

RAPPORT DU MINI PROJET

Diagnostic des maladies de peau avec le machine Learning :



Réalisé par :

- BASMA AITHOUNA
- LEILA BAHATI
- TAHA BENBELLA
- REDA HADDAN
- RAWAHA ELHOUSAYNI

ENCADRÉ PAR:

MR MOHAMMED KHALIL

Remerciements

En premier lieu, on tient à remercier notre professeur encadrant, MR. KHALIL MOHAMMED, pour nous avoir donner cette démarche de réalisation d'un projet de qualité et nous avoir mis à disposition le temps nécessaire à sa réalisation.

On remercie particulièrement toute l'équipe qui a développé cette application, et notamment rendu cette formation très sympathique.

Sommaire

CHAPITRE 1

I	Intr	roduction:	5
	1.1	Le psoriasis :	5
	1.2	L'acné :	5
	1.3	L'eczéma :	5
	1.4	La cellulite :	6
	1.5	Le mélanome :	6
	1.6	L'herpès :	6
L'o	bjecti	if :	6
2	Ana	alyse médicale :	7
,	2.1	Qu'est-ce que la dermatologie ?	7
2	2.2	Que font les dermatologues ?	7
2	2.3	Tests diagnostiques des maladies de la peau :	7
,	2.4	Instruments utilisés en dermatologie :	8
3	Ana	alyse automatique :	9
	3. I	Qu'est-ce que l'Apprentissage automatique ?	9
	3.2	Le diagnostic assisté par ordinateur	9
	3.3	Pourquoi utilise-t-on une machine pour le diagnostic ?	
4	La c	comparaison algorithme / médecin :	10
		CHAPITRE 2	
5	Syst	tème globale :	12
6	Des	scripteurs utilisés :	13
(6. I	Descripteur de couleur :	14
(6.2	Descripteur de texture :	14
7	Mét	thodes proposées :	15
-	7. I	Réseau neuronal convolutif CNN :	15
-	7.2	La matrice de cooccurrence de niveau de gris (GLCM) :	17
-	7.3	Les K plus proches voisins KNN :	18
		CHAPITRE 3	
8	La t	base de données :	21
9	Mes	sure d'évaluation :	23
10	R	Résultat final :	25
П	C	Conclusion :	27
12	В	Bibliographies :	28

CHAPITRE 1:

I Introduction:

La peau désigne les couches successives de tissus. A cause d'un virus, d'un champignon, d'une bactérie, d'un allergène, d'un dérèglement de ces tissus ou des gènes... La peau peut souffrir et faire souffrir parce que ces maladies sont largement taboues dans la société.

I.I Le psoriasis:

Avec près de 60 000 nouveaux cas chaque année, le psoriasis est une des affections dermatologiques les plus fréquentes. C'est une maladie inflammatoire de la peau qui peut toucher différentes parties du corps (coude, genoux, cuir chevelu...). Non contagieuse et le plus souvent bénigne, cette maladie se caractérise par l'apparition d'épaisses plaques de peau qui se détachent sous forme d'écailles blanches.



I.2 L'acné:

L'acné est une maladie de la peau caractérisée par l'éruption de boutons rouges, de points

noirs et de kystes. Elle touche le plus souvent le visage et le thorax, mais aussi le dos, le torse voire le cuir chevelu. Elle est très fréquente chez les adolescents : on estime que de 80 % à 90 % d'entre eux en sont atteints à un moment ou à un autre. Les garçons sont plus touchés que les filles. 10 % à 20 % d'entre eux consultent un médecin pour cette raison. Dans la majorité des cas, l'acné disparaît vers l'âge de 20 ou 30 ans, mais il arrive qu'elle persiste, on parle alors d'acné tardive.



1.3 L'eczéma:

Parmi toutes les maladies de la peau, l'eczéma est la plus fréquente : cette maladie motive

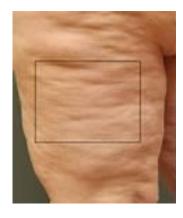
jusqu'à 30 % des consultations en dermatologie. Dans les pays industrialisés, elle atteindrait de 15 % à 30% des enfants et de 2 % à 10 % des adultes. Selon de récentes estimations, les cas d'eczéma auraient doublé et peut-être même triplé depuis les 30 dernières années. L'eczéma est une dermatose prurigineuse caractérisée par une inflammation non contagieuse de la peau qui s'accompagne de rougeurs, de fines vésicules, de squames et de démangeaisons. Il peut



commencer très tôt dans la vie, et s'observe même chez les nourrissons. Les personnes atteintes connaissent des périodes communément appelées « poussées d'eczéma », durant lesquelles les symptômes s'aggravent. Ces poussées, de durée variable, sont entrecoupées de périodes de rémission. L'eczéma est souvent associé à l'asthme ou à diverses réactions allergiques.

I.4 La cellulite:

La cellulite touche presque exclusivement les femmes, chez qui les médecins la considèrent comme un phénomène physiologique normal. Près de 9 femmes sur 10 en sont touchées à un moment ou l'autre de leur vie, pour l'homme sur 50. La cellulite est le résultat d'un changement de structure des tissus adipeux (réserves de gras) logés sous l'épiderme. Elle donne à la peau une apparence « bosselée », jugée disgracieuse. On l'observe surtout à l'arrière des cuisses et sur les fesses.



1.5 Le mélanome :

Le mélanome est un cancer de la peau, c'est-à-dire une tumeur maligne se développant à partir de cellules cutanées. Il s'agit dans ce cas des cellules fabriquant la mélanine (pigment donnant la couleur à la peau, aux poils et aux cheveux) : les mélanocytes.



1.6 L'herpès:

L'herpès génital est une infection transmissible sexuellement (ITS) caractérisée par l'apparition de petites vésicules douloureuses sur les organes sexuels. Ces vésicules sont transparentes et remplies de liquide. La majorité des infections surviennent à l'âge adulte, vers



20 ans à 40 ans. Au Canada, environ une personne sur 5 est atteinte d'herpès génital. En Europe, l'herpès génital est nettement moins répandu. Par contre, il est très commun en Afrique, où de 30 % à 40 % des gens sont infectés et même jusqu'à 90 % dans certaines populations.

L'objectif:

L'objectif principal de ce mini-projet est de développé une application qui permettra à son utilisateur de déterminer s'il souffre d'une des 6 maladies cidessus qui sont les plus fréquentes (le Psoriasis, l'Acné, l'Eczéma, la Cellulite, l'Herpes ou la Mélanome).

2 Analyse médicale:

2.1 Qu'est-ce que la dermatologie?

Le mot racine de la dermatologie est derme ou dermis. Cela vient du mot grec derma, qui signifie peau.

La dermatologie est un domaine spécial de la médecine qui se concentre sur les conditions qui affectent votre peau. En plus de votre peau, la dermatologie comprend également des affections qui affectent vos ongles, vos cheveux et la délicate muqueuse de vos paupières, de votre nez et de votre bouche.

1. Le médecin spécialisé pratiquant la dermatologie s'appelle le dermatologue ou le dermatologiste.

2.2 Que font les dermatologues ?

Le travail d'un dermatologue peut généralement être divisé en trois domaines. Ceux-ci inclus :

- Médical: Un dermatologue peut diagnostiquer et traiter les conditions qui affectent votre peau. Cela inclut des conditions telles que le psoriasis, la rosacée et l'acné. Un dermatologue peut également identifier des symptômes sur votre peau qui pourraient être des signes d'autres problèmes de santé par exemple : des démangeaisons et une peau sèche peuvent parfois indiquer des problèmes de reins.
- Chirurgie : De nombreux dermatologues pratiquent une chirurgie mineure, comme enlever les grains de beauté ou les verrues ou faire des biopsies cutanées. D'autres dermatologues se spécialiseront dans la chirurgie plus étendue.
- Cosmétique: Un dermatologue peut traiter les problèmes de peau qui affectent votre apparence. Cela peut inclure la perte de cheveux, des taches brunes ou des rides. De nombreux dermatologues sont également formés pour administrer des traitements cosmétiques. Ceux-ci incluent les charges, les peelings chimiques et l'épilation au laser.

2.3 Tests diagnostiques des maladies de la peau :

Divers tests cutanés peuvent être effectués pour diagnostiquer les allergies cutanées, les infections cutanées bactériennes, virales ou fongiques et d'autres problèmes affectant la peau. Des tests cutanés sont également effectués pour faire la différence entre les cellules malignes (cancéreuses) et les excroissances bénignes (non cancéreuses).

Les tests cutanés les plus courants comprennent :

- Test de patch : Les tests de patch sont utilisés pour aider à diagnostiquer les allergies cutanées. Les allergènes identifiés (substances auxquelles une personne peut être allergique) sont appliqués sur la peau du dos avec des patchs adhésifs et laissés pendant un certain temps. La peau est ensuite examinée pour toute réaction.
- Biopsie cutanée: Des biopsies cutanées sont effectuées pour diagnostiquer un cancer de la peau ou des troubles cutanés bénins. Lors d'une biopsie cutanée, la peau est prélevée (après l'application d'une anesthésie locale) et est emmenée dans un laboratoire pour analyse. La peau peut être retirée avec un scalpel, une lame bleue Gillette ou un outil de biopsie à poinçon cylindrique. Des points de suture peuvent être utilisés pour fermer la plaie.
- Culture : Une culture est un test qui est effectué pour identifier le microorganisme (bactérie, champignon ou virus) qui cause une infection. La peau (raclures de surface, biopsies, contenu des bosses de pus et des cloques), les cheveux ou les ongles peuvent être cultivés pour détecter des bactéries, des champignons ou des virus.

2.4 Instruments utilisés en dermatologie :

Voici quelques instruments utilisés par les dermatologistes :



DERMATOSCOPE

BOIS LUMIÈRE

POINÇONS DE BIOPSIE

DERMIQUE

3 Analyse automatique:

3.1 Qu'est-ce que l'Apprentissage automatique ?

L'apprentissage automatique (de l'anglais Machine learning ou ML), considéré comme étant un champ de l'intelligence artificielle, est un ensemble d'algorithmes destinés à s'améliorer au fil des expériences.

Ce champ de l'informatique trouve son usage dans plusieurs domaines dans lesquels une programmation explicite peut être difficile ou induire la machine dans la confusion, par exemple pour faire de la vision par ordinateur ou dans le domaine médical.

Cet ensemble de techniques permettent de construire des modèles en fournissant des données représentatives (données d'entrainement ou training data) en vue d'effectuer des prédictions sans explicitement programmer la machine à les faire.

3.2 Le diagnostic assisté par ordinateur

On se focaliseras dans ce rapport sur une application bien précise de l'apprentissage automatique qui est le diagnostic assisté par une machine.

Le diagnostic assisté par ordinateur est un ensemble de systèmes destinés à assister les soignants dans le processus de diagnostic, l'interprétation et la prise de décision à la base d'imagerie médicale (IRM, rayons X, ...) en assurant un compromis entre la rapidité et l'exactitude des informations.

3.3 Pourquoi utilise-t-on une machine pour le diagnostic?

Selon des recherches effectuées en 2018, il a été remarqué que les algorithmes de machine Learning étaient plus performantes en termes de détection des maladies de peau qu'un ensemble de médecins qualifiés (environ 50% de dermatologues certifiés et 50% de dermatologues et médecins divers).

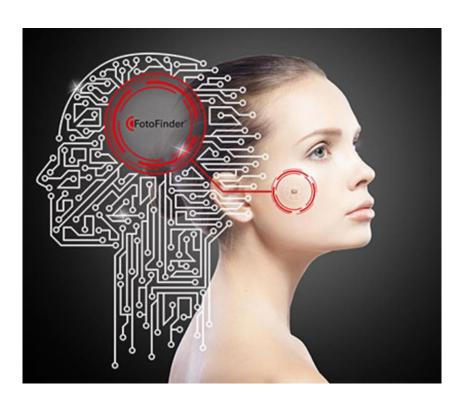
On constate donc une meilleure performance des machines par rapport à des professionnels moyennement expérimentés en termes de détection des lésions de la peau. Il faudra cependant noter que les algorithmes d'apprentissage automatique sont moins performants lorsqu'on est face à de nouveaux cas non traités lors de la phase d'apprentissage.

4 La comparaison algorithme / médecin :

Les progrès sont remarquables dans l'utilisation de l'intelligence artificielle (IA) en médecine. Dans ce cas de figure, soit la reconnaissance d'une affection dermatologique, l'IA apporte « des résultats comparables à ceux apportés par les médecins dermatologues », écrit l'auteur principal, le Dr Jung-Im Na, du Service de dermatologie de l'Université nationale de Séoul. Cependant, pour que ces systèmes soient sensibles, précis et finalement utiles, leurs performances doivent être testées dans un environnement similaire à la pratique réelle.

Les performances de l'algorithme ont été comparées à celles de 21 dermatologues, 26 étudiants en dermatologie et 23 usagers de santé non médecins. Les performances de l'algorithme apparaissent similaires à celles des étudiants en dermatologie mais légèrement inférieures à celles des dermatologues. Néanmoins, l'IA permet d'accroître de 77,4% à 86,8% la sensibilité du diagnostic de malignité des cliniciens et de 47,6% à 87,5% la sensibilité du diagnostic de malignité du « grand public ».

Des résultats qui pourraient permettre une habilitation des médecins généralistes en dermatologie diagnostique de première ligne, suggèrent les chercheurs. L'IA doit soutenir les humains pour permettre des diagnostics plus rapides et plus précis.



CHAPITRE 2:

5 Système globale:

Descripteur

- descripteur de couleur
- descripteur de texture

Methode de classification

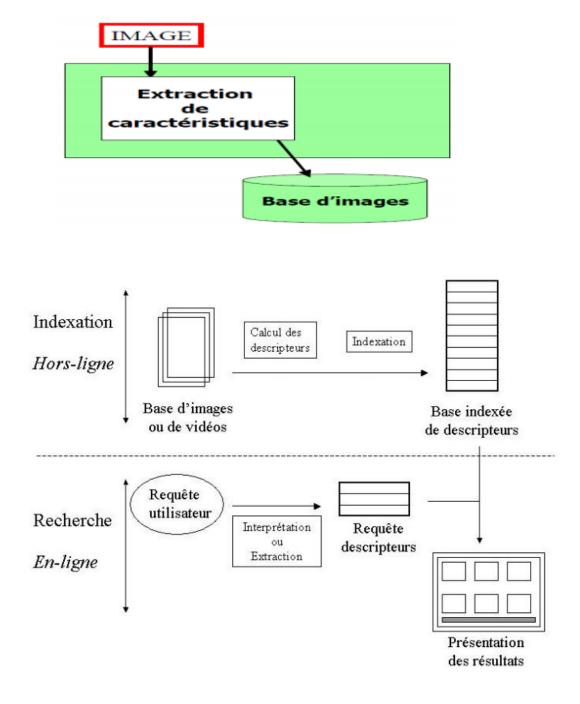
- CNN
- GLCM
- KNN

Systemes proposés

• une application web developpée par html et css et python

6 Descripteurs utilisés :

En vision par ordinateur, l'extraction de caractéristiques visuelles (ou visual features extraction en anglais) consiste en des transformations mathématiques calculées sur les pixels d'une image numérique. Les caractéristiques visuelles permettent généralement de mieux rendre compte de certaines propriétés visuelles de l'image, utilisées pour des traitements ultérieurs entrant dans le cadre d'applications telles que la détection d'objets ou la recherche d'images par le contenu. A partir des années 2010, la grande majorité des travaux en vision par ordinateur repose sur des caractéristiques qui sont apprises par des réseaux convolutifs sur de grandes bases d'images, conformément aux principes de l'apprentissage profond.



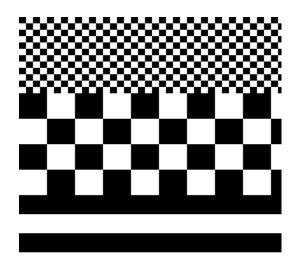
6.1 Descripteur de couleur :

La prise en compte de la couleur des images a été historiquement l'une des premières caractéristiques employées pour la recherche d'images par le contenu et produit encore des résultats parfois spectaculaires sur certaines bases. C'est en 1991 que Swain et Ballard ont proposé d'utiliser un histogramme couleur pour ce genre de tâche. Il représente la distribution globale des couleurs dans l'image. Son calcul consiste en une quantification de l'espace couleur choisi (RVB par exemple) suivie du calcul de l'histogramme des pixels ainsi transformés. Par exemple, si l'on considère une image RVB classique et que l'on quantifie chaque plan couleur sur 4 bins, l'histogramme résultant aura une dimension 4³=64. Si l'image est originellement codée sur 24 bits (la valeur de chaque plan est dans l'intervalle [0-255]), l'histogramme couleur sur 64 bins pourrait être représenté par un « cube » de 4*4*4.

			Rouge														
0-63			64-127 128-191 192-2			92-255	-255										
	Vert					Vert		Vert			Vert						
		0- 63	64- 127	128- 191	192- 255	0- 63	64- 127	128- 191	192- 255	0- 63	64- 127	128- 191	192- 255	0- 63	64- 127	128- 191	192- 255
	0-63		Х	х	х	Х	Х	х	х	Х	Х	х	х	Х	х	х	х
	64-127	X	Х	х	х	Х	Х	х	х	х	х	х	х	Х	х	х	х
Bleu	128- 191	х	х	х	х	х	х	х	х	х	×	х	х	х	x	х	х
	192- 255	х	х	x	х	х	х	х	х	х	х	х	х	х	x	х	х

6.2 Descripteur de texture :

Dans le traitement d'image, l'analyse de **texture** consiste à calculer une série de mesures dans le but de définir une texture perçue sur une image. L'analyse de texture renvoie des informations sur l'arrangement spatial des couleurs ou des intensités dans tout ou partie de cette image. Les données de texture d'une image peuvent être artificiellement créées (textures artificielles) ou résulter de l'analyse d'images filmées à partir de scènes ou d'objets réels (textures naturelles). L'analyse (ou caractérisation) de texture joue un rôle important dans la classification d'image. Les éléments apportant le plus de précisions dans la segmentation sont les fréquences spatiales et la moyenne du niveau de gris.



Textures artificielles



Textures naturelles

7 Méthodes proposées :

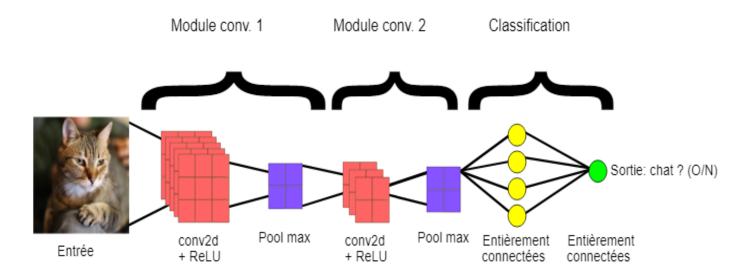
7.1 Réseau neuronal convolutif CNN:

Les réseaux de neurones convolutifs sont à ce jour les modèles les plus performants pour classer des images. Désignés par l'acronyme CNN, de l'anglais Convolutional Neural Network, ils comportent deux parties bien distinctes. En entrée, une image est fournie sous la forme d'une matrice de pixels. Elle a 2 dimensions pour une image en niveaux de gris. La couleur est représentée par une troisième dimension, de profondeur 3 pour représenter les couleurs fondamentales [Rouge, Vert, Bleu].

La première partie d'un CNN est la partie convolutive à proprement parler. Elle fonctionne comme un extracteur de caractéristiques des images. Une image est passée à travers une succession de filtres, ou noyaux de convolution, créant de nouvelles images appelées cartes de convolutions. Certains filtres intermédiaires réduisent la résolution de l'image par une opération de maximum local. Au final, les cartes de convolutions sont mises à plat et concaténées en un vecteur de caractéristiques, appelé code CNN.

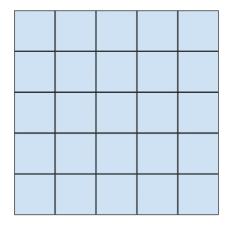
Ce code CNN en sortie de la partie convolutive est ensuite branché en entrée d'une deuxième partie, constituée de couches entièrement connectées (perceptron multicouche). Le rôle de cette partie est de combiner les caractéristiques du code CNN pour classer l'image.

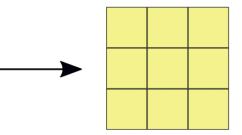
La sortie est une dernière couche comportant un neurone par catégorie. Les valeurs numériques obtenues sont généralement <u>normalisées</u> entre 0 et 1, de somme 1, pour produire une distribution de probabilité sur les catégories.



Carte de caractéristiques d'entrée

Carte de caractéristiques de sortie





Carte de caractéristiques d'entrée

s	Filtre convolutif

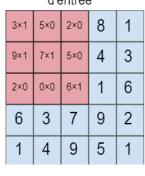
3	5	2	8	1
9	7	5	4	3
2	0	6	1	6
6	3	7	9	2
1	4	9	5	1

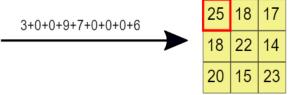
1	0	0
1	1	0
0	0	1



Carte de caractéristiques d'entrée

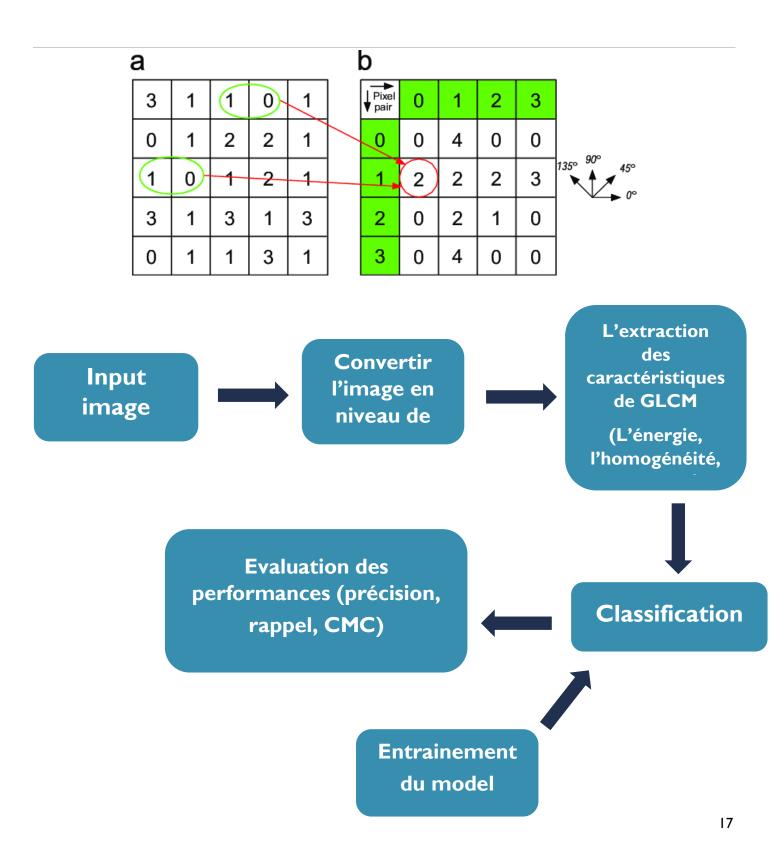
Carte de caractéristiques de sortie





7.2 La matrice de cooccurrence de niveau de gris (GLCM) :

Une matrice de cooccurrence ou distribution de cooccurrence (également appelée: matrices de cooccurrence de niveau de gris GLCM) est une matrice qui est définie sur une image comme étant la distribution des valeurs de pixel cooccurrentes (valeurs d'échelle de gris ou couleurs) à un décalage donné. Il est utilisé comme une approche d'analyse de texture avec diverses applications, en particulier dans l'analyse d'images médicales.



7.3 Les K plus proches voisins KNN:

En intelligence artificielle, plus précisément en apprentissage automatique, la méthode des k plus proches voisins est une méthode d'apprentissage supervisé. En abrégé k-NN ou KNN, de l'anglais k-nearest neighbors.

Dans ce cadre, on dispose d'une base de données d'apprentissage constituée de N couples « entrée-sortie ». Pour estimer la sortie associée à une nouvelle entrée x, la méthode des k plus proches voisins consiste à prendre en compte (de façon identique) les k échantillons d'apprentissage dont l'entrée est la plus proche de la nouvelle entrée x, selon une distance à définir. Puisque cet algorithme est basé sur la distance, la normalisation peut améliorer sa précision.

L'algorithme des k plus proches voisins (k-NN) est une méthode non paramétrique utilisée pour la classification et la régression. Dans les deux cas, il s'agit de classer l'entrée dans la catégorie à laquelle appartient les k plus proches voisins dans l'espace des caractéristiques identifiées par apprentissage. Le résultat dépend si l'algorithme est utilisé à des fins de classification ou de régression :

- En classification k-NN, le résultat est une classe d'appartenance. Un objet d'entrée est classifié selon le résultat majoritaire des statistiques de classes d'appartenance de ses k plus proches voisins, (k est un nombre entier positif généralement petit). Si k = I, alors l'objet est affecté à la classe d'appartenance de son proche voisin.
- En régression k-NN, le résultat est la valeur pour cet objet. Cette valeur est la moyenne des valeurs des k plus proches voisins.

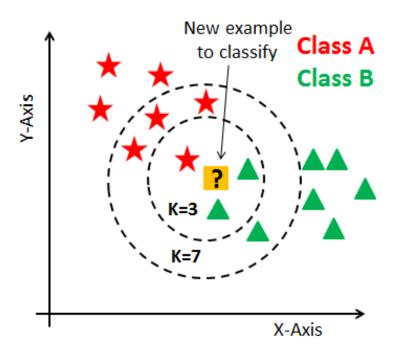


Image requête



Conversion en niveau de gris et calcule de l'histogramme



Calcule des distances entre les histogrammes de l'image requête et de chaque image de la base de donné en utilisant la distance euclidienne



Stocker les distances trouver dans une liste et la trier par ordre croissant (les images les plus proche ont les plus petits distances)



Localiser les K plus proches images à l'image requête

CHAPITRE 3:

8 La base de données :

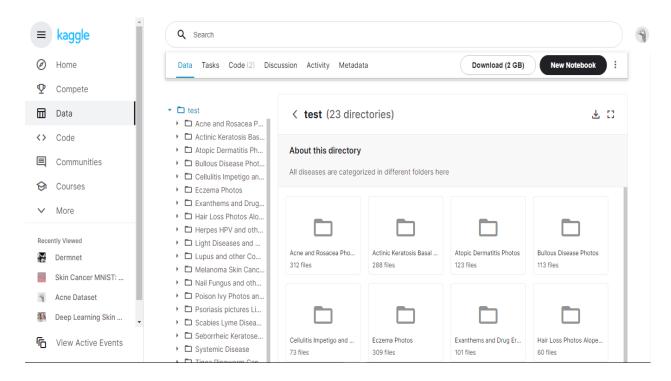
Notre programme est dédié pour détecter les six maladies de peau les plus fréquentes.

On a donc besoin d'une Database d'images classifiées par maladie de peau comme suite :

Nom	Exemple	Nombre d'images	Nombre d'image d'entrainement	Nombre d'image de test	Nombre d'image de validation
Acné		1154	841	313	313
Cellulites		363	289	74	74
Eczéma		1544	1235	309	309
Herpes		507	405	102	309
Mélanome (Skin cancer)		581	464	117	117
Psoriasis		1757	1405	352	352

La Database dont nous avons besoins se trouve au niveau du lien ci-dessous

https://www.kaggle.com/shubhamgoel27/dermnet



En téléchargeant cette database et en supprimant les données non nécessaire...on a fini par la création de notre base de données avec succès.

Problématique:

Notre base de données contient 6 maladies mais parmi eux seulement 3 classes (acné, mélanome, cellulites) donnent des résultats pertinents cela nous a poussé à réaliser une application qui détecte seulement l'acné, la mélanome et la cellulite.

9 Mesure d'évaluation :

Méthodes	Précision	Rappel	Accurcy
CNN	82,86%	82,12%	88,02%
GLCM			
KNN			

Avec:

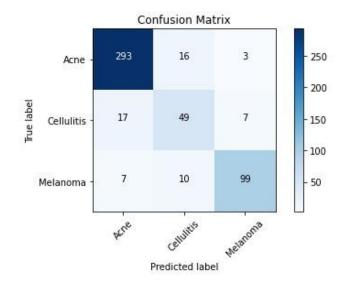
$$Pr\acute{e}cision = rac{VP}{VP + FP}$$
 $Rappel = rac{VP}{VP + FN}$

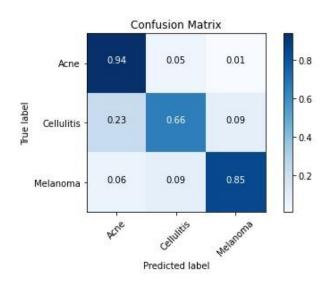
	Pertinents	Non pertinents
Récupérés	Vrais positifs (VP)	Faux positifs (FP)
Non récupérés	Faux négatifs (FN)	Vrais négatifs (VN)

Remarque:

Pour les deux méthodes GLCM et KNN on n'a pas obtenue les mesures d'évaluations adéquates.

Matrice de confusion par la méthode CNN pour les 3 maladies :

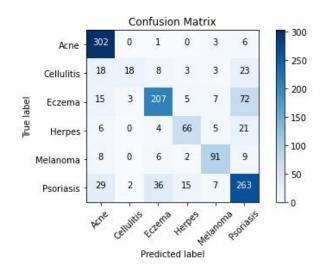


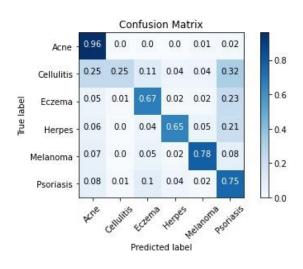


Matrice de confusion

Matrice de confusion

Matrice de confusion pour les 6 maladies :



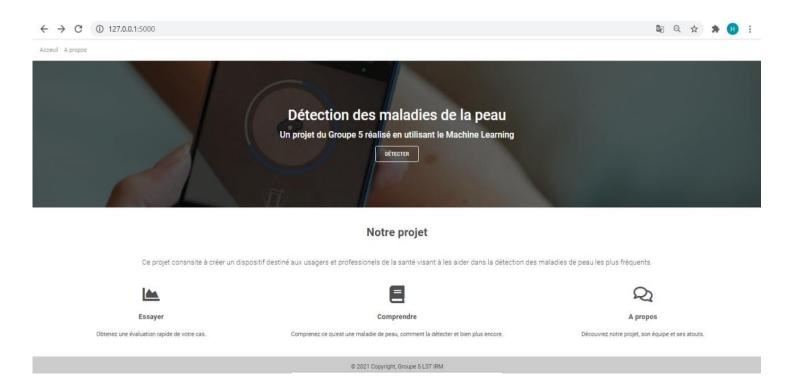


Matrice de confusion

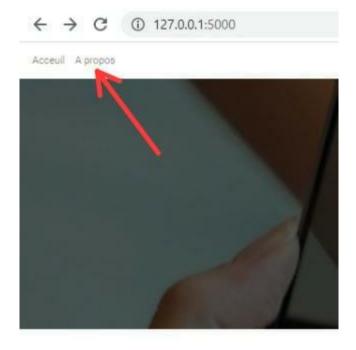
Matrice de confusion

10 Résultat final :

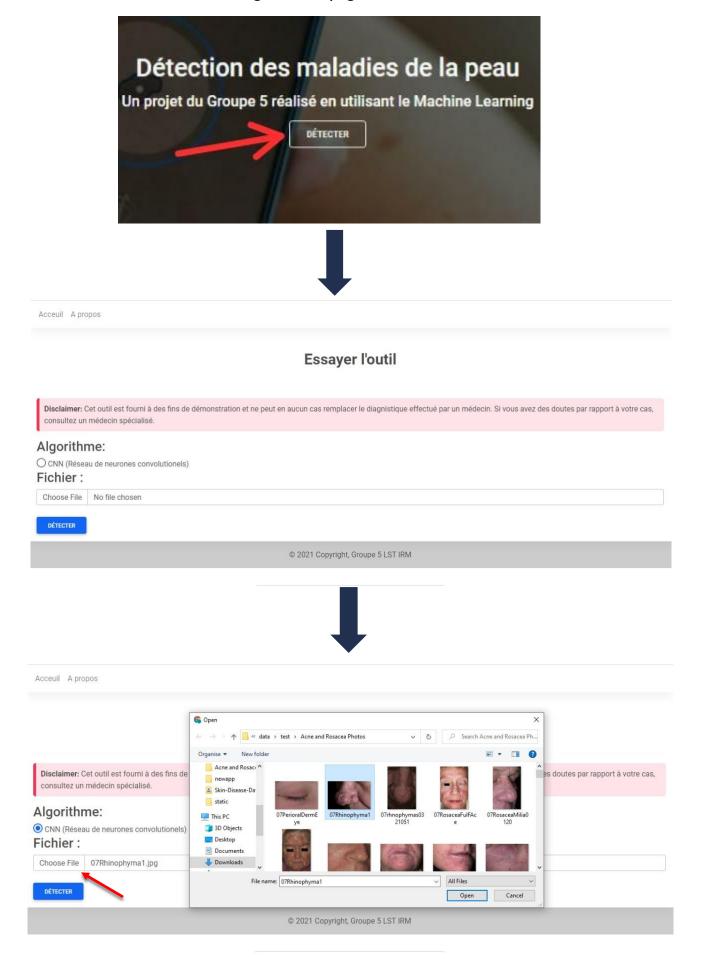
Finalement nous avons pu créer une application web dont la page d'accueil est la suivante :



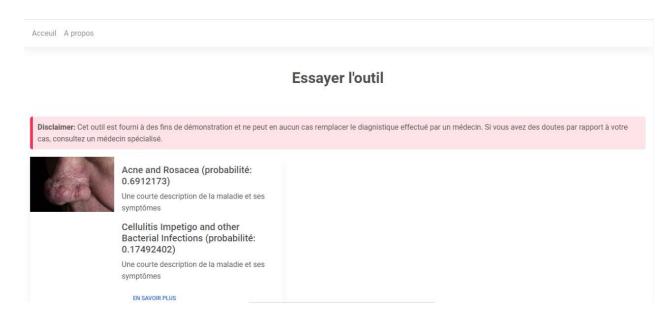
Le boutons « A propos » vous dirige vers une page qui présente les membres de notre groupe :



Le bouton « Détecter » vous dirige vers la page d'exécution :



Et voilà le résultat que donne notre projet :



II Conclusion:

Dans le parcours de soin, la phase de diagnostic est essentielle pour l'orientation du patient et son suivi. Le machine learning apporte de nouvelles solutions aux professionnels de santé pour gagner du temps et optimiser le bon diagnostic. Il ouvre de nouvelles perspectives dans le repérage des maladies.

Ce projet nous a permis de découvrir ce monde assez riche et large, et construire une idée générale et une vision propre sur ce métier.

12 **Bibliographies:**

https://www.doctissimo.fr/html/sante/mag 2002/sem02/mag0712/dossier/cellulite niv2.htm

https://www.passeportsante.net/fr/Maux/Problemes/Fiche.aspx?doc=eczema_pm

https://www.doctissimo.fr/html/dossiers/psoriasis.htm

https://www.passeportsante.net/fr/Maux/Problemes/Fiche.aspx?doc=melanome

https://www.passeportsante.net/fr/Maux/Problemes/Fiche.aspx?doc=herpes_genital_pm

https://www.passeportsante.net/fr/Maux/Problemes/Fiche.aspx?doc=cellulite_pm

https://www.passeportsante.net/fr/Maux/Problemes/Fiche.aspx?doc=acne_pm

https://github.com/bnsreenu/python for microscopists/blob/62f78fd433ee67146daaaadf8a223 d572d0654f1/200 image classification using GLCM.py

https://deeplizard.com/learn/video/OO4HD-IwRN8

https://deeplizard.com/learn/video/FNqp4ZY0wDY

https://deeplizard.com/learn/video/Zrt76Albeh4