République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

UNIVERSITE ECHAHID CHEIKH LARBI TEBESSI -TEBESSA-

Faculté des Sciences Exactes et des Sciences de la Nature et de la vie

Département d'Informatique

Mémoire de fin d'étude

Pour l'obtention du diplôme de MASTER

Domaine : Mathématiques et Informatique
Filière : Informatique
Option : Systèmes d'informations

Thème

Classification de catégorie d'âge par Deep Learning

Réalisé Par :
Adjal Chams

Devant le jury :

Dr. Menassel Rafik	*MCA*	*Université Larbi Tébessi*	*Président*
Dr. Aouine Mohamed	*MAA*	*Université Larbi Tébessi*	*Examinateur*
Mr. Abdelhakim Gharbi	*MAA*	*Université Larbi Tébessi*	*Encadreur*
Pr. Gattal Abdeldjalil	*Pr*	*Université Larbi Tébessi*	*Co-Encadreur*

Thème

Classification de catégorie d'âge par Deep Learning

Adjal Chams

Université de Larbi Tébessi

Remerciment

Avant toute chose, je tiens à exprimer ma profonde gratitude à Allah, pour m'avoir permis de mener à bien mon parcours de Master et de concrétiser ce travail, en me dotant de la volonté et de la patience nécessaires.

Mes remerciements les plus sincères et mon plus grand respect s'adressent à mes encadrants, ***Monsieur Abdelhakim Gharbi****. et* ***Monsieur Gattal Abdeldjalil*** *.Je leur suis infiniment reconnaissant pour la qualité exceptionnelle de leur encadrement, leur constante disponibilité, et la confiance inestimable qu'ils m'ont accordée tout au long de la réalisation de ce mémoire. Leur direction et leurs précieux conseils ont été déterminants pour l'aboutissement de mes travaux.*

Je souhaite également exprimer ma gratitude aux membres du jury pour avoir accepté d'examiner et d'évaluer ce travail. Leur expertise et leur temps sont grandement appréciés.

Je *remercie également sincèrement* ***les membres du Laboratoire de recherche LAVIA*** *pour leurs conseils avisés, leur accueil bienveillant, leur disponibilité, ainsi que pour m'avoir considérée comme l'une des leurs et soutenue tout au long de ce projet.*

Je tiens également à remercier chaleureusement mon amie ***Meriem Malak****, pour son soutien inconditionnel, ses encouragements précieux et son aide précieuse qui ont grandement contribué à maintenir ma motivation durant cette période.*

Dédicaces

إلى من زرعوا في قلبي بذور حب العلم والسعي نحو النجاح.. والدي ووالدتي الغاليين

إلى معلّمي الأوّل في هذه الحياة"أبي"

إلى الدفء والمحبة والحنان والعطف"أمّي"

إلى عزيزاتي ورفيقاتي اللاتي ساندني "أخواتي " أدامكن الله لي سندا لا يغيب

إلى عزّي في هذه الحياة" أخي"

إلى صغيرتي لونة

إلى من ارتاح قلبي برفقتها و طاب وقتي برفقتها "ريان"

إلى كل من زرع بسمة على وجهي في هذا المشوار ,شكرا لكم .

Résumé

La classification de l'âge et du genre à partir d'images faciales joue un rôle important dans divers domaines tels que la sécurité, l'interaction homme-machine et la personnalisation des services. Cette étude présente une approche basée sur l'apprentissage profond pour la classification de l'âge (en 8 catégories) et la prédiction du genre. En utilisant l'apprentissage par transfert, plusieurs modèles pré-entraînés (VGG16, ResNet101, DenseNet121) sont évalués sur l'ensemble de données UTKFace (23 708 images), selon deux protocoles. Afin d'améliorer les performances, un modèle hybride combinant VGG16 et DenseNet121 est proposé, tirant parti de la complémentarité des deux architectures. Ce modèle de fusion atteint une précision de 75,14 % pour la classification d'âge et 92,90 % pour la reconnaissance du genre, surpassant les modèles individuels.

Mots-clés : Classification de l'âge, Détection du genre, Apprentissage profond, Transfer Learning, Fusion de caractéristiques, CNN, UTKFace.

Abstract

The estimation of gender and age from faces is a basic necessity in its uses such as security, human-computer interaction, and personalized services. This work provides a deep learning method for age group classification (8 groups) and gender estimation. Transfer learning is adopted here for testing diverse pre-learned models such as VGG16, ResNet101, and DenseNet121 based on the UTKFace dataset (images: 23,708) for two protocols. This fusion between VGG16 and DenseNet121 is presented for achieving performance boost through their respective complementary feature representation. The fusion model achieves **75.14%** accuracy for age classification and **92.90%** for gender recognition, outperforming the individual models.

Keywords: Age classification, Gender recognition, Deep learning, Transfer learning, Feature fusion, CNN, UTKFace.

ملخص

يعدّ تصنيف العمر والجنس من صورة الوجه من المهام الأساسية في العديد من التطبيقات مثل الأمن، التفاعل بين الإنسان والآلة، والخدمات الشخصية. تهدف هذه الدراسة إلى تقديم إطار فعال قائم على التعلم العميق لتصنيف الفئات العمرية (8 فئات) والتعرف على الجنس انطلاقًا من صور الوجوه. تم الاعتماد على التعلم بالنقل (Transfer Learning) لتقييم أداء عدة نماذج مدرّبة مسبقًا من الشبكات العصبية الالتفافية (CNN) مثل VGG16 و ResNet101وDenseNet121، وذلك باستخدام قاعدة البيانات UTKFace التي تضم 23,708 صورة، وبالاعتماد على بروتوكولين تجريبيين. ولتحقيق أداء أفضل، تم اقتراح نموذج هجيني يعتمد على دمج معماريتي VGG16 وDenseNet121، للاستفادة من تكامل تمثيلاتهما المميزة للخصائص الوجهية. وقد حقق هذا النموذج دقة بلغت **75.14%**في تصنيف العمر و**92.90%** في تصنيف الجنس، متفوقًا بذلك على النماذج الفردية مثل ResNet101 الذي بلغ أداءه % 74.39 للعمر و % 92.24 للجنس.

الكلمات المفتاحية :تصنيف العمر، التعرف على الجنس، التعلم العميق، التعلم بالنقل، دمج الخصائص، الشبكات العصبية الالتفافية(CNN) ، قاعدة بيانات UTKFace .

Table des matières

Résumé ..... *i*
Abstract ..... *i*
ملخص ..... *ii*
Table des matières ..... *iii*
Liste de Figures ..... *vi*
Liste de Tableaux ..... *ix*
Introduction générale *1*
Chapitre 1 : Reconnaissance faciale pour l'estimation de l'âge *4*
1 Introduction *4*
2 Aperçu de la reconnaissance biométrique *4*
2.1 Définition 4
2.2 Les Différents Types de Biométrie 5
3 Reconnaissance faciale *6*
3.1 Définition 6
3.2 Historique 7
3.3 Étapes de reconnaissance facial 9
3.3.1 Détection du visage 9
3.3.2 Extraction des caractéristiques 9
3.3.3 Reconnaissance faciale 10
3.4 Classification des systèmes de reconnaissance faciale 10
3.4.1 Approches locales 10
3.4.2 Approches globales 10
3.4.3 Approches hybrides 11
3.5 Avantages de reconnaissance faciale 11
3.5.1 Efficacité accrue des systèmes modernes 11
3.5.2 Compatibilité avec les infrastructures existantes 12
3.5.3 Automatisation des processus 12
3.5.4 Amélioration de la sécurité globale 12
3.5.5 Élimination des fraudes liées au temps de travail 12
4 Estimation de l'âge à partir des images faciales *12*
4.1 Principe et importance 12
4.2 Algorithmes d'estimation de l'âge 12
4.2.1 Méthodes Basées sur la Classification 13
4.2.2 Méthodes Basées sur la Régression 14
4.2.3 Méthodes Basées sur le Classement (Ranking-Based) 14
4.2.4 Méthodes Hybrides (Classification + Régression) 14

4.2.5 Méthodes Basées sur la Classification Probabiliste (Soft Classification) ... 14
4.3 Applications réelles de l'estimation de l'âge ... 15
4.3.1 Simulation de l'âge ... 15
4.3.2 Gestion électronique de la relation client. ... 15
4.3.3 Contrôle d'accès et surveillance ... 15
4.3.4 Interaction humain-ordinateur spécifique à l'âge ... 15
4.3.5 Biométrie ... 15
4.3.6 Personnes disparues ... 16
4.4 Défis courants ... 16
4.4.1 Pose de la tête et alignement ... 17
4.4.2 Résolution de l'image ... 17
4.4.3 Modifications Faciales ... 18
4.4.4 Génétique ... 18
4.4.5 Manque de Données ... 18
4.4.6 Le mode de vie et l'état de santé ... 19
5 Conclusion ... 19
Chapitre 2 : État de l'art ... *21*
1 Introduction ... 21
2 Définition de l'intelligence artificielle (IA) ... 21
2.1 Apprentissage automatique (Machine learning) ... 22
2.2 Apprentissage Profond (Deep learning) ... 23
2.3 Réseaux de neurones convolutifs (CNN) ... 23
2.3.1 Apprentissage par transfert (Transfer Learning) ... 24
2.3.2 Les Principaux Modèles CNN Pré-entraînés en Apprentissage Profond ... 24
3 Jeux de données utilisés pour l'estimation de l'âge ... 25
3.1 FG-NET (Face and Gesture Recognition Network) ... 25
3.2 MORPH ... 26
3.3 IMDB-WIKI ... 26
3.4 AFAD ... 27
3.5 OIU-Adience ... 27
3.6 CACD ... 27
4 Travaux connexes ... 27
4.1 Approches basées sur l'apprentissage automatique ... 28
4.2 Approches basées sur l'apprentissage profond ... 29
5 Tableau récapitulatif des méthodes existantes ... 32
6 Synthèse ... 34
7 Conclusion ... 36
Chapitre3 : Méthode proposée pour la classification de l'âge et du genre à partir d'images faciales ... *38*
1 Introduction ... 38
2 Le corpus UTKFace ... 38
2.1 Description d'ensemble de données UTKFace ... 38

2.2 Protocoles de test utilisés 40
2.3 Les défis de UTKface 41
2.3.1 Données non équilibrées (déséquilibre de classes) 41
2.3.2 Qualité d'image variable 42
2.3.3 Mauvaise annotation 42
3 Prétraitement et organisation des données *44*
4 ENVIRONNEMENT DE TRAVAIL *45*
4.1 Kaggle Notebooks 45
4.2 Le langage de programmation Python 46
4.3 Les bibliothèques de programmation Python 46
5 Conception et implémentation de notre méthode *47*
5.1 Architecture du système de classification de l'âge et du genre proposée 47
5.1.1 Architecture proposée basée Transfer Learning 48
5.1.2 Architecture base de fusion des caractéristiques extraites par deux modèles 49
5.2 Résultats expérimentaux et discussions 51
5.2.1 Résultats expérimentaux des modèles de Transfer Learning 51
5.2.2 Résultats expérimentaux du modèle de fusion VGG16–DenseNet121 57
6 Comparaison des résultats avec les travaux existants *62*
7 Conclusion *63*
Chapitre4 : Système de détection des catégories d'âge et genre selon des images faciales *65*
1 Introduction *65*
2 Description du système automatique de détection d'âge et genre *65*
3 La structure de devine mon âge & genre *68*
3.1 L'interface de système 68
3.2 Résultats de prétraitement 69
3.3 Résultats Obtenus selon des images faciale téléchargé 69
3.4 Résultats Obtenus selon des images faciale en temps réel 71
4 Conclusion *71*
Conclusion générale *72*
Bibliographies *74*

Liste de Figures

Figure 1-1 vérification et identification biométrique.

Figure1-2 Classification des caractéristiques biométriques : physiologiques et comportementales

Figure 1-3 Technologie de reconnaissance faciale.

Figure 1-4 Historique de reconnaissance faciale**.**

Figure 1-5 Étapes de reconnaissance faciale**.**

Figure 1-6 Méthodes de reconnaissance faciale.

Figure 1-7 Système d'Estimation de l'Âge.

Figure 1-8 Défis de l'estimation de l'âge**.**

Figure 1-9 Impact de la pose de la tête et de l'alignement sur l'estimation de l'âge

Figure 1-10 Résolution de l'image et précision du modèle**.**

Figure 1-11 Modifications faciales.

Figure1-12 Facteurs influençant le vieillissement facial : génétique, mode de vie et état de santé.

Figure 2-1 Architecture de réseau de neurones convoulutifs.

Figure 3-1 Distribution des âges dans l'ensemble de données UTKFace.

Figure 3-2 Distribution du genre dans l'ensemble de données UTKFace.

Figure 3-3 Exemples de visages issus de différentes tranches d'âge dans la base UTKFace.

Figure 3-4 Répartition des images selon les groupes d'âge dans UTKFace.

Figure 3-5 Exemple des images floues, mal exposées ou de faible résolution.

Figure 3-6 Annotation erronée de l'âge : 39 ans en 2017, âge réel 37 ans en 2025.

Figure 3-7 Annotation incorrectes du genre.

Figure 3-8 Processus de préparation des données.

Figure 3-9 Architecture de classification des attributs faciaux : âge et genre.

Figure 3-10 Architecture des modèles CNN pré-entraînés utilisés dans l'étude.

Figure 3-11 Architecture combinée VGG16–DenseNet121 pour la prédiction du genre et de l'âge.

Figure 3-12 Comparaison de l'Accuracy des modèles CNN pré-entraînés pour la classification de l'âge (Protocole 1).

Figure 3-13 Comparaison de l'Accuracy des modèles CNN pré-entraînés pour la classification de l'âge (Protocole 2).

Figure 3-14 Comparaison de l'Accuracy des modèles CNN pré-entraînés pour la classification de genre (Protocole 2).

Figure 3-15 Résultats obtenus pour notre model.

Figure 3-16 Graphe de précision et de perte du notre modèle pour la classification de l'âge.

Figure 3-17 La matrice de confusion du notre modèle pour la classification de l'âge.

Figure 3-18 Rapport de classification de model fusion vgg16 et denceNet121 pour classification de genre**.**

Figure 3-19 Graphe de précision et de perte de notre modèle pour la classification du genre.

Figure 3-20 Matrice de confusion de notre modèle pour la classification du genre.

Figure 4-1 Présentation du système devine mon âge & genre.

Figure 4-2 Interface principale.

Figure 4-3 Traitement des images faciales par l'application

Figure 4-4 Quelques exemples de détection d’âge et genre selon des images faciale.

Figure 4-5 Exemple de détection d’âge et genre selon des images faciale en temps réel

Liste de Tableaux

Tableau 2-1 Tableau récapitulatif des travaux connexes.

Tableau 3-1 Distribution des classes d'âges.

Tableau 3-2 Distribution de genre.

Tableau 3-3 Configuration d'entraînement utilisée dans les différentes expérimentations.

Tableau 3-4 Résultats des modèles CNN pré-entraînés pour la classification de l'âge selon le Protocole.

Tableau 3-5 Performances des modèles CNN pré-entraînés pour la classification de l'âge selon le Protocole 2.

Tableau 3-6 Performances des modèles CNN pré-entraînés pour la classification du genre selon le Protocole 2.

Tableau 3-7 Tableau comparatifs des résultats obtenus par rapports aux travaux existants.

Introduction générale

La reconnaissance faciale constitue un domaine de recherche en pleine expansion, largement exploité dans divers secteurs tels que la sécurité, la santé, le divertissement, la cosmétique et la biométrie. Son principe repose sur la capacité à identifier une personne en analysant les caractéristiques physiques de son visage. Parmi ses nombreuses applications, l'estimation automatique de l'âge et du genre suscite un intérêt croissant en raison de ses implications dans des contextes variés tels que le contrôle d'accès, l'interaction homme-machine, le marketing ciblé, la surveillance et la personnalisation des services. Les êtres humains sont naturellement capables d'évaluer l'âge et le genre d'une personne à partir de son apparence, en se basant sur des éléments comme la structure faciale, les rides, la coiffure ou la texture de la peau. En revanche, cette tâche s'avère plus complexe pour une machine, nécessitant un processus d'apprentissage à partir d'un grand nombre d'exemples. Pour ce faire, il est essentiel de disposer d'un ensemble de données correctement annotées, ainsi que d'algorithmes capables de déceler et d'identifier les motifs visuels caractéristiques de chaque catégorie. Historiquement, les premières approches se basaient sur des méthodes dites "manuelles", où des caractéristiques spécifiques du visage étaient extraites et analysées. Toutefois, ces techniques présentaient des limites dans des conditions réelles, en raison des variations de pose, d'éclairage ou d'expression. Aujourd'hui, grâce à l'avancée des techniques d'apprentissage profond, notamment des réseaux de neurones convolutifs (CNN), il est possible d'automatiser l'extraction de ces caractéristiques et d'améliorer significativement la précision des systèmes de reconnaissance faciale.

Problématique

La problématique centrale de ce projet réside dans la difficulté de réaliser de façon automatique l'estimation de l'âge et du genre à partir d'images faciales. Cette tâche est fortement impactée par de nombreux facteurs tels que la variabilité des traits faciaux entre individus, sexes et ethnies, les conditions de capture des images (éclairage, angle, résolution), ainsi que des éléments génétiques et des modes de vie qui influent sur l'apparence physique. Dans ce contexte, plusieurs questions se posent:

- *Comment concevoir un modèle capable d'estimer avec précision l'âge et le genre malgré la grande variabilité des visages humains ?*
- *Comment assurer une bonne généralisation du système face aux variations des conditions de prise de vue ?*
- *Quels types d'architectures d'apprentissage profond permettent de mieux capturer les caractéristiques pertinentes pour ces deux tâches ?*
- *Et enfin, comment intégrer efficacement ce type de solution dans des systèmes intelligents réels, tout en maintenant un équilibre entre performance et complexité ?*

Objective

L'objectif de ce travail est de développer un système intelligent et précis capable de classer automatiquement les individus selon leur tranche d'âge et leur genre, en s'appuyant sur des techniques d'apprentissage profond. Plus précisément, il s'agit de comparer et d'évaluer les performances de différentes architectures de réseaux de neurones dans le cadre de la classification de l'âge, tout en intégrant la reconnaissance du genre comme composante complémentaire du système, d'améliorer la robustesse et la précision du système dans des conditions réalistes, et de proposer une solution pratique pouvant être intégrée dans des applications intelligentes relevant de domaines variés tels que la sécurité, le marketing ou l'interaction homme-machine.

Plan du mémoire

Ce mémoire est structuré en quatre chapitres principaux :

- ***Chapitre 1 :*** *Reconnaissance faciale pour l'estimation de l'âge.*

- ***Chapitre 2 :*** *État de l'art.*

- ***Chapitre 3 :*** *Méthode proposée pour la classification de l'âge et du genre à partir d'images faciales.*

- ***Chapitre 4 :*** *Système de détection des catégories d'âge et de genre selon des images faciale*.

Chapitre 1 Reconnaissance faciale faciale pour l'estimation de l'âge

Chapitre 1

Chapitre 1 : Reconnaissance faciale pour l'estimation de l'âge

1 Introduction

La reconnaissance faciale est l'une des pierres angulaires de la biométrie, car elle permet d'identifier ou de vérifier une personne uniquement par son visage. Aujourd'hui, grâce aux progrès de l'intelligence artificielle et de la vision par ordinateur, elle s'est glissée dans presque tous les secteurs, de la sécurité au contrôle des appareils. Parmi toutes ces utilisations, estimer l'âge d'un utilisateur reste l'un des défis les plus délicats. Bien que les rides, les joues et d'autres traits révèlent une certaine maturité, les différences entre chaque visage, l'impact du temps et même l'éclairage rendent le travail incertain. L'idée est donc d'étudier ces caractéristiques avec des modèles d'apprentissage automatique et d'apprentissage profond pour donner une estimation aussi juste que possible de l'âge biologique ou apparent de la personne.

Ce chapitre présente un aperçu général de la reconnaissance biométrique et ses différentes approches, en mettant particulièrement l'accent sur la reconnaissance faciale. Nous explorerons son fonctionnement, ses méthodes d'analyse et son évolution historique. Ensuite, nous détaillerons les méthodes d'estimation de l'âge basées sur les images faciales, en abordant leurs principes, les algorithmes utilisés, ainsi que les défis et applications associés.

2 Aperçu de la reconnaissance biométrique

2.1 Définition

La biométrie est une discipline qui mesure et analyse les caractéristiques biologiques d'un individu pour son authentification ou identification. Elle repose sur l'exploitation d'attributs uniques, qu'ils soient physiologiques ou comportementaux, afin d'assurer une reconnaissance fiable et sécurisée [1].

Deux approches sont utilisées :

- La vérification biométrique (1:1) compare les données d'un individu à un seul profil enregistré pour confirmer son identité.
- L'identification biométrique (1:M) confronte ces données à l'ensemble des profils d'une base pour retrouver l'identité correspondante.

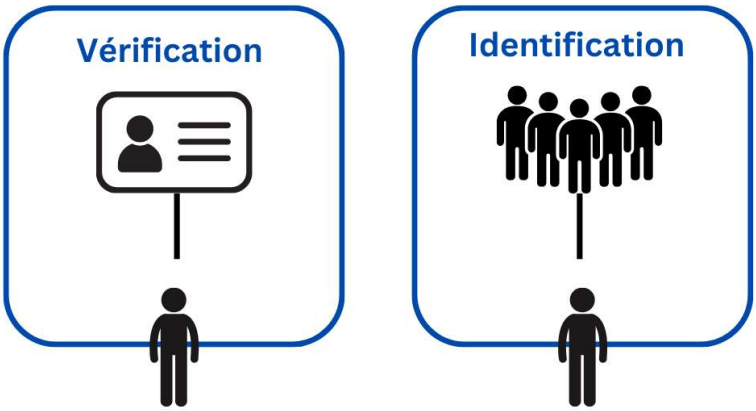


Figure 1-1 vérification et identification biométrique

Ces méthodes garantissent une identification précise et adaptée aux besoins de sécurité et de contrôle d'accès[2].

2.2 Les Différents Types de Biométrie

Les systèmes biométriques sont généralement divisés en deux catégories principales : la biométrie physiologique et la biométrie comportementale.

L'identification et l'authentification précises d'un individu reposent sur l'analyse de ses attributs biométriques, qui sont des caractéristiques uniques et mesurables. L'adoption de la technologie biométrique permet une reconnaissance et une validation automatiques des individus en se basant sur leurs traits physiologiques et comportementaux. Parmi ces caractéristiques, on retrouve notamment la reconnaissance faciale, l'empreinte palmaire, l'iris, la rétine, l'empreinte digitale, la voix, la signature manuscrite, la démarche, la géométrie de la main, l'écriture manuscrite, et la dynamique de frappe au clavier [2], comme illustré dans la figure 1-2.

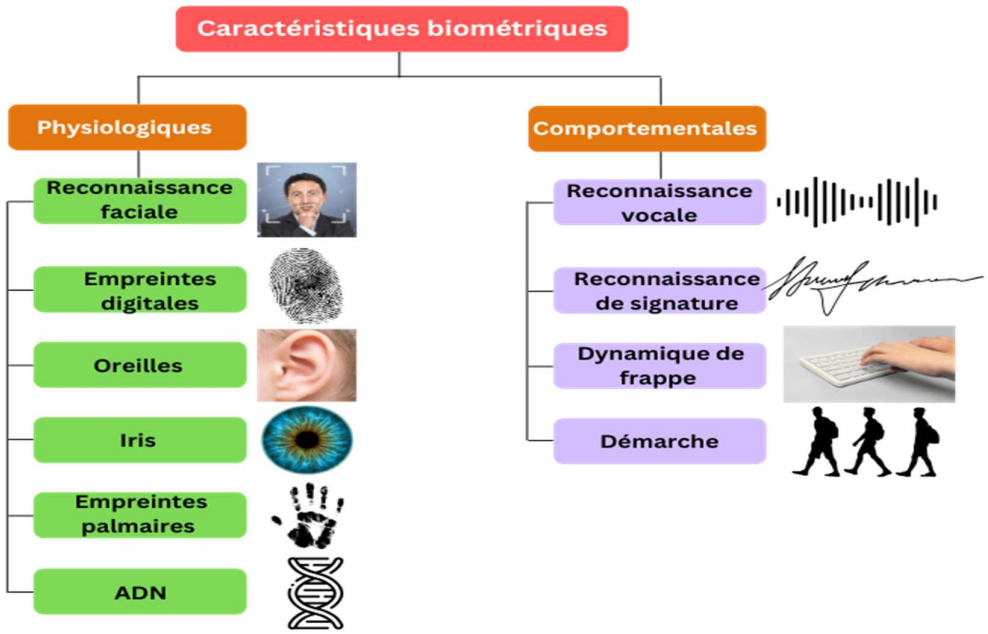


Figure 1- 2 Classification des caractéristiques biométriques : physiologiques et comportementales

3 Reconnaissance faciale

3.1 Définition

La reconnaissance faciale est une technologie biométrique avancée permettant l'identification automatique des individus tout en limitant l'intervention humaine. Largement adoptée dans les services financiers, elle optimise la sécurité et l'efficacité des processus en intégrant des solutions comme la vérification faciale et l'identification des clients. Utilisée pour

contrôler l'accès aux espaces physiques et numériques, elle analyse et compare les traits du visage afin de confirmer une identité à partir d'images ou de vidéos. Son fonctionnement sans contact et sa capacité d'authentification à distance en font un outil essentiel dans des domaines tels que la finance intelligente et les services numériques, garantissant ainsi une gestion sécurisée et fluide des interactions.[3][4]



Figure 1-3 Technologie de reconnaissance faciale

3.2 Historique

La reconnaissance faciale a connu une évolution marquante depuis ses débuts dans les années 1960 jusqu'à son intégration dans diverses applications contemporaines.

- 1964 : Naissance de la reconnaissance faciale

  Première tentative de programmation informatique pour identifier des visages.Woodrow Bledsoe, Helen Chan et Charles Bisson développent une méthode semi-automatique : des opérateurs saisissent manuellement des mesures faciales (taille des yeux, bouche, etc.).

- 1991 : Révolution des Eigenfaces

  Alex Pentland et Matthew Turk (MIT) présentent Eigenfaces, une méthode basée sur l'analyse en composantes principales (ACP). Automatisation partielle grâce à des modèles statistiques.

- 2011 : Ère du Deep Learning

Le deep learning et les réseaux de neurones artificiels remplacent les méthodes manuelles.L'ordinateur apprend seul à identifier les traits pertinents, boosté par des millions d'images.

- 2014 : Facebook et DeepFace

Facebook lance DeepFace, un algorithme atteignant 97 % de précision, rivalisant avec l'œil humain. Popularisation de la reconnaissance faciale via les réseaux sociaux.

- 2018 : Surveillance de masse en Chine

La police chinoise arrête un suspect dans une foule de 50 000 personnes grâce à une reconnaissance faciale en direct.Utilisation massive dans un cadre sécuritaire et contrôle social.

- 2019 : Biométrie obligatoire

Des expérimentations de reconnaissance faciale ont été menées à Nice, et la Chine a exigé la numérisation des visages pour l'achat de nouveaux téléphones .[5]

- 2024 : Meta (anciennement Facebook) annonce l'intégration de la reconnaissance faciale dans sa modération automatisée pour détecter les publicités frauduleuses utilisant l'image de célébrités.[6]

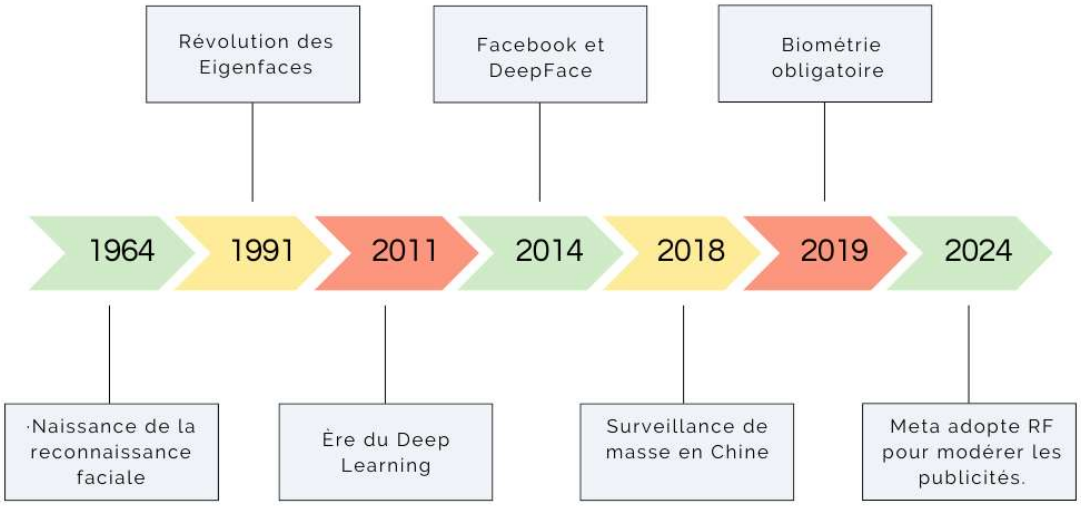


Figure 1-34 Historique de reconnaissance faciale

La reconnaissance faciale continue d'évoluer, s'intégrant dans de nombreux aspects de la vie quotidienne, tout en soulevant des questions éthiques et légales concernant la vie privée et la surveillance.

3.3 Étapes de reconnaissance facial

La reconnaissance faciale repose sur trois étapes fondamentales :

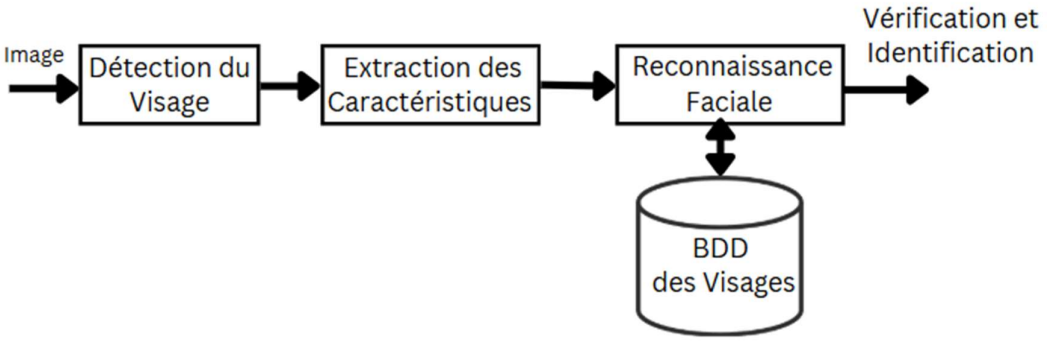


Figure1-5 Étapes de reconnaissance faciale

3.3.1 Détection du visage

Cette étape permet d'identifier et de localiser un visage humain dans une image. Plusieurs techniques sont utilisées, comme le détecteur de Viola-Jones, l'Histogramme des Gradients Orientés (HOG) et l'Analyse en Composantes Principales (PCA). La détection du visage peut aussi servir pour la classification d'images et de vidéos, la reconnaissance d'objets ou encore l'identification de zones spécifiques dans une image.[7]

3.3.2 Extraction des caractéristiques

Une fois le visage détecté, cette étape consiste à analyser et extraire les traits importants du visage, comme les yeux, le nez et la bouche, ainsi que leur disposition. Ces informations sont stockées sous forme d'une "empreinte faciale" unique. Différentes méthodes permettent d'extraire ces caractéristiques, notamment Eigenface, ICA, LDA, SIFT, les filtres de Gabor, les ondelettes de Haar et les motifs binaires locaux (LBP).[7]

3.3.3 Reconnaissance faciale

Étape compare l'empreinte faciale extraite avec celles enregistrées dans une base de données pour identifier la personne. Parmi les méthodes utilisées, on retrouve les filtres de corrélation, les réseaux de neurones convolutifs (CNN) et l'algorithme des k-plus proches voisins (KNN).[7]

3.4 Classification des systèmes de reconnaissance faciale

Comparé à d'autres systèmes biométriques tels que la reconnaissance de l'iris, des yeux ou des empreintes digitales, le système de reconnaissance faciale n'est pas le plus performant ni le plus fiable. De plus, malgré ses nombreux avantages, il est confronté à plusieurs contraintes et défis. Les systèmes de reconnaissance faciale peuvent être classés en trois catégories principales en fonction de leur méthode de détection et de reconnaissance :

3.4.1 Approches locales

Ces approches reposent sur l'extraction de caractéristiques spécifiques à certaines régions du visage, telles que les yeux, la bouche et le nez. Elles se basent sur des techniques de traitement d'image permettant d'identifier des points d'intérêt et d'en extraire des descripteurs. Parmi les méthodes les plus courantes, on retrouve le Local Binary Pattern (LBP), le Histogram of Oriented Gradients (HOG), ainsi que des algorithmes robustes comme le Scale-Invariant Feature Transform (SIFT) et le Speeded-Up Robust Features (SURF). Bien que ces approches offrent une certaine robustesse aux variations d'éclairage et aux changements de pose, elles restent sensibles aux occlusions et aux expressions faciales.[8]

3.4.2 Approches globales

Contrairement aux approches locales, les approches globales exploitent l'ensemble du visage comme entrée et projettent les données dans un espace de dimensions réduites pour en extraire les principales caractéristiques. Parmi les techniques les plus utilisées, on retrouve l'Analyse en Composantes Principales (PCA)**,** l**'**Analyse Linéaire Discriminante (LDA) et l**'**Analyse en Composantes Indépendantes (ICA). Plus récemment, les réseaux de neurones convolutifs (CNNs) ont démontré une grande efficacité dans l'apprentissage des représentations

faciales. Toutefois, ces approches restent sensibles aux variations d'éclairage, aux expressions faciales et aux différences de pose.[8]

3.4.3 Approches hybrides

Afin d'améliorer la précision et la robustesse des systèmes de reconnaissance faciale, certaines méthodes combinent les avantages des approches locales et holistiques. Ces approches hybrides exploitent à la fois les caractéristiques locales et globales du visage pour affiner les résultats de classification. Parmi les exemples notables, on trouve l'association des Gabor Filters avec LDA, l'intégration des CNNs avec des Auto-Encoders, ou encore l'utilisation conjointe de descripteurs locaux et de méthodes d'apprentissage profond**.**[8]

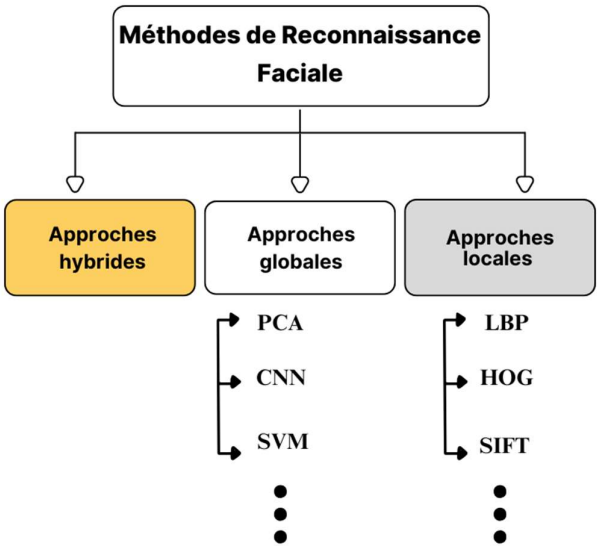


Figure 1-6 Méthodes de reconnaissance faciale

3.5 Avantages de reconnaissance faciale

3.5.1 Efficacité accrue des systèmes modernes

Les technologies de reconnaissance faciale actuelles, notamment celles utilisant des modèles 3D, atteignent des taux de précision élevés. Cette fiabilité réduit les risques de contournement du système, permettant un suivi rigoureux des présences et une gestion sécurisée des accès.

3.5.2 Compatibilité avec les infrastructures existantes

Ces systèmes s'adaptent facilement aux environnements logiciels déjà en place dans les organisations, minimisant les coûts et les perturbations liés à leur mise en œuvre.

3.5.3 Automatisation des processus

L'absence de nécessité d'une surveillance humaine constante en fait une solution pratique pour les entreprises, réduisant les erreurs et les coûts opérationnels.

3.5.4 Amélioration de la sécurité globale

Au-delà du suivi des employés, ces systèmes permettent d'enregistrer et de contrôler les déplacements des visiteurs dans des zones définies. L'accès est strictement limité aux personnes préalablement authentifiées, éliminant les intrusions non autorisées.

3.5.5 Élimination des fraudes liées au temps de travail

La vérification biométrique par scan facial supprime les possibilités de fraude, comme l'enregistrement de présence par un tiers, garantissant une traçabilité individuelle et immuable.

4 Estimation de l'âge à partir des images faciales

4.1 Principe et importance

L'estimation de l'âge à partir du visage repose sur l'analyse des traits faciaux pour estimer l'âge d'un individu. Le visage contient une richesse d'informations telles que l'individualité, l'émotion, l'attitude, le niveau de maturité, l'ethnicité et le genre, ce qui en fait un indicateur clé pour la modélisation de l'âge. Cette identification repose sur l'extraction et l'analyse de caractéristiques spécifiques du visage, permettant de classer les individus dans différentes tranches d'âge.[9]

4.2 Algorithmes d'estimation de l'âge

L'estimation de l'âge consiste à attribuer automatiquement un âge précis ou une catégorie d'âge à un visage humain, qu'il s'agisse de l'âge biologique, de l'âge perçu ou d'une approximation. Cette technologie s'adapte à différentes approches algorithmiques lors de sa

phase finale. Selon le modèle choisi, le problème peut être traité comme une classification en plusieurs catégories. Inversement, il peut être abordé sous forme de régression lorsque les âges sont interprétés comme des valeurs continues. La combinaison de ces deux méthodes permet de concevoir un système hybride plus fiable[10].

D'autres approches existent également, comme le classement ordinal, qui exploite la relation d'ordre entre les âges et la classification douce qui attribue une distribution de probabilités aux âges au lieu d'une valeur unique. Ces différentes méthodes permettent d'optimiser l'estimation de l'âge en fonction des contraintes et des objectifs du système développé.[11]. Une illustration schématique de ce processus est présentée dans la figure 7.

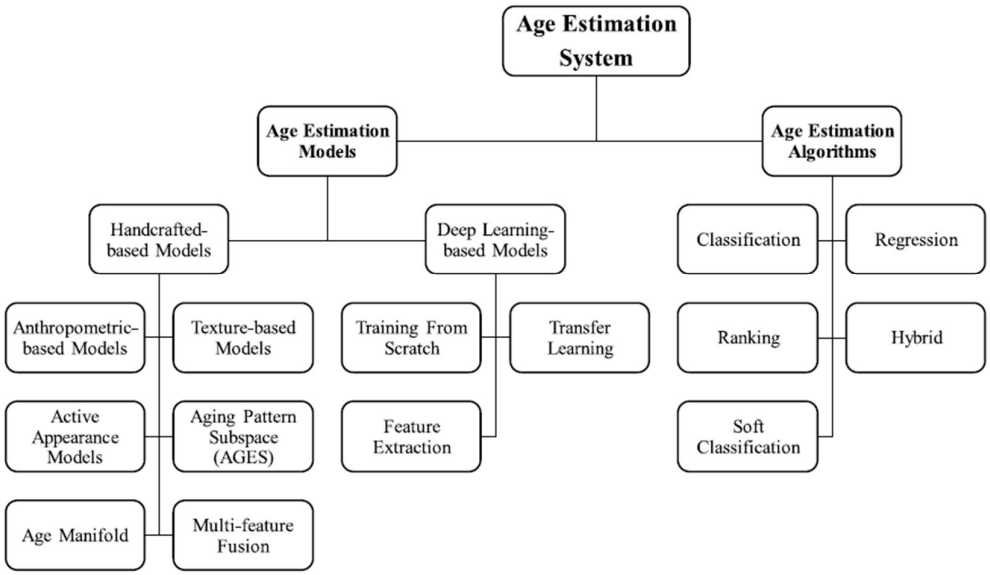


Figure 1-7 Système d'Estimation de l'Âge [10]

4.2.1 Méthodes Basées sur la Classification

Les algorithmes de classification considèrent l'âge comme une classe discrète et entraînent un modèle pour classer un visage dans un groupe d'âges prédéfini. Par exemple, machine à vecteurs de support (SVM) et réseaux de neurones convolutifs (CNN) sont souvent utilisés pour cette tâche. Dans ce cas, le modèle apprend à reconnaître les motifs faciaux associés à chaque tranche d'âge et prédit l'âge en sélectionnant la classe la plus probable.

4.2.2 Méthodes Basées sur la Régression

Contrairement à la classification, les algorithmes de régression estiment l'âge comme une valeur continue au lieu d'une classe**.** Les modèles comme la régression linéaire, la régression par machines à vecteurs de support (SVR) et les réseaux neuronaux profonds apprennent une fonction qui mappe les traits du visage à un âge numérique. Cette approche est plus précise que la classification stricte, car elle tient compte de la continuité du vieillissement.

4.2.3 Méthodes Basées sur le Classement (Ranking-Based)

Ces méthodes formulent l'estimation de l'âge comme une série de décisions binaires ("L'âge est-il supérieur à X ans ?"). Plutôt que de prédire un âge directement, les algorithmes de classement par hyperplans ordonnés (Ordinal Hyperplane Ranking) séparent progressivement les âges en différents niveaux, permettant ainsi une meilleure modélisation des relations entre classes. Cette approche est plus précise que la classification simple et plus robuste que la régression pure, mais elle nécessite un entraînement plus complexe et plus de ressources de calcul.

4.2.4 Méthodes Hybrides (Classification + Régression)

Les méthodes hybrides combinent classification et régression pour améliorer la robustesse des prédictions. Un exemple courant consiste à utiliser un CNN pour extraire des caractéristiques faciales, suivi d'un classificateur SVM pour estimer la tranche d'âge, puis d'un réseau neuronal régressif pour affiner l'estimation de l'âge exact. Une autre approche hybride consiste à classifier d'abord le visage dans une grande catégorie (enfant, adulte, senior) avant d'affiner la prédiction par régression.

4.2.5 Méthodes Basées sur la Classification Probabiliste (Soft Classification)

Classification Probabiliste est une approche avancée où chaque âge est représenté par une distribution de probabilités plutôt qu'une seule classe. Par exemple, un modèle peut prédire qu'un individu a 60 % de chances d'avoir 25 ans, 30 % d'avoir 24 ans, et 10 % d'avoir 26 ans. Cette technique est souvent implémentée avec un réseau de neurones avec une sortie softmax, qui génère une probabilité pour chaque âge possible. L'estimation finale peut être obtenue par une moyenne pondérée des âges prédictifs.

4.3 Applications réelles de l'estimation de l'âge

4.3.1 Simulation de l'âge

La simulation de l'âge requiert l'analyse de l'apparence faciale d'un individu à diverses étapes de sa vie. Ces données permettent de reconstituer numériquement l'évolution du visage au fil du temps. L'estimation de l'âge contribue ainsi à identifier les schémas de vieillissement propres à chaque personne, optimisant la précision des modèles prédictifs.

4.3.2 Gestion électronique de la relation client.

Dans les plateformes commerciales en ligne, où l'interaction client est dématérialisée, l'estimation de l'âge devient un outil stratégique. En déterminant l'âge des utilisateurs, elle permet d'adapter les services et les offres à leurs attentes démographiques, renforçant ainsi l'engagement et la personnalisation des expériences .[11]

4.3.3 Contrôle d'accès et surveillance

Les enjeux de contrôle d'accès et de surveillance revêtent une importance croissante dans un monde hyperconnecté. L'AE offre des solutions concrètes, comme le filtrage de l'accès à des produits réglementés (ex. : tabac, alcool) ou à des contenus inadaptés aux mineurs [35]. Dans le secteur médical, son intégration à des systèmes intelligents (robots infirmiers, unités de soin) permettrait d'ajuster les protocoles en fonction de l'âge des patients, combinant sécurité et personnalisation .[11]

4.3.4 Interaction humain-ordinateur spécifique à l'âge

Les exigences technologiques diffèrent significativement selon les tranches d'âge. L'estimation automatique de l'âge facilite la création d'interfaces adaptatives, où les paramètres ergonomiques, cognitifs ou visuels sont optimisés en temps réel pour répondre aux besoins spécifiques de chaque utilisateur. Cette approche améliore l'accessibilité et l'efficacité des systèmes interactifs .[11]

4.3.5 Biométrie

L'estimation de l'âge est classée parmi les biométries douces secondaires, fournissant des informations complémentaires sur l'identité des utilisateurs. Ces données permettent d'améliorer

les performances des systèmes biométriques en complétant les caractéristiques principales, comme la reconnaissance faciale, les empreintes digitales ou l'iris.[11]

4.3.6 Personnes disparues

Dans les systèmes d'identification, l'estimation de l'âge couplée à la simulation d'âge peut faciliter la recherche de personnes disparues sur de longues périodes. En analysant les photos actuelles d'un individu et en les comparant à des images antérieures, cette technologie permet de prédire l'évolution de son apparence au fil du temps, optimisant ainsi les chances de reconnaissance .[11]

4.4 Défis courants

Les chercheurs rencontrent beaucoup de difficultés quand ils créent des modèles pour estimer l'âge à partir de photos. Certains de ces problèmes sont "contrôlables", c'est-à-dire qu'on peut les résoudre en modifiant le modèle ou en ajoutant de nouvelles techniques. D'un autre côté, il y a des problèmes "non contrôlables", que les scientifiques ne savent pas encore gérer avec les technologies actuelles.



Figure 1-8 Défis de l'estimation de l'âge.

4.4.1 Pose de la tête et alignement

5L'un des principaux défis contrôlables dans l'estimation de l'âge à partir d'images faciales est la pose de la tête et son alignement. En effet, dans des conditions réelles, les visages peuvent apparaître sous différents angles (de face, de profil, inclinés, etc.), ce qui affecte la précision des modèles. Pour résoudre ce problème, une étape de prétraitement est essentielle ; elle consiste à détecter le visage et à le réaligner pour le placer dans une position standard.



Figure 1-9 Impact de la pose de la tête et de l'alignement sur l'estimation de l'âge

4.4.2 Résolution de l'image

La résolution des images est un facteur clé dans la performance des modèles d'estimation de l'âge. En effet, des images de faible qualité ne permettent pas de capturer des détails importants comme les rides ou la structure du visage, ce qui limite la capacité du modèle à distinguer les différents âges. Cette problématique est souvent liée à la diversité des résolutions des appareils photo ou des caméras utilisés pour capturer les images.



Figure 1-10 Résolution de l'image et précision du modèle

4.4.3 Modifications Faciales

Les modifications faciales, telles que la présence d'accessoires (lunettes, piercings) ou de pilosité (barbe, moustache), peuvent masquer certaines caractéristiques essentielles du visage et altérer la précision des modèles d'estimation de l'âge. Ces variations étant difficilement maîtrisables, il serait pertinent d'explorer des approches de prétraitement visant à normaliser les images, par exemple en supprimant virtuellement ces éléments.

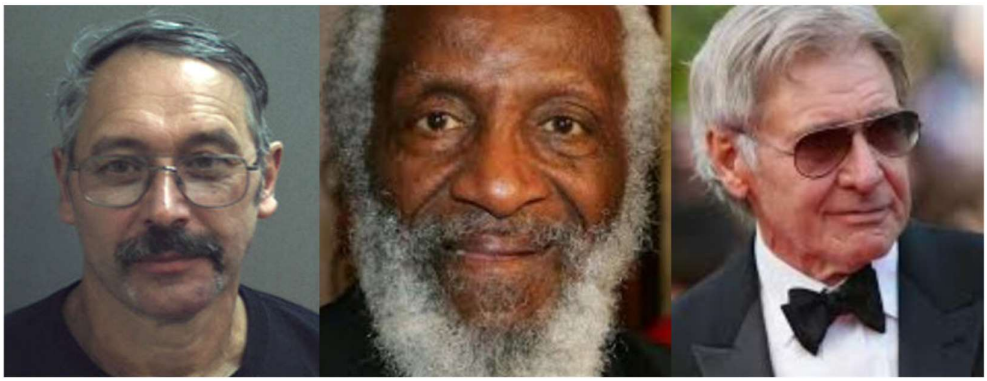


Figure 1-11 Modifications faciales

4.4.4 Génétique

Le vieillissement facial est fortement influencé par des facteurs génétiques, rendant son estimation particulièrement complexe pour les modèles d'apprentissage automatique. L'origine ethnique, le genre et les conditions environnementales dans lesquelles un individu a évolué jouent un rôle déterminant.

4.4.5 Manque de Données

Le surapprentissage représente un défi majeur dans les modèles d'estimation de l'âge, principalement en raison du manque de diversité des ensembles de données d'entraînement. Bien que de vastes bases de données faciales existent, elles sont souvent mal étiquetées ou insuffisamment représentatives en termes de genre, d'âge et d'origine ethnique.

4.4.6 Le mode de vie et l'état de santé

Le mode de vie et l'état de santé d'une personne sont des facteurs non contrôlables qui influencent fortement l'apparence physique et, par conséquent, la performance des modèles d'estimation de l'âge. Par exemple, une personne en bonne santé avec un mode de vie sain peut paraître plus jeune que son âge réel, tandis qu'une personne souffrant de problèmes de santé ou ayant un mode de vie stressant peut paraître plus âgée. Ces variations rendent difficile la tâche des modèles, qui peuvent être trompés par ces différences d'apparence. [12]



Figure 1-12Facteurs influençant le vieillissement facial : génétique, mode de vie et état de santé

5 Conclusion

Dans ce chapitre, nous avons présenté les bases de la reconnaissance faciale et son rôle dans l'estimation de l'âge. Nous avons exploré son fonctionnement, ses méthodes d'analyse et son évolution historique, tout en mettant en évidence les différentes approches biométriques utilisées. Ensuite, nous avons détaillé les étapes essentielles de la reconnaissance faciale et les principales techniques permettant d'extraire et d'analyser les traits du visage. Enfin, nous avons abordé l'estimation de l'âge, ses méthodes, ses applications et les défis qu'elle soulève.

Chapitre 2

État de l'art

Chapitre 2

Chapitre 2 : État de l'art

1 Introduction

L'identification de l'âge à partir du visage est une tâche compliquée, même pour les personnes. Depuis plus de trente ans, la reconnaissance faciale suscite un vif intérêt en raison de son rôle clé dans l'analyse d'image et la reconnaissance de motifs. Cette évolution s'explique par l'essor des besoins commerciaux et juridiques, ainsi que par la disponibilité de technologies avancées telles que les smartphones, les caméras numériques et les processeurs graphiques (GPU).

Même si les algorithmes modernes de l'apprentissage automatique et de la reconnaissane faciale ont atteint un stade avancé de développement, leur performance est fortement dépendante des conditions concrètes d'emploi.Déceler des visages dans des contextes dynami ques et non maîtrisés (variations de lumière, modifications de posture et d'expression, obstruct ions partielles ou déplacements de caméra) demeure toujours une problématique de taille.

Le choix des modèles et algorithmes joue un rôle déterminant dans la fiabilité des résultats et dépend des performances des différentes approches appliquées à la reconnaissance faciale. Une estimation plus précise de l'âge ouvre de nouvelles perspectives dans plusieurs domaines.[13] [14]

2 Définition de l'intelligence artificielle (IA)

L'Intelligence Artificielle est le domaine de l'informatique qui se consacre à la création de machines qui effectuent des tâches qui nécessitent un niveau d'intelligence humaine, telles que l'apprentissage, le raisonnement, la perception, la prise de décisions et la créativité. Selon son pionnier John McCarthy, il l'a défini comme « la science et l'ingénierie d'élaboration de machines intelligentes. L'IA couvre un large éventail de technologies, y compris des algorithmes de recommandation comme les playlists personnalisées sur Spotify.[15]

2.1 Apprentissage automatique (Machine learning)

L'apprentissage automatique est une sous-catégorie de l'intelligence artificielle qui permet aux ordinateurs d'apprendre automatiquement à partir de données, sans être programmés de manière explicite. Il repose sur des techniques statistiques pour améliorer les performances des systèmes face à diverses tâches.

On distingue plusieurs types d'apprentissage selon la nature des données :

- **Apprentissage supervisé (Supervised Learning)**

L'apprentissage supervisé utilise des données étiquetées pour entraîner des modèles à classer ou prédire des résultats. Il ajuste ses paramètres pour optimiser la précision du modèle. Des techniques courantes incluent les réseaux neuronaux, la régression, les SVM, et les arbres décisionnels.

- **Apprentissage non supervisé (Unsupervised Learning)**

L'apprentissage non supervisé analyse des données non étiquetées pour identifier des structures cachées, comme des clusters. Il est utile pour la segmentation de la clientèle, la réduction de dimensionnalité, et la reconnaissance de formes, avec des méthodes comme le clustering k-means et l'analyse en composantes principales (ACP).

- **Apprentissage semi-supervisé (Semi-Supervised Learning)**

L'apprentissage semi-supervisé combine des données étiquetées limitées et des données non étiquetées abondantes. Il est utilisé lorsque l'étiquetage des données est coûteux ou insuffisant pour un apprentissage supervisé complet.

- **Apprentissage par renforcement (Reinforcement Learning)**

L'apprentissage par renforcement est une approche dans laquelle un agent apprend à maximiser une fonction de récompense en interagissant avec son environnement, par essais et erreurs. Contrairement à l'apprentissage supervisé, l'algorithme n'est pas formé sur des données étiquetées, mais apprend progressivement en recevant des récompenses ou des punitions en fonction des actions entreprises.[16]

2.2 Apprentissage Profond (Deep learning)

Le Deep Learning, une branche du Machine Learning, est une technique d'apprentissage automatique qui s'inspire du fonctionnement du système nerveux des êtres vivants. Les algorithmes de Deep Learning traitent l'information de manière comparable aux réseaux de neurones biologiques, qui réagissent aux signaux nerveux. En fonction de la nature et de la fréquence des signaux reçus, certains réseaux de neurones peuvent se renforcer, tandis que d'autres peuvent se désactive. [17]

2.3 Réseaux de neurones convolutifs (CNN)

Un réseau de neurones convolutifs (CNN), également connu sous le nom de ConvNet**,** est un type spécialisé d'algorithme d'apprentissage profond spécialement conçu pour traiter des données organisées sous forme de grille, telles que les images ou les vidéos. Cette architecture est particulièrement efficace pour extraire des caractéristiques pertinentes à partir de données visuelles, éliminant ainsi le besoin de l'ingénierie manuelle des caractéristiques. Les CNN sont utilisés dans des tâches nécessitant la reconnaissance d'objets, telles que la classification d'images, la détection et la segmentation. Ils sont devenus un pilier des applications en vision par ordinateur, permettant à des systèmes d'exécuter des tâches qui nécessitaient autrefois une perception humaine, comme la reconnaissance faciale, l'analyse d'images médicales et l'analyse de vidéos.[18][19]

Un CNN se compose généralement de plusieurs couches, chacune jouant un rôle spécifique dans le traitement de l'information :

1. **Couche de convolution (ConV)** : applique des filtres aux données d'entrée afin de générer des cartes d'activation contenant les caractéristiques les plus significatives
2. **Couche d'activation ReLU (Rectified Linear Unit)** : introduit de la non-linéarité en remplaçant les valeurs négatives par zéro, ce qui améliore la capacité d'apprentissage du modèle
3. **Couche de pooling (POOL)** : réduit la dimensionnalité des cartes d'activation tout en conservant l'essentiel de l'information, ce qui diminue la complexité du modèle
4. **Couche entièrement connectée (Fully Connected – FC)** : relie tous les neurones entre eux et génère un vecteur de sortie représentant les probabilités des classes à prédire.[20]

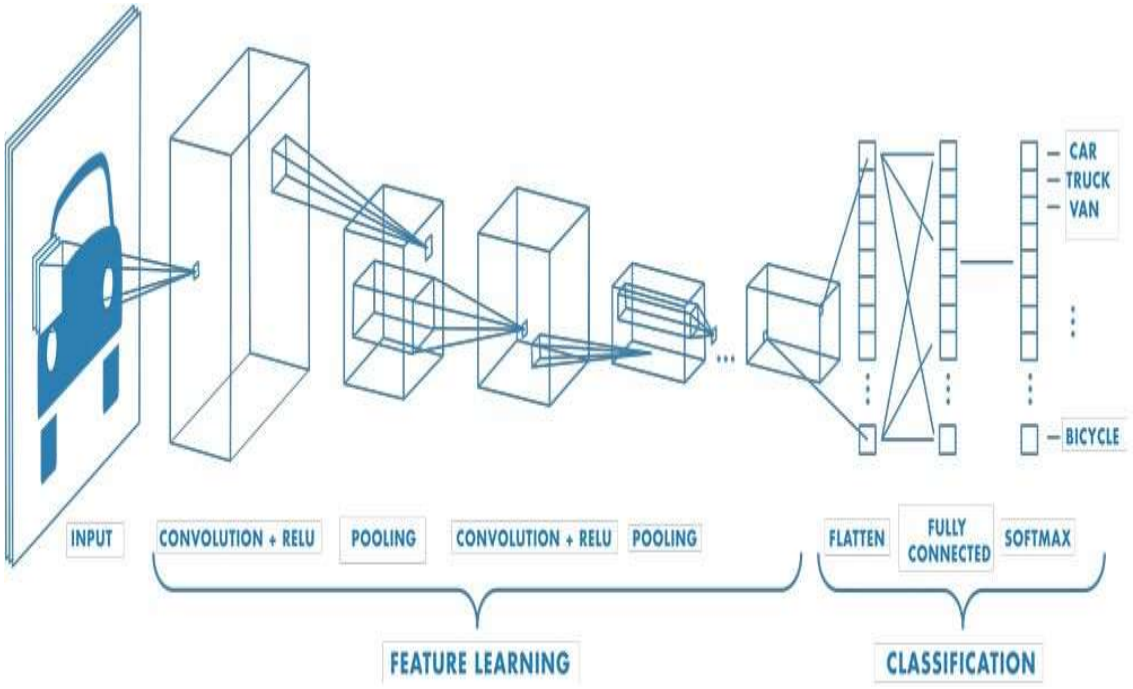


Figure 2-1 Architecture de réseau de neurones convoulutifs

2.3.1 Apprentissage par transfert (Transfer Learning)

L'apprentissage par transfert est une méthode d'apprentissage consistant à exploiter les connaissances acquises lors de l'entraînement d'un modèle sur une tâche donnée pour en améliorer la performance sur une autre tâche liée, souvent avec moins de données. L'idée est de transférer les représentations apprises sur un grand jeu de données (comme ImageNet) vers un nouveau contexte, en adaptant uniquement certaines couches du réseau.[21]

2.3.2 Les Principaux Modèles CNN Pré-entraînés en Apprentissage Profond

Plusieurs architectures de réseaux de neurones convolutifs (CNN) ont été développées pour résoudre efficacement des problèmes complexes de vision par ordinateur. Ces modèles, souvent pré-entraînés sur de larges bases de données de référence telles qu'**ImageNet**, sont particulièrement utiles dans des contextes où les ensembles de données disponibles sont de taille réduite. Cette approche, connue sous le nom d'apprentissage par transfert, permet de réutiliser les connaissances acquises par ces modèles sur des tâches génériques pour les adapter à des tâches spécifiques, comme la classification de l'âge et du genre à partir d'images faciales.

- ***MobileNetV2 :*** Est une version optimisée de MobileNetV1, utilisant une architecture résiduelle inversée pour réduire le nombre de paramètres et améliorer l'efficacité, avec 53 couches convolutives. Cette approche permet de réduire les paramètres de 30 % et les opérations de 2 % par rapport à MobileNetV1, rendant le modèle plus léger et adapté aux environnements à ressources limitées. [22]
- ***ResNet :*** Introduit en 2015, est une architecture de réseau de neurones qui intègre l'apprentissage résiduel pour résoudre le problème de la disparition du gradient, couramment rencontré dans les architectures comme VGG-Net. Cette approche permet de créer des réseaux plus profonds sans compromettre les performances, en réutilisant les caractéristiques déjà apprises à chaque étape. En utilisant des connexions résiduelles, ResNet peut apprendre de manière plus efficace et réduire le temps d'entraînement en évitant les répétitions inutiles. Cependant, l'ajout de nombreuses couches peut rallonger le temps d'apprentissage, parfois jusqu'à plusieurs semaines.[23]
- ***VGG-Net :*** Introduit en 2014, améliore la précision d'entraînement en augmentant la profondeur du réseau et en utilisant des noyaux plus petits. Cette architecture est largement utilisée pour des tâches telles que l'estimation de l'âge facial grâce à sa capacité à capturer des caractéristiques complexes. VGG-Net est souvent préféré en raison de la disponibilité de ses réseaux pré-entraînés, facilitant ainsi son utilisation dans divers domaines de la vision par ordinateur.[23]
- ***AlexNet :*** Proposé en 2012 par Alex Krizhevsky, est l'un des premiers modèles CNN à exploiter les GPU pour accélérer l'entraînement. Il intègre ReLU, dropout et le pooling avec chevauchement pour améliorer la vitesse et réduire la perte de caractéristiques. Ce modèle a été utilisé pour la classification et l'estimation de l'âge facial, mais nécessite encore une bonne structuration pour atteindre des performances optimales.[23]

3 Ensemble de données utilisés pour l'estimation de l'âge

3.1 FG-NET (Face and Gesture Recognition Network)

L'ensemble de données FG-NET est un jeu d'images faciales publiques dédié à la recherche sur le vieillissement du visage. Elle contient 1 002 images (en couleur et en niveaux de gris) de 82 individus, avec des âges variant de 0 à 69 ans, majoritairement concentrés entre

0 et 40 ans. Les images proviennent de photographies personnelles numérisées, ce qui introduit une forte variabilité en termes de résolution, d'éclairage, d'expression faciale et de point de vue. Certaines images présentent également des occlusions partielles. Le jeu est enrichi par des fichiers d'annotations indiquant la position de 68 points de repère faciaux ainsi que l'âge exact du sujet dans chaque image. Cette base est largement utilisée pour l'évaluation des modèles de régression d'âge.[20][24]

3.2 MORPH

L'ensemble de données MORPH, développée par Ricanek et Tesafaye (2006) au sein du laboratoire Face Aging Group de l'Université de Caroline du Nord à Wilmington, est l'une des bases publiques les plus utilisées pour l'étude du vieillissement facial. Elle fournit, en plus des images, des **attributs annotés** tels que l'âge, le genre, l'origine ethnique, le poids, la taille et l'ascendance, elle est divisée en deux sous-ensembles :

- **Album I** : comprend 1 724 images de 515 individus, âgés de 27 à 68 ans, dont 1 430 images d'hommes et 294 de femmes. Les écarts d'âge entre les prises de vue d'un même individu varient de 46 jours à 29 ans.
- **Album II** : plus vaste, contient 55 134 images de plus de 13 000 personnes, offrant ainsi un échantillon riche et diversifié pour les modèles de régression et de classification de l'âge.[20]

3.3 IMDB-WIKI

L'ensemble de données IMDB-WIKI est l'une des plus grandes bases publiques utilisées pour l'estimation de l'âge à partir d'images de visages. Elle contient environ 524 230 images de personnes âgées de 0 à 100 ans, collectées automatiquement depuis les sites IMDB et Wikipedia. Elle regroupe des photos de célébrités, mais comprend aussi de nombreuses images de mauvaise qualité, comme des dessins, des visages masqués, des images avec plusieurs personnes ou encore des photos floues ou incomplètes. Ces imperfections nécessitent un bon prétraitement avant l'entraînement des modèles.[20]

3.4 AFAD

L'ensemble de données AFAD (Asian Face Age Database) est une grande base publique créée en 2016. Elle est utilisée pour tester les performances des modèles d'estimation de l'âge. Elle contient 164 432 images de visages de personnes d'origine asiatique, avec les informations d'âge et de genre. Parmi ces images, on trouve 63 680 visages de femmes et 100 752 visages d'hommes, avec des âges compris entre 15 et 40 ans. Les photos ont été collectées à partir du réseau social asiatique RenRen, qui est surtout utilisé par les étudiants.[20]

3.5 OIU-Adience

L'ensemble de données OIU-Adience contient environ 26 580 images de visages appartenant à 2 284 individus, réparties en huit tranches d'âge : 0–2, 4–6, 8–13, 15–20, 25–32, 38–43, 48–53 et 60+. Les images ont été prises à l'aide de smartphones puis publiées sur le site Flickr, sans filtrage ni traitement préalable. Elles présentent donc une grande diversité en termes de bruit, de pose, de qualité d'image et d'apparence générale. Cette diversité reflète fidèlement les difficultés rencontrées dans les environnements réels et non contrôlés. C'est pourquoi cette base est souvent utilisée dans les travaux de classification de l'âge et du genre.[24]

3.6 CACD

La base de données CACD (Cross-Age Celebrity Dataset) est une base publique contient plus de 160 000 images de visages de 2 000 célébrités, avec des âges allant de 16 à 62 ans. Elle est principalement utilisée pour l'étude de la variation des visages à travers différentes tranches d'âge chez une même personne**.**[24]

4 Travaux connexes

L'estimation de l'âge à partir d'images faciales est un domaine de recherche actif qui a évolué avec le temps, passant des méthodes traditionnelles d 'apprentissage automatique aux approches récentes basées sur l'apprentissage profond. Plusieurs études ont exploré différentes techniques pour améliorer la précision et la robustesse des modèles.

4.1 Approches basées sur l'apprentissage automatique

Dans cette étude, Steel .[25] présente une méthode pour la détection des mineurs à partir d'images faciales en utilisant les ratios anthropométriques du visage. Au lieu d'extraire des caractéristiques complexes comme dans le Deep Learning, cette approche analyse les proportions entre différentes parties du visage (yeux, nez, bouche) et les utilise comme des indicateurs du vieillissement. Un classificateur Naïve Bayes est ensuite appliqué pour classer les individus en mineurs ou adultes.

L'approche a été testée sur la base de données FG-NET, où elle a atteint une précision d'environ 86% pour la détection des mineurs. Bien que cette méthode soit rapide et interprétable, elle est limitée lorsqu'il s'agit de prédire un âge précis.

Wang et al. [26] ont proposé raSVM+ (Relative Attribute SVM+), une amélioration des SVM classiques destinée à l'estimation de l'âge, intégrant des attributs relatifs (rides, morphologie faciale, poches sous les yeux) comme informations privilégiées pour mieux capturer les signes de vieillissement. Testé sur les ensemble de données FG-NET et MORPH Album2, ce modèle réduit les erreurs entre classes d'âges proches et atteint une erreur absolue moyenne (MAE) de 4,07 ans sur FG-NET.

Priyadarshni et al**.**[27] ont présenté un système de classification de l'âge basé sur le classificateur K-Nearest Neighbors (K-NN). Leur approche se compose de trois étapes : extraction des caractéristiques faciales et classification. L'extraction des caractéristiques repose sur la Local Gabor Binary Pattern Histogram (LGBPH) pour la texture de la peau et une analyse des rides basée sur la densité des rides. Le classificateur K-NN est ensuite utilisé pour catégoriser les visages en trois groupes : enfant, adulte et âgé. Testé sur un ensemble de données privée de 270 images, le modèle a atteint une précision globale de 96,6%, avec des taux de reconnaissance de 93,3% pour les enfants, 96,6% pour les adultes et 100% pour les personnes âgées.

Madhavi et al. [28]ont présenté une approche d'estimation de l'âge basée sur les Machines à Vecteurs de Support (SVM) combinées à l'Analyse Discriminante Linéaire (LDA). Leur méthode suit trois étapes principales : prétraitement des images**,** extraction des caractéristiques et classification . Après le prétraitement, les caractéristiques faciales sont ensuite extraites et

utilisées dans un modèle SVM avec LDA, qui a démontré une meilleure précision que d'autres approches comme PCA et LBP. Testé sur l'ensemble de données FG-NET, le modèle a atteint une précision de 84%.

4.2 Approches basées sur l'apprentissage profond

Noor et al. [29]ont développé une méthode innovante pour prédire l'âge et classifier le genre à partir d'images faciales. Leur approche repose sur l'extraction de caractéristiques profondes via VGG-19, suivie d'un CNN personnalisé pour la classification. Entraîné sur l'ensemble de données UTKFace, le modèle inclut des étapes de prétraitement, d'extraction de caractéristiques et de régularisation pour éviter le surajustement. Il atteint une précision de 78,67 % pour l'âge et 97,02% pour le genre**.**

Z. Lyu [13] a proposé une méthode pour estimer l'âge à partir de photos de visages en combinant trois modèles : EfficientNetB0, DenseNet121 et Inception V3. Entraîné sur l'ensemble de données UTKFace, ce modèle fusionne les caractéristiques extraites par chaque réseau pour améliorer la précision. Il atteint une précision de 65 %, ce qui est mieux que les modèles pris individuellement. L'ajout d'une couche Transformer a été testé mais n'a pas amélioré les résultats. Cette étude montre que combiner plusieurs modèles peut rendre la reconnaissance de l'âge plus efficace.

Nguyen et al. [30] ont exploré la prédiction de l'âge et du genre à partir d'images faciales en utilisant trois modèles de deep learning : Unet, MobileNets et EfficientNets. Leurs expériences, menées sur l'ensemble de données UTKFace**,** ont montré que Unet était le plus performant pour la classification du genre avec une précision de 97.22 %, tandis que MobileNets a fourni la meilleure estimation d'âge avec une erreur moyenne absolue (MAE) de 2.248. L'étude met en avant l'importance du choix du modèle et des hyperparamètres pour améliorer la précision des prédictions

Sathyavathi et al. [31] ont proposé un modèle basé sur un réseau neuronal convolutif profond (DCNN) combiné à l'algorithme Cuckoo Search (CS) pour améliorer la précision de la prédiction de l'âge à partir d'images faciales. Le modèle est entraîné sur l'ensemble de données UTKFace, et différentes architectures ont été testées : CNN, DNN, LSTM et leur modèle SLSTM-DNN, qui atteint la meilleure précision avec 91.45%. Leur approche optimise le temps

d'exécution et l'utilisation des ressources en appliquant des techniques de prétraitement, d'extraction de caractéristiques (PCA) et de détection des contours (Gabor Filter).

Rahman et al.[22] ont proposé un modèle de prédiction d'âge basé sur un réseau neuronal convolutif (CNN) avec des modules d'attention (CBAM, SENet, Self-Attention). Leur objectif était d'améliorer la précision en se concentrant sur les parties les plus importantes du visage. Leur modèle a été entraîné sur un dataset fusionné (UTKFace + Facial-Age) et testé sur FG-NET et CACD. Grâce à l'apprentissage par transfert et aux modules d'attention, ils ont réduit l'erreur moyenne absolue (MAE) à 3.5.

Bekhouche et al.[32] ont proposé une nouvelle approche pour l'estimation de l'âge basée sur un réseau de neurones profonds multi-étapes (MSDNN). Leur modèle exploite différentes couches de EfficientNet et MobileNetV3 pour extraire des caractéristiques plus détaillées et améliorer la précision des prédictions. Ils ont entraîné et testé leur modèle sur trois ensemble de données bien connues : MORPH2, qui contient 55 134 images d'individus âgés de 16 à 70 ans, CACD, qui comprend 159 449 images de célébrités âgées de 14 à 62 ans, et AFAD, composé de 165 501 images de jeunes adultes asiatiques âgés de 15 à 40 ans. Les résultats montrent une erreur moyenne absolue (MAE) de 2.59 ans sur MORPH2**,** 3.25 ans sur AFAD et 4.90 ans sur CACD.

George et al.[33] ont proposé une approche basée sur un réseau neuronal convolutif (CNN) personnalisé pour l'estimation de l'âge à partir d'images faciales. Leur modèle est conçu pour traiter l'âge de deux manières : en régression, où il prédit un âge exact, et en classification, où il classe les individus en trois groupes : enfants, adolescents et adultes.

Leur modèle a été entraîné et testé sur UTKFace, en utilisant 11 700 images annotées avec l'âge, le genre et l'ethnicité. Les résultats montrent que le modèle de régression atteint une erreur moyenne absolue (MAE) de 2.15 ans, tandis que le modèle de classification atteint une précision de 93%.

Kumar et al. [34] ont proposé un modèle CNN optimisé pour la classification de l'âge et du genre à partir d'images faciales réelles et non contraintes. Le modèle est pré-entraîné sur IMDb-WIKI, puis affiné avec OIU-Adience, un ensemble de données contenant 26 580 images annotées avec des tranches d'âge et le genre. L'architecture du CNN comprend plusieurs couches de convolution, max-pooling, normalisation et dropout, avec une fonction softmax

pour la classification. Les résultats montrent que le modèle atteint une précision de 84.8% pour la classification de l'âge et une erreur moyenne absolue (MAE) de 2.26 ans.

Garain et al.[35] ont proposé GRA_Net, un modèle d'apprentissage profond basé sur un réseau d'attention résiduel avec mécanisme de gating pour améliorer l'estimation de l'âge et la classification du genre à partir d'images faciales. Contrairement aux méthodes traditionnelles, leur approche combine classification et régression pour l'âge, permettant une meilleure précision. Leur modèle a été entraîné et testé sur cinq ensembles de données publiques : FG-NET, Wikipedia, AFAD, UTKFace et AdienceDB, couvrant une large diversité de visages et d'âges. Les résultats montrent une erreur moyenne absolue (MAE) de 3.35 ans pour l'estimation de l'âge et une précision de 99.2% pour la classification du genre sur UTKFace et 82% sur AdienceDB.

Kim et al.[36] ont développé un modèle basé sur les GANs (réseaux antagonistes génératifs), appelé GRA-GAN, pour améliorer la reconnaissance faciale en transformant les caractéristiques du visage liées à l'âge, au genre et à la race. L'objectif est de corriger le déséquilibre des bases de données et de générer des visages réalistes dans différentes catégories. Le modèle utilise une fusion d'informations basée sur les canaux pour intégrer les caractéristiques extraites par l'encodeur et le décodeur, préservant ainsi les détails importants des images entraîné sur plusieurs ensembles de données (MORPH, AFAD, AAF et UTKFace), il montre une amélioration significative par rapport aux méthodes traditionnelles, les résultats indiquent un MAE (erreur moyenne absolue) de 3.77 à 6.12 ans pour l'estimation de l'âge.

Dey et al. [37]ont proposé un modèle CNN personnalisé pour prédire l'âge et le genre à partir d'images faciales. Leur approche repose sur un prétraitement avancé des images, incluant la normalisation, l'augmentation des données et l'alignement des visages pour améliorer la précision du modèle.

Le modèle a été entraîné et testé sur les ensembles de données Adience et UTKFace, qui contiennent des images variées en termes d'expressions, de poses et d'éclairages. L'architecture du CNN comprend des couches de convolution, de pooling et des couches entièrement connectées pour extraire et classifier les caractéristiques faciales, les résultats montrent une précision de 86.42% pour l'âge sur Adience et 81.96% sur UTKFace.

Jain et al.[38] ont conçu un modèle CNN personnalisé pour la prédiction de l'âge et du genre à partir d'images faciales. Leur approche repose sur un prétraitement avancé des images, comprenant la normalisation, l'augmentation des données et l'extraction automatique des caractéristiques faciales,le modèle a été entraîné et testé sur les ensembles de données Google et IMFDB, contenant un large éventail d'images représentant différentes tranches d'âge et divers groupes ethniques. L'architecture CNN utilisée comporte plusieurs couches convolutionnelles et entièrement connectées, optimisées pour extraire des informations discriminantes du visage, les résultats montrent une précision de 80% pour la classification de l'âge et du genre.

5 Tableau récapitulatif des méthodes existantes

Publication	**Architecture**	**Taille de l'image**	**Nombre de classe**	**Dataset**	**Performance**	
					Age	**Genre**
Steel, 2013	Naïve Bayes basé sur les ratios anthropométriques	/	2	FG-NET	Accu : 82% Précision : 89% Rappel : 87%	N/A
Wang et al., 2016	raSVM+ (Relative Attribute SVM+)	/	70 Âge réel	FG-NET	4.07(MAE)	N/A
			61 Âge réel	MORPH Album 2 **(5 395 images)**	5.05(MAE)	
Priyadarshni et al. 2019	K-NN avec LGBPH et analyse des rides	/	3	Base privée (270 images)	96,6%(accu)	N/A
Jain et al., 2020 [38]	CNN "from scratch"	70 x 70	55 (5-60)	Google et IMFDB	80%(accu)	80%(accu)
Madhavi et al.,2021	SVM avec LDA	/	7	FG-NET	84%(accu)	N/A
Garain et al., 2021 [35]	GRA_Net		70 Âge réel	FG-NET	3.35(MAE)	/
			25 Âge réel	AFAD	3.10(MAE)	/
			/	Wikipedia	5.45(MAE)	/
			116 Âge réel	UTKFace	1.07(MAE)	99.2% (Accu)
			8 Âge réel	AdienceDB	10.57 (MAE)	82% (Accu)
	GRA-GAN	256x256	47	MORPH	3.77 - 6.12	96.31–98.44%

Kim et al., 2022 [36]			Âge réel	AFAD AAF UTKFace	(MAE)	(Accu)
Sathyavathi et al.2023 [31]	DCNN + Algorithme Cuckoo Search (DCNN-CS)		6	UTKFace	Accu :91.45% Précision :85.13% Rappel : 78.12% F1-score :83.19%	N/A
Rahman et al., 2023 [22]	CNN from scratch avec Self-Attention	200x200	70 Âge réel	FG-NET	2.54(MAE)	N/A
			116 Âge réel	UTKFace	3.85(MAE)	
			54 Âge réel	CACD	4.24(MAE)	
			116 Âge réel	Dataset combiné (FG-NET, UTKFace)	3.5(MAE)	
Noor et al. 2024 [29]	VGG19+ CNN "from scratch"	224 x224	8	UTKFace	Accu: 78.67% Précision :83.08% Rappel : 74.1% F1-score : 78.3% AUC : 98.23%	97.02%(Accu)
Zixuan Lyu, 2024 [13]	Fusion(EfficientNetB0, DenseNet121 Inception V3)	200 x200	7	UTKFace	Accu : 65% précision : 69% Rappel : 67% F1-score : 68%	N/A
Nguyen et al. 2024 [30]	Unet	224 x224	116 Âge réel	UTKFace	4.722 (MAE)	97.02%(Accu)
	MobileNets				2.248 (MAE)	87.45%(Accu)
	EfficientNets				2.306 (MAE)	96.52%(Accu)
Bekhouche et al., 2024[32]	Multi-Stage Deep Neural Network (MSDNN) basé sur EfficientNet et MobileNetV3	224x224	54 Âge réel	MORPH2	2.59(MAE)	N/A
			25 Âge réel	AFAD	3.25(MAE)	
			48 Âge réel	CACD	4.90(MAE)	
George et al., 2024[33]	CNN "from scratch" (Deep CNN)	224x224	3	UTKFace (11 700 images)	93%(Accu)	
			116 Âge réel		2.15(MAE)	

Kumar et al., 2024 [34]	CNN "from scratch" (pré-entraîné sur IMDb-WIKI)	224x224 256x256	8	OIU-Adience IMDb-WIKI	84.8%(Accu) 2.26(MAE)	
Dey et al., 2024 [37]	CNN "from scratch"	/	8	Adience,	Accu : 86.42% Précision : 88.16% Rappel : 86.38% F1-score : 87.26%	Accu : 97.65% Précision: 97.54% Rappel : 97.5% F1-score : 97.52%
				UTKFace	Accu : 81.96% Précision : 82.27% Rappel : 81.75% F1-score : 82.01%	Accu : 96.32% Précision : 97.5% Rappel : 97.5% F1-score : 97.5%

Tableau 2-1 Tableau récapitulatif des travaux connexes

6 Synthèse

L'estimation de l'âge à partir d'images faciales repose sur différentes méthodes qui ont évolué au fil du temps. Les premières approches étaient basées sur des modèles classiques de Machine Learning (ML), tandis que les méthodes récentes exploitent le Deep Learning (DL) pour améliorer la précision et la robustesse des prédictions.

Les méthodes traditionnelles utilisaient des algorithmes tels que SVM (Support Vector Machines), KNN (K-Nearest Neighbors), Random Forest (RF) et Naïve Bayes (NB) pour classer les âges en groupes prédéfinis ou pour prédire une valeur d'âge à partir de caractéristiques extraites manuellement du visage (rides, contours, texture de la peau). Steel (2013) a proposé une approche utilisant Naïve Bayes basée sur les ratios anthropométriques du visage, atteignant 86 % de précision sur l'ensemble de données FG-NET. De même, Wang et al. (2016) a utilisé une méthode de régression basée sur SVM (raSVM+), obtenant une MAE de 4.07 ans sur MORPH Album 2. Cependant, ces méthodes présentent des limites importantes, car elles dépendent fortement de la qualité des caractéristiques sélectionnées et généralisent mal sur des images variées.

Avec l'essor du Deep Learning, les réseaux de neurones convolutifs (CNN) sont devenus la méthode dominante. Contrairement aux méthodes classiques, les CNN apprennent automatiquement les caractéristiques les plus pertinentes à partir d'un grand volume d'images. Des architectures populaires comme VGG-19, EfficientNet, MobileNet et ResNet ont été largement utilisées pour améliorer la précision de l'estimation de l'âge. Noor et al. (2024) a

proposé une approche combinant VGG19 et un CNN personnalisé, atteignant une précision de 78.67 % sur UTKFace. Bekhouche et al. (2024) a développé un Multi-Stage Deep Neural Network (MSDNN) basé sur EfficientNet et MobileNetV3, atteignant une MAE comprise entre 2.59 et 4.90 ans sur MORPH2, AFAD et CACD. Ces modèles sont capables de traiter des variations complexes du visage, mais ils nécessitent de grandes bases de données et une forte puissance de calcul.

Pour améliorer encore la précision, les chercheurs ont introduit des techniques avancées comme les mécanismes d'attention (CBAM, SENet, Self-Attention), qui permettent au modèle de se concentrer sur les zones les plus informatives du visage. Kim et al. (2022) a expérimenté une approche basée sur GRA-GAN (Generative Adversarial Network) pour enrichir les bases de données et améliorer la robustesse des modèles, obtenant une MAE entre 3.77 et 6.12 ans sur MORPH, AFAD et UTKFace.

En parallèle, certaines recherches ont exploré des approches hybrides, combinant l'extraction automatique des caractéristiques avec un classificateur ML. Par exemple, Garain et al. (2021) a introduit GRA_Net, qui intègre un CNN avec des mécanismes d'attention et des techniques de Machine Learning, obtenant une MAE de 3.35 ans sur FG-NET, AFAD, Wikipedia, UTKFace et AdienceDB. Nguyen et al. (2024) a proposé une approche combinant Unet, MobileNets et EfficientNets, atteignant une MAE de 2.248 ans sur UTKFace. Ces modèles hybrides permettent d'améliorer la précision tout en optimisant les ressources computationnelles.

En conclusion, les méthodes d'estimation de l'âge sont passées d'une approche basée sur l'extraction manuelle des caractéristiques à des modèles entièrement automatisés grâce au Deep Learning. Les recherches actuelles se concentrent sur l'amélioration de la précision, la réduction des biais et l'optimisation des performances des modèles pour une utilisation en conditions réelles. Les approches hybrides, qui tirent parti des avantages du ML et du DL, apparaissent comme une alternative prometteuse, combinant robustesse, précision et efficacité computationnelle.

7 Conclusion

Dans ce chapitre, nous avons présenté un état de l'art sur les approches d'estimation de l'âge et du genre à partir d'images faciales. Nous avons introduit les concepts de l'intelligence artificielle, du machine learning et du deep learning, en mettant l'accent sur les réseaux de neurones convolutifs (CNN). Plusieurs architectures pré-entraînées, telles que VGG19, ResNet et EfficientNet, ont été décrites. Nous avons également analysé les principales ensembles de données utilisées, comme UTKFace, FG-NET, MORPH et IMDB-WIKI, en soulignant leurs spécificités et limites. Ensuite, nous avons passé en revue des travaux antérieurs basés sur des approches classiques et profondes. Enfin, une synthèse comparative des méthodes a permis d'identifier les plus performantes, posant ainsi les fondations de notre propre méthode présentée dans le chapitre suivant.

Chapitre 3 : Méthode proposée pour la classification de l'âge et du genre à partir d'images faciales

Chapitre 3

Chapitre3 : Méthode proposée pour la classification de l'âge et du genre à partir d'images faciales

1 Introduction

Dans ce chapitre, nous présentons la démarche méthodologique adoptée pour traiter la problématique de la classification conjointe de l'âge et du genre à partir d'images faciales. Nous nous appuyons sur une base de données publique contenant une grande diversité de visages, présentant des défis liés notamment au déséquilibre des classes et à la variabilité des traits. Nous exploitons les avantages du Transfer Learning à travers l'utilisation de modèles pré-entraînés tels que VGG16 et DenseNet121, intégrés dans une architecture de fusion conçue pour améliorer la précision des prédictions. Nous abordons également l'organisation des données, les conditions expérimentales, ainsi que l'évaluation comparative des performances entre les modèles individuels et le modèle hybride, dans le but d'analyser les apports concrets de notre approche.

2 Le corpus *UTKFace*

2.1 Description d'ensemble de données UTKFace

L'ensemble de données UTKFace est une base publique largement utilisée pour les tâches d'estimation de l'âge et de classification du genre à partir d'images faciales. Il contient plus de 23 000 images de visages humains, couvrant une large plage d'âges (de 0 à 116 ans), de genres et d'origines ethniques, les images présentent une variété de poses, d'expressions faciales, et de conditions d'éclairage. Chaque fichier image est nommé selon la convention suivante : [age]_[gender]_[ethnicity]_[timestamp].jpg

- **Age** : valeur numérique entre 0 et 116.
- **Genre** : 0 pour homme, 1 pour femme.
- **Origines ethniques** : Blanc= 0, Noir =1 , Asiatique =2, Indien = 3, Autres =4.

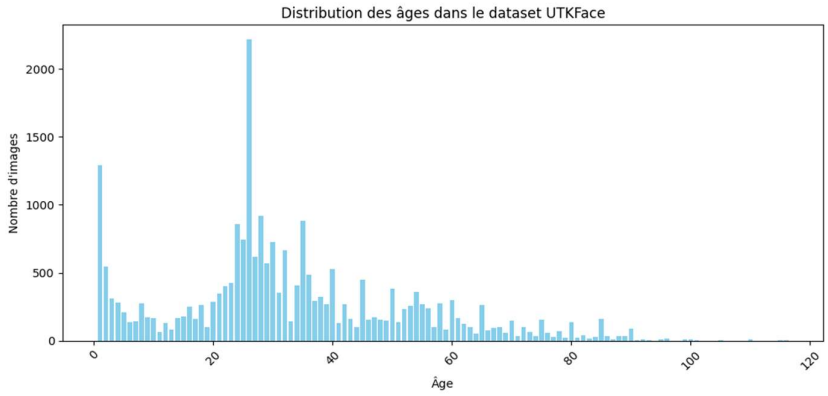


Figure 3-1 Distribution des âges dans l' ensemble de données UTKFace

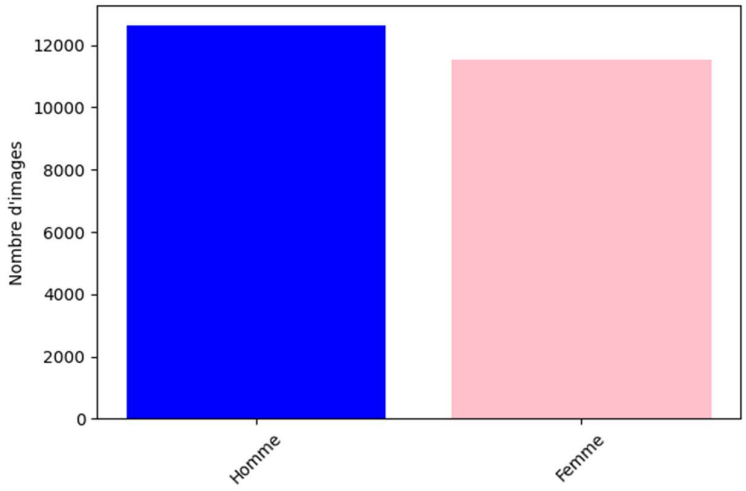


Figure 3-2 Distribution du genre dans l'ensemble de données UTKface

Dans ce travail, nous avons procédé à une classification en huit groupes d'âge (0–2), (4–6), (8–13), (15–20), (25–32), (38–43), (48–53) et (60+), ainsi qu'une classification binaire du genre (homme / femme).

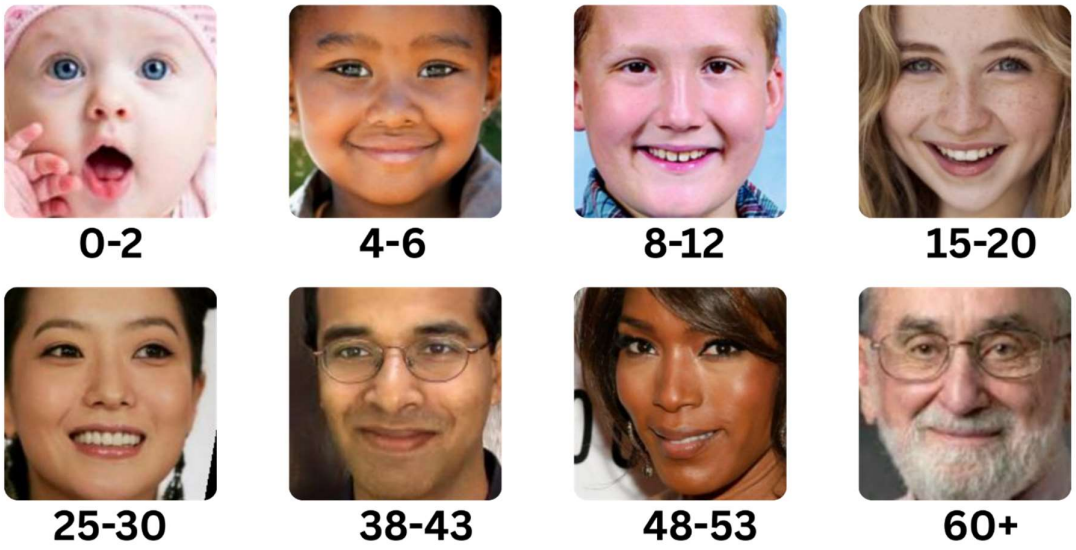


Figure 3-3 Exemples de visages issus de différentes tranches d'âge dans la base UTKFace

2.2 Protocoles de test utilisés

Afin d'évaluer les performances des modèles proposés pour la classification de l'âge et du genre à partir d'images faciales, deux protocoles de test ont été adoptés en utilisant l' ensemble de données UTKFace :

1. **Premier protocole** : Les données sont divisées en 80 % pour l'entraînement**,** dont 20 % sont réservées à la validation**,** et 20 % pour le test.
    - **10633** des images ont été utilisées pour l'**entraînement** (*training*).
    - **2658** des images ont été utilisées pour la **validation**.
    - **3323** des images ont été utilisées pour le **test** (*testing*).

Ce protocole a été utilisé pour l'évaluation des modèles dans la classification de **l'âge uniquement.**

2. **Deuxième protocole** : Une division plus simple a été utilisée avec **80 %** des données pour l'entraînement **et** 20 % pour le test.
    - ✓ **13291** des images ont été utilisées pour l'**entraînement** (*training*).
    - ✓ **3323** des images ont été utilisées pour le **test** (*testing*).

Ce protocole a été utilisé pour l'évaluation conjointe de la classification de **l'âge et du genre.**

Groupe d'âge	Entraînement	Test
0–2	1284	321
4–6	480	120
8–12	624	156
15–20	981	245
25–32	5417	1355
38–43	1337	334
48–53	1034	259
60+	2134	533
Total	**13291**	**3323**

Tableau 3-1 Distribution des classes d'âges

Genre	Train	Test
Homme	6804	1739
Femme	6487	1584
Total	**13291**	**3323**

Tableau 3-2 Distribution de genre

2.3 Les défis de UTKface

2.3.1 Données non équilibrées (déséquilibre de classes)

La classe **25–32 ans** contient à elle seule **6 772 images**, ce qui est nettement plus que les autres classes. En comparaison, les classes comme **4–6**, **8–12** ou **15–20** ans sont moins représentées, avec un nombre d'images beaucoup plus faible. Ce déséquilibre peut avoir un impact négatif sur l'apprentissage du modèle, car il aura tendance à mieux reconnaître les classes majoritaires et à négliger les classes minoritaires. Cela peut réduire la précision globale et créer un biais dans les prédictions.

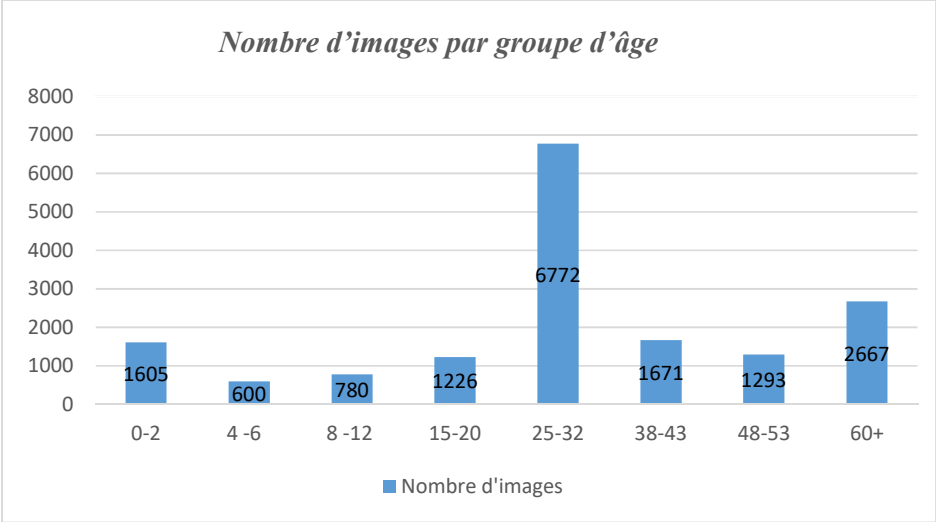


Figure 3-4 Répartition des images selon les groupes d'âge dans UTKFace

2.3.2 Qualité d'image variable

Certaines images dans UTKFace sont de faible qualité (floues, mal exposées ou de faible résolution), ce qui peut entraîner des difficultés pour les modèles d'extraire des caractéristiques utiles.



Figure3-5 Exemple des images floues, mal exposées ou de faible résolution

2.3.3 Mauvaise annotation

- ***Annotations incorrectes de l'âge***

L'erreur d'annotation de l'âge constitue une problématique significative dans les bases de données telles que UTKFace. En effet, l'estimation de l'âge à partir de l'apparence physique conduit fréquemment à des étiquetages erronés, ne reflétant pas

fidèlement l'âge réel de l'individu représenté sur la photographie. Cette disparité peut être attribuée à divers facteurs, notamment les caractéristiques génétiques, le mode de vie, ou encore la qualité de l'image.

En 2025, l'âge réel de ce joueur est de 38 ans, pourtant il a été annoté comme ayant 39 ans en 2017.



Figure 3-6 Annotation erronée de l'âge : 39 ans en 2017, âge réel 37 ans en 2025

- ***Annotations incorrectes du genre***

Dans le cadre de l'exploitation du jeu de données UTKFace, il est important de noter que certaines annotations peuvent présenter des inexactitudes, en particulier celles liées au genre (homme/femme). Ces erreurs d'annotation se produisent lorsque l'apparence de la personne dans l'image ne correspond pas à l'étiquette de genre qui lui a été attribuée.



Figure 3-7 Annotation incorrectes du genre

3 Prétraitement et organisation des données

Dans le cadre de cette étude, une phase cruciale de préparation des données a été mise en place pour optimiser la qualité des entrées fournies aux modèles de classification. En premier lieu, un processus de prétraitement automatique des images extraites de la base UTKFace a été mis en œuvre. Cela impliquait l'utilisation de l'algorithme Haar Cascade d'OpenCV pour détecter les visages. Chaque visage détecté a été extrait, recadré et redimensionné à une résolution uniforme de 224 × 224 pixels afin de garantir des dimensions d'entrée homogènes pour les réseaux de neurones convolutifs.

Ensuite, les images prétraitées ont été organisées dans des sous-répertoires représentant les huit catégories d'âge prédéfinies, à savoir : (0-2), (4-6), (8-13), (15-20), (25-32), (38-43), (48-53) et (60+). Cette structure hiérarchique permet un étiquetage implicite des données, facilitant ainsi leur exploitation par les générateurs de données, qui infèrent automatiquement les classes à partir des noms de dossiers.

Enfin, les données ont été partitionnées en deux ensembles distincts : un ensemble d'entraînement comprenant 80 % des images, et un ensemble de test constitué des 20 % restants. Cette répartition a été effectuée de manière à préserver la représentativité de chaque classe, assurant ainsi une évaluation équitable et rigoureuse des performances des modèles lors des expérimentations.

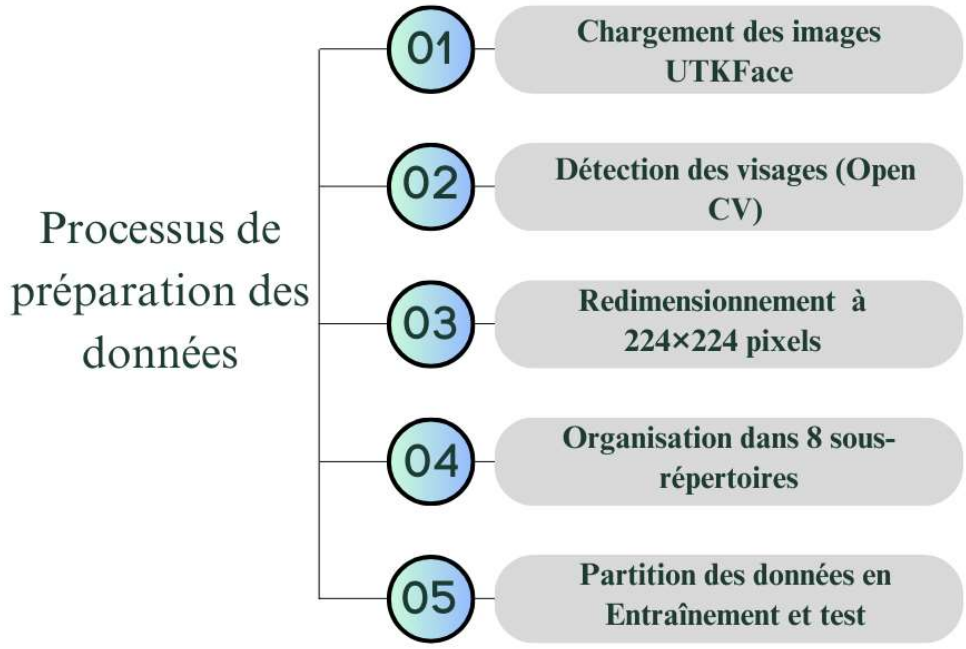


Figure 3-8 Processus de préparation des données

4 ENVIRONNEMENT DE TRAVAIL

4.1 Kaggle Notebooks

Kaggle Notebooks est un environnement de programmation en ligne proposé par la plateforme **Kaggle**, appartenant à Google. Il permet le développement d'applications en Python directement dans le cloud, sans nécessiter d'installation locale ni de configuration préalable. Ce service met à disposition des ressources matérielles puissantes, notamment des **GPU (comme NVIDIA Tesla P100 ou T4)** et **CPU**, utilisables gratuitement avec certaines limitations de temps.

Ce service nous permet :

- Le codage en langage **Python** dans un environnement interactif basé sur **Jupyter Notebook**.
- Le développement d'applications en **Deep Learning** en utilisant les bibliothèques populaires telles que **TensorFlow**, **Keras**, **PyTorch**, **OpenCV** et **scikit-learn**.
- L'exécution et la visualisation des expériences directement dans le cloud, sans besoin de configuration locale.

- L'importation facile de jeux de données via Kaggle Datasets ou des fichiers externes, ainsi que la gestion d'un espace de travail pour sauvegarder les notebooks et les résultats.

4.2 Le langage de programmation Python



Python est un langage de programmation interprété, orienté objet et de haut niveau, reconnu pour sa syntaxe claire et sa facilité d'apprentissage. Il est largement utilisé dans divers domaines tels que la programmation compétitive, le développement web, ainsi que la création de logiciels. Grâce à ses nombreuses applications dans les secteurs modernes comme la science des données, l'apprentissage automatique et l'automatisation, Python connaît une croissance rapide en popularité et figure parmi les langages les plus utilisés à travers le monde.[39]

4.3 Les bibliothèques de programmation Python

Pour la réalisation de notre système de classification d'âge et de genre basé sur des techniques d'apprentissage profond, nous avons exploité plusieurs bibliothèques Python essentielles facilitant le développement et la mise en œuvre des différentes phases du projet.

- ✓ **TensorFlow** : est une bibliothèque open source de Machine Learning développée par Google, qui permet de concevoir, entraîner et déployer des modèles d'apprentissage automatique et d'apprentissage profond (Deep Learning).
- ✓ **Keras** : Keras est une API de deep learning développée en Python, capable de s'exécuter sur des frameworks tels que JAX, TensorFlow ou PyTorch. Elle se distingue par sa simplicité d'utilisation, permettant aux développeurs de se concentrer sur les éléments essentiels sans être freinés par la complexité technique. Sa flexibilité offre un équilibre entre rapidité d'exécution pour les tâches simples et possibilités étendues pour des workflows plus avancés.[40]
- ✓ **NumPy** : constitue une bibliothèque fondamentale pour la manipulation efficace de tableaux multidimensionnels, offrant un large éventail de fonctions mathématiques et algébriques essentielles au traitement des données.
- ✓ **OpenCV** : est une bibliothèque libre dédiée à la vision par ordinateur et à l'apprentissage automatique. Elle fournit une infrastructure commune pour le

développement d'applications de traitement d'images et accélère la mise en œuvre des techniques avancées de vision artificielle.

- ✓ **Tkinter** : est la bibliothèque standard de Python pour créer des interfaces graphiques (GUI — Graphical User Interface). Elle permet de développer facilement des fenêtres, des boutons, des menus, des zones de texte, des labels, et bien d'autres éléments d'interface utilisateur.

5 Conception et implémentation de notre méthode

5.1 Architecture du système de classification de l'âge et du genre proposée

Dans cette étude, nous avons exploré deux approches complémentaires pour la classification automatique de l'âge et du genre à partir d'images faciales.

Dans le premier scénario, nous avons opté pour l'utilisation de modèles de Transfer Learning, en tirant parti de réseaux pré-entraînés tels que **VGG16**, **ResNet50**, **DenseNet121**, **MobileNetV2**, entre autres, initialement formés sur de grandes ensembles de données génériques comme ImageNet. Cette stratégie permet de capitaliser sur les capacités de généralisation de ces architectures tout en les adaptant à notre tâche spécifique par le réajustement des couches finales.

Dans le second scénario, nous avons adopté une approche hybride basée sur la fusion de deux modèles pré-entraînés. Plus précisément, notre système repose sur la concaténation des caractéristiques extraites par les modèles **VGG16** et **DenseNet121**, tous deux pré-entraînés sur ImageNet. L'objectif de cette fusion est de **t**irer parti de la complémentarité des descripteurs générés par ces deux architectures, afin d'améliorer la qualité des représentations faciales extraites et, par conséquent, la précision de la classification. La sortie fusionnée est ensuite transmise à un ensemble de couches entièrement connectées (Fully Connected**)** assurant la classification finale selon deux attributs : l'âge et le genre**.**

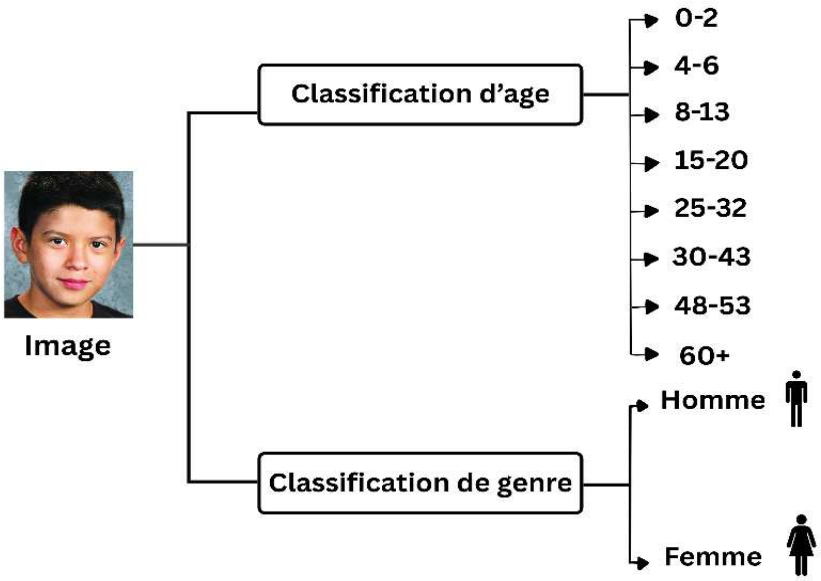


Figure 3-9 Architecture de classification des attributs faciaux : âge et genre

5.1.1 Architecture proposée basée Transfer Learning

L'architecture proposée dans ce scénario repose sur l'utilisation de modèles de Transfer Learning, en exploitant des réseaux CNN pré-entraînés tels que VGG16, ResNet50, DenseNet121, MobileNetV2, EfficientNetB0, entre autres.Ces modèles, initialement formés sur de larges bases de données telles que ImageNet, permettent d'extraire automatiquement des caractéristiques visuelles discriminantes à partir des images faciales.

Dans notre approche, les couches convolutives des modèles sont conservées pour servir d'extracteurs de caractéristiques, tandis que les couches de classification d'origine sont remplacées par un ou plusieurs blocs de couches entièrement connectées (Fully Connected), adaptés à la tâche ciblée : la classification de l'âge et du genre. Cette stratégie permet de combiner la robustesse d'un entraînement préalable massif avec une adaptation efficace aux spécificités du ensemble de données UTKFace.

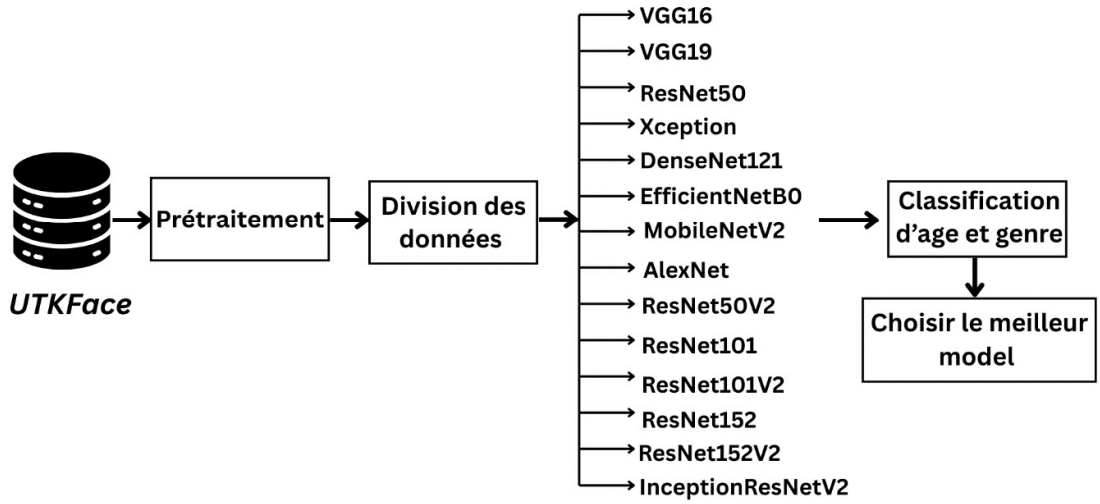


Figure 3-10 Architecture des modèles CNN pré-entraînés utilisés dans l'étude

5.1.2 Architecture base de fusion des caractéristiques extraites par deux modèles

Dans le cadre de cette étude, nous avons conçu un modèle hybride basé sur la fusion de deux réseaux de neurones convolutifs pré-entraînés : VGG16 et DenseNet121, tous deux initialement entraînés sur la base de données ImageNet. L'objectif principal de cette fusion est de combiner les forces respectives de ces deux architectures afin de produire une représentation faciale plus riche, plus robuste et plus discriminante.

L'image d'entrée, redimensionnée à 224×224×3, est transmise simultanément aux deux modèles. Le VGG16, reconnu pour sa capacité à capturer les structures globales de l'image, fournit un vecteur de caractéristiques par applatissement (Flatten) des cartes d'activation. En parallèle, DenseNet121, qui exploite des connexions denses entre ses couches pour renforcer la propagation de l'information et limiter la perte de gradient, applique une opération de GlobalAveragePooling2D pour générer un vecteur de caractéristiques compact.

Les deux vecteurs extraits sont ensuite fusionnés à l'aide d'une opération de concaténation (`Concatenate`) pour former une représentation fusionnée, qui est transmise à une couche entièrement connectée (512 neurones, activation ReLU), suivie d'une couche de Dropout (taux de 0.4) pour limiter le surapprentissage. La dernière couche dépend du type de tâche traitée :

- Pour la classification de l'âge, une couche de sortie Softmax composée de 8 neurones permet d'assigner l'image faciale à l'une des huit tranches d'âge définies dans l'ensemble de données UTKFace : (0–2), (4–6), (8–13), (15–20), (25–32), (38–43), (48–53), et (60+). Cette approche multi-classes exploite des indices visuels propres à chaque tranche pour modéliser les étapes du vieillissement.
- Pour la classification du genre, le même modèle est réutilisé, à l'exception de la couche finale, remplacée par un neurone unique avec une activation « sigmoid », qui retourne la probabilité que le visage soit celui d'une femme d'un homme). La fonction de perte utilisée est « binary_crossentropy », adaptée à cette tâche binaire.

.

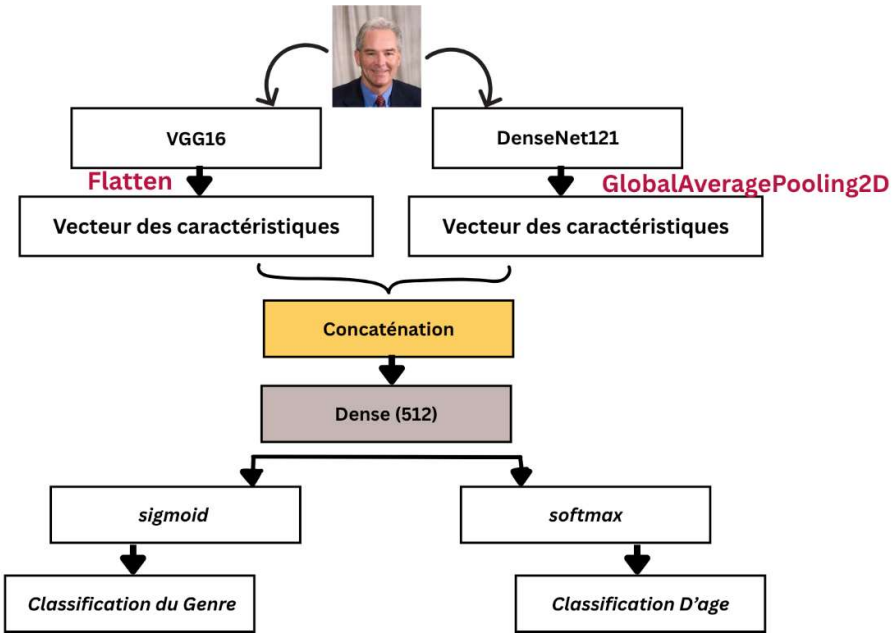


Figure 3-11 Architecture combinée VGG16–DenseNet121 pour la prédiction du genre et de l'âge

Ce modèle montre une grande capacité d'adaptation, car il est capable de traiter efficacement deux types de classification (multi-classes et binaire) à partir de la même base de caractéristiques extraites. Les résultats obtenus confirment la pertinence de cette approche pour la prédiction automatique des attributs faciaux dans des conditions variées.

5.2 Résultats expérimentaux et discussions

Afin de garantir une comparaison équitable entre les différents modèles testés (VGG16, VGG19, ResNet50, Xception, AlexNet, EfficientNetB0, MobileNetV2,ResNet101, ResNet101V2, ResNet50V2,DenseNet121, InceptionResNetV2, ResNet152V2, ResNet152 ,notre modèle hybride VGG16–DenseNet121), nous avons conservé des paramètres d'entraînement constants tout au long des expérimentations. Tous les modèles ont été entraînés dans les mêmes conditions expérimentales, en utilisant le même environnement, le même ensemble de données (UTKFace), et les mêmes hyperparamètres d'optimisation. Ces paramètres sont résumés dans le tableau suivant :

Paramètre	Valeur
Taille d'image	224 × 224 × 3
Taille du batch (Batch size)	32
Nombre d'époques	50
Optimiseur	adam
Taux d'apprentissage (LR)	0.0001
Fonction de perte (âge)	Categorical Crossentropy
Fonction de perte (genre)	Binary Crossentropy
Métrique	Accuracy, Recall, F1 score, precision
Augmentation de données	Non

Tableau 3-3Configuration d'entraînement utilisé dans les différentes expérimentations

Ces réglages ont permis d'assurer une stabilité dans l'entraînement et une comparabilité directe entre les différentes architectures testées. Les performances de chaque modèle sont détaillées dans les sous-sections suivantes.

5.2.1 Résultats expérimentaux des modèles de Transfer Learning

Les performances des modèles de réseaux de neurones convolutifs (CNN) pré-entraînés ont été évaluées dans le cadre de ce travail pour deux tâches de classification distinctes : **la prédiction de l'âge** et **la prédiction du genre** à partir d'images faciales. Pour ce faire, les expérimentations ont été menées en suivant les deux protocoles de test décrits précédemment sur l'ensemble de données **UTKFace**.

L'évaluation des performances s'est appuyée sur les métriques standards de classification, à savoir : l'accuracy**,** la précision (precision)**,** le rappel (recall) et le F1-score. Tous les modèles ont été entraînés avec une taille de lot (batch size**)** fixée à 32, afin d'assurer une comparaison équitable entre les différentes architectures testées.

5.2.1.1 Protocole 1 (80% entraînement, 20% validation, 20% test)

1) ***Classification d'âge*** Les performances de la **classification de l'âge** obtenues selon le **Protocole 1** sont synthétisées dans le tableau ci-dessous :

Model	Accuracy	Precision	Recall	F1-score
VGG16	65.48	69.93	65.48	67.10
VGG19	65.12	71.22	65.12	67.09
ResNet50	68.22	70.84	68.22	69.26
Xception	67.35	68.87	67.35	67.99
DenseNet121	67.98	70.67	67.98	69.01
EfficientNetB0	41.80	49.39	41.80	42.67
MobileNetV2	55.43	61.57	55.43	57.55
AlexNet	66.78	62.85	66.78	60.34
ResNet50V2	68.28	69.83	68.28	68.93
ResNet101	70.54	71.76	70.54	71.06
ResNet101V2	69.94	71.87	69.94	70.75
ResNet152	65.00	67.02	65.00	65.86
ResNet152V2	68.10	70.25	68.10	69.00
InceptionResNetV2	66.93	69.50	66.93	67.97

Tableau 3-4 Résultats des modèles CNN pré-entraînés pour la classification de l'âge selon le Protocole

2) Analyse des résultats de la classification de l'âge (Protocole 1)

L'analyse des résultats de la **classification de l'âge** selon le Protocole 1 révèle plusieurs points clés :

- Le modèle **ResNet101** se distingue en affichant la meilleure précision (**70,54 %**) et le meilleur F1-score (**71,06 %**), montrant sa capacité à capturer les variations d'âge de manière plus robuste.
- Les architectures issues de la famille **ResNet** (ResNet50, ResNet50V2, ResNet101V2, ResNet152V2) occupent globalement le haut du classement, ce qui souligne l'efficacité des connexions résiduelles pour la prédiction d'attributs faciaux continus comme l'âge.
- Les modèles plus légers ou spécialisés, tels que **MobileNetV2** et **EfficientNetB0**, présentent des performances sensiblement inférieures (respectivement 55,43 % et 41,80 % d'accuracy), ce qui met en évidence leurs limites lorsque la profondeur du réseau et la richesse des caractéristiques sont déterminantes.

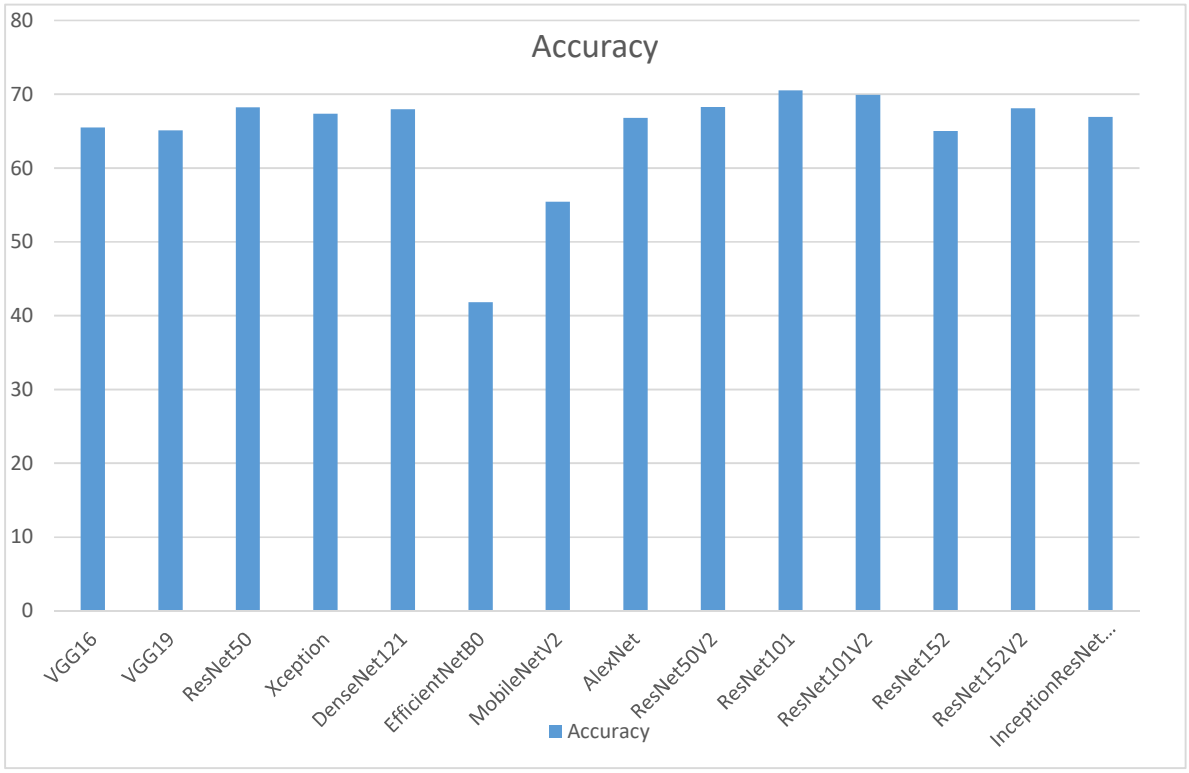


Figure 3-12 Comparaison de l'Accuracy des modèles CNN pré-entraînés pour la classification de l'âge (Protocole 1)

5.2.1.2 Protocole 2 *(80% entraînement, 20% test)*

1) Classification d'âge :

Les performances de la classification de l'âge obtenues selon le Protocole 2 sont synthétisées dans le tableau ci-dessous

Model	Accuracy	Precision	Recall	F1-score
VGG16	74.00	72.54	74.00	72.95
VGG19	73.13	72.60	73.13	72.34
ResNet50	72.62	71.62	72.65	71.72
Xception	66.60	61.05	66.60	61.45
DenseNet121	73.94	72.61	73.94	73.04
EfficientNetB0	67.65	62.94	67.65	62.40
MobileNetV2	69.97	68.08	69.97	68.60
AlexNet	71.86	68.57	71.86	69.01
ResNet50V2	72.31	70.18	72.31	70.82
ResNet101	73.37	71.91	73.37	72.22
ResNet101V2	72.80	70.90	72.80	71.55
ResNet152	74.39	73.85	74.39	73.80
ResNet152V2	73.16	72.55	73.16	72.53
InceptionResNetV2	71.98	70.09	71.98	70.50

Tableau 3-5Performances des modèles CNN pré-entraînés pour la classification de l'âge selon le Protocole 2

2) Analyse des résultats de la classification de l'âge (Protocole 2)

- L'évaluation des performances selon le Protocole 2, qui repose sur une division directe en 80 % pour l'entraînement et 20 % pour le test sans ensemble de validation, montre des résultats cohérents avec ceux obtenus dans le Protocole 1, tout en soulignant quelques différences notables.

- Le modèle ResNet152 se démarque en obtenant la meilleure précision (accuracy) de 74,39 % et le meilleur F1-score de 73,80 %, suivi de très près par VGG16 (74,00 %) et DenseNet121 (73,94 %). Ces modèles confirment leur capacité à extraire des caractéristiques discriminantes à partir des visages, même sans validation intermédiaire.
- Les modèles de la famille ResNet (ResNet101, ResNet50V2, ResNet152V2...) présentent globalement des performances stables et robustes, avec des F1-scores situés entre 70 % et 73 %, ce qui renforce leur fiabilité dans des contextes pratiques.
- À l'inverse, les architectures plus légères comme EfficientNetB0 (67,65 %) et MobileNetV2 (69,97 %) affichent des résultats moins compétitifs, ce qui confirme leurs limites face à des tâches complexes telles que l'estimation de l'âge à partir d'images faciales.

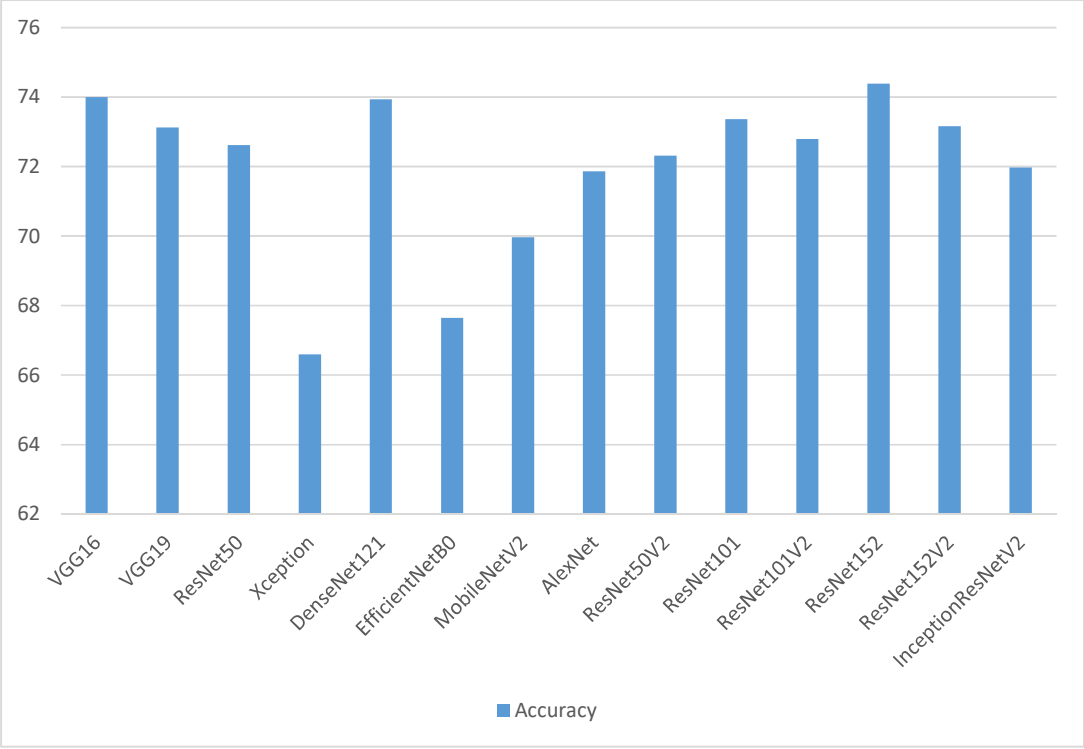


Figure 3-13 Comparaison de l'Accuracy des modèles CNN pré-entraînés pour la classification de l'âge (Protocole 2)

3) Classification de genre

Les performances de la **classification de genre** obtenues selon le **Protocole 2** sont synthétisées dans le tableau ci-dessous :

Model	Accuracy	Precision	Recall	F1-score
VGG16	91.30	91.30	91.30	91.30
VGG19	91.63	91.64	91.63	91.64
ResNet50	92.03	92.04	92.03	92.02
Xception	91.00	91.00	91.00	91.00
DenseNet121	91.66	91.67	91.66	91.66
EfficientNetB0	60.76	73.00	60.76	53.35
MobileNetV2	87.75	87.76	87.75	87.75
AlexNet	89.77	89.79	89.77	89.76
ResNet50V2	91.21	91.22	91.21	91.21
ResNet101	92.24	92.24	92.23	92.24
ResNet101V2	91.90	91.91	91.90	91.90
ResNet152	92.06	92.07	92.06	92.06
ResNet152V2	92.21	92.21	92.21	92.20
InceptionResNetV2	91.45	91.47	91.45	91.45

Tableau 3-6Performances des modèles CNN pré-entraînés pour la classification du genre selon le Protocole 2

4) Analyse des résultats de la classification du genre (Protocole 2)

Les résultats obtenus pour la classification du genre à partir d'images faciales selon le Protocole 2 révèlent un très bon niveau de performance pour la majorité des modèles testés. Le modèle ResNet101 enregistre les meilleures performances avec une accuracy de 92,24 % et un F1-score équivalent, ce qui témoigne de sa capacité à distinguer efficacement les caractéristiques faciales propres à chaque genre. D'autres architectures de la même famille, telles que ResNet152V2**,** ResNet50 et ResNet101V2, atteignent également des scores élevés, dépassant les 91 % de précision, ce qui confirme la stabilité et l'efficacité des réseaux résiduels dans cette tâche.

Les modèles **VGG16**, **VGG19**, **DenseNet121**,et **InceptionResNetV2** montrent également de très bonnes performances, oscillant entre 91,30 % et 91,66 %**,** ce qui les rend tout à fait compétitifs pour la reconnaissance du genre.

En revanche, le modèle EfficientNetB0 affiche une baisse significative de performance avec une accuracy de seulement **60,76 %** et un F1-score de **53,35 %**, soulignant ses limitations dans la détection du genre à partir de visages. MobileNetV2, bien que plus performant qu'EfficientNetB0, reste en retrait par rapport aux autres modèles avec **87,75 %** de précision.

En résumé, les modèles ResNet et VGG se distinguent nettement dans cette tâche, confirmant leur adaptabilité aux problèmes de classification binaire liés aux attributs faciaux.

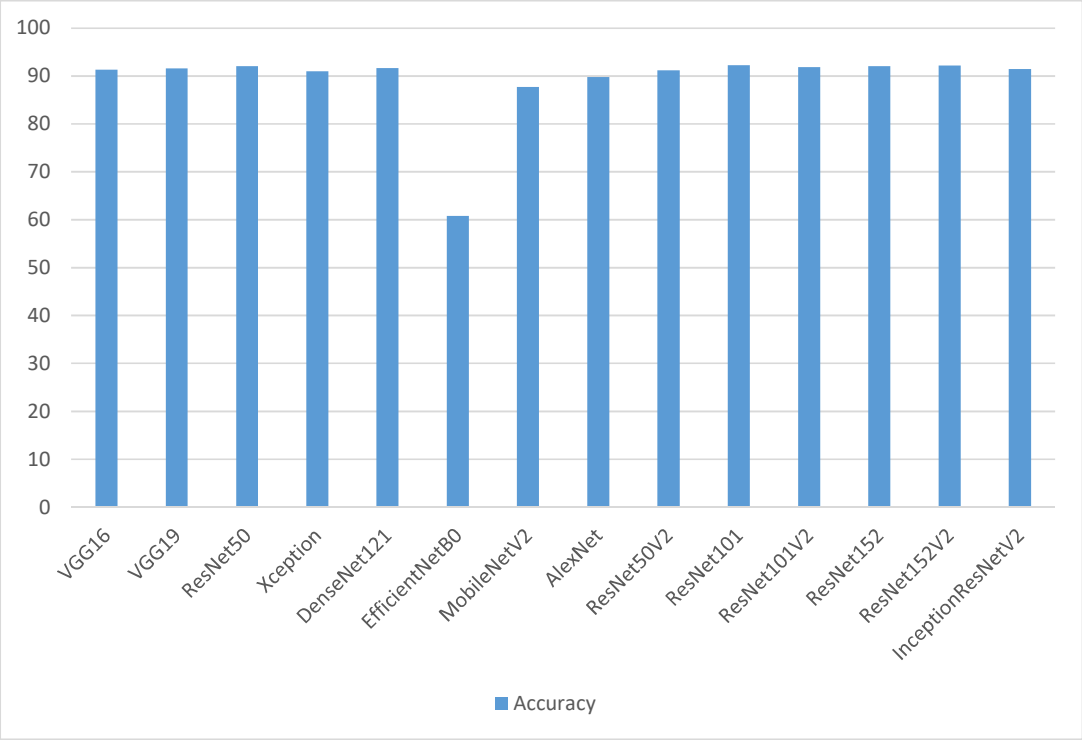


Figure 3-14 Comparaison de l'Accuracy des modèles CNN pré-entraînés pour la classification de genre (Protocole 2)

5.2.2 Résultats expérimentaux du modèle de fusion VGG16–DenseNet121

Cette section expose de manière détaillée les performances expérimentales obtenues à l'aide de notre modèle hybride basé sur la fusion des architectures VGG16 et DenseNet121. L'ensemble des expérimentations a été conduit conformément au protocole 2, reposant sur une division des données en 80 % pour l'apprentissage et 20 % pour les tests, ce qui permet d'assurer

une évaluation fiable, cohérente et représentative de la capacité généralisante du modèle proposé.

5.2.2.1 Classification d'âge

Pour évaluer les performances du modèle sur la tâche de classification de l'âge**,** différents outils d'analyse ont été mobilisés, tels que le rapport de classification, les courbes d'évolution de la fonction de perte et de la précision, ainsi que la matrice de confusion.

	precision	recall	f1-score	support
0-2	0.9284	0.9688	0.9482	321
15-20	0.6000	0.5143	0.5538	245
25-32	0.7902	0.9033	0.8430	1355
38-43	0.4468	0.3144	0.3691	334
4-6	0.7347	0.6000	0.6606	120
48-53	0.4564	0.3436	0.3921	259
60+	0.8318	0.8724	0.8516	533
8-12	0.7394	0.6731	0.7047	156
accuracy			0.7514	3323
macro avg	0.6910	0.6487	0.6654	3323
weighted avg	0.7313	0.7514	0.7374	3323

Figure 3-15 Résultats obtenus pour notre model

La **Figure 3-15** présente un rapport de classification mettant en évidence les performances globales du modèle de fusion **VGG16–DenseNet121** appliqué à la classification d'âge. Le modèle atteint une précision (accuracy) de **0.7514** et un F1-score pondéré de **0.7374**, traduisant une capacité de généralisation satisfaisante sur l'ensemble des 3323 échantillons de test. Toutefois, une analyse plus fine par classe révèle une variabilité significative des performances selon les groupes d'âge. Les classes **« 0–2 »**, **« 25–32 »** et **« 60+ »** obtiennent des F1-scores élevés (respectivement **0.9482**, **0.8430** et **0.8516**), indiquant une bonne capacité du modèle à les distinguer, en particulier la classe « 25–32 » qui, bénéficiant du plus grand nombre d'instances (**1355**), présente une prédiction particulièrement robuste. En revanche, les classes **« 38–43 »** et **« 48–53 »** enregistrent des F1-scores nettement plus faibles (**0.3691** et **0.3921**), révélant des difficultés de classification attribuables à leur sous-représentation dans le jeu de données et à une similarité morphologique accrue avec d'autres tranches d'âge, ce qui complique leur différenciation.

Pour la visualisation des performances du model proposé, nous pouvons tracer les courbes des métriques d'apprentissage et les met à jour automatiquement après chaque époque. La bibliothèque « *Matplotlib* » propose de nombreux outils et fonctions pour ce processus de

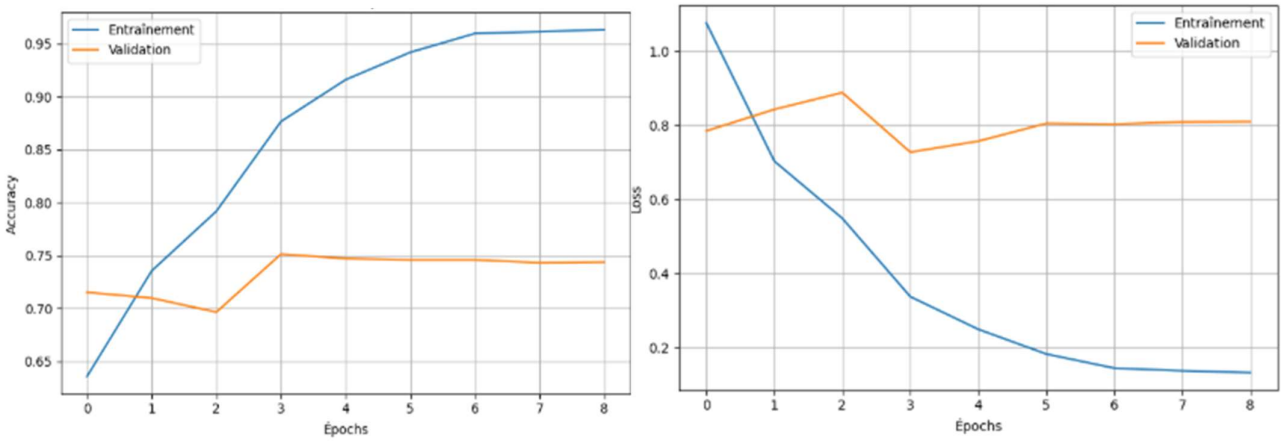


Figure 3-16 Graphe de précision et de perte du notre modèle pour la classification de l'âge

visualisation. Dans la figure 3.10, nous présentons un exemple des courbes pour les deux métriques : Précision « *accuaracy* » et Perte « *loss* » pendant l'apprentissage du modèle.

L'analyse des courbes d'évolution de la précision et de la perte montre que le modèle commence à surapprendre dès la 3e époque. Durant les premières itérations, la précision de validation s'améliore et la perte diminue, ce qui indique un apprentissage efficace et une bonne généralisation. Toutefois, à partir de la 3e époque, la précision de validation stagne tandis que la perte augmente, contrastant avec la progression continue sur les données d'entraînement.

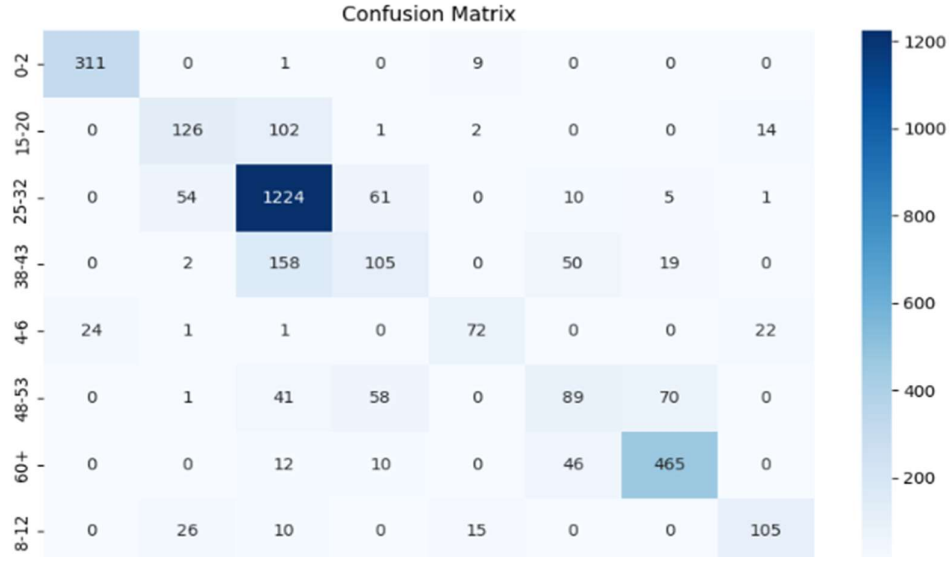


Figure 3-17 La matrice de confusion du notre modèle pour la classification de l'âge

Cette divergence entre les performances d'entraînement et de validation traduit une mémorisation excessive des données d'apprentissage au détriment de la généralisation, ce qui limite les performances du modèle sur des données inédites.

L'analyse de la matrice de confusion montre que certaines classes d'âge comme "0-2", "25-32" et "60+" sont bien reconnues par le modèle, avec un grand nombre de prédictions correctes et peu de confusions. En revanche, d'autres classes intermédiaires comme "38-43", "48-53" et "15-20" posent davantage de difficultés, étant souvent confondues avec des classes voisines. Par exemple, la classe "38-43" est fréquemment prédite comme "25-32" ou "48-53", et "48-53" est souvent prise pour "38-43" ou "60+", ce qui reflète la difficulté du modèle à bien délimiter les frontières entre ces tranches d'âge proches. Ces erreurs peuvent être accentuées par la proximité visuelle entre les âges et par un déséquilibre de représentation entre les classes. Globalement, les confusions se produisent majoritairement entre tranches d'âge adjacentes, ce qui indique que, malgré ses performances solides sur certaines classes, le modèle montre des limites à bien généraliser les nuances d'âge dans les intervalles intermédiaires.

5.2.2.2 Classification de genre

De la même manière, pour la tâche de classification du genre, plusieurs outils d'analyse ont été utilisés afin d'évaluer les performances du modèle, notamment le rapport de classification, les courbes d'évolution de la précision et de la perte, ainsi que la matrice de confusion.

	precision	recall	f1-score	support
Male	0.9408	0.9224	0.9315	1739
Female	0.9166	0.9362	0.9263	1584
accuracy			0.9290	3323
macro avg	0.9287	0.9293	0.9289	3323
weighted avg	0.9292	0.9290	0.9290	3323

Figure 3-18 Rapport de classification de model fusion vgg16 et denceNet121 pour classification de genre

Le rapport de classification (Figure31) du modèle de détection du genre révèle de très bonnes performances avec une précision globale de 92.9 %. Les deux classes (Homme et Femme) affichent des scores F1 équilibrés, respectivement de 93.15 % et 92.63 %, ce qui témoigne d'une bonne capacité de généralisation. L'ensemble de données est relativement

équilibré (8543 hommes et 8071 femmes), ce qui renforce la fiabilité des résultats. En somme, le modèle parvient à distinguer efficacement les deux genres avec une faible marge d'erreur.

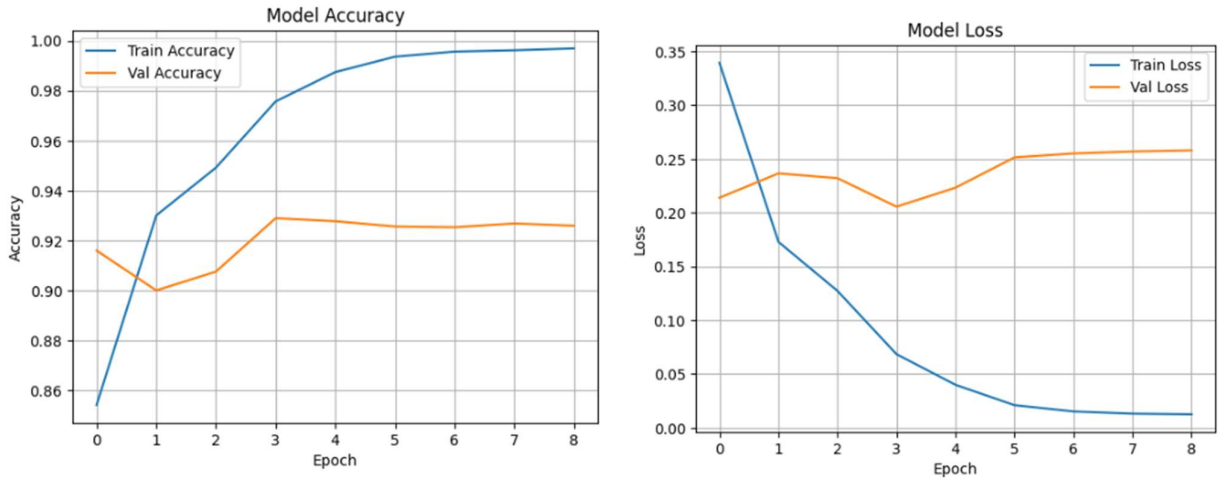


Figure 3-19 Graphe de précision et de perte de notre modèle pour la classification du genre

L'analyse de l'évolution de la précision et de la perte montre que le modèle atteint une excellente performance dès les premières époques, avec une précision de validation maximale de 92.90 % obtenue à l'époque 4. La perte d'entraînement diminue régulièrement, atteignant 0.0162 à l'époque 9, tandis que la perte de validation commence à augmenter après l'époque 4, indiquant un début de surapprentissage. Ainsi, bien que l'précision d'entraînement continue de s'améliorer jusqu'à 99.57 %, la stabilité de la précision de validation autour de 92.5 % suggère que les meilleures performances de généralisation sont atteintes entre les époques 4 et 5.

La figure 3-20 illustre la matrice de confusion obtenue après l'évaluation du modèle, montrant les performances de classification du genre.

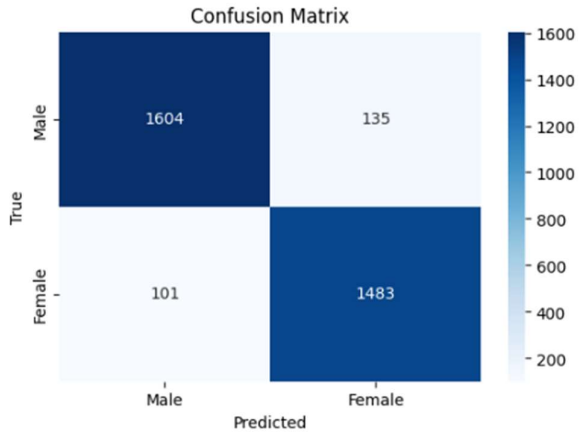


Figure 3-20 Matrice de confusion de notre modèle pour la classification du genre

6 Comparaison des résultats avec les travaux existants

Les résultats issus de l'état de l'art montrent des performances variables selon les méthodes utilisées pour la classification d'âge et de genre. Par exemple, Zixuan Lyu [13] rapporte une précision de 65 % pour la classification d'âge, sans indication pour le genre. L'approche par fusion de modèles VGG16 et DenseNet121 améliore ces performances, atteignant 75,14 % pour l'âge et 92,90 % pour le genre, dépassant ainsi les résultats obtenus avec VGG16 (74,00 % âge, 91,30 % genre) ou DenseNet121 (73,94 % âge, 91,66 % genre) seuls. Les travaux de Dey et al. obtiennent une précision notable de 81,96 % pour la classification d'âge et 96,32 % pour le genre, tandis que Noor et al. atteignent 78,67 % pour l'âge et 97,02 % pour la classification de genre. Enfin, Kumar et al. [34] obtiennent la meilleure performance pour l'âge avec 84,8 %, bien que leurs résultats pour le genre ne soient pas reportés.

Ces variations de performance peuvent être liées aux différences de datasets utilisés, aux protocoles d'entraînement, ainsi qu'aux techniques d'optimisation et de prétraitement appliquées.

État de l'art	***Performance (Accuracy%)***	
	Age	**Genre**
Zixuan Lyu. [13]	65.00	/
Model fusion (vgg16 /denseNet121)	75.14	92.90
Vgg16	74.00	91.30
DenseNet 121	73.94	91.66
Dey et al.	81.96	96.32
Kumar et al. [34]	84.8	/
Noor et al.[29]	78.67	97.02

Tableau 3-7 Tableau comparatifs des résultats obtenus par rapports aux travaux existants

7 Conclusion

Dans ce chapitre, nous avons conçu et mis en œuvre un modèle de fusion basé sur les architectures VGG16 et DenseNet121 pour la classification conjointe de l'âge et du genre à partir d'images faciales. En utilisant l'ensemble de données UTKFace, préalablement prétraitée et organisée selon des tranches d'âge et des catégories de genre, le modèle a atteint des niveaux de performance satisfaisants, avec une précision de **75,14 % pour l'estimation de l'âge** et **92,90 % pour la classification du genre**.

Les résultats obtenus confirment l'efficacité de la stratégie de fusion, en combinant les forces respectives des deux réseaux pour améliorer la représentation des traits faciaux. Le modèle s'est montré particulièrement performant pour distinguer le genre, et relativement robuste dans la reconnaissance des tranches d'âge, malgré la complexité et la variabilité naturelle des visages humains. Néanmoins, certaines limites subsistent, notamment en ce qui concerne la gestion des classes d'âge déséquilibrées et la sensibilité au surapprentissage.

À l'avenir, ce modèle pourrait être étendu par l'intégration de modules d'attention, l'exploration de techniques de régularisation avancées, ou encore son adaptation à d'autres bases de données plus variées pour renforcer sa généralisation dans des contextes réels.

Chapitre4 : Système de détection des catégories d'âge et genre selon des images faciales

Chapitre4

Chapitre4 : Système de détection des catégories d'âge et genre selon des images faciales

1 Introduction

Dans ce chapitre, nous présentons un système intelligent permettant de prédire l'âge et le genre d'une personne à partir d'une image faciale. Ce système s'appuie sur des modèles de deep learning pré-entraînés ainsi qu'une interface graphique conviviale développée avec Tkinter. Cette application constitue une preuve de concept de l'intégration entre l'apprentissage profond et l'interaction utilisateur.

2 Description du système automatique de détection d'âge et genre

Nous proposons la conception d'un système automatique capable de prédire l'âge et le genre d'une personne à partir d'une image faciale, en temps réel, en s'appuyant sur des réseaux de neurones convolutifs (CNN) pré-entraînés. Ce système intelligent permet de détecter le visage depuis une image importée ou via une webcam, d'en extraire les caractéristiques pertinentes, puis de fournir une estimation de la tranche d'âge et du genre, composée de plusieurs étapes fondamentales :

1) ***Acquisition de l'image :*** L'image faciale peut être obtenue de deux manières :

   - **Importation depuis le disque local** : l'utilisateur sélectionne une image existante à analyser.
   - **Capture en temps réel via webcam** : le système se connecte à la caméra de l'ordinateur pour capturer une image instantanée. Cette option permet une interaction directe et dynamique avec l'utilisateur.

   Cette phase constitue l'entrée du système et conditionne la qualité des prédictions futures. Une bonne résolution et un éclairage adéquat améliorent sensiblement la performance.

2) ***Détection du visage***

Une fois l'image acquise, le système applique un algorithme de détection de visages basé sur les classificateurs Haar Cascade, une méthode proposée par Viola et Jones et implémentée dans OpenCV. Elle repose sur des caractéristiques Haar et une cascade de classificateurs entraînés, permettant de localiser rapidement la région du visage dans une image, même en présence de variations d'orientation ou d'arrière-plan.

Si aucun visage n'est détecté, le système affiche un message d'erreur à l'utilisateur afin de corriger l'entrée.

3) ***Prétraitement de l'image :*** Le visage détecté est extrait puis préparé pour être compatible avec les modèles de deep learning utilisés :

- **Redimensionnement** : l'image du visage est redimensionnée à **224x224 pixels**, une taille standard pour les réseaux convolutifs pré-entraînés tels que VGG16 ou ResNet.
- **Normalisation** : les valeurs des pixels sont converties en **valeurs flottantes** dans l'intervalle [0,1] ou transformées selon une fonction de prétraitement spécifique (`preprocess_input`).
- **Conversion de format** : l'image est convertie en un tenseur compatible avec TensorFlow/Keras, avec ajout éventuel d'une dimension batch.

Ce prétraitement est crucial pour garantir la cohérence des entrées avec les poids appris par les modèles.

4) ***Prédiction (âge et genre)***

Le cœur du système repose sur deux modèles de deep learning, chacun entraîné pour une tâche spécifique :

- **Modèle de prédiction de l'âge** : ce modèle classe l'image dans l'une des 8 tranches d'âge définies : ['0-2', '4-6', '8-12', '15-20', '25-32', '38-43', '48-53', '60+'].Le modèle retourne un vecteur de probabilités correspondant à chaque classe, dont la plus élevée détermine la prédiction finale.

- **Modèle de prédiction du genre** : ce modèle binaire distingue entre **"Homme"** et **"Femme"**. Il applique une couche « ***sigmoïde*** » finale pour produire une probabilité d'appartenance à chaque classe.

Les résultats des deux modèles sont ensuite affichés dans l'interface graphique.

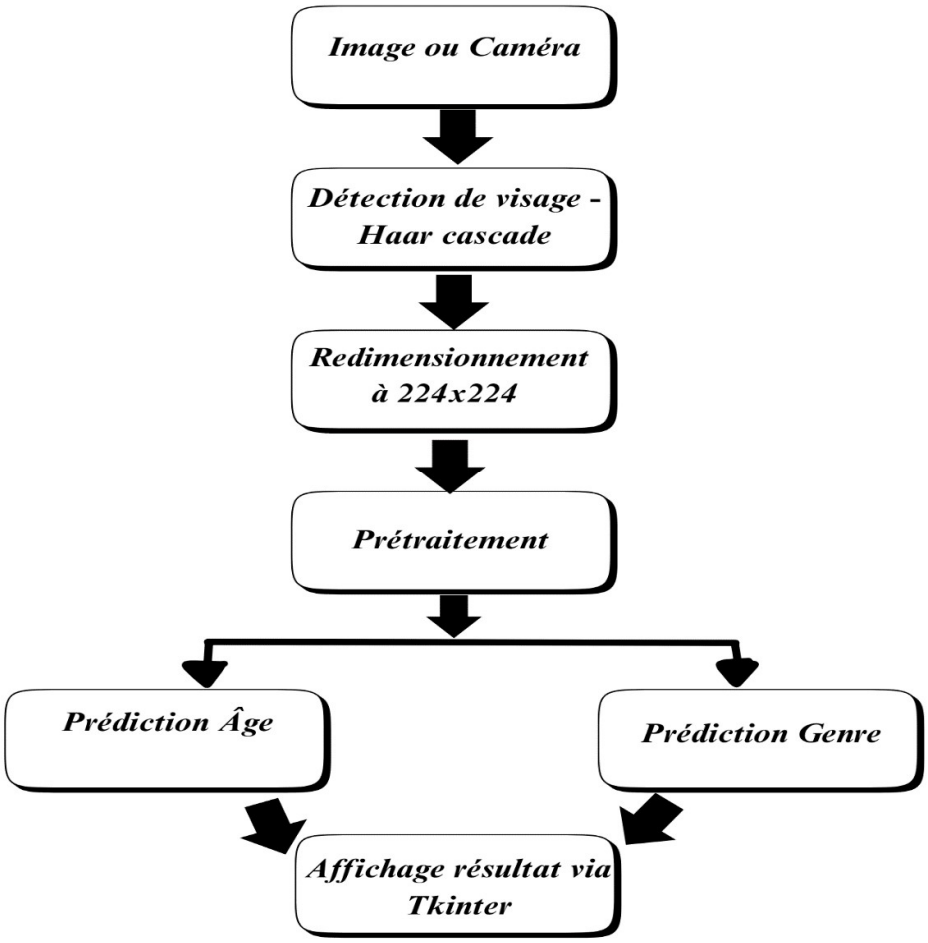


Figure 4-1 Présentation du système devine mon âge & genre

3 La structure de devine mon âge & genre

3.1 L'interface de système

L'interface graphique a été conçue avec la bibliothèque Tkinter pour Python, dans un souci d'ergonomie et de simplicité d'utilisation. Elle permet à l'utilisateur d'interagir directement avec le système sans connaissances techniques préalables.

Fonctionnalités principales :

- Importation d'une image à partir de fichiers locaux.
- Activation de la webcam pour capturer une image en temps réel.
- Affichage dynamique de l'image sélectionnée ou capturée.
- Présentation textuelle des prédic tions (genre et tranche d'âge).

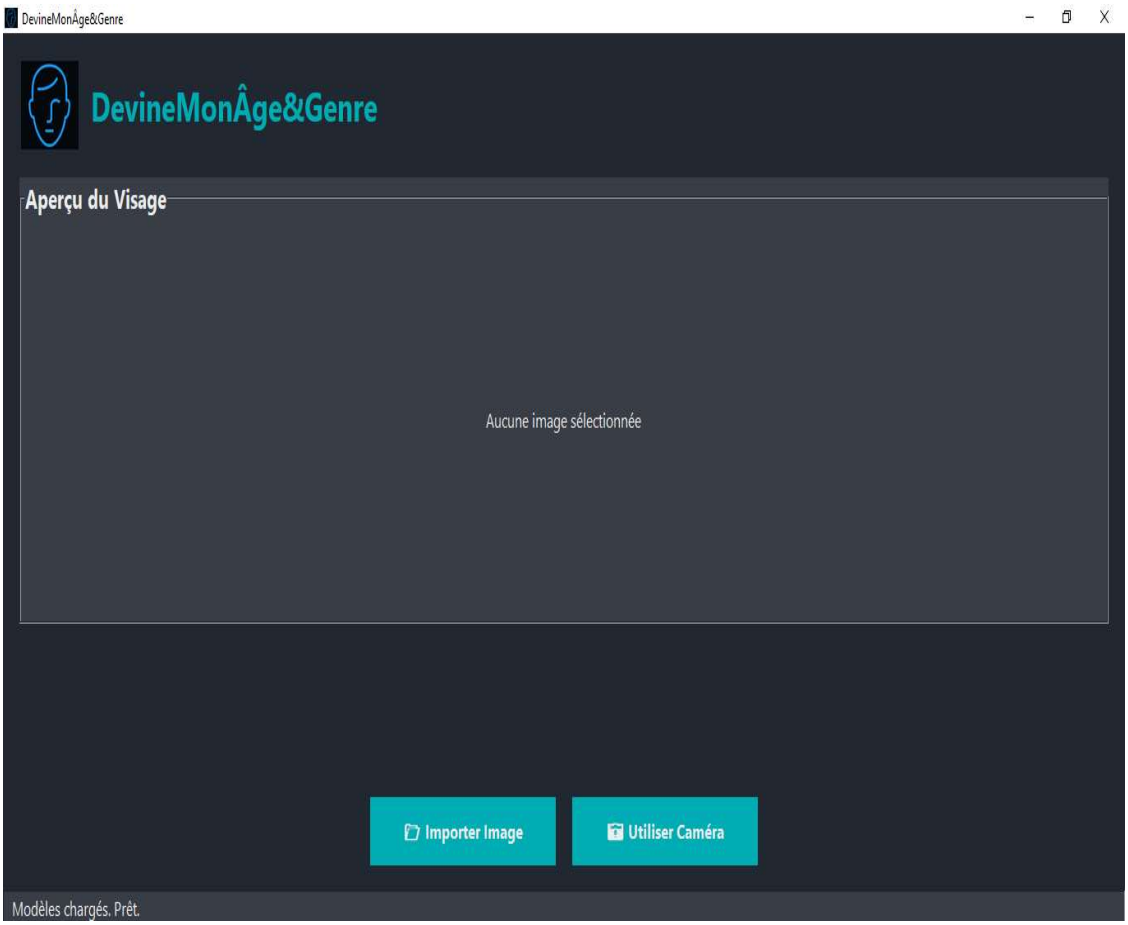


Figure 4-2 Interface principale de système

Des efforts particuliers ont été apportés à l'aspect visuel : thème sombre pour le confort visuel, intégration de logos/icônes, et gestion des erreurs par des boîtes de dialogue conviviales.

3.2 Résultats de prétraitement

Le prétraitement d'images est une étape cruciale pour l'analyse faciale. Il utilise l'algorithme Haar Cascade pour détecter et recadrer les visages, éliminant le superflu et ciblant la zone d'intérêt. Chaque visage est ensuite redimensionné à 224x224 pixels tout en conservant ses proportions, ce qui standardise les entrées pour les réseaux neuronaux. Ces manipulations réduisent la complexité de calcul et améliorent la qualité des données, permettant aux modèles d'âge et de genre de produire des prédictions plus fiables grâce à des caractéristiques faciales claires et distinctes (Figure 4-3).

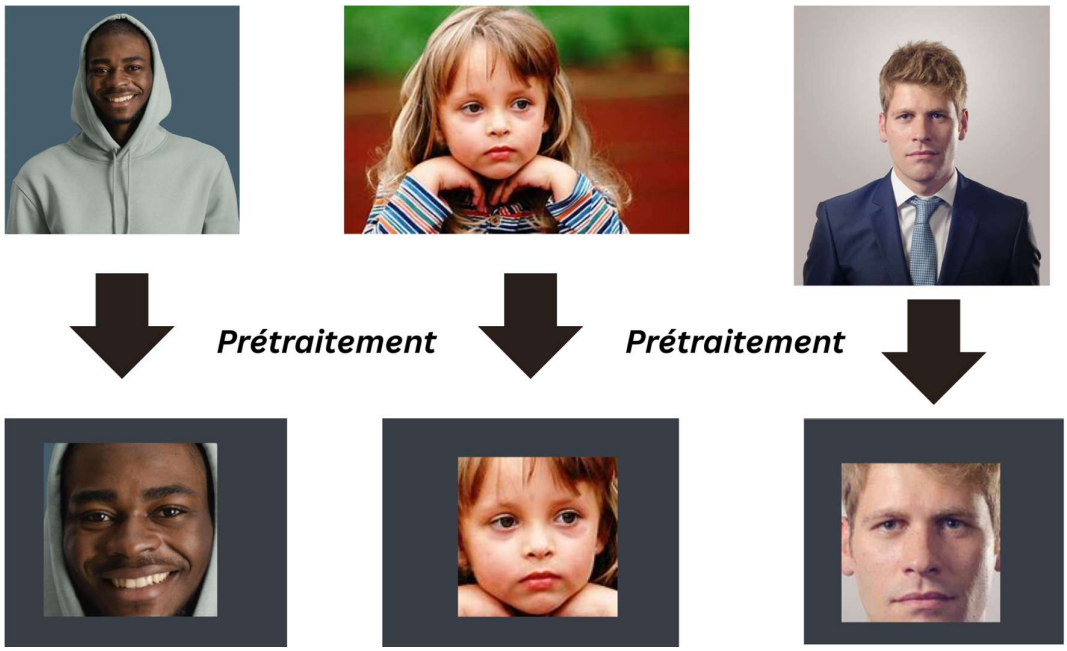


Figure 4-3 Traitement des images faciales par l'application

3.3 Résultats Obtenus selon des images faciale téléchargé

Cette section explore les résultats obtenus par le système d'analyse faciale en utilisant des images statiques importées par l'utilisateur. Après le téléchargement d'un fichier image (JPEG,

PNG, etc.), le processus commence par une étape cruciale de détection de visage. L'algorithme embarqué scanne l'image pour localiser un ou plusieurs visages, identifiant la région la plus pertinente.

Le système a été testé sur un ensemble d'images variées (adultes, enfants, visages masculins et féminins), démontrant une capacité raisonnable à estimer correctement les tranches d'âge et le genre. Le taux de réussite dépend notamment de la qualité de l'image et de l'angle du visage.

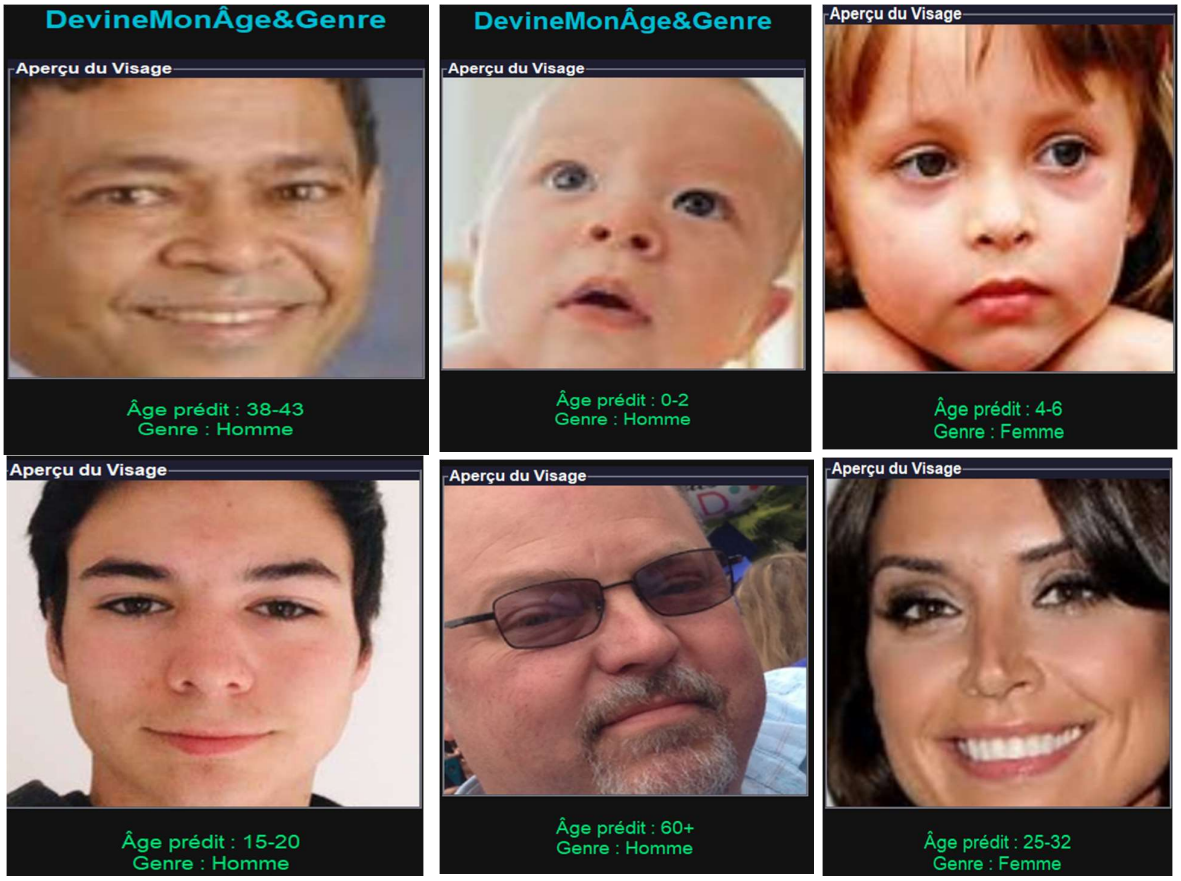


Figure 4-4 Quelques exemples de détection d'âge et genre selon des images faciale

3.4 Résultats Obtenus selon des images faciale en temps réel

Cette section affiche les performances du système développé pour l'analyse des images faciales capturées en temps réel, notamment via la fonctionnalité d'accès à la caméra.

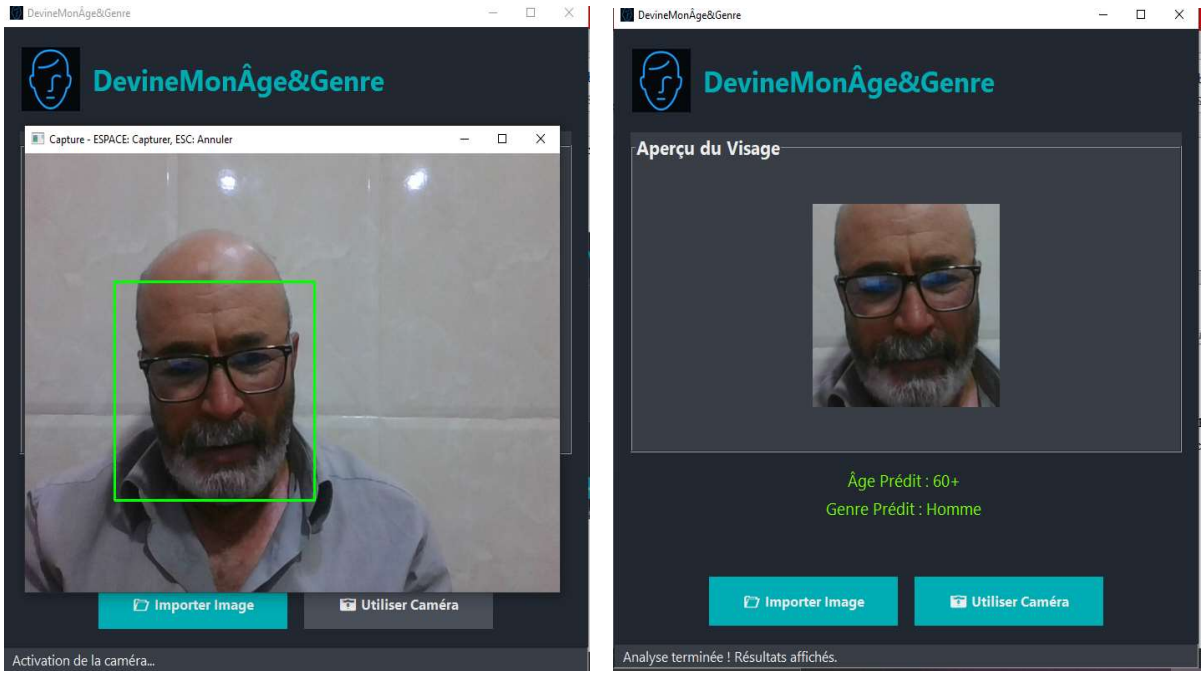


Figure 4-5 Exemple de détection d'âge et genre selon des images faciale en temps réel

4 Conclusion

Ce système représente une application concrète de la vision par ordinateur et de l'intelligence artificielle dans le domaine de l'estimation démographique. Il démontre comment des modèles de deep learning peuvent être intégrés dans une interface conviviale afin d'offrir des prédictions rapides et relativement fiables de l'âge et du genre. Bien que perfectible, cette solution pose les bases d'un outil pouvant être utilisé dans divers contextes : analyse marketing, sécurité, ou encore dans le domaine éducatif. Les perspectives d'amélioration sont nombreuses et ouvrent la voie à des applications plus complexes et plus précises à l'avenir.

Conclusion générale

Dans ce travail, nous avons abordé un défi majeur dans le domaine de la vision par ordinateur et de l'apprentissage profond, à savoir la classification de l'âge et du genre à partir d'images faciales. Ce sujet revêt une grande importance dans de nombreuses applications pratiques telles que l'analyse démographique, le marketing ciblé et les systèmes de reconnaissance d'identité.

Nous avons commencé par étudier la base de données ***UTKFace,*** *qui contient plus de 23 000 images de visages humains variés en termes d'âge, de genre et d'origine ethnique. Plusieurs défis ont été rencontrés, notamment le déséquilibre des classes, la qualité variable des images et les erreurs dans les annotations, que nous avons traités par des étapes successives de nettoyage et de préparation des données.*

Pour atteindre notre objectif, nous avons adopté deux approches principales :

- ***L'apprentissage par transfert (Transfer Learning)*** *: en utilisant des modèles de réseaux neuronaux convolutifs préentraînés tels que VGG16, ResNet et DenseNet121, avec une adaptation des couches finales pour répondre à notre tâche.*
- ***Le modèle hybride*** *: combinant les caractéristiques extraites des modèles VGG16 et DenseNet121, afin d'améliorer la précision de la classification.*

Les résultats obtenus montrent que le modèle hybride a atteint des performances remarquables, avec une précision de 75,14 % pour la classification de l'âge et de 92,90 % pour la classification du genre, surpassant les modèles individuels. Les analyses ont également révélé la difficulté à distinguer certaines tranches d'âge proches, comme 38–43 et 48–53, en raison de la similitude des traits et du faible nombre d'exemples pour certaines classes.

Comparé aux travaux antérieurs, comme celui de Zixuan Lyu (65 % pour la classification de l'âge) ou Dey et al. (81,96 %), notre modèle se distingue comme une solution équilibrée entre précision et efficacité. Toutefois, des axes d'amélioration demeurent, tels que le traitement du déséquilibre des données et l'augmentation de la précision pour les tranches d'âge sous-représentées.

En conclusion, cette recherche constitue une avancée significative vers le développement de systèmes intelligents capables d'analyser avec précision les caractéristiques humaines. Elle ouvre également la voie à des améliorations futures, comme l'intégration de techniques de renforcement profond (Deep Reinforcement Learning) ou l'élargissement du corpus à des données plus diversifiées. Nous espérons que ces efforts contribueront au développement d'applications plus intelligentes et performantes dans des domaines tels que la sécurité, la santé et le marketing intelligent.

Bibliographies

[1] S. M. Kadhim, J. Koh, S. Paw, Y. C. Tak, and S. Ameen, “Deep Learning Models for Biometric Recognition based on Face , Finger vein , Fingerprint , and Iris : A Survey,” vol. 2024, no. 01, pp. 117–157, 2024, doi: 10.2478/jsiot-2024-0007.

[2] B. Zahid and H. R. Hatem, “Biometrics based on deep learning : A survey,” *Int. J. Nonlinear Anal. Appl. ·*, no. August, pp. 1–14, 2024, doi: 10.22075/ijnaa.2024.33566.5007.

[3] Y. Zhan, X. Zhan, T. Xu, and J. Bai, “Machine Learning-Based Facial Recognition for Financial Fraud Prevention,” vol. 1, no. 1, pp. 77–84, 2024.

[4] N. E. L. Fadel, “Facial Recognition Algorithms : A Systematic Literature Review,” *J. Imaging*, vol. 11, no. 2, 2025.

[5] G. Thierry, “9 dates qui ont marqué l’essor de la reconnaissance faciale - WE DEMAIN.” Accessed: Feb. 27, 2025. [Online]. Available: https://www.wedemain.fr/inventer/9-dates-qui-ont-marque-l-essor-de-la-reconnaissance-faciale_a4568-html/

[6] “Meta veut améliorer la détection des publicités frauduleuses grâce à la reconnaissance faciale.” Accessed: Feb. 27, 2025. [Online]. Available: https://www.lemonde.fr/pixels/article/2024/10/22/meta-veut-ameliorer-la-detection-des-publicites-frauduleuses-grace-a-la-reconnaissance-faciale_6358245_4408996.html

[7] Y. Kortli, M. Jridi, M. Merzougui, A. Alasiry, and M. Atri, “Comparative Study of Face Recognition Approaches,” *Proc. Int. Conf. Adv. Syst. Emergent Technol. IC_ASET 2020*, no. February 2021, pp. 300–305, 2020, doi: 10.1109/IC_ASET49463.2020.9318305.

[8] Y. Kortli, M. Jridi, A. Al Falou, and M. Atri, “Face recognition systems: A survey,” *Sensors (Switzerland)*, vol. 20, no. 2, 2020, doi: 10.3390/s20020342.

[9] S. Haseena, S. Saroja, R. Madavan, A. Karthick, B. Pant, and M. Kifetew, “Prediction of the Age and Gender Based on Human Face Images Based on Deep Learning Algorithm,” *Comput. Math. Methods Med.*, vol. 2022, 2022, doi: 10.1155/2022/1413597.

[10] O. O. Victoria and O. O. Kayode, “Age Estimation from Facial Images of Human Beings,” vol. 7, no. 6, pp. 116–120, 2023, doi: 10.13140/RG.2.2.20238.78402.

[11] A. S. Al-Shannaq and L. A. Elrefaei, “Comprehensive Analysis of the Literature for Age Estimation from Facial Images,” *IEEE Access*, vol. 7, pp. 93229–93249, 2019, doi: 10.1109/ACCESS.2019.2927825.

[12] K. ELKarazle, V. Raman, and P. Then, “Facial Age Estimation Using Machine Learning Techniques: An Overview,” *Big Data Cogn. Comput.*, vol. 6, no. 4, 2022, doi: 10.3390/bdcc6040128.

[13] Z. Lyu, “Age Estimation from Facial Photos: A CNN-Based Approach with Multi-Model Feature Fusion,” *Highlights Sci. Eng. Technol.*, vol. 85, pp. 363–375, 2024, doi: 10.54097/g24fan25.

[14] I. Adjabi, A. Ouahabi, A. Benzaoui, and A. Taleb-Ahmed, “Past, present, and future of face recognition: A review,” *Electron.*, vol. 9, no. 8, pp. 1–53, 2020, doi: 10.3390/electronics9081188.

[15] “Intelligence Artificielle : Définition, histoire, enjeux.” Accessed: Apr. 29, 2025. [Online]. Available: https://datascientest.com/intelligence-artificielle-definition

[16] “Qu’est-ce que le machine learning (ML) ? | IBM.” Accessed: May 09, 2025. [Online]. Available: https://www.ibm.com/fr-fr/think/topics/machine-learning

[17] “Machine Learning vs Deep Learning : Quelles différences ?” Accessed: May 09, 2025. [Online]. Available: https://datascientest.com/quelle-difference-entre-le-machine-learning-et-deep-learning

[18] “An Introduction to Convolutional Neural Networks: A Comprehensive Guide to CNNs in Deep Learning | DataCamp.” Accessed: May 10, 2025. [Online]. Available:

https://www.datacamp.com/tutorial/introduction-to-convolutional-neural-networks-cnns

[19] "Introduction to Convolution Neural Network | GeeksforGeeks." Accessed: May 10, 2025. [Online]. Available: https://www.geeksforgeeks.org/introduction-convolution-neural-network/

[20] O. Abhulimen and E. Ogunti, "Facial Age Estimation Using Deep Learning : A Review," *J. Multidiscip. Eng. Sci. Technol.*, vol. 8, no. 5, pp. 13927–13946, 2021.

[21] A. Panda, D. Panigrahi, S. Mitra, S. Mittal, and S. Rahimi, "Transfer Learning Applied to Computer Vision Problems: Survey on Current Progress, Limitations, and Opportunities," vol. 1, no. 1, pp. 1–16, 2024, [Online]. Available: http://arxiv.org/abs/2409.07736

[22] M. A. Rahman, S. S. Aonty, K. Deb, and I. H. Sarker, "Attention-Based Human Age Estimation from Face Images to Enhance Public Security," *Data*, vol. 8, no. 10, Oct. 2023, doi: 10.3390/data8100145.

[23] H. A. Dahlan, "A Survey on Deep Learning Face Age Estimation Model : Method and Ethnicity," vol. 12, no. 11, pp. 86–101, 2021.

[24] O. Agbo-Ajala and S. Viriri, *Deep learning approach for facial age classification: a survey of the state-of-the-art*, vol. 54, no. 1. Springer Netherlands, 2021. doi: 10.1007/s10462-020-09855-0.

[25] C. M. Steel, "Naï ve Bayesian Minor Detection Using Facial Anthropometric Ratios," 2013, doi: 10.13140/RG.2.2.11689.29280.

[26] S. Wang, D. Tao, and J. Yang, "Relative Attribute SVM+ Learning for Age Estimation," *IEEE Trans. Cybern.*, vol. 46, no. 3, pp. 827–839, 2016, doi: 10.1109/TCYB.2015.2416321.

[27] A. Solanki and V. Priyadarshni, "Human Age Classification System Using K-NN Classifier," *Recent Patents Comput. Sci.*, vol. 12, no. February 2019, 2019, doi: 10.2174/2213275912666190226161715.

[28] B. Reddy and A. Reddy, *Machine Learning Technologies and Applications*, no. March. 2021. doi: 10.1007/978-981-33-4046-6.

[29] A. Noor, W. Ahmad, and S. Adnan, "Deep Faces: Advancing Age and Gender Classification using Facial Images with Deep Features," *IJIST*, vol. 6, no. 2, pp. 808–818, 2024, [Online]. Available: https://www.researchgate.net/publication/383863794

[30] H. T. Nguyen, L. T. T. Pham, D. T. Dang, S. N. Huynh, P. H. Dang, and Q. T. H. Nguyen, "Age Prediction from Facial Images Using Deep Learning Architecture," *Appl. Comput. Syst.*, vol. 29, no. 2, pp. 22–29, Dec. 2024, doi: 10.2478/acss-2024-0018.

[31] S. Sathyavathi and K. R. Baskaran, "An Intelligent Human Age Prediction from Face Image Framework Based on Deep Learning Algorithms," *Inf. Technol. Control*, vol. 52, no. 1, pp. 245–257, Mar. 2023, doi: 10.5755/j01.itc.52.1.32323.

[32] S. Eddine Bekhouche *et al.*, "Facial Age Estimation Using Multi-Stage Deep Neural Networks Academic Editors: Prayag Tiwari, Facial Age Estimation Using Multi-Stage Deep Neural Networks," 2024, doi: 10.3390/electron.

[33] G. George, S. A. Adeshina, and M. Mahammat Boukar, "A Deep CNN technique for Age Estimations from Facial Images."

[34] R. Kumar, K. Singh, D. P. Mahato, and U. Gupta, "Face-based age and gender classification using deep learning model," *Procedia Comput. Sci.*, vol. 235, pp. 2985–2995, 2024, doi: 10.1016/j.procs.2024.04.282.

[35] A. Garain, B. Ray, P. K. Singh, A. Ahmadian, N. Senu, and R. Sarkar, "GRA_Net: A Deep Learning Model for Classification of Age and Gender from Facial Images," *IEEE Access*, vol. 9, pp. 85672–85689, 2021, doi: 10.1109/ACCESS.2021.3085971.

[36] Y. H. Kim, S. H. Nam, S. B. Hong, and K. R. Park, "GRA-GAN: Generative adversarial network for image style transfer of Gender, Race, and age," *Expert Syst. Appl.*, vol. 198, no. January, p. 116792, 2022, doi: 10.1016/j.eswa.2022.116792.

[37] P. Dey, T. Mahmud, M. S. Chowdhury, M. S. Hossain, and K. Andersson, "Human

Age and Gender Prediction from Facial Images Using Deep Learning Methods," *Procedia Comput. Sci.*, vol. 238, no. 2019, pp. 314–321, 2024, doi: 10.1016/j.procs.2024.06.030.

[38] K. Jain, M. Chawla, A. Gadhwal, R. Jain, and P. Nagrath, *Age and Gender Prediction Using Convolutional Neural Network*, vol. 121, no. Ic4s. 2020. doi: https://doi.org/10.1007/978-981-15-3369-3_19.

[39] "What is Python? Its Uses and Applications | GeeksforGeeks." Accessed: May 15, 2025. [Online]. Available: https://www.geeksforgeeks.org/what-is-python/

[40] "About Keras 3." Accessed: May 15, 2025. [Online]. Available: https://keras.io/getting_started/about/