

# Reconnaissance Races de Chiens

MEHDI DOUBIANI





1. PROBLEMATIQUE

2. ÉTAPES SUIVIES

4. APPROCHE CLASSIQUE

5. APPROCHE CNN CLASSIQUE

3. DONNÉES

6. APPROCHE CNN TRANSFERT LEARNING

7. CONCLUSIONS ET PISTES D'AMÉLIORATIONS

# PROBLEMATIQUE

- Une association de protection des animaux aimerait avoir un algorithme capable de classer les images en fonction de la race du chien présent sur l'image.



# ÉTAPES SUIVIES

## DONNÉES

Classification d'image de  
races de chiens:

Stanford Dogs Dataset  
20580 pour 120 races

## APPROCHE CLASSIQUE

- Prétraitement des images (filtres)
- Extraction de features (SIFT)
- Réduction de dimension (clustering)
  - Visual Bag-Of-Words
  - Classification

## APPROCHE CNN CLASSIQUE

- Réseaux de neurones convolutionnels
- CNN à partir de zéro (from scratch)
- Data Augmentation



# ÉTAPES SUIVIES

## APPROCHE CNN TRANSFERT LEARNING

Test de 5 réseaux de neurones pré-entraînés :

1. VGG16
2. ResNet50v2
3. Inception V3
4. Inception V4  
(InceptionResNet50v2)
5. Xception

## APPROCHE CNN TRANSFERT LEARNING

1. En bloquant la base convolutionnelle et sans data augmentation.
2. En bloquant la base convolutionnelle et avec data augmentation.
3. En bloquant uniquement les premières couches convolutionnelle et avec data augmentation.

COMPARAISON DES  
MODELES ET  
CONCLUSIONS

# DONNÉES

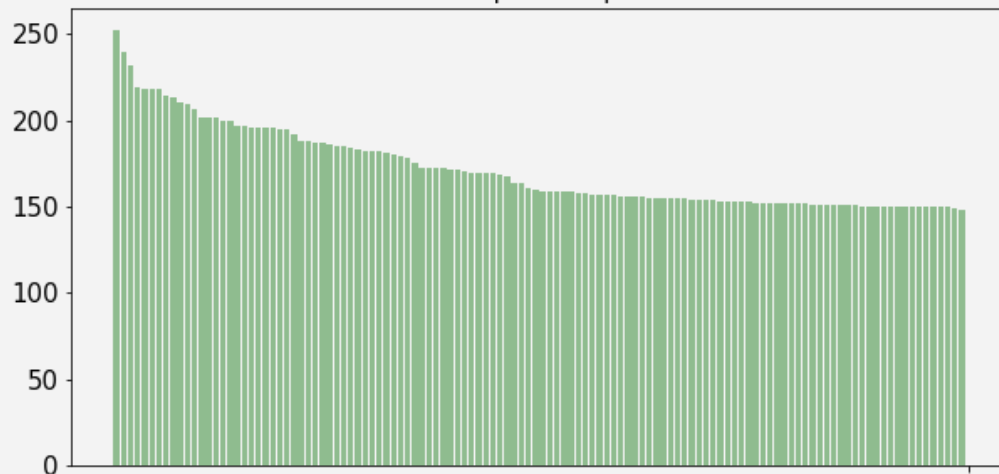
20580

DOG PHOTOS

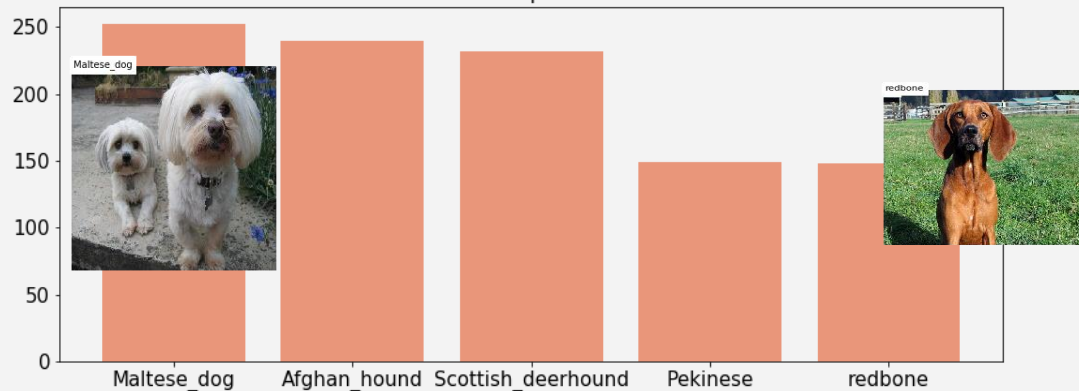
120

BREEDS

Number of pictures per breed



Most and least represented breeds





# APPROCHE CLASSIQUE

Pré-traitement :

- Redimensionnement 500x375px
- Passage en niveaux de gris
- Test de plusieurs filtres :
  - ✓ Egalisateur d'histogrammes
  - ✓ Filtres moyenneur, gaussien, médian, non-local mean
  - ✓ Whitening



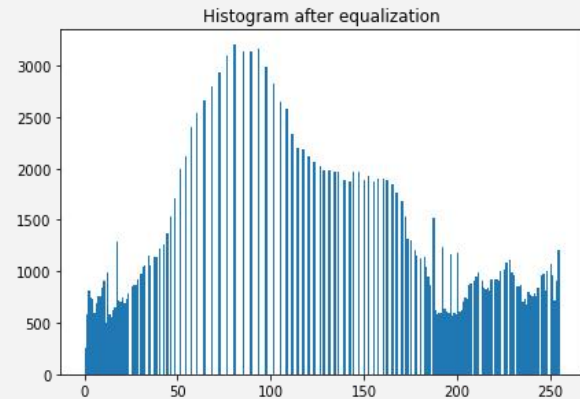
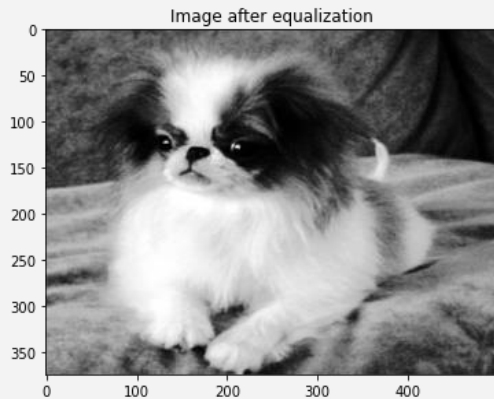
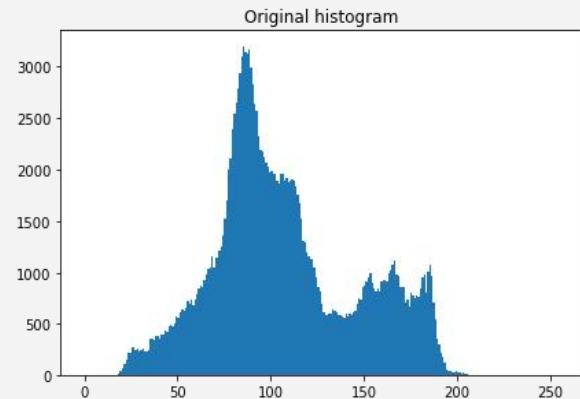
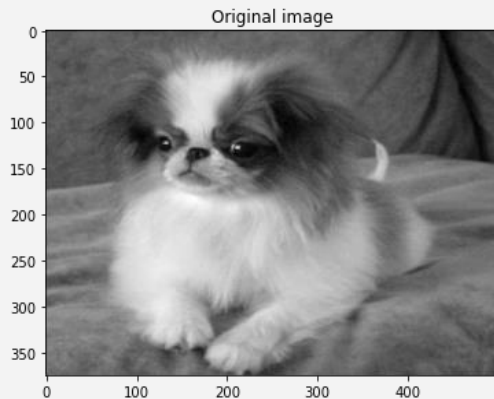
# APPROCHE CLASSIQUE

Pré-traitement :

Egalisation d'histogrammes

Améliore le contraste

Donne de la clarté aux détails

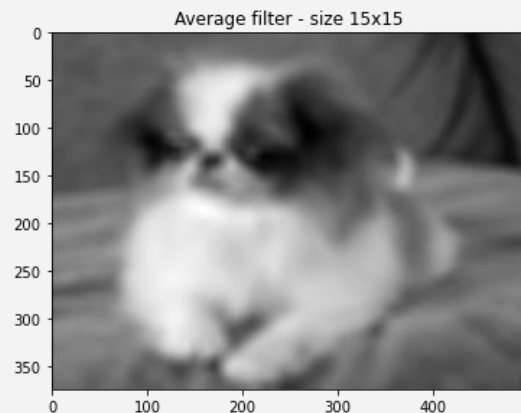
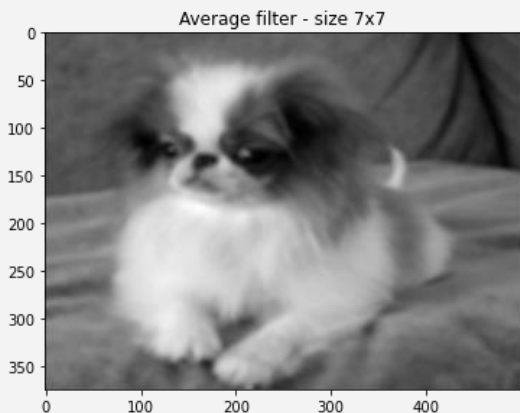
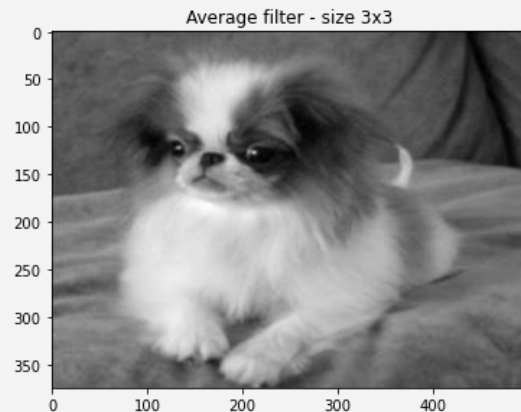
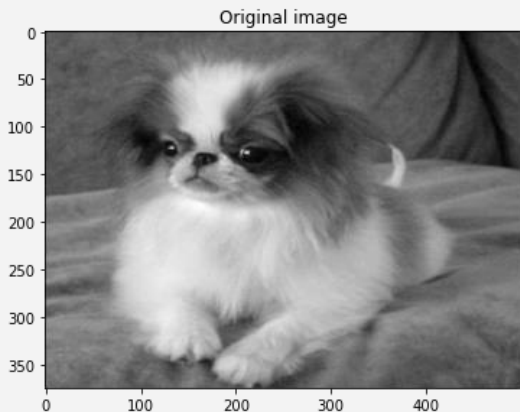




# APPROCHE CLASSIQUE

Pré-traitement :

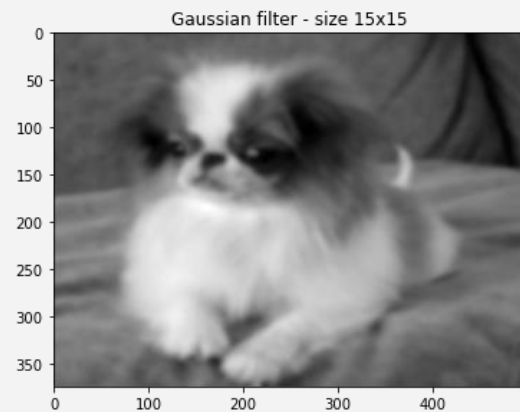
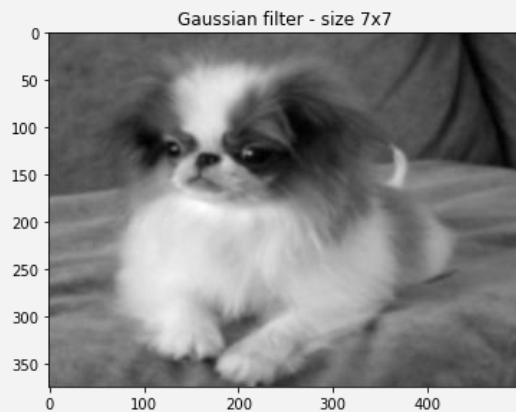
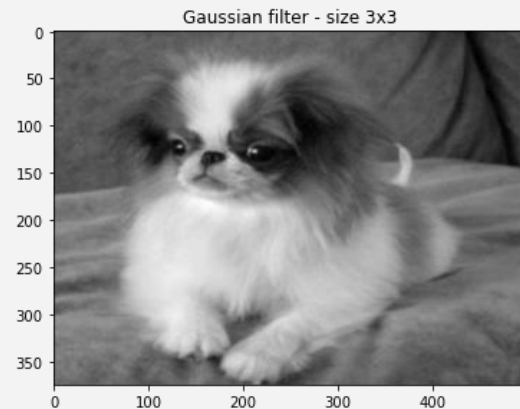
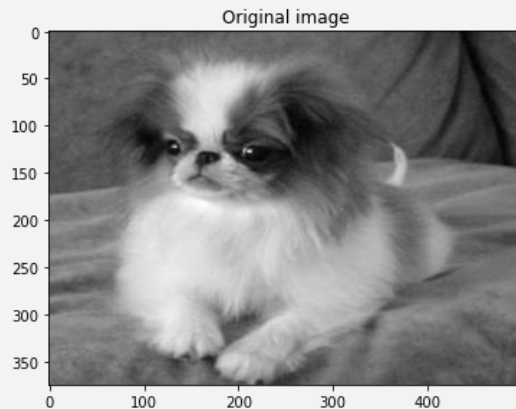
Filtre Moyenneur



# APPROCHE CLASSIQUE

Pré-traitement :

Filtre Gaussien

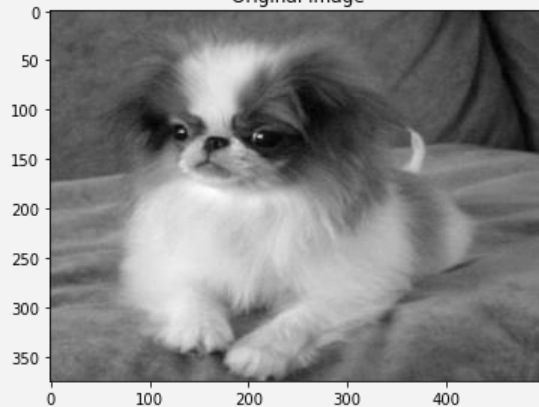


# APPROCHE CLASSIQUE

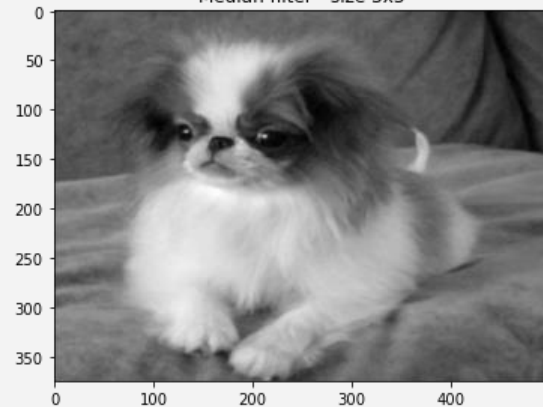
Pré-traitement :

Filtre Médian

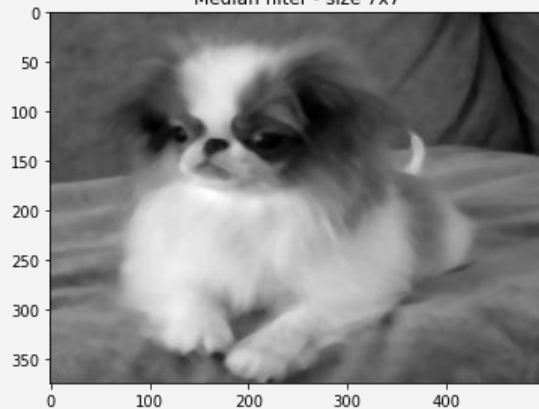
Original image



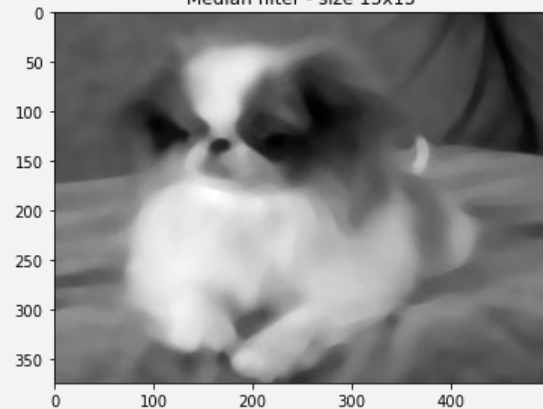
Median filter - size 3x3



Median filter - size 7x7

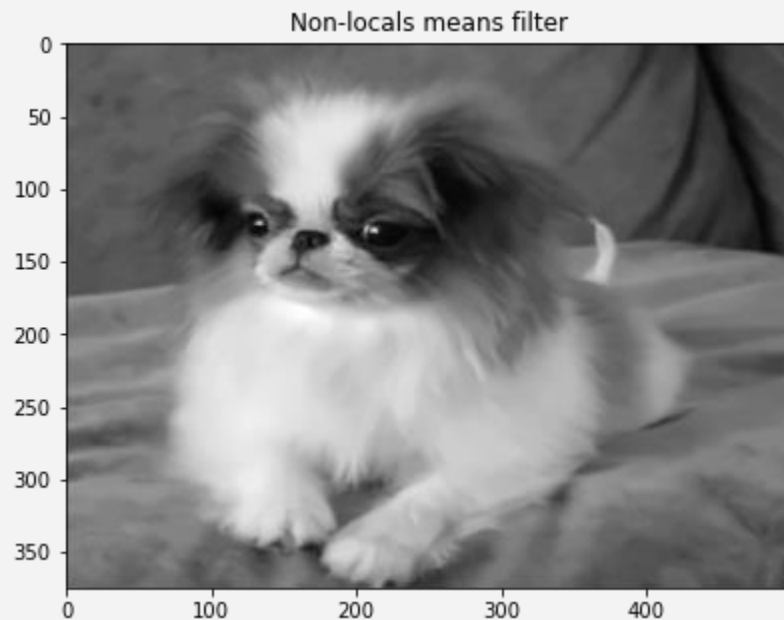
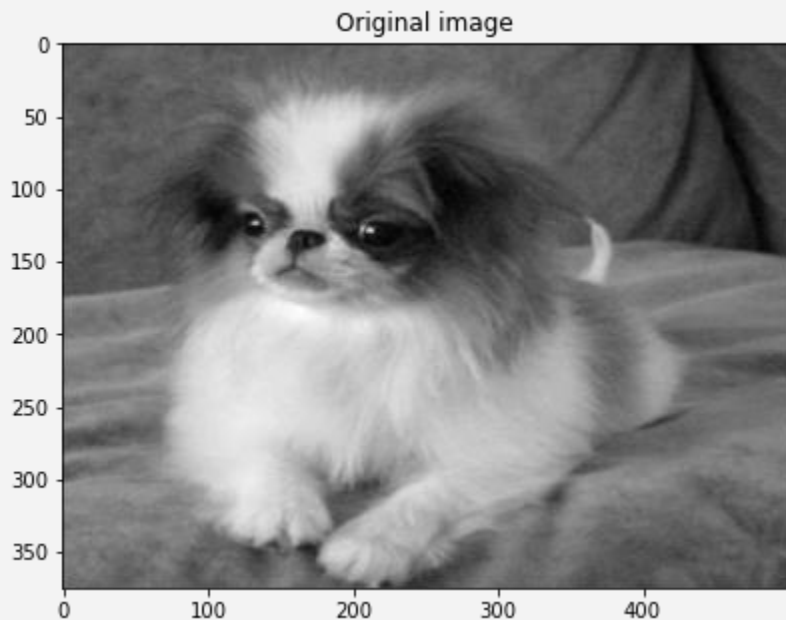


Median filter - size 15x15



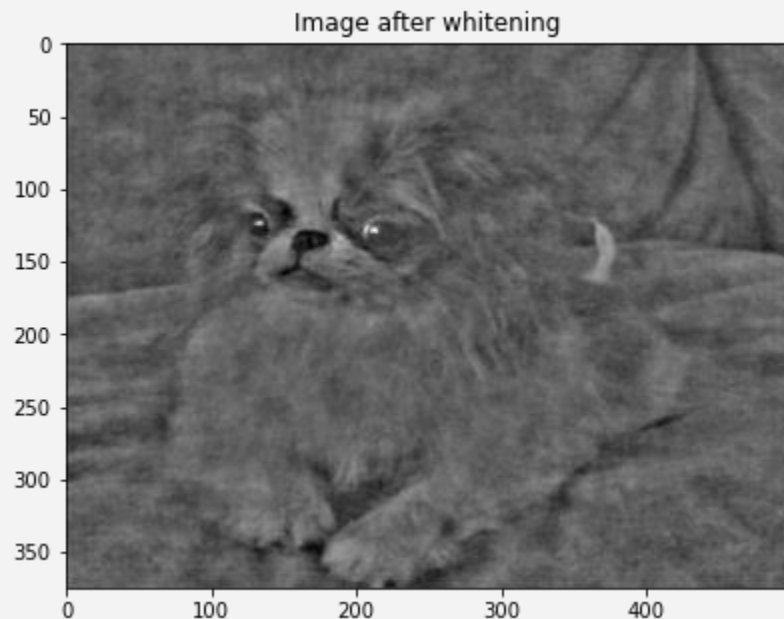
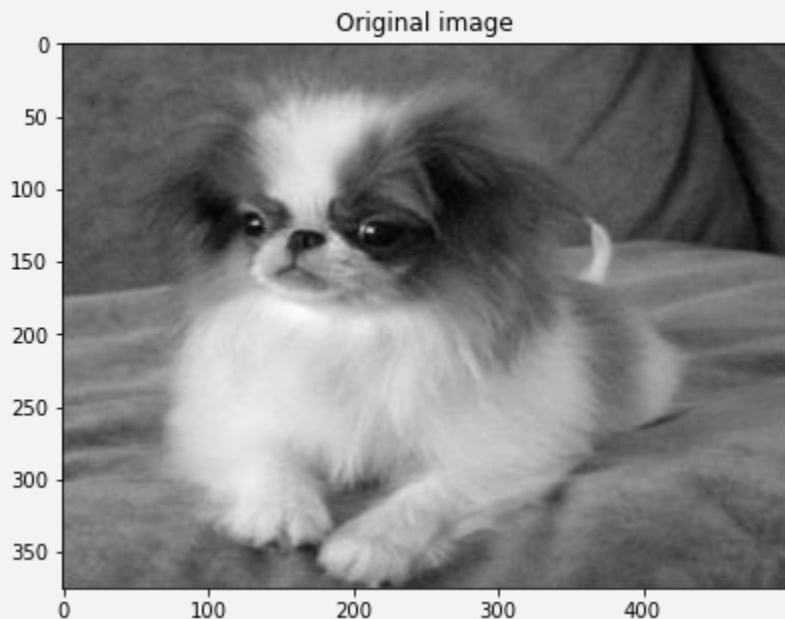
# APPROCHE CLASSIQUE

Pré-traitement :  
Non Local Means

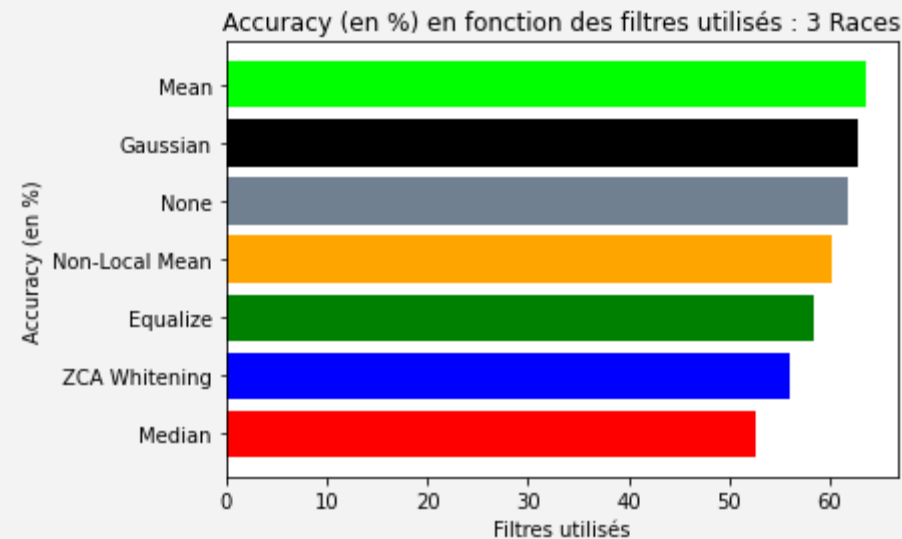
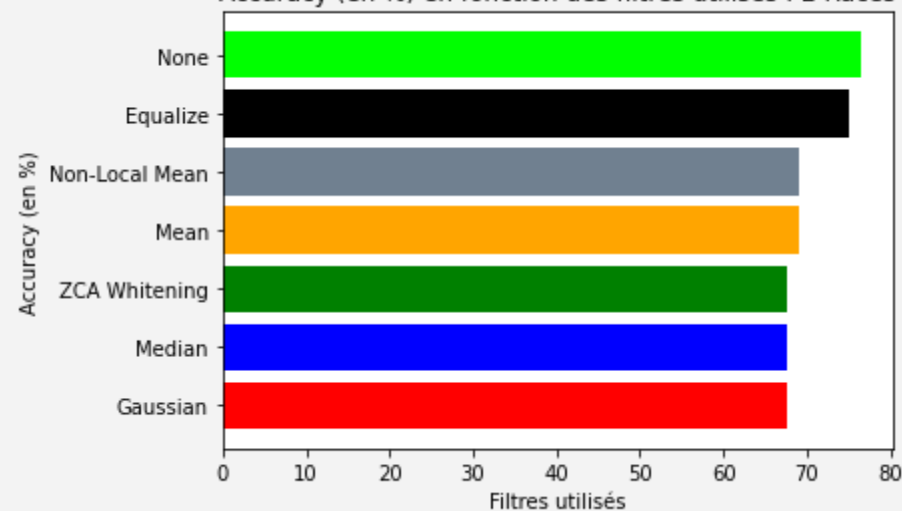
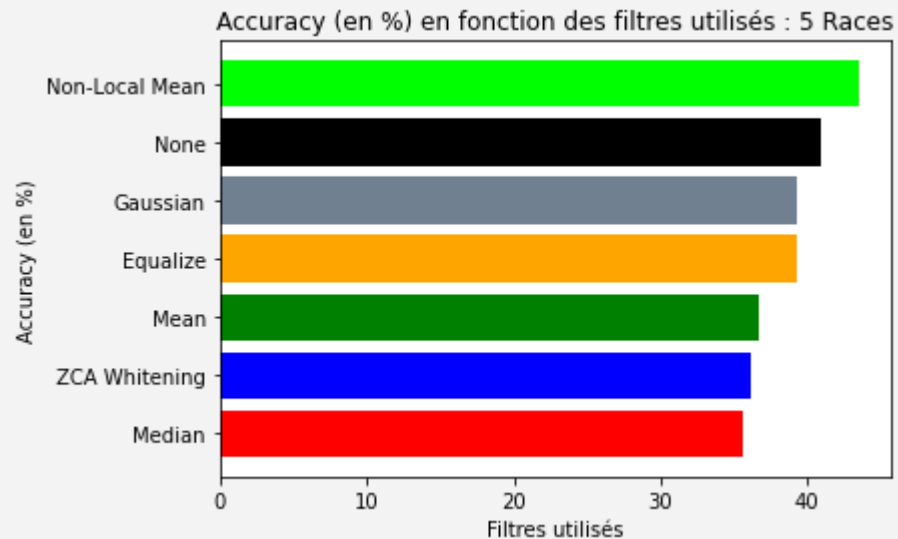


# APPROCHE CLASSIQUE

Pré-traitement :  
Whitening



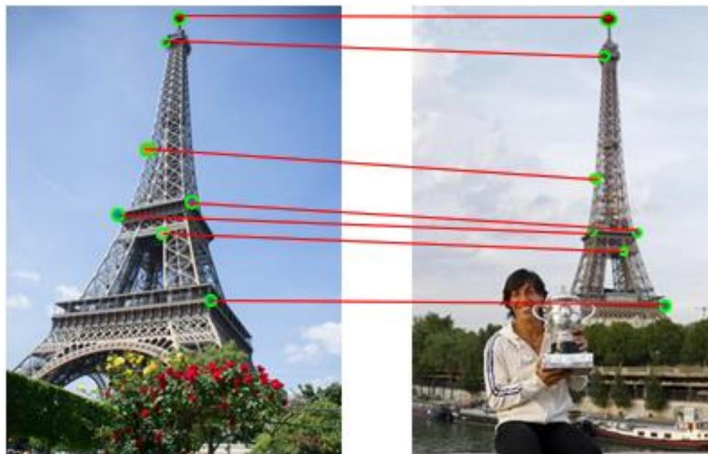
# APPROCHE CLASSIQUE





# APPROCHE CLASSIQUE

## Extraction de features

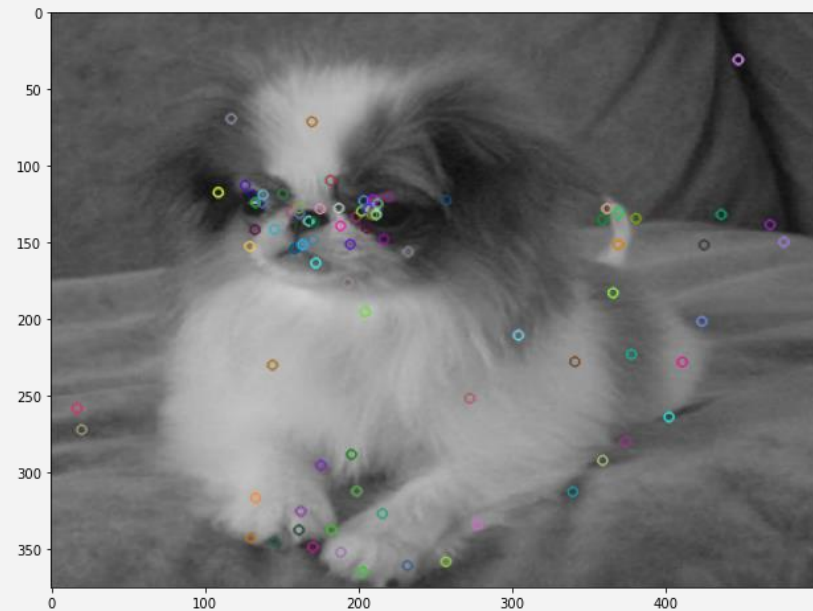
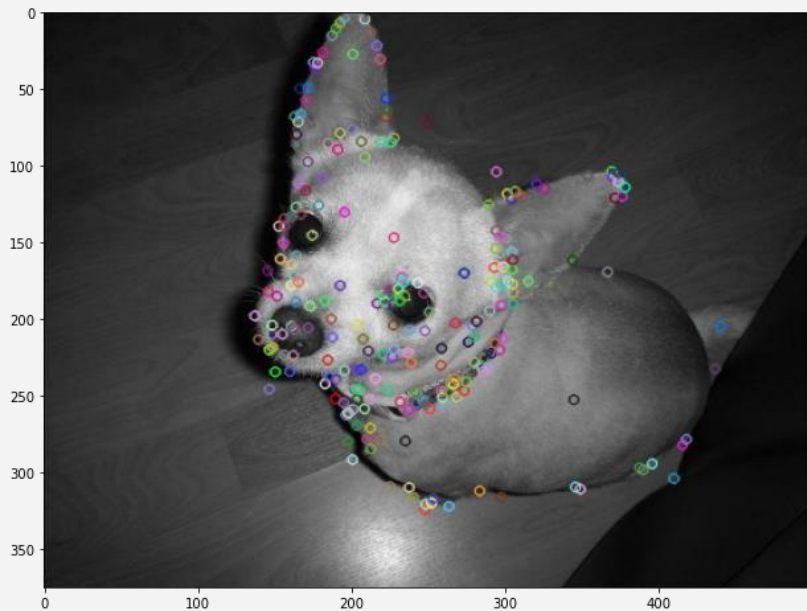


Caractéristiques d'une « bonne » feature:

- Répétable
- Distinctive
- Locale

# APPROCHE CLASSIQUE

## EXTRACTION DE FEATURES : SIFT



# APPROCHE CLASSIQUE

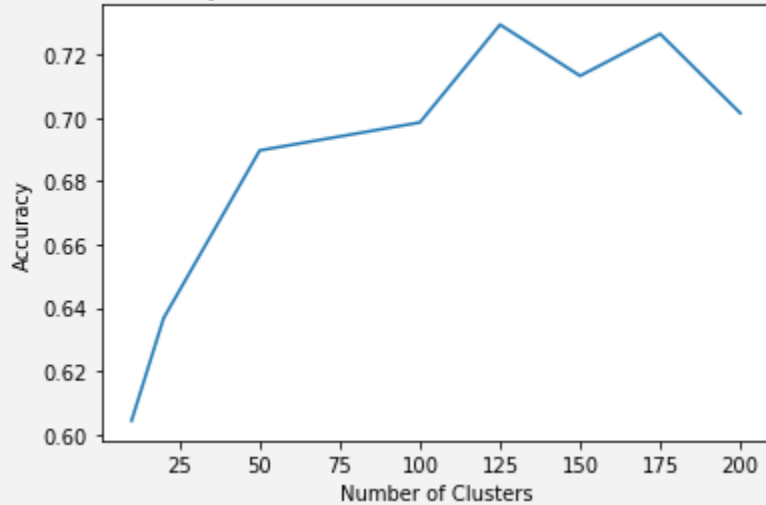
## CLUSTERING - KMEANS

- Nombre de descripteurs SIFT élevés :
  1. 2 races : 272515
  2. 3 races : 400 348
  3. 5 races : 659 616
- Utilisation de K-MEANS pour réduire la dimension
- Nombre de clusters testés:
  - 2 races : [10, 20, 50, 100, 125, 150, 175, 200]
  - 3 races : [150, 225, 300, 375, 450, 600, 750]
  - 5 races : [150, 225, 300, 375, 450, 600, 750]
- Pour chaque image, on détermine ensuite le nombre de descripteur dans chaque cluster : bag-of-words visuels

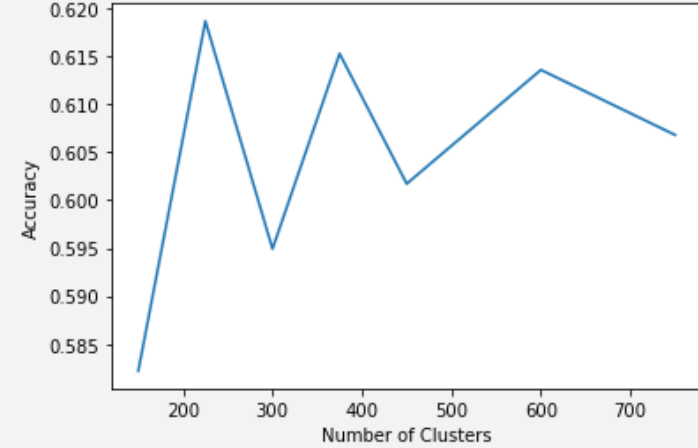
# NOMBRE DE CLUSTERS OPTIMAL

## APPROCHE CLASSIQUE

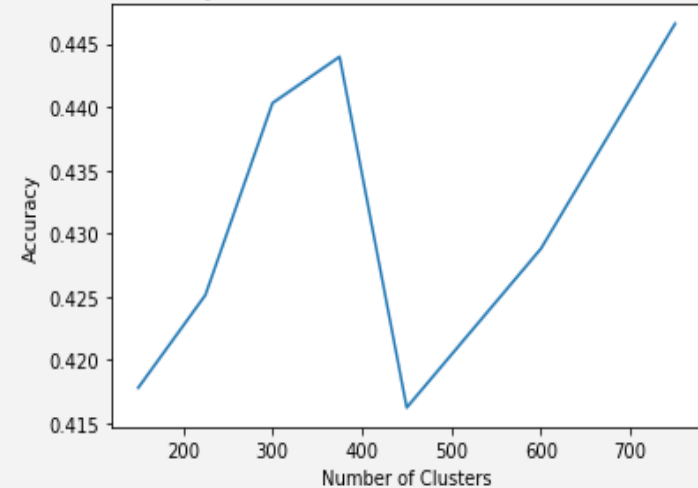
Accuracy en fonction du nombre de clusters : 2 Races



Accuracy en fonction du nombre de clusters : 3 Races



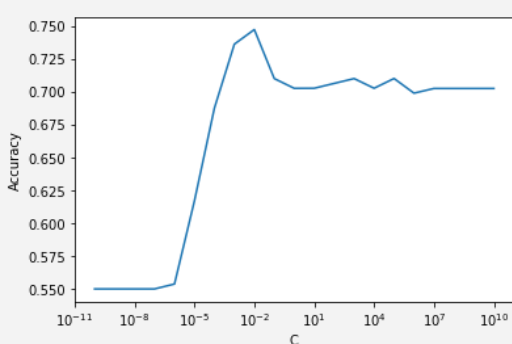
Accuracy en fonction du nombre de clusters : 5 Races



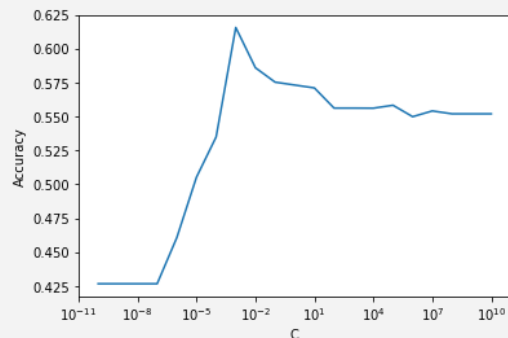
# APPROCHE CLASSIQUE

## CLASSIFICATION – RÉGRESSION LOGISTIQUE

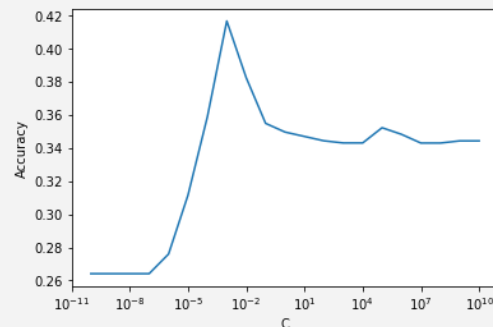
GRIDSEARCH pour trouver la meilleure valeur de C



2 RACES



3 RACES

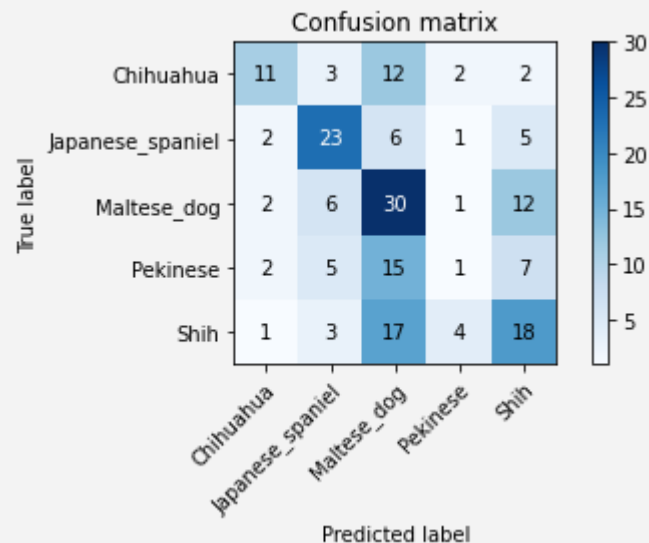
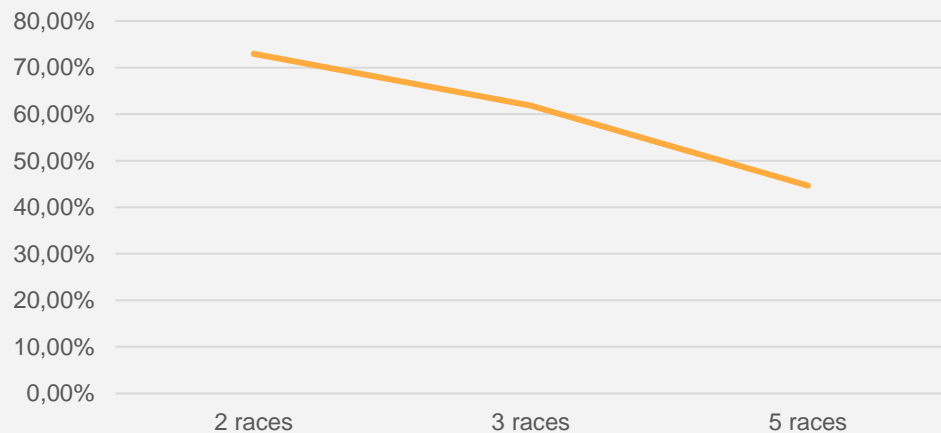


5 RACES

# APPROCHE CLASSIQUE

## CLASSIFICATION – RÉGRESSION LOGISTIQUE

Accuracy en fonction du nombre de Races (en %)



Matrice de confusions 5 races – Filtre Non Local Means

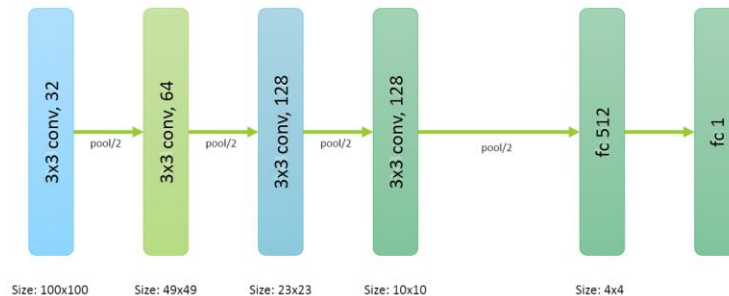


# APPROCHE CNN CLASSIQUE

CNN From Scratch , construction d'un réseau neuronal depuis zéro :

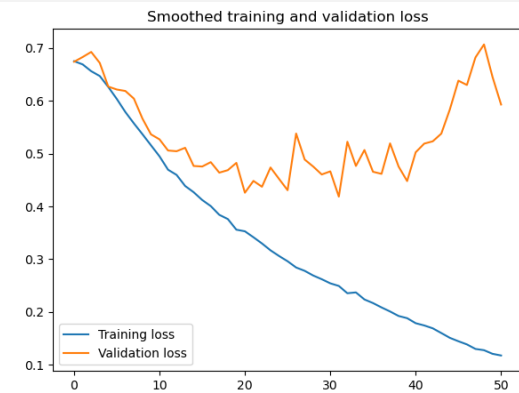
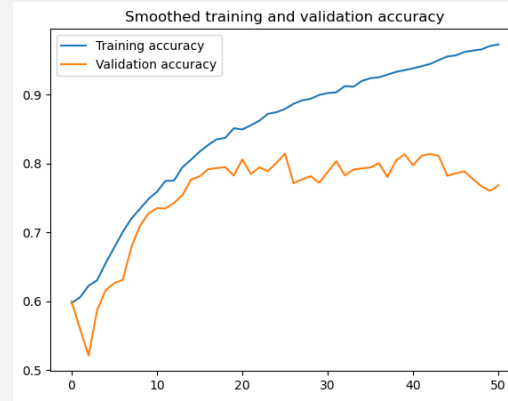
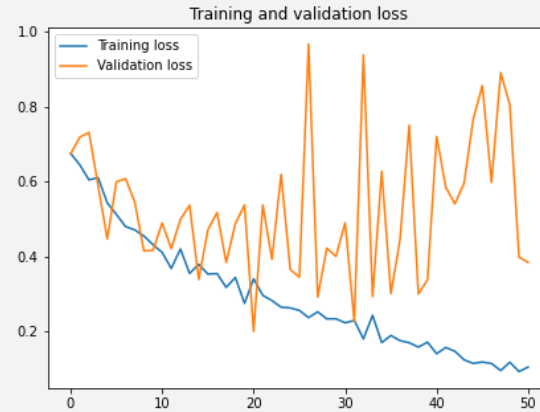
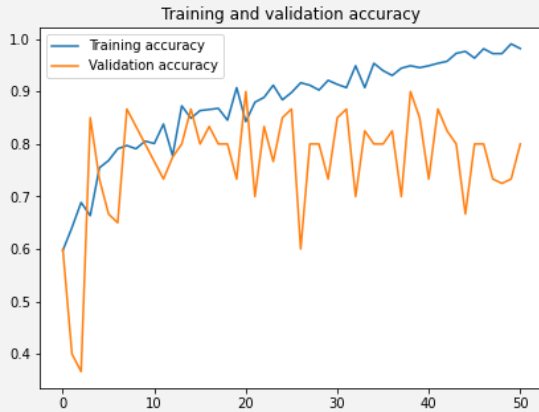
- Redimensionnement en  $224 \times 224$
- Utilisation des images en couleurs
- Séparation Train / Validation / Test : 70 % / 15 % / 15 %

## CNN from scratch



# APPROCHE CNN CLASSIQUE

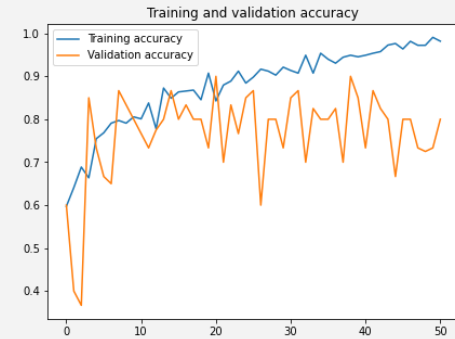
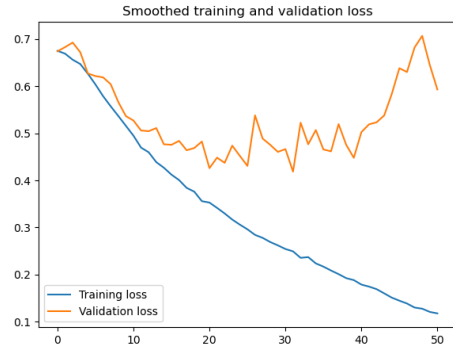
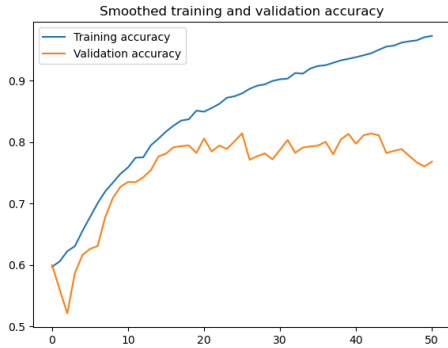
## CNN From Scratch - 2 Races



Accuracy sur le jeu de test : 78,4%

# APPROCHE CNN CLASSIQUE

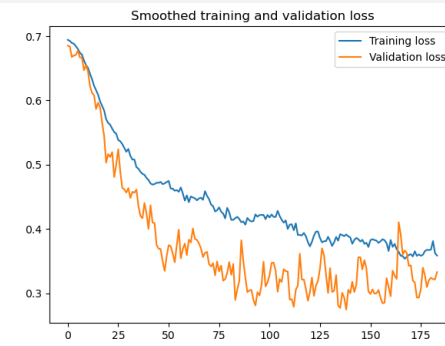
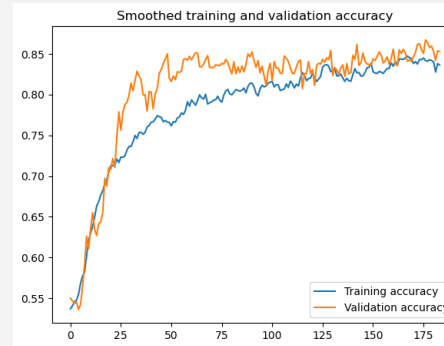
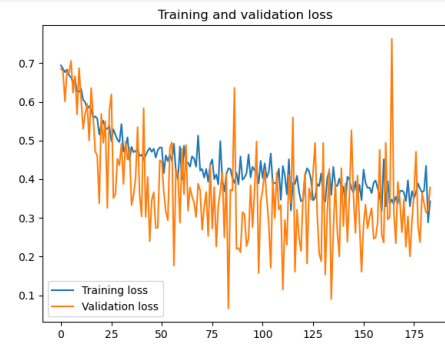
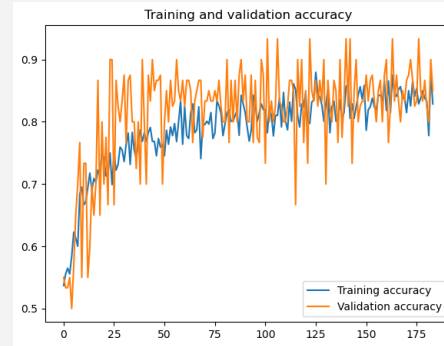
## SANS DATA AUGMENTATION



Accuracy sur le jeu de test : 78,4%

## CNN From Scratch – 2 Races OVERFITTING

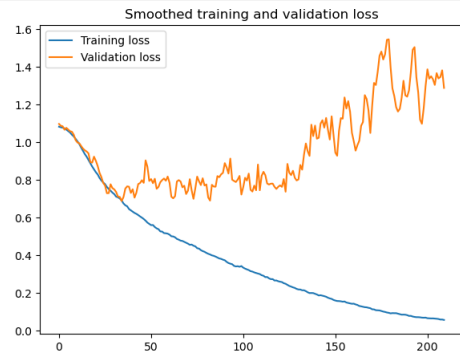
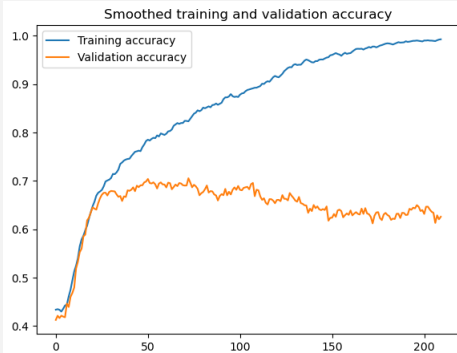
## AVEC DATA AUGMENTATION



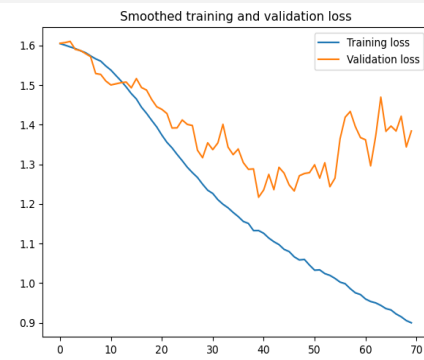
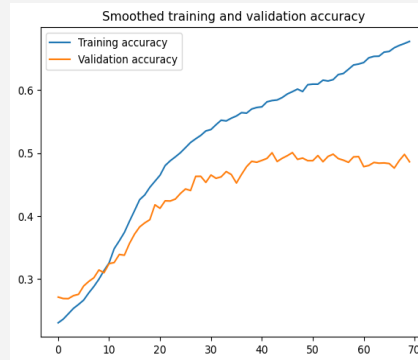
Accuracy sur le jeu de test : 80,3%

# APPROCHE CNN CLASSIQUE

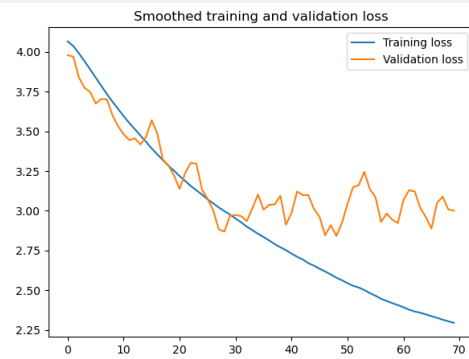
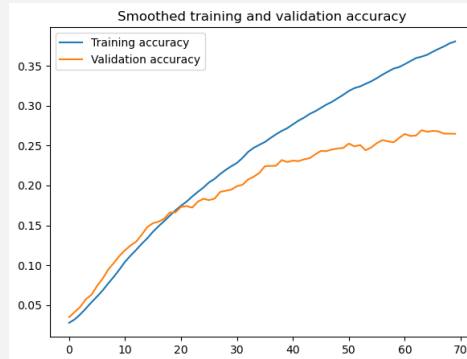
## 3 RACES : TEST ACCURACY 65,2%



## 5 RACES : TEST ACCURACY 42,6%

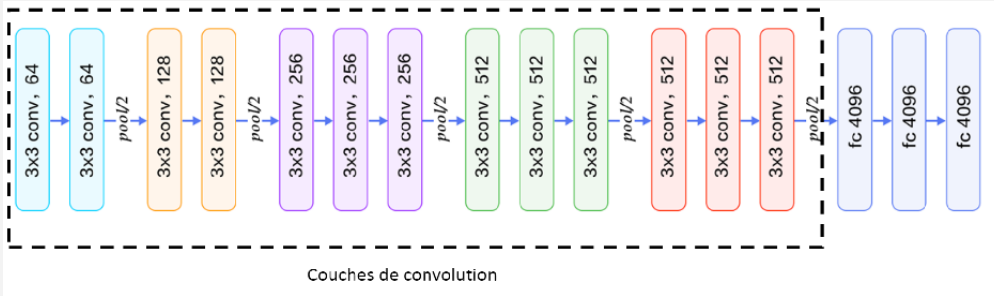


## 120 RACES : TEST ACCURACY 23,2%



# APPROCHE CNN TRANSFER LEARNING

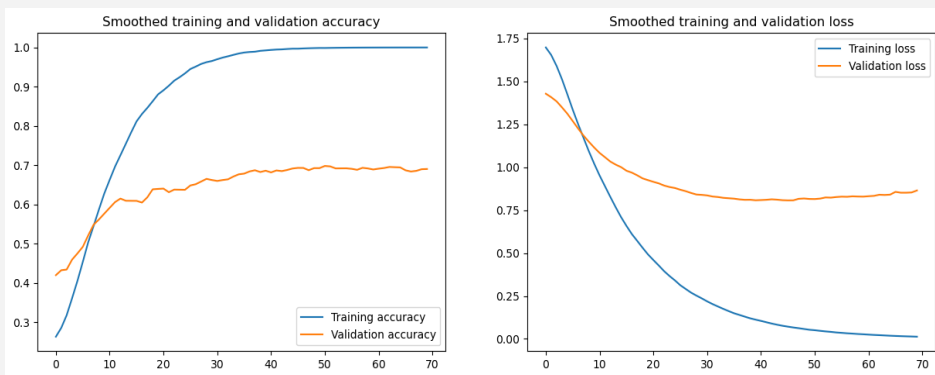
## VGG-16



# APPROCHE CNN TRANSFER LEARNING

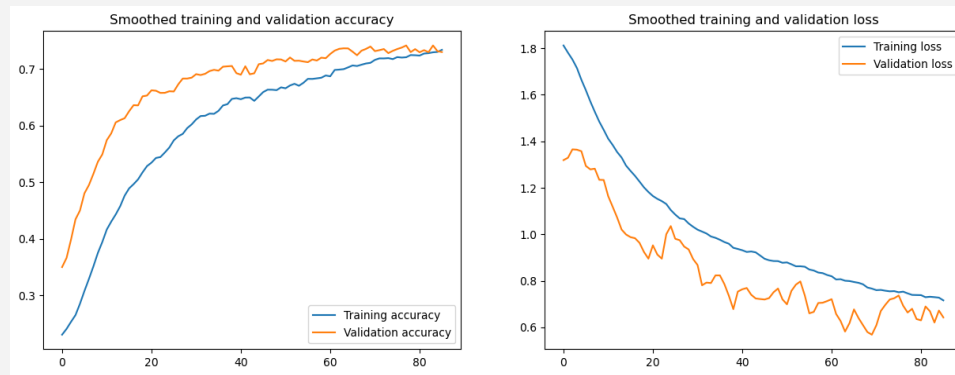
VGG-16

EXTRACTION DE FEATURES: 5 - RACES  
SANS DATA AUGMENTATION



Test accuracy : 65,7%

EXTRACTION DE FEATURES: 5 - RACES  
AVEC DATA AUGMENTATION



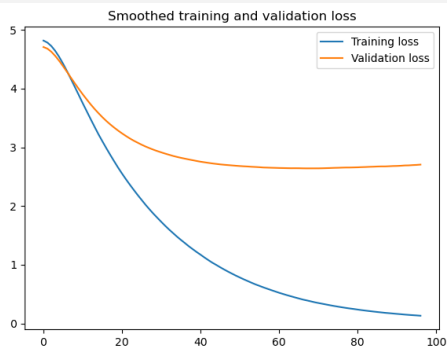
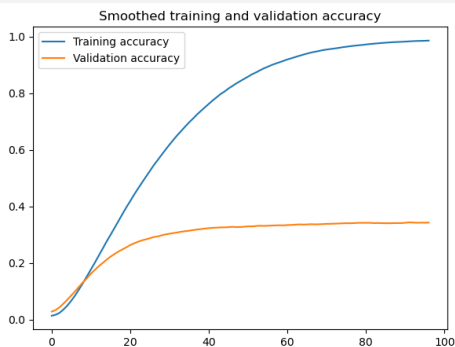
Test accuracy : 71,3%



# APPROCHE CNN TRANSFER LEARNING

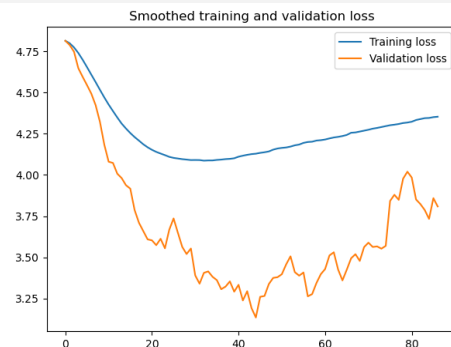
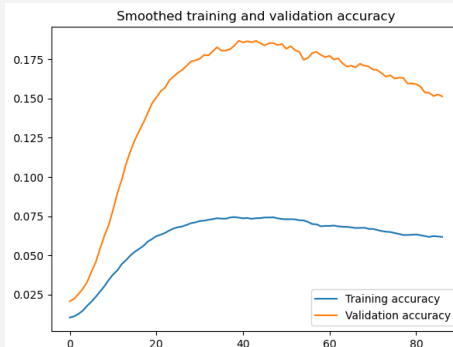
VGG-16

EXTRACTION DE FEATURES: 120 - RACES  
SANS DATA AUGMENTATION



Test accuracy : 34,4%

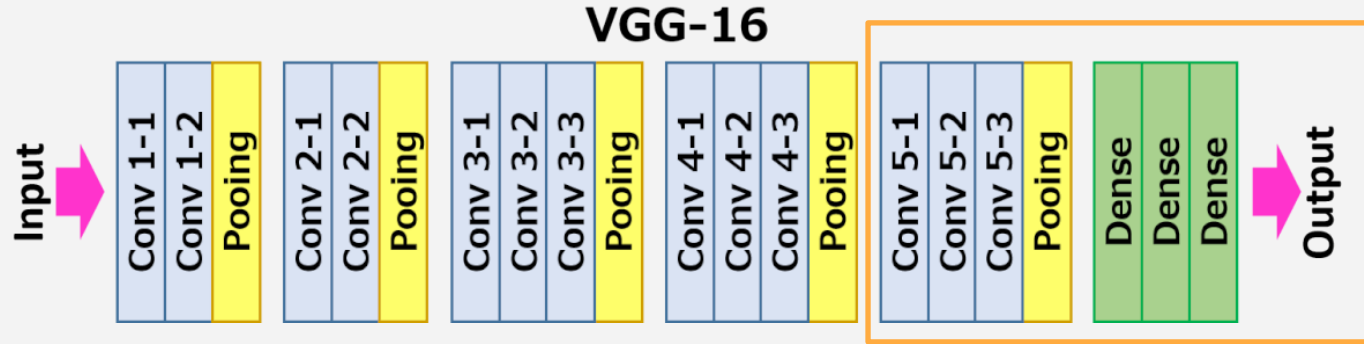
EXTRACTION DE FEATURES: 120 - RACES  
AVEC DATA AUGMENTATION



Test accuracy : 17,8%

# APPROCHE CNN TRANSFER LEARNING

## FINE TUNING



1. On repart des poids obtenus précédemment avec l'extraction de features.
2. On réentraîne les dernières couches de convolutions (Block 5)

# APPROCHE CNN TRANSFER LEARNING

## FINE TUNING

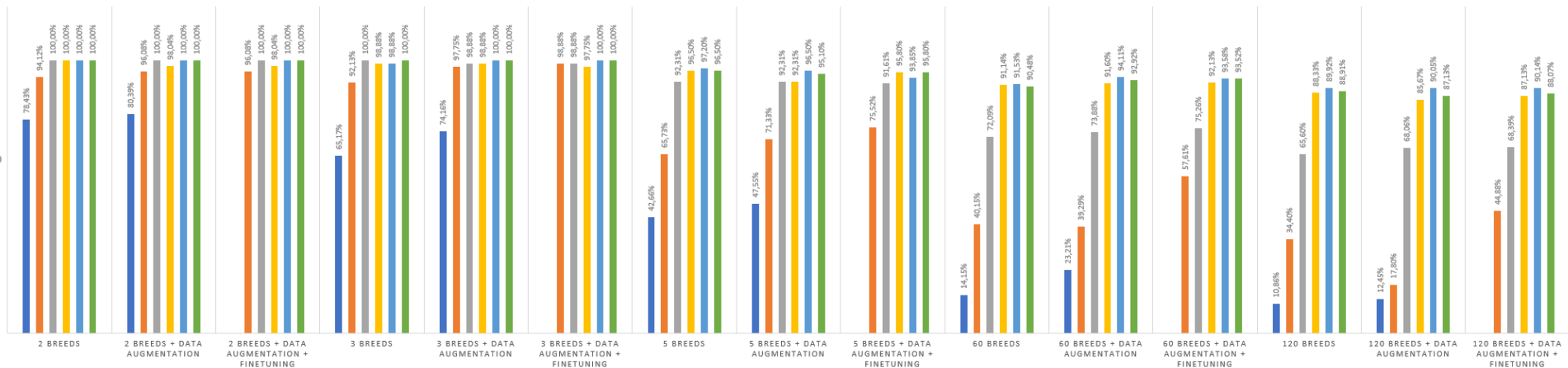
1. Même principe pour les autres modèles ResNet50v2 , Inceptionv3, Inceptionv4 et Xception.
2. Sauf que la structure de ces réseaux étant plus complexe, l'ont « freeze » 90% des couches et on réentraîne avec nos propres couches + 10% restant du modèle

# APPROCHE CNN TRANSFER LEARNING

## COMPARAISON DES MODÈLES

ACCURACY TEST SCORE PAR MODÈLES TESTÉS ET PAR NOMBRE DE RACES DE CHIENS (EN %)

■ CNN From Scratch ■ VGG16 ■ ResNet50V2 ■ InceptionV3 ■ InceptionV4 (InceptionResNet50V2) ■ Xception



# CONCLUSIONS ET PISTES D'AMÉLIORATIONS

1. LES RESEAUX CONVOLUTIONNELS SONT PLUS EFFICACES QUE LES METHODES CLASSIQUES.
2. DATA AUGMENTATION PERMET D'ÉVITER L'OVER-FITTING SI LE NOMBRE DE DONNEES EST FAIBLE.
3. DATA AUGMENTATION PEUT AVOIR DES EFFETS DÉLÉTÈRES SI LES TRANSFORMATIONS APPLIQUÉES SONT TROP IMPORTANTES.
4. LE FINE TUNING PERMET D'AMÉLIORER LES PERFORMANCES DU MODELE.
5. ALLER PLUS LOIN DANS LE FINE TUNING EN BLOQUANT MOINS DE COUCHES
6. FAIRE VARIER LE NOMBRE ET LA STRUCTURE DES COUCHES (DROPOUT, DENSE,etc...) SUPPLEMENTAIRES EN SORTIE DES MODELES PRÉ-ENTRAÎNÉS.
7. FAIRE DE LA DATA AUGMENTATION SUR QUELQUES EPOCHS PUIS CONTINUER D'ENTRAINER SANS DATA AUGMENTATION DANS LA FOULÉE.
8. AVOIR PLUS DE DONNÉES EN ENTRÉE.





MERCI DE VOTRE ATTENTION