Republic of Iraq

Ministry of Higher Education and Scientific Research

University of Diyala

College of Science

Gender and Age Classification Based on Face Images by Using Deep Neural Networks

A Thesis
Submitted to the Department of Computer Science\ College of Science\ University of Diyala in a Partial Fulfillment of the Requirements for the Degree of Master in Computer Science

By

Omar Amer Mohammed

Supervised By

Dr. Jamal Mustafa Al-Tuwaijari

Assistant Professor

2022 A.C. **1444 A.H.**

بسم الله الرحمن الرحيم

وَيَسْأَلُونَكَ عَنِ الرُّوحِ
قُلِ الرُّوحُ مِنْ أَمْرِ رَبِّي
وَمَا أُوتِيتُم مِنَ الْعِلْمِ إِلَّا قَلِيلًا

صدق الله العظيم

سورة الاسراء. اية (58)

Acknowledgment

First of all, praise is to Allah, the lord of the whole creation, on all the blessing was the help in achieving this thesis to its end.

I wish to express my thanks to my college (college of science), my supervisor, Ass- Prof. Dr. Jamal Mustafa Al-Tuwaijari for supervising this research and for the generosity, patience and continuous guidance throughout the work. It has been my good fortune to have the advice and guidance from him. My thanks to the academic and administrative staff at the Department of the computer sciences.

I would like to thank my family who has supported me throughout and encouraging me to further my studies and help to complete this project, and I wish to express my thanks to my friends. Especially the special Farah.

Omar Amer Mohammed

Abstract

Age estimation and human gender determination from visual data (like face images) are still have been fascinating research topics due to the variety of potential used applications. The most important applications include developing intelligent human-machine interfaces and improving safety and security in various domains such as transportation, security, medical, etc. However, the diversity of characteristics like facial plastic surgery, facial hair, wrinkles, skin condition, different human races, and external factors are affected the face appearance which is crucial for gender determination and age estimation. These characteristics make these areas a challenging topic and are regarded as fertile ground for researchers. However, the deep neural networks represent the most important techniques that have recently been utilized to determine human gender and age estimation.

In this thesis, an efficient system for gender determination and human age estimation from face images based on external appearance using deep neural network technique (Multitasking Convolutional Neural Network Algorithm (MCNN)) is proposed. The general construction of the proposed system consists of several stages; Firstly, the stage of image acquisition; Secondly, the stage of pre-processing (using the region of interest technique, image resizing, data augmentation); Thirdly, the feature extraction stage; Finally, the classification stage which produces two results; age estimation, and gender recognition.

The proposed system is implemented using the Internet Movie Database (IMDB) and IMDB-WIKI in combination. The obtained results showed that the proposed system introduces an accuracy of 98.2% for human gender recognition and an accuracy of 98.4% for age estimation. Furthermore, this proposed system achieved better results than the previous work.

List of Contents

	Contents	*Page No*
	Chapter One: General Introduction	**1-7**
1.1	Introduction	1
1.2	Related Work	2
1.3	Problem Statement	5
1.4	Aim of Thesis	6
1.5	Thesis Layout	7
	Chapter Two: Theoretical Background	**8-33**
2.1	Introduction	8
2.2	Face System Types and Challenges	8
2.3	Biometric System	10
2.3.1	Soft Biometrics	11
2.4	Face System Estimation	13
2.5	Age and Gender in Face System	14
2.6	Deep Learning Techniques	15
2.7	Augmentation	16
2.7.1	Type of Data Augmentation	19
2.8	Methods of Feature Extraction and Classification	19
2.8.1	Convolutional Neural Networks (CNN)	20
2.9	Basic Structure of CNN	21
2.10	Cross –Validation	26
2.10.1	Holdout	27
2.10.2	k-Fold	28
2.11	Model Architecture Efficient Net	29
2.12	Confusion Matrix	31
	Chapter Three: The Proposed System	**34-47**
3.1	Introduction	34
3.2	The Proposed System	34

3.2.1	Image Acquisition in the Proposed System	36
3.2.2	Region of Interest (ROI)	38
3.2.3	Resizing the Image	40
3.2.4	Augmentation	41
3.2.5	Holdout Technique	42
3.2.5.1	Training Set	43
3.2.6	Feature Extraction and Classification	43
3.2.6.1	Optimum CNN parameters	43
	Chapter Four: The Experimental Results	**48-74**
4.1	Introduction	48
4.2	Software and Hardware	48
4.3	Dataset in the Proposed System	48
4.4	Image Acquisition	50
4.5	Images Dataset Description of Ages and Genders	51
4.6	Pre-Processing Proposed System	52
4.6.1	Region of Interest (ROI).	53
4.6.2	Resizing Image (128*128)	54
4.6.3	Data Augmentation (DA)	55
4.7	Classification Using Multiple-Task CNN	56
4.8	Classification Evaluation	57
4.9	Evaluation System	62
4.9.1	Confusion Matrix for Gender	63
4.9.2	Confusion Matrix for Age	67
4.10	Previous Studies Comparing	73
	Chapter Five: Conclusions and Suggestions for Future Work	**75-76**
5.1	Conclusions	75
5.2	Suggestions for Future Work	76
	References	**77-92**

List of Figures

Figure No.	*Caption*	*Page No.*
2.1	General Catachresis of Estimation System Based on Face	13
2.2	Sequence of Learning Technical	16
2.3	Data Augmentations Taxonomy	18
2.4	The general structure of the CNN system	20
2.5	Pooling Layers and the Convolution	21
2.6	An example of convolution operation	22
2.7	Two Classic Pooling Methods	23
2.8	An Example of ReLU Transformation	24
2.9	Hold Out Technical	27
2.10	Dataset splitting into k-fold	29
2.11	different between k-fold and hold out	29
2.12	Architecture for EfficientNet-B3	30
2.13	Schematic representation of Efficient Net	31
2.14	Confusion Matrix	32
3.1	System Architecture for Age Estimation and Gender Recognition	35
3.2	Identification of facial landmarks using Dlib	38
3.3	The structure of ROI process.	39
3.4	Type and value of augmentation	41
4.1	Samples of an images trimming from the dataset	50
4.2	Final result in an example of the dataset	50
4.3	Histogram age description in proposed system dataset	51
4.4	Description genders in proposed system dataset	52
4.5	Image example and location in the dataset	53
4.6	Steps of ROI	54

4.7	Result of resizing image	55
4.8	Steps of data augmentation	55
4.9	Horizontal step	56
4.10	Classification result	57
4.11	Accuracy of train model	58
4.12	Train model loss result	58
4.13	Result of training (70%) vs. validation (30%)	60
4.14	Result of training (80%) vs. validation (20%).	61
4.15	Result of training (90%) vs. validation (10%)	62
4.16	Chart of gender classification	64
4.17	Chart of gender support classification	65
4.18	Histogram of gender results in the proposed system	65
4.19	Chart of age class support	68
4.20	Chart of three measures for age class result	69
4.21	Histogram result of age system	69

List of Tables

Table No.	***Caption***	***Page No.***
3.1	Summary of the proposed CNN layers.	44
4.1	Dataset number before trimming.	49
4.2	Dataset number after trimming and using in proposed system.	49
4.3	Finally number of multiple–task CNN structures.	56
4.4	Result of the three best models in the proposed system.	59
4.5	Result of proposed system gender.	64

4.6	Description of the result.	66
4.7	The result of evaluating ten classes in the test part.	68
4.8	Description of the result.	70
4.9	Description of range age between (91-100).	71
4.10	Comparison with previous studies.	73

List of Algorithms

Caption	***Page No.***
Algorithm (3.1): Image Selection	36
Algorithm (3.2): Region of Interest (ROI)	40
Algorithm (3.3): Data Augmentation	42
Algorithm (3.4): Efficient-Net-B3 Training	46

List of Abbreviations

Abbreviations	Meaning
2D	Tow Dimension
3D	Three Dimension
AI	Artificial Intelligence
CNN	Convolutional Neural Network
CNNs	Convolutional Neural Networks
DA	Data Augmentation
DCNNs	Deep Convolutional Neural Networks
DL	Deep Learning
DMTL	Data Mining Template Library
DRNNs	Deep Recursive Neural Networks
F	Female
FERET	Face Recognition Technology
GANs	Generative Adversarial Networks
HCI	Human-Computer Interaction
IMDB	Internet Movie Database
JPG OR JPEG	Joint Photographic Expert Group
M	Male
MAE	Mean Absolute Error
ML	Machine Learning
MTCNN	Multi-Task Cascaded Convolutional Neural Network
MTL	Multi-Task Learning
PBT	Pixel Brightness Transformations
RELU	Rectified Linear Unit
RGB	Red Green Blue
ROI	Region of Interest
ROR	RSSI-Based Omni-Directional Routing
VGG	Visual Geometry Group

Chapter One

General Introduction

1.1 Introduction

Face system is a computer technology for detecting and recognition of human faces age and gender determination in a digital image. It is utilized in a variety of applications [1]. Face detection and recognition have developed into a very active and significant area of image processing research. The bulk of current face detection algorithms focused on frontal human face identification and face recognition is a well-studied issue in computer vision [2]. It is a challenging vision issue with several practical applications, including identity verification, intelligent visual surveillance, and automated immigration screening systems [3]. According to numerous application scenarios, recognizing faces in real-world applications remains a challenging process [4]

Face recognition age and gender determination has been the most commonly used application of image analysis. The breadth of its commercial law enforcement applications and the availability of cutting-edge methodology have all contributed to its popularity. Additionally, it may be utilized for image retrieval based on content, video coding, video conferencing, crowd monitoring, and intelligent human-computer interactions [5]. The face systems include another important part, which is an estimation and classification, and it falls under the name age and gender and play fundamental roles in social interactions [6]. Despite the essential roles that attribute play in our day-to-day lives, the ability to automatically estimate them accurately and reliably from face images is still far from meeting the needs of commercial applications [7].

Recently, deep learning techniques had a lot of success in face systems. Convolutional Neural Network (CNN) is the most well-known deep learning example in which the training data represent a crucial issue. It is easier to obtain a greater network generalization with more training data. However, using merely

faces, appropriately labeled data for age and gender recognition may be limited and difficult to obtain. Overshoot is more likely to occur with insufficient training images. A variety of solutions can be used to alleviate the issue of overfitting. Working in a multi-tasking learning style and increasing the feature of including the age workbook is a modern and high-accuracy method for estimating age and gender recognition [8].

1.2 Related Work

Many publications in the field of human gender and age classification estimation based on deep learning algorithms have been published recently, and this thesis highlights a few of them:

- **K. Zhanc *et al.* (2017)** [9] Proposed a system using residual networks of residual networks (RoR). Furthermore, two simple procedures based on observation of age group features are described to improve the accuracy of age estimation. To increase performance and avoid over fitting, the RoR model is first pre-trained on Image Net, then fine-tuned on the internet movie database (IMDB-WIKI-101) to learn more about the features of face images. Then, it is utilized to fine-tune the audience dataset using two modest mechanisms, pre-training by gender and training with a weighted loss layer to improve age estimation performance. The accuracy of the proposed system in IMDB-WIKI is 66.74% in age and 93.24 % in gender. The proposed system's limitation is that because there is a need to investigate the application of RoR on a large scale and high-resolution picture classifications, this work does not considerably improve the age group and gender classification performance. The percentage of use of the data set was not mentioned.
- **Seok. Lee *et al.*** (2018) [10] Proposed a system that presents a deep residual learning model for estimating age and gender. This method detects faces

from the images entered into it from the dataset. Three deep neural networks are used in the estimate approach, residual learning methods are used, and train the model with images from the IMDB-WIKI database; in this dataset, there are few images of human subjects younger than 20 years old, and so have augmented the set by collecting images from the Internet. According to the results achieve , the suggested system has an age accuracy of around 52.2% and a gender accuracy of approximately 88.5%. Where in the proposed system, a portion of the database was utilized, and the system's training process was augmented by internet images, but images from a global and approved dataset were not utilized.

- **Koichi. Ito *et al.* (2018)** [11] Proposed a system using CNN for predicting age and gender from facial images. The CNN architectures, regression and classification, and STL/ Data Mining Template Library (DMTL) were investigated. When compared against other architectures, WideResNet has the most remarkable performance on age and gender prediction, with age determined by regression and gender predicted by classification. The introduction of DMTL allows CNN to perform better in accuracy and computation time and use the IMDB-WIKI dataset to evaluate the system. The result of the proposed system is that age Mean Absolute Error (MAE) is 8.59, gender accuracy is 93.86%. A robust data cleaning mechanism does not support the limitation of the system, and the limitation of proposed method's accuracy must be improved by adjusting the DMTL methodology.
- **Paolo. Giammatteo Lee *et al.*** (2019) [12] Proposed a system that has two models of CNN (Visual Geometry Group VGG16 type), with a prediction layer change that may be useful for edge computing devices. Both networks are intended to classify a human face according to gender and age. The first (VGG16/10) views gender and age as two intertwined features, with the final neurons designed to retain both qualities together simultaneously. In the prediction layer, the second one (VGG16/8+1) has a neuron for gender

prediction and another eight for age prediction. Such networks were designed to provide information on the gender and age of the individual detected in an image without the need for two separate systems and by using IMDB-WIKI dataset to evaluate the algorithm. The result of the proposed system in age accuracy is 57.0% , and gender accuracy is 88.4%. The system's limitation does not consider other essential edge computing requirements, such as processing time and accuracy.

- **Cao hong. Nga *et al***. (2020) [13] Proposed a system that uses the images from the IMDB-WIKI dataset presents a transfer learning pipeline to gender and age prediction. To begin with, freeze all layers in Image Net models that have already been trained. The models are then trained for four stages with predetermined learning rates, and the layers' blocks are unlocked in a predetermined order. Apply a multi-output neural network paradigm to simultaneously predict age and gender, with the final loss function depending on the sum of age and gender losses. The result of the proposed system is VGG-19 classification models have the highest gender accuracy (nearly 91%). Age's MAE is 15.23, and the system's limitation is on using a weighted sum of individual losses that can be applied to minimize the effect of age loss on the total loss. And the VGG-19 classification-based models can be trained further as their losses are still on a downward trend.
- **Mohammed benkaddour *et al.*** (2021) [7] Proposed a system that uses CNN to produce a gender prediction and age estimation system for a face image or real-time video. Three CNN network models with varied architectures (number of filters, number of convolution layers) were built and verified on the IMDB and WIKI datasets. The results revealed that CNN networks considerably enhance the system's performance and recognition accuracy. In IMDB, the suggested system has an age accuracy of 86.20 % and a gender accuracy of 94.49 %, whereas in IMDB-WIKI, the age accuracy is 83.97

percentage, and the gender accuracy is 93.56 percentage. The suggested network improves age and gender classification precision significantly.

1.3 Problem Statement

The problems can be summarized in two type (first type is detection, identification) and second type (estimation systems). Detection, identification and estimation systems are biometric systems, and the range of problems that face such systems are demonstrated in the following points:

• **Problem of dimensional image**: Faces are difficult to distinguish from patterns in the visual field. When it comes to identifying a three-dimensional item like a face, though, a two-dimensional picture is the best way.

• **Problem of human composition**: Face is a non-rigid body, where age and gender estimation systems suffer from an important problem, which is the different human races. As a result of this great discrepancy, it is difficult to determine and estimate with high efficiency, especially in ethnic minorities.

• **Problem of angles and backgrounds**: Estimation systems depend on the concept of working on clear and high-quality images in general, as the systems are less efficient in images with non-front angles or simple rotation and backgrounds with overlapping meanings. This problem is due to how the images are captured by the system hardware .

• **Problem of disparity in features**: Problem of disparity in human features in terms of extreme closeness between shapes or changing shapes by simply applying cosmetics and facial plastic surgery.

• **Problem of traditional systems**: traditional systems work on a system of classifying ages into groups in order to avoid the difficulty of estimating in childhood or old age.

1.4 Aim of Thesis

The aim of this thesis is to design and implement a system that can be used to recognize the human gender and age estimation with a high accuracy rate. The system is based on face image and utilizes computer facilities techniques such as deep learning and applies a Multitasking Convolutional Neural Network Algorithm (MCNN)) to achieve a high accuracy percentage of gender recognition and age estimation based on facial appearance. In order to achieve this aim, there are several objectives should be accomplished:

- Solve the problem of simple angles and backgrounds that selects only the image of faces and determining the areas of appreciation in the face, which are the forehead, eyes, nose, mouth and temples, and making a processing process that supports simple angles.
- Solve the problem of human composition or races in estimation systems by training and testing the suggested system on a very large data set with people of various races and ages ranging from one year to 100 years old.
- Solve the problem of contrast in images in terms of extreme closeness, different twins, cosmetic issues and Problem of multiple face by employing a multi-tasking deep learning system that uses the Convolutional Neural Network (CNN) technique to harness the advantages that support age and gender at the same time.
- Solve the Problem of traditional by design a system based on deep learning using the principle of classification instead of (clustering) and training the system to estimate age and gender without giving information about these categories based on facial appearance only.

1.5 Thesis Layout

In addition to chapter one, this thesis consists of four other chapters and are as follows:

Chapter Two: The theoretical underpinning of the general algorithm and approaches employed in this thesis is presented in this chapter.

Chapter Three: This chapter describes the suggested developed system and its associated algorithm in detail and the actions involved.

Chapter Four: This chapter contains the results acquired after using the suggested system on the data set in question and a commentary on the results.

Chapter Five: This chapter summarizes the study's findings and makes recommendations for future research.