

UNIVERSITÉ KASDI MERBAH -OUARGLA  
Faculté Des Nouvelles Technologies De L'information Et De la Communication  
Département D'électronique Et De Télécommunication



**MÉMOIRE PRESENTE POUR L'OBTENTION DU DIPLOME DE  
MASTER**

**Domaine :** Science et technologie

**Option :** Télécommunications

**Spécialité :** Système des télécommunications

**Présenté par :**

❖ M<sup>elle</sup>. HADDAOUI Anfal Meriem

❖ M<sup>elle</sup>. RAHMOUNI Yasmina

**Thème :**

**La Prédiction De La Beauté Faciale Basée  
Sur Réseaux De Neurone Convolutif**

Soutenu publiquement le : 22/06/2023

Devant le jury composé de :

|                            |              |      |              |
|----------------------------|--------------|------|--------------|
| Pr. Bettayeb Nadjla        | Président    | Prof | Univ Ouargla |
| Dr. Benlamoudi Azeddine    | Encadreur    | MCA  | Univ Ouargla |
| Mr. Boukhari Djamel-Eddine | Co-Encadreur | MR   | Univ El oued |
| Dr. Korichi Maarouf        | Examineur    | MCA  | Univ Ouargla |

**Année universitaire : 2022/2023**

## Résumé

L'évaluation de la beauté du visage est désormais possible grâce aux progrès récents de l'intelligence artificielle et de la vision par ordinateur. La prédiction de la beauté faciale est un problème de reconnaissance visuelle important pour l'évaluation de l'attractivité du visage, qui correspond à la perception humaine. Les méthodes d'apprentissage en profondeur ont récemment démontré une capacité étonnante pour la représentation et l'analyse des caractéristiques.

Dans ce travail, nous avons étudié la prédiction de la beauté faciale, en particulier pour les méthodes d'apprentissage profond en utilisant les réseaux de neurones convolutifs.

Dans le cadre de cette mémoire, nous travaillons sur la synthèse de travaux récents basée sur la prédiction de la beauté du visage en fonction de techniques d'apprentissage en profondeur et d'ensembles de données.

Les résultats obtenus par ces techniques d'apprentissage en termes d'évaluation des performances PC, MAE et RMSE obtenus par les réseaux de neurones convolutifs sont améliorés surtout en entraînés avec différentes fonctions de perte. Enfin, nos résultats sont comparables à ceux obtenus par les AlexNet, ResNet-18, ResNeXt-50, CNN – SCA, R3CNN et Semi-supervised.

**Mots-clés :** *Réseaux de neurones convolutifs, prédiction de la beauté du visage, apprentissage en profondeur, évaluation des performances.*

## Abstract

Assessment of facial beauty is now possible because of recent advancements in artificial intelligence and computer vision. Facial beauty prediction (FBP) is an important visual recognition problem for the assessment of facial attractiveness, which is consistent with human perception. A deep learning method has recently been demonstrating an amazing ability for feature representation and analysis.

In this work, we studied the prediction of facial beauty, in particular for deep learning methods using convolutional neural networks.

As part of this thesis, we work on the synthesis of recent work based on the prediction of facial beauty based on deep learning techniques and data sets.

The results obtained by these learning techniques in terms of evaluation of PC, MAE and RMSE performance obtained by convolutional neural networks are improved especially when trained with different loss functions. Finally, our results are comparable to those obtained by AlexNet, ResNet-18, ResNeXt-50, CNN – SCA, R3CNN and Semi-supervised.

***Key-words:*** *Convolutional Neural Networks, Facial beauty prediction, Deep learning, Performance Evaluation.*

## ملخص

أصبح تقييم جمال الوجه ممكناً الآن بفضل التطورات الحديثة في الذكاء الاصطناعي ورؤية الكمبيوتر. يعد التنبؤ بجمال الوجه مشكلة مهمة في التعرف على البصر لتقييم جاذبية الوجه، والتي تتوافق مع الإدراك البشري. أظهرت طرق التعلم المتعمق مؤخرًا قدرة مذهلة على تمثيل الميزات وتحليلها.

في هذا العمل، درسنا التنبؤ بجمال الوجه، خاصة بالنسبة لطرق التعلم العميق باستخدام الشبكات العصبية التلافيفية. كجزء من هذه الأطروحة، نعمل على توليف الأعمال الحديثة بناءً على التنبؤ بجمال الوجه بناءً على تقنيات التعلم العميق ومجموعات البيانات.

تم تحسين النتائج التي حصلت عليها تقنيات التعلم هذه من حيث تقييم الأداء الارتباط مع بيرسون، متوسط الخطأ المطلق و الخطأ التربيعي الواسطي التي حصلت عليها الشبكات العصبية التلافيفية خاصة في التدريب مع وظائف فقدان مختلفة.

أخيراً، يمكن مقارنة نتائجنا بتلك التي حصلت عليها AlexNet و ResNet-18 و ResNeXt-50 و CNN – SCA و R<sup>3</sup>CNN و شبه الخاضعة للإشراف.

**المفاتيح :** الشبكات العصبية التلافيفية، تنبؤ جمال الوجه، التعلم العميق، تقييم الأداء.

## Remerciements

*On remercie dieu le tout puissant de nous avoir donné la santé et la volonté d'entamer et de terminer ce mémoire.*

*Tout d'abord, ce travail ne serait pas aussi riche et n'aurait pas pu avoir le jour sans l'aide et l'encadrement de **Mr. BOUKHARI Djamel-Eddine**, on le remercie pour la qualité de son encadrement exceptionnel, pour sa patience, sa rigueur et sa disponibilité durant notre préparation de ce mémoire*

*On remercie tout particulièrement **Dr. BENLAMOUDI Azeddine**, pour son soutien moral et ses encouragements au cours de nos travaux de mémoire.*

*On remercie **Pr. KORICHI Maarouf** pour son aide précieuse et d'avoir accepté de juger ce travail.*

*On remercie **Pr. BETTAYEB Nadjla** qui nous fait l'honneur de présider ce jury.*

## *Dédicaces*

*Avec l'expression de ma reconnaissance, je dédie ce travail*

*A ma mère Aïcha Righi et à mon père Houcine Haddaoui aucune parole ne peut exprimer leurs vraies valeurs de gratitude et d'amour, qui sont les deux personnes les plus chers au monde, que Dieu les protège et prends soin d'eux pour moi.*

*A mes frères : Abdrehman et Abdelaziz et Yacine et à mes tantes, mes oncles, Puisse Dieu vous donne santé, bonheur, courage et surtout réussite*

*A mes chers amis : Amira et Kelthoum, Nour Elhouda, Maroua, Yasmine, Achraf, Mohamed, Pour leurs aides et supports dans les moments difficiles.*

*Je tiens tout d'abord à exprimer ma profonde gratitude et mes sincères remerciements pour votre soutien, votre guidance et votre encouragement tout au long de la réalisation de mon mémoire.*

*Votre présence bienveillante et votre expertise ont été d'une aide inestimable. Votre soutien constant m'a permis de surmonter les défis, de repousser mes limites et de donner le meilleur de moi-même.*

*Que Dieu leur donne une longue et joyeuse vie*

*Anfal Meriem*

## *Dédicaces*

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

*A mon très cher père tu as toujours été pour moi un exemple du père respectueux, honnête, de la personne méticuleuse, je tiens à honorer l'homme que tu es, je voudrais te remercier pour ton amour, ta générosité, ta compréhension... ton soutien fut une lumière dans tout mon parcours. Aucune dédicace ne saurait exprimer l'amour l'estime et respect que j'ai toujours eu pour toi, ce modeste travail est le fruit de tous les sacrifices que tu as déployés pour mon éducation.*

*Je t'aime papa et j'implore le tout-puissant pour qu'il t'accorde une bonne santé et une vie longue et heureuse.*

*A ma chère mère je vous remercie pour tout le soutien et l'amour que vous me portez depuis mon enfance et j'espère que votre bénédiction m'accompagne toujours.*

*Que ce modeste travail soit l'exaucement de vos vœux tant formulés, le fruit de vos innombrables sacrifices. Puisse dieux, le très haut, vous accorder santé, bonheur et longue vie.*

*A mes chères sœurs Fouzia, Farida, Hanene, Samra, Siham et à mon cher frère Salah Eddine pour leurs soutiens tout au long de mon parcours.*

*A mes amies je ne peux pas trouver les mots justes et sincères pour vous exprimer mon affection et mes pensées, vous êtes pour moi des sœurs et des amies sur qui je peux compter. En témoignage de l'amitié qui nous unit et des souvenirs de tous les moments que nous avons passés ensemble, je vous dédie ce travail et je vous souhaite une vie pleine de santé et de bonheur.*

*A tous ceux que j'aime et ceux qui m'aiment merci d'être là pour moi.*

*Yasmína*

# Table Des matières

|   |                                    |
|---|------------------------------------|
| Liste Des Figures .....   | <b>Erreur ! Signet non défini.</b> |
| Liste Des Tables.....   | <b>Erreur ! Signet non défini.</b> |
| Introduction générale .....   | <b>Erreur ! Signet non défini.</b> |
| <b>Abstract</b> .....   | iii                                |
| <b>Chapitre 1 : Etat de l’art des méthodes de prédictions la beauté faciale</b> ..... | <b>Erreur ! Signet non défini.</b> |
| 1 Introduction .....  | 4                                  |
| 2 Beauté du visage.....   | 5                                  |
| 3 Historique sur la beauté du visage.....   | 5                                  |
| 4 Les problèmes et difficultés .....  | 8                                  |
| 4.1 LES PROBLEMES.....  | 8                                  |
| 4.2 DIFFICULTES.....  | 9                                  |
| 5 Les méthodes de prédictions de la beauté faciale .....                              | 10                                 |
| 5.1 L’APPRENTISSAGE PROFOND.....  | 10                                 |
| 5.2 ACQUISITION DE BASE DE DONNEES.....   | 13                                 |
| A. Ensemble de données du FBP .....   | 13                                 |
| B. Ensemble de données Multi-Modality Beauty (M2B).....                               | 13                                 |
| C. Ensemble de données SCUT-FBP .....   | 13                                 |
| D. Ensemble de données SCUT-FBP5500 .....   | 14                                 |
| 6 Conclusion.....   | 15                                 |
| <b>Chapitre 2 : Les réseaux de neurones convolutifs</b> .....                         | <b>Erreur ! Signet non défini.</b> |
| 1 Introduction .....  | 17                                 |
| 2 Définition CNN .....  | 17                                 |
| 3 Historique CNN.....   | 18                                 |
| 4 Architecture CNN .....  | 18                                 |
| 4.1 COUCHE DE CONVOLUTION(CONV).....  | 19                                 |
| 4.2 COUCHE DE POOLING (POOL) .....  | 20                                 |
| 4.3 COUCHES DE CORRECTION (RELU).....   | 21                                 |
| 4.4 COUCHES ENTIEREMENT CONNECTEE (FC).....   | 21                                 |
| 4.5 COUCHES DE PERTE (LOSS) .....   | 22                                 |
| 5 Paramètre des CNN .....   | 22                                 |
| 6 Avantages de CNN.....   | 23                                 |
| 7 Exemples de CNN et transfert d’apprentissage .....                                  | 23                                 |
| 7.1 LENET .....   | 23                                 |
| 7.2 ALEX NET.....   | 24                                 |
| 7.3 VGG.....  | 25                                 |
| 7.4 GOOGLE NET .....  | 25                                 |
| 7.5 RESNET .....  | 26                                 |
| 8 Conclusion.....   | 27                                 |

|  |                                    |
|--|------------------------------------|
| <b>Chapitre 3 : Résultats et Discussions.....</b>              | <b>Erreur ! Signet non défini.</b> |
| 1 Introduction .....   | 29                                 |
| 2 L'évaluation des performances .....                          | 29                                 |
| 2.1 L'ERREUR ABSOLUE MOYENNE MAE .....                         | 29                                 |
| 2.2 LA RACINE DE L'ERREUR QUADRATIQUE MOYENNE (RMSE) .....     | 29                                 |
| 2.3 CORRELATION DE PEARSON .....                               | 30                                 |
| 3 Outils et langage de développement .....                     | 30                                 |
| 3.1 BIBLIOTHEQUES.....   | 31                                 |
| 4 Identification de l'ensemble de données (dataset) :.....     | 32                                 |
| 5 Expérimentations et discussion les résultats obtenus .....   | 33                                 |
| 5.1 EXPERIMENTATIONS .....                                     | 33                                 |
| A. FRACTIONNEMENT DE L'ENSEMBLE D'ENTRAINEMENT ET DE TEST..... | 33                                 |
| B. ÉVALUATION DE REFERENCE .....                               | 34                                 |
| C. DENSENET201.....  | 35                                 |
| 6 Comparaison avec les méthodes actuels.....                   | 39                                 |
| 7 Discussion .....   | <b>Erreur ! Signet non défini.</b> |
| 8 Conclusion générale .....                                    | 41                                 |

# Liste Des Figures

|   |    |
|---|----|
| Figure 1.1 : Ratio d'or.....  | 6  |
| Figure 1.2 : la proportion d'or.....  | 6  |
| Figure 1.3 : Architecture GPNet [48].....   | 10 |
| Figure 1.4 : Un modèle de régression d'ensemble basé sur les DCNN [49].....   | 11 |
| Figure 1.5 : la structure de CNN-ER.....  | 11 |
| Figure 1.6 : l'ensemble de données M2B .....  | 13 |
| Figure 1.7 : l'ensemble de données SCUT-FBP .....   | 14 |
| Figure 1.8 : l'ensemble de données SCUT-FBP 5500 .....  | 14 |
| Figure 1.9 : Distribution de scores l'ensemble de données SCUT-FBP 5500 .....   | 15 |
| Figure 2.1 : Architecture sur un réseau de neurones convolutifs .....   | 19 |
| Figure 2.2 : Exemple de filtre de taille (3*3) .....  | 20 |
| Figure 2.3 : Max pooling.....   | 20 |
| Figure 2.4 : Average pooling. ....  | 21 |
| Figure 2.5 : allure de la fonction ReLU. ....   | 21 |
| Figure 2.6 : Padding de 2 autour d'une image de taille 32×32.....   | 22 |
| Figure 2.7 : Architecture de Le Net.....  | 24 |
| Figure 2.8 Architecture d'Alex Net.....   | 24 |
| Figure 2.9: structure de VGG.....   | 25 |
| Figure 2.10 : Architecture de Google Net .....  | 26 |
| Figure 2.11 : Architecture un bloc de ResNet .....  | 26 |
| Figure 3.1 : Logo du Google Colab .....   | 31 |
| Figure 3.2 : Notebook Jupyter et Python.....  | 31 |
| Figure 3.3 : charger le dataset .....   | 33 |
| Figure 3.4: Images de divers traits du visage et cotes de beauté de l'ensemble de données de référence SCUT-FBP5500. ....                             | 33 |
| Figure 3.5 : La courbe bleue représente les rangs de la vérité terrain. La courbe rouge représente les rangs de prédiction model EfficientNetB7. .... | 38 |
| Figure 3.6 : La relation entre une vérité de terrain et la prédiction model EfficientNetB7. ....  | 39 |
| Figure 3.7 : Comparaison des performances de la validation croisée en cinq fois.....  | 40 |
| Figure 3.8 : Comparaison des performances de différentes méthodes par fractionnement de 60 à 40 % .....   | 40 |

# Liste Des Tables

|  |    |
|--|----|
| Table 3.1 : Comparaison des performances de la validation croisée à 5 fois .....                       | 34 |
| Table 3.2 : Comparaison des performances de 60 % de formation et de 40 % de test.....                  | 34 |
| Table 3.3 : Résultat des performances de différentes méthodes par fractionnement de 80 à 20 % .....    | 38 |
| Table 3.3 : Comparaison des performances de la validation croisée en cinq fois .....                   | 40 |
| Table 3.4 : Comparaison des performances de différentes méthodes par fractionnement de 60 à 40 % ..... | 41 |

# Liste des abréviations

|        |  |
|--------|--|
| MSE    | Mean Square Error.   |
| MAE    | Mean Absolute Error  |
| RMSE   | Root Mean Squared Error                                    |
| PC     | Pearson Correlation  |
| ACC    | Accuracy   |
| FBP    | Facial Beauty Prediction                                   |
| CNN    | Convolutional Neural Network                               |
| DCNN   | Deep Convolutional Neural Network                          |
| MSMFME | Multi Source Manifold Flexible Manifold Embedding          |
| FSCLDE | Feature Selection and Cascaded Deep Discriminant Embedding |
| M2B    | Multi-Modality Beauty                                      |
| MLP    | Multi-Layer Perceptron                                     |
| GPU    | Graphical Processing Unit                                  |
| NFME   | Nonlinear Flexible Manifold Embedding                      |
| AM     | Homme Asiatique  |
| AF     | Femme Asiatique  |
| CM     | Homme Caucasien  |
| CF     | Femme Caucasienne  |
| LSVRC  | Large Sqaule Visual Recognition Challenge                  |
| VGG    | Visual Geometry Group                                      |
| SCA    | Spatial and Channel-wise Attention                         |
| LDL    | Label distribution learning                                |
| ResNet | Residual Neural Network                                    |
| HTML   | Hypertext Markup Language                                  |
| CPU    | Central Processing Unit                                    |
| NumPy  | Numerique Python   |
| SciPy  | Scientific Python  |
| PanDas | Panel Data   |
| LR     | Linear Regression  |
| GR     | Gaussian Regression  |
| SVR    | Support Vector Regression                                  |

# *Introduction générale*

Le visage humain joue un rôle important dans notre vie quotidienne. Un comportement humain instinctif est la quête de la beauté, en particulier la beauté du visage est la nature des êtres humains. Comme la demande de chirurgie esthétique a considérablement augmenté au cours des dernières années, la compréhension de la beauté devient de la plus haute importance pour les milieux médicaux [1]. Il y a 4000 ans d'exploration de la beauté physique humaine [2]. Afin de créer des œuvres belles et harmonieuses, les artistes de l'Égypte ancienne ont défini une série de règles de ratio du corps humain. Les sculpteurs de la Grèce antique et de la Rome antique ont suivi les règles égyptiennes. La beauté physique du visage influence de nombreuses décisions sociales telles que : les choix du partenaire et les décisions d'embauche [1].

L'attractivité faciale est un trait physique hautement souhaitable pour tous les membres de la société [3]. Philosophes, artistes et scientifiques tentent de percer les secrets de la beauté depuis des centaines d'années [4]. En tant que sujet émergent, la prédiction de la beauté du visage reçoit de plus en plus d'attention de la part des chercheurs et des utilisateurs, une question primordiale de reconnaissance faciale et de compréhension faciale [5]. La beauté est considérée dans l'analyse informatisée de la beauté du visage comme un type d'information. Il est lié à la façon dont les gens perçoivent l'attractivité du visage.

Dans le domaine de la psychologie, un certain nombre de théories sur la façon dont les gens observent l'attractivité du visage ont été établies. L'étude de l'attractivité faciale à l'aide d'ordinateurs est un domaine académique récent avec des ressources de recherche limitées considérant que peu d'articles ont été publiés sur ce sujet. La plupart des publications existantes se concentrent sur quelques caractéristiques sporadiques de l'analyse de l'attractivité du visage [6, 7]. Le problème incarne deux points principaux : premièrement, la complexité de la perception humaine et la grande variété des traits du visage rendent difficile la construction de modèles robustes et pratiques pour évaluer la beauté. Deuxièmement, plusieurs bases de données de référence de visage sont principalement configurées pour les problèmes de reconnaissance faciale et sont utiles pour la prédiction de l'attractivité [8]. La majorité des études sur la beauté du visage portent sur la conception de descripteurs de beauté faciale [9]. En effet, la symétrie faciale et les caractéristiques sexuelles secondaires

influencent l'attractivité perçue d'un visage [10]. La plupart des recherches FBP de ces dernières années ont été basées sur l'apprentissage en profondeur [11, 12]. Les réseaux de neurones artificiels profonds constitués de nombreuses couches de fonctions non linéaires qui peuvent être enseignées par rétro propagation sont généralement appelés systèmes d'apprentissage en profondeur. De plus, les développements récents dans l'architecture d'apprentissage en profondeur ont été motivés par la force et l'adaptabilité de ces algorithmes [13, 14]. La renaissance des techniques d'apprentissage en profondeur, en particulier les réseaux de neurones convolutifs (CNN), offre une nouvelle perspective sur le problème FBP. Pour plusieurs applications de vision par ordinateur, telles que la reconnaissance faciale, l'identification d'objets, la segmentation sémantique, la classification d'images, l'analyse biomédicale, le sous-titrage et la biométrie, les DCNN fonctionnent bien mieux [15, 16].

Notre mémoire est organisée selon les chapitres suivants :

- Le premier chapitre est dédié à un État de l'art sur la prédiction de la beauté faciale.
- Le deuxième chapitre commence par une présentation détaillée de l'algorithme d'apprentissage profond utilisant les réseaux de neurones convolutifs appliqué sur les prédictions de la beauté faciale.
- Le troisième chapitre présente les résultats obtenus sur différents datasets. Une analyse des résultats obtenus ainsi qu'une étude comparative de ces résultats.

Nous terminons ce mémoire par une conclusion générale et les perspectives envisagées concernant notre travail.

# ***Chapitre 1 :***

## ***Etat de l'art des méthodes prédictions de la beauté faciale***

### **Résumé**

Nous nous limiterons, dans ce chapitre à la présentation des principales méthodes de prédictions de la beauté faciale proposées dans la littérature. On commence par décrire l'historique de la recherche sur la beauté du visage Ensuite un bref sur les problèmes principaux et difficultés. La grande partie de ce chapitre sera consacrée aux méthodes de prédictions de la beauté faciale basées sur l'apprentissage profond, plus spécifiquement les réseaux de neurones convolutifs.

## 1 Introduction

L'attractivité faciale ou la prédiction de la beauté faciale (FBP) est une étude récente avec de multiples applications potentielles. Il s'agit d'un problème majeur dans la vision informatique en raison de la rareté des bases de données publiques liées à la FBP et de ses essais expérimentaux à petite échelle sur les bases de données. En outre, l'évaluation de la beauté du visage est intrinsèquement individuelle, et les gens ont des préférences personnelles pour la beauté [1]. Les techniques d'apprentissage profond ont montré des capacités remarquables dans l'analyse et la représentation des caractéristiques. Les études précédentes se sont principalement concentrées sur les composants discrets de l'esthétique faciale, avec peu de comparaisons entre les différentes techniques [17].

L'attractivité faciale est un domaine de grand intérêt dans toute la communauté scientifique. Au cours de la période de 20 ans allant de 2003 à 2023, une simple recherche Google Scholar sur l'« attrait facial » a donné lieu à environ 17 800 articles, ce qui représente une augmentation d'environ 14 000. Ces perspectives qui conduisent à une meilleure compréhension de l'histoire et le développement de la beauté humaine Perception et les théories sur lesquelles il est basé. Premièrement, nous passons la Grèce antique. Deuxièmement, nous partons d'un point de vue biologique et non biologique. Nous expliquons la relation entre la valeur mate, la préférence mate et l'attractivité. Enfin, nous allons discuter des traits faciaux qui sont considérés comme des déterminants de l'attrait. Nous élargissons la portée de cette étude : L'attractivité faciale pour les expressions faciales statiques et dynamiques [18].

Il y a de la recherche ces sections décrivent les questions de recherche il y a un dicton populaire que la beauté est dans l'œil du spectateur. Tandis que certains croient que cette phrase est vraie, d'autres croient que la beauté peut être mesurée en utilisant une équation mathématique basée principalement sur le ratio d'or [19]. Alors regardons quel est le ratio d'or du visage, et si le ratio d'or des sports du visage est la meilleure façon de mesurer la beauté et le charme du visage.

## 2 Beauté du visage

La beauté est définie comme « la combinaison de qualités telles que la forme, la couleur ou la forme ravissent les sens esthétiques, en particulier la vue » et les scientifiques depuis des siècles et est intégrée dans la vie quotidienne de la vie sociale [20].

Dion, Berscheid Walster (1972) a mené un test empirique de la « bonté » humaine Belle, et les personnes attrayantes éprouvées sont mieux perçues comme une personne peu attirante [21]. L'attractivité a également un impact important sur le contexte social Les relations, en tant que personnes attrayantes, bénéficient d'un traitement positif dans différents contextes sociaux, plus de succès dans l'accouplement et de meilleures possibilités sur le marché du travail (Langlois et coll., 2000 ; Penton-Voak, 2011, Little et al, 2011). Lors de l'interaction avec les gens, la chose la plus importante est de faire face à un canal important de communication interpersonnelle [22]. Il exprime les sentiments et leur communique des informations émotionnelles. Le visage est la caractéristique la plus importante du corps.

## 3 Historique sur la beauté du visage

Historiquement, la formation des praticiens qui traitent de l'esthétique a été enracinée dans des concepts dépassés tels que le ratio d'or. Ou phi et les canons néoclassiques. Le nombre d'or est d'environ 1,618. Si la distance entre certaines régions est plus proche de 1,618 que d'autres régions définies, alors elle est considérée comme idéale. Golden Ratio Face utilise la symétrie faciale, la structure faciale et le rapport d'or pour calculer la beauté de chaque individu (Voir la figure 1). L'attractivité dépend de zones spécifiques du visage que d'autres aiment, et ces zones sont considérées comme des facteurs attrayants, soulignant l'importance de l'opinion personnelle dans la détermination de l'attractivité. Différentes personnes ont des tendances différentes et des formes de visage différentes. La façon dont nous voyons un visage est holistique. Les différents facteurs du visage que nous voyons sont étroitement liés, (Oxford Dictionary, 2015).

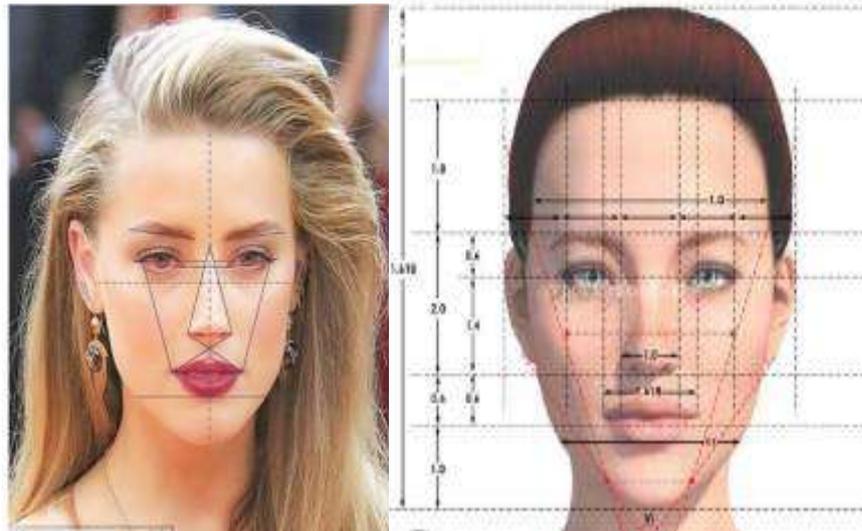


Figure 1.1 : Ratio d'or.

Nous avons vu que tout au long des âges, peintres et sculpteurs ont tenté d'établir des proportions idéales pour la forme humaine. Cependant, le plus célèbre de tous les axiomes sur les proportions « idéales » est probablement celui de la proportion d'or [42].

La proportion d'or est une proportion géométrique dans laquelle une ligne AB est divisée à un point C de telle manière que  $AB/AC = AC/CB$  (Figure 2). Autrement dit, le rapport entre la section la plus courte et la section la plus longue de la ligne est égal au rapport entre la section la plus longue et la ligne entière. Cela donne à  $AC/AB$  la valeur 0.618, appelée le nombre d'or [43]. Le point auquel la ligne est divisée est connu comme la section d'or et est représenté par le symbole  $\Phi$  (phi) dérivé du nom du sculpteur grec Phidias [44]. Phidias a dirigé la construction et la conception du Parthénon, le temple principal de la déesse grecque Athéna sur la colline de l'Acropole à Athènes [45]. On dit que le Parthénon lui-même et les statues qu'il contient se conforment à des proportions « idéales », Phidias incorporant peut-être la proportion d'or dans la conception architecturale, bien que ce ne soit, au mieux, que des conjectures [46, 47].

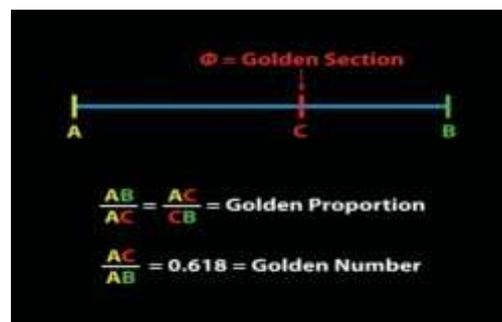


Figure 1.2 : la proportion d'or

Le domaine de la recherche d'attractivité faciale est vaste, y compris les données qui invalident la notion de phi et le canon néoclassique comme des outils qui peuvent être utilisés pour évaluer l'attractivité à travers les races [22]. De plus, la perception esthétique influencée par l'âge, le sexe et l'ethnicité du sujet et du spectateur est prise en compte. Les outils objectifs de classification de l'attractivité sont complexes, de sorte que des systèmes de notation subjectifs et objectifs sont utilisés.

Histoire superficielle, chaque époque semble offrir une interprétation différente du concept de beauté humaine et des vertus qu'elle propose. Certains commentaires sont d'accord, tandis que d'autres sont tout à fait en désaccord [23].

- Saint Thomas Aquinas (1225 – 1274) séparait la beauté physique et la beauté métaphysique, mais croyait que les deux existaient (*Summa Theologiae*) [24]. La beauté du corps est constituée de membres et de traits
- La beauté de l'esprit consiste en des conversations et des actions bien formées et utilisées avec intelligence. »
- L'écrivain Michel de Montaigne (1533 - 1592) : le mouvement intellectuel européen des XVIIe et XVIIIe siècles connus sous le nom des Lumières [25].
- Le philosophe Voltaire (1694 – 1778) a décrit la beauté humaine comme étant culturellement déterminée, sans existence objective, c'est-à-dire que la beauté est dans la « culture » du spectateur. [26]
- David Hume (1711 – 1776) : estime que la beauté n'est pas seulement déterminée sur le plan culturel, mais aussi subjective sur le plan individuel, c'est-à-dire l'idée que « la beauté est dans l'œil du spectateur » [27].
- Immanuel Kant (1724 – 1804) : (dans sa *Critique du jugement* (1790), rejette Hume et retourne à Platon : “Le beau est le symbole du bien moral”. [28]
- Friedrich Schiller (1759 – 1805) était un adepte de Kant ; il estimait que la beauté procurait « du plaisir sans avantage pratique ». [28] Les philosophes et leurs opinions continuaient de s'estomper. [29]
- Ralph Waldo Emerson (1803-1802) a rédigé deux essais intitulés « Beauty » - l'un dans « Nature » (1836) et l'autre dans « The Conduct of Life » (1860). Dans le premier essai, Emerson soutient que la vraie beauté est inhérente à la nature, et que la simple perception des formes naturelles apporte le plaisir [30].

- Subrahmanyam Chandrasekhar (1910-1995) a expliqué que les arts et les sciences poursuivent la même qualité insaisissable - la beauté. Il a défini la beauté comme ce qui suscite la réponse la plus profonde et la plus profonde de l'esprit humain [31].

### 4 Les problèmes et difficultés

Dans l'analyse informatisée de la beauté du visage, la beauté est considérée comme une sorte d'information, telle que l'identité, l'âge et l'expression. Les progrès réalisés dans les domaines de recherche connexes peuvent favoriser l'étude de la beauté du visage. Toutefois, cette dernière présente également des problèmes particuliers. Cette section aborde les principaux problèmes et défis de l'analyse de la beauté faciale par ordinateur. L'analyse informatisée de la beauté du visage [32].

#### 4.1 Les problèmes

Une meilleure compréhension de la manière dont les gens perçoivent l'attrait d'un visage. Dans la discipline de la psychologie, un certain nombre de théories sur la façon dont les gens perçoivent l'attrait d'un visage ont été avancées. Des études de perception sont fréquemment menées pour tester ces hypothèses en utilisant des stimuli produits par des méthodes de déformation d'image et le contenu de l'hypothèse [33].

Cependant, les travaux existants présentent les limites suivantes. Premièrement, le calcul de la forme des visages n'est pas normalisé et est donc affecté par les variations de translation, d'échelle et de rotation. Deuxièmement, les stimuli utilisés dans les expériences de perception existantes ne sont basés que sur un petit nombre de visages [34]. Des expériences basées sur une grande base de données donneraient des résultats plus convaincants. En outre, le traitement d'images et les méthodes d'apprentissage automatique permettent de trouver de nouvelles règles à partir des données, ce qui peut faire progresser notre compréhension de la perception de la beauté des visages [35].

- Prédiction plus précise de la beauté faciale. Le cœur de ce problème est de construire un modèle qui mappe la biométrie faciale à un score de beauté faciale. Jusqu'à présent, il n'y avait pas de base de données publique pour l'étude de la beauté faciale.
- Embellissement du visage. C'est une application typique des règles et des modèles de beauté du visage appris. Les œuvres existantes se concentrent uniquement sur

l'embellissement de la géométrie du visage. Évidemment, la texture du visage joue également un rôle important dans la beauté du visage.

- Élaborer des systèmes d'application. Un avantage de l'analyse de beauté faciale basée sur ordinateur est la commodité de développer des systèmes d'application avec les connaissances apprises.

### 4.2 Difficultés

Selon le contexte de recherche actuel, l'analyse informatisée de la beauté faciale présente les défis suivants :

- L'analyse informatisée de la beauté faciale est un nouveau sujet de recherche qui a peu de ressources de recherche. Le nombre d'ouvrages publiés sur ce sujet est faible. Il n'existe pas de cadre de recherche unifié ni de problèmes reconnus. Les travaux existants se concentrent sur des aspects épars de l'analyse de la beauté du visage, et il y a peu de comparaisons entre différentes méthodes.
- Il n'existe pas de base de données publique pour l'étude de la beauté faciale. Un ensemble de données pour l'étude de la beauté faciale est nécessaire pour avoir une variabilité suffisante de l'attractivité.
- L'objectif de l'analyse de la beauté faciale est différent des autres tâches d'analyse faciale, comme la reconnaissance faciale, l'estimation de l'âge et la reconnaissance de l'expression faciale, qui visent à augmenter le taux de reconnaissance ou à réduire l'erreur des carrés moyens. Lors de la construction de modèles de beauté faciale, la précision de prédiction n'est qu'un objectif. L'intelligibilité et la praticabilité des modèles sont également importantes et nécessitent davantage d'enquêtes.
- Dans l'étude de la beauté faciale, des expériences de perception doivent souvent être menées, ce qui exige beaucoup de travail et de temps. C'est une contrainte de l'étude de la beauté faciale [36].

## 5 Les méthodes de prédictions de la beauté faciale

### 5.1 L'apprentissage profond

#### A. Apprentissage supervisé dans la prédiction de la beauté du visage

La prédiction de la beauté du visage est une tâche difficile et subjective qui a reçu une attention considérable ces dernières années. L'une des approches les plus populaires pour cette tâche est l'apprentissage supervisé, où le modèle est formé sur un ensemble de données étiquetées d'images faciales et leurs scores de beauté correspondants [37, 38, 39].

GPNet géométrique a priori est un réseau hybride proposé par Peng Tianhao et al en 2023 [48]. Ce modèle utilise une structure à deux branches avec une branche CNN locale et une branche mondiale Swin Transformer. Les deux branches sont des modules de fusion de caractéristiques multi-échelles qui utilisent la régularisation géométrique pour améliorer la précision de la prédiction de la beauté du visage. La figure 1.3 illustration de l'architecture GPNet.

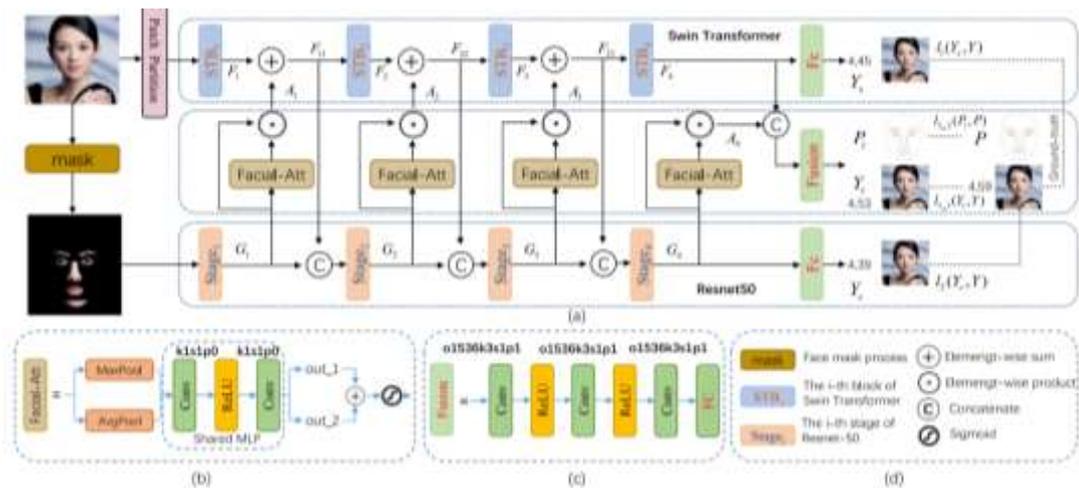


Figure 1.3 : Architecture GPNet [48].

Saïd et al. (2023), [49] ont proposé un modèle de régression d'ensemble basé sur les DCNN pour la prédiction de la beauté du visage. L'architecture du modèle se compose de deux CNN préformés, à savoir AlexNet et VGG16, et d'un réseau entièrement construit à partir de zéro. La figure 1.4 illustre l'architecture de l'ensemble des DCNN.

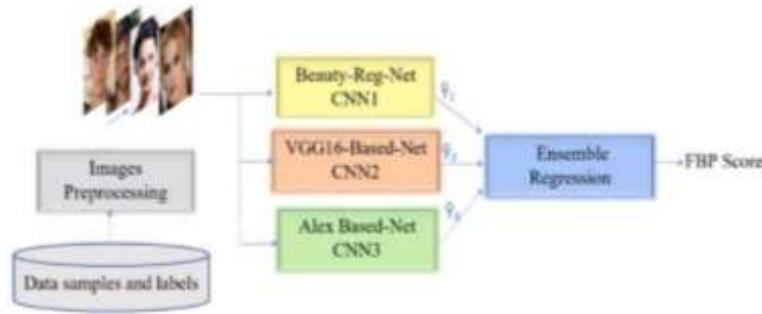


Figure 1.4 : Un modèle de régression d'ensemble basé sur les DCNN [49]

F. Bougourzi et al. (2022) [50] ont proposé un ensemble CNN nommé CNN-ER pour l'estimation de la beauté du visage. Ce modèle utilise deux branches des réseaux ResneXt-50 et Inception-v3 et est formé avec quatre fonctions de perte, à savoir dynamique ParamSmoothL1, dynamique Huber, dynamique Tukey et MSE. La figure 1.5 fournit une illustration de l'architecture CNN-ER.

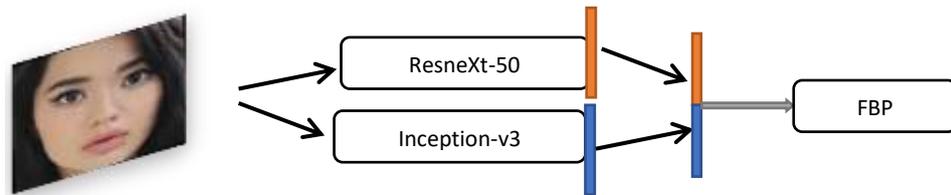


Figure 1.5 : la structure de CNN-ER

Cao et al. [66] ont introduit une architecture de réseau plus profonde utilisant des groupes résiduels dans les résidus (RIR). Les auteurs ont également présenté une base de données sur la beauté du visage, SCUT-FBP5500 [67], qui se compose de 5 500 images faciales avec des annotations pour les scores de beauté. Deux protocoles d'évaluation, validation croisée 5 fois (80%-20%) et fractionnement 60%-40%, ont été utilisés pour tester l'architecture proposée avec trois architectures CNN : Alexnet [54], Resnet-18 [62] et ResneXt -50 [62]. Le mécanisme d'attention combiné spatial et canal a également été introduit pour une meilleure compréhension des caractéristiques, ce qui a permis d'améliorer la compréhension des caractéristiques et de meilleures performances pour FBP.

R3CNN proposé par Lin et al. [68] intègre le classement relatif dans la régression pour améliorer les performances de FBP. Cette architecture peut être mise en œuvre de manière flexible en utilisant les CNN existants comme réseau fédérateur. L'architecture proposée fournit de meilleurs résultats que les jeux de données SCUT-FBP [69] et SCUT-FBP5500 [67].

### **B. Apprentissage semi-supervisé**

L'apprentissage semi-supervisé est un type d'apprentissage automatique qui consiste à former un modèle sur une petite quantité de données étiquetées et une grande quantité de données non étiquetées. Cette technique est de plus en plus appliquée dans le domaine de la prédiction de la beauté du visage, où la disponibilité des données étiquetées est souvent limitée. Dans ce contexte, F. Dornaika et ses collègues ont proposé plusieurs techniques semi-supervisées qui ont montré des résultats prometteurs.

MSMFME une technique est le Multi-Source Manifold Flexible Manifold Embedding (MSMFME), qui a été proposé par Dornaika en 2023 [70]. Cette approche utilise plusieurs graphiques pour créer un modèle d'intégration de variété flexible qui peut être entraîné sur des données étiquetées et non étiquetées. Le modèle a été testé à l'aide d'une validation croisée quintuple sur l'ensemble de données SCUT FBP-5500.

FSCLDE, en 2020, Dornaika et al ont proposé une autre technique semi-supervisée appelée Feature Selection and Cascaded Deep Discriminant Embedding (FSCLDE) [71]. Cette approche utilise une architecture d'extraction et de sélection de caractéristiques en cascade pour transformer les descripteurs bruyants et faibles en descripteurs forts. La structure qui en résulte peut transformer toute approche linéaire en une variation profonde. La technique a été évaluée sur trois ensembles de données de beauté du visage (SCUT-FBP5500, SCUT-FBP et M2B) à l'aide d'un classificateur 1-NN avec un protocole d'évaluation de validation croisée quintuple.

NFME une autre technique proposée par Dornaika et al en 2020 est le Graph-based Nonlinear Flexible Manifold Embedding (NFME) [72]. Cette technique est spécifiquement conçue pour la prédiction de la beauté du visage et est basée sur la texture. Pour gérer le scénario de propagation des scores réels, les auteurs ont modifié et noyauté une technique linéaire existante d'incorporation de collecteur flexible. La technique a été évaluée sur les trois mêmes ensembles de données de beauté du visage à l'aide d'un classificateur 1-NN avec un protocole d'évaluation de validation croisée quintuple.

Dans l'ensemble, ces techniques semi-supervisées ont montré des résultats prometteurs dans la prédiction de la beauté du visage. Des recherches supplémentaires sont nécessaires pour évaluer leur efficacité sur des ensembles de données plus vastes et dans différents contextes.

## 5.2 Acquisition de base de données

C'est une étape cruciale car elle fonctionne comme une base pour les modèles informatiques. Dans la recherche d'attractivité, l'acquisition des données de visage est de façons à partir des bases de données de visage public.

### A. Ensemble de données du FBP

Le choix d'une base de données de taille appropriée est essentiel à la généralisation et à la validation du modèle prédit. Normalement, la base de données devrait comprendre différents genres et visages de tous les niveaux d'attrait, d'âge et d'ethnicité [76]. La plupart des modèles FBP récents sont fondés sur des données, ce qui fait de l'ensemble de données de référence l'un des éléments essentiels pour FBP [19]. Des échantillons d'images de visage générés à partir de diverses bases de données employées dans la classification et la prédiction de la beauté de visage sont montrés.

### B. Ensemble de données Multi-Modality Beauty (M2B)

Il comprend des fichiers audios, des images de pansement et des images faciales de 1240 femmes. Ces images ont été classées en utilisant différents scores dans un intervalle [1-10]. Ils ont été regroupés en deux groupes ethniques : les gens de l'Est et les gens de l'Ouest (chaque groupe comprenait 620 personnes). Les cotes ont été choisies parmi 40 participants.



Figure 1.6 : l'ensemble de données M2B

### C. Ensemble de données SCUT-FBP

Il comprend environ 500 femmes asiatiques d'origines et de classements d'attractivité différents. Ceux-ci avaient été validés en fonction de l'auto-cohérence, de la cohérence, de

l'écart-type et de la répartition des côtes. Les scores de beauté (classements) sont tombés dans un intervalle de [1-5], résultant d'une moyenne de scores divers. Les scores ont été générés à partir de 75 personnes via un dispositif Web, en prenant 70 évaluateurs par image en moyenne.



Figure 1.7 : l'ensemble de données SCUT-FBP

### D. Ensemble de données SCUT-FBP5500

Il comprend 5500 visages, des expressions neutres et des personnes « non voilées » âgées de 15 à 60 ans. Ils ont été classés en quatre sous-groupes avec des genres et des traces variés. Il s'agit de 2000 hommes asiatiques (AM), 2000 femmes asiatiques (AF), 750 hommes caucasiens (CM) et 750 femmes caucasiennes (CF). Chaque image a été étiquetée avec des notes de beauté dans un intervalle de [1-5]. Les notes ont été recueillies parmi 60 participants.



Figure 1.8 : l'ensemble de données SCUT-FBP 5500

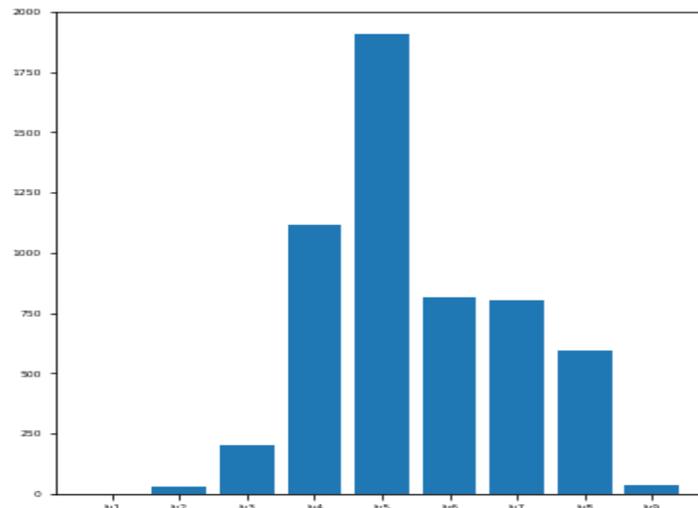


Figure 1.9 : Distribution de scores l'ensemble de données SCUT-FBP 5500

## 6 Conclusion

Selon l'application recherchée prédictions de la beauté faciale basées sur l'apprentissage profond, on peut choisir entre algorithme supervisées ou semi-supervisées. Plusieurs algorithmes pour les deux cas, sont proposés dans la littérature. Cependant, les méthodes basées sur supervisées et le plus utilisé dans la prédiction de la beauté faciale spécifiquement les réseaux de neurones convolutifs. Dans le chapitre suivant en vas détailler sur toutes les applications réseaux de neurones convolutifs, permettent d'avoir des résultats meilleurs par rapport à ceux qui utilisent Les algorithmes d'apprentissage semi-supervisées sont Local and Global Consistency, Flexible Manifold Embedding et Kernel Flexible Manifold Embedding.

# *Chapitre 2 :*

## *Les réseaux de neurones convolutifs*

### **Résumé**

Dans ce chapitre, nous avons détaillé le principe de la méthode réseaux de neurones convolutifs. Une étude des exemples de CNN et transfert d'apprentissage.

Notre travail consiste en l'étude des algorithmes l'apprentissage profond utilisant les réseaux de neurones convolutifs appliquée sur prédictions de la beauté faciale.

### 1 Introduction

La méthode de classification supervisée par les réseaux de neurones multicouches (perceptron) effectue un apprentissage à partir d'un ensemble de données prototypes (exemples des classes) où chaque donnée est décrite par un ensemble d'attributs. La difficulté de choisir les attributs et le nombre de couches, généralement faible (1 à 2 couches), fait que les performances atteintes par le perceptron multicouches sont limitées. D'autres méthodes d'apprentissage profond ont été proposées sont actuellement très en vogue.

La plupart des méthodes d'apprentissage profond utilisent des architectures de réseaux de neurones, ce qui explique pourquoi il est souvent question de réseaux de neurones profonds pour désigner des modèles d'apprentissage profond. Le terme « profond » revient au nombre de couches cachées du réseau de neurones.

Un des types de réseaux de neurones profonds le plus populaire est le réseau de neurones à convolution (Convolutional Neural Network ou CNN). Un CNN utilise des couches à convolution afin d'extraire des caractéristiques directement à partir des données. Son architecture est parfaitement adaptée au traitement des données matricielles telles que les images. Nous allons décrire dans ce chapitre les réseaux de neurones à convolution et leurs applications. La définition du réseau de neurones à convolution (cf. section 2)

### 2 Définition CNN

Les réseaux neuronaux convolutifs sont un type de réseau neuronal artificiel largement utilisé dans les domaines de l'analyse d'images et de vidéos et de la reconnaissance d'objets. Ce type de réseau neuronal est conçu pour traiter des données qui sont généralement bidimensionnelles et utilise de nombreuses couches différentes pour extraire différents profils d'images et déterminer et classer leur contenu.

Les réseaux de neurones convolutionnels (CNN) sont une extension des MLP permettant de répondre efficacement aux principaux défauts des MLP. [51] Ils sont conçus pour extraire automatiquement les caractéristiques des images d'entrée, sont invariants à de légères distorsions de l'image, et implémentent la notion de partage des poids permettant de réduire considérablement le nombre de paramètres du réseau. Ce partage des poids permet en outre de prendre en compte de manière forte les corrélations locales contenues dans une image. [52] Les poids sont forcés à être égaux pour détecter des lignes, des points ou des coins à tous les endroits possibles de l'image, implémentant de fait l'idée du partage des poids. [53]

### 3 Historique CNN

Le premier réseau de neurones convolutif a été en réalité développé en 1998 par le chercheur français Yann LeCun. Ce réseau appelé LeNet a permis d'atteindre de très bonne performance en reconnaissance des caractères, Bien que cette approche donne des résultats, ses progrès et son évolution ont été limités par les progrès technologiques en matière de micro-processeurs, de puissance de calculs, et du manque d'accessibilités à des données afin de pouvoir entraîner les neurones. [54] Cependant certains chercheurs ont continué à travailler sur ce modèle pendant environ deux décennies. [54] Et, avec l'aide des évolutions en matière de technologies mais surtout avec la disponibilité toujours plus grande de données, ont pu améliorer cette technique. [55] Ce n'est qu'en 2012 que l'apprentissage profond est remis au goût du jour en remportant avec succès le concours de reconnaissance d'image fondé par l'université de Stamford (Large Scale Visual Recognition Challenge : ILSVRC) grâce à l'évolution en matière de technologies et à la disponibilité toujours plus grande de données. [56] Un nouvel algorithme d'apprentissage profond explose les records Il s'agit d'un réseau de neurones convolutif appelé AlexNet, largement inspiré du réseau LeNet. ImageNet regroupant 15 000 000 d'images naturelles comportant différents objets et diverses scènes (véhicules, animaux...). [14] Aujourd'hui, les réseaux de neurones convolutifs, sont toujours les modèles les plus performants pour la classification d'images. [57] Google, Microsoft, Facebook, Baidu (le moteur de recherche chinois), Alibaba (site marchand chinois), Nvidia (géant du processeur graphique) ...utilisent les CNN dans leurs applications. [58]

### 4 Architecture CNN

Les réseaux de neurones à convolution sont basés sur le perceptron multicouche (MLP), et inspirés du comportement du cortex visuel des vertébrés. Bien qu'efficaces pour le traitement d'images, les MLP ont beaucoup de mal à gérer des images de grande taille, ce qui est dû à la croissance exponentielle du nombre de connexions avec la taille de l'image. [59] Un réseau de neurone convolutif se compose de plusieurs couches comme le montre la figure 2.1

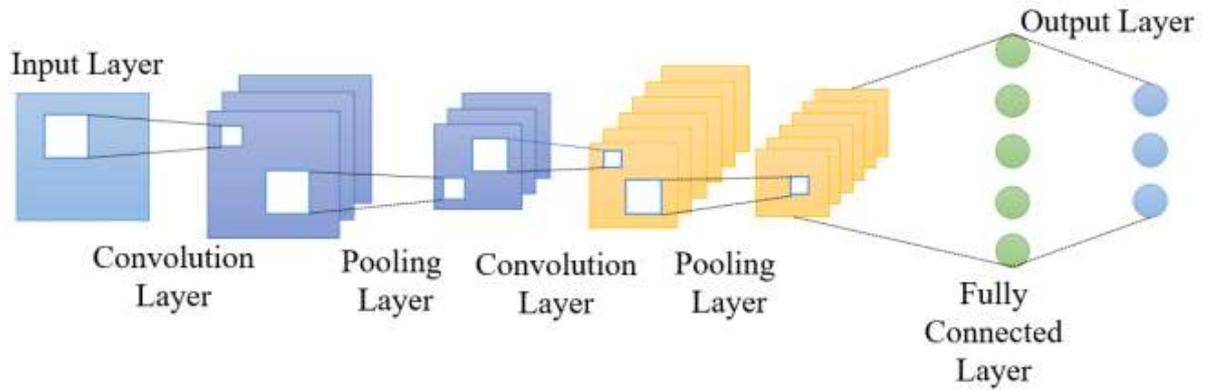


Figure 2.1 : Architecture sur un réseau de neurones convolutifs

Les couches CNN : Une architecture CNN est formée par un empilement de couches de traitement indépendantes [60] :

- La couche de convolution (CONV) qui traite les données d'un champ récepteur.
- La couche de pooling (POOL), qui permet de compresser l'information en réduisant la taille de l'image intermédiaire (souvent par sous-échantillonnage).
- La couche de correction (ReLU), souvent appelée par abus 'ReLU' en référence à la fonction d'activation (Unité de rectification linéaire).
- La couche "entièrement connectée" (FC), qui est une couche de type perceptron.
- La couche de perte (LOSS).

#### 4.1 Couche de convolution (CONV)

La couche de convolution, on applique un produit de convolution entre l'image originale ou celle obtenue dans la couche précédente avec un filtre dont des coefficients sont assimilés à des poids synaptiques. Le produit de convolution d'une matrice (image ou vecteur de mot dans ce cas)  $I$  avec un filtre  $K$  de taille  $(s*s)$  est :

$$O(x, y) = \sum_{i=0}^{s-1} \sum_{j=0}^{s-1} K(i, j) \times I(x - 1 + j, y - 1 + i) \quad (1)$$

Il y' a trois paramètres permettent de dimensionner le volume de la couche de convolution la profondeur, le pas et la marge :

- **Profondeur de la couche** : Nombre de noyaux de convolution (ou nombre de neurones associés à un même champ récepteur). [61]
- **Le pas** : le contrôle le chevauchement des champs récepteurs. Plus le pas est petit, plus les champs récepteurs se chevauchent et plus le volume de sortie sera grand. [62]

- **La marge (à 0) ou zero padding:** Parfois, il est commode de mettre des zéros à la frontière du volume d'entrée. La taille de ce 'zero-padding' est le troisième hyper paramètre. [18] Cette marge permet de contrôler la dimension spatiale du volume de sortie. En particulier, il est parfois souhaitable de conserver la même surface que celle du volume d'entrée. [62]

Cette marge permet de contrôler la dimension spatiale du volume de sortie. En particulier

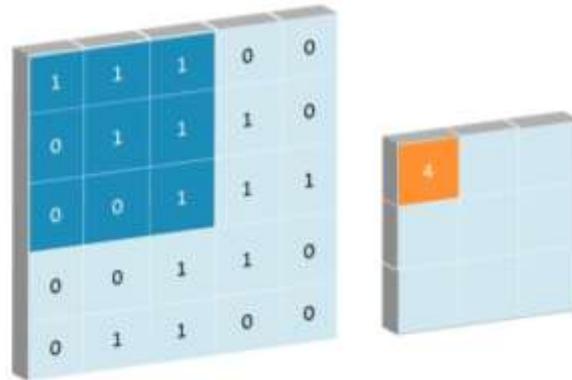


Figure 2.2 : Exemple de filtre de taille (3\*3)

## 4.2 Couche de pooling (POOL)

Après chaque couche convolutive, il peut y avoir une couche de mise en commun. Cette couche est sous-échantillonnée le long de la dimension spatiale d'une entrée donnée, ce qui réduira encore le nombre de paramètres dans l'activation. Il existe plusieurs façons d'effectuer cette mise en commun, comme la prise de valeur moyenne ou maximale, ou la combinaison linéaire de neurones dans le bloc. [62]

Les types de pooling les plus populaires sont le max et l'average pooling.

### 4.2.1 Max pooling

Chaque opération de pooling sélectionne la valeur maximale de la surface. [1]

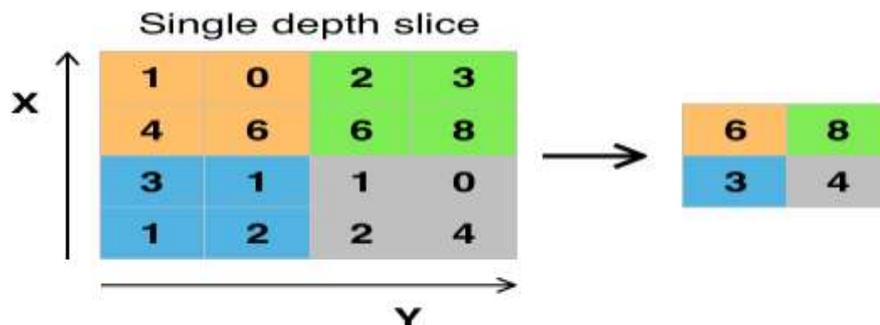


Figure 2.3 : Max pooling.

### 4.2.2 Average pooling

Chaque opération de pooling sélectionne la valeur moyenne de la surface. [1]

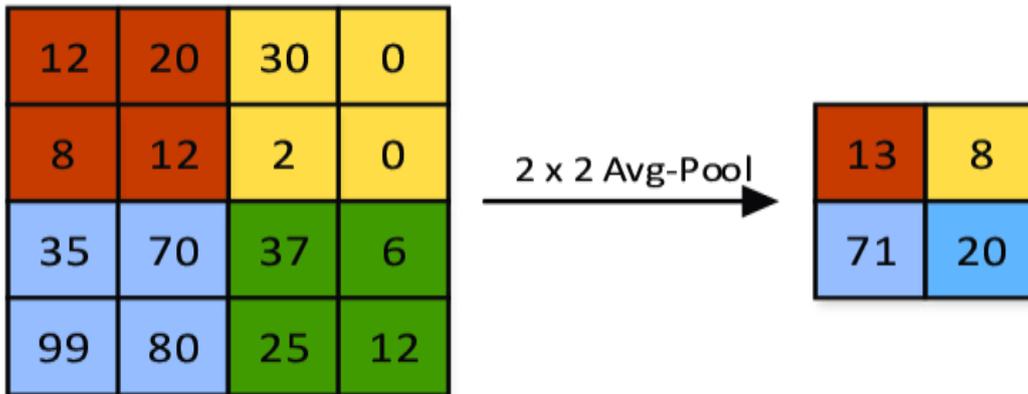


Figure 2.4 : Average pooling.

### 4.3 Couches de correction (RELU)

Il est possible d'améliorer l'efficacité du traitement en intercalant une couche entre les couches de traitement qui effectuera une fonction mathématique (fonction d'activation) sur les signaux de sortie.

La fonction ReLU (abréviation d'Unités Rectifié linéaires) :  $F(x) = \max(0, x)$  Cette fonction force les neurones à renvoyer des valeurs positives. [63]

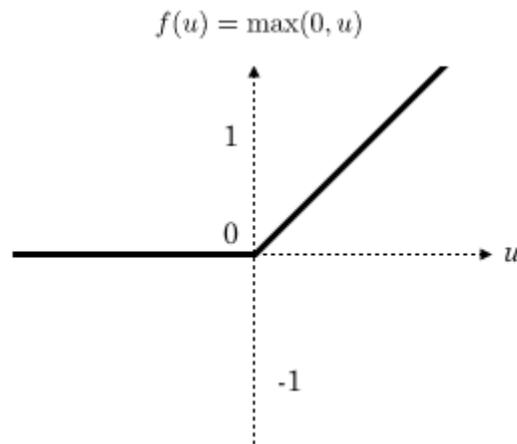


Figure 2.5 : allure de la fonction ReLU.

### 4.4 Couches entièrement connectée (FC)

Après plusieurs couches de convolution et de max-pooling, le raisonnement de haut niveau dans le réseau neuronal se fait via des couches entièrement connectées. Les neurones d'un canapé (module) entièrement connecté sont connectés à toutes les sorties du canapé



## 6 Avantages de CNN

- Les CNNs n'ont pas besoin de supervision humaine pour identifier les caractéristiques importantes.
- Ils sont très précis dans la reconnaissance et la classification des images.
- Le partage du poids est un autre avantage majeur des CNNs.
- Les réseaux neuronaux convolutionnels minimisent également le calcul par rapport à un réseau neuronal régulier. [68]

## 7 Exemples de CNN et transfert d'apprentissage

Il est plus usuel d'utiliser des modèles de CNNs déjà entraînés et de les réadapter pour le problème de classification ou autre, c'est ce que l'on appelle le transfert d'apprentissage. Il s'agit de transférer l'apprentissage d'un modèle traitant un problème vers un autre type de problème. On peut effectuer cette opération de deux manières :

- L'extraction de variables du CNN : ici, le CNN est utilisé comme un extracteur, c'est-à-dire qu'un vecteur est extrait à partir d'une certaine couche du modèle sans rien modifier à sa structure ou son poids et le vecteur précédemment extrait est utilisé pour une nouvelle tâche.
- Le réglage fin du modèle de CNN : ici, le nouveau réseau est initialisé avec les poids et la structure du modèle pré-entraîné à utiliser. La structure du modèle pré-entraîné est légèrement modifiée pour la nouvelle tâche, et enfin le nouveau modèle est entraîné pour la nouvelle tâche.

Il existe un grand nombre de réseaux CNN pré entraînés, les plus connus sont :

### 7.1 LeNet

LeNet [55] un réseau de convolution à 7 niveaux pionnier de LeCun et al en 1998, qui classifie les chiffres, a été appliqué par plusieurs banques pour reconnaître les numéros manuscrits sur les chèques numérisés en 32x32 pixels. La capacité à traiter des images à plus haute résolution nécessite des couches plus convolutives et plus grandes, cette technique est donc limitée par la disponibilité des ressources informatiques. (Fig.2.7)

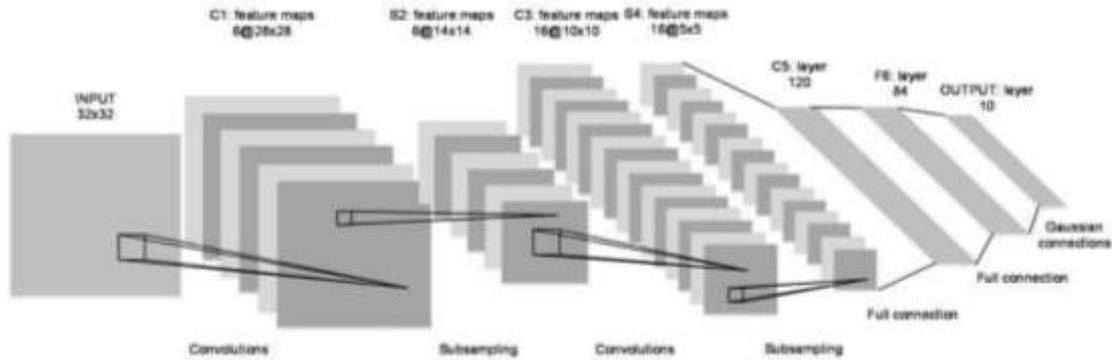


Figure 2.7 : Architecture de Le Net

## 7.2 Alex Net

Alex Net [54] est le nom d'un réseau neuronal convolutif, initialement écrit avec CUDA pour fonctionner avec le support du GPU, qui a participé au défi de reconnaissance visuelle à grande échelle Image Net en 2012. Alex Net a largement surperformé tous ses concurrents précédents et a remporté le défi en réduisant l'erreur de top 5 de 26% à 15,3% il a été conçu par le groupe Supervision, composé d'Alex krizhevsky. Alex Net a eu un grand impact dans le domaine de l'apprentissage automatique, en particulier dans l'application de l'apprentissage en profondeur à la vision artificielle. Alex Net contenait huit couches les cinq premières étaient des couches convolutionnelles et les trois dernières étaient des couches entièrement connectées il a parallèlement deux lignes CNN formées sur deux GPU avec des connexions croisées et Il a utilisé la fonction d'activation non saturante de Relu, qui a montré une amélioration des performances d'entraînement par rapport au Tanh et au sigmoïde. Il a été formé pendant 6 jours simultanément sur deux GPU Nvidia Geforce GTX 580, ce qui explique pourquoi le réseau est divisé en deux pipelines. (fig 2.8)

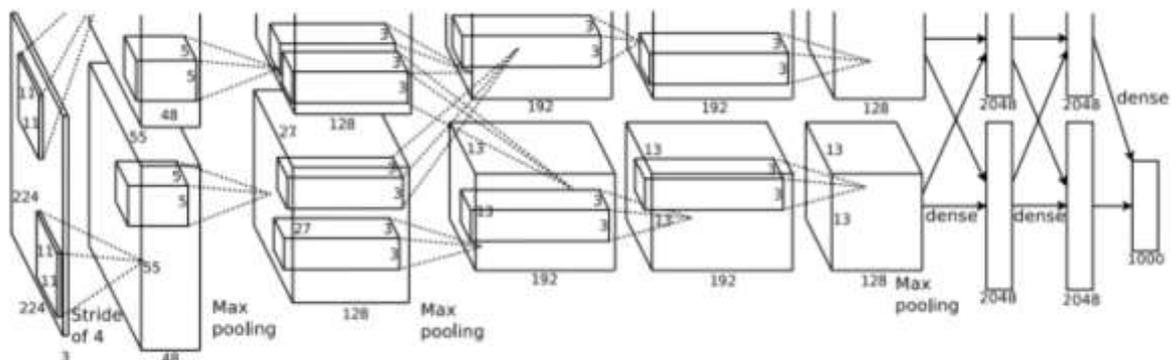


Figure 2.8 Architecture d'Alex Net

### 7.3 VGG

VGG [56] est un réseau neuronal très performant dans le cadre du challenge de reconnaissance visuelle à grande échelle Image Net. Il a obtenu la première place dans la tâche de localisation des images et la deuxième dans la tâche de classification des images. VGG veut donner une réponse à « comment concevoir la structure du réseau » parmi de nombreux choix. VGG montre également que la profondeur du réseau joue un rôle important et les réseaux plus profonds donnent le meilleur résultat. Un inconvénient de VGG est que ce réseau est généralement gros. Il contient environ 160 paramètres. (fig 2.9)

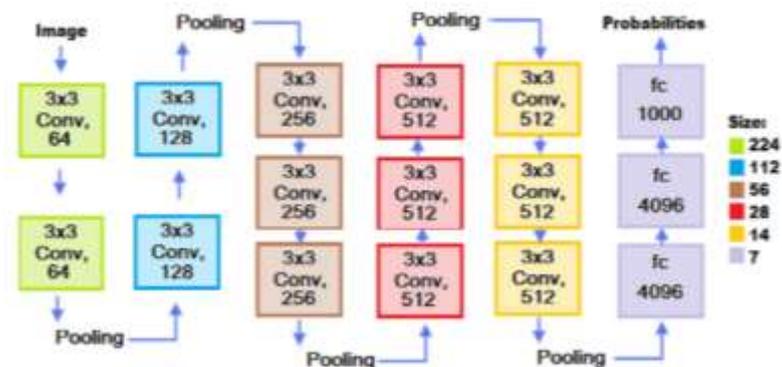


Figure 2.9: structure de VGG

### 7.4 Google Net

Google Net [57] a 22 couches de profondeur, et presque 12 fois moins de paramètres (donc plus rapide et moins que Alex Net et beaucoup plus précis. Il réduit le nombre de paramètres de 60 millions (Alex Net) à 4 millions. Leur idée était de créer un modèle qui pourrait également être utilisé sur un Smartphone (conserver un budget de calcul d'environ 1,5 milliard de multiplications par prévision) Couche de lancement : l'idée de la couche de lancement est de couvrir une plus grande surface, mais aussi de conserver une résolution fine pour les petites informations sur les images. L'idée est donc de convoluer en parallèle différentes tailles, des tailles plus précises.

(1\*1) à un plus grand (5\*5). Le moyen le plus simple d'améliorer les performances en matière d'apprentissage en profondeur consiste à utiliser plus de couches et plus de données. Google Net utilise 9 modules de démarrage. Le problème est que plus de paramètres signifient également que votre modèle est plus enclin à sur-adapter. Ainsi, pour éviter une explosion de paramètres sur les couches initiales. (Fig.2.10)

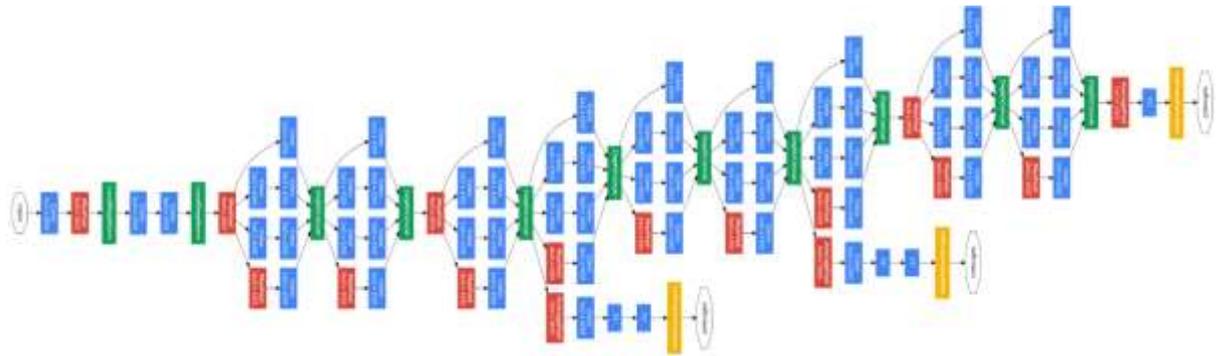


Figure 2.10 : Architecture de Google Net

## 7.5 ResNet

Le réseau Residual Neural Network (ResNet) [62] a introduit une architecture innovante avec des « sauts de connexions » et une lourde normalisation par lots. Ces connexions sont également appelées unités gated ou unités récurrentes gated et ont une forte similitude avec les éléments récents réussis appliqués dans les RNN. Grâce à cette technique, ils ont pu former un NN avec 152 couches tout en ayant une complexité inférieure à celle de VGGNet. Il atteint un taux d'erreur de 3,57% parmi les cinq premiers, ce qui est supérieur à la performance au niveau humain sur cet ensemble de données. (Fig.2.11)

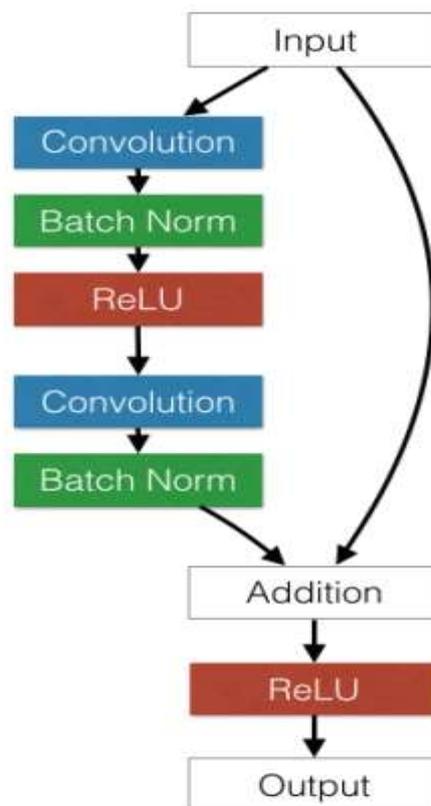


Figure 2.11 : Architecture un bloc de ResNet

## 8 Conclusion

Dans ce chapitre, nous avons étudié le CNN qui est unique dans sa façon de travailler mathématiquement, qui se permet d'être mieux pour la résolution de problèmes spécifiques. En général, CNN tend à être une manière plus puissante et plus précise de résoudre des problèmes de classification et d'alignement. CNN demeure dominant lorsque les ensembles de données sont limités et qu'il n'est pas nécessaire d'entrer des images. CNN est la solution la plus courante pour les problèmes de vision et de machine learning qui dépend de la prédiction de la beauté du visage.

# *Chapitre 3 :*

## *Résultats et Discussions*

### **Résumé**

On commence ce chapitre par un rappel des différents paramètres d'évaluation des performances utilisés dans notre travail (cf. section 2).

Dans expérimentations et discussion les résultats obtenus, pour valider les approches utilisées dans la prédiction de la beauté du visage, nous avons d'abord testé toutes les méthodes en section 5. Enfin, nous avons une comparaison avec les algorithmes actuels de compression la prédiction de la beauté du visage est présentée en section 6.

### 1 Introduction

Après l'étude des réseaux de neurones convolutifs appliquée sur les prédictions de la beauté faciale de ce dernier (cf. chapitre 2). Nous effectuons dans ce chapitre une étude détaillée des résultats obtenus par les algorithmes appliqués sur la prédiction de la beauté du visage fréquemment utilisés dans la littérature. Une étude comparative avec les algorithmes de la prédiction de la beauté du visage actuels comme PI-CNN, CNN avec SCA, CNN + LDL, ResNet-18 based AaNet, R<sup>3</sup>CNN et CNN-ER est ainsi effectuée.

### 2 L'évaluation des performances

En fonction de l'application recherchée, l'algorithme de la prédiction beauté est évalué. On doit pouvoir vérifier un certain nombre de critères de qualité de prédiction, entre autres, on peut citer : l'erreur absolue moyenne (MAE), de l'erreur quadratique moyenne (RMSE) et de la corrélation de Pearson (PC) [79, 80]. À ce stade et après avoir obtenu nos résultats de prédiction, nous devons évaluer les performances de notre système à partir des données de test. Et puisque le modèle appartient aux problèmes de régression on va appliquer les métriques spécifiées à ce type de problèmes [81]. Les métriques utilisées sont :

#### 2.1 L'erreur absolue moyenne MAE

C'est une mesure d'erreur populaire pour les problèmes de régression, elle est simplement définie comme la moyenne de la différence absolue entre la sortie prédite et la sortie réelle [82]. L'erreur quadratique est couramment utilisée car elle est indépendante du fait que la prédiction était trop élevée ou trop faible, elle signale simplement que la prédiction était incorrecte. L'erreur absolue moyenne (MAE) est définie par :

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (2)$$

Où  $y_i$  : la valeur prédite,  $\hat{y}_i$  : la valeur réelle de base and  $N$  : le nombre de toutes les valeurs prédites

#### 2.2 La racine de l'erreur quadratique moyenne (RMSE)

RMSE (Root Mean Squared Error) est une mesure de la précision d'un modèle de régression. Il est couramment utilisé pour évaluer à quel point les prédictions d'un modèle

correspondent aux valeurs réelles [83, 84]. Pour calculer le RMSE, vous devez d'abord obtenir les prédictions du modèle et les valeurs réelles correspondantes pour un ensemble de données. Ensuite, vous effectuez les étapes suivantes :

- 1-Calculez la différence entre chaque prédiction et la valeur réelle correspondante.
- 2-Pour chaque différence, élevez-la au carré.
- 3-Calculez la moyenne des carrés des différences.
- 4-Prenez la racine carrée de la moyenne calculée à l'étape précédente.

Voici la formule pour le calcul du RMSE :

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|} \quad (3)$$

Où  $y_i$  : la valeur prédite,  $\hat{y}_i$  : la valeur réelle de base and  $N$  : le nombre de toutes les valeurs prédites [85].

### 2.3 Corrélation de Pearson

Est une mesure statistique d'ajustement pour tester l'ajustement du modèle, représente la proportion de variance dans le résultat que notre modèle est capable de prédire en fonction de ses caractéristiques. Une valeur de 0 à 1 est considérée comme un pourcentage et plus la valeur est élevée, Corrélation de Pearson (PC) est définie par :

$$Pc = \frac{\sum_{i=1}^N (y_i - \bar{y}_i) (\hat{y}_i - \bar{\hat{y}}_i)}{\sqrt{\sum_{i=1}^N (y_i - \bar{y}_i)^2} \sqrt{\sum_{i=1}^N (\hat{y}_i - \bar{\hat{y}}_i)^2}} \quad (4)$$

Ces paramètres permettent d'estimer l'efficacité d'algorithme et indiquent de meilleures performances obtenues par le système FBP [86].

## 3 Outils et langage de développement

Nous avons réalisé notre système en utilisant :

**Colaboratory** : ("Colab" en abrégé) est un outil d'analyse de données et d'apprentissage automatique qui vous permet de combiner du code Python exécutable et du texte enrichi avec des graphiques, des images, HTML, LaTeX et plus encore dans un seul document stocké dans Google Drive. Il se connecte aux puissants environnements d'exécution de Google Cloud Platform et vous permet de partager facilement votre travail et de collaborer avec d'autres.



Figure 3.1 : Logo du Google Colab

Google Colab est un notebook Jupyter gratuit qui permet d'exécuter Python dans le navigateur sans avoir besoin d'une configuration complexe. Il est livré avec Python installé et toutes les principales bibliothèques Python sont installées. Il est également intégré à des GPU gratuits. Python est un langage idéal pour l'écriture de scripts de façon très simples et le développement rapide d'applications dans de nombreux domaines et sur la plupart des plateformes, tout ça grâce à ses nombreuses bibliothèques.

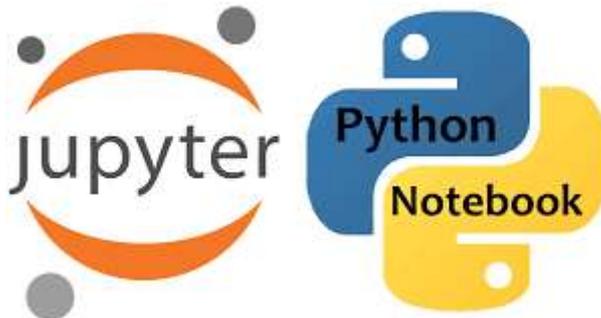


Figure 3.2 : Notebook Jupyter et Python

### 3.1 Bibliothèques

Dans notre programme, on a utilisé un ensemble de librairies pour utiliser des fonctions prédéfinies. Ils sont appelés au début de programme.

**Keras** : est une bibliothèque open source écrite en python qu'est exécutée sur la plateforme d'apprentissage automatique TensorFlow et aussi elle est facile à utiliser pour développer et évaluer des modèles d'apprentissage profond car Elle permet de créer très facilement des layers pour les réseaux de neurones ou de mettre en place des architectures complexes.

**TensorFlow** : est une bibliothèque open source de l'apprentissage automatique, créée par Google (la deuxième génération du système de Google Brain) permettant de développer et d'exécuter des applications de l'apprentissage automatique et profond. Il s'agit d'une boîte à outils permettant de résoudre des problèmes mathématiques extrêmement complexes et le calcul numérique haute performance avec aisance. Son architecture flexible permet de déployer facilement le calcul sur une variété de plateformes (CPU, GPU, TPU).

**Scikit-learn "Sklearn"** : est une importante bibliothèque d'outils dédiés au l'apprentissage automatique et à la science des données dans l'univers Python.

**Numpy** : Le terme NumPy est en fait l'abréviation de Numérique Python. Il s'agit d'une bibliothèque Open Source en langage Python. On utilise cet outil pour la programmation scientifique en Python, et notamment pour la programmation en sciences des données sachant qu'elle propose un grand nombre de routines pour un accès rapide aux ces données. Elle est destinée à manipuler des matrices ou tableaux multidimensionnels (array) ainsi que des fonctions mathématiques opérant sur ces tableaux.

**Matplotlib** : Une bibliothèque Python, destinée à tracer et visualiser des données sous formes de graphiques. Elle peut être combinée avec les bibliothèques python de calcul scientifique NumPy et SciPy et elle est notamment utilisée sur des serveurs d'application web, des shells et des scripts Python.

**Pandas** : Le nom Pandas est en fait la contraction du terme Panel Data, désignant les ensembles de données incluant des observations sur de multiples périodes temporelles. La bibliothèque logicielle open-source Pandas est spécifiquement conçue pour la manipulation et l'analyse de données en langage Python, elle propose en particulier des structures de données et des opérations de manipulation de tableaux numériques et de séries temporelles pour rendre le travail avec les données plus facile.

#### 4 Identification de l'ensemble de données (dataset) :

Pour notre application, nous avons utilisé le dataset SCUT-FBP5500, la figure 3.4 présent des échantillons féminins asiatiques et les scores correspondants sont de droite à gauche : (1,56 ; 2,45 ; 3,51 ; 4,28 ), les échantillons masculins asiatiques et les scores correspondants sont de droite à gauche : (1,53 ; 2,46 ; 3,53 ; 4,23 ), les échantillons féminins caucasiens et les scores correspondants sont de droite à gauche : (1,93 ; 2,45 ; 3,66 ; 4,45 ) et les échantillons Homme Caucasien et les scores correspondants sont de droite à gauche : (1,53 ; 2,66 ; 3,45 ; 4,2).

Voici le code pour télécharger le dataset SCUT-FBP5500

```
!git clone https://github.com/LapTQ/facial_beauty_perception
!git clone https://github.com/LapTQ/mlapplications.git
!pwd

Cloning into 'facial_beauty_perception'...
remote: Enumerating objects: 13309, done.
remote: Counting objects: 100% (2279/2279), done.
remote: Compressing objects: 100% (1790/1790), done.
remote: Total 13309 (delta 495), reused 2266 (delta 486), pack-reused 11030
Receiving objects: 100% (13309/13309), 545.25 MiB | 21.95 MiB/s, done.
Resolving deltas: 100% (505/505), done.
Updating files: 100% (12178/12178), done.
Cloning into 'mlapplications'...
remote: Enumerating objects: 87, done.
remote: Total 87 (delta 0), reused 0 (delta 0), pack-reused 87
Unpacking objects: 100% (87/87), 221.09 MiB | 6.83 MiB/s, done.
Updating files: 100% (15/15), done.
/content
```

Figure 3.3 : charger le dataset

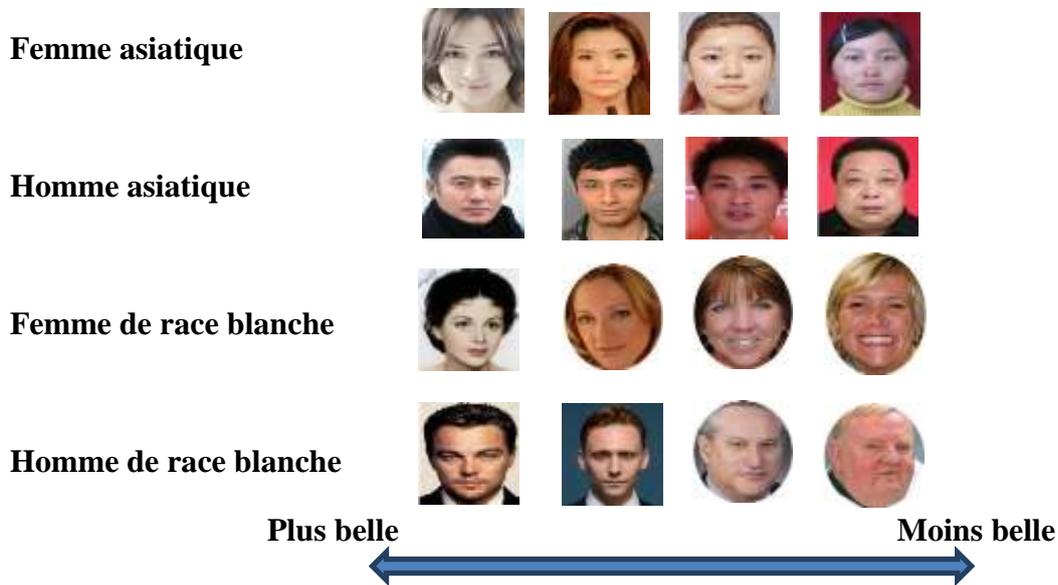


Figure 3.4: Images de divers traits du visage et cotes de beauté de l'ensemble de données de référence SCUT-FBP5500.

## 5 Expérimentations et discussion les résultats obtenus

### 5.1 Expérimentations

#### A. Fractionnement de l'ensemble d'entraînement et de test

Nous utilisons des types de paramètres expérimentaux pour évaluer les méthodes de prédiction de la beauté du visage sur la référence SCUT-FBP5500, sont :

- Validation croisée 5 fois. Pour chaque validation, 80 % des échantillons (4400 images) sont utilisés pour l'entraînement et le reste (1100 images) est utilisé pour les tests.
- La répartition de 60 % de formation et de 40 % de test. 60 % des échantillons (3300 images) sont utilisés pour l'entraînement et le reste (2200 images) est utilisé pour les tests.

### B. Évaluation de référence

Nous définissons ResNet-18, ResNet-50, ResNeXt-50 comme références de l'ensemble de données SCUT-FBP5500, et nous évaluons la référence sur diverses métriques de mesure, notamment : la corrélation de Pearson (PC), l'erreur absolue maximale (MAE) et la racine erreur quadratique moyenne (RMSE). Les résultats de l'évaluation sont présentés ci-après (voir le code dans l'annexe).

| <b>MAE</b>  | <b>1</b>      | <b>2</b>      | <b>3</b>      | <b>4</b>      | <b>5</b>      | <b>Moyenne</b> |
|-------------|---------------|---------------|---------------|---------------|---------------|----------------|
| ResNet-18   | 0.2066        | 0.2125        | <b>0.2494</b> | <b>0.1815</b> | 0.2263        | 0.2153         |
| ResNet-50   | 0.2508        | 0.2242        | 0.2643        | 0.2145        | 0.2715        | 0.2451         |
| ResNeXt-50  | <b>0.1889</b> | <b>0.1823</b> | 0.2542        | 0.2136        | <b>0.2138</b> | <b>0.2106</b>  |
| <b>RMSE</b> | <b>1</b>      | <b>2</b>      | <b>3</b>      | <b>4</b>      | <b>5</b>      | <b>Moyenne</b> |
| ResNet-18   | 0.2576        | 0.2658        | <b>0.306</b>  | <b>0.2411</b> | 0.2925        | 0.2726         |
| ResNet-50   | 0.3047        | 0.2774        | 0.3359        | 0.2604        | 0.3305        | 0.3018         |
| ResNeXt-50  | <b>0.2337</b> | <b>0.2338</b> | 0.3082        | 0.2594        | <b>0.2706</b> | <b>0.2611</b>  |
| <b>PC</b>   | <b>1</b>      | <b>2</b>      | <b>3</b>      | <b>4</b>      | <b>5</b>      | <b>Moyenne</b> |
| ResNet-18   | 0.9184        | 0.9306        | 0.9024        | <b>0.9231</b> | 0.9034        | 0.9156         |
| ResNet-50   | 0.8866        | 0.919         | 0.8743        | 0.9102        | 0.8743        | 0.8929         |
| ResNeXt-50  | <b>0.9289</b> | <b>0.9404</b> | <b>0.903</b>  | 0.9109        | <b>0.9182</b> | <b>0.9203</b>  |

Table 3.1 : Comparaison des performances de la validation croisée à 5 fois

| <b>Méthodes</b>   | <b>PC</b>     | <b>MAE</b>    | <b>RMSE</b>   |
|-------------------|---------------|---------------|---------------|
| <b>ResNet-18</b>  | <b>0.8905</b> | 0.2571        | 0.3514        |
| <b>ResNet-50</b>  | 0.8741        | 0.2528        | 0.3623        |
| <b>ResNeXt-50</b> | 0.8865        | <b>0.2458</b> | <b>0.3471</b> |

Table 3.2 : Comparaison des performances de 60 % de formation et de 40 % de test

### C. DenseNet201

DensNet201\_model

Model: "model\_1"

| Layer (type)  | Output Shape         | Param # | Connected to  |
|---|----------------------|---------|---|
| input_1 (InputLayer)                                | (None, 350, 350, 3)  | 0       | []  |
| zero_padding2d (ZeroPadding2D)                      | (None, 356, 356, 3)  | 0       | ['input_1[0][0]']   |
| conv1/conv (Conv2D)                                 | (None, 175, 175, 64) | 9408    | ['zero_padding2d[0][0]']  |
| conv1/bn (BatchNormalization)                       | (None, 175, 175, 64) | 256     | ['conv1/conv[0][0]']  |
| conv1/relu (Activation)                             | (None, 175, 175, 64) | 0       | ['conv1/bn[0][0]']  |
| zero_padding2d_1 (ZeroPadding2D)                    | (None, 177, 177, 64) | 0       | ['conv1/relu[0][0]']  |
| pool1 (MaxPooling2D)                                | (None, 88, 88, 64)   | 0       | ['zero_padding2d_1[0][0]']  |
| conv2_block1_0_bn (BatchNormalization)              | (None, 88, 88, 64)   | 256     | ['pool1[0][0]']   |
| conv2_block1_0_relu (Activation)                    | (None, 88, 88, 64)   | 0       | ['conv2_block1_0_bn[0][0]']                                       |
| conv2_block1_1_conv (Conv2D)                        | (None, 88, 88, 128)  | 8192    | ['conv2_block1_0_relu[0][0]']                                     |
| conv2_block1_1_bn (BatchNormalization)              | (None, 88, 88, 128)  | 512     | ['conv2_block1_1_conv[0][0]']                                     |
| conv2_block1_1_relu (Activation)                    | (None, 88, 88, 128)  | 0       | ['conv2_block1_1_bn[0][0]']                                       |
| conv2_block1_2_conv (Conv2D)                        | (None, 88, 88, 32)   | 36864   | ['conv2_block1_1_relu[0][0]']                                     |
| conv2_block1_concat (Concatenate)                   | (None, 88, 88, 96)   | 0       | ['pool1[0][0]',<br>'conv2_block1_2_conv[0][0]']                   |
| conv5_block32_2_conv (Conv2D)                       | (None, 11, 11, 32)   | 36864   | ['conv5_block32_1_relu[0][0]']                                    |
| conv5_block32_concat (Concatenate)                  | (None, 11, 11, 1920) | 0       | ['conv5_block32_1_concat[0][0]',<br>'conv5_block32_2_conv[0][0]'] |
| bn (BatchNormalization)                             | (None, 11, 11, 1920) | 7680    | ['conv5_block32_concat[0][0]']                                    |
| relu (Activation)                                   | (None, 11, 11, 1920) | 0       | ['bn[0][0]']  |
| global_average_pooling2d_1 (GlobalAveragePooling2D) | (None, 1920)         | 0       | ['relu[0][0]']  |
| dense_1 (Dense)                                     | (None, 512)          | 983552  | ['global_average_pooling2d_1[0][0]']                              |
| dropout_1 (Dropout)                                 | (None, 512)          | 0       | ['dense_1[0][0]']   |
| final_output (Dense)                                | (None, 5)            | 2565    | ['dropout_1[0][0]']   |

Total params: 19,308,101  
Trainable params: 19,079,045  
Non-trainable params: 229,056

### Programme de DensNet201\_model

```
input_shape=(350,350, 3)
model_input =tf.keras.Input(shape=Input_shape)

Base_model1 =tf.keras.applications.DenseNet201(input_shape=input_shape, input_tensor=model_input, include_top=False, weights="imagenet")
for layer in Base_model1.layers:
    layer.trainable = True

[ ] Base_model1_last_layer = Base_model1.get_layer('relu')
print('last layer output shape:',Base_model1_last_layer.output_shape)
Base_model1_last_output = Base_model1_last_layer.output

last layer output shape: (None, 11, 11, 1920)

[ ] x1 =tf.keras.layers.GlobalAveragePooling2D()(Base_model1_last_output)
x1 =tf.keras.layers.Dense(512, activation='relu')(x1)
x1 =tf.keras.layers.Dropout(0.25)(x1)
final_output1 =tf.keras.layers.Dense(1, activation='softmax', name='final_output')(x1)
DensNet201_model =tf.keras.models.Model(model_input, final_output1)
optimizer = tf.keras.optimizers.legacy.Adam(lr=0.0001, beta_1=0.9, beta_2=0.999, epsilon=None, decay=0.0, amsgrad=True)
DensNet201_model.compile(loss='mean_squared_error',
                        optimizer=optimizer,
                        metrics=['mean_absolute_error'])

/usr/local/lib/python3.10/dist-packages/keras/optimizers/legacy/adam.py:117: UserWarning: The `lr` argument is deprecated, use `learning_rate` instead.
  super().__init__(name, **kwargs)
```

## D. InceptionV3

### InceptionV3\_model

```
=====
====
Total params: 22,852,385
Trainable params: 22,817,953
Non-trainable params: 34,432
```

### Programme de InceptionV3\_model

```
[ ] Base_model2 =tf.keras.applications.InceptionV3(input_shape=input_shape, input_tensor=model_input, include_top=False, weights="imagenet")
for layer in Base_model2.layers:
    layer.trainable = True

[ ] Base_model2_last_layer = Base_model2.get_layer('mixed10')
print('last layer output shape:', Base_model2_last_layer.output_shape)
Base_model2_last_output = Base_model2_last_layer.output

[ ] last layer output shape: (None, 9, 9, 2048)

[ ] x2 =tf.keras.layers.GlobalAveragePooling2D()(Base_model2_last_output)
x2 =tf.keras.layers.Dense(512, activation='relu')(x2)
x2 =tf.keras.layers.Dropout(0.25)(x2)
final_output2 =tf.keras.layers.Dense(1, activation='softmax', name='final_output2')(x2)
InceptionV3_model =tf.keras.models.Model(model_input, final_output2)
optimizer = tf.keras.optimizers.legacy.Adam(lr=0.0001, beta_1=0.9, beta_2=0.999, epsilon=None, decay=0.0, amsgrad=True)
InceptionV3_model.compile(loss='mean_squared_error',
                        optimizer=optimizer,
                        metrics=['mean_absolute_error'])
```

## E. MobileNetV2

### MobileNetV2\_model

```
=====
====
```

Total params: 2,914,369  
Trainable params: 2,880,257  
Non-trainable params: 34,112

### Programme de MobileNetV2\_model

```
Base_model3 =tf.keras.applications.MobileNetV2(input_shape=input_shape, input_tensor=model_input, include_top=False, weights="imagenet")
for layer in Base_model3.layers:
    layer.trainable = True

WARNING:tensorflow: 'input_shape' is undefined or non-square, or 'rows' is not in [96, 128, 160, 192, 224]. Weights for input shape (224, 224)

Base_model3_last_layer = Base_model3.get_layer('out_relu')
print('last layer output shape:', Base_model3_last_layer.output_shape)
Base_model3_last_output = Base_model3_last_layer.output

last layer output shape: (None, 11, 11, 1280)

x3 =tf.keras.layers.GlobalAveragePooling2D()(Base_model3_last_output)
x3 =tf.keras.layers.Dense(512, activation='relu')(x3)
x3 =tf.keras.layers.Dropout(0.25)(x3)
final_output3 =tf.keras.layers.Dense(1, activation='softmax', name='final_output3')(x3)
MobileNetV2_model =tf.keras.models.Model(model_input, final_output3)
optimizer = tf.keras.optimizers.legacy.Adam(lr=0.0001, beta_1=0.9, beta_2=0.999, epsilon=None, decay=0.6, amsgrad=True)
MobileNetV2_model.compile(loss='mean_squared_error',
                          optimizer=optimizer,
                          metrics=['mean_absolute_error'])
```

## F. EfficientNetB7

### EfficientNetB7\_model

```
=====  
=====  
Total params: 3,177,025  
Trainable params: 3,142,913  
Non-trainable params: 34,112
```

### Programme de EfficientNetB7\_model

```
Base_model4 =tf.keras.applications.EfficientNetB7(input_shape=input_shape, input_tensor=model_input, include_top=False, weights="imagenet")
for layer in Base_model4.layers:
    layer.trainable = True

Base_model4_last_layer = Base_model4.get_layer('top_activation')
print('last layer output shape:', Base_model4_last_layer.output_shape)
Base_model4_last_output = Base_model4_last_layer.output

last layer output shape: (None, 11, 11, 2560)

x4 =tf.keras.layers.GlobalAveragePooling2D()(Base_model4_last_output)
x4 =tf.keras.layers.Dense(512, activation='relu')(x3)
x4 =tf.keras.layers.Dropout(0.25)(x4)
final_output4 =tf.keras.layers.Dense(1, activation='softmax', name='final_output4')(x4)
EfficientNetB7_model =tf.keras.models.Model(model_input, final_output4)
optimizer = tf.keras.optimizers.legacy.Adam(lr=0.0001, beta_1=0.9, beta_2=0.999, epsilon=None, decay=0.6, amsgrad=True)
EfficientNetB7_model.compile(loss='mean_squared_error',
                             optimizer=optimizer,
                             metrics=['mean_absolute_error'])
```

## 5.2 Résultat des Model

Le résultat de la prédiction de la beauté du visage à l'aide de la capacité du réseau par comparaison, qui contient 0,8 de l'ensemble de données est utilisé pour l'entraînement, tandis que 0,2 est utilisé pour les tests. Cela signifie que 20 % des instances de l'ensemble de données sont choisies au hasard pour les tests, tandis que les 80 % restants sont choisis au hasard pour l'entraînement dans le tableau 1.

| Méthodes              | PC            | MAE           | RMSE          |
|-----------------------|---------------|---------------|---------------|
| <b>DenseNet201</b>    | 0.8794        | 0.2512        | 0.3315        |
| <b>InceptionV3</b>    | 0.8897        | 0.2456        | 0.3280        |
| <b>MobileNetV2</b>    | 0.8939        | 0.2435        | 0.3178        |
| <b>EfficientNetB7</b> | <b>0.8981</b> | <b>0.2305</b> | <b>0.3056</b> |

Table 3.3 : Résultat des performances de différentes méthodes par fractionnement de 80 à 20 %

De plus, nous analysons la distribution des prédicteurs du réseau, comme le montre la figure 3.5. La courbe montre l'estimation des scores de prédiction. On peut voir qu'il y a un décalage sur la valeur moyenne des prédictions de l'ensemble de données. Cependant, on peut noter que les valeurs de prédiction suivent le comportement général de la distribution de la vérité terrain.

La figure 3.6 illustre la relation entre la vérité terrain et la prédiction. Étant donné que notre objectif est de prédire les scores de prédiction de la beauté du visage, ce modèle estime ses paramètres à l'aide de données provenant de scores de beauté du visage compris dans une certaine plage. On peut en déduire que la vérité-terrain corrèle le plus de valeurs de prédiction.

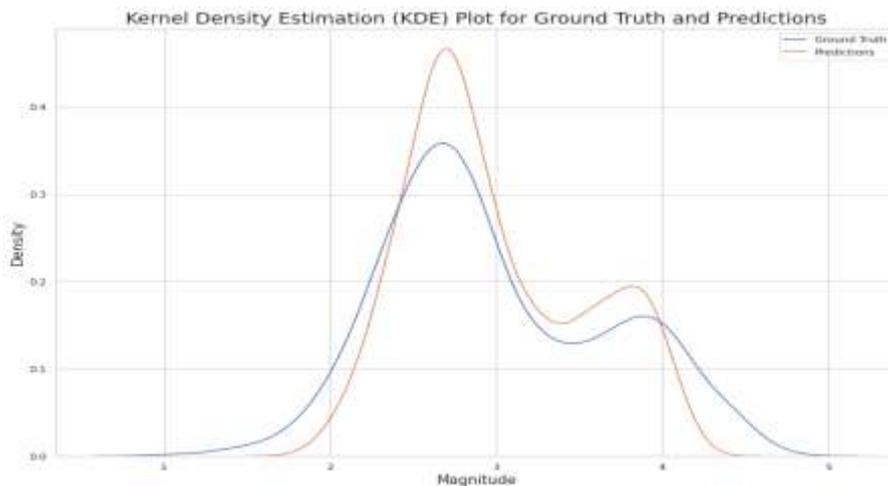


Figure 3.5 : La courbe bleue représente les rangs de la vérité terrain. La courbe rouge représente les rangs de prédiction model EfficientNetB7.

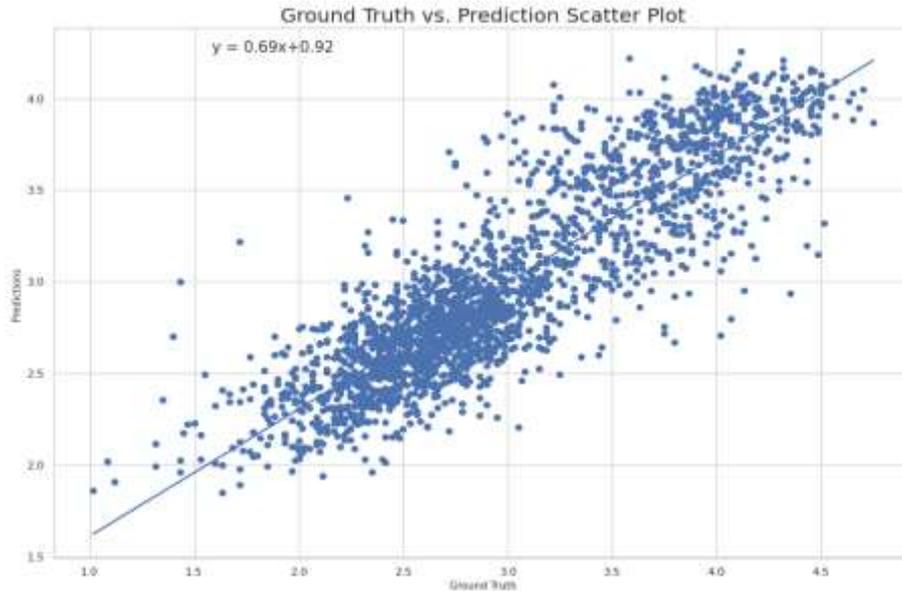


Figure 3.6 : La relation entre une vérité de terrain et la prédiction model EfficientNetB7.

## 6 Comparaison avec les méthodes actuelles

Nous effectuons des comparaisons en utilisant une gamme de techniques, y compris des techniques basées sur les caractéristiques géométriques et sur l'apprentissage en profondeur, telles que LR, GR, SVR, AlexNet, ResNet-18 et ResNeXt-50, etc. MAE, RMSE et PC sont choisis comme les métriques.

La validation croisée quintuple de la prédiction de la beauté du visage à l'aide témoigne de la capacité du réseau via la comparaison, qui contient une division de 80% à 20 % pour chaque pli de la figure 3.5, table 3.5 et 0,6 de l'ensemble de données est utilisé pour l'entraînement, tandis que 0,4 est utilisé pour les tests. Cela signifie que 40 % des instances de l'ensemble de données sont choisies au hasard pour les tests, tandis que les 60 % restants sont sélectionnés au hasard pour l'entraînement dans la figure 3.6 et table 3.4.

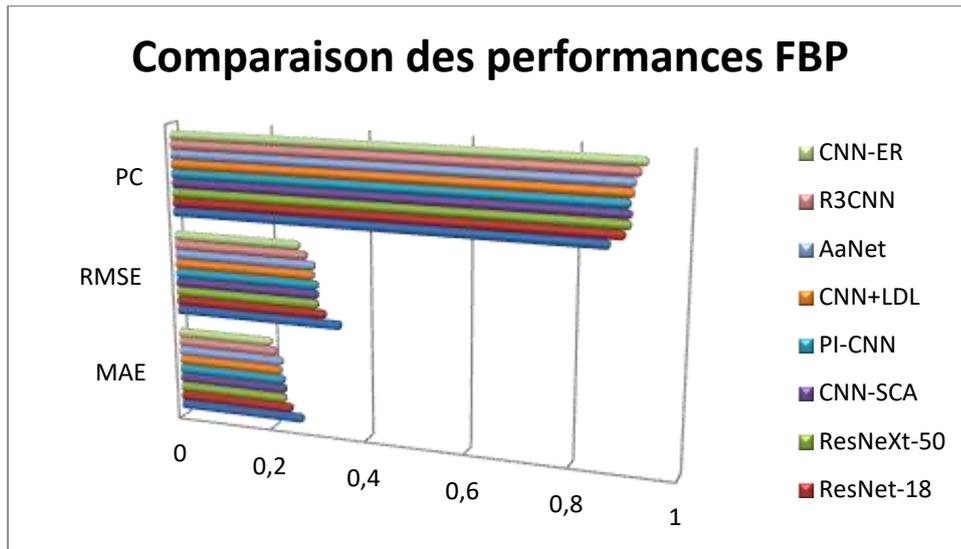


Figure 3.7 : Comparaison des performances de la validation croisée en cinq fois

| Méthodes              | MAE ↓  | RMSE ↓ | PC ↑   |
|-----------------------|--------|--------|--------|
| AlexNet               | 0.2651 | 0.3481 | 0.8634 |
| ResNet-18             | 0.2419 | 0.3166 | 0.89   |
| ResneXt-50            | 0.2291 | 0.3017 | 0.8997 |
| CNN with SCA          | 0.2287 | 0.3014 | 0.9003 |
| PI-CNN                | 0.2267 | 0.3016 | 0.8978 |
| CNN + LDL             | 0.2201 | 0.294  | 0.9031 |
| ResNet-18 based AaNet | 0.2236 | 0.2954 | 0.9055 |
| R <sup>3</sup> CNN    | 0.212  | 0.28   | 0.9142 |
| CNN-ER                | 0.2009 | 0.265  | 0.925  |

Table 3.3 : Comparaison des performances de la validation croisée en cinq fois

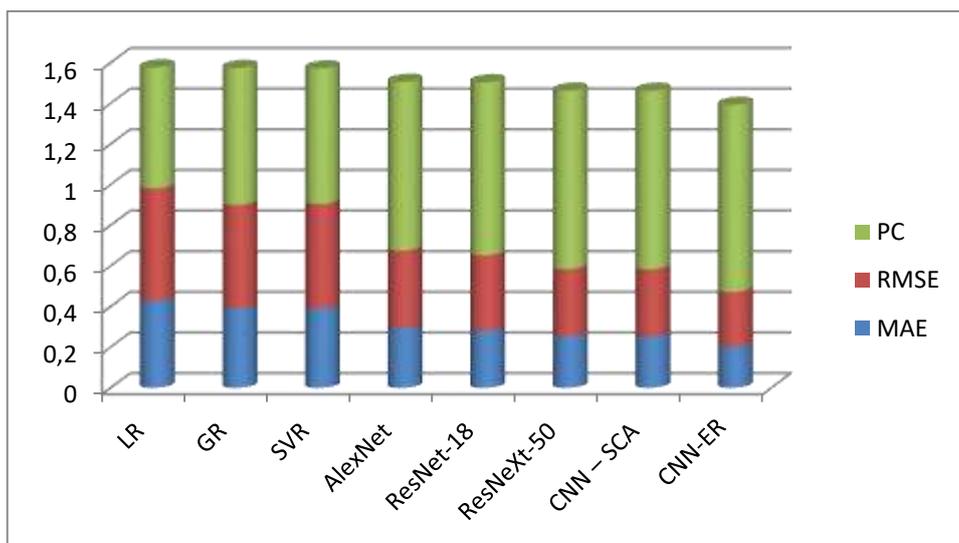


Figure 3.8 : Comparaison des performances de différentes méthodes par fractionnement de 60 à 40 %

| Méthodes   | MAE ↓  | RMSE ↓ | PC ↑   |
|------------|--------|--------|--------|
| LR         | 0.4289 | 0.5531 | 0.5948 |
| GR         | 0.3914 | 0.5085 | 0.6738 |
| SVR        | 0.3898 | 0.5152 | 0.6668 |
| AlexNet    | 0.2938 | 0.3819 | 0.8298 |
| ResNet-18  | 0.2818 | 0.3703 | 0.8513 |
| ResNeXt-50 | 0.2518 | 0.3325 | 0.8777 |
| CNN – SCA  | 0.2517 | 0.332  | 0.878  |
| CNN-ER     | 0.2032 | 0.2683 | 0.9207 |

Table 3.4 : Comparaison des performances de différentes méthodes par fractionnement de 60 à 40 %

## 7 Conclusion

Dans ce travail, nous avons étudié certains modèles de réseaux de neurones convolutifs

En conclusion, les réseaux de neurones convolutifs sont un choix solide pour la prédiction de la beauté faciale basée sur l'apprentissage profond. Leur capacité à capturer des traits faciaux complexes et des performances supérieures dans les problèmes de vision par ordinateur en font une approche préférée.

## *Conclusion générale*

Dans ce mémoire, nous avons examiné de manière approfondie les recherches sur la prédiction de la beauté du visage menées au cours des dernières années, démontrant que les réseaux de neurones à convolution profonde ont fait leurs preuves dans cette tâche.

Dans ce mémoire, nous avons étudié réseaux de neurones convolutifs (CNN) profonds pour la prédiction de la beauté du visage. L'étude de la puissance des approches d'apprentissage par transfert standard (DenseNet201, InceptionV3, MobileNetV2 et EfficientNetB7) sur le problème de prédiction de la beauté du visage, les résultats obtenus par les modèles CNN étudié et entraînés avec des fonctions de perte, nous présentons une régression d'ensemble pour l'estimation de la beauté du visage. Décrire et optimiser un ensemble d'hyper-paramètres pour régressive la beauté du visage. En utilisant toutes les techniques d'apprentissage par transfert mentionnées précédemment, le modèle étudié a été développé pour prédire les scores de beauté du visage. Les résultats expérimentaux indiquent que les réseaux étudiés peuvent fonctionner mieux que les approches de référence CNN précédentes et ils ont montré que le réseau était plus performant que plusieurs études en littérature ouverte disponibles (AlexNet, ResNet-18, ResNeXt-50, CNN - SCA). Il améliore la congruence de l'évaluation avec le jugement humain.

En perspective, nous proposons d'élargir la portée de notre base de données et d'améliorer les réseaux en utilisant différentes architectures collectées à partir de ResNeSt et Inception-ResNet-v2.

# Annexes

## Programme Facial Beauty Prédiction

```
%cd /content
!git clone https://github.com/etrain-xyz/facial-beauty-prediction.git
%cd /content/facial-beauty-prediction
!pip install -r requirements.txt
!pip install gdown
!gdown https://drive.google.com/uc?id=1w0TorBfTIqbquQVd6k3h_77ypnrvfGwf
!unzip -q 'SCUT-FBP5500_v2.1.zip'
```

```
import cv2
from matplotlib import pyplot as plt
import numpy as np
from sklearn.model_selection import train_test_split
import pandas as pd

from keras.layers import Input, Activation, Conv2D, Flatten, Dense, Max
Pooling2D, Dropout, Add, LeakyReLU
from keras.models import Model, load_model
from keras.preprocessing.image import ImageDataGenerator
from keras.callbacks import ModelCheckpoint
import datetime
from tensorflow.keras.optimizers import Adam
import os, sys, glob
plt.style.use('dark_background')
```

### Load dataset

```
labels = pd.read_csv('/content/facial-beauty-prediction/SCUT-
FBP5500_v2/train_test_files/All_labels.txt', sep=" ", header=None)
labels.head()
labels.describe()
labels_np = labels.values
print(labels_np[:5])
```

```
imgs = np.empty((len(labels_np), 350, 350, 3), dtype=np.uint8)
```

```
for i, (img_filename, rating) in enumerate(labels_np):
    img = cv2.imread(os.path.join('/content/facial-beauty-
prediction/SCUT-FBP5500_v2/Images/', img_filename))
```

```
    if img.shape[0] != 350 or img.shape[1] != 350:
        print(img_filename)
```

```
    imgs[i] = img
```

```

x_train, x_val, y_train, y_val = train_test_split(imgs, labels_np[:, 1]
, test_size=0.1)
np.save('x_train.npy', x_train)
np.save('x_val.npy', x_val)
np.save('y_train.npy', y_train)
np.save('y_val.npy', y_val)
print(x_train.shape, x_val.shape)
print(y_train.shape, y_val.shape)

```

```

import
config
    import torchvision
    import torch.nn as nn
    import torch
    def load_model(model_arch):
        model = getattr(torchvision.models, model_arch)(pretrained=True)
        num_ftrs = model.fc.in_features
        model.fc = nn.Linear(in_features=num_ftrs, out_features=1)
        model = model.to(torch.device(config.device))
        return model

```

```

#config.py
import torch
import torchvision.transforms as transforms

model_arch = 'resnet18'
epochs = 50

if torch.cuda.is_available():
    device = 'cuda:0'
else:
    device = 'cpu'

transform = transforms.Compose([
    transforms.Resize(256),
    transforms.CenterCrop(224),
    transforms.ToTensor(),
    transforms.Normalize(mean=[0.485, 0.456, 0.406], std=[0.229, 0.224,
0.225]),])

data_root = './SCUT-FBP5500_v2'

models_dir = './models'
#train.py
...

```

```
if __name__ == '__main__':
    # The split of 60% training and 40% testing
    train_dir = os.path.join(config.data_root, 'train_test_files/split_of_60%training
and 40%testing/train.txt')
    val_dir = os.path.join(config.data_root, 'train_test_files/split_of_60%training and
40%testing/test.txt')
    saved_path = os.path.join(config.models_dir, config.model_arch+'_best_state.pt')
    train(train_dir, val_dir, model_saved_path=saved_path)
```

# Références bibliographiques

- [1] Zhang, David, Fangmei Chen, and Yong Xu. Computer models for facial beauty analysis. Switzerland: Springer International Publishing, 2016.
- [2] Fan, Jintu, et al. "Prediction of facial attractiveness from facial proportions." *Pattern Recognition* 45.6 (2012): 2326-2334.
- [3] Knight, Helen, and Olly Keith. "Ranking facial attractiveness." *The European Journal of Orthodontics* 27.4 (2005): 340-348.
- [4] El Sayed, Abdul Rahman, et al. "3D face detection based on salient features extraction and skin colour detection using data mining." *The Imaging Science Journal* 65.7 (2017): 393-408.
- [5] Bruce, Vicki, and Andy Young. "Understanding face recognition." *British journal of psychology* 77.3 (1986): 305-327.
- [6] Leyvand, Tommer, et al. "Data-driven enhancement of facial attractiveness." *ACM SIGGRAPH 2008 papers*. 2008. 1-9.
- [7] Jain, Anil K., and Stan Z. Li. *Handbook of face recognition*. Vol. 1. New York: springer, 2011.
- [8] Saeed, Jwan, and Adnan Mohsin Abdulazeez. "Facial beauty prediction and analysis based on deep convolutional neural network: a review." *Journal of Soft Computing and Data Mining* 2.1 (2021): 1-12.
- [9] Yan, Haibin. "Cost-sensitive ordinal regression for fully automatic facial beauty assessment." *Neurocomputing* 129 (2014): 334-342.
- [10] Little, Anthony C., et al. "Self-perceived attractiveness influences human female preferences for sexual dimorphism and symmetry in male faces." *Proceedings of the Royal Society of London. Series B: Biological Sciences* 268.1462 (2001): 39-44.
- [11] Kagian, Amit, et al. "A machine learning predictor of facial attractiveness revealing human-like psychophysical biases." *Vision research* 48.2 (2008): 235-243.
- [12] Luthfi, Muhammad, et al. "Mobile Device Facial Beauty Prediction using Convolutional Neural Network as Makeup Reference." *2022 International Conference on Computer Engineering, Network, and Intelligent Multimedia (CENIM)*. IEEE, 2022.

- [13] Vakhshiteh, Fatemeh, Farshad Almasganj, and Ahmad Nickabadi. "Lip-reading via deep neural networks using hybrid visual features." *Image Analysis & Stereology* 37.2 (2018): 159-171.
- [14] Nagpal, Chaitanya, and Shiv Ram Dubey. "A performance evaluation of convolutional neural networks for face anti spoofing." 2019 International Joint Conference on Neural Networks (IJCNN). IEEE, 2019.
- [15] Xu, Jie, et al. "A new humanlike facial attractiveness predictor with cascaded fine-tuning deep learning model." arXiv preprint arXiv:1511.02465 (2015).
- [16] Gan, Junying, et al. "Deep self-taught learning for facial beauty prediction." *Neurocomputing* 144 (2014): 295-303.
- [17] Wu, Xiang, et al. "A light CNN for deep face representation with noisy labels." *IEEE Transactions on Information Forensics and Security* 13.11 (2018): 2884-2896.
- [18] Ghorbanzadeh, Omid, et al. "Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection." *Remote Sensing* 11.2 (2019): 196.
- [19] Albawi, Saad, Tareq Abed Mohammed, and Saad Al-Zawi. "Understanding of a convolutional neural network." 2017 international conference on engineering and technology (ICET). Ieee, 2017.
- [20] Laurentini, Aldo, and Andrea Bottino. "Computer analysis of face beauty: A survey." *Computer Vision and Image Understanding* 125 (2014): 184-199.
- [21] Bottino, Andrea, and Aldo Laurentini. "The analysis of facial beauty: an emerging area of research in pattern analysis." *Image Analysis and Recognition: 7th International Conference, ICIAR 2010, Póvoa de Varzim, Portugal, June 21-23, 2010. Proceedings, Part I 7*. Springer Berlin Heidelberg, 2010.
- [22] Johnston, Victor S., and Melissa Franklin. "Is beauty in the eye of the beholder?." *Ethology and Sociobiology* 14.3 (1993): 183-199.
- [23] Cunningham, Michael R., et al. "Their ideas of beauty are, on the whole, the same as ours: Consistency and variability in the cross-cultural perception of female physical attractiveness." *Journal of personality and social psychology* 68.2 (1995): 261.
- [24] Gaiger, Jason. "The aesthetics of Kant and Hegel." *A companion to art theory* (2002): 127-138.
- [25] Dickerson, Adam B. "IMMANUEL KANT 1724–1804." *Fifty major thinkers on education*. Routledge, 2002. 60-64.

- [26] Burnham, Douglas. *An introduction to Kant's critique of judgement*. Edinburgh University Press, 2019.
- [27] Wicks, Robert. *Routledge philosophy guidebook to Kant on judgment*. Routledge, 2007.
- [28] Cooper, David E. "Immanuel Kant (1724–1804): German Philosopher." *Key Writers on Art: From Antiquity to the Nineteenth Century*. Routledge, 2005. 121-125.
- [29] Luo, Wei. "Aching for the altered body: Beauty economy and Chinese women's consumption of cosmetic surgery." *Women's Studies International Forum*. Vol. 38. Pergamon, 2013.
- [30] Marlowe, Cynthia M., Sandra L. Schneider, and Carnot E. Nelson. "Gender and attractiveness biases in hiring decisions: Are more experienced managers less biased?." *Journal of applied psychology* 81.1 (1996): 11.
- [31] Vegter, Florine, and J. Joris Hage. "Clinical anthropometry and canons of the face in historical perspective." *Plastic and reconstructive surgery* 106.5 (2000): 1090-1096.
- [32] Bashour, Mounir. "History and current concepts in the analysis of facial attractiveness." *Plastic and reconstructive surgery* 118.3 (2006): 741-756.
- [33] Farkas, Leslie G., et al. "Vertical and horizontal proportions of the face in young adult North American Caucasians: revision of neoclassical canons." *Plastic and Reconstructive Surgery* 75.3 (1985): 328-338.
- [34] Farkas, L. G., J. C. Kolar, and I. R. Munro. "Geography of the nose in an attractive face: a morphometric study." *Madrid: International Society of Aesthetic Plastic Surgery* (1985): 36.
- [35] Farkas, L. G., C. R. Forrest, and L. Litsas. "Revision of neoclassical facial canons in young adult Afro-Americans." *Aesthetic Plastic Surgery* 24 (2000): 179-184.
- [36] Thidar, Aye Mya, et al. "Assessing facial beauty of Sabah ethnic groups using Farkas principles." *Health Sciences Research* 3.1 (2016): 1-9.
- [37] Atiyeh, B. S., and S. N. Hayek. "Numeric expression of aesthetics and beauty." *Aesthetic plastic surgery* 32 (2008): 209-216.
- [38] Farkas, Leslie G., et al. "Anthropometric proportions in the upper lip-lower lip-chin area of the lower face in young white adults." *American journal of orthodontics* 86.1 (1984): 52-60.
- [39] Edler, R. J. "Background considerations to facial aesthetics." *Journal of orthodontics* (2014).
- [40] Dunlap, Richard A. *The golden ratio and Fibonacci numbers*. World Scientific, 1997.

- [41] Burkert, Walter. Lore and science in ancient Pythagoreanism. Harvard University Press, 1972.
- [42] Meisner, Gary B. The golden ratio: The divine beauty of mathematics. Race Point Publishing, 2018.
- [43] Kepler, Johannes. The harmony of the world. Vol. 209. American Philosophical Society, 1997.
- [44] Huntley, Herbert Edwin. The divine proportion. Courier Corporation, 2012.
- [45] Akhtaruzzaman, Md, and Amir A. Shafie. "Geometrical substantiation of Phi, the golden ratio and the baroque of nature, architecture, design and engineering." *International Journal of Arts* 1.1 (2011): 1-22.
- [46] Prokopakis, Emmanuel P., et al. "The golden ratio in facial symmetry." *Rhinology* 51.1 (2013): 18-21.
- [47] Thapa, Gyan Bahadur, and Rena Thapa. "The relation of Golden Ratio, mathematics and aesthetics." *Journal of the Institute of Engineering* 14.1 (2018): 188-199.
- [48] Peng, Tianhao, et al. "Geometric prior guided hybrid deep neural network for facial beauty analysis." *CAAI Transactions on Intelligence Technology* (2023).
- [49] Saeed, Jwan Najeeb, Adnan Mohsin Abdulazeez, and Dheyaa Ahmed Ibrahim. "An Ensemble DCNNs-Based Regression Model for Automatic Facial Beauty Prediction and Analyzation." *Traitement du Signal* 40.1 (2023): 55-63.
- [50] Bougourzi, Fares, Fadi Dornaika, and Abdelmalik Taleb-Ahmed. "Deep learning based face beauty prediction via dynamic robust losses and ensemble regression." *Knowledge-Based Systems* 242 (2022): 108246.
- [51] Taigman, Yaniv, et al. "Deepface: Closing the gap to human-level performance in face verification." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2014.
- [52] Aamir, Muhammad, et al. "ML-DCNNNet: multi-level deep convolutional neural network for facial expression recognition and intensity estimation." *Arabian Journal for Science and Engineering* 45.12 (2020): 10605-10620.
- [53] Tan, Mingxing, and Quoc V. Le. "Mixconv: Mixed depthwise convolutional kernels." *arXiv preprint arXiv:1907.09595* (2019).
- [54] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

- [55] LeCun, Yann, et al. "Gradient-based learning applied to document recognition." *Proceedings of the IEEE* 86.11 (1998): 2278-2324
- [56] Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." *arXiv preprint arXiv:1409.1556* (2014).
- [57] Szegedy, Christian, et al. "Going deeper with convolutions." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015.
- [58] Szegedy, Christian, et al. "Inception-v4, inception-resnet and the impact of residual connections on learning." *Thirty-first AAAI conference on artificial intelligence*. 2017.
- [59] Ioffe, Sergey, and Christian Szegedy. "Batch normalization: Accelerating deep network training by reducing internal covariate shift." *International conference on machine learning*. pmlr, 2015.
- [60] Szegedy, Christian, et al. "Rethinking the inception architecture for computer vision." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
- [61] Chollet, François. "Xception: Deep learning with depthwise separable convolutions." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.
- [62] He, Kaiming, et al. "Deep residual learning for image recognition." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
- [63] Hinton, Geoffrey E., Simon Osindero, and Yee-Whye Teh. "A fast learning algorithm for deep belief nets." *Neural computation* 18.7 (2006): 1527-1554.
- [64] Byeon, Wonmin, et al. "Scene labeling with lstm recurrent neural networks." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015.
- [65] Xie, Saining, et al. "Aggregated residual transformations for deep neural networks." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.
- [66] Cao K, Choi KN, Jung H, Duan L. Deep learning for facial beauty prediction. *Information*. 2020 Aug 10;11(8):391.
- [67] Liang, Lingyu, et al. "SCUT-FBP5500: A diverse benchmark dataset for multi-paradigm facial beauty prediction." *2018 24th International conference on pattern recognition (ICPR)*. IEEE, 2018.
- [68] Lin, LuoJun, Lingyu Liang, and Lianwen Jin. "Regression Guided by Relative Ranking Using Convolutional Neural Network (R3CNN) for Facial Beauty Prediction." *IEEE Transactions on Affective Computing* 13.1 (2019): 122-134.

- [69] Xie, Duorui, et al. "Scut-fbp: A benchmark dataset for facial beauty perception." 2015 IEEE International Conference on Systems, Man, and Cybernetics. IEEE, 2015.
- [70] Dornaika, F. "Multi-similarity semi-supervised manifold embedding for facial attractiveness scoring." *Soft Computing* (2023): 1-10.
- [71] Dornaika F, Moujahid A, Wang K, Feng X. Efficient deep discriminant embedding: Application to face beauty prediction and classification. *Engineering Applications of Artificial Intelligence*. 2020 Oct 1;95:103831.
- [72] Dornaika F, Wang K, Arganda-Carreras I, Elorza A, Moujahid A. Toward graph-based semi-supervised face beauty prediction. *Expert Systems with Applications*. 2020 Mar 15;142:112990.
- [73] Laurinavičius, Donatas, Rytis Maskeliūnas, and Robertas Damaševičius. "Improvement of Facial Beauty Prediction Using Artificial Human Faces Generated by Generative Adversarial Network." *Cognitive Computation* (2023): 1-18.
- [74] Yang CT, Wang YC, Lo LJ, Chiang WC, Kuang SK, Lin HH. Implementation of an Attention Mechanism Model for Facial Beauty Assessment Using Transfer Learning. *Diagnostics*. 2023 Mar 29;13(7):1291.
- [75] Lebedeva, Irina, Fangli Ying, and Yi Guo. "Personalized facial beauty assessment: a meta-learning approach." *The Visual Computer* 39.3 (2023): 1095-1107.
- [76] Dornaika, Fadi, and Abdelmalik Moujahid. "Multi-view graph fusion for semi-supervised learning: application to image-based face beauty prediction." *Algorithms* 15.6 (2022): 207.
- [77] Gan, Junying, et al. "Facial beauty prediction fusing transfer learning and broad learning system." *Soft Computing* (2022): 1-14.
- [78] J Iyer, Tharun, et al. "Machine learning-based facial beauty prediction and analysis of frontal facial images using facial landmarks and traditional image descriptors." *Computational Intelligence and Neuroscience* 2021 (2021).
- [79] Wei, Wei, et al. "Assessing facial symmetry and attractiveness using augmented reality." *Pattern Analysis and Applications* (2021): 1-17.
- [80] Xu J. Mt-resnet: a multi-task deep network for facial attractiveness prediction. In 2021 2nd International Conference on Computing and Data Science (CDS) 2021 Jan 28 (pp. 44-48). IEEE.
- [81] Zhai Y, Yu C, Qin C, Zhou W, Ke Q, Gan J, Labati RD, Piuri V, Scotti F. Facial beauty prediction via local feature fusion and broad learning system. *IEEE Access*. 2020 Oct 20;8:218444-57.

- [82] Gan J, Jiang K, Tan H, He G. Facial beauty prediction based on lighted deep convolution neural network with feature extraction strengthened. Chinese Journal of Electronics. 2020 Mar;29(2):312-21.
- [83] Lin, LuoJun, et al. "Attribute-Aware Convolutional Neural Networks for Facial Beauty Prediction." IJCAI. 2019.
- [84] Zhai, Yikui, et al. "BeautyNet: Joint multiscale CNN and transfer learning method for unconstrained facial beauty prediction." Computational intelligence and neuroscience 2019 (2019).
- [85] Xu, Lu, Jinhai Xiang, and Xiaohui Yuan. "CRNet: classification and regression neural network for facial beauty prediction." Advances in Multimedia Information Processing–PCM 2018: 19th Pacific-Rim Conference on Multimedia, Hefei, China, September 21-22, 2018, Proceedings, Part III. Cham: Springer International Publishing, 2018.
- [86] Xu, Jie, et al. "Facial attractiveness prediction using psychologically inspired convolutional neural network (PI-CNN)." 2017 IEEE international conference on acoustics, speech and signal processing (ICASSP). IEEE, 2017.