



CLASSEZ DES IMAGES À L'AIDE D'ALGORITHMES DE DEEP LEARNING

BOURBON Vicente

PROBLÉMATIQUE

- Association de protection des animaux
- Augmentation de leur nombre de pensionnaires
- Mission: Algorithme de classification d'images de chien en fonction de la race
- Données: Stanford Dogs Dataset

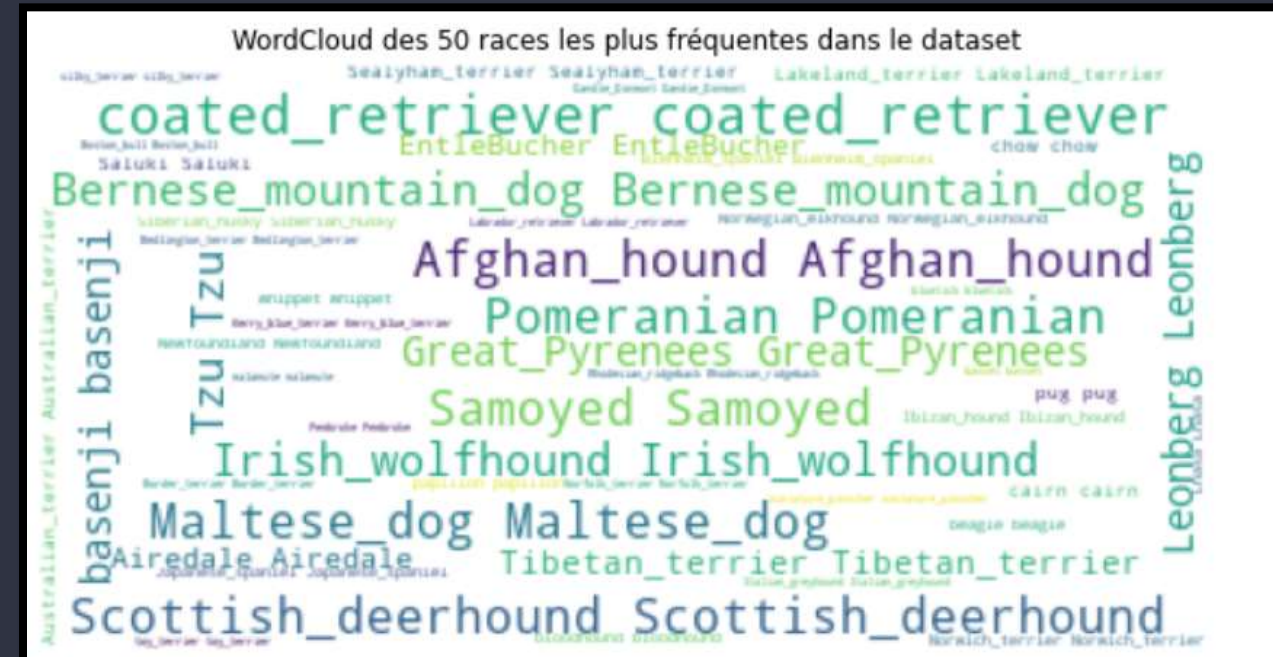


SOMMAIRE

- Exploration des données
- Pré-processing des images
- Classification d'images avec SIFT
- Construction d'un réseau CNN
- CNN et transfer learning
- Conclusion

I. EXPLORATION DES DONNÉES

- 20580 photos de chiens, 119 races différentes
- Entre 150 et 300 photos par race
- Photos couleurs avec dimensions différentes
- Suppression d'une photo disposant d'une couche Alpha





II. PRÉ- PROCESSING DES IMAGES

Exemples de
traitement

Image originale



Image réduite



Image originale en couleurs



Image transformée en noir et blanc



- Réduction de la taille des images
 - Réduction des coûts mémoire et calcul
 - Diminution qualité image
- Passage en noir et blanc
 - Pas d'apport d'information de la couleur

Image originale



Image assombrie



Image originale

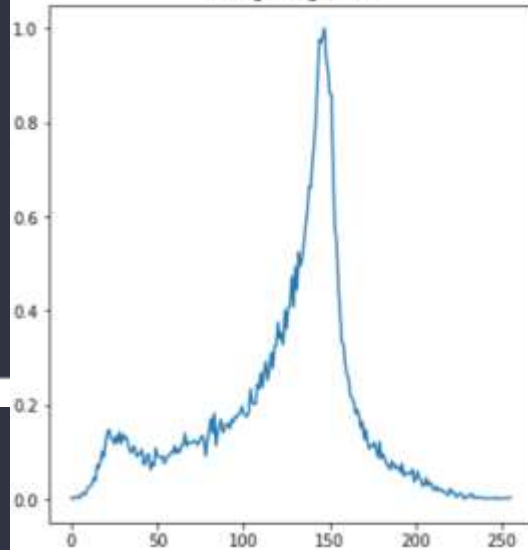
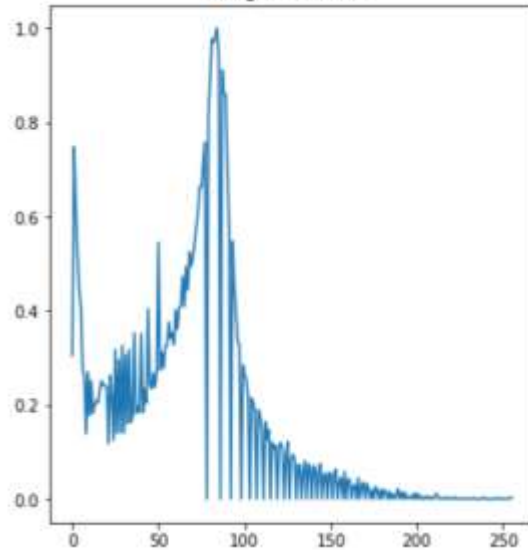


Image assombrie



- Exposition/Luminosité: répartition des valeurs des pixels sur la plage de valeurs
- Bonne exposition: répartition homogène sur toute la plage de valeurs

Image originale



Image après réduction du bruit



- Réduction du bruit: corriger la valeur des pixels
- Dé-bruitage de Bergman: Minimisation de la variation totale des valeurs de pixels de l'image

- Data Augmentation: création superficielle d'images à partir d'une seule image
- Changement orientation, échelle, centrage

Image d'exemple



Augmentation 0



Augmentation 1



Augmentation 2





III. CLASSIFICATION D'IMAGES AVEC SIFT

- Algorithme populaire pour la vision par ordinateur
- Extraction et description des features d'une image
- Utilisation limitée avec 3 races de chiens

- Détection de features: points d'intérêts et caractéristiques d'une image
- Description des features: vecteurs de représentation
- Création de visual words: clustering sur les features
- Construction de bag of features: matrice de comptage des visual words dans une image

Image originale avec réduction du bruit

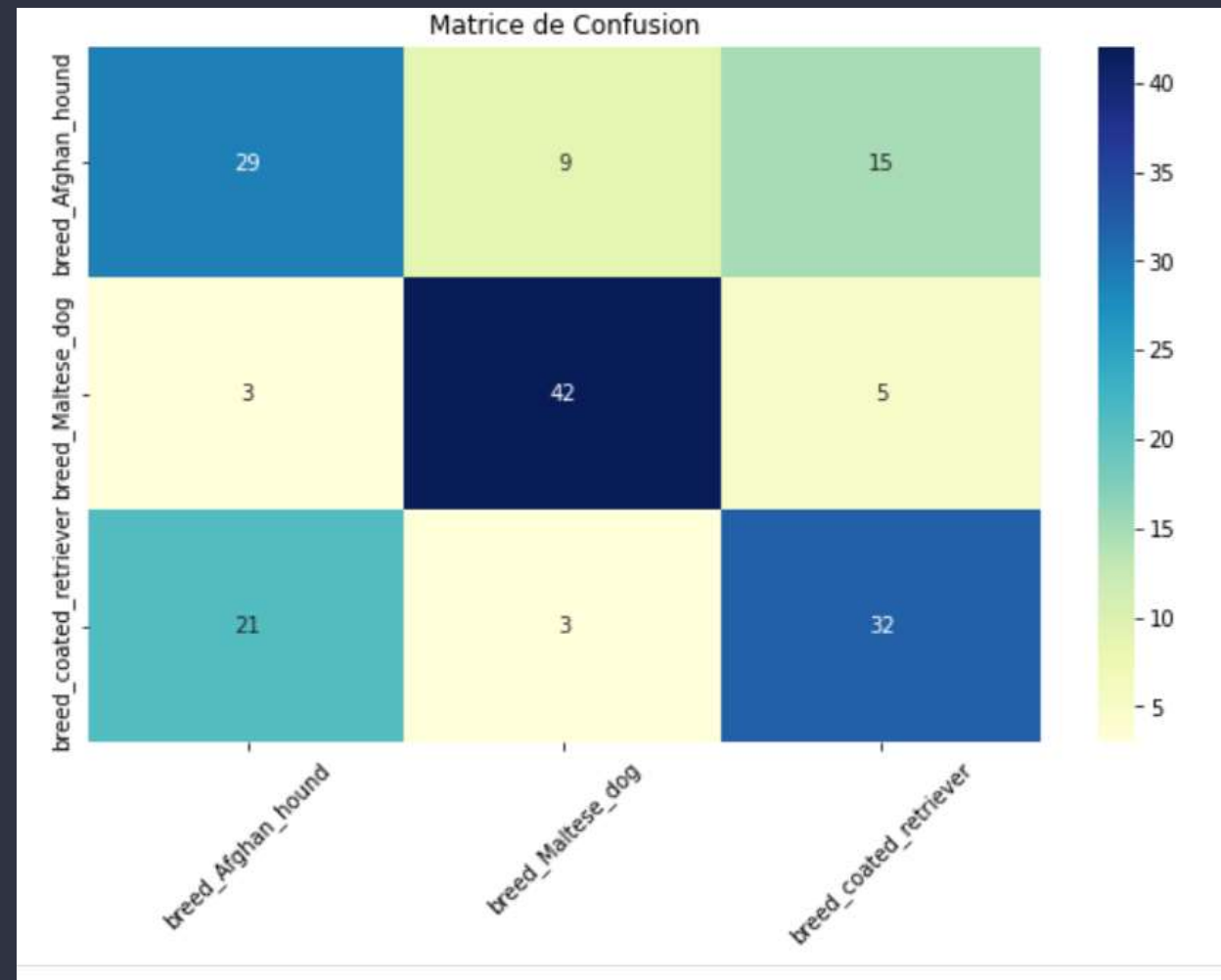


Image après détection des features



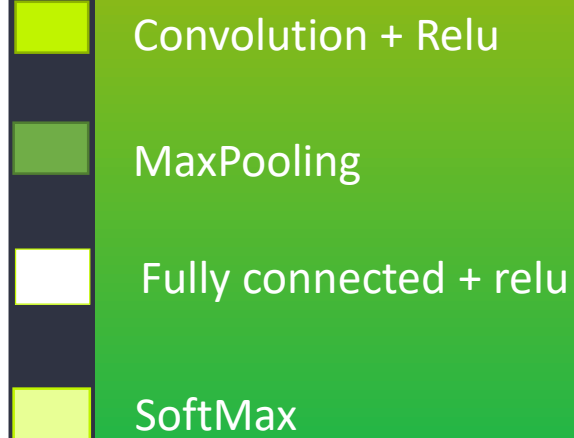
RÉSULTATS SIFT

- Classification avec méthode SVC multi-classes
- Visualisation des résultats avec matrice de confusion
- Pas de calcul d'accuracy car pas modèle final retenu
- Résultats correctes mais uniquement 3 races assez différentes



IV. CONSTRUCTION D'UN RÉSEAU CNN

- Extraction et description des features automatiques
- Structure en deux blocs principaux
 - Extraction, description des features
 - Classification
- Optimisation d'hyperparamètres
 - Fonction d'activation
 - Fonction d'optimisation
 - Nombre d'épochs

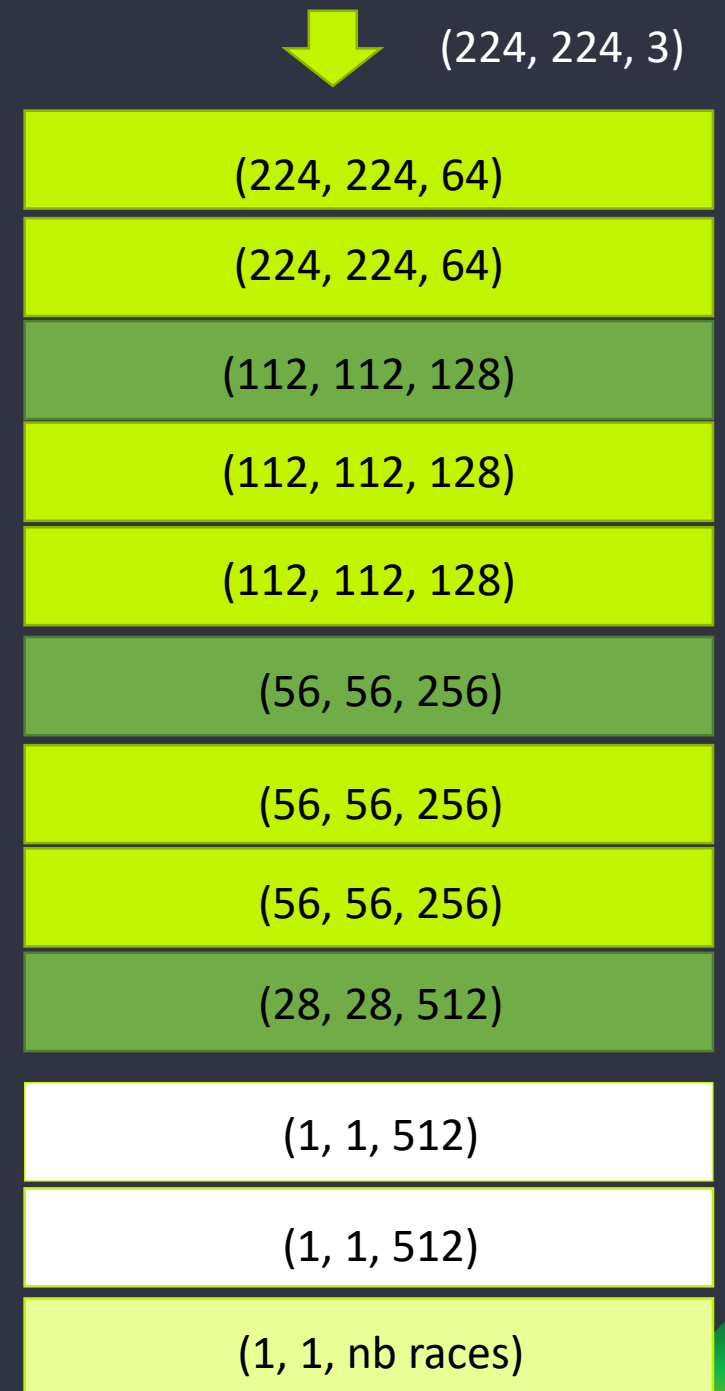


Convolution + Relu

MaxPooling

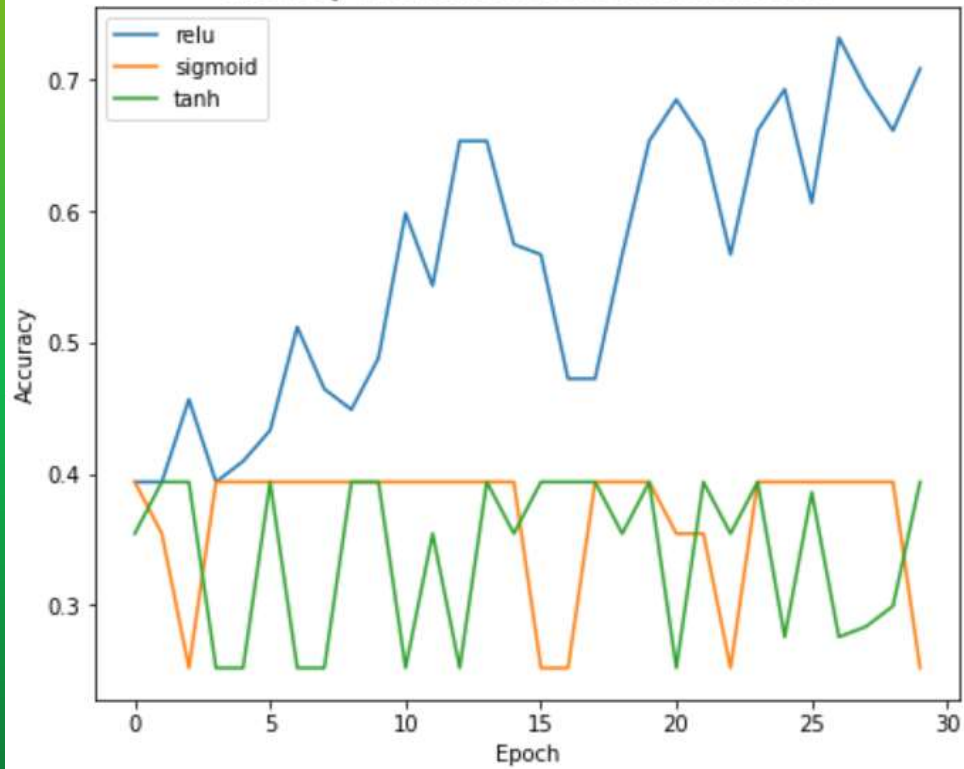
Fully connected + relu

SoftMax

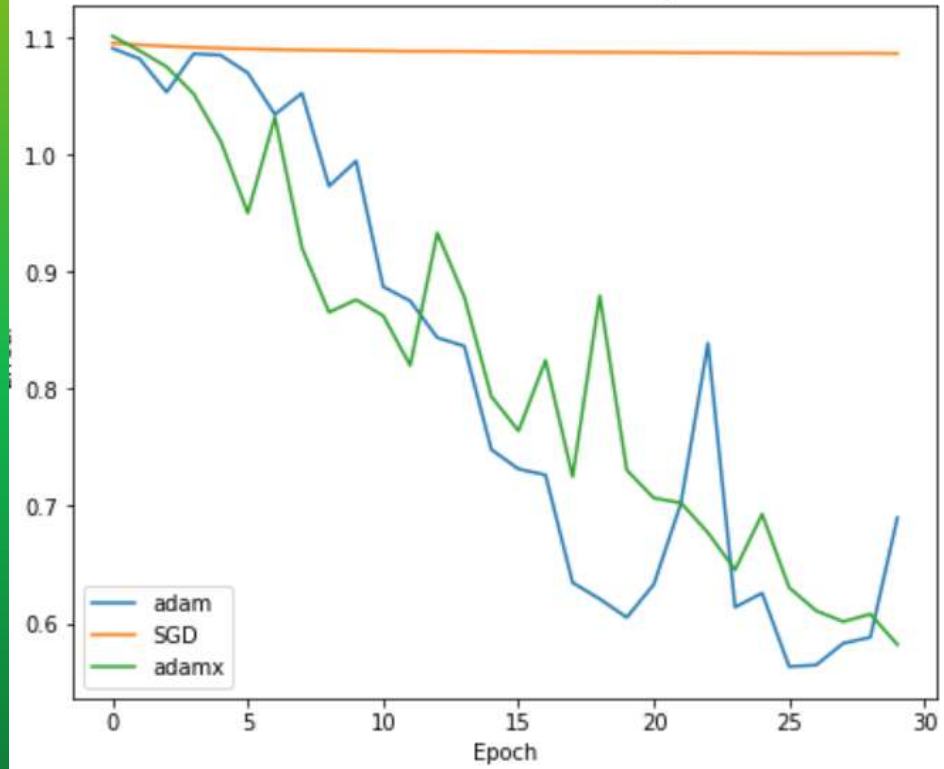


OPTIMISATION GRAPHIQUE

Accuracy en fonction de la fonction d'activation

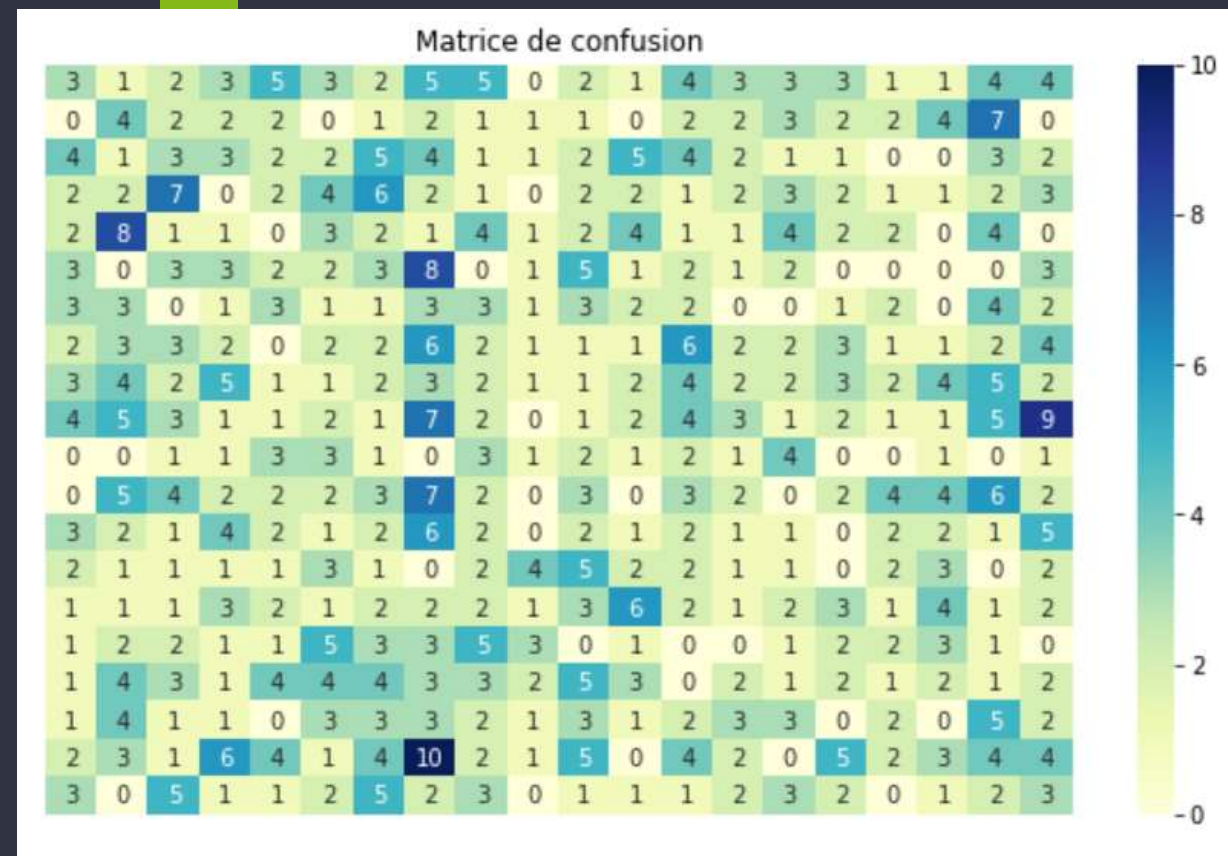


Erreur en fonction de la fonction d'optimisation



RÉSULTATS CNN

- Séparation jeux entraînement/test
- Visualisation des résultats avec matrice de confusion
- Accuracy:
 - 0.69 avec 3 races différentes
 - 0.28 avec 20 races
- Résultats insuffisants





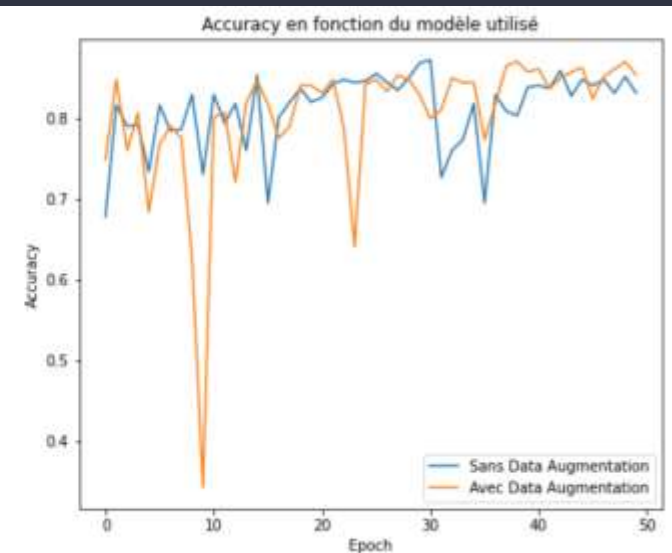
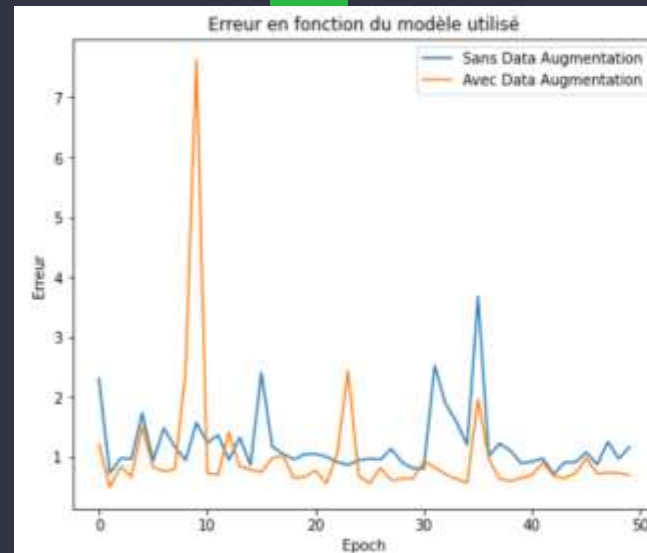
V. CNN ET TRANSFER LEARNING

- Entraînement CNN très couteux en temps calcul
- Transfer learning: utilisation d'un modèle pré-entraîné et modification d'un nombre plus ou moins important de couches
- 3 modèles comparés avec 15 races:
 - VGG16, ResNet50, Xception
- Comparaison des modèles:
 - Méthode transfer learning, accuracy, temps de calcul

MODÈLE RETENU

- Xception avec fine-tuning partiel
- Test de l'apport de la data augmentation:
 - Résultats similaires
 - Erreur plus faible avec
- Xception avec fine-tuning partiel et data augmentation
- Entraîné sur Kaggle avec toutes les photos: 5h

	Temps entraînement	accuracy train	accuracy val	accuracy test
VGG16	1379.809457	1.0	0.638734	0.661699
ResNet50	513.711118	0.746157	0.078212	0.117735
Xception	742.636825	0.993945	0.828678	0.862891



VI. CONCLUSION

- Nouvelles compétences
 - Travail sur images
 - SIFT
 - Deep learning
- Difficultés
 - Limites liées aux temps de calcul
- Pistes d'amélioration(avec ressources suffisantes)
 - Modèle SIFT optimisé
 - Travail de recadrage des photos