Doctoral Dissertation

Thesis Advisor: Kim TaeYong

A Combined Classifier of Neural Networks with Decision Fusion for Age and Gender Classification

결정 융합과 신경망이 결합된 연령 및 성별 분류기에 관한 연구

February 2020

The Graduate School of Advanced Imaging Science, Multimedia and Film

Chung-Ang University

Major in Imaging Engineering-Game Engineering,

Department of Imaging Science and Art

James Rwigema

A Combined Classifier of Neural Networks with Decision Fusion for Age and Gender Classification

결정 융합과 신경망이 결합된 연령 및 성별 분류기에 관한 연구

Presented to the Faculties of the Chung-Ang University in Partial Fulfillment of the Requirement of the Degree of Doctor of Philosophy

February 2020

The Graduate School of Advanced Imaging Science, Multimedia and Film, Chung-Ang University

Major in Imaging Engineering-Game Engineering,

Department of Imaging Science and Art

James Rwigema

A Combined Classifier of Neural Networks with Decision Fusion for Age and Gender Classification

By

James Rwigema

Department of Imaging Science and Arts
Chung-Ang University

Approved:

채영호 ______________________

이원형 ______________________

하동환 ______________________

홍현기 ______________________

김태용 ______________________

Dissertation submitted in partial fulfilment of the requirements of the degree of Doctor of Philosophy in the Department of Imaging Science and Arts in the Graduate School of Advanced Imaging, multimedia & Film
Chung–Ang University

February 2020

Table of Contents

Abstract.

Age and Gender are identified as very important attributes in human identification and these attributes are used in various fields of Human and Computer Interaction (HCI) such as security systems, video-surveillance systems, online purchasing systems, judicial systems, transport, medicine, and so many others. In recent years, age and gender estimation based on facial feature analysis have been articulated as a challenging research topic by many researchers in the HCI field. In this research, we aim to present a combined classifier of neural networks with decision fusion for age and gender classification. The novelty of our research is the fusion of the decisions obtained by the two neural networks to increase the accuracy of age and gender estimation. We used the probabilistic decision fusion techniques such as Majority Voting decision fusion, Naïve – Bayes Combination decision fusion and Sum Rule decision fusion for better recognition accuracy rate. Among these technics used, the sum rule decision fusion provided the highest accuracy rate of 86.133 % which is higher compared to the state of art because of reducing the adjacent classes' likelihoods during decision classifications.

Key words: *Age and Gender estimation, Gabor filters, Artificial Neural Networks, Probabilistic Decision Fusion, Convolutional Neural Networks, and Support Vector Machine*.

List of Tables

List of Figures

I. Introduction

1.1. Background

Lots of research have been undertaken in the field of age, gender, and race estimation which covers many technical areas such as image processing, surveillance and security, telecommunication and human-computer interaction. The world-wide range of commercial and law enforcement applications are a sign of its huge economic significance. Therefore, there is a high demand of building automatic systems capable of processing facial image object and extract the most useful information from these biometric features that could be used by these systems. Since the most important and impressive biometric features of any human being is the face, facial images are praying a vital role in providing the appropriate features for gender recognition and age estimation.

In this thesis, we aimed to study on gender recognition and estimating the corresponding age of human facial images.

Developing a model with the capability of recognizing the gender while estimating the age of the presented facial image and extracting their gender and age information is a challenging task. This is because, there is a big necessity of creating a general model capable of extracting useful biometric features for all kinds of human facial image subjects. Therefore, since each person's facial image has his/her personal unique distinctive features that vary in different ways from individual to individual [1], developing a model that extracts useful features to discriminate between individual facial images requires in-depth studies of human face objects.

Processing human faces requires considering many aspects. One such aspect is to analyze face structure and determine the exact location of face elements, such as eyes, noses, mouths, eyebrows, lips and cheeks. Another is the extraction of relevant information from these detected elements, which can provide us with useful information regarding the identity, age, ethnicity and gender of a person [2].

However, human gender recognition and age estimation currently presents a very active field of research in today’s researchers and scholar whom belongs into this field of research.

1.2. Motivation

Regardless of the major efforts deployed in the age and gender prediction, the results still indicate that until to date it is seen as a challenging area of research due to inadequate training datasets where every person needs a variate of facial images in a defined wide range of age group, also its challenging for the Age identification due to the skin texture and illusion of their skin color.

With the development of human-computer interaction (HCI), the research on facial image has been extensively carried out in many areas including image processing, pattern recognition and computer vision. The human face contains an amount of important information related to personal characteristics, the identity, emotional state, ethnic origin, gender, age, and head orientation of a person are all shown in a face image [3].

The major objective of our research is to increase the recognition rate of gender and age accuracy through the distinctive human facial features used to describe the gender and the age on any human being.

Identifying persons to allow them access to or control of facilities, tools and information are amongst the most common applications of gender and gender recognition through their respective face facial images. As an example, human facial recognition technology is currently being used by hotels and casinos to identify a blacklisted under age individuals who are illegitimate from accessing these facilities. Among all these human facial information, age and gender are among of the most significant characteristics which is widely used in many applications, such as human-computer interaction, surveillance monitoring, and video content analysis. For example, an automatic age and gender estimation system can not only improve the human–computer interface, but also prevent under ages from accessing cigarettes, alcohol, and pornographic websites, gaming privileges etc. Therefore, facial age and gender estimation has attracted increasing attentions from scholars in the field of computer vision and pattern recognition (human age recognition is a sub- element of face recognition).

The performance requirements of our gender recognition and age classification are as follows

- Flexibility: the developed model should adopt to the available gender and age databases irrespective of the variation in dataset numbers with respect to age categories and produce an accuracy rate of classification.
- Accuracy: Gender Recognition and Age Classification should be improved above 80%. The

classifier should be able to provide reliable accuracy classification results

- Scalability: Reduce the processing time the size of classification vocabulary can be increased without hindering the processing time.
- User independence: the classifiers should be able to classify new input images from other new databases without affecting the accuracy rates.

1.3. Problem Description

In this thesis, we aim to propose a novel multi classifier decision fusion for age and gender classification which will increase the classification accuracy of gender and age classification. As many previous researches in the this field of gender and age classification had mentioned about age bias with respect to the available public databases used in their research, this have affected the experimental accuracy rate, we used the decision fusion methods to solve the lack of enough image in age based databases. Thus, this research is subjective to the field of human face recognition using facial features. Therefore the common known procedures used in this research includes; facial image acquisition (input image), image pre-processing, feature extraction and output classification which is presented in figure 1.1.

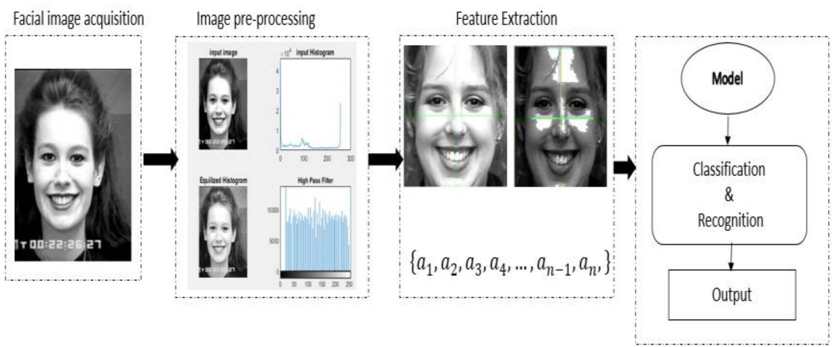

Figure I. 1. General image processing procedures.

For any image classification system, it requires the four major steps mentioned in the figure I.1. Where under each step, there several operations applied to any image to be classified depending on the intended output of the researcher. Therefore, Image processing system includes treating images as two dimensional signals while applying already set signal processing methods to them.

The detailed description of these steps are described in the figure 1.2 below.

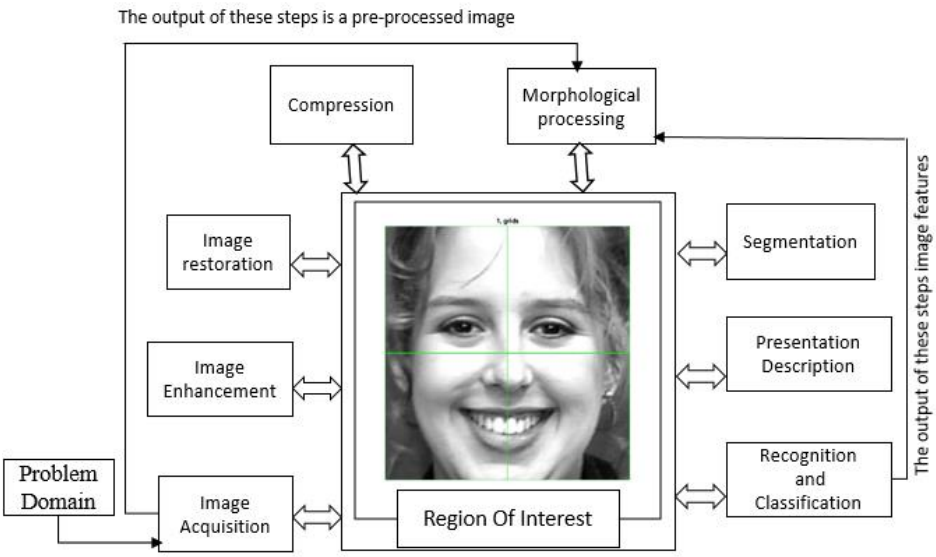

Figure I. 2. Image processing block diagram.

1.3.1. Terminologies

Feature extraction

This is a mechanism dimensional reduction of raw data to a more meaningful and manageable classes for better processing. Feature extraction is very useful in many image processing applications where it helps to select the appropriate data without losing the most important or relevant information.

Feature classification

This is a pattern recognition technique that categorically places a huge number of processed data into different classes or groups with respect to the research targets.

Decision fusion

This is a mechanism of data concatenation that combines multiple decisions from a number of classifiers into a common decision received or generated by multiple classifiers.

An overview the image processing scheme is present in figure I.2. An input colored or grayscale facial image is inserted into the system in the image acquisition stage [4].

Image Acquisition

This is the first step or process of the fundamental steps of digital image processing. Image acquisition could be as simple as being given an image that is already in digital form. Generally, the image acquisition stage involves preprocessing, such as scaling etc.

Image Enhancement

Image enhancement is among the simplest and most appealing areas of digital image processing. Basically, the idea behind enhancement techniques is to bring out detail that is obscured, or simply to highlight certain features of interest in an image. Such as, changing brightness & contrast etc.

Image Restoration

Image restoration is an area that also deals with improving the appearance of an image. However, unlike enhancement, which is subjective, image restoration is objective, in the sense that restoration techniques tend to be based on mathematical or probabilistic models of image degradation.

Compression

Compression deals with techniques for reducing the storage required to save an image or the bandwidth to transmit it. Particularly in the uses of internet it is very much necessary to compress data.

Morphological Processing

Morphological processing deals with tools for extracting image components that are useful in the representation and description of shape.

Segmentation

Segmentation procedures partition an image into its constituent parts or objects. In general, autonomous segmentation is one of the most difficult tasks in digital image processing. A rugged segmentation procedure brings the process a long way toward successful solution of imaging problems that require objects to be identified individually.

Representation and Description

Representation and description almost always follow the output of a segmentation stage, which usually is raw pixel data, constituting either the boundary of a region or all the points in the region itself. Choosing a representation is only part of the solution for transforming raw data into a form suitable for subsequent computer processing. Description deals with extracting attributes that result in some quantitative information of interest or are basic for differentiating one class of objects from another.

Recognition and classification

Recognition is the process that assigns a label, such as, "vehicle" to an object based on its descriptors.

Region of Interest

Knowledge may be as simple as detailing regions of an image where the information of interest is known to be located, thus limiting the search that has to be conducted in seeking that information. The knowledge base also can be quite complex, such as an interrelated list of all major possible defects in a materials inspection problem or an image database containing high-resolution satellite images of a region in connection with change-detection applications.

This step involves the computation of the pre-processed image and this is the step in which the local and global features of the input images are extracted depending on the regions of interest with respect to the output classifier of the system, the extracted features are computed using various computer applications such as Active Appearance Mode and Gabor wavelet transform [4, 5 ,6], Principal component analysis [7,8] for global features extraction, and Gabor filter [6], Linear Discriminant analysis(LDA) [7], local binary pattern (LBP) , Histogram of oriented Gradient (HOG) Speeded-Up Robust Features (SURF) etc. For local feature extraction, these extracted features are stored as feature vectors to be used during the verification.

Then the next step is verification where the test image passes through the previous processes and the extracted features are compared with the temperate stored feature vectors during the enrolment process for matching store. After verification the next step is the classification process which is done using different classifiers such as Support Vector machine (SVM) [8], nearest neighbor classification [9, 10], Extreme learning machine (ELM) [11], AdaBoost classification [12], random forests [13] artificial Neural networks [14, 15], decision trees [10] and so many other.

The last step during image processing is the output display which show the intended result of the system. With due to previous research the image processing field, human facial face detection is on of dominating field of research which have took an overwhelming interest of different researchers, among others it includes face detection and recognition, human emotion recognition, age estimation, gender recognition, race recognition and even a combination of two or three research areas.

1.4. Contributions

This thesis proposes a Gender recognition and age classification model using a combined classifier of neural networks with decision fusion for age and gender classification. The fundamental idea is to increase the classification accuracy rate of both gender and age of the input images from a defined database. Decision fusion methods is used to fuse the nearest likelihood decisions provided by the classifiers hence resulting into an increase of the classification accuracy.

The original contributions of this thesis are:

I. Proposing a hybrid approach of neural networks for classifying gender and age of the same image at the same time

II. Using decision fusion methods such as; majority voting, naive-bayes probabilistic method and the sum rule method to fuse the decision provided by the classifiers in order to increase the classification accuracy rate of age.

III. Through the proposed approach, we overcome the overfitting problem of age classes during classification being caused by the age bias found in public age database due to a big number of image sets compared to the others which finally leads to difficulties in the training phase.

IV. Provided a bigger option of classification through probabilistic decision fusion.

1.5. Thesis organization

This thesis includes 5 sections.

1. Section I: Introduces the background and motivation of this work.
2. Section II: Contains an over view of relevant researches in the field of gender and age classification. A comprehensive review of different pattern recognitions methodologies often used in gender recognition and age classification is presented in the section.

3. Section III: Presents the architecture of the proposed model, firstly recognition the gender of the input image using the Gabor filters as feature extractor and using the Simple Vector Machine as the Gender classifier, then how we used the cropped image as an input of the CNN while the feature extracted by the Gabor filters are the input of the conventional-artificial neural networks, then finally how we applied the decision fusion methods to fuse the decisions provided by the two classifiers.
4. Section IV: Shows the effectiveness of the proposed model as far as gender and age classification is concern. The proposed method is compared to other state of art methods. The challenges faced during age classification, and how we managed to solve them is also discussed in the section.
5. Section V: concludes the thesis by summarizing the contribution of this work and proposes the possible improvements in the feature work.

II. Related Research

2.1. Gender and age estimation

For any gender and age classification system, the most primary requirement is the facial part of the image to be classified. Therefore, gender and age classification is among the research fields involved in face recognition systems.

Face processing has long been recognized as an important module for many computer vision applications. Face recognition, and the classification of the age and gender of face objects are two interesting field of research in this area. With such a face analysis component, it becomes possible to identify a person in order to allow access to private facilities or to display targeted information in advertising based on demographic category of individuals in public places. In this chapter we provide a brief review of some of existing methods in face, gender and age classification and discuss their strengths and weaknesses.

When given an image to process, the face recognition system detects a human subject through face detection techniques and the face regions are segmented from the respective images. Then, the facial features are identified in order to align them into a conical way. Thereafter, the face representation is extracted from the face region that are finally fed into the classification model to find the face in the pre-trained database, which matches to the extracted facial features [16].

For every face recognition model is divided into two major phases which are; Face identification and Face verification as presented in the diagram 2.1 below.

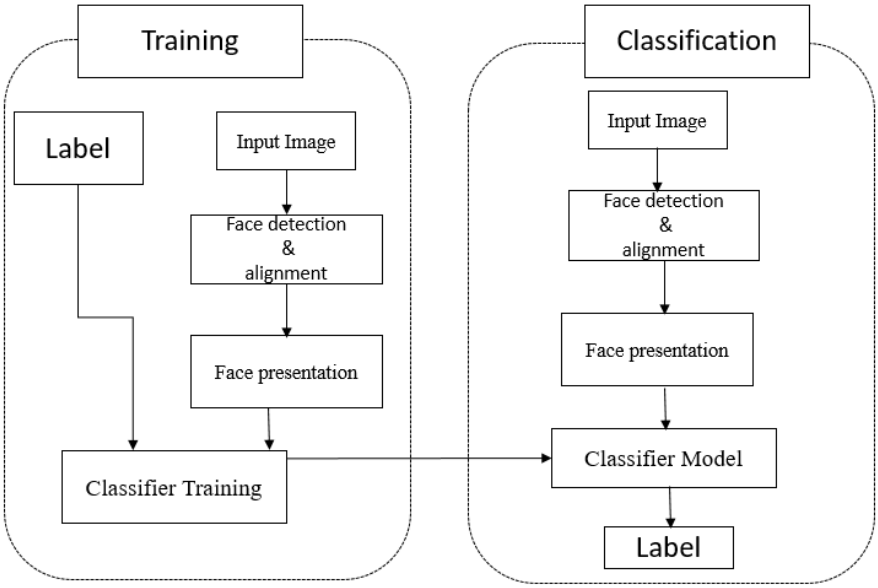

Figure II. 1. Face recognition system [79].

Face Identification: The designed system has to acknowledge the unknown face received from a pool of candidate faces.

Face Verification: The designed system must either accept or reject the claimed faces as belonging to a specific person. This type of application is used in different applications, mostly those related to access control with respect to recognition goals.

Our research is grouped into two categories, which are gender recognition and age classification. Where in gender recognition we used the Gabor filters for feature extraction and Simple Vector Machine (SVM) for classification purposes.

2.1.1. Gender Recognition

A comprehensive review of the methods have been provided to recognize the gender based on facial images [17], they focused on developments related on 2-D based systems which included a concise section of techniques involving 3-D data.

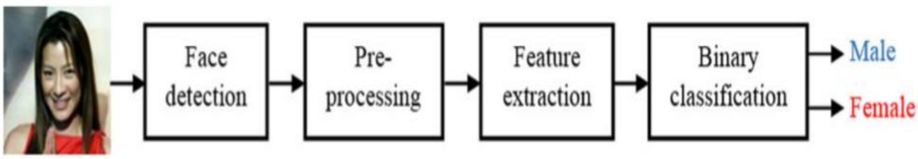

Figure II. 2. General framework for gender recognition system [17]

Yu et al [18], presented a study and analysis of gender classification based on human gait. it have been seen through their research that, human can recognize gender based on the gaits information and the contributions of different body components vary. The most significant body parts for gender cognition and the head and hair, back, chest and thigh. However, this research suffers some challenges such as view variations, clothing and shoes as well as carrying objects. In [19]. They used to SVMS for gender classification through thumbnail facial images in comparison with the traditional classifiers such as linear, Quadratic, Fisher Linear Discriminant, nearest neighborhood and more modern techniques such as RBF networks and large assemble- RBF classifiers. In [20], a systematic study on gender classification with automatically detected and aligned faces was presented. This findings has indicated that, automatic alignment of images would be useful in gender classification once the alignment there is a further improvements in the image alignment methods. In [21], proposed a feature selection method by using genetic algorithms to select features extracted by PCA. They compared different classifiers such as Bayesian, NN, LDS and SVM and demonstrated that using a SVM classifier is a better approach for classifying gender. In [22], they suggested "SEXNET" Neural Network model to recognize gender. The network uses the faces' raw pixels to compress the face and then estimates their sex in subsequent layers of their proposed network. In 1995, Brunelli et al. [23], achieved a 79% recognition rate for gender by using the HyperBF network on a set of geometrical features extracted from faces. In [24, 25]. They mentioned that facial landmarks constitute the most compressed representation of faces and are known to preserve information such as pose, gender and facial structure present in the faces.

In [26], a comprehensive experimental study was carried out on gender classification using non-distorted and distorted faces, and two approaches comparison were considered (local and global), where they considered three types of features and three classifiers provided by three statistical tests applied on two performance measures. A gender recognition approach was proposed which combined Haar-like wavelets with Ada+SVM classifier, the Haar-like features provided a higher speed when recognizing the gender from the face images based on fast calculations algorithms [27].

Alexandre [28] proposed a multi-class gender recognition system, where he based on shape of the image, texture and plain intensity features gathered at different scales. Features where extracted at different image resolutions and obtain the classification based on the extracted features and finally fuse the obtained decisions.

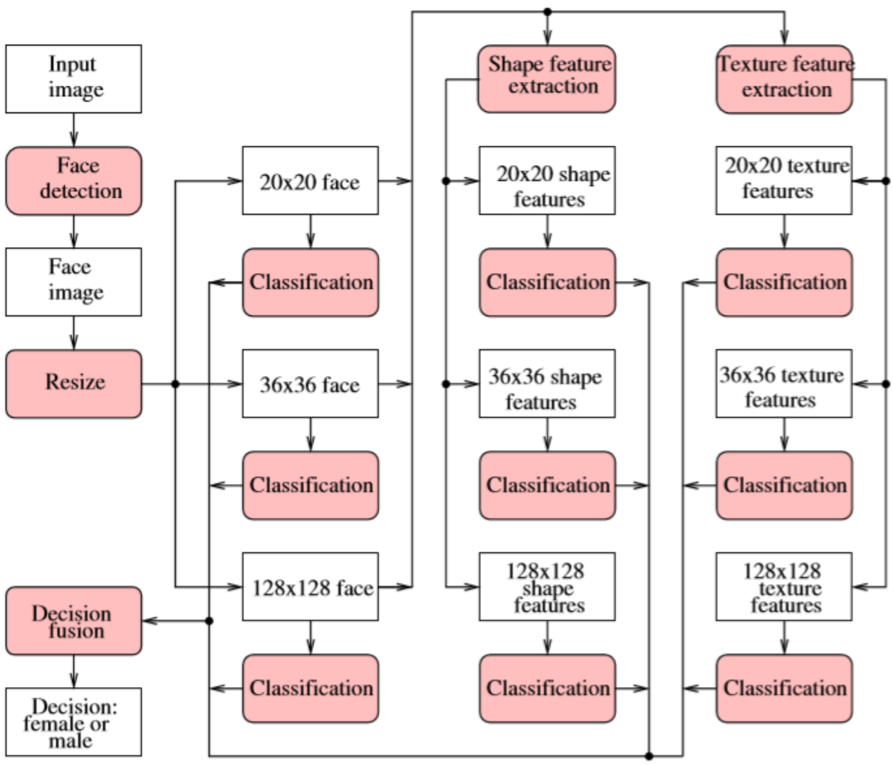

Figure II. 3. Multiscale decision fusion approach for Gender recognition [28].

2.1.2. Age Classification

Human faces are mostly affected by age where it is not automatically identified during the aging periods, most of the changes happens due to a variety of factor such as healthy, the race group, weather conditions, gender, lifestyles, alcohol, foods, drugs used and so many others. In [15], they made a survey on the complete state-of-art the techniques in the face images-based synthesis where they mentioned two major fundamental problems which inspired them to develop the techniques which are:

- Face image synthesis defines as rendering face images with customized single or mixed facial attributes (identity, expression, gender, age, ethnicity, pose, etc.).
- Face image analysis which they defined as interpreting face images in terms of facial attributes as mentioned above.

Several research approaches have been used deployed on facial images for different research targets. Facial images were used in an ethnic estimation research where they used both the convolution neural networks in comparison with the artificial neural network [4] a frame work of integrating multiple models (AAM, LBP, GW, LPQ) for facial age estimation based on multi features was proposed [29], and an overall result of 64.8% of age prediction was achieved, a deep learning algorithms of automatic age estimation (deep convolution neural networks) used to extract high-level complex age related visual features and predict age range of input facial image was proposed in [4].

Loss functions and age encoding strategies are another source of variation between different AE CNNs. Some papers address AE as an ordinal regression problem [30]. A facial feature detection for age classification in [31], where an explicit feature extraction and analysis was proposed, facial wrinkles were considered as the most important features to be considered for age classification. In [32] used the deep-Convolution neural networks (CNN), which provided a significant increase in classification performance. The evaluation was made on FG-Net for age and gender classification. Also in [33], a fast and robust system is proposed for age group classification. In [34] they presented the most Known aging pattern subspace (AGES) which uses Active Appearance Mode (AAM), the basic idea of AGES is to model the aging pattern, which can be defined as a sequence of a particular individual's face images sorted in time order, by constructing a representative subspace. A fuzzy version LDA was introduced through defined age memberships to solve intrinsic age ambiguity problem and they used the Gabor features and fuzzy LDA to achieve a classification precision in the consumer images [5]. However, these techniques do only use the local features foe the age classification yet the global features can play a vital role in the age classification process.

The basic Hough Transform method has been applied to detect straight lines [35] and was later extended in [36] to be able to compute shape analysis and identify arbitrary shapes. This approach was applied for the purposes of face recognition in [35] and proved its robustness against various noises. It also has high level of efficiency in terms of its memory usage. In [37], proposed a method for age classification that first extracted specific features of the face elements such as eyes, noses, mouths and chins. It then compute the ratios estimated between the top of sides of the head before, finally, processing skin wrinkle information in order to classify people in three classes: babies, young adults and seniors. Wen Bing Horng et al [33], employed Sobel edge detector with a back-propagation neural network to classify human face subjects into four classes: babies, young adults, middle adults and old adults.

In [38], shows a detailed survey of several approaches which could be used in age and gender recognition. These studies outlines the appropriate models or algorithms that could be fitting in the extraction and classification of Gender and age with respected to specified classes. M. Wiggins et al [39], proposed a model of classifying patients based on the age. In their model, they used the naïve-Bayesian classifier that provides an accurate rate of 84%. It was indicated that, the methodology for evolving the Bayesian classifier could be used to evolve Bayesian networks in general thereby identifying the dependencies among the variables of interest. Such a classifier can then be used for medical applications for diagnosis and prediction purposes. In [40], the presented combined method of Active Appearance model as a feature extractor and support vector regulation as a classifier. Zhang et al [41], presented a Multi-task Warped Gaussian process [MTWGP] to personalize the age estimation. Liu et al. [42] used a hierarchical age grouping to train an AE CNN reporting the currently best score on MORPH-II following the well-established protocol from [43].

Age and gender recognition is considered as one of the crucial parts for many computer vision applications including demographic data collections, visual surveillances and others. The current trending research in age and gender recognitions was discussed in [17]. In [77], they presented a result of 42.9 % of age classification and 74.1 % of gender recognition. In their research they use a combination of facial appearance and context. Among the early algorithms in the field of age and gender recognition, Cottrel and Met-calfe [44] extracted the whole-face features, which were fed deployed into a back propagation network model to classify males and female. Nguyen et al. [45, 46], proposed an age and gender classification model which used EEG paralinguistic features for the classification and learning local binary patterns for gender classification in the real world face images. In [50] an age and gender recognition model was presented, their research used boosted Gabor features for feature extraction where they reached an accurate rate of 50.3% of age classification and

75.7% of gender classification.

A fine-tuned age range method was applied to estimate the age using a private database was conducted in [14], also in [47], an image based age group classification was proposed purposely for three major age group estimation namely child, adult and elderly.

In 2013, Chen, Y. et al. [48], introduced a new method based on subspace learning that operates as a set of constrained optimization problems to characterize age-related features. By employing semi-supervised learning techniques, they applied the Support Vector Regression (SVR) methods onto the features to create an age estimators model.

A novel method was proposed for facial expression recognition with convolution neural networks coping with data by adjusting their respective position hence forming another dataset [49], and Gabor filter were used to extract both local and global features [32], where these features are fused together into feature vector being used as a face descriptor for recognition. Fusion was applied in case the global features were not clear during the extraction. It have been observed that, all studies [51 and 52] which train gender/age CNNs use shallow architectures, while the works employing deeper architectures (like AlexNet [53], or VGG-16/19 [54]) fine-tune already 150 pre-trained CNNs [55, 56 and 57]. In addition, several studies [58, 59] compared single

-task training for GR and AE versus simultaneous multi-task training [70]. However, all the above mentioned research methods have suffered a low accurate rate in the combined Age and Gender research field due to two major challenges which are :

- In abundancy datasets leads to overfitting once age estimation is not a true/false experiment instead it's a multi-class decision experiment.
- Irrespective of the available age databases, still these databases are age biased where some images has a big number of image sets compared to the others, which finally leads to difficulties in the training phase.

In section III. The proposed approach is a combined classifier of neural networks with decision fusion for age and gender classification**.** After image normalization, we extracts the local features of the normalized image using Gabor filters and these features were trained by SVM to determine the gender of the image. In addition, after gender classification, these features are further passed through artificial neural networks with respect to the gender of the image while on the other side, the normalized image is cropped to a passport size and the whole image is passed with in the Convolutional neural network as presented in Figure III.1.

III. A Combined Classifier of Neural Networks with Decision Fusion for Age and Gender Classification.

We proposed a combined classifier of neural networks with decision fusion for age and gender classification. After image normalization, we extract the local features using Gabor filters and these features were trained by Support Vector Machine (SVM) to determine the gender of the image. These features are further passed through conventional artificial neural networks with respect to the gender to determine the age of the image depending on the facial local features of the image such as facial wrinkles and skin texture. On the other side, the normalized image is cropped to a passport size and the whole image is passed within the convolutional neural network. In our research we implemented decision fusion techniques [60, 47] after classifier's decision of neural networks, this was done to overcome the overlapping of age categories hence improving the accuracy rate. Decision fusion techniques were applied to the decisions obtained through the neural networks during age classification; this was done by fusing the decisions provided by the neural networks depending on the neighboring likelihoods of class labels. The proposed model is presented in the figure III.1 below.

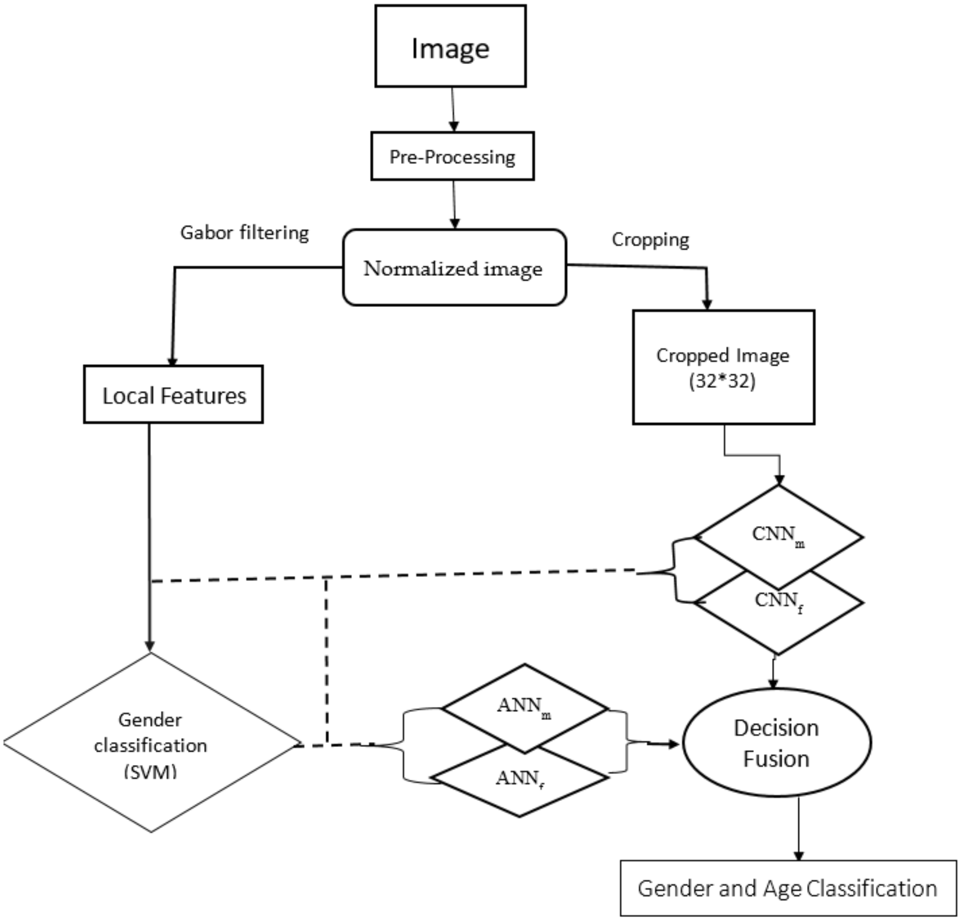

Figure III. 1. A flow chart for the proposed model.

Most of the previous works, have been focused on single field either age estimation or gender classification [61], here they practiced the age estimation, age-based and a sequential study of rank-based age estimation methods by using a divide and rule age estimator. In [14 and 62], it indicates the impact of image pre-processing operations in order to achieve the desired accuracy. There is clear dependences between gender and age, in [15], it have been indicated that age is an instance of gender where gender plays a significate role during age estimation of which it is a great importance to consider the gender of a given image in order to be able to classify the age.

3.1. Feature Extraction for Gender and Age Classification

In image processing field, Pre-processing is a very critical factor where in most databases, the size of the images in the database varies and contains a variety of background information in many cases. During the pre-processing, we used the matlab tool to pre-process the image while identifying the most reliable points of interest that includes eyes, nose point, mouth, chine and forehead. Also, the images are resized and further cropped for clear feature extraction. This information is irrelevant and to avoid it, this image normalization is performed.

During the normalization process, several operations were applied upon the images in the databases such as; image re-sizing, grayscale conversion, histogram equalization and image cropping as shown in the figure III.2.

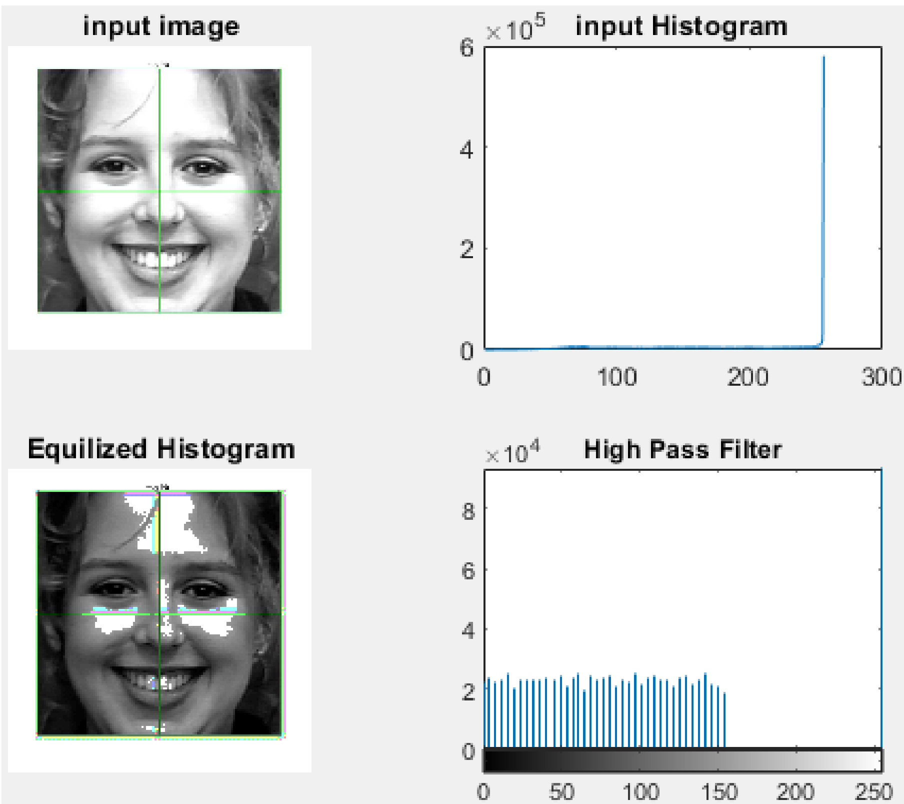

Figure III. 2. Image pre-processing.

Local features such as wrinkles, skin, hairs and geometric features have been commonly used to classify the age groups in many studies [62].

3.1.1. Feature Extraction for Gender Classification.

Various feature extractors have been used in computer vision application, in image processing, features are defined as the most distinctive information which is extracted from the images in form of numerical values that are mostly not understandable by human beings. These features are categorically describe into two classes based on the application to be used i.e local and global features. Local features are used as descriptors mainly for object recognition and I identification while global features are used as descriptors for image retrievals object detection and classification.

The major challenge in facial images is the robustness of local feature such the facial wrinkles and skin illuminations, many algorithms have been used in the extraction of these features like

Local Binary Patterns (LBP) [49], Wavelet Decomposition (WD) and Sobel edge magnitude [63]. However, for Gender recognition and age estimation purposes, Gabor filter are the most reliable method to extract the local features. In this research, Gabor filters were used to extract the dominant direction / regional wrinkles and the skin textures [3]

The two dimension Gabor filter are defined as follows

$$g(x,y) = \left(\frac{1}{2\pi\sigma_x\,\sigma_y}\right) exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2}+\frac{y^2}{\sigma_y^2}\right)+2\pi jwx\right] \quad (1)$$

Where σ_x and σ_y are the standard deviation of X- and y- axes and W is the radial frequency.

$$G(u,v) = exp\left[-\frac{1}{2}\left(\frac{(u-W^2)}{\sigma_u^2}+\ \frac{(v^2)}{\sigma_v^2}\right)\right]. \quad (2)$$

Where $\sigma_u = \frac{1}{2\pi\sigma_x}$ *and* $\sigma_v = \frac{1}{2\pi\sigma_y}$

The Gabor Wavelet is made by the dilations and rotations of $g(x, y)$.

$$g_s(x,y) = a^{-m} g(x', \tilde{y}) \quad (3)$$

$$\begin{cases} x' = x\cos\theta + y\sin\theta \\ y' = -x\sin\theta + y\cos\theta \end{cases} \tag{4}$$

Where θ is the filter orientation expressed by $\theta = n\pi/K$ where K is the number of the filter's orientation and a^{-m} is the filter's scale, m = 0 … S, where S is the number of scales.

$$a = \left(\frac{U_h}{U_l}\right)^{\frac{1}{(S-1)}} \tag{5}$$

Where U_l and U_h are the lower and upper average frequency. The figure III.3, below presents a human facial face image with a 2D feature vector components.

Figure III. 3. Human facial face and its 2D Gabor presentations

Thus, the Gabor function can be thought of as being a Gaussian function shifted in frequency to position. From the eq. (1) and eq. (2), applied for texture feature extraction by using Gabor functions and wavelets. The Gabor functions built a complete basis set but non-orthogonal. Expanding a signal using this basis provides a localized frequency description. A class of self-similar functions referred to as Gabor wavelets is considered by letting g(x, y) to be the mother Gabor wavelet. Then this self-similar filter dictionary can be obtained by appropriate dilations and rotations of g(x, y) through the generating function.

All filters in our research were allocated with respect to facial region of interest. The allocation of these filters depends on the direction of the facial wrinkles. To determine the wrinkle features, the mean and variance of the magnitude response of the Gabor filter in each defined wrinkle area was calculated, due to the fact that the mean and the variance of the magnitude represent both the strength and the quality of the wrinkles.

3.1.2. Feature Extraction for age Classification

In an aging human facial images, ageing features are more difficult to extract, in our research we used both the local features extracted by the Gabor filters and these features are distributed into the artificial neural networks for age classification. While on the other side we used the global features which were collectively represented by the entire facial image cropped in the pixel range of 32 X 32 which was distributed into the convolutional neural network for both feature extraction and age classification.

3.2. Gender classification

Gender is an important demographic attribute of human beings. In computer vision, gender recognition is one of the challenging standalone research field which attracts a number of researchers. As any other research in the identification of human demographic attributes, gender recognition can play a vital role in numerous applications such as Human Computer Interaction (HCI), surveillance, content-based indexing, biometrics , demographic studies and targeted advertisements and so many others. In [64, 65 and 66], indicated that gender detection can be useful for human-computer interaction, such as the designation of individuals where several algorithms have been designed for this purpose and the proportion of each of these issues has been resolved, they based on Gabor filters and Local Binary Patterns (LBP) for extracting facial features that these characteristics are robust against interference in order to achieve an appropriate classification. Therefore, Local Binary Patterns (LBP) are among the most basic and popular handcrafted features which were used for GR [45, 67]. As image processing continue to image continues to be an important research field, Microsoft in 2010 released a facial recognition application named "Azure" which is used as both commercial and public for facial image recognition. In [80], they have carried out an assessment of the four most popular face recognition tools with reference to gender, age and race recognition which includes Face++, IBM, Amazon and MS Azure, MS azure have presented a row average recognition percentage accuracy of 45% of age classification and 97% of gender recognition . Also in [81], they compared there developed model for gender, age and race recognition to face++ , MS Cognitive Services (Azure) and Hype face

In our research approach, we used the Gabor filters to extract the local facial features [68], we used the lower and upper frequencies, and the radial frequency was equal to the upper frequency. Our Gabor was set to 6 orientations and scales of 4 which makes it 24 Gabor filters, then $U_l = 0.025$ and$U_h = 0.05$. We used these values to calculate the mean, standard deviation, root mean square value and finally, these parameters helped us to form a one-dimension feature vector of the extracted local features. The SVM classifier was used for gender classification.

The figure III.4 below indicates the gender classification flow chat.

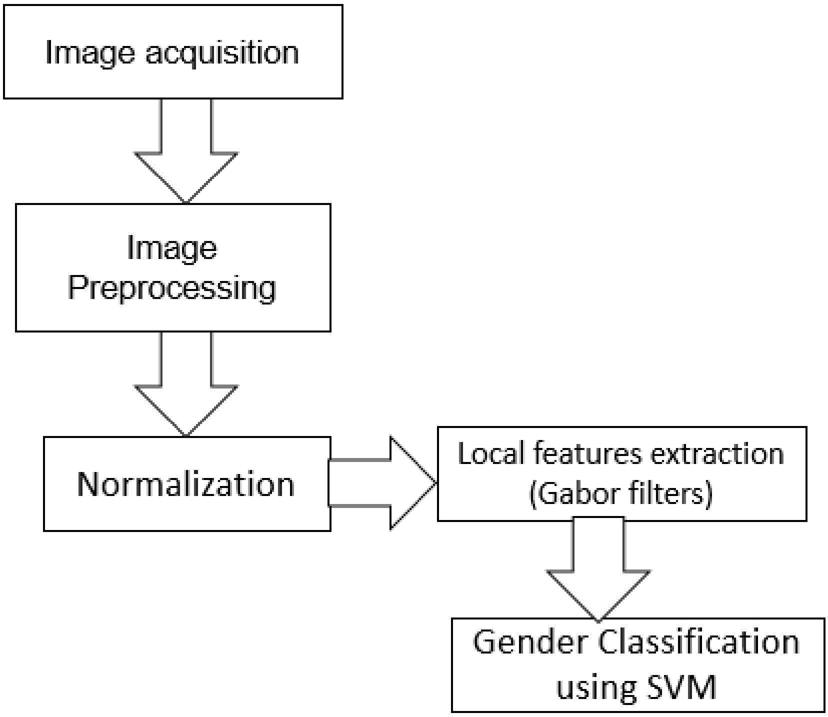

Figure III. 4. Gender classification flowchart.

3.3. Age Classification

Recently, a human facial age classification has drown a lot of attention in the field of computer vision due to its important applications in age-based Artificial Intelligence (AI), biometrics, Human Computer Interaction (HCI). Yi el al. [69]. Has indicted that a minor difference of mono-task and multi-task training. In [38], classified the age by considering the two geometric features and three wrinkle features obtained from the facial images of the human beings. In this research, we referred to the related previous researches to enhance the age classification accuracy by comparing of both age classification with decision fusion and age classification without decision fusion of the hybrid neural network.

3.3.1. Age classification by Artificial Neural Networks

As we aim to produce a gender and age recognition as the final output, we used a hybrid combination of artificial neural networks and convolution neural networks for age classification after gender recognition. After extracting the facial features of all the sub blocks and constructs the feature vector, the feature vector is saved in the database. The SVM Classifier was used to these values, compares these values with the feature vector of the trained images, and returns the closely related image. Finally, the best matching image is returned as the result to the application device.

3.3.2. Artificial Neural Networks Architecture

The artificial neural network was set up with the reference of that one which was used in [14], we used the Multi-layer Perception MLP model of the ANN for training the extracted features, where 24 facial feature were used as inputs in this Neural network, the network structure includes 3 hidden layers and each layer had 50 neurons with respect to gender category. The MLP was trained for 2000 epochs.

After the feature extraction from Gabor filters, to maintain the extra information which is obtained from the Gabor filters along with the original image, the weighted sum of image and Gabor responses is used as the input to the Artificial neural networks were used to for further local feature extraction of the recognized gender output, and these extracted features were trained in order to contribute in the age estimation process of the designed system. Since the gender of the facial image have been recognized as presented in Figure III.4, the artificial neural networks were subdivided into two respective gender categories as presented below.

$w_{ann(f),0} - w_{ann(f),R-1}$ for ANN_f , $w_{ann(m),0} - w_{m,R-1}$ for ANN_m ,

$w_{m,0} - w_{m,R-1}$ refers to the weights of the male confidence values of ANN_m,

$w_{f,0} - w_{f,R-1}$ refers to the weights of the female confidence values of ANN_f.

Confidence values of the two convolution neural networks are presented as:

$c_{ann(f),0} - c_{ann(f),R-1}$ for ANN_f , $c_{ann(m),0} - c_{ann(m),R-1}$ For ANN_m.

3.4. Age Classification for Convolutional neural networks.

Convolution neural networks is well known deep learning algorithm capable of analyzing facial images with respect to the targeted output. In our research we applied convolution neural networks for age estimation by training the whole facial passport image after pre-processing and cropping the image to a minimized size of 32×32 full passport image size as shown in the figure III.5, below.

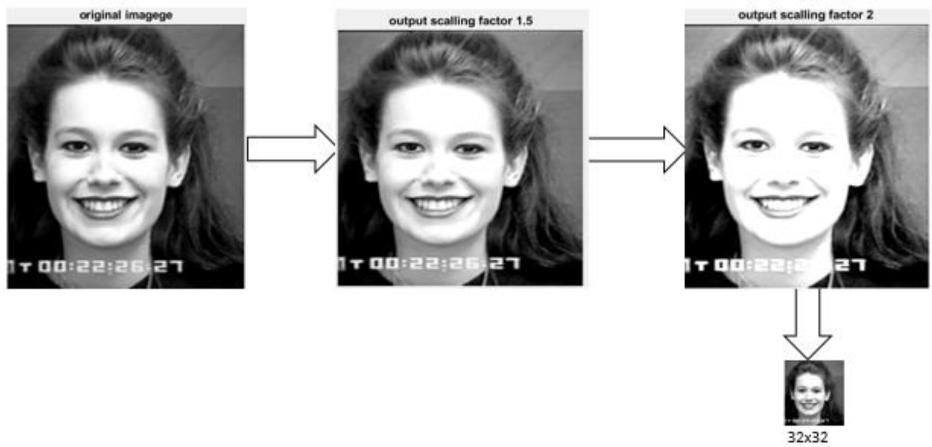

Figure III. 5. Image resizing for CNN

3.4.1. Convolutional Neural Networks Architecture

We used the Convolution neural network for age estimation by training the whole facial passport image after pre-processing and cropping the image to a minimized size of 32×32 pixel full passport image size with respect to the gender category as presented in section 3. We used a narrow convolution neural network set-up as discussed in [70] as shown in table II.1.

Table III. 1. The Depth of convolutional neural networks used in our proposal

$CNN_{(m,f)}$ Depth	2 convolution layers
	4 convolution layer
	6 convolution layers

As indicated in table1, a four CNN $_{(m,f)}$ architectures of different depth (*Fast*_$CNN_{(m,f)}$_n) was used, where n ϵ {2,4,6} is the number of convolutional layers are compared for age estimation task. All the convolutional layers were composed of the kernel size of 3x3 pixels and 2 max-pooling layers which helped reduce both height and width of the feature maps. To prevent convergence and overfitting, a batch normalization was employed and a 0:5

dropout module was used. Then we used equation (8) to concatenate feature vectors of ANN and CNN for the final gender and age output. To overcome the overfitting problem, we applied equation (15) the soft boundary technique of [4].

This has provided a wider range of feature's extraction which results into a reliable results of features during training process of the proposed algorithm. Since there is a very big relativity of feature dependencies, therefore considering both image subdivision and entire face feature extraction, it gives us enough room for high accuracy approximation during gender and age estimation of an image.

These images were trained according to their gender classes respectively as shown below.

$w_{cnn(f),0} - w_{cnn(f),R-1}$ for CNN_f and $w_{cnn(m),0} - w_{cnn(m),R-1}$ for CNN_m where ; R is the number of testers to be Recognized.

$w_{cnn(f),0} - w_{cnn(f),R-1}$ for CNN_f and $w_{cnn(m),0} - w_{cnn(m),R-1}$ for CNN_m refers to the weights of the Image confidence values of CNN_f .

Confidence values of the two convolution neural networks are presented as:

$c_{cnn(f),0} - c_{cnn(f),R-1}$ for CNN_f , and $c_{cnn(m),0} - c_{cnn(m),R-1}$ for CNN_m .

For the loss function, we used the equation

$$f(x) = \left(C(I,W)_{age} - L_{age}\right)^2 + \alpha ln\left(e^{-2c(I,W)_{gender}L_{gender}} + 1\right) \quad (6)$$

Where C hyper parameter for tuning (*I, W*) denotes the function of the networks. I is the input face image. W is the warping function used in both Artificial and Convolutional Neural Networks [66]. The subscripts "age", and "gender" denote the 2 dimensions of output. L is the 2 dimensional label of training set. $L_{gender} \in (-1,1)$, -1 denotes Male and 1 denotes Female. α is hyper-parameters to tune the importance of each term. Where α was set between 0.01 to 0.1.

In order to reduce the complexity of the network and training time, they only use the regression-based age estimation deep CNN. Euclidean loss function collectively known as Mean Absolute Error

$$E(w) = \frac{1}{2N} \sum_{i=1}^{N} \|\tilde{y}_n - y_n\|_2^2 \quad (7)$$

Where W is the parameter of neural networks $\tilde{y}_n$ is the age prediction value by the neural networks and y_n is the actual age value, and N is the batch number [61].

3.5. Decision Fusion of the Hybrid Neural Network

In pattern recognition, decision fusion is used as a classifier combination that enables us to achieve a better classification accuracy during the intended accuracy goals. Thus, we aimed to use the decision fusion method to fuse the decisions obtained from combined classifiers (CNN and ANN), known as hybrid neural network as mentioned in the previous sections. The key point of our contribution in this research based on the decisions for the classified labels and we fused the nearest decisions concerning their classifiers while enabling us to increase the classification and recognition accuracy rate of our experimental results. The figure III.6, below, presents our main concept of age classification with decision fusion

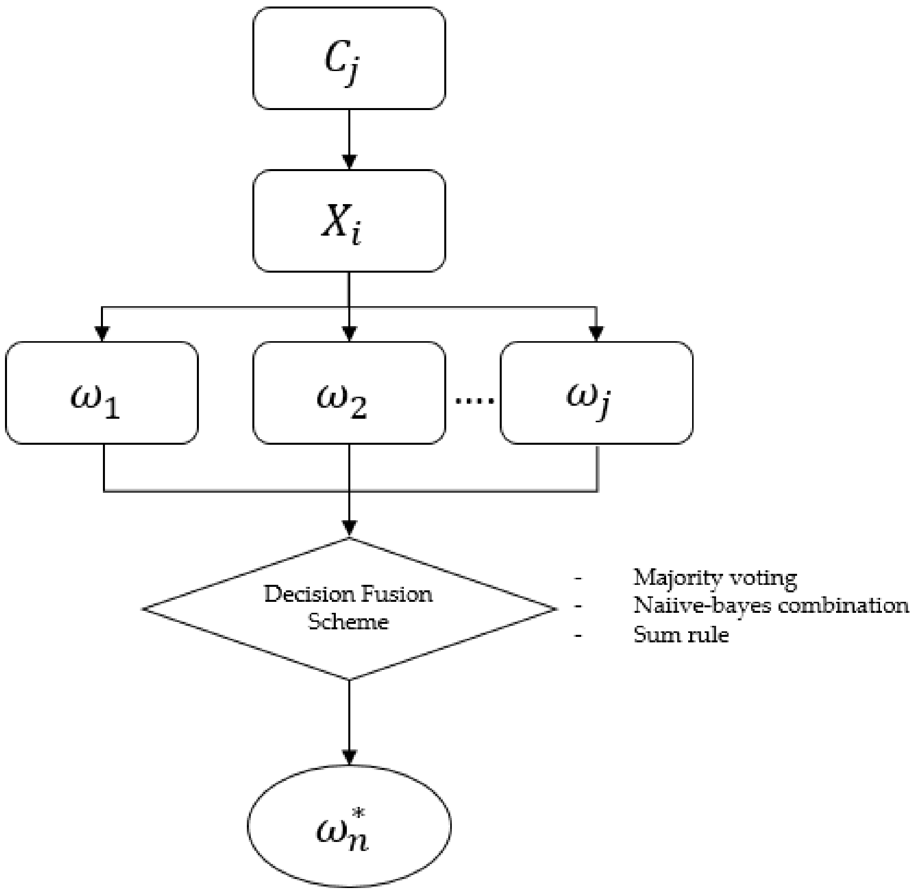

Figure III. 6. Block diagram of Decision Fusion.

In [71, 72, 73 and 28], they developed a theoretical framework of fusing decisions from multiple classifiers using schemes like the sum rule, product rule, max rule, min rule, median rule and majority voting.

3.6. Classifier Combination Strategies

Many commonly used classifier combination strategies can be developed from these rules such as product rules, sum roles, max and min rules, medium rules and majority voting rule.

Considering the posteriori probabilities yielded by the classifiers in equation (8) below,

$$p^{-(R-1)}(\omega_j)\prod_{i=1}^{R} P(\omega_j|x_i) = max_{k=1}^{m}\, p^{-(R-1)}(\omega_j)\prod_{i=1}^{R} P(\omega_n|x_i) \quad (8)$$

The decision rule (7) quantifies the likelihood of a hypothesis by combining the a posteriori probabilities generated by the individual classifiers by means of a product rule. It is effectively a severe rule of fusing the classifier outputs, as it is sufficient for a single recognition engine to inhibit a particular interpretation by outputting a close to zero probability for it.

In our research we used three decision fusion schemes which are majority voting, naive-bayes combination and the sum rule. We considered an input pattern Z into one of j possible class labels $(C_{1,} C_2 ,..., C_i, ... ,, C_n)$ where $j\epsilon[1 ..... , n]$. Let x_i be the i^th classifier which receives a vector of features from the input pattern Z, where $i\epsilon[1, .. , R]$. Therefore, the output of i^th classifier will be the decision ω_n^*. After the classifier's decisions respectively, we fused the decisions of the classifiers using the decision fusion schemes mentioned above with respect to posteriori probability (ω_j / x_i).

3.6.1. Majority Voting.

Majority voting commonly known as hard decision fusion is used when the classifier receives the highest number of votes. Where the ensemble chooses a class when (any one of the situations are considered):

(i) All classifiers agree on the specific class (unanimous voting),

(ii) Predicted by at least one more than half the number of classifiers (simple majority).

(iii) It receives the highest number of votes, whether or not the sum of those votes exceeds 50% (majority voting or plurality voting). The ensemble decision for the majority voting can be described as follows: choose class ω_j

Given that as $d_{i,j} = \begin{cases} 1 \; if \; P(\omega_n^*/\, x_i) = \; max_{j=1}^{n} P(\omega_j/\, x_i) \\ 0 \; otherwise \end{cases}$ (9)

With the assumption of equal priors and by hardening the probabilities according to Eq (9) we can conclude that Z is assigned to $\omega_j$when;

$$\sum_{i=1}^{R} d_{i,j} = max_{n=1}^{j} \sum_{i=1}^{R} d_{i,n} \quad (10)$$

Where for each classω_n, the sum of the right hand side of

Eq (10) simply counts the votes received by the individual classifiers. In our experiments we used two classifiers (i) where $i\epsilon[1,2]$ and we used three age class labels $j\epsilon[1,2 \; and \; 3]$. Thus, the class which receives the largest number of votes is then selected as the majority decision.

3.6.2. Naïve – Bayes Combination probabilistic decision fusion

This method takes the classifiers as mutually independent for given class labels [73]. Where ω_n is the n^{th}

Class label, $i = 1,..,R$ and $j = 1,..,n$, the Naïve – Bayes Combination decides the class using maximum likelihood and under the assumptions that Z is assigned to ω_j, the naïve-bayes combination is presented as follows;

$$max_{i=1}^{R} P(\omega_j / x_i) = max_{k=1}^{j} max_{i=1}^{R} P(\omega_n / x_i) \quad (11)$$

3.6.3. Sum rule decision fusion

As already mentioned above that $P(\omega_j / x_i)$ presents the expecting posteriori probability the sum rule can be used to computes the soft class label vectors using [74,28].

The sum rule can be presented by assigning Z to ω_j as follows;

Considering equation (8), in more detail. In some applications it may be appropriate further to assume that the a posteriori probabilities computed by the respective classifiers will not deviate dramatically from the prior probabilities. This is a rather strong assumption but it may be readily satisfied when the available observational discriminatory information is highly ambiguous due to high levels of noise. In such a situation we can assume that the a posteriori probabilities can be expressed as;

$$P(\omega_n|x_i) = P(\omega_n)(1+\delta_{ni}) \tag{12}$$

Where δ_{ni} satisfies $\delta_{ni} << 1$

Substituting (12) for the a posteriori probabilities in (8), we find;

$$p^{-(R-1)}(\omega_n)\prod_{i=1}^{R} P(\omega_n|x_i) = P(\omega_n)\prod_{i=1}^{R}(1+\delta_{ni}) \tag{13}$$

If we expand the product and neglect any terms of second and higher order, we can approximate the right-hand side of (13) as;

$$P(\omega_n)\prod_{i=1}^{R}(1+\delta_{ni}) = P(\omega_n) + P(\omega_n)\sum_{i=1}^{R}\delta_{ni} \tag{14}$$

Substituting (14) and (12) into (8), we obtain a sum decision rule

$$(1-R)P(\omega_j) + \sum_{i=1}^{R} P(\omega_j/x_i) = max_{i=1}^{R}\,[(1-R)P(\omega_n) + \sum_{i=1}^{R} P(\omega_n/x_i)] \tag{15}$$

For each test sample, the expert outputs are combined using the Sum rule and the resulting value compared against the decision threshold of 0.5.

IV.Experiment Results and Discussions

4.1. Databases

As shown in figure III.1, our research aim to provide an improved accurate rate of both gender and age accuracy using a combined algorithm of Gabor filters for Gender Recognition, ANN and CNN for age estimation hence getting the estimated accuracy of the final output. In our experiment, we used two commonly known public aged database, MORPH Album 2 whose sample images are shown in Figure IV.1

Figure Ⅳ. 1. Sample images of MORPH Album 2 datasets.

The MORPH Album 2 database contains a total of 55,000 pictures of 13,000 volunteers aged 16 to 77 years, 45,000 of which are used for network training, and the remaining 10,000 are used for testing [75].

We also used FG-NET whose sample images are shown in Figure Ⅳ .2

Figure IV. 2. Sample images of FG-Net aging datasets.

The FG-NET face database consists of 1002 images of 82 different individuals with different expressions, illumination, and attitude changes. Each one has 6 to 18 images of different ages, ranging from0 to 69 years old [60]

For experimental results we used our own private datasets to test the performance our proposed method as shown in Figure IV.3

Figure IV. 3. Sample images for our Private database.

In forming our database, we collected different images from the internet of 2000 different individuals, 1000 of these individual are women while the other 1000 are men, for each gender category, every category is composed of 3 age classes (young, old and adult). For the young age class is between 1 -24 years old which has 250 images per gender category, the adult age is between 25 – 49 years which has 450 images per gender category and the old age class is from 50 years and above which has 300 images per gender category. Generally in our database we used three common races or skin colors, we used 60 images with black skin colors (Africans/Negros), then we used 120 images for Asians race/ origin, the remaining 1820 images are for the white peoples with origin in Europe/ America and other surrounding countries.

4.2. Gender classification

In our experiments we used the Gabor filters to extract the local facial features [3], we used the lower (U_l) and upper frequencies (U_h) , and the radial frequency (W) was equal to the upper frequency. Our Gabor were set to 6 orientations (k) and the scale(s) of 4 which makes it 24 Gabor filters, then U_l = 0.025 and $U_h = 0.05$. We used these values to calculate the mean, standard deviation, root mean square value and finally these parameters helped us to form a one dimension feature vector of the extracted local features. An SVM classifier which is a commonly used classifier was used for gender classification.

4.3. Experimental results for gender classification

Our Experiments were done using Matlab, firstly we carried the experiment of gender recognition using Gabor filters as local features extractor and simple vector machine as classifier.

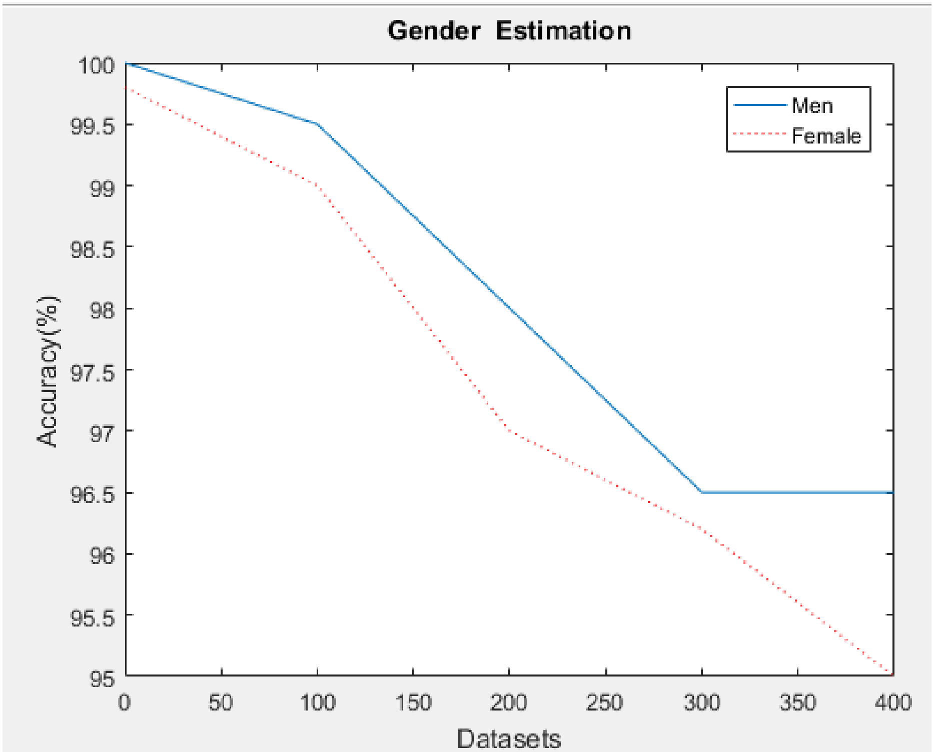

Figure IV. 4. Gender recognition using 400 images

Figure IV.4. Indicates the results of gender estimation both men and women, as it is indicated by the results, male recognition accuracy is high than that of women, this is due to a number factors such as make-ups, skin textures, aging conditions etc. also it have been seen that as the number of training datasets increases, there is a slight decrease in recognition accuracy.

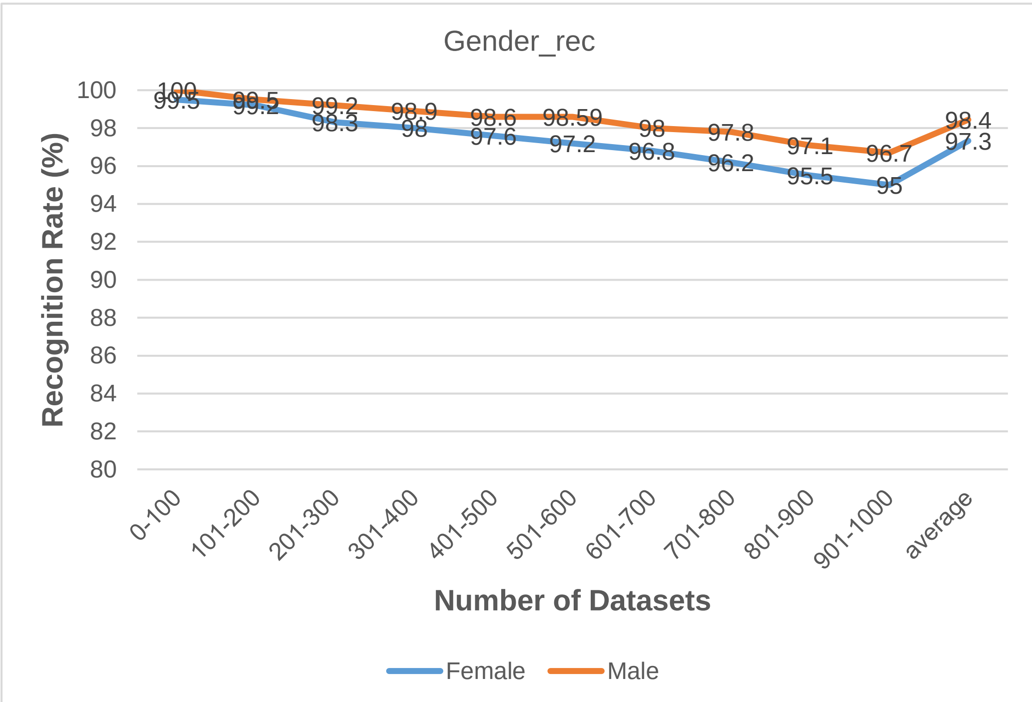

Figure IV. 5. Gender Recognition using 1000 images.

Figure IV.5. Show the average recognition accuracy of 98.4% of male and 97.3% of women which makes 97.8% overall gender recognition rate of our model. Compared to other state-of art for gender recognition, our proposal produce a low age accurate recognition rate, this is due to the fact that the images we used for our approach are aging images purposely produced for age estimation. Therefore during gender features extraction, it's very hard to get appropriate features more especially for the Adult and old age groups. as our target for this research is to produce an improved accuracy of the combination of both gender and age instead on focusing just on either gender or age recognition.

In table IV.1 we present the comparison of Gender and age classification of previous researchers to our proposed model. From our observation and results presented, our model provided a well improved accuracy rate of gender classification compared to the rest. For [77, 50, 79 and 31], we had the same research goal of both gender and age classification. Looking at their results gender classification, our model was far better that theirs. For [43, 45, 46, 24 and 25], their aim was to classify gender only. however even if they had only a single target of gender classification our model also provided better recognition accuracy compared to most them only [43] we had the same gender recognition accuracy of 98.7%.

Table IV. 1. Comparison of the proposed methods for gender recognition to the previous works

S/N	**Reference**	**year**	**Approach (Classifiers)**	**Database used & accuracy**		
				Morph II	**LFW**	**Private DB**
1	[77]	2009	Context/appearance	N/A	N/A	74.1%
2	[50]	2010	Gabor filters	75.7 %	N/A	N/A
3	[43]	2010	BIF +OLPP	98.7%	N/A	N/A
4	[45]	2012	LBP +SVM	N/A	N/A	94.8 %
5	[46]	2013	EEG+ SVM	N/A	N/A	97. 0 %
6	[79]	2014	LBP/SIFT/CH	N/A	N/A	90.1 %
7	[24]	2017	HF-RESNET	N/A	94.0 %	N/A
8	[25]	2018	GP-GAN	N/A	93.1 %	N/A
9	[31]	2018	GF + Wide CNN	N/A	N/A	88.9 %
10	**OURS**	**2019**	**SVM**	N/A	N/A	**98.7 %**

4.4. Age classification without Decision Fusion

We extracted the local features to classify gender, and these features were used by the conventional artificial neural network for the age classification output. As mention in section 3. We used two categories of conventional neural networks with respect to the gender categories. To overcome the overfitting problem, we applied Eq(15) the soft boundary technique of [4]. During our experiment of age classification, we formed 3 bins of the age group which is "Young", "Adult", and "Old".

During our experiment of age classification, we used several decision fusion techniques where among others we used simple fusion (majority voting) by applying Eq (10) for age class label classification with respect to a gender category. Then we apply both Combining Probabilistic (Soft) outputs and the Naïve – Bayes Combination for soft age class label fusion with respect to gender category by applying Eqs (11 and 15). Before using Decision fusion technique, we first used the ANN and CNN for age classification as presented in figure IV. 6 and figure IV.7, respectively.

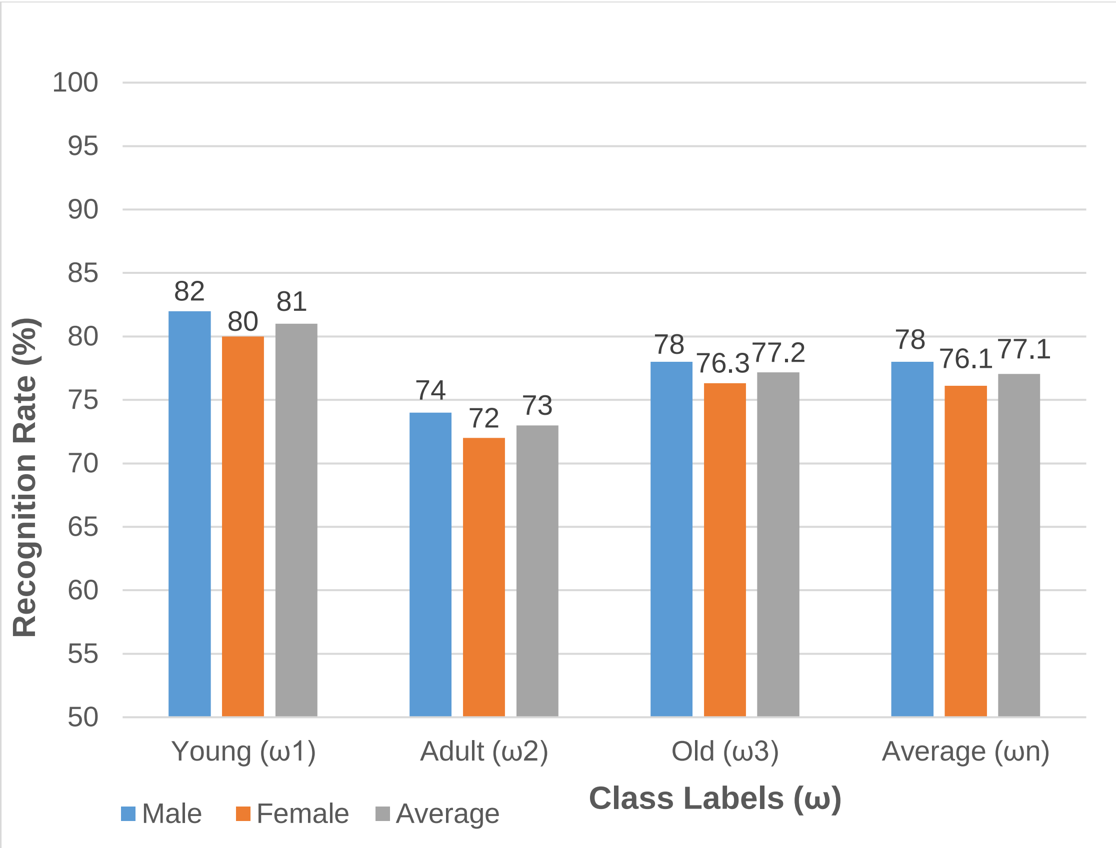

Figure IV. 6. Age Estimation by ANN only

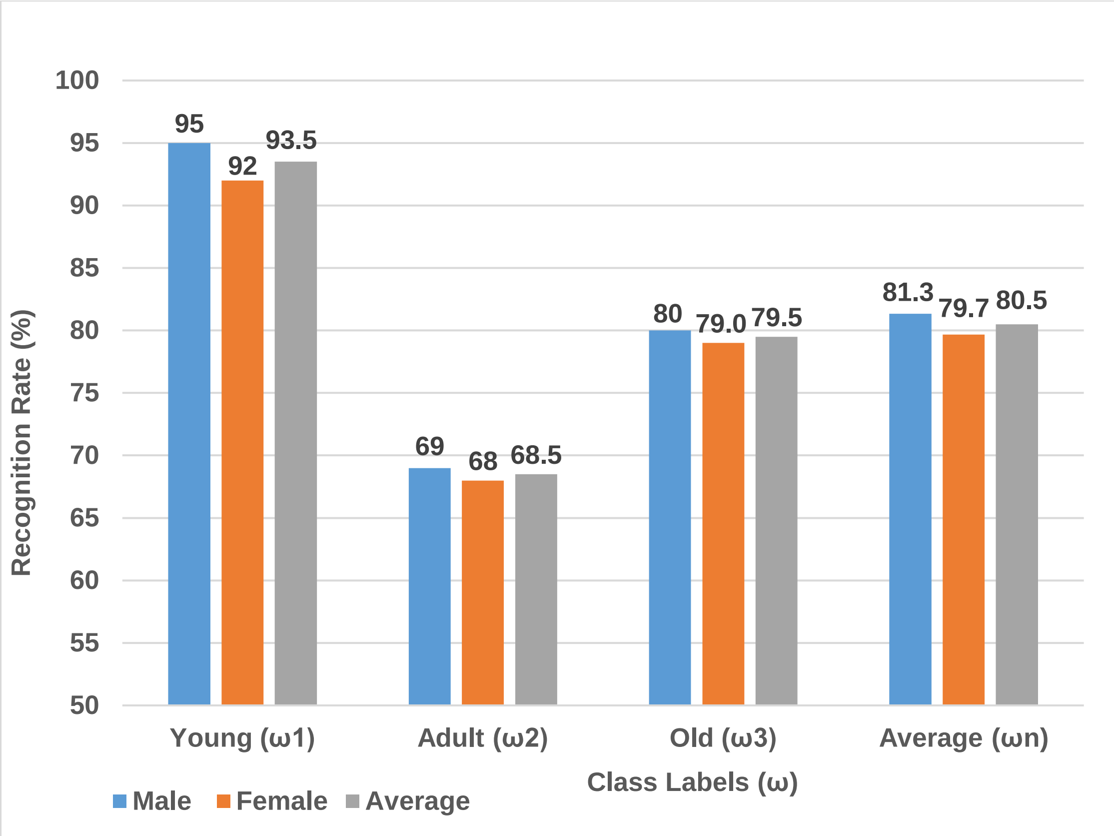

Figure IV. 7. Age estimation using CNN only

As presented in the results of figure IV.6 and figure IV.7, an average accuracy rate of 77.167% of estimation was achieved by using conventional artificial neural networks only, then an accurate rate of 80.5% of age classification was achieved by using convolutional neural networks only before applying the decision fusion the proposed algorithm. However, for the age classification of the adult label, ANN provided a higher accuracy compared to the CNN, this was because, for this age label, training the whole image does not provide enough feature due to the fact that, there is a lot of facial modifications due to very many factors like , usage of make-ups and other skin transformation factors. Therefore this makes the ANN best classifier since it uses a lot of local features. All these experiments were carried out on our private database.

4.5. Age Classification with Decision Fusion

We aim to use a hybrid approach of neural networks through decision fusion techniques, we applied a simple decision fusion (majority voting) and combining probabilistic (soft) outputs. An average accuracy rate of 81.733% of age classification after using majority voting as presented in figure IV.8.

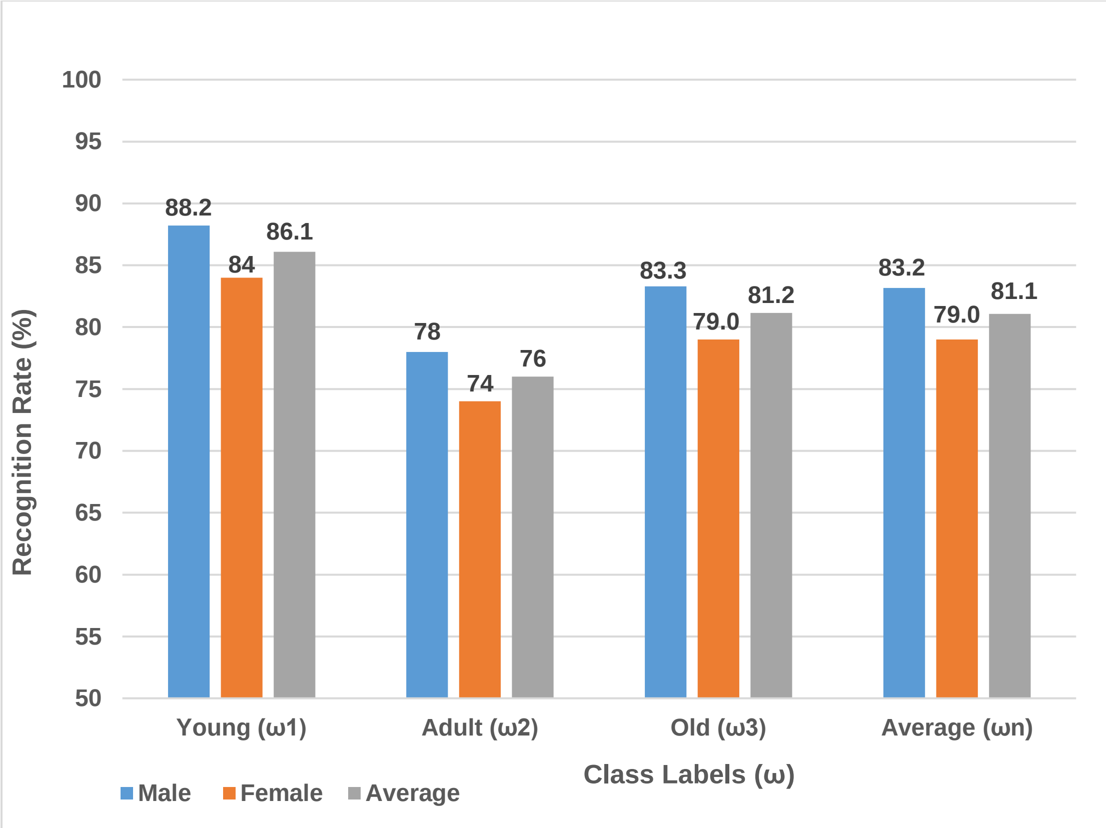

Figure IV. 8. Age classification using majority voting decision

As it is presented in the architectural diagram of our model in section 3, we first recognized the gender of the image and then we further used the convolution neural network for both female and male ($CNN_{(m,f)}$) respectively and artificial neural network for both male and female ($ANN_{(m,f)}$), respectively, for age estimation. According to results presented in figure IV.8, the young class label has the highest accuracy rate of both age recognition. This due to the fact that their skin texture is still smooth and original, then the old class has the second-highest estimation rate and the adult class presents the least accuracy rate. This is due to a lot of skin transforms that occurs during this age class due to usage of too much cosmetics, developing skin rashes, beards to the male gender and other factors. Also in all age classes the male gender possesses the highest age accuracy rate than females.

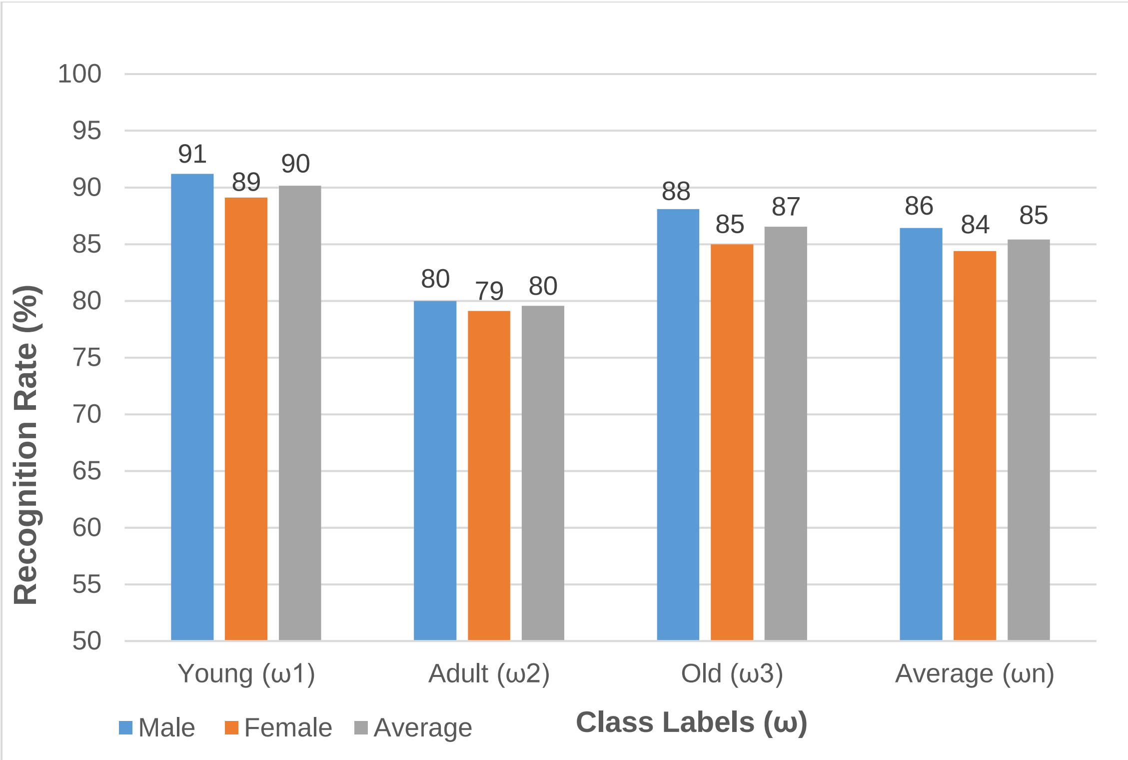

Figure IV. 9. Age classification using Naïve – Bayes Combination decision fusion

As shown by figure IV.9, we also used decision fusion of classified class ω by applying Naïve – Bayes Combination decision fusion presented in Eq (15). The Naïve – Bayes Combination decision fusion provides relatively fewer options for decisions to be fused compared to the during decision classification. However, the marginal error is not big as the results show the average recognition rate is 85.4% which is only 1% less than that of sum-rule decision fusion.

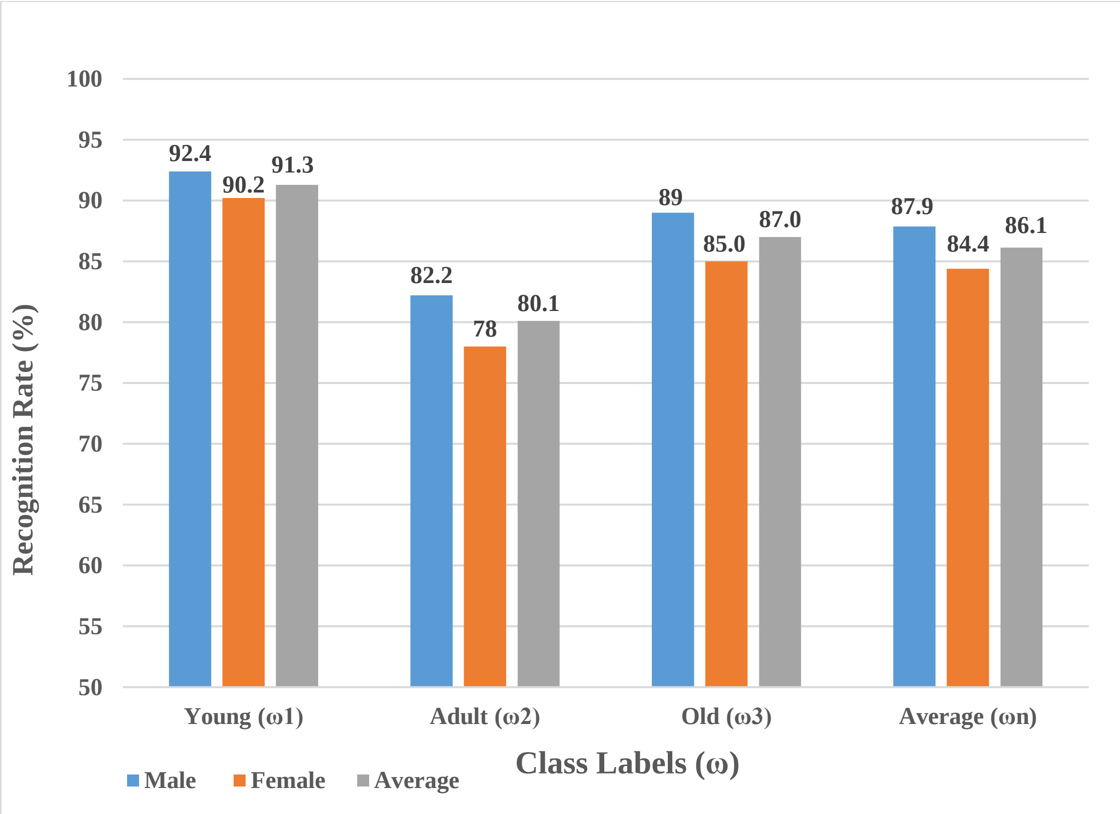

Figure IV. 10. Age classification using sum rule decision fusion.

Figure IV.10, shows a 5% improvement accuracy rate of age classification using sum-rule decision fusion compared to majority voting decision fusion. Sum-rule decision fusion is one of the probabilistic decision fusion methods. Then a soft decision fusion was applied to the recognized age classes which resulted in an average accuracy rate of 86.1%. Since using CNN only in age classification have indicated a remarkable overfitting of class labels. The ANN had helped us to use the skin texture and facial wrinkles during age classification. Therefore, combining these decisions of both neural networks using the sum rule have indicated good results in solving the overfitting problem encountered during age classification. This is because, the sum rule decision fusion provides enough selection of the nearest neighborhood decisions hence enabling the right decision of the classified age class.

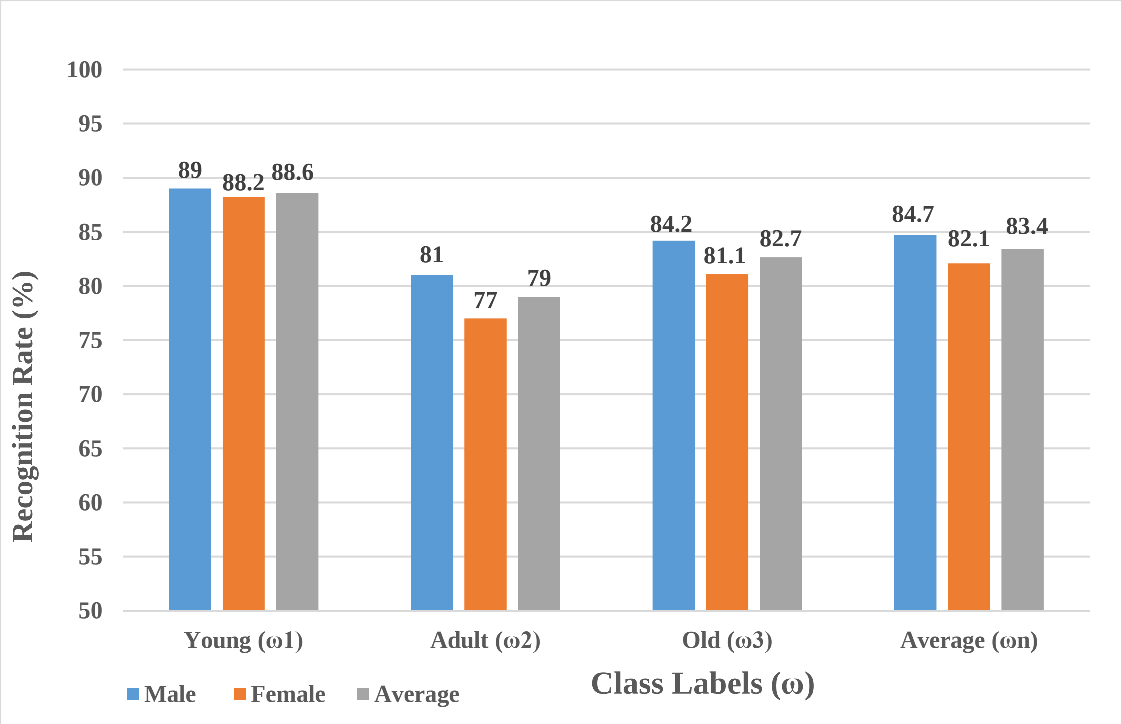

Figure IV. 11. Age classification using sum rule decision fusion for Adience landmark aging public Database.

Sum-rule decision fusion was used for the experiment which was carried out on a public database called FG-Net aging database. Our experiments were carried out on 6000 images of the FG-Net aging database. Each age category was presented by 2000 images and 1000 are female while 1000 are male. As results indicated in figure 10, an average accuracy rate of 84.6% age classification using our proposed model.

Table IV. 2. Method used in our proposed research model of age and Gender recognition.

Methods used in our research model			
S/N	**Methods**	**Recognition Accuracy (%)**	**Time(Microseconds)**
1	ANN Only	77	62.453219
2	CNN Only	80.5	39.246490
3	Majority Voting	81.2	69.253542
4	Naïve-Bayes	85.5	153.258971
5	Sum Rule	86.2	132.563892

Table IV.1. Presents all the techniques used in our research, we used different classification methods and decision fusions. As presented in the figure above, the probability decision fusion technique has shown an improvement of 5% compared to the voting decision fusion. And among the two probabilistic decision fusion, the Naïve – Bayes Combination decision fusion provided a 1% superior results compared to the sum rule of the decision profile probability technique.

Table IV. 3. Comparison of the proposed methods for age classification to the previous works.

S/N	References	year	Approach (Classifiers)	Database used & Accuracy			
				Morph II	FG-Net	Adience	Private DB
1	[78]	1999	Deformable templets	N/A	N/A	N/A	67.0 %
2	[77]	2009	Context/appearance	N/A	N/A	42 %	N/A
3	[40]	2009	AAM + SVM	N/A	81.0 %	N/A	N/A
4	[50]	2010	Gabor filters	50.3%	N/A	N/A	N/A
5	[79]	2014	LBP/SIFT/CH	N/A	N/A	N/A	55.88 %
6	[76]	2014	ANN	N/A	N/A	N/A	70.5 %
7	[75]	2014	GF + ULBP	N/A	85.0 %	N/A	N/A
8	[30]	2016	Ordinal CNN	67.3 %	N/A	N/A	N/A
9	[31]	2018	GF + Wide CNN	N/A	N/A	61.3%	N/A
10	**OURS**	**2019**	**ANN + CNN**	N/A	N/A	**83.5%**	**86.1 %**

The average estimation percentage accuracy rate of the previous methods and our proposed method is presented in table IV.2. According the results presented, our proposed method using a combined algorithm of Gabor filters, Conventional artificial neural networks and Convolution neural networks with decision fusion have performed well compared to the presented state of art. Compared to other models, we have achieved a higher accuracy rate than there results, for estimation of both gender and age. For researches presented in [31, 50, 77 and 79], we all had the same intension of classifying both the gender and age of facial images. Compared to their results, our method achieved a higher accuracy rate of both gender and age. Therefore, we are fully convinced that our proposal performed well compared to existing state of art.

I also made a comparison of our developed model to those of the facial recognition developed tools which includes Face++ [82] and MS cognitive Services [83] for age and gender classification, the results of face Face++ and MS Cognitive Services were adopted from [81]. Our results a far better than those of recognition developed tools as presented in the table IV.3 below.

Table IV. 4. Comparative experimental results of our proposed model to the developed tools for age and gender classification.

S/N	Model	Gender recognition (%)	Age Classification (%)
1	Our Model	97.5	86.1
2	Face++ [82]	91.1	38.8
3	MS Cognitive Services (Azure) [83]	92.9	59.3

As results presented in table IV.4 above indicates, the recognition accuracy of our model for both gender and age classification is very high compare to those of recognition tools developed to facial image recognitions.

V. Conclusion and Recommendation

In the research, we propose an age and gender estimation model using a combined algorithm of Gabor, ANN and CNN with decision fusion. We first extracted the local features of the images using the Gabor filters and we based on those local features to recognize the gender, where gender was recognized to an accuracy rate of 98.4% as presented in the results of Figure IV.5, we further continued with our experiment after knowing the gender of the image to collectively estimate both gender and age as primary motive of our research, we took the extracted feature by the Gabor filters and pass then to the conventional artificial neural networks and also training a whole cropped images of the passport 32x32 size using the Convolutional neural networks with respect to their gander classes, then combining together these Neural networks, we managed to estimate the age in the images with their respective gender classes. The age was grouped into three major bins (Young, Adult and old) accordingly. We used the decision fusion technique by applying majority voting decision fusion, Naïve – Bayes Combination decision fusion and decision profile (sum rule). As results indicates, Naïve – Bayes Combination decision fusion provide the highest age and gender recognition accuracy rate of 86.133% compared to other decision fusion techniques Figure IV.10. Our model have achieved a promising estimation result of 86.133% age and gender accuracy rate compared to the state of art presented in table IV.2.

In our feature research we aim to expand our model and include in other facial image while extracting the local features under various conditions for recognition research fields like human race and facial emotions.

References

[1] Abdulalla, F.Q. and Shaker, S.H., 2018. A Survey of human face detection methods. Journal of Al-Qadisiyah for computer science and mathematics, 10(2), pp.Page-108.

[2] Huang, Chung-Lin, and Ching-Wen Chen. "Human facial feature extraction for face interpretation and recognition." Pattern recognition 25, no. 12 (1992): 1435-1444.

[3] Choi, S. E., Lee, Y. J., Lee, S. J., Park, K. R., & Kim, J. (2011). Age estimation using a hierarchical classifier based on global and local facial features. Pattern Recognition, 44(6), 1262-1281.

[4] Pontes, J.K., Britto Jr, A.S., Fookes, C. and Koerich, A.L., 2016. A flexible hierarchical approach for facial age estimation based on multiple features. Pattern Recognition, 54, pp.34-51.

[5] Gao, F. and Ai, H., 2009, June. Face age classification on consumer images with gabor feature and fuzzy lda method. In International Conference on Biometrics (pp. 132-141). Springer, Berlin, Heidelberg

[6] Lian, H.C., Lu, B.L., Takikawa, E. and Hosoi, S., 2005, August. Gender recognition using a min-max modular support vector machine. In International Conference on Natural Computation (pp. 438-441). Springer, Berlin, Heidelberg.

[7] F. Mahmud, M. T. Khatun, S. T. Zuhori, S. Afroge, M. Aktar and B.Pal, "Face recognition using Principle Component Analysis and Linear Discriminant Analysis", 2015 International Conference on Electrical Engineering and Information Communication Technology (ICEEICT), Dhaka, pp. 1-4, 2015.

[8] Mahmud, M. E. Haque, S. T. Zuhori and B. Pal, "Human face recognition using PCA based Genetic Algorithm", 2014 International Conference on Electrical Engineering and Information & Communication Technology, Dhaka, pp. 1-5, 2014.

[9] Zhang, B., Liu, G. and Xie, G., 2016, October. Facial expression recognition using LBP and LPQ based on Gabor wavelet transform. In 2016 2nd IEEE International Conference on Computer and Communications (ICCC) (pp. 365-369). IEEE.

[10] Islam, B., Mahmud, F., Hossain, A., Mia, M.S. and Goala, P.B., 2018, September. Human Facial Expression Recognition System Using Artificial Neural Network Classification of Gabor Feature Based Facial Expression Information. In 2018 4th International Conference on Electrical Engineering and Information & Communication Technology (iCEEiCT) (pp. 364-368). IEEE.

[11] Liu, Z.T., Sui, G.T., Li, D.Y. and Tan, G.Z., 2015, July. A novel facial expression recognition method based on extreme learning machine. In 2015 34th Chinese Control Conference (CCC) (pp. 3852-3857). IEEE

[12] Y. Xing and W. Luo, "Facial expression recognition using local Gabor features and adaboost classifiers," 2016 International Conference on Progress in Informatics and Computing (PIC), Shanghai, pp. 228-232, 2016.

[13] Jia, J., Xu, Y., Zhang, S. and Xue, X., 2016, August. The facial expression recognition method of random forest based on improved PCA extracting feature. In 2016 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC) (pp. 1-5). IEEE.

[14] Kohail, S.N., 2012, March. Using artificial neural network for human age estimation based on facial images. In 2012 International Conference on Innovations in Information Technology (IIT) (pp. 215-219). IEEE.

[15] Verma, K. and Khunteta, A., 2017, August. Facial expression recognition using Gabor filter and multi-layer artificial neural network. In 2017 International Conference on Information, Communication, Instrumentation and Control (ICICIC) (pp. 1-5). IEEE

[16] Jafri, R., & Arabnia, H. R. (2009). A survey of face recognition techniques. Jips, 5(2), 41-68.

[17] Ng, C.B., Tay, Y.H. and Goi, B.M., 2015. A review of facial gender recognition. Pattern Analysis and Applications, 18(4), pp.739-755.

[18] Yu, S., Tan, T., Huang, K., Jia, K., & Wu, X. (2009). A study on gait-based gender classification. IEEE Transactions on image processing, 18(8), 1905-1910.

[19] Moghaddam, B., & Yang, M. H. (2000, March). Gender classification with support vector machines. In Proceedings Fourth IEEE International Conference on Automatic Face

and Gesture Recognition (Cat. No. PR00580) (pp. 306-311). IEEE.

[20] Makinen, E., & Raisamo, R. (2008). Evaluation of gender classification methods with automatically detected and aligned faces. IEEE Transactions on Pattern Analysis and Machine Intelligence, 30(3), 541-547.

[21] Sun, Z., Bebis, G., Yuan, X., & Louis, S. J. (2002, December). Genetic feature subset selection for gender classification: A comparison study. In Sixth IEEE Workshop on Applications of Computer Vision, 2002.(WACV 2002). Proceedings. (pp. 165-170). IEEE.

[22] Golomb, B. A., Lawrence, D. T., & Sejnowski, T. J. (1990, October). Sexnet: A neural network identifies sex from human faces. In NIPS (Vol. 1, p. 2).

[23] Boser, B. E., Guyon, I. M., & Vapnik, V. N. (1992, July). A training algorithm for optimal margin classifiers. In Proceedings of the fifth annual workshop on Computational learning theory (pp. 144-152). ACM.

[24] Ranjan, R., Patel, V. M., & Chellappa, R. (2017). Hyperface: A deep multi-task learning framework for face detection,

landmark localization, pose estimation, and gender recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 41(1), 121-135.

[25] Di, X., Sindagi, V. A., & Patel, V. M. (2018, August). Gp-gan: Gender preserving gan for synthesizing faces from landmarks. In 2018 24th International Conference on Pattern Recognition (ICPR) (pp. 1079-1084). IEEE.

[26] Andreu, Y., García-Sevilla, P. and Mollineda, R.A., 2014. Face gender classification: A statistical study when neutral and distorted faces are combined for training and testing purposes. Image and Vision Computing, 32(1), pp.27-36.

[27] Lu, H. and Lin, H., 2007, August. Gender recognition using adaboosted feature. In Third International Conference on Natural Computation (ICNC 2007) (Vol. 2, pp. 646-650). IEEE.[28]. Kittler, Josef, Mohamad Hatef, Robert PW Duin, and Jiri Matas. "On combining classifiers." IEEE transactions on pattern analysis and machine intelligence 20, no. 3 (1998): 226-239.

[29] Luo, Haowen, Taorong Qiu, Chao Liu, and Peifan Huang. "Research on fatigue driving detection using forehead EEG

based on adaptive multi-scale entropy." Biomedical Signal Processing and Control 51 (2019): 50-58.

[30] Niu, Z., Zhou, M., Wang, L., Gao, X., & Hua, G. (2016). Ordinal regression with multiple output cnn for age estimation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4920-4928).

[31] Hosseini, S., Lee, S.H., Kwon, H.J., Koo, H.I. and Cho, N.I., 2018, January. Age and gender classification using wide convolutional neural network and Gabor filter. In 2018 International Workshop on Advanced Image Technology (IWAIT) (pp. 1-3). IEEE.

[32] Levi, G. and Hassner, T., 2015. Age and gender classification using convolutional neural networks. In Proceedings of the iEEE conference on computer vision and pattern recognition workshops (pp. 34-42).[33]. Horng, Wen-Bing, Cheng-Ping Lee, and Chun-Wen Chen. "Classification of age groups based on facial features." no. 3 (2001): 183-192.

[34] Geng, X., Zhou, Z. H., & Smith-Miles, K. (2007). Automatic age estimation based on facial aging patterns. IEEE

Transactions on pattern analysis and machine intelligence, 29(12), 2234-2240.

[35] Hough, P. V. (1959). Machine analysis of bubble chamber pictures. In Conf. Proc. (Vol. 590914, pp. 554-558).

[36] Duda, R. O., & Hart, P. E. (1971). Use of the Hough transformation to detect lines and curves in pictures (No. SRI-TN-36). Sri International Menlo Park Ca Artificial Intelligence Center.

[37] Chiam, T. C. (2012). Age classification from facial images for detecting retinoblastoma (Doctoral dissertation).

[38] Ng, C. B., Tay, Y. H., & Goi, B. M. (2012). Vision-based human gender recognition: A survey. arXiv preprint arXiv:1204.1611.

[39] Wiggins, M., Saad, A., Litt, B., & Vachtsevanos, G. (2008). Evolving a Bayesian classifier for ECG-based age classification in medical applications. Applied soft computing, 8(1), 599-608.

[40] Luu, K., Ricanek, K., Bui, T. D., & Suen, C. Y. (2009, September). Age estimation using active appearance models and support vector machine regression. In 2009

IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems (pp. 1-5). IEEE.

[41] Zhang, Y., & Yeung, D. Y. (2010, June). Multi-task warped gaussian process for personalized age estimation. In 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (pp. 2622-2629). IEEE.

[42] Liu, H., Lu, J., Feng, J., & Zhou, J. (2017). Group-aware deep feature learning for facial age estimation. Pattern Recognition, 66, 82-94.

[43] Guo, G., & Mu, G. (2010, June). Human age estimation: What is the influence across race and gender? In 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops (pp. 71-78). IEEE.

[44] Cottrell, G.W. and Metcalfe, J., 1991. EMPATH: Face, emotion, and gender recognition using holons. In Advances in neural information processing systems (pp. 564-571).

[45] Shan, C. (2012). Learning local binary patterns for gender classification on real-world face images. Pattern recognition letters, 33(4), 431-437.

[46] Nguyen, P., Tran, D., Huang, X., & Ma, W. (2013, November). Age and gender classification using EEG paralinguistic features. In 2013 6th International IEEE/EMBS Conference on Neural Engineering (NER) (pp. 1295-1298). IEEE.

[47] Duan, Mingxing, Kenli Li, Canqun Yang, and Keqin Li. "A hybrid deep learning CNN–ELM for age and gender classification." Neurocomputing 275 (2018): 448-461.

[48] Chen, Y., Zheng, W. S., Xu, X. H., & Lai, J. H. (2013). Discriminant subspace learning constrained by locally statistical uncorrelation for face recognition. Neural Networks, 42, 28-43.

[49] Sikka, Karan, Tingfan Wu, Josh Susskind, and Marian Bartlett. "Exploring bag of words architectures in the facial expression domain." In European Conference on Computer Vision, pp. 250-259. Springer, Berlin, Heidelberg, 2012

[50] Shan, C. (2010, October). Learning local features for age estimation on real-life faces. In Proceedings of the 1st ACM international workshop on Multimodal pervasive video analysis (pp. 23-28). ACM.

[51] Levi, G., & Hassner, T. (2015). Age and gender classification using convolutional neural networks. In Proceedings of the iEEE conference on computer vision and pattern recognition workshops (pp. 34-42).

[52] Zhang, Y., & Xu, T. (2018). Landmark-Guided Local Deep Neural Networks for Age and Gender Classification. Journal of Sensors, 2018.

[53] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).

[54] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

[55] R. Rothe, R. Timofte, L. Van Gool, Deep expectation of real and apparent age from a single image without facial landmarks, IJCV (2016).

[56] Zhu, Y., Li, Y., Mu, G., & Guo, G. (2015). A study on apparent age estimation. In Proceedings of the IEEE International Conference on Computer Vision Workshops (pp. 25-31).

[57] Ozbulak, G., Aytar, Y., & Ekenel, H. K. (2016, September). How transferable are CNN-based features for age and gender classification?. In 2016 International Conference of the Biometrics Special Interest Group (BIOSIG) (pp. 1-6). IEEE.

[58]Yi, D., Lei, Z., & Li, S. Z. (2014, November). Age estimation by multi-scale convolutional network. In Asian conference on computer vision (pp. 144-158). Springer, Cham.

[59] Yang, H. F., Lin, B. Y., Chang, K. Y., & Chen, C. S. (2013). Automatic age estimation from face images via deep ranking. Networks, 35(8), 1872-1886.

[60] Jain, Anil K., Karthik Nandakumar, and Arun Ross. "50 years of biometric research: Accomplishments, challenges, and opportunities." Pattern Recognition Letters 79 (2016): 80-105.

[61] Liao, Haibin, et al. "Age Estimation of Face Images Based on CNN and Divide-and-Rule Strategy." Mathematical Problems in Engineering 2018 (2018).

[62] Lopes, A.T., de Aguiar, E., De Souza, A.F. and Oliveira-Santos, T., 2017. Facial expression recognition with

convolutional neural networks: coping with few data and the training sample order. Pattern Recognition, 61, pp.610-628

[63] Chitaliya, N.G. and Trivedi, A.I., 2010, March. Feature extraction using wavelet-pca and neural network for application of object classification & face recognition. In 2010 Second International Conference on Computer Engineering and Applications (Vol. 1, pp. 510-514). IEEE.

[64] Mäkinen, Erno, and Roope Raisamo. "An experimental comparison of gender classification methods." pattern recognition letters 29.10 (2008): 1544-1556.

[65] Barani, Milad Jafari, Karim Faez, and Fooad Jalili. "Implementation of gabor filters combined with binary features for gender recognition." Int. J. Electr. Comput. Eng 4 (2014): 108-115.

[66] Ren, Haoyu, and Ze-Nian Li. "Gender recognition using complexity-aware local features." 2014 22nd International Conference on Pattern Recognition. IEEE, 2014.

[67] Tapia, J. E., & Perez, C. A. (2013). Gender classification based on fusion of different spatial scale features selected

by mutual information from histogram of LBP, intensity, and shape. IEEE transactions on information forensics and security, 8(3), 488-499.

[68] Fu, Yun, Guodong Guo, and Thomas S. Huang. "Age synthesis and estimation via faces: A survey." IEEE transactions on pattern analysis and machine intelligence 32, no. 11 (2010): 1955-1976.

[69] Yang, H. F., Lin, B. Y., Chang, K. Y., & Chen, C. S. (2013). Automatic age estimation from face images via deep ranking. networks, 35(8), 1872-1886.

[70] Antipov, G., Baccouche, M., Berrani, S.A. and Dugelay, J.L., 2017. Effective training of convolutional neural networks for face-based gender and age prediction. Pattern Recognition, 72, pp.15-26

[71] Mangai, Utthara Gosa, Suranjana Samanta, Sukhendu Das, and Pinaki Roy Chowdhury. "A survey of decision fusion and feature fusion strategies for pattern classification." IETE Technical review 27, no. 4 (2010): 293-307

[72] Giannakakis, G., Matthew Pediaditis, Dimitris Manousos, Eleni Kazantzaki, Franco Chiarugi, Panagiotis G. Simos,

Kostas Marias, and Manolis Tsiknakis. "Stress and anxiety detection using facial cues from videos." Biomedical Signal Processing and Control 31 (2017): 89-101.

[73] Kuncheva, Ludmila I. Combining pattern classifiers: methods and algorithms. John Wiley & Sons, 2014

[74] Choi, Young-Jae, In-Sik Choi, and Dae-young Chae. "Decision-level Fusion Scheme of SVM and naive Bayes Classifier for Radar Target Recognition." In 2018 International Symposium on Antennas and Propagation (ISAP), pp. 1-2. IEEE, 2018.

[75] Panis, G. and Lanitis, A., 2014, September. An overview of research activities in facial age estimation using the FG-NET aging database. In European Conference on Computer Vision (pp. 737-750). Springer, Cham.

[76] Kalansuriya, T.R. and Dharmaratne, A.T., 2014. Neural network based age and gender classification for facial images. ICTer, 7(2)

[77] Gallagher, A. C., & Chen, T. (2009, June). Understanding images of groups of people. In 2009 IEEE Conference on

Computer Vision and Pattern Recognition (pp. 256-263). IEEE.

[78] Kwon, Y. H., & da Vitoria Lobo, N. (1999). Age classification from facial images. Computer vision and image understanding, 74(1), 1-21.

[79] Fazl-Ersi, E., Mousa-Pasandi, M. E., Laganiere, R., & Awad, M. (2014, October). Age and gender recognition using informative features of various types. In 2014 IEEE International Conference on Image Processing (ICIP) (pp. 5891-5895). IEEE.

[80] Jung, S. G., An, J., Kwak, H., Salminen, J., & Jansen, B. J. (2018, June). Assessing the accuracy of four popular face recognition tools for inferring gender, age, and race. In Twelfth International AAAI Conference on Web and Social Media.

[81] Vasileiadis, M., Stavropoulos, G., & Tzovaras, D. (2019). Facial Soft Biometrics Detection on Low Power Devices. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops.

[82] Megvii Technology. Face++ cognitive services. https://www.faceplusplus.com. 5, 6

[83] Microsoft. Microsoft azure cognitive services. https://azure.microsoft.com/en-us/services/cognitive-services/. 6

국문초록

연령과 성별을 이용한 사용자 인식은 보안 시스템, 비디오 기반 감시 시스템, 온라인 결제 시스템, 재판 과정 및 적정 의약품 처방 등 다수의 Human and Computer Interaction(HCI) 분야에서 중요한 분야로 대두되고 있다. 최근의 얼굴 특징 분석에 기반한 연령 및 성별 추정 연구는 많은 HCI 분야 연구원들의 도전 과제로써 입지가 분명하다. 본 논문은 연령 및 성별 분류를 위한 융합 결정 기술을 이용해 Conventional Artificial Neural Networks(C-ANN)과 Convolution Neural Networks(CNN)을 접목한 알고리즘의 제시 내용에 중점을 두고 있다. 연구의 차별성은 C-ANN 과 CNN 두 종류 신경망의 결정 과정 융합을 통해 연령 및 성별 추정의 정확도 향상을 이루어 냈음에 있다. 사용자 인식 정확도 향상을 위해 사용되는 3 가지 결정 융합 방법인 Majority Voting, Naïve – Bayes Combination 그리고 Sum Rule 을 비교한 결과 Sum Rule 이 결정 분류 과정에서 다른 2 가지 방법에 비해 인접 클래스들의 likelihood 를 다수 감소시켜 최대 86.133% 분류 정확도를 보여, 기존 연구인 State of Art 방법 대비 향상되었다.

Acknowledgements

I would like to take this opportunity to thank all those who have contributed to this Thesis. First and foremost, my deepest gratitude goes to my Supervisor Prof. TaeYong Kim, who gave me the opportunity to pursue this research. I deeply appreciate his overwhelming guidance, encouragement, and most importantly the most valuable helpful discussions he offered me during this research.

I would like also to extend my gratitude to the GSAIM Professors who contributed a lot to my education career. Their efforts are highly appreciated.

Thanks to my colleagues in the Gametech Lab for being so helpful and friendly to me.

My appreciation and respect goes to my Wife, Our children and the extended family for their endless encouragements, love and affection.

Sincerely yours,

James Rwigema