

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université Saad Dahlab- Blida 1
Faculté des Sciences
Département d'Informatique



Mémoire de Master en Informatique

Option : Ingénierie des logiciels

Thème

Impact de la variation de la distribution de luminosité des données sur les modèles d'apprentissage profond : Application à la prédiction de l'Âge, du genre et l'ethnicité d'un individu à partir d'une image

Réalisé par :

Khadidja Boudjafer & Ouafaa Bourguieg

Promoteurs :

Pr. Djamel BOUCHAFFRA, Directeur de Recherche, CDTA, Alger.

Dr. Fayçal YKHLEF, Maître de Recherche A, CDTA, Alger.

Co-promoteur :

Dr. Soraya Cheriguene, Maître de Conférences B, Université de Blida 1, Blida

Année universitaire : 2022-2023

Remerciements

Nous exprimons notre profonde gratitude à Allah le Tout-Puissant pour nous avoir accordé la santé et la volonté nécessaires pour mener à bien ce projet.

Nous exprimons notre sincère reconnaissance à toutes les personnes qui ont contribué à la réalisation de ce projet. Tout d'abord, nous souhaitons remercier nos promoteurs, le professeur Djamel Bouchaffra, directeur de la division Architecture des Systèmes et Multimédia, et le Dr Faycal Yekhlef, Maître de Recherche A au CDTA, pour leur orientation, leurs compétences et leur confiance en nous. Grâce à leurs conseils et à leurs critiques constructives, nous avons élaboré et développé des solutions innovantes. Je tiens à exprimer ma profonde gratitude à Rania Boutrif pour le soutien essentiel dans la réalisation de notre projet universitaire. Ta créativité, détermination et collaboration ont été remarquables.

Nous exprimons une grande reconnaissance à notre Co-promotrice, le Dr Cheriguene, Maître de Conférences B à l'université Saad Dahlab, Blida 1, pour sa disponibilité, sa précieuse contribution, ses conseils et son expertise.

Nous exprimons notre gratitude à l'ensemble du personnel du CDTA, en particulier le personnel de la division Architecture des Systèmes et Multimédia, pour leur hospitalité et leur assistance, ainsi qu'au corps enseignant et administratif de l'université Saad Dahlab 1 pour leur contribution au succès de notre parcours universitaire.

Nous tenons également à remercier les membres du jury d'avoir accepté d'examiner et d'évaluer notre travail. Nous exprimons notre profonde reconnaissance et notre gratitude envers tous ceux qui nous ont apporté leur aide et leur encouragement pendant les moments difficiles.

Dédicace

Je dédie ce travail a :

Moi-même comme hommage et à ma détermination, à mon amour pour la connaissance et à ma volonté de toujours aller plus loin. Je m'engage à continuer d'apprendre, de grandir et de contribuer au monde qui m'entoure. Je me félicite d'avoir eu l'opportunité de mener cette recherche,

Je tiens à dédier ces mots à ma très chère maman, qui m'a soutenu de manière inconditionnelle et m'a prodigué des conseils précieux. Son écoute attentive et son suivi permanent ont toujours été présents, me donnant la certitude que je pouvais compter sur elle. Son sourire lumineux a su raviver l'espoir en moi lorsque je l'avais parfois perdu. Je perçois encore sa voix résonner dans ma tête, m'encourageant à me surpasser et à donner le meilleur de moi-même. Je vous remercie du fond du cœur de m'avoir offert une éducation excellente, qui me sert énormément et continuera à me servir tout au long de ma vie.

Je souhaite également exprimer ma gratitude envers le meilleur des pères mon très chère papa, qui a consenti d'innombrables sacrifices pour veiller à ce que je ne manque de rien. Son acharnement à me transmettre sa patience, son sens des responsabilités et son éthique du travail est inestimable. Je réalise désormais toute la richesse qui se cache derrière ses paroles.

Vous, mes parents, êtes les piliers de ma vie. Votre amour, votre soutien et vos enseignements ont façonné la personne que je suis aujourd'hui, et je vous en suis infiniment reconnaissant(e). Votre dévouement et vos valeurs m'inspirent chaque jour à être la meilleure version de moi-même.

A Mes frères et sœurs Nassima ,Adel, Walid et Zakaria.

A mes neveux et nièces adorés : Iselm, Iyade , Sadjia , Khawla et Achewak .

A mon cher fiancé Amine, Merci pour tout ton soutien morale et physique ,

Mes amis et collègues de travail :

Feriel Kessi, Rania, Boutrif Fadhila Brihi, Manel Derrouazine, Ouafaa Bourguieg, Azhar Boudissa, Nour El Yakine Krim. Qui partagent avec moi les meilleurs souvenirs de ma vie

J'ai une profonde reconnaissance pour leur inestimable appui et assistance.

Mes amis je vous dédie ce travail et je vous souhaite une vie pleine de santé et de bonheur.

Khadidja

Dédicace Du profond de mon cœur, je dédie ce travail à tous ceux qui me sont chers,

Au meilleur des pères

Tu me dirigeais toujours vers le bon chemin, Merci pour les valeurs nobles, l'éducation et le soutien permanent venu de toi

A ma très chère maman

Tu représentes pour moi le symbole de la bonté par excellence, la source de tendresse et l'exemple du dévouement, tu n'as pas cessé de m'encourager et de prier pour moi.

Que vous trouvez en moi la source de leur fierté, Qu'allah le tout puissant vous préserve, vous accorde santé et bonheur.

À mes chers frères Abd Djalil, Abd El Kader, Abd Salam et Bachir Vous êtes mes compagnons de vie, mes soutiens inébranlables et mes amis les plus proches. À travers les hauts et les bas, vous avez été là pour moi, me prodiguant votre soutien, vos conseils et votre réconfort.

À mon cher fiancé, Merci d'être là pour moi, de me soutenir dans mes rêves et mes aspirations, et de partager les hauts et les bas de la vie. Tu es mon pilier, ma source de réconfort et ma plus grande inspiration.

À tous ma famille, mes grands-parents, oncles, tantes, cousins et cousines, Chaque membre de ma famille joue un rôle unique et irremplaçable dans ma vie.

À mes copines d'amour Hayat et Neila À toutes les deux, je tiens à vous remercier du fond du cœur pour votre amitié sincère et votre présence constante.

À mes camarades de la promo informatiques : Pour tous les bons moments qu'on a passé ensemble (Rania,Manel,Fadila,Marwa,khadidja). Et a tous mes collègues.

À mon binôme khadidja et sa famille.

Ouafaa

Résumé

La puissance de généralisation des classifieurs est directement influencée par les données d'entraînement. Il est bien établi que les modèles de classification donnent de meilleurs résultats lorsqu'ils sont entraînés et évalués sur des données suivant la même loi de distribution. Cependant, en pratique, les données utilisées ne sont pas toujours conformes à une même loi de distribution, et un tel changement peut impacter la puissance de généralisation des modèles. Notre objectif dans ce mémoire est de mesurer l'impact de la variation de la loi de distribution de la luminosité des images sur les modèles d'apprentissage profond. Nous nous sommes focalisés sur la prédiction simultanée de l'âge, du genre et de l'ethnicité d'un individu à partir des images faciales. Nous avons limité notre projet à l'utilisation des modèles neuronaux à plusieurs sorties connus sous le nom anglo-saxon de Convolutional Neural Network (CNN) et Residual Network (ResNet). Les résultats obtenus en utilisant trois bases de données (UTKFace, FairFace et FFHQ) nous montrent que le changement de distribution a un impact significatif sur la prédiction de l'âge, du genre et de l'ethnicité en utilisant les modèles d'apprentissage profond.

Mots clés :

Distribution de données, Apprentissage profond, prédiction de l'âge, du genre, et de l'ethnicité, Réseau de neurones convolutifs, Réseaux résiduels profonds.

Abstract

The generalization power of classifiers is directly influenced by the training data. It is well established that classification models perform better when trained and evaluated on data that follows the same distribution. However, in practice, the data used is not always consistent with a single distribution, and such a change can impact the generalization power of the models. Our objective in this thesis is to measure the impact of varying the distribution of image brightness on deep learning models. We focus on the simultaneous prediction of age, gender, and ethnicity of an individual based on facial images. We have limited our project to the use of neural models with multiple outputs, known as Convolutional Neural Networks (CNN) and Residual Networks (ResNet). The results obtained using three databases (UTKFace, FairFace, and FFHQ) show that distributional changes have a significant impact on the prediction of age, gender, and ethnicity using deep learning models.

Keywords:

Data distribution, Deep learning, Age prediction, Gender prediction, Ethnicity prediction, Convolutional neural networks, Deep residual networks.

ملخص

قوة تعميم المصنفات متأثرة مباشرة ببيانات التدريب. من المعروف جيداً أن نماذج التصنيف تعطي نتائج أفضل عندما يتم تدريبها وتقييمها على بيانات تتبع نفس قانون التوزيع. ومع ذلك، في الواقع، لا تكون البيانات المستخدمة دائماً متوافقة مع نفس قانون التوزيع، ويمكن أن يؤثر مثل هذا التغيير على قوة تعميم النماذج. هدفنا في هذه الأطروحة هو قياس تأثير تغيير قانون توزيع سطوح الصور على نماذج التعلم العميق. ركزنا على التنبؤ المتزامن بالعمر والجنس والأصل العرقي لفرد ما باستخدام الصور الوجيهة. حددنا مشروعنا لاستخدام النماذج العصبية متعددة النواتج المعروفة باسم شبكات التعصب المحاكية للتحويلات تظهر النتائج التي تم الحصول عليها باستخدام ثلاث قواعد بيانات (ResNet) وشبكات التعصب المتبقية (CNN) المكررة أن تغيير التوزيع يؤثر بشكل كبير على التنبؤ بالعمر والجنس والأصل العرقي (UTKFace و FairFace و FFHQ) باستخدام نماذج التعلم العميق.

كلمات مفتاحية

توزيع البيانات، التعلم العميق، تنبؤ العمر، تنبؤ الجنس والأصل العرقي، شبكات العصبونات التكرارية، الشبكات العصبية العميقة المتبقية.

Table de matières

Remerciement	III
Dédicace	III
Résumé	V
Abstract	VI
ملخص	VII
Table de matières	VIII
Liste des figures	X
List des équations	XI
List des tableaux	XII
Liste des acronymes et abréviations	XIII
Introduction générale	16
Chapitre 1 : Exploration de l'apprentissage automatique, analyse des distributions des données et la prédiction de l'âge, genre et ethnicité	17
1.1 Introduction	17
1.2 Exploration de l'apprentissage automatique	18
1.2.1 Aperçu sur l'apprentissage automatique	18
1.2.2 Types d'algorithmes d'apprentissage automatique	19
1.1.1 Réseaux de neurones	20
1.2 Analyse des distributions de données.....	25
1.2.1 Biais.....	26
1.2.3 Manipulation d'image	26
1.2.4 Variation des contextes.....	26
1.2.5 Variation des attributs visuels	26
1.2.6 Distribution de luminosité	26
1.3 Détection de l'âge, le genre et l'ethnicité à partir des attributs faciaux	27
1.3.1 Détection de l'âge.....	27
1.3.2 Détection du genre.....	27
1.3.3 Détection de l'ethnicité.....	28
1.3.4 Etat de l'art sur la prédiction de l'âge, du genre et l'ethnicité	28
1.3.5 Discussions et limitation.....	30
1.6 Conclusion	30

Chapitre2 : Mesure de l'impact de la distribution de donnée sur les modèles d'apprentissage profond 31

2.1	Introduction	31
2.2	Bases de données	31
2.2.1	UTKFace	32
2.2.2	Fairface.....	32
2.2.3	Flickr-Faces-HQ dataset.....	33
2.3	Méthodologie adoptée	34
2.3.1	Choix des données.....	34
2.3.2	Prétraitement des données	35
2.3.3	Calcul de luminosité	35
2.3.4	Sélection des modèles d'apprentissage profond	36
2.3.5	Évaluation des modèles avec différentes distributions.....	36
2.3.6	Analyse des résultats	37
2.5	Conclusion.....	Erreur ! Signet non défini.

Chapitre 3 : Implémentation résultats et expérimentations. 39

3.1	Introduction	39
3.2	Matériels et outils de développement	39
3.2.1	Langage de programmation.....	39
3.2.2	Roboflow	40
3.2.3	Plateforme de développement	40
3.3	Implémentation du système	41
3.3.1	Choix et acquisition des données	41
3.3.2	Préparation des données	41
3.3.3	Modélisation des réseaux de neurones	48
3.4	Résultats et tests.....	51
3.4.1	CNN multi-output.....	51
3.4.2	Résultat de ResNet multi-output	55
4.2	Interprétations.....	56
4.3	Conclusion.....	57

Conclusion Général.....	58
-------------------------	----

Bibliographie	58
---------------------	----

Liste des figures

<i>Figure 1-1</i> :Schéma présentatif résumant l'apprentissage automatique.....	18
<i>Figure 1-2</i> : Présentation de l'apprentissage supervisé.....	19
<i>Figure 1-3</i> : Parties du CNN.....	22
<i>Figure 1-4</i> : Présentation de l'architecture des couches d'un CNN-Multi-output.	23
<i>Figure 1-5</i> : Architecture d'un ResNet 34.....	25
<i>Figure 2-1</i> : Échantillon des photos de la dataset UTKFace.....	32
<i>Figure 2-2</i> : Échantillon des photos de la dataset FairFace.....	33
<i>Figure 2-3</i> : Échantillon des photos de la dataset FFHQ.	33
<i>Figure 2-4</i> : Présentation d'ordre des étapes logiques de méthodologie.....	34
<i>Figure 3-1</i> : Échantillon des images de mauvaise qualité.	42
<i>Figure 3-2</i> : Image représentatif d'une image aberrantes.....	43
<i>Figure 3-3</i> : Représentation d'une image d'un visage partiellement caché.	43
<i>Figure 3-4</i> : Représentation d'images double avec étiquette différente.....	43
<i>Figure 3-5</i> : Répartitions des étiquettes de classement des images.....	45
<i>Figure 3- 6</i> : Histogrammes de la densité de probabilité de la luminosité des pixels.....	47
<i>Figure 3-7</i> : Représentations des distributions et de l'analyse des données avant et après le prétraitement.....	52
<i>Figure 3-8</i> : Graphes d'apprentissage pour l'entraînement du CNN fait sur la data UTKFace....	52
<i>Figure 3-9</i> : Graphes d'apprentissage pour l'entraînement du CNN fait sur la data FairFace	52
<i>Figure 3-10</i> : Graphes d'apprentissage pour l'entraînement du CNN fait sur la data FFHQ	52
<i>Figure 3-11</i> : Graphes d'apprentissage pour l'entraînement du ResNet fait sur la data UTKFace	52
<i>Figure 3-12</i> : Graphes d'apprentissage pour l'entraînement du ResNet fait sur la data UTKFace	52

List des équations

Equation (2.1) : Accuracy	36
Equation (2.2) : F1-Score	36

List des tableaux

Tableau 1-1: Comparaison entre un CNN un avec une seul sortie et plusieurs.....	24
Tableau 1-2 : Parcours des travaux connexes	28
Tableau 3-1 : Répartition des âges par classe d'âge	44
Tableau 3-2 : Répartition ethnique par classe d'ethnicité	44
Tableau 3-3 : Répartition des genres par classe de genre	44
Tableau 3-4 : Les couches et les hyper paramètre de l'architecture utilisé du CNN	49
Tableau 3-5 : Résultats d'exactitude pour la prédiction de l'âge, du genre et de l'ethnicité au niveau de train et test sur les ensembles de données UTK, FairFace et FFHQ	53
Tableau 3-6 : Résultats du F-score pour la prédiction de l'âge, du genre et de l'ethnicité sur les ensembles de données UTK, FairFace et FFHQ	54
Tableau 3-7 : Résultats d'exactitude pour la prédiction de l'âge, du genre et de l'ethnicité au niveau de train et test sur les ensembles de données UTK, FairFace	56
Tableau 3-8 : Résultats F1 pour la prédiction de l'âge, du genre et de l'ethnicité au niveau de train et test sur les ensembles de données UTK, FairFace	56

Liste des acronymes et abréviations

ACC: Accuracy.

ANN: Artificial neural network.

BD : Base de données.

CelebAMask-HQ : Celebrity Attribute Mask High Quality.

CELEB A : Celebrity Attribute.

CNN: Convolutional neural network.

DEX : Deep EXtraction.

FFHQ : Flickr-Faces-HQ.

Flickr YFCC-100M : Yahoo Flickr Creative Commons 100 Million.

FFHQ: (Flickr-Faces-HQ).

IA: Intelligence Artificielle.

IMDB-Wiki : Internet Movie Database-Wikipedia.

GPU: Graphics Processing Unit.

LFW : Labelled Faces in the Wild.

MAE : Mean Absolute Error

NVIDIA FFHQ : NVIDIA Flickr-Faces-HQ (Images haute qualité de visages provenant de Flickr, développé par NVIDIA).

PNG: Portable Network Graphics.

ResNet: Résiduel Network.

RELU : Rectified Linear Unit.

RVB: Rouge, Vert, Bleu.

SVM : Support Vector Machine.

TPU : Tensor Processing Unit.

UTK: University of Tennessee at Knoxville.

Introduction générale

Ces dernières années, les modèles d'apprentissage spécifiquement les réseaux de neurones suscitent un vif intérêt dans le contexte de renforcer, consolider et éliminer la sensibilité de ces modèles par rapport à la distribution de données. Conjointement, la reconnaissance automatique des traits démographiques tels que le l'âge, le genre et l'ethnicité, à partir des images faciales, présente un domaine important. Elle a attiré l'attention des chercheurs dans plusieurs domaines.

De ce fait, de nombreux travaux ont été consacrés au développement de nouvelles techniques pour améliorer les performances de la prédiction automatique de ces caractéristiques humaines. Cependant, de nombreux facteurs peuvent affecter les performances de ces systèmes de classification. L'estimation de l'âge, de genre et de l'ethnicité à partir des images de visage est sensible à certains facteurs, tels que la pose et l'éclairage variable et la résolution des images faciales. En outre, la performance de prédiction des attributs démographiques dépend aussi de la qualité et de la distribution des données d'entraînement et de généralisation et la sensibilité des modèles choisis.

Ce mémoire s'intéresse à l'impact de la variation de la distribution des données sur la performance des modèles de deep learning dédiés à la prédiction de l'âge, du genre et de l'ethnicité à partir de visage d'un individu. La problématique de ce projet est plus spécifique, car elle se concentre sur un cas plus restreint qui traite la diversité de la luminosité des images et son influence sur la performance des modèles. Cette étude apportera une contribution importante à la recherche en apprentissage automatique.

Afin d'atteindre le but de notre travail, l'organisation de notre mémoire sera comme suit :

- **Chapitre 01** : introduit des généralités sur l'apprentissage automatique, en mettant l'accent sur les approches de l'apprentissage profond. Ce chapitre concentre aussi sur la distribution des données et la prédiction de l'âge, du genre et de l'ethnicité en exposant quelques travaux antérieurs.
- **Chapitre 02** : présente la méthodologie adoptée ainsi que les différentes méthodes utilisées pour la conception et l'évaluation de notre travail.
- **Chapitre 03** : est consacré à l'implémentation et la démonstration des résultats obtenus.

Enfin, nous concluons par une conclusion générale et des perspectives du projet.

Chapitre 1: Exploration de l'apprentissage automatique, analyse des distributions des données et la prédiction de l'âge, genre et ethnicité

1.1 Introduction

La distribution des données a un impact primordial sur l'apprentissage des algorithmes. Il est donc nécessaire de prendre en considération tous les aspects possibles de la distribution. Le but de ce chapitre est de présenter diverses approches de l'apprentissage automatique et leur application dans la prédiction des données démographiques telles que l'âge, le genre et l'ethnicité. Dans cette étude, nous nous concentrons sur deux modèles multi-sortie : le CNN et le ResNet, en mettant en contexte quelques travaux connexes comme référence de base.

Le chapitre sera structuré en sections principales suivant un ordre chronologique dans le but de produire des résultats logiques et cohérents

Commencent en premier lieu par l'exploration de l'apprentissage automatique En vue d'atteindre une vue d'ensemble sur les méthodes et les techniques utilisé dans le contexte de la prédiction par une exploitation des algorithmes d'apprentissage automatique couramment utilisé.

Puis en deuxième lieu il faut fournir une analyse sur la distribution des données

En dernier cette section sera dédiée à la prédiction de l'âge, du genre et l'ethnicité à partir es images faciales ainsi que les résultats obtenus dans des travaux antérieurs dans le but de fournir des pistes de recherche et exploiter les perspectives futures dans ce domaine.

1.2 Exploration de l'apprentissage automatique

1.2.1 Aperçu sur l'apprentissage automatique

L'apprentissage automatique (Machine Learning en anglais) c'est faire apprendre une machine à réaliser plusieurs tâches distinctes simultanément sans programmation et avoir la capacités d'extraire des connaissances à partir des données et assister les décideurs humains dans le processus complexe de prise de décision [1].

En général, le processus de l'apprentissage automatique se réalise en deux phases (Voir Figure 1-1).

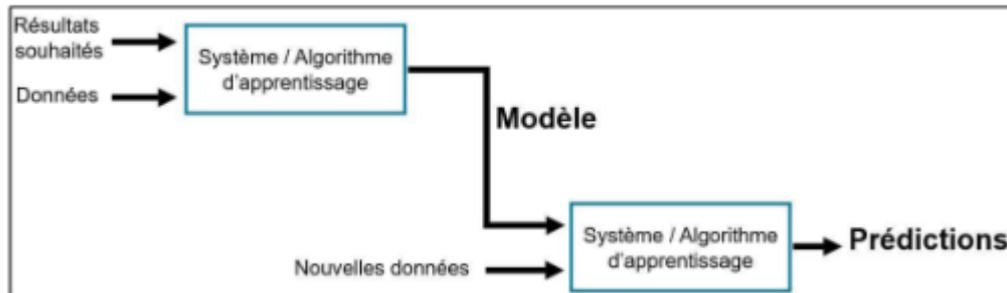


Figure 1-1 :Schéma présentatif résumant l'apprentissage automatique[2].

La première phase est la conception du système qu'on appelle entraînement. Il consiste à estimer un modèle à partir de l'analyse des données. La seconde phase est la mise en production. Après la détermination du modèle, on teste la seconde partie de nouvelle données utile pour la réalisation de la tâche désirée [3].

L'apprentissage automatique est utilisé dans différents domaines tels que la reconnaissance d'objets (visages, langages naturels, formes...) [4], la détection de fraude à la carte de crédit [5], et la génération automatique d'images [6].

1.2.2 Types d'algorithmes d'apprentissage automatique

Différents types d'algorithmes ont émergé en réponse à la diversité des besoins et des caractéristiques des informations disponibles, ainsi qu'à la qualité et à la quantité des données. Parmi les plus connus, citons les algorithmes suivants :

1.2.2.1 Apprentissage supervisé

L'apprentissage supervisé consiste à entraîner un modèle dans le but d'avoir un comportement précis à partir d'un ensemble de données étiquetées et structurées dans le même contexte en entrée. Une analyse est faite sur les données afin d'étudier ces exemples et de les comprendre et d'ainsi de rendre le modèle capable de prédire et déterminer correctement les étiquettes pour toute entrée qui lui convient où traiter les nouvelles données (voir figure 1-2).

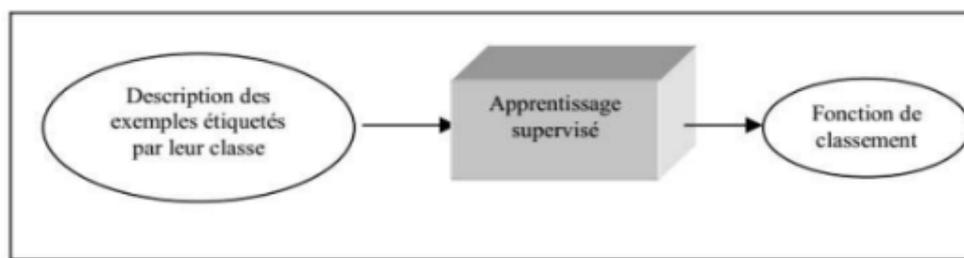


Figure 1-2 : Présentation de l'apprentissage supervisé [2].

Il existe une grande variété d'algorithmes d'apprentissage supervisé. Cependant, certains sont plus fréquemment utilisés que d'autres. Pour les données étiquetées, divers algorithmes sont utilisés comme le k plus proche voisins [7], l'Arbre de décision [8], les Support vecteurs Machines (SVM) [9], le Bayes naïf [10] et la régression linéaire [11].

1.1.1.1 Apprentissage non supervisé

Les algorithmes de ce type d'apprentissage apprennent à partir d'ensemble de données inconnues, qui ne sont pas accompagnées d'annotations contrairement à l'apprentissage supervisé ce qui distingue les deux types.

Les algorithmes d'apprentissage non supervisé utilisent des techniques de clustering basées sur la similarité et la dimensionnalité pour effectuer la classification des données, leur permettant ainsi de découvrir des structures sans utiliser d'étiquettes. Différents algorithmes sont adoptés dans cette approche de classification, tels que le k-means (Clustering) [12], la réduction de dimension par factorisation [13], et les modèles de Markov cachés [14].

1.1.1.2 Apprentissage profond

L'apprentissage profond est un type d'apprentissage automatique qui se base plus spécifiquement sur la notion de réseau de neurones artificiel. Ces réseaux sont composés de plusieurs couches de neurones cachées ou chacune recevant et interprétant les données de la couche précédente, d'où la notion de profondeur. Cette particularité leur permet de détecter des propriétés précises à partir de données non structurées. Le développement de ce type des réseaux a été conçu pour effectuer une extraction automatique de primitives à partir d'un ensemble de données brutes volumineuses, sans recourir à des techniques préétablies de l'extraction de caractéristiques. Ceci a permis de résoudre des problèmes d'apprentissage beaucoup plus complexes que ne le font les réseaux de neurones moins profonds et représente une différence majeure par rapport à l'apprentissage automatique traditionnel [15].

Il existe un grand nombre de variables d'architectures d'apprentissage profondes. Nous allons choisir de définir les plus connus et adéquats par rapport à l'étude de performance de classifieurs.

1.1.1 Réseaux de neurones

Les réseaux de neurones artificiels (Artificial neural network, ANN) sont une inspiration approximative des réseaux de neurones biologiques, avec une structure et un fonctionnement partiellement semblable. ANN présente un modèle de transformation de données constitué d'un ensemble d'unités (représentant les neurones), superposées en couches et interconnectées via des connecteurs [16,17].

Pour rendre la machine capable d'apprendre à partir des données fournies en entrée, un réseau de neurones repose sur un grand nombre de neurones fonctionnant en parallèle et organisés en couches. La première couche reçoit des informations brutes en entrée, puis, chaque couche reçoit les sorties d'informations des couches précédentes. Cette organisation en couches permet un traitement progressif des données et des transformations complexes [17].

Différents types de réseaux de neurones sont distingués en fonction de divers critères. Ils sont catégorisés en fonction du nombre des couches cachées qui séparent les données d'entrée de la production des résultats, en fonction du nombre de nœuds cachés du modèle, ou encore par la nature de leur fonction d'activation. La variance des réseaux de neurones est déterminée selon la façon de transmission des informations entre les différentes couches de neurones. Il existe de nombreuses architectures de réseaux de neurones différentes qui peuvent être utilisées pour simuler cette propagation de l'information. Dans la variante la plus simple des réseaux de neurones

acycliques, appelée : propagation avant (en anglais feed-forward"), les informations passent directement de l'entrée aux nœuds de traitement puis vers les sorties [17].

1.1.1.1 Réseaux de neurone convolutif

Les réseaux de neurones convolutifs, abrégés CNN (Convolutional Neural Network en anglais), font partie des modèles les plus performants, notamment pour la tâche de classification d'images. Ces réseaux se basent tout d'abord sur une entrée, généralement une image, qui est fournie en tant qu'input. Cette image est ensuite passée à travers une succession de filtres et est représentée sous la forme d'une matrice de pixels avec une résolution définie, présentée en termes de hauteur, largeur et dimension. Cela permet son intégration dans le processus d'extraction des caractéristiques par la partie convolutive, qui fonctionne comme un extracteur de caractéristiques. Ainsi, de nouvelles images appelées cartes de convolutions sont créées, puis elles sont concaténées pour former un vecteur de caractéristiques.

Ensuite, le flux de données passe à la couche de pooling, qui réduit la taille des caractéristiques. Ensuite, l'activation est appliquée pour convertir les valeurs négatives en zéro, avant de passer aux couches entièrement connectées. Ces dernières transforment le vecteur de caractéristiques afin de produire les sorties finales. Finalement, les prédictions du réseau sont fournies. Un avantage majeur des réseaux convolutifs est l'utilisation d'un poids unique associé aux signaux entrant dans tous les neurones d'un même noyau de convolution. Cette méthode réduit l'empreinte mémoire, améliore les performances et permet une invariance du traitement par translation. L'absence de paramétrage initial et d'intervention humaine est un atout majeur des CNN [18].

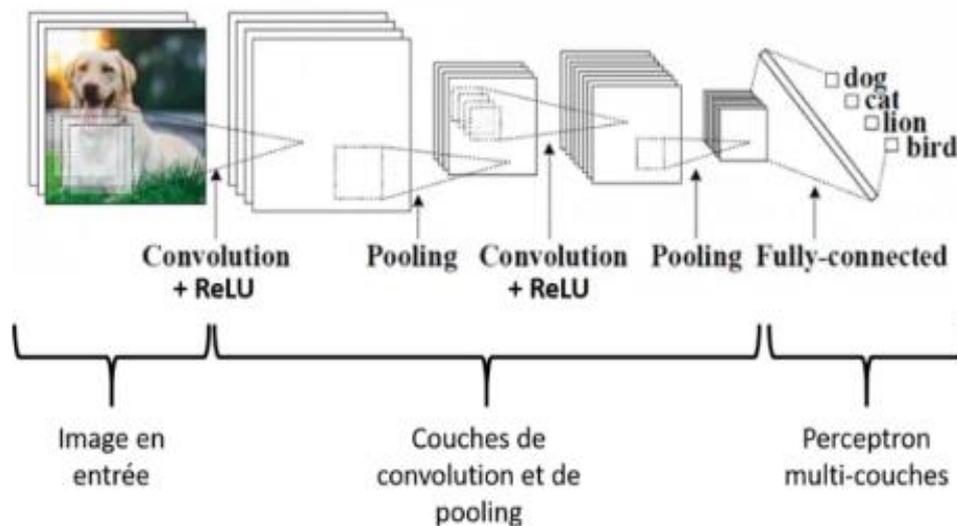


Figure 1-3 : Parties du CNN [18].

Un réseau de neurones convolutif comprend quatre couches : la couche de convolution, la couche de rassemblement, la couche de correction ReLU et la couche complètement connectée [18].

A. Couche de convolution

La convolution est une opération mathématique importante pour les données de pré-traitement. Le filtre de convolution est utilisé pour identifier les caractéristiques de l'image en utilisant certains filtres. Le filtre de convolution est utilisé pour identifier les caractéristiques de l'image. Lorsqu'une image traverse une série de filtres, elle produit des cartes de convolutions, également appelées noyaux de convolution. En utilisant une méthode de maximum local, des filtres intermédiaires sont connectés à la résolution de l'image. Finalement, les cartes de convolutions sont mises à plat et combinées pour créer un vecteur de caractéristiques [18].

B. Couche de mise en commun : pooling

Le pooling est une technique utilisée pour réduire la taille d'une image en ne gardant que les informations les plus importantes, comme la valeur maximale ou la valeur moyenne pour chaque groupe de 4 pixels.

Pooling est une étape indispensable dans la construction de réseaux de convolution. L'agrégation spatiale aide à réduire la taille spatiale des cartes de caractéristiques. Il est judicieux d'insérer une couche de pooling entre deux couches de convolution successives d'une architecture CNN pour contrôler le sur-apprentissage (l'overfitting). Il existe plusieurs fonctions de pooling, comme pooling-moyenne qui prend la valeur moyenne de toutes les valeurs couvertes par la zone, mais le max-pooling qui prend la valeur maximale[18].

C. Rectified Linear Unit

Les cartes de convolution sont passées à travers une couche d'activation non linéaire comme l'unité d'activation rectifiée linéaire (ReLU), qui utilise des zéros pour remplacer les nombres négatifs des images filtrées[18].

D. Couche Dense

La couche dense ou bien la couche entièrement connectée est la dernière couche d'un réseau de neurones. Cette couche accepte un vecteur en entrée et sort un nouveau, appliquant une combinaison linéaire et une fonction d'activation aux valeurs entrées. La couche connectée permet de classer l'image entrante dans le réseau en renvoyant un vecteur de taille N (N est le nombre de classes) [18].

1.1.1.2 CNN a sortie multiple

Le réseau de neurones convolutifs, ou le CNN Multi-Output a sorties multiples, a été élaboré pour répondre aux besoins de sorties multiples simultanées. Cette architecture permet d'éviter de créer des réseaux distincts pour chaque tâche [19].

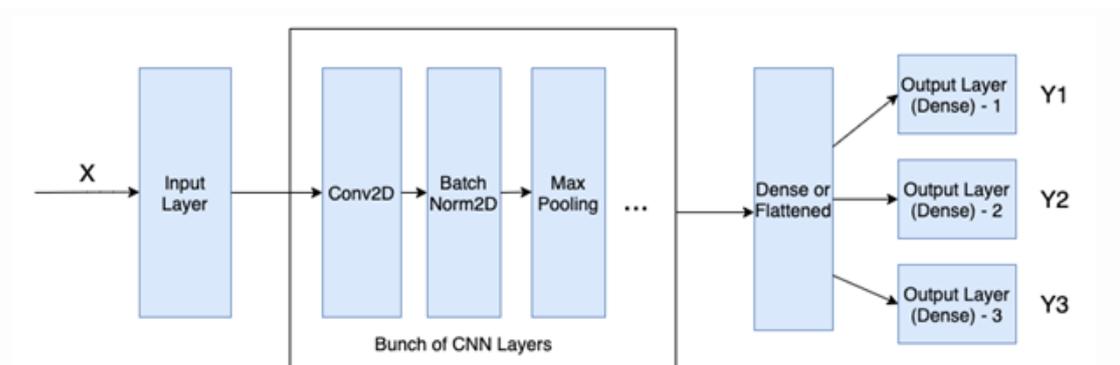


Figure 1-4 : Présentation de l'architecture des couches d'un CNN-Multi-output [19].

L'architecture d'un CNN multi-sorties est basée sur un CNN simple mais diffère en termes d'ordonnancement et de nombre de couches. En ajoutant une branche pour chaque sortie, nous

pouvons mettre en évidence la distinction entre un CNN simple et un CNN multi-sorties. Pour illustrer cette différence, nous présentons un tableau comparatif dans la figure suivante.

Tableau 1-1 : Comparaison entre les CNN à une seule sortie et multi sortie [19].

Critères	CNN à une seule sortie	CNNs à multiples sorties
Sorties	Une seule sortie pour une tâche spécifique	Plusieurs sorties pour tâches multiples
Architecture	Architecture standard avec couches convolutionnelles, de pooling et entièrement connectées	Architecture similaire avec des branches supplémentaires pour chaque sortie
Fonctionnement	Effectue des prédictions ou classifications pour une tâche unique	Effectue des prédictions et des classifications simultanées pour différentes tâches
Coût	Coût en termes des ressources informatiques dépendant de la taille et de la complexité du modèle	Potentiellement plus coûteux en raison de la complexité accrue et du traitement parallèle des différentes sorties

1.1.1.3 Réseaux résiduels profonds

Les réseaux résiduels profonds (Deep Residual Networks, ResNets en anglais) représentent une architecture de réseau de neurones plus profonde qui s'inspire des réseaux de neurones convolutifs, avec une augmentation du nombre de couches, pour résoudre le problème de dégradation des performances des réseaux. Cette dégradation est due à la difficulté d'apprendre des fonctions d'identité à travers de nombreuses couches de neurones. Pour résoudre ce problème, ResNet utilise des blocs résiduels empilés les uns sur les autres, généralement représentés par des formules mathématiques prédéfinies.

Les ResNets permettent aux informations de contourner certaines couches du réseau, ce qui facilite l'apprentissage de fonctions d'identité à travers des réseaux profonds. Ces modèles ont été largement utilisés dans la communauté de la vision par ordinateur pour résoudre une variété de tâches, notamment la classification d'images, la détection d'objets et la segmentation sémantique [20].

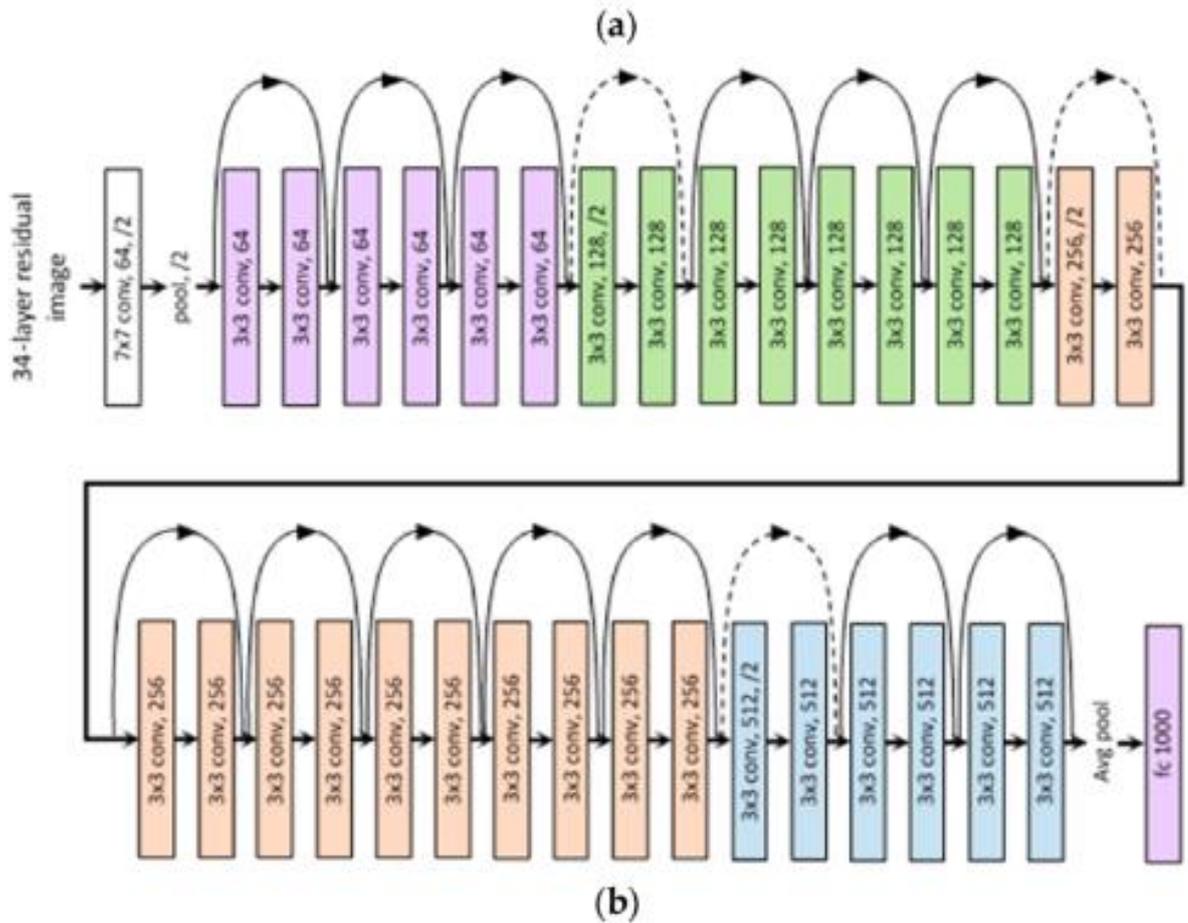


Figure 1-5 : Architecture d'un ResNet 34 [20].

1.2 Analyse des distributions de données

Les distributions des données regroupées dans le même contexte afin de créer une base de données peuvent varier en fonction de plusieurs facteurs, tels que la source des données, la qualité des données. Voici quelques exemples de variations possibles de la distribution des données.

1.2.1 Biais

Les données peuvent être biaisées en faveur d'une ou plusieurs classes, ce qui peut conduire à des prédictions erronées du modèle d'apprentissage [21].

1.2.2 Déséquilibre

Les classes déséquilibrées posent de réelles difficultés aux algorithmes de Machine Learning et de Deep Learning [22].

1.2.3 Manipulation d'image

Modifier les images en introduisant différents arrière-plans, superpositions ou occultations. Cela peut impliquer de placer des objets sur différents arrière-plans, de superposer d'autres éléments ou de masquer partiellement certaines parties de l'image. Ces manipulations peuvent simuler différentes conditions environnementales ou introduire des objets ou des distractions supplémentaires, augmentant ainsi la complexité de l'ensemble de données [23].

1.2.4 Variation des contextes

Les images peuvent provenir de différentes sources, environnements ou contextes d'acquisition. On cite les conditions de luminosité différentes, environnements urbains ou naturels.

1.2.5 Variation des attributs visuels

Les images peuvent varier en termes de couleurs, de textures, d'éclairage, de perspectives, et de déformations. Ces variations visuelles peuvent être liées à des conditions d'acquisition différentes, à des environnements variés ou à des manipulations spécifiques.

1.2.6 Distribution de luminosité

Chaque dataset est composé d'un certain nombre d'images, et chaque image possède un histogramme qui représente la probabilité de densité de luminosité des pixels. Il est important de noter que cette distribution spécifique influence la perception et l'interprétation des informations lors de l'entraînement des machines [24].

En conclusion, il est important de prendre en compte les variations de distribution des données, les changements de distribution au fil du temps et les variations des contextes et des

attributs visuels. Chacun de ces éléments peut entraîner une distribution complètement différente par rapport à d'autres ensembles de données.

1.3 Détection de l'âge, le genre et l'ethnicité à partir des attributs faciaux

La détection des attributs faciaux tels que l'âge, le genre et l'ethnicité est devenue une tâche courante dans le domaine du deep learning. Elle est réalisée en traitant une grande quantité de données d'images, ce qui permet de reconnaître les motifs et les caractéristiques présents dans ces images.

Cette capacité de détection peut être utilisée dans de nombreuses industries grâce à des algorithmes d'analyse d'images qui traitent une grande quantité de données. Dans le cadre de notre projet de fin d'étude, nous avons choisi de nous concentrer sur la détection de l'âge, du genre et de l'ethnicité en raison du nombre important de travaux de recherche réalisés sur ce sujet, ainsi que des variations significatives de la distribution des images et de la sensibilité des performances des modèles aux changements de données. Ces facteurs ont une influence sur l'apprentissage des modèles et, par conséquent, sur leurs résultats.

1.3.1 Détection de l'âge

La détection de l'âge à partir d'une image qui contient l'apparence d'un être humain est une tâche très souvent demandée et dans plusieurs Domaines, on peut citer par exemple sécurité comme la gestion et control d'accès, la vérification de l'âge en ligne, la vente de produits soumis à des restrictions d'âge, le marketing, la sécurité routière, les attractions familiales, l'analyse démographique, les rencontres en ligne et beaucoup d'autre Domaines de la vie quotidienne a fin de faciliter le mode de vie et faire gagniez le temps.

Cette dernière présente une difficulté pour la machine. Dans ce contexte, l'amélioration des performances de prédiction de l'âge a été un objectif visé par plusieurs travaux cités dans la littérature [25].

1.3.2 Détection du genre

Tout comme le concept d'âge, les concepts de "genre" et de "sexe" ne sont pas équivalents. Le terme "genre" se réfère aux différences sociales apprises entre les hommes et les femmes, tandis que le terme "sexe" est utilisé pour désigner les différences biologiques entre les "mâles" et les "femelles" [26]. La détection du genre à partir d'une image est un domaine de recherche axé sur l'apparence. Cette tâche présente des défis uniques en raison de la variabilité des traits faciaux entre les individus et des stéréotypes de genre associés à certaines caractéristiques visuelles. Il est

nécessaire d'extraire des caractéristiques discriminantes du visage, telles que la forme du visage, les traits faciaux et les textures.

La détection du genre à partir d'une image présente un large éventail d'applications potentielles, allant de la recherche en sciences sociales à la sécurité et à la surveillance, en passant par le marketing et la personnalisation des services. Un nombre considérable de travaux ont été effectués dans la littérature scientifique pour développer des méthodes de détection du genre. Parmi eux, on peut citer la conférence intitulée "Apprentissage en profondeur du genre des images faciales à l'aide de TensorFlow" en 2019 [27], le journal publié sous le titre "Détection de la tendance criminelle à partir d'images faciales et effet de biais de genre" [28], ainsi que l'article de journal sur les performances de reconnaissance faciale et le rôle des informations démographiques.

1.3.3 Détection de l'ethnicité

Il existe plusieurs manières de définir l'ethnicité. En général, elle fait référence à la façon dont une personne s'identifie au sein d'un groupe en fonction de caractéristiques telles que la couleur de la peau, la forme des yeux, la symétrie du visage, la forme du nez et l'apparence en général, etc. Cette caractéristique a été abordée dans plusieurs travaux de recherche, notamment lors de la conférence de J. Buolamwini et T. Gebru en 2018 intitulé : Gender shades : Disparités d'exactitude intersectionnelle dans la classification commerciale du genre. Qui traite des disparités de précision intersectionnelle dans la classification commerciale des genres [29].

Par conséquent, prédire l'ethnicité pour les trois ensembles démographiques distincts tout en maintenant un taux de performance élevé représente un défi à relever et à intégrer dans la littérature scientifique.

1.3.4 Etat de l'art sur la prédiction de l'âge, du genre et l'ethnicité

La prédiction de l'âge, du genre et de l'ethnicité dans le cadre de la classification multitâche suscite un intérêt croissant dans le domaine de l'intelligence artificielle. De nombreuses recherches, études et travaux connexes ont été réalisés dans ce contexte afin de maximiser les performances et de prendre en compte tous les facteurs qui influencent ces prédictions. Afin d'établir des bases solides, nous avons mené une étude comparative (Tableau 2.2).

Tableau 1-2 : Parcours des travaux connexes.

Référence	Data	Données démographiques			classifieur	Accuracy
		Age	Genre	Ethnicity		
[25]	UTK Face	X			CNN	90%
[30]	LFW CELEB A IMDB-Wiki		X		GoogleNet	96%
[31]	UTKFace			X	CNN	77.5%
[32]	UTK Face	X	X		CNN	83.57%
[33]	FairFace	X	X	X	ResNet34	90%

1.3.5 Discussions et limitation

En ce qui concerne les comparaisons entre les travaux connexes, il est important de reconnaître les limitations principales :

- Les bases de données utilisées diffèrent d'autre. Chaque base de données possède ses propres caractéristiques et biais inhérents, ce qui peut avoir un impact sur les performances des modèles.
- Le classifieur utilisé impacte forcément la performance de prédiction parce que chaque classifieurs diffère de l'autre par la Capacités de discrimination, Capacité d'apprentissage, Complexité et temps de calcul, Adaptation au problème, Prétraitement et ingénierie des caractéristiques et Interprétabilité.
- Architecture de chaque modèle (type et nombre des couches intégrés au niveau du modèle)

Dans l'article [25] « Détection d'âge et de genre en utilisant réseaux de neuronaux convolutionnels (CNN) », le classificateur UTK Face atteint une précision de 90 % dans la détection de l'âge. L'article [30] utilise également une architecture personnalisée pour la classification de genre, avec une précision de 96%. L'article [31] détecte également l'ethnicité en utilisant des réseaux neuronaux profonds, avec une précision de 77,5%. L'article [32] prédit l'âge et du genre à l'aide de réseaux neuronaux convolutionnels profonds, avec une précision de 83,5%. L'article [33] emploie également une approche d'apprentissage profond avec le modèle ResNet34 pour la prédiction de l'âge, du genre et de la nationalité, avec une précision de 90%. Ces résultats démontrent l'efficacité des différentes architectures de réseaux neuronaux et des bases de données dans la détection et la prévision des caractéristiques démographiques.

1.6 Conclusion

Ce chapitre nous a fourni une base solide pour aborder le chapitre suivant sur la méthodologie. Les trois sections nous ont apporté une compréhension approfondie de l'exploration de l'apprentissage automatique, de l'analyse des distributions des données et de la prédiction de l'âge, du genre et de l'ethnicité. Nous sommes maintenant prêts à appliquer ces connaissances pour analyser et concevoir des objectifs qui répondent aux besoins spécifiques de notre projet.

Chapitre2 : Mesure de l'impact de la distribution de donnée sur les modèles d'apprentissage profond

2.1 Introduction

Ce chapitre se concentre sur la mesure de l'impact de la distribution des données sur les performances des modèles d'apprentissage profond. Comprendre comment la distribution des données affecte les résultats obtenus est crucial lors de l'utilisation de modèles d'apprentissage profond pour résoudre des problèmes de classification, de régression ou de génération. Les variations dans la distribution des données peuvent résulter de changements dans les conditions de collecte des données, d'erreurs d'attribution ou de préjugés inhérents. Ces variations peuvent causer des problèmes lors de la formation et de l'évaluation des modèles d'apprentissage profond.

Dans ce chapitre, nous aborderons tout d'abord les bases de données que nous avons utilisées, en décrivant leurs caractéristiques, leur distribution, leurs variations et leur structuration. Nous examinerons également la méthode qui supervise et guide l'ensemble du processus, garantissant ainsi sa cohérence et sa fiabilité.

Ensuite, nous présenterons les différentes méthodes de conception que nous avons entreprises pour atteindre notre objectif de manière explicite. Nous discuterons des approches et des techniques utilisées pour évaluer la répartition des données, ainsi que pour résoudre les problèmes liés aux changements de distribution des données.

Ce chapitre sera une étape cruciale pour comprendre comment la distribution des données peut influencer les performances des modèles d'apprentissage profond. En analysant les bases de données utilisées, leurs caractéristiques et leurs variations, nous serons en mesure de mieux appréhender les résultats obtenus et de prendre des décisions éclairées sur les approches à adopter.

2.2 Bases de données

Les bases de données jouent un rôle essentiel dans l'apprentissage automatique et l'apprentissage profond, car elles permettent aux utilisateurs et aux modèles d'accéder à un grand volume de données structurées, organisées et étiquetées, facilitant ainsi l'apprentissage. Dans le domaine du Deep Learning, les bases de données peuvent prendre la forme d'images, de vidéos, de textes, de signaux audios, de graphiques ou de données tabulaires. Il est important que les bases de données utilisées soient générales et ne soient pas spécifiques à une seule personne ou à une

seule région. Dans cette section, nous présenterons les bases de données les plus couramment utilisées pour la prédiction de l'âge, du genre et de l'ethnicité.

2.2.1 UTKFace

Le jeu de données UTKFace [34] est un ensemble de données d'images faciales relativement important, comprenant des individus de 0 à 116 ans. Il contient 20 000 images de visages qui présentent une grande variété de poses, d'éclairages, d'occlusions, de résolutions et d'expressions faciales. Les images sont étiquetées en termes d'âge, de genre et d'origine ethnique, et comprennent cinq catégories ethniques : blanc, noir, asiatique, indien et autres. Les étiquettes de cet ensemble de données ont été estimées à l'aide de l'algorithme Deep EXpectation (DEX) [35] puis vérifiées par des annotateurs humains.



Figure 2-1 : Échantillon des photos de la dataset UTKFace [35].

2.2.2 Fairface

Cet ensemble de données est constitué de plus de 108 501 images de visages de personnes, qui varient considérablement en termes de pose, d'éclairage, de résolution et d'expression. Le jeu de données FairFace est utilisé pour la reconnaissance des attributs faciaux tels que l'âge, le genre et l'ethnicité [36].

Les annotations divisent les données en sept catégories ethniques : Blanc, Noir, Indien, Asiatique de l'Est, Asiatique du Sud-Est, Moyen-Orient et Latino. Les images de cet ensemble de données sont créées à partir de l'ensemble de données Flickr YFCC-100M [34]. La différence entre les ensembles de données UTKFace et FairFace réside uniquement dans la manière dont les étiquettes d'âge sont présentées. Dans UTKFace, l'âge est indiqué de manière précise, tandis que FairFace utilise des intervalles d'âge.



Figure 2-2 : Echantillon des photos de la dataset FairFace [36].

2.2.3 Flickr-Faces-HQ dataset

L'ensemble de données se compose de 70 000 images PNG de haute qualité à une résolution de 1024×1024 et contient des variations considérables en termes d'âge, FFHQ-Aging est un ensemble de données de visages humains conçu pour comparer les algorithmes de transformation de l'âge ainsi que de nombreuses autres tâches de vision possibles. Ce jeu de données est une extension du jeu de données NVIDIA FFHQ, en plus des 70 000 images FFHQ originales, il contient également les informations suivantes pour chaque image: * Informations sur le sexe (homme / femme avec score de confiance) * Informations sur le groupe d'âge (10 classes avec score de confiance) * Pose de la tête (tangage, roulis et lacet) * Type de lunettes (aucune, normale ou foncée) * Score d'occlusion oculaire (0-100, score différent pour chaque œil) * Carte sémantique complète (19 classes, basées sur les étiquettes CelebAMask-HQ) [37].



Figure 2-3 : Échantillon des photos de la dataset FFHQ.

2.3 Méthodologie adoptée

Une fois que le système à concevoir est délimité et que les principales lignes directrices sont établies, il est important de choisir une méthodologie appropriée pour concevoir une solution adaptée. La méthodologie choisie doit répondre aux besoins et aux spécificités du système à concevoir, tout en facilitant l'optimisation et en garantissant une approche de conception de solution efficace. Dans notre étude, la méthodologie choisie s'adapte aux objectifs spécifiques qui consistent à analyser l'impact de la variation et de la distribution des données d'entraînement et de test sur les performances des modèles d'apprentissage profond, en tenant compte des distributions de luminosité. Les étapes de la méthodologie sont illustrées dans la Figure 2-4. Notre choix est justifié par l'absence de références ultérieures sur ce point de recherche dans la littérature scientifique. En identifiant cette lacune et en l'exploitant, notre étude permet de combler un vide dans la littérature existante et d'élargir notre compréhension du sujet. Elle offre ainsi une nouvelle perspective qui enrichit le domaine des modèles d'apprentissage et des distributions de données.

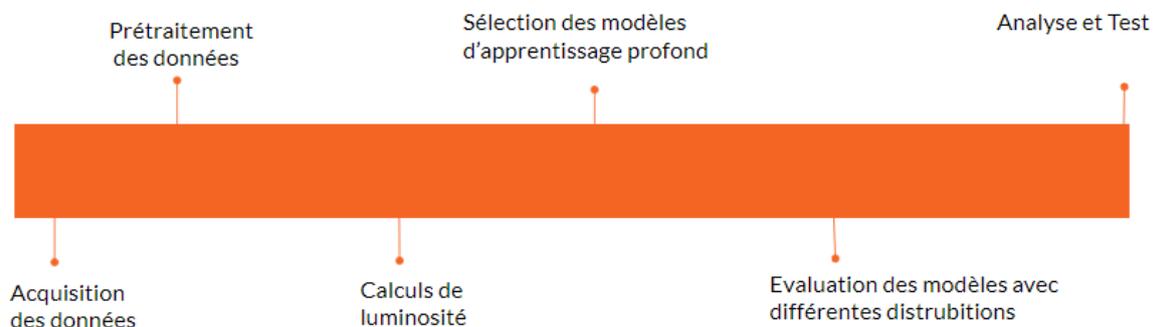


Figure 2-4 : Présentation d'ordre des étapes logiques de méthodologie.

2.3.1 Choix des données

La deuxième étape de notre méthodologie consiste à sélectionner des ensembles de données comportant des informations sur l'âge, le genre et l'ethnicité. Ce choix est effectué en respectant les protocoles éthiques appropriés, tout en garantissant la confidentialité et l'anonymat des individus représentés dans ces ensembles de données provenant de diverses sources. Cela permet d'assurer une représentativité adéquate de l'âge, du genre et de l'ethnicité.

Dans cette optique, nous avons choisi trois ensembles de données spécifiques pour valider notre étude. Tout d'abord, nous utilisons la base de données d'images UTKFace, qui a été largement utilisée dans la littérature scientifique. Ensuite, nous avons opté pour les ensembles de données FairFace et FFHQ. Ces bases de données offrent une grande variation dans la représentation des visages, avec une collecte équilibrée prenant en compte la diversité de l'âge, du genre et de

l'ethnicité au sein de la population étudiée. Elles fournissent également des annotations précises pour les étiquettes de classement et sont de taille importante.

Il est essentiel de prendre en compte les éventuels biais lors de l'acquisition de données afin de prévenir toute forme de discrimination ou de préjugé. C'est pourquoi nous avons choisi ces trois ensembles de données open source, qui ont été téléchargés depuis Kaggle, dans le but de garantir une approche plus large et de minimiser les biais potentiels.

2.3.2 Prétraitement des données

Une fois les données acquises, il est nécessaire de les prétraiter afin de les rendre appropriées pour les objectifs fixés. Cette étape peut inclure plusieurs opérations, tant avant le lancement de l'apprentissage que pendant celui-ci. Parmi ces opérations, on retrouve la normalisation, le redimensionnement, l'élimination des valeurs aberrantes en identifiant d'abord les classes correctes en parcourant l'ensemble de données, ainsi que la gestion des données manquantes, soit par suppression. Il est également important de comparer et de structurer toutes les bases de données de manière à avoir une égalité de taille des classes. Le prétraitement des données vise à garantir la qualité et la cohérence des données utilisées, ce qui aura un impact direct sur les performances du modèle.

2.3.3 Calculé de luminosité

Dans cette étude, nous avons identifié la distribution de la luminosité comme un facteur clé ayant un impact sur les performances des modèles. Nous utilisons des techniques de traitement d'image, de calcul numérique et des formules mathématiques pour analyser et quantifier la répartition de la luminosité dans les pixels images de notre ensemble de données [38].

Pour calculer la densité de luminosité, nous examinons la luminosité de chaque pixel de chaque image individuellement. En convertissant les images dans un espace colorimétrique en niveaux de gris, Ensuite, nous calculons la moyenne de ces valeurs de niveaux de gris pour obtenir la luminosité de l'ensemble de données.

La loi de Lambert-Beer [39], une loi fondamentale en optique, pour tenir en compte de la répartition de la lumière à travers l'image. Cette méthode utilise les valeurs de niveaux de gris pour obtenir une mesure quantitative de la luminosité, Cela nous permet de quantifier la répartition de la luminosité dans notre ensemble de données et d'obtenir des informations précieuses sur les caractéristiques visuelles des images.

2.3.4 Sélection des modèles d'apprentissage profond

La quatrième étape de la méthodologie consiste à sélectionner les modèles d'apprentissage profond appropriés pour effectuer la détection en utilisant une sortie multiple signifie que les modèles seront entraînés pour prédire simultanément l'âge, le genre et l'ethnicité, deux modèles spécifiques ont été choisis : ResNet-34 et CNN (Convolutional Neural Network).

La sélection de ces deux modèles permettra d'évaluer l'impact de la distribution et la sensibilité de la performance des modèles. Ces modèles ont été choisis en raison de leur efficacité prouvée et de leur adaptabilité à la tâche spécifique de détection de ces caractéristiques.

2.3.5 Évaluation des modèles avec différentes distributions

La cinquième étape de la méthodologie consiste à entraîner et à évaluer les modèles d'apprentissage profond en utilisant différentes distributions de données.

Cela peut être réalisé en divisant l'ensemble de données en ensembles d'entraînement, de validation et de test, et refaire le test avec un échantillon d'une dataset différente. Les performances des modèles sont ensuite mesurées par plusieurs métriques d'évaluations afin d'évaluer et voir la qualité du modèle par rapport à un ensemble de distribution différente.

Les métriques d'évaluation les plus couramment utilisées comprennent la précision, le rappel, la F-mesure et la matrice de confusion. Chaque métrique offre une perspective différente sur les performances du modèle. Dans cette partie, nous allons définir les métriques que nous avons utilisées pour évaluer notre système.

2.3.5.1 Le taux de bonnes classifications

L'accuracy (ou précision en français) est une mesure couramment utilisée pour évaluer la performance d'un modèle de classification. Elle représente la proportion de prédictions correctes parmi l'ensemble des prédictions effectuées par le modèle. Représenter comme suit :

$$\text{Accuracy} = \frac{\text{Prédictions correctes}}{\text{Total des prédictions}} \quad (2.1)$$

-La prédiction correcte : le nombre d'observations correctement classées par un modèle par rapport à l'ensemble total d'observations.

-Totale des prédictions : ensemble des prédictions effectuées par un modèle, c'est-à-dire le nombre total d'observations classifiées par le modèle, qu'elles soient correctes ou incorrectes.

2.3.5.2 Le F1 - Score

Le score F1 peut être considéré comme une mesure combinée de la précision et du rappel, avec une valeur optimale de 1 et une valeur minimale de 0.

$$\text{F1 Score} = \frac{2 \times \text{Précision} \times \text{Rappel}}{\text{Précision} + \text{Rappel}} \quad (2.2)$$

La précision : mesure le taux des positifs parmi les exemples prédits positives,

Le rappel : définit le taux de vrais positifs parmi toutes les exemples réellement positifs.

Pour le cas multi-classes, ces métriques sont mesurées pour chaque classe individuellement, puis la moyenne est généralement utilisée comme une mesure agrégée.

2.3.5.3 Mesure globale de l'erreur ou de la perte d'un modèle

La mesure globale de l'erreur ou de la perte d'un modèle, appelée "overall loss" en anglais est une métrique qui varie en fonction du type de modèle et du problème afin de calculer l'erreur totale du modèle sur l'ensemble des données d'entraînement ou de test.

L'erreur est calculée en comparant les prédictions du modèle avec les valeurs cibles réelles et en mesurant la différence entre les deux. Utilisée pour évaluer la qualité générale des prédictions du modèle [40].

2.3.5.4 Erreur moyenne absolue

Le graphe de l'erreur absolue moyenne (mean absolute error) MAE utilisé dans les modèles de prédiction pour évaluer la précision des prédictions par rapport aux valeurs réelles. L'erreur absolue moyenne est une mesure de la différence moyenne entre les prédictions d'un modèle fil du temps des valeurs réelles correspondantes [41]. Plus précisément dans notre études l'Erreur moyenne facilite l'interprétation de la caractéristique démographique d'Age du modèle de régression. Les modèles réaliser comporte les graphe d'analyse dans la phase d'analyse des résultat dans le chapitre 3.

2.3.6 Analyse des résultats

Enfin, la méthodologie comprend une étape d'analyse des résultats obtenus. Les performances des modèles avec différentes distributions sont comparées et interprétées pour comprendre l'impact de la distribution des données sur l'apprentissage profond et sur les modèles convolutifs. Des graphiques, des tableaux et d'autres outils d'analyse statistique peuvent être utilisés pour visualiser et présenter les résultats de manière claire et concise.

2.5 Conclusion

Nous avons présenté dans ce chapitre les différentes étapes de la méthodologie adoptée pour la conception de notre étude, les méthodes utilisées et les objectifs visés. Commencant par les datasets, ensuite leurs prétraitements en utilisant plusieurs techniques telles que redimensionnement des images et la normalisation et d'autres. Nous avons décrit par la suite le choix des modèles d'apprentissage profond et les métriques d'évaluation utilisées.

Nous avons également exposé dans ce chapitre les différentes méthodes de calcul de la luminosité, ainsi que des approches de leur égalisation. Ces approches vont nous aider pour analyser l'impact de la distribution de la luminosité des images sur la performance des modèles d'apprentissage profond. Dans le prochain chapitre nous allons voir les démarches d'implémentation et la discussion des résultats obtenus.

Chapitre 3 : Implémentation résultats et expérimentations.

3.1 Introduction

Le chapitre précédent a présenté en détail la méthodologie et le cadre conceptuel de notre étude. Dans le présent chapitre, nous allons présenter le volet pratique de notre recherche en définissant les méthodes et les outils utilisés. Nous allons fournir des implémentations détaillées pour deux méthodes suggérées : le CNN multi-output et le ResNet34.

Nous commencerons par expliquer en détail la mise en œuvre de ces deux méthodes. Pour le CNN multi-output, nous expliquerons comment nous avons configuré le réseau, les couches utilisées, et les fonctions d'activation. De même, nous fournirons des détails sur la configuration du modèle ResNet34, y compris les couches résiduelles, et les paramètres d'apprentissage.

Ensuite, nous présenterons les tests expérimentaux que nous avons réalisés en utilisant ces différents modèles. Nous expliquerons les ensembles de données utilisés, les prétraitements effectués et les métriques utilisées pour évaluer les performances des modèles. Nous discuterons également en profondeur des résultats établis, en analysant les avantages et les limitations de chaque méthode.

3.2 Matériels et outils de développement

3.2.1 Langage de programmation

Le choix du langage de programmation est crucial pour mettre en œuvre le paradigme requis et s'adapter aux besoins spécifiques. Dans notre cas, nous avons opté pour Python, qui s'est avéré être le meilleur choix pour notre implémentation.

3.2.1.1 Python

Python est un langage programmation open source, programmation multi-paradigme qui fonctionne sur plusieurs plateformes. Il favorise la programmation impérative structurée, fonctionnelle et orientée objet. Python fournit des outils de haute qualité et une syntaxe simple à utiliser [43], ce qui permet de maximiser la productivité des programmeurs. Il peut s'utiliser dans de nombreux contextes et s'adapter à tout type d'utilisation grâce à des bibliothèques spécialisées

en citant ci-dessous les bibliothèques que nous avons utilisées afin de mettre en œuvre notre travail :

NUMPY : est le package le plus simple et puissant pour le calcul mathématique et la manipulation de données en Python.

PANDAS : est utilisée pour la manipulation et l'analyse des données.

CV2 : OpenCV est une bibliothèque de haute performance pour le traitement des images numériques et la vision.

Matplotlib : Python est utilisé pour l'apprentissage automatique et l'évaluation des performances.

Globe : est utilisé pour rechercher des chemins de fichiers en utilisant des caractères génériques.

Cibole : la bibliothèque Python "cibole" fait référence à un projet spécifique.

Sklearn : Python est utilisé pour l'apprentissage automatique et l'évaluation des performances.

TENSORFLOW : TENSORFLOW est une bibliothèque de logiciels gratuite et open-source pour programmer diverses tâches. Peut l'utiliser dans les réseaux de neurones.

3.2.2 Roboflow

Roboflow est une plate-forme de vision informatique offrant des outils et des services pour la gestion, l'annotation et la création de modèles d'apprentissage automatique pour l'analyse d'image et de vidéo. Il simplifie la préparation des données, y compris l'augmentation, l'étiquetage et la gestion, afin de simplifier la modélisation de la vision par ordinateur.

3.2.3 Plateforme de développement

3.2.3.1 Google colaboratory

C'est un environnement de développement qui permet d'exécuter du code Python et développer des modèles de machine learning.

Google Colab nous a offert un accès gratuit à un environnement Jupyter Notebook, hébergé sur les serveurs de Google. Cela nous a permis de collaborer avec d'autres utilisateurs et de profiter des puissantes ressources de Google, telles que les unités de traitement graphique (GPU) et les unités de traitement tensoriel (TPU) gratuites, afin d'accélérer l'exécution de code intensif en calcul. Dans notre cas, nous avons utilisé le GPU Tesla K80 et Tesla T4.

3.2.3.2 Kaggle

Kaggle est une plateforme en ligne qui permet d'explorer des jeux de données, participer à des compétitions, partager des notebooks, exécuter des codes en ligne. Il spécifie pour la résolution de problèmes liés à l'analyse de données et à l'apprentissage automatique.

3.2.3.3 Google Drive

Google Drive permet aux utilisateurs de sauvegarder, synchroniser et utiliser les fichiers en ligne de stocker, partager, modifier et visualiser différents types de fichiers sur différents appareils et plateformes. Il offre des fonctionnalités telles que la version de fichier, l'accès hors ligne et l'intégration avec les services Google tels que Docs, Sheets et Slides.

3.3 Implémentation du système

L'implémentation de notre système a connu plusieurs phases importantes. Dans cette étude, nous abordons dans ce qui suit les étapes clés de l'implémentation, telles que la collecte et la préparation des données d'entraînement, la conception de l'architecture des réseaux, l'entraînement du modèle et l'évaluation de sa performance.

3.3.1 Choix et acquisition des données

Une fois que les données ont été soigneusement sélectionnées en fonction des objectifs et de la méthodologie du projet, il est important de les télécharger à partir de la source éthique pour les rendre plus utiles et exploitables. Voici quelques étapes que vous pouvez suivre pour enrichir les données dans le contexte de l'acquisition des ensembles de données UTKFace, FairFace et FFHQ.

3.3.2 Préparation des données

3.3.2.1 Prétraitement

Suppression des images corrompues : Le téléchargement des jeux de données différerait d'une source à l'autre pour chaque ensemble de données. Et pour cela les images peuvent être corrompues pour diverses raisons, telles que des erreurs de téléchargement, stockage ou des erreurs de traitement d'image.

Suppression des images de mauvaise qualité : Les images de mauvaise qualité, par exemple les images floues, surexposées ou sous-exposées, peuvent nuire aux performances du modèle.



Figure 3-1 : Echantillon des images de mauvaise qualité.

Suppression des images aberrantes : Les images aberrantes peuvent être des images qui ne représentent pas des visages humains, des images contenant des artefacts.



Figure 3-2 : Image représentatif des images aberrantes.

Suppression des images avec des visages partiellement cachés : Les images où une partie importante du visage est cachée (par exemple, des images avec des lunettes de soleil, des chapeaux ou des écharpes ou de l'écriture supposée sur le visage).



Figure 3-3 : Représentation d'une image d'un visage partiellement caché.

Suppression des images en double avec des étiquettes différentes : la dataset peut contenir des images identiques, mais qui ont été étiquetées différemment.



Figure 3-4 : Représentation d'images double avec étiquette différente.

Traitement des données déséquilibrées : Chaque BD a sa propre structuration par rapport à la taille de chaque classe donc il peut y avoir un nombre disproportionné d'exemples dans une ou plusieurs classes par rapport aux autres pour avoir une droiture au niveau des modèles il faut rendre les trois ensembles de données dans la même distribution de classes.

Restructuration de dataset : Restructurer nôtres dataset tel qu'on doit avoir la même distribution de données au niveau de classes (regroupement d'âge et ethnicité) de telle sorte d'avoir 9 classes d'âge et 4 classes d'ethnicité et 2 classe pour genre voir (Tableau 3-1 , Tableau 3-2 et Tableau 3-3).

Tableau 3-1 : Répartition de la classe âge.

Age	
Classe	ID classe
de 0 à 2 ans	2
de 3 à 9 ans	39
de 10 à 19	1019
de 20 à 29 ans	2029
de 30 à 39	3039
de 40 à 49	4049
de 50 à 59	5059
de 60 à 69	6069
plus 70	70

Tableau 3-2 : Répartition de la classe ethnicité.

Ethnicité	
Classe	ID classe
Blanc	0
Noir	1
Indien	2
Asiatique	3

Tableau 3-3 : Répartition de la classe de genre.

Genre	
Classe	ID classe
Femelle	1
Male	0

Augmentation des données : L'augmentation des données entre les jeux permet de générer des ensembles de données plus importants à partir des données déjà préexisté, ce qui améliore la prévision de l'apprentissage, dans notre cas après avoir effectué tous les traitements nécessaires la taille et redimensionner la taille des datasets dans le cadre de les rendre dans la même distribution la taille de chacune se démultiplie jusqu'à 19000 photos ce qui est compatible avec le nombre des couches de CNN mais très peu par rapport aux 34 couches du modèle ResNet ce qui va réduire les performances de ce dernier et pour éviter tout cela la meilleure solution était de faire une augmentation jusqu'à 44000 images.

Unifié id étiquette de classe : Unification de l'identifiant d'étiquette de classe : Afin d'intégrer les informations de données au niveau du jeu de données et d'optimiser le code, nous avons étiqueté les images en utilisant le nom suivant : [âge][genre][ethnicité]date comme illustré dans la Figure 0-5 , en nous basant sur le fichier CSV associé aux bases de données FairFace et FFHQ.

Cette étiquette d'identifiant est construite en combinant les informations d'âge, de genre, d'ethnicité et de date, ce qui permet une meilleure organisation et une référence facile aux caractéristiques des images dans le jeu de données. En utilisant le fichier CSV fourni, vous pouvez associer chaque image à son étiquette correspondante.

Cette unification des étiquettes facilite également l'accès et la manipulation des données, en particulier lors de l'utilisation du code pour le traitement, l'entraînement et l'évaluation des modèles.

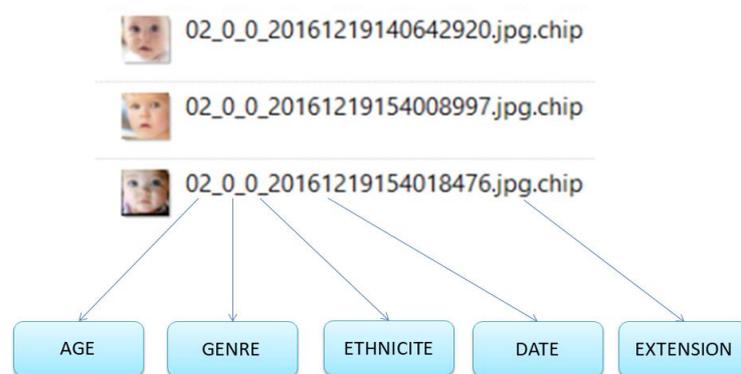


Figure 3-6 : Répartitions des étiquettes de classement des images

Unifier les classes des dataset : Les ensembles de bases de données d'images de visages publics déjà existants présentent un biais marqué en faveur de certains groupes ethniques, tandis que d'autres races, comme les personnes d'origine latino, sont significativement sous-représentées. Dans le but d'avoir des classifications identiques et d'adapter les données à une structure commune, nous avons conservé les classes communes et les classes qui représentent la vaste majorité de la population mondiale, à savoir Blanc, Noir, Asiatique et Indien. Nous avons exclu toutes les classes qui sont sous-représentées dans ces ensembles de données.

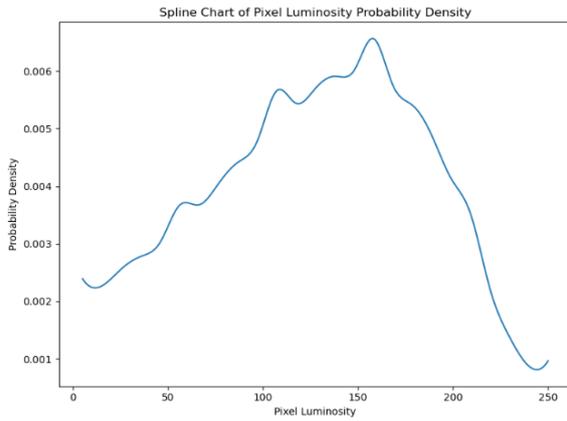
3.3.2.2 Calculer la luminosité pour chaque dataset :

Comme étape principale pour examiner les distributions de densité de luminosité des pixels pour chaque base de données (DB) dans leur intégralité, nous allons implémenter un algorithme qui passe par des phases successives dans l'ordre suivant :

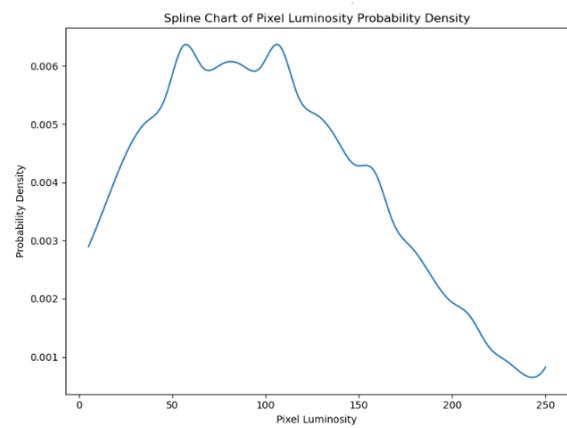
- Charger la dataset d'images dans un environnement de programmation.
- Convertir les images de la dataset en niveaux de gris en utilisant des fonctions de conversion de couleurs telles que `cv2.cvtColor()` de la bibliothèque OpenCV.
- Calculer la luminosité moyenne de chaque image en utilisant une formule qui prend en compte

Les valeurs de chaque pixel de l'image. Répéter cette étape pour chaque image dans la base de données.

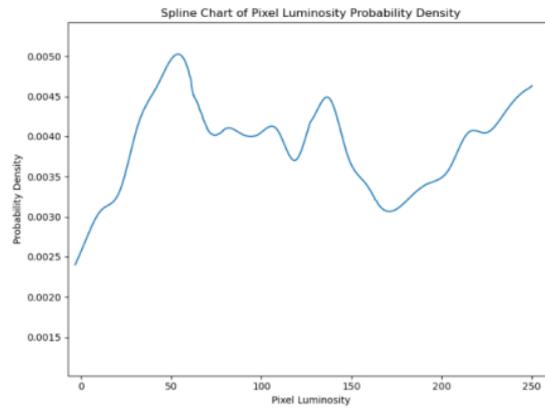
- Enregistrer les valeurs de luminosité moyenne pour chaque image dans un fichier ou dans une base de données.
- Calculer la moyenne de luminosité pour toute la dataset.
- Appliquer ces étapes nous permettra de visualiser les diagrammes et de remarquer les différentes distributions des datasets représentées dans les histogrammes (voir Figure 3-7) pour UTKFace, FairFace et FFHQ respectivement...



(a)



(b)



(c)

Figure 3-6: Les histogrammes de la densité de probabilité de la luminosité des pixels.

- (a) UTK
- (b) Fairface
- (c) FFHQ

Considérablement en termes de forme et de concentration. Cela peut être dû à des facteurs tels que les sources des images, les conditions de capture ou les caractéristiques

Spécifiques des visages inclus dans cette base de données.

Il est important de prendre en compte ces différences lors de l'analyse et de l'utilisation des ensembles de données pour des tâches telles que la modélisation ou l'apprentissage automatique.

En comprenant ces variations de distribution, nous pouvons adapter nos approches et nos modèles pour tenir compte des spécificités de chaque ensemble de données.

Il est également essentiel de garder à l'esprit que ces différences peuvent influencer les performances et les résultats des modèles appliqués aux ensembles de données. Par conséquent, une compréhension approfondie des caractéristiques de chaque ensemble de données est cruciale pour prendre des décisions éclairées dans le processus d'analyse et de modélisation.

3.3.3 Modélisation des réseaux de neurones

La modélisation d'une architecture appropriée pour nos modèles est essentielle pour garantir de bonnes performances et obtenir des résultats précis. Pour cela, nous avons défini les différentes couches nécessaires ainsi que les hyperparamètres spécifiques à chaque modèle choisi.

3.3.3.1 Modélisation CNN

La modélisation du modèle CNN nécessite des réseaux de neurones, ce qui implique de définir le type de couches à utiliser, leur nombre, et d'estimer les meilleurs hyperparamètres pour optimiser les performances du modèle et éviter des problèmes tels que le surajustement (overfitting) [44] ou le sous-ajustement (underfitting) [45]. Il est nécessaire de faire des choix judicieux dès le début ou de reconfigurer les paramètres en fonction des résultats obtenus afin d'améliorer le modèle. Il est également important de comprendre chaque paramètre pour chaque couche.

Dans notre cas, nous allons définir un modèle Keras à sorties multiples, composé de trois branches, chacune traitant et représentant une caractéristique démographique séparément.

Notre conception du modèle convolutionnel comprend, pour chaque branche, une couche Conv2D avec une activation ReLU, une couche de BatchNormalization, un MaxPooling, une couche de Dropout et enfin une couche Dense. Ces couches sont définies dans une méthode réutilisée pour chaque branche.

Tableau 3-4 : Les couches et les hyperparamètres de l'architecture utilisé du CNN.

Couche	Forme de sortie	Param #	Connecté
input_1 (InputLayer)	(None, 198, 198, 3)	0	
conv2d_1 (Conv2D)	(None, 198, 198, 16)	448	input_1[0][0]
conv2d_4 (Conv2D)	(None, 198, 198, 16)	448	input_1[0][0]
conv2d_7 (Conv2D)	(None, 198, 198, 16)	448	input_1[0][0]
activation_1 (Activation)	(None, 198, 198, 16)	0	conv2d_1[0][0]
activation_5 (Activation)	(None, 198, 198, 16)	0	conv2d_4[0][0]
activation_9 (Activation)	(None, 198, 198, 16)	0	conv2d_7[0][0]
batch_normalization_1	(None, 198, 198, 16)	64	activation_1[0][0]
batch_normalization_5	(None, 198, 198, 16)	64	activation_5[0][0]
batch_normalization_9	(None, 198, 198, 16)	64	activation_9[0][0]
max_pooling2d_1	(None, 66, 66, 16)	0	batch_normalization_1[0][0]
max_pooling2d_4	(None, 66, 66, 16)	0	batch_normalization_5[0][0]
max_pooling2d_7	(None, 66, 66, 16)	0	batch_normalization_9[0][0]
dropout_1	(None, 66, 66, 16)	0	max_pooling2d_1[0][0]
dropout_5	(None, 66, 66, 16)	0	max_pooling2d_4[0][0]
dropout_9	(None, 66, 66, 16)	0	max_pooling2d_7[0][0]
conv2d_2 (Conv2D)	(None, 66, 66, 32)	4640	dropout_1[0][0]
conv2d_5 (Conv2D)	(None, 66, 66, 32)	4640	dropout_5[0][0]
conv2d_8 (Conv2D)	(None, 66, 66, 32)	4640	dropout_9[0][0]

3.3.3.2 Modélisation ResNet34

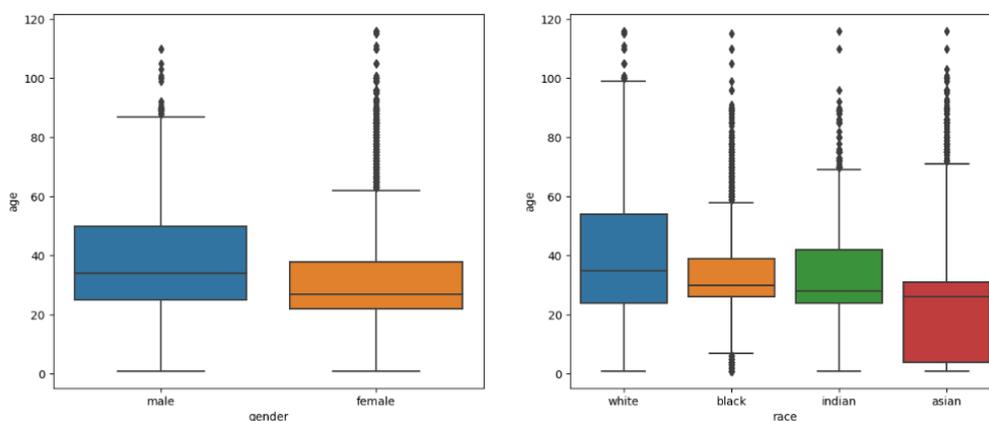
Le modèle ResNet34 est déjà prêt à 80% en utilisant les 34 couches déjà définies. Le modèle est construit à l'aide de Keras pour avoir des sorties multiples. ResNet34 utilise une architecture résiduelle, donc il a fallu adapter uniquement les hyperparamètres pour éviter le problème de dégradation du réseau. Les hyperparamètres de ResNet34 comprennent le taux d'apprentissage, qui contrôle la vitesse à laquelle le modèle ajuste ses poids, ainsi que la taille des lots utilisés lors de l'entraînement. Ces hyperparamètres et les fonctions d'activation ont été les mêmes que pour le CNN afin de maintenir une égalité entre les modèles.

3.3.3.3 Répartition des données

Après avoir spécifié l'architecture des réseaux de neurones que nous allons utiliser, il est essentiel de charger les données prétraitées dans le format approprié pour notre modèle. Assurons-nous que les données sont divisées de manière typique en trois ensembles : un ensemble d'entraînement de 70%, un ensemble de validation de 10% et un ensemble de test de 20%.

3.3.3.4 Analyse et visualisation des données

L'analyse et la visualisation des données sont essentielles dans les modèles d'apprentissage pour comprendre les données, préparer les données, sélectionner les caractéristiques, valider et justifier les résultats, ainsi que communiquer efficacement les informations. Cette étape permet d'obtenir des modèles plus robustes et d'exploiter au mieux les données disponibles voir Figure 3-7.



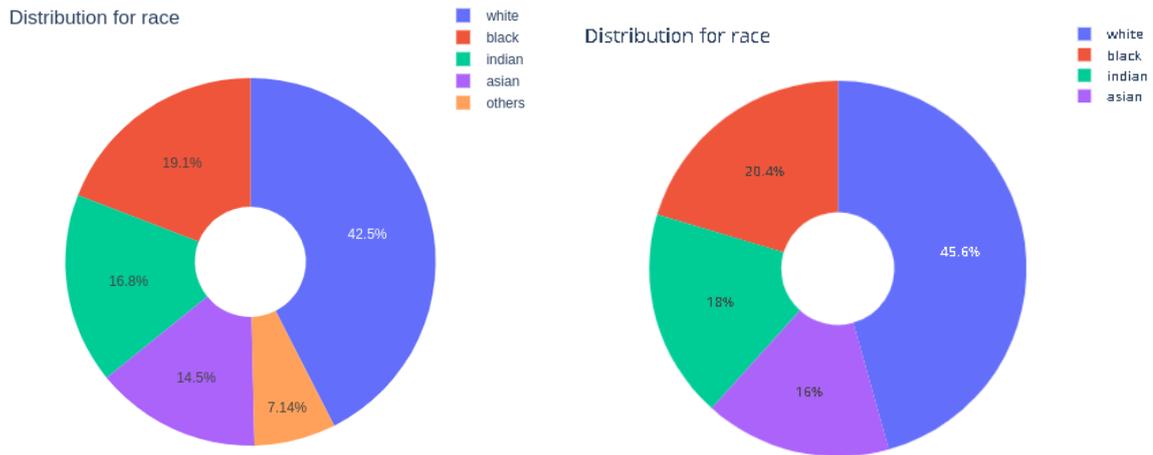


Figure 3-7 : Représentations des distributions et de l'analyse des données avant et après le prétraitement.

3.3.3.5 Lancer l'entraînement

Après avoir correctement chargé les données prétraitées dans le format approprié pour notre modèle, nous allons compiler notre modèle en utilisant l'optimiseur Adam. Assurons-nous de diviser les données en ensembles d'entraînement, de validation et de test. Choisissons également une taille de lot (batch size) appropriée et déterminons le nombre d'époques. Ensuite, nous pourrions lancer l'entraînement de notre modèle à sorties multiples.

3.4 Résultats et tests

Nous présentons ci-dessous les résultats obtenus après les tests des algorithmes d'apprentissage profond CNN multi-output et ResNet sur toutes les distributions des données.

NB : Au niveau des tableaux 3-5 et 3-7, les résultats d'exactitude pour la variable 'Âge' ont été représentés à l'aide de la métrique R2 (coefficient de détermination). En conclusion, pour traiter la prédiction de l'âge au niveau du modèle, elle a été considérée comme un problème de régression, puis les résultats ont été convertis en pourcentage afin de différencier et de mettre en évidence les performances.

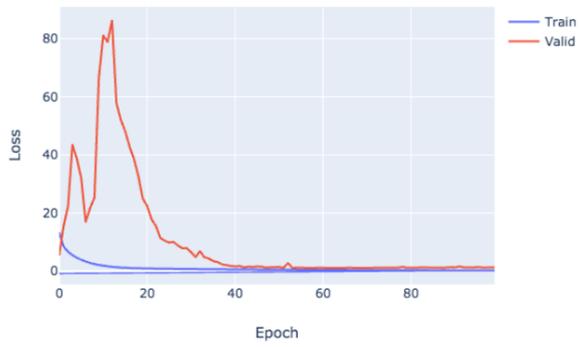
3.4.1 CNN multi-output

3.4.1.1 Courbes d'apprentissage

Les figures ci-dessus présentent les graphes dans l'ordre suivant : Graphique de la Perte Totale et Graphique de l'Erreur Moyenne Absolue pour l'âge.

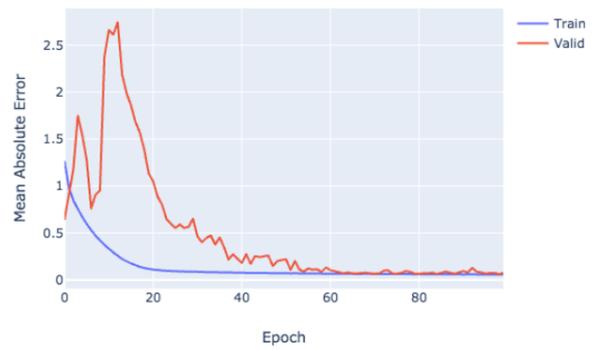
A. UTKFace

Overall loss



(a)

Mean Absolute Error for age feature



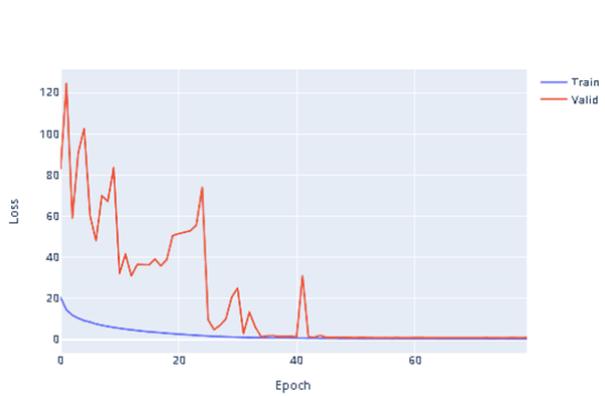
(b)

Figure 3-8 : les graphes d'apprentissage pour l'entraînement fait sur la data UTKFace.

(a) Overall loss

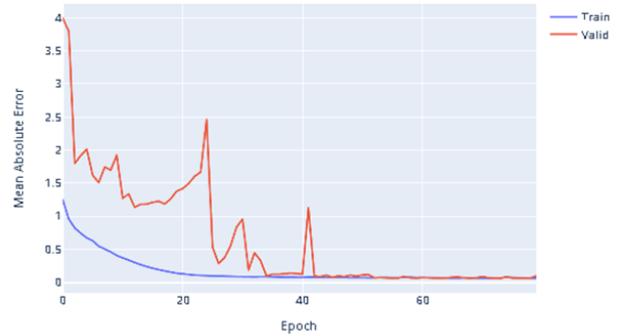
(b) Mean Absolute Error for age

B. FairFace



(a)

Mean Absolute Error for age feature



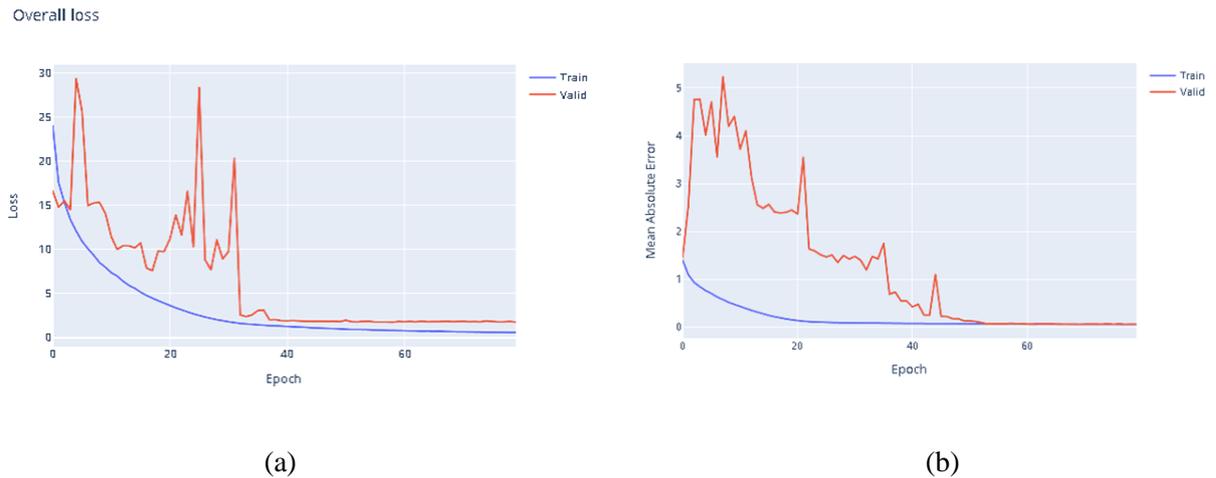
(b)

(a) Overall loss

(b) Mean Absolute Error for age

Figure 3-9 : les graphes d'apprentissage pour l'entraînement fait sur la data FairFace.

C. FFHQ



(a) Overall loss

(b) Mean Absolute Error for age

Figure 3-10 : Les graphes d'apprentissage pour l'entraînement fait sur la data FFHQ.

En analysant les graphes de la perte globale (Overall loss), les courbes d'entraînement et de test pour chaque base de données présentent une tendance décroissante, ce qui indique que l'entraînement progresse et que le modèle apprend progressivement à mieux prédire les étiquettes. Vers l'époque 56, les courbes se stabilisent, ce qui suggère une convergence efficace vers une solution optimale.

De même, le mean absolute error présente une différence presque nulle entre la valeur prédite et la valeur réelle, ce qui indique que la prédiction est précise à environ 95%.

3.4.1.2 Taux de classification

Une fois les modèles entraînés sur chaque base de données individuellement, nous avons procédé à une phase d'évaluation en effectuant des tests sur la performance de classification. Nous avons varié les échantillons de données à partir des données disponibles et calculé le taux de classification pour chaque modèle et chaque test effectué. Cette approche nous a permis d'obtenir le tableau ci-dessus.

Tableau 3-5 : Résultats d'exactitude pour la prédiction de l'âge, du genre et de l'ethnicité au niveau de train et test sur les ensembles de données UTK, FairFace et FFHQ.

		Entraîner								
		UTK			FairFace			FFHQ		
		Âge	Genre	Ethnicité	Age	Genre	Ethnicité	Âge	Genre	Ethnicité
Tester	UTK	64%	90%	86%	42%	76%	69%	42%	76%	X
	FairFace	45%	90%	78%	49%	79%	77%	49%	79%	X
	FFHQ	53%	78%	X	38%	78%	X	47%	91%	X

3.4.1.3 F1 Score

Tableau 3-6 : Résultats du F1-score pour la prédiction de l'âge, du genre et de l'ethnicité sur les ensembles de données UTK, FairFace et FFHQ.

		Entraîner									
		UTK				FairFace				FFHQ	
		Male	0,91	Blanc	0,88	Male	0,79	Blanc	0,79	Male	0,79
Tester	UTK	Male		Noir	0,85			Noir	0,62		
			Asiatique	0,86	Asiatique	0,6	Femelle	0,72	Femelle	0,72	
		Femelle		Indien	0,77			Indien	0,58		
			Male	0,9	Blanc	0,8	Male	0,78	Blanc	0,88	male
	FairFace	Male		Noir	0,69			Noir	0,85		
			Asiatique	0,61	Asiatique	0,86	Femelle	0,77	Femelle	0,77	
		Femelle		Indien	0,58			Indien	0,78		
			Male	0,8	X		Male	0,78 <th colspan="2">X</th> <th>Male</th> <td>0,78</td>	X		Male
	FFHQ	Femelle	0,76	X		Femelle	0,77	X		Femelle	0,77

3.4.2 ResNet multi-output

3.4.2.1 Courbes d'apprentissage

A. UTKFace

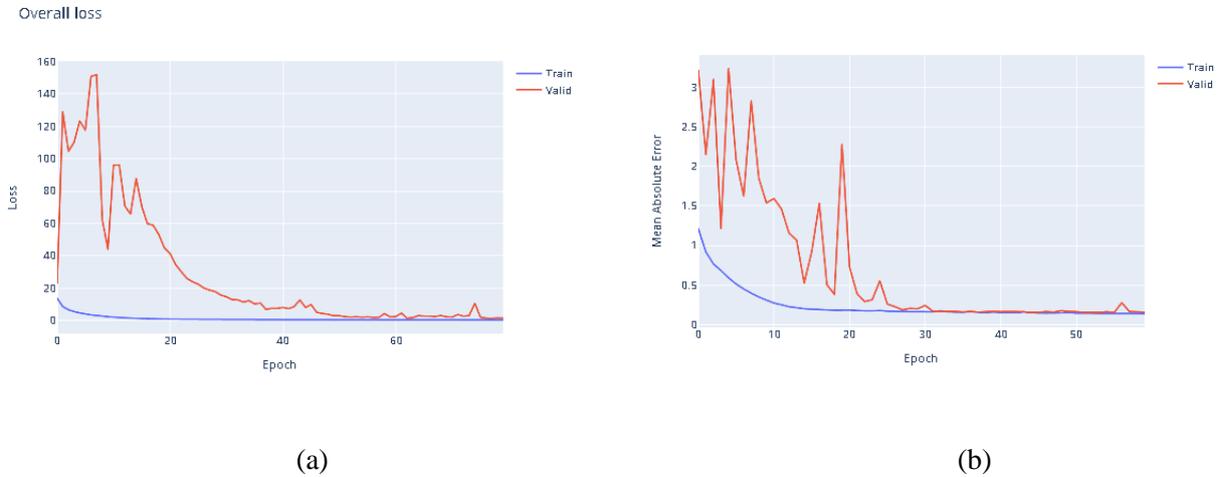
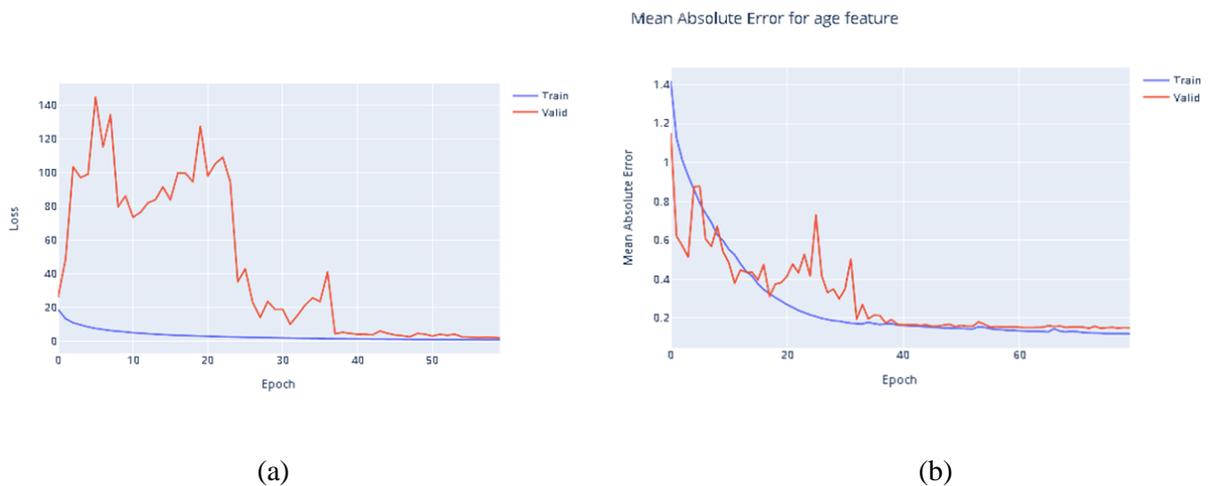


Figure 3-11 : Les graphes d'apprentissage pour l'entraînement fait sur la data UTKFace.

B. FairFace



(a) Overall loss

(b) Mean Absolute Error for age

Figure 3-12 : Les graphes d'apprentissage pour l'entraînement fait sur la data FairFace.

les graphes de la perte globale (Overall loss) et les courbes d'entraînement représente pratiquement presque les mêmes résultats obtenus au niveau du CNN .

3.4.2.2 Taux de classification

Tableau 3-7 : Résultats d'exactitude pour la prédiction de l'âge, du genre et de l'ethnicité au niveau de train et test sur les ensembles de données UTK, FairFace.

		Entraîner					
		UTK			FairFace		
		Âge	Genre	Ethnicité	Âge	Genre	Ethnicité
Tester	UTK	44%	91%	84%	42%	76%	69%
	FairFace	38%	90%	71%	49%	78%	86%

3.4.2.3 F1 Score

Tableau 3-9 : Résultats F1 pour la prédiction de l'âge, du genre et de l'ethnicité au niveau de train et test sur les ensembles de données UTK, FairFace.

		UTKFace			FairFace			
			Blanc	Noir		Blanc	Noir	Asiatique
UTK	Male	0,91	0,88	0,85	Male	0,79	0,79	0,65
			0,86				0,6	
	Femelle	0,91	0,77	Femelle	0,72	0,57		
			0,8			0,89		
FairFace	Male	0,9	0,69	Male	0,78	0,86		
			0,61			0,87		
	Femelle	0,89	0,58	Femelle	0,77	0,78		

3.5 Interprétations

En analysant et en comparant les résultats obtenus pour chaque ensemble de données à l'aide des métriques appropriées figurer dans les tableaux 3-9,3-8,3.-7et 3-6, nous avons pu tirer les interprétations suivantes.

En observant les performances des deux modèles, on remarque que le taux de variation des performances est toujours plus faible lors des tests par rapport à l'entraînement pour tous les ensembles de données et les modèles, à l'exception de la classe "genre" qui ne présente pas d'écart remarquable. Une analyse approfondie nous a conduit à conclure que la détection du genre n'est pas influencée par l'effet d'une nouvelle distribution ni par la distribution de luminosité.

En ce qui concerne la détection de l'âge, l'écart de variation est notable. Dans la majorité des cas, UTKFace était classé en premier, suivi de FFHQ, puis de FairFace. Cette inégalité

s'explique par le fait que l'ensemble de données UTKFace était initialement conçu pour prédire l'âge exact, ce qui lui confère une certaine précision.

En comparant les résultats des modèles CNN et ResNet, les deux types présentent des résultats presque similaires malgré leurs architectures différentes (ResNet étant une architecture plus profonde et plus complexe que CNN). Ainsi, on peut interpréter que la prédiction des caractéristiques démographiques ne nécessite pas nécessairement une architecture profonde avec de nombreuses couches.

Il convient de justifier la différence de résultats dans la prédiction de l'âge en mentionnant que l'ensemble de données UTKFace était spécifiquement destiné à estimer l'âge exact, tandis que FairFace et FFHQ étaient regroupés par tranches d'âge. Il est également important de souligner que les plages d'âge de FairFace et FFHQ n'étaient pas les mêmes.

Il est également pertinent de mentionner que les données de FFHQ ne sont pas seulement les plus récentes, mais elles présentent également une meilleure qualité d'image.

L'hypothèse avancée dans la thèse est pratiquement et théoriquement validée à 100 %, ce qui confirme que le changement de distribution et de luminosité a un impact important sur les mesures de performance des deux architectures CNN et ResNet.

3.6 Conclusion

Dans ce dernier chapitre, nous avons présenté les résultats obtenus à partir des modèles d'apprentissage dédiés à la prédiction des caractéristiques démographiques.

Dans notre étude, nous avons examiné l'impact de la distribution des données sur les modèles d'apprentissage afin de démontrer leur sensibilité. Cette analyse a été réalisée dans le contexte de la prédiction simultanée de l'âge, du genre et de l'ethnicité. Les résultats obtenus par nos modèles, présentés dans ce chapitre, se sont révélés très satisfaisants après plusieurs tentatives infructueuses et diverses modifications apportées à la mise en œuvre des modèles, comme le confirment les tests effectués.

Ce que nous avons accompli permet de conclure sur la sensibilité des modèles d'apprentissage profond et ouvre de nouvelles perspectives de recherche pour améliorer leurs performances. L'évaluation des résultats, accompagnée d'une interprétation, confirme que la variation de la distribution des données par rapport à la luminosité impacte de manière significative les résultats de prédiction simultanée de l'âge, du genre et de l'ethnicité.

Conclusion Général

Notre projet vise à démontrer de manière objective l'impact de la variation de la distribution des données sur les modèles d'apprentissage profond, tels que les CNN multi-output et ResNet, en particulier, pour la détection et la prédiction de l'âge, du genre et de l'ethnicité.

Pour atteindre nos objectifs, nous avons commencé par présenter une introduction générale sur l'apprentissage automatique afin d'acquérir les bases de la détection et de la prédiction. Ensuite, nous avons effectué une recherche approfondie et une analyse des travaux antérieurs portant sur tous les aspects liés à notre sujet. Dans un deuxième temps, nous avons exposé la méthodologie de notre travail, ainsi que les différentes méthodes utilisées pour concevoir et évaluer notre système. Enfin, nous avons présenté notre implémentation et discuté des résultats obtenus.

Pour mener à bien notre travail, plusieurs étapes ont été réalisées. La première étape consistait à acquérir et traiter les ensembles de données afin de les adapter à notre problématique tout en préservant l'intégrité de l'apprentissage. La deuxième étape comprenait le calcul de la distribution de luminosité des jeux de données, ainsi que le choix des modèles d'apprentissage profond CNN et ResNet, et leur entraînement. Ensuite, chaque modèle a été évalué sur un autre ensemble de données provenant d'une autre distribution, afin de mesurer les performances des classifieurs sur des données non vues et provenant d'une autre distribution de luminosité. Les deux systèmes ont été testés et évalués à l'aide des métriques d'évaluation les plus courantes, et les résultats obtenus étaient très satisfaisants. La présence d'une variation a clairement démontré l'influence de la distribution, et plus spécifiquement de la distribution de luminosité, ce qui a renforcé notre hypothèse sur le sujet.

Tout au long de ce travail, nous avons rencontré de nombreuses difficultés et obstacles, notamment la difficulté de trouver des données éthiques et étiquetées pour toutes les caractéristiques démographiques visées, l'utilisation d'ordinateurs obsolètes par rapport aux dernières technologies, les limitations du GPU dans Google Colab et Kaggle, les entraînements chronophages, et une connexion internet instable.

Ce projet est particulièrement intéressant car il aborde un sujet qui n'a pas encore été largement exploré dans la littérature spécialisée sur l'apprentissage profond des modèles et leurs sensibilités. Nous avons ainsi la possibilité d'apporter une contribution originale et novatrice à ce domaine de recherche.

Dans le cadre de nos travaux futurs, nous nous efforcerons de développer une méthode robuste ou des adaptateurs qui permettront d'éliminer la sensibilité aux variations de distribution des données au sein du modèle, sans modifier les données elles-mêmes. Nous continuerons à explorer de nouvelles approches et à affiner nos modèles afin de garantir des résultats fiables et reproductibles.

Bibliographie :

- [1] Vincent Barra, Antoine Cornuéjols, Laurent Miclet, "Apprentissage artificiel : Concepts et algorithmes - de Bayes et Hume au deep learning," Eyrolles, [2021].
- [2] « Boughaba_Boukhris.pdf ». Consulté le : 26 février 2023. [En ligne]. Disponible sur : https://dspace.univ-ouargla.dz/jspui/bitstream/123456789/17195/1/Boughaba_Boukhris.pdf?fbclid=IwAR3eu8e33QrTzhLZXQ4kjt3n1vrMQ29ny9R4tsKKUx79_27e7C51CIDjshY.
- [3] J. M. John-Mathews, "L'interprétabilité en apprentissage machine. Un regard sur les réseaux de neurones artificiels profonds," 10.13140/RG.2.2.36513.02402, 2018.
- [4] Z. Si et al., « Enhancing Recommendation with Search Data in a Causal Learning Manner », ACM Trans. Inf. Syst., vol. 41, no 4, p. 111:1-111:31, avr. 2023, doi: 10.1145/3582425.
- [5] I. Sadgali, N. Sael, F. Benabbou, et A. E. Y. E. Foutayeni, « Etat d'art des techniques d'analyse de comportement client pour la détection de fraude par carte de crédit », Int. Conf. Res. Appl. Math. Comput. Sci., vol. 2020, 2020, Consulté le : 26 février 2023. [En ligne]. Disponible sur : <https://icramcs2020.sciencesconf.org/data/>.
- [6] Creswell, Antonia & White, Tom & Dumoulin, Vincent & Arulkumaran, Kai & Sengupta, Biswa & Bharath, Anil. (2017). Generative Adversarial Networks: An Overview. IEEE Signal Processing Magazine. 35. 10.1109/MSP.2017.2765202.
- [7] E. Mathieu-Dupas, « Algorithme des k plus proches voisins pondérés et application en diagnostic », présenté à 42èmes Journées de Statistique, 2010. Consulté le : 2 mars 2023. [En ligne]. Disponible sur : <https://inria.hal.science/inria-00494814>.
- [8] R. Rakotomalala, "Arbres de décision," Revue Modulad, vol. 33, pp. 163-187, 2005.
- [9] R. Pupale, « Support Vector Machines(SVM) — An Overview », Medium. Consulté le: 2 mars 2023. [En ligne]. Disponible sur: <https://towardsdatascience.com/https-medium-com-pupalerushikesh-svm-f4b42800e989>
- [10] S. Patel, « Chapter 1 : Supervised Learning and Naive Bayes Classification — Part 1 (Theory) », Machine Learning 101. Consulté le: 4 mars 2023. [En ligne]. Disponible sur: <https://medium.com/machine-learning-101/chapter-1-supervised-learning-and-naive-bayes-classification-part-1-theory-8b9e361897d5>.
- [11] R. Sonmez, « Conceptual cost estimation of building projects with regression analysis and neural networks », Can. J. Civ. Eng., vol. 31, no 4, p. 677 -683, août 2004, doi: 10.1139/104-029.
- [12] E. Aytaç, « Unsupervised learning approach in defining the similarity of catchments: Hydrological response unit based k-means clustering, a demonstration on Western Black Sea Region of Turkey », Int. Soil Water Conserv. Res., vol. 8, no 3, p. 321 -331, sept. 2020, doi: 10.1016/j.iswcr.2020.05.002.

- [13] L. Spyrou, S. Kouchaki, et S. Sanei, « Multiview Classification and Dimensionality Reduction of Scalp and Intracranial EEG Data through Tensor Factorisation », *J. Signal Process. Syst.*, vol. 90, févr. 2018, doi: 10.1007/s11265-016-1164-z..
- [14] S. R. Eddy, « What is a hidden Markov model? », *Nat. Biotechnol.*, vol. 22, no 10, Art. no 10, oct. 2004, doi: 10.1038/nbt1004-1315.
- [15] J.-M. John-Mathews, « L'interprétabilité en apprentissage machine. Un regard sur les réseaux de neurones artificiels profonds. », 2018. doi: 10.13140/RG.2.2.36513.02402..
- [16] Z. Elloumi, B. Lecouteux, O. Galibert, et L. Besacier, « Prédiction de performance des systèmes de reconnaissance automatique de la parole à l'aide de réseaux de neurones convolutifs », *Rev. TAL*, 2018, Consulté le: 9 mars 2023. [En ligne]. Disponible sur: <https://hal.science/hal-01976284/>.
- [17] A. Schmitt, B. Le Blanc, M.-M. Corsini, C. Lafond, et J. Brůžek, « Les réseaux de neurones artificiels », *Bull. Mém. Société D'Anthropologie Paris BMSAP*, vol. 13, no 1-2, Art. no 1-2, juin 2001, doi: 10.4000/bmsap.4463..
- [18] Y. Sun, B. Xue, M. Zhang, G. G. Yen and J. Lv, "Automatically Designing CNN Architectures Using the Genetic Algorithm for Image Classification," in *IEEE Transactions on Cybernetics*, vol. 50, no. 9, pp. 3840-3854, Sept. 2020, doi: 10.1109/TCYB.2020.2983860.
- [19] S. Subedi, « Multi output neural network in Keras (Age, gender and race classification) », *Sanjaya's Blog*. Consulté le: 26 mars 2023. [En ligne]. Disponible sur: <https://sanjayasubedi.com.np/deeplearning/multioutput-keras/>.
- [20] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in *Proc. Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part IV, 2016*, pp. 630-645.
- [21] « Apprentissage automatique et biais : Impacts et solutions », IBM-France. Consulté le: 15 avril 2023. [En ligne]. Disponible sur: <https://www.ibm.com/blogs/ibm-france/2019/09/26/apprentissage-automatique-et-biais/>
- [22] M. Lango and J. Stefanowski, "What makes multi-class imbalanced problems difficult? An experimental study," *Expert Systems with Applications*, vol. 199, p. 116962, 2022. doi: 10.1016/j.eswa.2022.116962..
- [23] Georg-un, « Answer to “What does it mean the term variation for an image dataset?” », *Data Science Stack Exchange*. Consulté le: 15 avril 2023. [En ligne]. Disponible sur: <https://datascience.stackexchange.com/a/64303>.
- [24] « Traitement d'images ». Consulté le: 15 avril 2023. [En ligne]. Disponible sur: <https://web.maths.unsw.edu.au/~lafaye/CCM/video/traiting.htm>.
- [25] « Age Detection Model using CNN — a complete guide | by Skillcate AI | Medium ». Consulté le: 15 avril 2023. [En ligne]. Disponible sur: <https://medium.com/@skillcate/age-detection-model-using-cnn-a-complete-guide-7b10ad717c60>.
- [26] « Memoire Online - Etude des méthodes de reconnaissances d'empreinte digitale a l'aide du deep learning - Jean-Edmond DASSE ». Consulté le: 20 avril 2023. [En ligne]. Disponible sur:

<https://www.memoireonline.com/08/21/12173/Etude-des-methodes-de-reconnaissances-dempreinte-digitale-a-laide-du-deep-learning.html>

- [27] V. Janahiraman and P. Subramaniam, "Gender Classification Based on Asian Faces using Deep Learning," pp. 84-89, 2019. doi: 10.1109/ICSEngT.2019.8906399..
- [29] J. Buolamwini and T. Gebru, "Gender shades: Intersectional accuracy disparities in commercial gender classification," in Proc. Conference on Fairness, Accountability, and Transparency, 2018, pp. 77-91.
- [30] F. H. K. Zaman, "Gender classification using custom convolutional neural networks architecture," International Journal of Electrical and Computer Engineering (IJECE), vol. 10, no. 6, pp. 5758-5771, 2020.
- [31] A. Belhadi, Y. Djenouri, G. Srivastava, D. Djenouri, J. C. W. Lin, and G. Fortino, "Deep learning for pedestrian collective behavior analysis in smart cities: A model of group trajectory outlier detection," Information Fusion, vol. 65, pp. 13-20, 2021.
- [32] V. Sheoran, S. Joshi, and T. R. Bhayani, "Age and gender prediction using deep CNNs and transfer learning," in Proc. Computer Vision and Image Processing: 5th International Conference, CVIP 2020, Prayagraj, India, December 4-6, 2020, Revised Selected Papers, Part II, 2021, pp. 293-304.
- [33] K. Karkkainen and J. Joo, "Fairface: Face attribute dataset for balanced race, gender, and age for bias measurement and mitigation," in Proc. Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2021, pp. 1548-1558.
- [34] « UTKFace | Large Scale Face Dataset ». Consulté le: 29 avril 2023. [En ligne]. Disponible sur: <https://susanqq.github.io/UTKFace/>.
- [35] R. B. Yousuf, S. Biswas, K. K. Kaushal, J. Dunham, R. Gelles, S. Muthiah, N. Ramakrishnan, "Lessons from Deep Learning applied to Scholarly Information Extraction: What Works, What Doesn't, and Future Directions," arXiv preprint arXiv:2207.04029, 2022..
- [36] « FairFace Dataset | Papers With Code ». Consulté le: 29 avril 2023. [En ligne]. Disponible sur: <https://paperswithcode.com/dataset/fairface>.
- [37] H. Bai, D. Kang, H. Zhang, J. Pan, and L. Bao, "FFHQ-UV: Normalized Facial UV-Texture Dataset for 3D Face Reconstruction," in Proc. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 362-371.
- [38] PIV and Water Waves, vol. Volume 9. in Advances in Coastal and Ocean Engineering, no. Volume 9, vol. Volume 9. WORLD SCIENTIFIC, 2004. doi: 10.1142/5581.
- [39] R. Mathevet et al., « Une approche quantitative de la loi de Beer- Lambert avec un smartphone. Seconde partie. », Bull. Union Profr. Phys. Chim., déc. 2019, Consulté le: 03 mai 2023. [En ligne]. Disponible sur: <https://hal.science/hal-02355802>
- [40] R. Malamut, « Understanding the Profit and Loss Distribution of Trading Algorithms », janv. 2005, Consulté le: 03 mai 2023. [En ligne]. Disponible sur: https://www.academia.edu/71128669/Understanding_the_Profit_and_Loss_Distribution_of_Trading_Algorithms.

- [41] C. J. Willmott and K. Matsuura, "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance," *Climate Research*, vol. 30, no. 1, pp. 79-82, 2005.
- [42] A. M. Reza, "Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement," *Journal of VLSI Signal Processing Systems for Signal, Image and Video Technology*, vol. 38, pp. 35-44, 2004.
- [43] S. Kelly, "Python, PyGame and Raspberry Pi game development," Apress, Niagara Falls, ON, Canada, 2019.
- [44] D. M. Hawkins, "The problem of overfitting," *Journal of Chemical Information and Computer Sciences*, vol. 44, no. 1, pp. 1-12, 2004.
- [45] P. Cunningham and S. J. Delany, "Underestimation Bias and Underfitting in Machine Learning," in *Proceedings of the [nom de la conférence ou de la revue]*, vol. 12641, pp. 20-31, 2021.