



**République Algérienne Démocratique et populaire**

Ministère de l'Enseignement Supérieur et  
de la Recherche Scientifique

**Université des Sciences et de la Technologie**

**Houari Boumediene**

Faculté d'Electronique et Informatique

Département Informatique

**Thème du projet :**

---

**Classification des lésions cutanées  
Par convolution des réseaux de  
neurones**

---

Réalisé par :

➤ MERBOUTI Souad

Matricule : 171731076879

**Section : M2 IV**

## Table des matières

<b>Introduction générale</b> .....	3
<b>Problématique</b> .....	3
<b>Objectif</b> .....	4
<b>1. Etude préliminaire</b> .....	5
<b>1.1 Epidémiologie</b> .....	5
<b>1.2 Diagnostique</b> .....	5
<b>1.3 Formes</b> .....	6
<b>1.4 Pronostic de survie</b> .....	7
<b>1.5 Travaux précédents :</b> .....	7
<b>2. Outils de réalisation :</b> .....	7
<b>2.1 CNN</b> .....	7
<b>2.2 TSM12</b> .....	8
<b>2.3 InceptionV3</b> .....	8
<b>2.4 ResNet152V2</b> .....	8
<b>2.5 VGG16</b> .....	8
<b>2.6 Métriques de comparaison</b> .....	8
<b>2.7 La base de données</b> .....	9
<b>3. Méthodologie de réalisation</b> .....	14
<b>3.1 Prétraitement</b> .....	14
<b>3.2 Data splitting</b> .....	14
<b>3.3 Normalisation</b> .....	14
<b>3.4 Data augmentation</b> .....	14
<b>3.5 Modèles</b> .....	15
<b>4. Résultats</b> .....	18
<b>5. Conclusion</b> .....	22
<b>6. Références</b> .....	23

## Introduction générale

La peau est l'organe le plus volumineux dont nous soyons dotés. C'est elle qui fait barrière entre nos organes internes et l'environnement extérieur, nous protégeant ainsi des blessures et des rayonnements solaires. Cependant, il arrive qu'elle tombe malade, notamment lorsqu'elle n'est pas bien protégée. Près de 80 000 nouveaux cas de cancer de la peau sont diagnostiqués chaque année en France.

Le cancer de la peau est défini par la croissance anormale des cellules de la peau et qui peut avoir différentes causes, notamment l'exposition au soleil. Cette maladie peut toucher différentes parties du corps humain, principalement les zones du visage, du cuir chevelu, des mains et des jambes.

Le cancer de la peau est divisé en deux grandes catégories : **les non-mélanomes et les mélanomes**. L'incidence de mélanome est très importante en Australie, aux États-Unis, au Canada, au Sud de l'Amérique Latine, dans le Sud de l'Afrique et en Europe.

## Problématique

Les cancers en général et plus particulièrement les cancers de la peau représentent un problème majeur de santé publique de par le nombre de décès dus à ces maladies. Le nombre de nouveaux cas annuels de mélanomes augmente de 5% par an depuis 50 ans, une croissance supérieure à celle de tout autre cancer. Si le diagnostic intervient tôt dans l'évolution de la lésion, le taux de survie relative à 5 ans est de 88% pour les stades localisés. A contrario, la survie à 5 ans d'une personne atteinte d'un mélanome à un stade avancé, avec des métastases est inférieure à 20%. Détecter un mélanome à un stade précoce, alors que relativement peu de signes de malignité sont visibles sur la lésion est donc très souhaitable, mais cette opération est difficile. Les travaux de recherche menés précédemment dans l'équipe I&M ont montré que les classifieurs automatiques basés sur des modèles de machines à vecteurs de support (MVS) sont plus efficaces que les humains pour le diagnostic des lésions atypiques.

A travers ce projet nous allons établir un classifieur qui puisse différencier à partir d'une image une lésion maligne d'une lésion bénigne.

## Objectif

Nous allons utiliser les réseaux de neurones convolutifs (CNN) pour détecter et classer avec précision les lésions cutanées malignes le plus tôt possible. Les CNN désignent une sous-catégorie de réseaux de neurones qui sont à ce jour un des modèles de classification d'images réputés être les plus performant (ils réduisent très efficacement le nombre de paramètres sans perdre la qualité des modèles).

Quatre CNN différents modèles ont été utilisés pour analyser et prédire ce type de cancer : **CNN "TSM12", InceptionV3, ResNet152V2 et VGG16**.

Les modèles ont été entraînés puis comparés à l'aide des métriques décrites par la suite.

La base de données utilisée est **HAM10000**, une grande collection de données multi-sources des images dermoscopiques de lésions cutanées pigmentées courantes.

Ce travail est divisé en 5 chapitres :

- **Etude préliminaire** : Une petite introduction sur le cancer de la peau et les solutions existantes pour le détecter.
- **Outils de réalisation** : Les techniques et outils utilisés.
- **Méthodologie de réalisation** : Les étapes de réalisation.
- **Résultats** : Résultats commenté de l'implémentation.
- **Conclusion**

## 1. Etude préliminaire

Il y a trois grands types de cancer de la peau :

- **Le carcinome basocellulaire** : le cancer cutané le plus fréquent (70% des cas) et le moins agressif.
- **Le carcinome spinocellulaire** : qui représente 15% des cancers de la peau et peut également s'étendre à d'autres parties du corps.
- **Le mélanome malin** : il se développe dans 80% des cas à partir d'une peau saine et dans 20% des cas, à partir d'un grain de beauté. Connu comme étant la forme de cancer de la peau la plus dangereuse, le mélanome peut évoluer et toucher d'autres organes rapidement.

### 1.1 Epidémiologie

Il existe deux types de cancers de la peau : les carcinomes cutanés et les mélanomes. Les carcinomes représentent 90% des cancers de la peau et ne se propagent que rarement. Le second type, les mélanomes, sont des tumeurs malignes du système pigmentaire qui se développent à partir des **mélanocytes** (cellules qui fabriquent la mélanine responsable de la pigmentation de la peau humaine).

Le mélanome représente entre 2 et 3 % de l'ensemble des cancers. Le mélanome cutané se situe au 8ème rang des cancers chez l'homme et au 6ème rang chez la femme. Les deux tiers de ces cancers sont liés à des expositions excessives au soleil, principalement des expositions intermittentes et intenses pendant l'enfance.

Le mélanome est une tumeur potentiellement agressive surtout lorsqu'elle est prise à un stade tardif et qui peut donner des métastases mettant en jeu le pronostic vital.

### 1.2 Diagnostique

On détecte les mélanomes grâce à :

- ***L'interrogatoire du patient***

On recherche les facteurs de risque suivants :

– Environnementaux : exposition solaire intense, intermittente, dès un âge jeune, sur des régions du corps non exposées habituellement (tronc, jambes)

– Constitutionnels : antécédents personnels ou familiaux de mélanomes, peau claire ou sensible (taches de rousseur, sujets roux), naevus congénital géant...

- ***L'examen clinique***

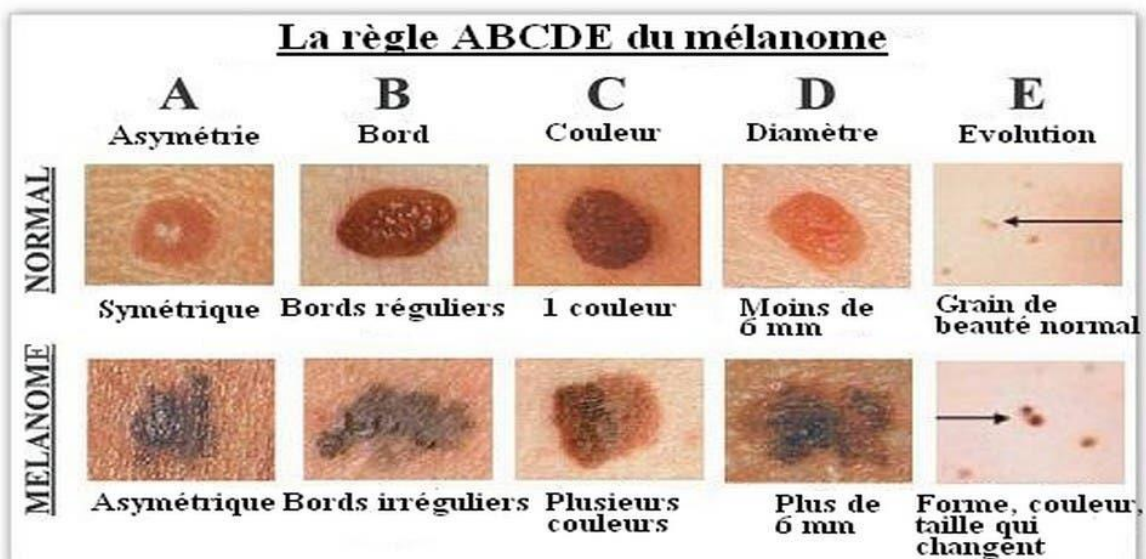
– Une lésion cutanée est suspecte lorsqu'elle répond aux critères suivants, réunis dans la règle ABCDE :

- **A** : asymétrie,
- **B** : bords irréguliers,
- **C** : couleur hétérogène,
- **D** : diamètre supérieur à 6 mm,
- **E** : extension en taille récente.

### **1.3 Formes**

***Le mélanome peut prendre différentes formes***

- A extension superficielle (sujet jeune, extension horizontale lente puis verticale)
- Nodulaire (lésion en relief)
- Lentigineux acral (atteint paumes, plantes, ongles chez des sujets noirs ou asiatiques)
- Muqueux (au niveau des muqueuses)



## 1.4 Pronostic de survie

Dans n'importe quel type de cancer de la peau, le pronostic va dépendre de la **précocité** du diagnostic et du stade de la maladie dont un des éléments est l'épaisseur de la lésion qui aura été enlevé chirurgicalement. Ces critères sont essentiels et sont définis par l'analyse au microscope. Il s'agit de **l'épaisseur de la lésion en mm** qui s'appelle l'indice de **Breslow** et du **niveau histologique** qui s'appelle l'indice de Clark.

Le pronostic et la survie pour la plupart des cancers de la peau autres que le mélanome sont excellents.

Par contre, Le mélanome représente **1,1 %** de l'ensemble des décès par cancer se classant ainsi au **16e rang des cancers les plus mortels**. Il est à lui seul responsable de 90 % des décès liés aux cancers cutanés.

## 1.5 Travaux précédents :

De nombreuses recherches ont été menées à l'aide de différentes techniques d'apprentissage automatique pour la classification des lésions cutanées.

En 2017, N. Codella et al. ont utilisé des machines à vecteurs de support (SVM) avec des réseaux de neurones pour la reconnaissance et la segmentation des mélanomes. Ce système a démontré une précision de classification supérieure à la précision de 8 experts en dermatologie.

En 2019, U. Yildiz et V. Kilic ont utilisé des classifieurs en extractant les couleurs et les textures caractérisant la base de données. Le pourcentage de l'accuracy était de 97%.

## 2. Outils de réalisation :

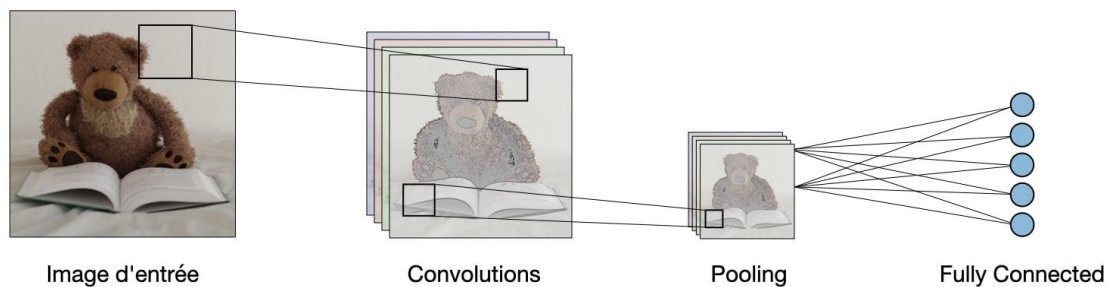
Dans ce chapitre, nous allons donner des informations sur les méthodes utilisées

### 2.1 CNN

Les réseaux de neurones convolutifs désignent une sous-catégorie de réseaux de neurones. Cependant, les CNN sont spécialement conçus pour traiter des images en entrée.

Comme pour les réseaux de neurones ordinaires, les paramètres des couches sont déterminés par rétropropagation du gradient : l'entropie croisée est minimisée lors de la phase d'entraînement. Mais dans le cas des CNN, ces paramètres désignent en particulier les features des images. Voir maintenant les différents types de couches d'un CNN. Il existe

quatre types de couches pour un réseau de neurones convolutif : la couche de **convolution**, la couche de **pooling**, la couche de **correction ReLU** et la couche **fully-connected**.



## 2.2 TSM12

Une implémentation de l'architecture CNN composée de 8 couches et 4 blocs. Le nombre 12 indique le nombre total de couches.

## 2.3 InceptionV3

Un modèle à 48 couches qui a été initialement présenté lors du défi de reconnaissance ImageNet en 2015.

## 2.4 ResNet152V2

Un réseau neuronal à 152 couches développé par l'équipe de recherche de Microsoft.

## 2.5 VGG16

C'est un modèle CNN à 16 couches qui utilise une architecture à profondeurs croissantes avec de faibles filtres de convolution.

## 2.6 Métriques de comparaison

5 métriques ont été utilisées pour comparer entre les performances des modèles :

- 1. La matrice de confusion** : La matrice de confusion est en quelque sorte un résumé des résultats de prédiction pour un problème particulier de classification. Elle compare les données réelles pour une variable cible à celles prédites par un modèle.
- 2. L'accuracy** : Ce paramètre fait la somme de tous les vrais positifs et vrais négatifs qu'il divise par le nombre total d'instances. Il permet d'apporter une réponse à la question suivante : de toutes les classes positives et négatives, combien parmi elles ont été prédites correctement ?
- 3. Le rappel (Recall)** : Le rappel est un paramètre qui permet de mesurer le nombre de prévisions positives correctes sur le nombre



total de données positives. Il permet de répondre à la question suivante : sur tous les enregistrements positifs, combien ont été correctement prédits ?

4. **La précision** : La précision indique le rapport entre les prévisions positives correctes et le nombre total de prévisions positives. Ce paramètre répond donc à la question suivante : sur tous les enregistrements positifs prédits, combien sont réellement positifs ?
5. **Le score F1** : Le score F1 (ou F-measure en anglais) est une moyenne harmonique de la précision et du rappel. Il équivaut au double du produit de ces deux paramètres sur leur somme. Sa valeur est maximale lorsque le rappel et la précision sont équivalents.

## 2.7 La base de données

La base de données utilisée est HAM10000. C'est une base de données large avec une grande collection d'images dermoscopiques multi sources représentant des lésions cutanées pigmentées courantes. Elle est composée de 10015 images de lésions cutanées de dimensions 600\*450 et stockées sous format jpg.

- ✓ D'abord, un traitement de données est effectué.
- 1. La base de données contient initialement 7 attributs :
  - **Patient's Age [age]**
  - **Patient's Sex [sex]**
  - **A Lesion ID [lesion id]**
  - **An Image ID [image id]**
  - **A Technical Validation Field Type [dx type]**
  - **The Localization of the Skin Lesion [localization]**
  - **A Diagnostic Skin Lesion Category [dx]**

```
✓ [22] skin_df.describe(exclude=[np.number])
0 s
```

	lesion_id	image_id	dx	dx_type	sex	localization
count	10015	10015	10015	10015	10015	10015
unique	7470	10015	7	4	3	15
top	HAM_0003789	ISIC_0028402	nv	histo	male	back
freq	6	1	6705	5340	5406	2192

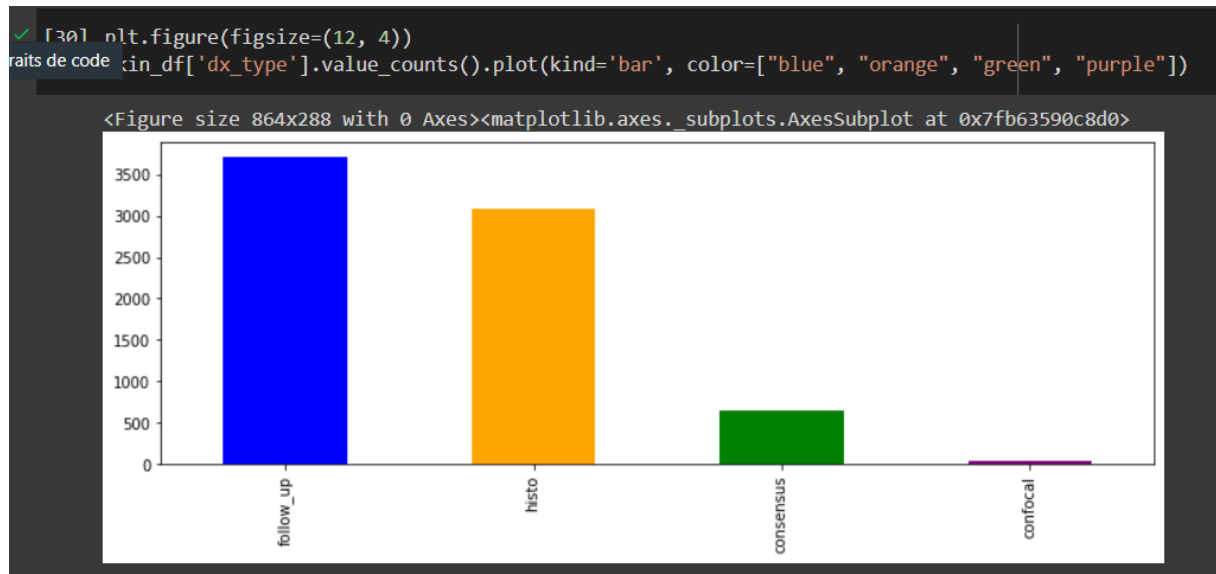
A partir de cet affichage nous remarquons :

- Il existe plusieurs images pour la même lesion\_id -> **Duplication**
- La classe **nv** a **dominé** toutes les autres catégories de lésions avec 6705 images sur 10015.

2. L'attribut technique de validation désigne la manière ou la méthode par laquelle la lésion a été diagnostiquée.

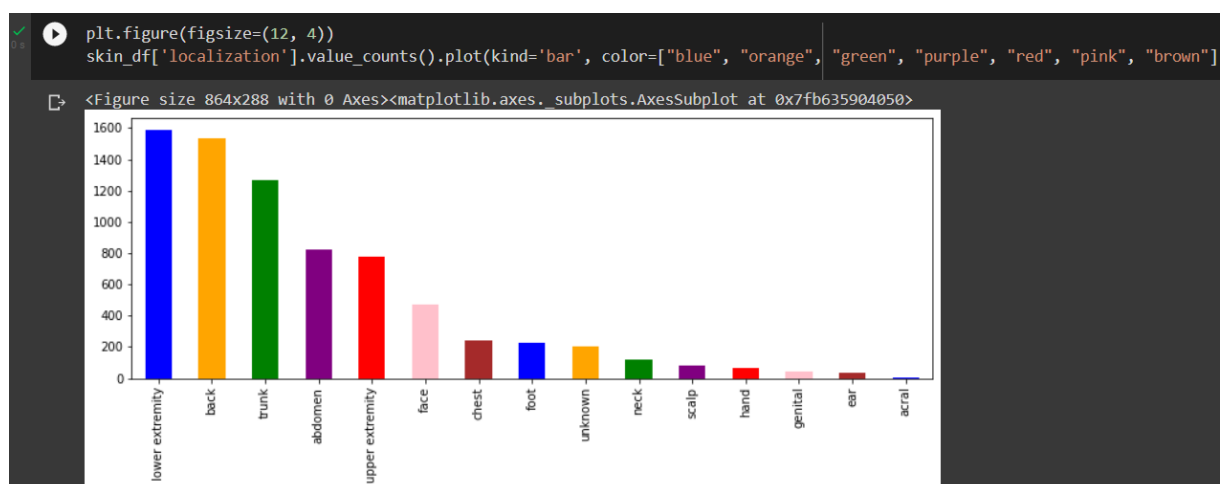
Selon le fichier csv, il existe 4 méthodes :

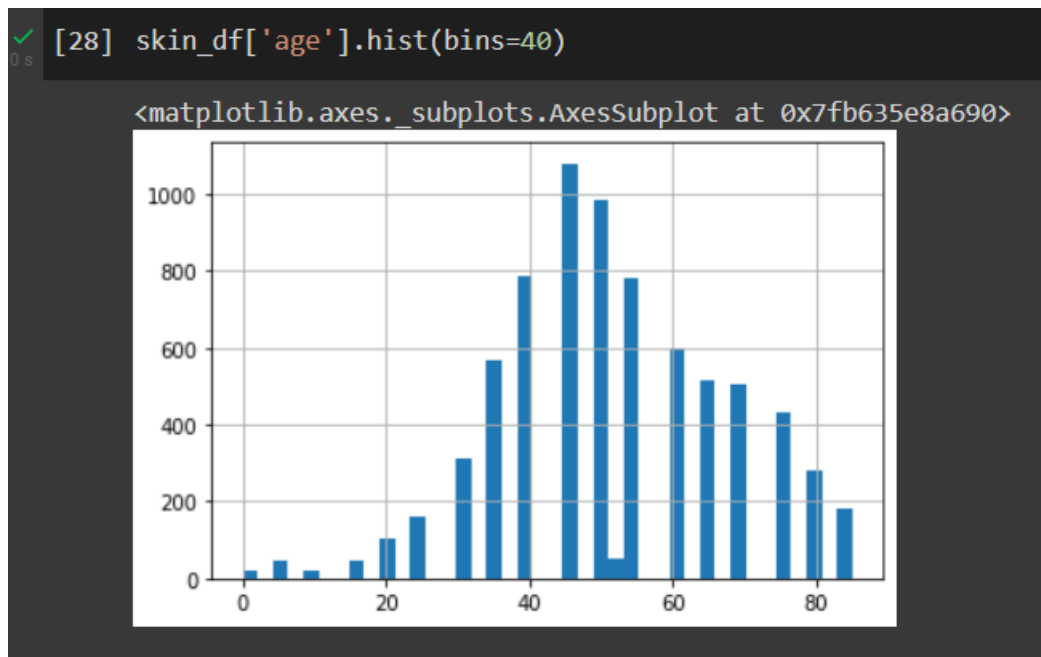
- **Histopathology (histopathologie)** : réalisé par des dermatologistes spécialisés.
- **Confocal** : La microscopie confocale (MC) par réflectance est une technique d'imagerie non invasive qui permet d'obtenir en temps réel des images de différents tissus de la peau avec une résolution proche de l'histologie.
- **Follow-up** : Le suivi des lésions.
- **Consensus** : Pour les cas bénins typiques.



- D'après la figure, la plupart des lésions ont été diagnostiquées après un **suivi** de la lésion.
- Plus de **50%** des lésions ont été diagnostiquées à travers **l'histopathologie**.

3. La figure suivante montre la localisation des lésions dans le corps humain.





La figure ci-dessus montre le nombre de patients par âge.

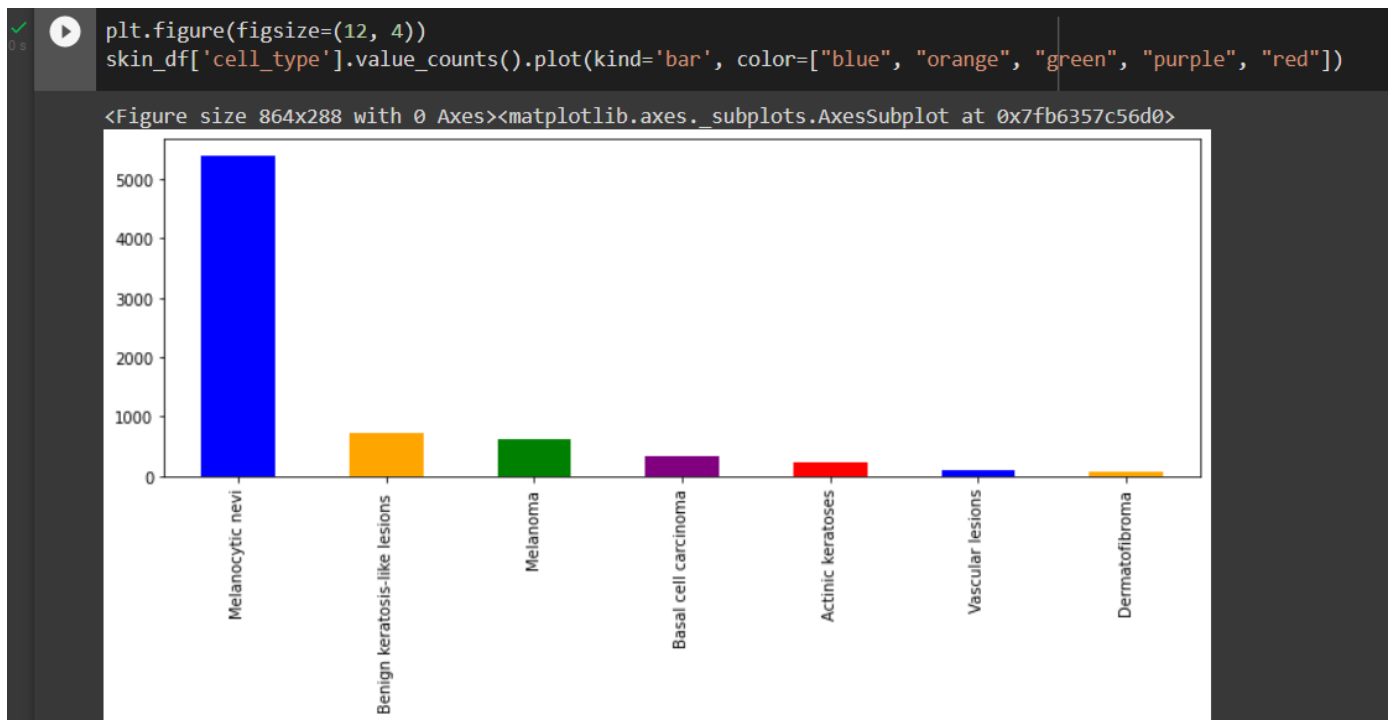
On remarque que la plupart des patients sont âgés entre 35 et 70 ans.

4. La figure suivante montre la fréquence de chaque type de lésions.

Il y a 7 types de lésions :

- **Melanocytic Nevi [nv]** : est une lésion cutanée pigmentée constituée de mélanocytes et à haut risque de transformation maligne. **[6705 images]**
- **Melanoma [mel]** : est le plus grave des cancers de la peau étant donné sa forte capacité à métastaser. **[1113 images]**
- **Benign Keratosis-like Lesions [bkl]** : C'est des lésions pigmentées bénignes. La cause est inconnue. Elles ont tendance à se développer chez les personnes âgées et ont une apparence collée avec une surface verruqueuse, lisse, cireuse, squameuse ou croûteuse. **[1099 images]**
- **Basal Cell Carcinoma [bcc]** : un cancer de la peau dont l'évolution est très favorable sous traitement. **[514 images]**
- **Actinic Keratoses [akiec]** : est la lésion précancéreuse la plus courante qui se forme sur la peau endommagée par une exposition chronique aux rayons ultraviolets (UV) du soleil et/ou d'un bronzage artificiel. **[327 images]**
- **Vascular Lesions [vasc]** : Type de problème de peau qui est très courant. **[142 images]**

- **Dermatofibroma [df]** : est une petite masse dure dont la couleur varie du rose ou rouge au brun. **[115 images]**



On remarque que le **Melanocytic Nevi** est le type de lésions le plus courant.

### 3. Méthodologie de réalisation

#### 3.1 Prétraitement

- Création de deux dictionnaire, le premier contient les noms d'images, le deuxième contient les catégories de diagnostics pour les lésions cutanées.
- Nettoyage de données : éliminer les duplications.
- Des codes numériques uniques ont été créés pour chaque catégorie de lésions, car les entiers sont plus faciles à gérer que les chaînes de caractères.
- Pour assurer que l'apprentissage soit plus rapide les images ont été redimensionnées par un facteur de 1/4. Les dimensions des images sont devenues 150 \* 112,5.
  - Ensuite les images ont été stockées sous forme de liste d'images numériques.

#### 3.2 Data splitting

- Convertir les étiquettes numériques en une matrice binaire (**one hot matrix**)
- La base de données a été divisée en 70% training 20% test et 10% validation. Cette division est fréquente lorsque le nombre d'entrées ne dépasse pas 10 000.

#### 3.3 Normalisation

L'équation utilisée dans la normalisation est :

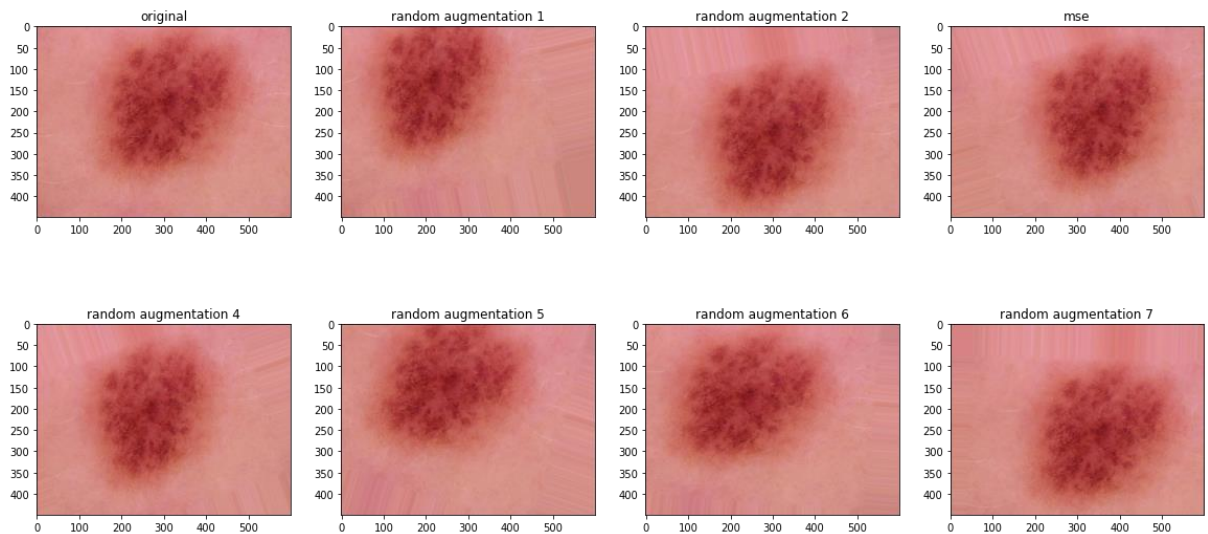
$$x_{new} = \frac{x - u}{\sigma}$$

#### 3.4 Data augmentation

L'augmentation de données était une étape nécessaire pour obtenir des résultats plus exactes et plus robustes. Les méthodes appliquées sont :

- Une rotation aléatoire de 20%
- Une translation aléatoire horizontale de 20%
- Une translation aléatoire verticale de 20%
- Un shear aléatoire de 10%
- Un zoom aléatoire de 10%

La figure suivante représente un exemple réel de l'augmentation de données



### 3.5 Modèles

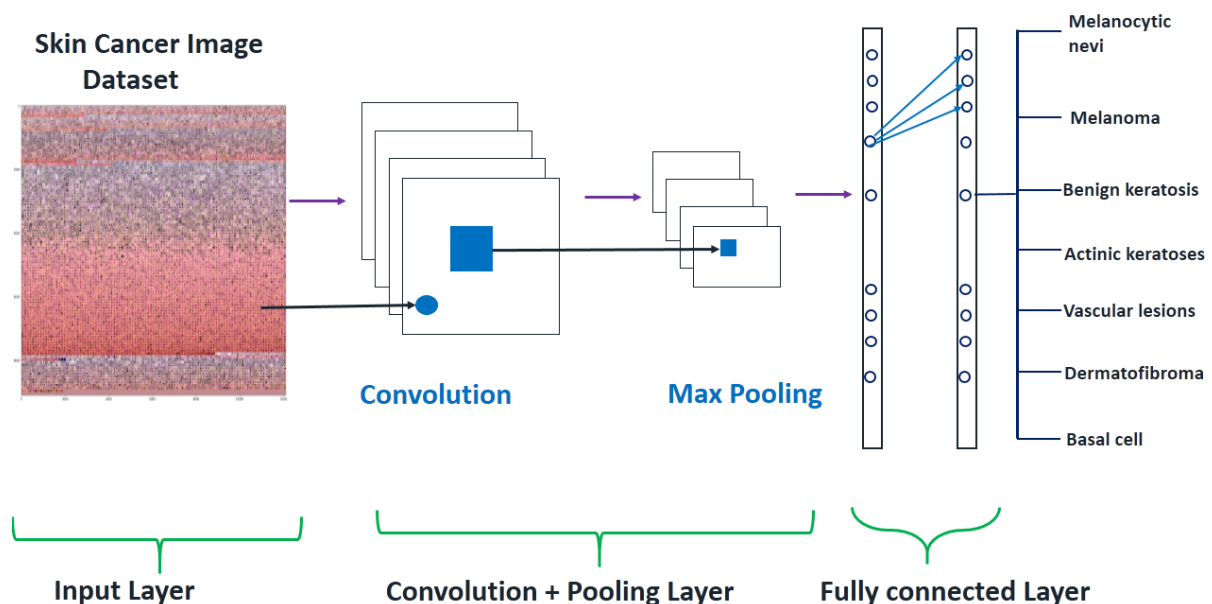
Les modèles ont été appliqués en utilisant les bibliothèques Tensorflow et keras.

#### TSM12

La figure suivante montre l'architecture de ce modèle et ses différentes couches.

- Les **2 premières couches** sont des couches convolutifs 2D.
- 2 couches ont été utilisées au lieu d'une, pour que le modèle soit plus flexible pour exprimer les transformations non-linéaires.
- **La couche 3** est **Max Pooling** qui supprime les informations du signal, elle agit comme un filtre de sous échantillonnage.
- **La couche 4 (Dropout)** est une technique de régularisation où les neurones sélectionnés au hasard sont ignorés. Elle utilise une représentation distribuée. 1/4 des neurones ont été ignorés.
- 32 Filtres ont été utilisés
- ReLU a été utilisé dans pratiquement toutes les couches sauf la dernière couche.
- Les couches de **5 jusqu'à 8** sont une duplication des couches de 1 à 4 avec de différents paramètres.

- Le nombre de filtres utilisés ici est 64 au lieu de 32 et la couche Dropout a des valeurs de neurones plus élevée.
- **La couche 9** transforme l'entrée en un vecteur 1D.
- **La couche 10** est une couche dense d'entrée FC avec 128 unités.
- **La couche 11** est une couche dropout ,1/2 des neurones ont été ignoré.
- **La couche 12 est** une couche dense de sortie FC avec 7 unités, le même nombre des catégories de lésion. La fonction d'activation est softmax au lieu de ReLU car une classification binaire a été utilisée et fonctionne bien avec l'encodage one-hot.



Layer #	Layer	Hyperparameters
1	Conv2D	32 Filters, 3x3 Filter Size ReLU Activation, Same Padding
2	Conv2D	32 Filters, 3x3 Filter Size ReLU Activation, Same Padding
3	MaxPool2D	2x2 Pool Size
4	Dropout (Core Layer)	0.25 Neurons
5	Conv2D	64 Filters, 3x3 Filter Size ReLU Activation, Same Padding
6	Conv2D	64 Filters, 3x3 Filter Size ReLU Activation, Same Padding
7	MaxPool2D	2x2 Pool Size
8	Dropout (Core Layer)	0.4 Neurons
9	Flatten (Core Layer)	-
10	Dense	128 Units, ReLU Activation
11	Dropout (Core Layer)	0.5 Neurons
12	Dense	7 Units, Softmax Activation



## InceptionV3, ResNet152V2 et VGG16

Ces modèles ont été implémenté via la bibliothèque Keras, des couches ont été ajoutées pour ajuster les paramètres, les entrées et les sorties.

### Les hyperparamètres

Nous avons utilisé certaines valeurs d'hyperparamètres communes pour avoir un meilleur modèle d'évaluation. La figure suivante met en évidence les valeurs de ces hyperparamètres :

- **Optimiseur (Adam)** : est l'optimiseur le plus utilisé et le plus simple à implémenter. Très efficace et peut manipuler de larges données.
- **Fonction de perte (Categorical Cross-Entropy)** : est utilisé lorsqu'une seule catégorie est appliquée pour chaque point de données ce qui était très optimale pour notre projet.
- **Epochs** : De multiples essais avec les valeurs : 20, 50, 100, 150 et 200 ont été effectués pour trouver le meilleur résultat. 50 était la valeur la plus optimale.
- **Batch size** : De multiples essais avec les valeurs : 5, 10, 20 et 40 ont été effectués pour trouver le meilleur résultat. 10 était la valeur la plus optimale.
- **Learning rate (taux d'apprentissage)** : Décroissant qui commence de 0.001 pour TSM12 et InceptionV3 et commence de 0.0001 pour ResNet152V2 et VGG16.

Hyperparameter	Value
Optimizer	Adam
Loss Function	Categorical Cross-Entropy
Epochs	50
Batch Size	10
Learning Rate	0.001 - 0.00001 (Reduces on Plateau)

### Les poids :

Des poids ont été ajoutés pour rendre les modèles plus sensibles à la catégorie des lésions cutanées des naevus mélanocytaires [nv] en raison de la déséquilibre des classes, puisqu'il représentait environ les deux tiers des données. Les pondérations étaient les facteurs par lesquels la valeur de la perte était multiplié en interne pour être utilisé dans l'algorithme de rétropropagation. L'équation suivante montre comment cela a fonctionné

$$weighed\ loss\ class[i] = loss[i] \_ class\ weights[i]$$

Skin Lesion Category	Weight
Actinic keratoses [akiec]	1.0
Basal cell carcinoma [bcc]	1.0
Benign keratosis-like lesions [bkl]	1.0
Melanoma [mel]	1.0
Melanocytic nevi [nv]	3.0
Vascular lesions [vasc]	1.0
Dermatofibroma [df]	1.0

## 4. Résultats

Ce projet a été exécuté dans un environnement google colab qui permet l'accès gratuit à un gpu.

### ➤ Précision des modèles, perte, erreurs et pour les ensembles d'apprentissage et de validation

		TSM12	InceptionV3	Resnet152V2	VGG16
<b>Exactitude (%)</b>	train	80.04	77.07	<b>99.43</b>	<b>94.56</b>
	val	77.67	76.25	78.51	76.83
<b>Loss</b>	train	0.55	0.64	0.02	0.17
	val	0.78	0.76	1.35	1.41
<b>RMSE</b>	train	0.11	0.11	0.005	0.034
	val	0.12	0.12	0.11	0.11
<b>MSE</b>	train	0.038	0.044	0.001	0.01
	val	0.044	0.049	0.05	0.055

### Remarques :

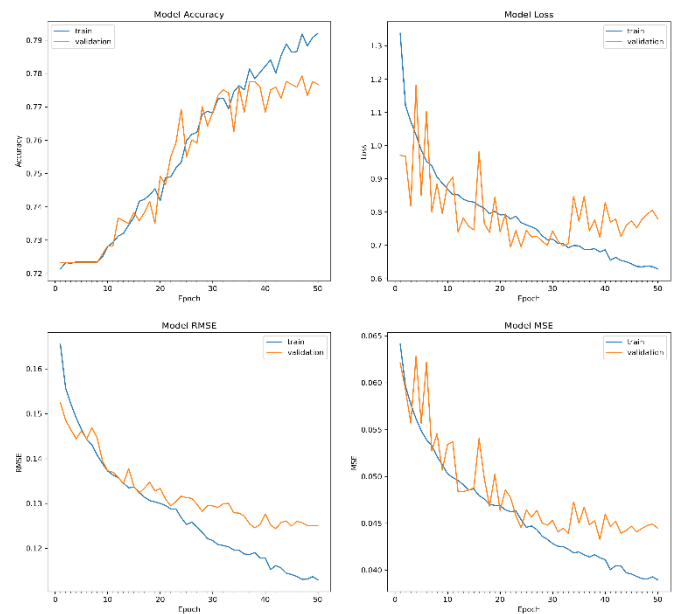
- On a remarqué que le modèle TSM12 était le plus rapide par contre ResNet152V2 était le plus lent, cela est dû au fait que ResNet152V2 a plus de couches que TSM12(152 couches).
- Selon les résultats, il y a eu un **overfitting** pour les modèles VGG16 et ResNet152V2(high bias).
- Un low bias pour TSM12 et InceptionV3 (ni overfitting ni underfitting).

## ➤ Historique des modèles

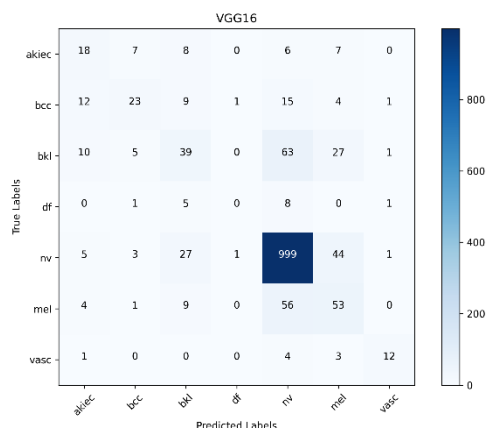
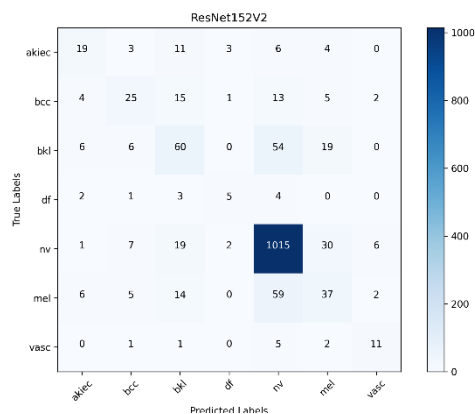
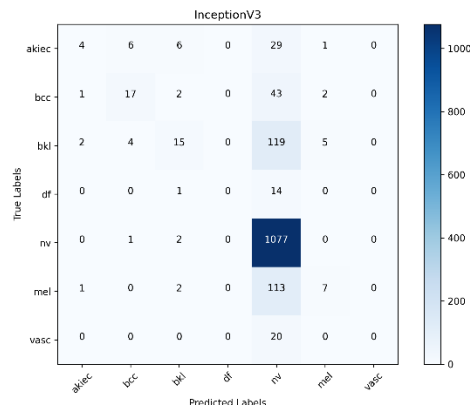
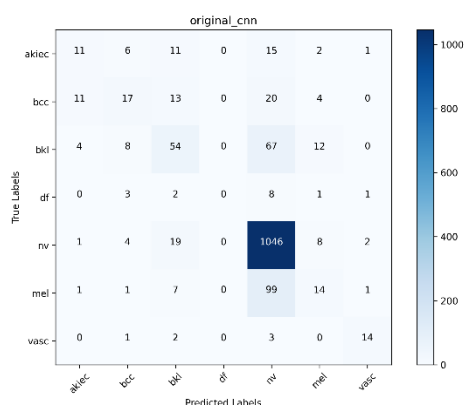
Ces historiques sont très importants pour savoir si le modèle a continué son apprentissage sans erreurs.

Ces historiques sont disponibles dans le code.

La figure montre l'historique du modèle TSM12.



## ➤ Matrices de confusion



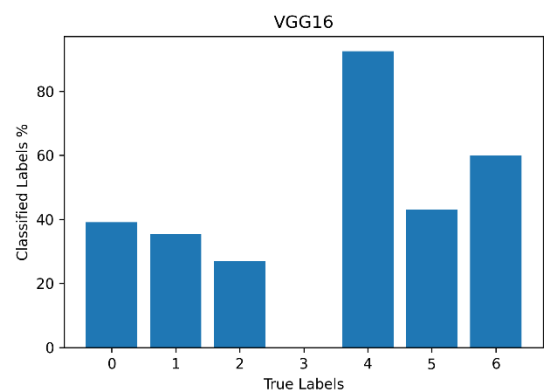
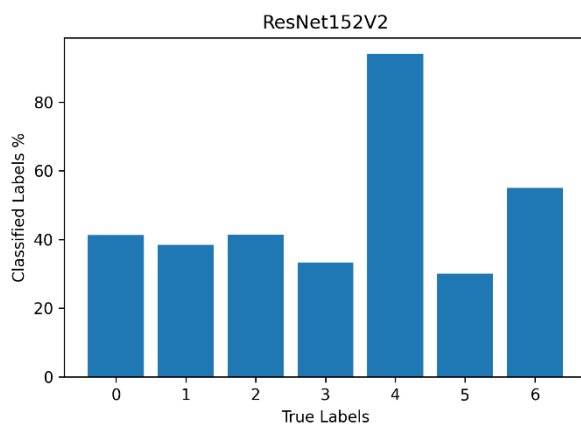
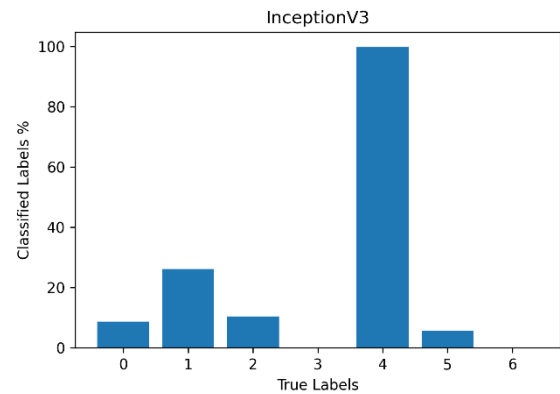
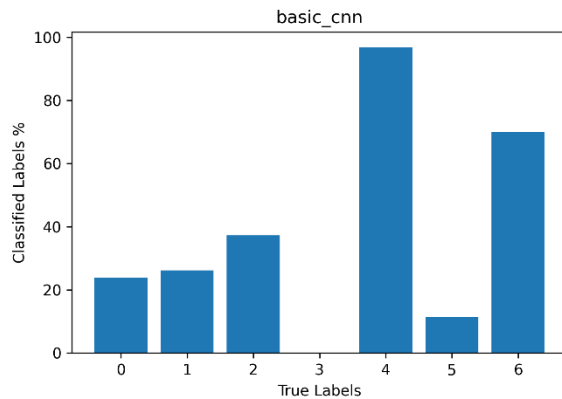
Métriques	TSM12	InceptionV3	ResNet152V2	VGG16
Accuracy (%)	77	75	78	77
Précision	72	68	76	74
Recall	77	75	78	77
Score F1	74	67	77	75

Le tableau précédent montre les valeurs des métriques pour les 4 modèles, et les figures montrent les matrices de confusion des 4 modèles.

### **Remarques :**

- On a remarqué que les valeurs de test accuracy sont presque les mêmes que validation accuracy, mais puisque pour l'apprentissage les valeurs sont différentes : on a un overfitting pour ResNet152V2 et VGG16.
- Selon l'article et selon leurs résultats (pas d'overfitting) Le modèle VGG16 est le mieux classé avec 88.55% d'accuracy, 87.76% de précision, 88.74% de Recall et 87.69 de F1 score. Suivi de ResNet152V2, TSM12 et enfin InceptionV3.
- Cependant on peut utiliser une approche naïve pour classer nos résultats : **On ignore l'overfitting et on classe selon les performances.**
- On obtient l'ordre suivant : **ResNet152V2, VGG16, TSM12 et enfin InceptionV3**
- Il faut noter que Resnet152V2 est plus lent que tous les autres car il contient 152 couches.

## ➤ L'accuracy pour les modèles de classification



### Remarques :

- Les 4 modèles ont une accuracy supérieure à **85%** pour la classe **Melanocytic Nevi [nv] (Classe avec nombre maximum d'images)**. Pour les autres classes, les valeurs sont différentes pour chaque modèle.
- Bien au contraire, les 4 modèles ont des valeurs d'accuracy **minimes** pour la classe **Dermatofibroma [df] (Classe avec nombre minimum d'images)**.
- On voit que le modèle InceptionV3 a eu des résultats inférieurs aux autres (le moins performant).
- On peut conclure que **plus il y a d'images**, plus les modèles **sont plus performants**.

## 5. Conclusion

- Ce projet avait pour but de comparer les différentes architectures Cnn dans la prédiction des lésions cutanées sur des images de lésions cutanées.
- La base de données a été prétraité par : le nettoyage de données-la division des données-la normalisation- l'augmentation et le redimensionnement de données.
- Nous avons construit une nouvelle architecture Cnn appelée TSM12. Nous avons utilisé d'autres modèles existants : InceptionV3, ResNet152V2 et VGG16.
- Des hyperparamètres et des métriques ont été utilisés pour comparer entre les modèles.
- Selon l'article le modèle VGG16 était le plus performant, mais en considérants nos résultats et en ignorant l'overfitting, le modèle ResNet152V2 était plus performant que VGG16.
- Et enfin nous pouvons dire que plus les données sont plus larges pour une classe, plus l'apprentissage est meilleur.
- **Il est conseillé d'augmenter les données pour la classe des lésions malignes pour une meilleur prédiction de ces tumeurs.**

Réalisé par :  
MERBOUTI SOUAD  
M2IV



## 6. Références

<https://www.e-cancer.fr/Comprendre-prevenir-depister/Se-faire-depister/Depistage-des-cancers-de-la-peau/Les-cancers-de-la-peau>

<https://www.cancer-environnement.fr/340-Melanome-cutane.ce.aspx>

<https://www.kaggle.com/kmader/skin-cancer-mnist-ham10000>

Skin Lesion Classification using Convolutional  
Neural Networks

Tareq Tayeh, Member, IEEE, Sulaiman Aburakhia, Member, IEEE,  
Moataz ElKhishen, Member, IEEE