy Images

Auteurs : A. Kalaivani, S. Karpagavalli

Date de publication : 13/04/2022

(Kalaivani & Karpagaballi, 2022)

Méthodologie :

La méthode utilisée repose sur l'approche Bootstrapping Ensembles based Convolutional Neural Network (BE-CNN). Le processus commence par traiter l'image d'entrée à l'aide de la structure des réseaux de neurones convolutionnels basée sur un réseau de segmentation intensif en calcul (CI-SN). Cela produit des caractéristiques intermédiaires et des prédictions de segmentation. Ensuite, ces prédictions segmentées sont combinées pour obtenir la prédiction finale.

Le modèle utilise des blocs de bootstrapping et de convolution pour créer des cartes de prédiction à différentes échelles. Ces blocs effectuent des opérations courantes pour rendre le modèle flexible. Le bootstrapping est une technique qui permet de générer plusieurs sous-ensembles d'entraînement, chacun contenant des exemples d'images différents. Il vise ainsi à réduire le surapprentissage.

Résultats :

Percentage (%) of performance measures	Standard CNN Model	BE-CNN Model
Precision	89.3	91.12
Re-call	89.21	92.24
F-measure	91.28	93.13
Accuracy	91.68	93.8

Table 9 : Résultats de l'étude

Etude 7 :

2023 :[Skin lesion classification using HG-PSO and YOLOv7 based convolutional network in real time - Hera Shaheen, Maheshwari Prasad Singh, 2023 \(sagepub.com\)](#)

Titre : Skin lesion classification using HG-PSO and YOLOv7 based convolutional network in real time

Auteur : Hera Shaheen

Date de publication : 15/10/2023

(Shaheen, 2023)

Méthodologie :

L'étude présente une méthode pour l'identification du mélanome, utilisant une approche combinant l'optimisation génétique et la recherche de particule HG-PSO, ainsi qu'un réseau de convolution (CNN) basé sur YOLOv7 pour la classification du cancer de la peau.

HG-PSO (Hybrid of Genetic and Particle Swarm Optimization) est une méthode qui combine 2 optimisations :

- L'optimisation génétique est une technique qui tente de trouver la meilleure solution à un problème en utilisant des opérations génétiques telles que la sélection, la recombinaison et la mutation.
- L'optimisation par essaims de particules est une méthode basée sur le comportement d'un ensemble de particules ou solutions potentielles, qui évoluent dans un espace de recherche en fonction de leur propre expérience et celle de leurs voisins.

YOLOv7 (You Only Look Once version 7) est une famille de modèles de réseau de neurones convolutifs utilisées pour la détection d'objet en temps réel dans des images ou des vidéos. Il se caractérise par sa capacité à effectuer la détection et la classification des objets en une seule passe, en prédisant directement les classes des objets détectés.

Dans l'étude, YOLOv7 est utilisé pour détecter les régions infectées de mélanome dans les images et HG-PSO est utilisé pour optimiser les paramètres du modèle YOLOv7.

La méthode commence par localiser la région infectée grâce à YOLOv7, puis un seuil de couleur est appliqué pour segmenter cette région, et enfin cela passe par le réseau de convolution.

Résultats :

Les tests ont été effectués sur 3 ensembles de données différents dont HAM10000 dataset. Les scores obtenus sont :

Accuracy : 98.86%, average precision : 99.00%, average recall : 98.85%, F1-score : 98.85%.

(Ozkan & Koklu, 2017)

Voici un résumé des études précédentes (*Table 8*) :

Date	Base de données	Modèle	accuracy	precision	recall	F1-score
2017	PH2	ANN	92,50%	92,38%	/	90,45%
2018	HAM10000	DenseNet	88,2%	78,67%	84,61%	/
2019	HAM10000	MobileNet (CNN)	97%	90%	91%	/
2020	HAM10000	ResNet50 modifié	93%	88%	86%	84%
2021	HAM10000	DCNN	91,43%	96,57%	93,66%	93,10%
2022	HAM10000	BE-CNN	93,8%	91,12%	92,24%	93,13%
2023	HAM10000	Yolov7 (CNN)	98,86%	99%	98,85%	99,85%

Table 10 : Récapitulatif des modèles et de leurs performances

[Data augmentation for improving deep learning in image classification problem | IEEE Conference Publication | IEEE Xplore](#)

[On Data Augmentation for GAN Training | IEEE Journals & Magazine | IEEE Xplore](#)

[A survey on Image Data Augmentation for Deep Learning \(springeropen.com\)](#)

[Using GANs for Data Augmentation | Baeldung on Computer Science](#)