Mohamed Khider University of Biskra
Faculty of Science and Technology
Department of Electrical Engineering

MASTER DISSERTATION

science and technology
telecommunication
Networks and Telecommunication

Réf. : Entrez la référence du document

Submitted and Defended by:
Mr.Ammari Ahmed

In : dimanche 27 septembre 2020

Facial age estimation using Multidimensional TEDA method

Board of Examiners :

Mr.	DEBILOU Abderrazak	Pr	University of Biskra	Président
Mr.	BAARIR Zine-Eddine	Pr	University of Biskra	Examinateur
Mr.	OUAMANE Abdelmalik	MCA	University of Biskra	Supervisor

Academic year: 2019 - 2020

Mohamed Khider Biskra University

Faculty of Sciences and Technology

Electrical Engineering Department

Field: Telecommunication

Option: Networks and Telecommunication

Graduation Dissertation

for obtaining a Master's degree

Theme

Facial age estimation using Multidimensional TEDA method

Presented by: Favorable opinion of the supervisor:

Ammari Ahmed   Dr.Ouamane Abdelmalik

Positive opinion of the President of the Jury

Mr. Debilou abderrazak

Stamp and signature

Table of Contents

Chapter 1: Basic concepts and terminology for facial age estimation

Chapter 2: Multidimensional analysis and TEDA

Chapter 3: System design and results

List of Figures

List of Tables

List of Abbreviations and acronyms

TEDA = Tensor Exponential Discriminant Analysis

PAL = Productive Aging Lab Face

AA = Actual Age

PA = Perceived Age

EA = Estimated Age

ECRM = Electronic Customer Relationship Management

ICU = Intensive Care Unit

MAE = Mean Absolute Error

CS = Cumulative Score

AAMs = Active Appearance Models

AAM = Active Appearance Model

PCA =Principal Component Analysis

AGES = AGing pattErn Subspace

LBP = Local Binary Patterns

BIF = Biologically Inspired Features

NN = Nearest Neighbor

SVMs = Support Vector Machines

ANNs = Artificial Neural Networks

2DLDA = Two Dimensional Linear Discriminant Analysis

LDA = Linear Discriminant Analysis

EM = Expectation Maximization

SVR = Support Vector Regression

PLS = Partial Least Squares

CCA = Canonical Correlation Analysis

IBR = Image Based Regression

LARR = Locally Adjusted Robust Regressor

2D = 2 Dimension

3D = 3 Dimension

MSL = Multilinear Subspace Learning

VVP = Vector-to-Vector Projection

TTP = Tensor-to-Tensor Projection

TVP = Tensor-to-Vector Projection

EDA = Exponential Discriminant Analysis

List of Abbreviations and acronyms

SSS = Small Sample Size

MDA = Multilinear Discriminant Analysis

WCCN = Within Class Covariance Normalization

CNN = Convolutional Neural Network

ERT = Ensemble of Regression Trees

VGG = Visual Geometry Group

ILSVRC =The ImageNet Large Scale Visual Recognition Challenge

LRN = Local Response Normalization

ReLU = Rectification Layer

DEX = Deep Expectation

GHPF = Gaussian High Pass Filter

CAM = Contourlet Appearance Model

BSIF = Binarized Statistical Image Features

DRF = Deep Random Forest

Abstract:

Human age, as an important personal trait, can be directly inferred by distinct patterns emerging from the facial appearance. Derived from rapid advances in computer graphics and machine vision, computer-based age estimation via faces have become particularly prevalent topics recently because of their emerging real-world applications.

For many practical applications, relying on humans to supply demographic information from face images is not feasible. Hence, there has been a growing interest in the automatic extraction of demographic information from face images. Here we focus on age estimation, whose objective is to determine the specific age or age range of a subject based on a facial image.

Our work confronted the challenge of age estimation using a multidimensional analysis (TEDA method), to analyze the natural images for the purpose of facial age estimation.

We are testing facial age estimation application on the PAL database.

Keywords : *biometrics, soft biometrics, multilinear approach, facial age estimation.*

ملخص :

يمكن استنتاج عمر الإنسان، باعتباره سمة شخصية مهمة، بشكل مباشر من خلال أنماط متميزة ناشئة عن مظهر الوجه. ونتيجة للتقدم السريع في رسومات الكمبيوتر ورؤية الآلة، أصبح تقدير العمر استناداً إلى الكمبيوتر عبر الوجوه موضوعاً شائعاً بشكل خاص مؤخراً بسبب التطبيقات الناشئة في العالم الحقيقي.

وبالنسبة للعديد من التطبيقات العملية، فإن الاعتماد على البشر في توفير المعلومات الديموغرافية من الصور المباشرة ليس بالأمر الممكن. ومن ثم، كان هناك اهتمام متزايد بالاستخراج التلقائي للمعلومات الديموغرافية من الصور المباشرة. وهنا نركز على تقدير العمر، الذي يتمثل هدفه في تحديد العمر أو الفئة العمرية لشخص ما بناءً على صورة الوجه.

وقد واجه عملنا تحدي تقدير العمر باستخدام أسلوب تحليل متعدد الأبعاد (طريقة TEDA)، لتحليل الصور الطبيعية لغرض تقدير عمر الوجه.

نحن نختبر تطبيق تقدير عمر الوجه على قاعدة بيانات (PAL).

الكلمات المفتاحية : *القياسات الحيوية، القياسات الحيوية اللينة، تحليل متعدد الأبعاد ، تقدير عمر الوجه.*

General Introduction

Age is a human attribute which grows alongside an individual. Estimating human age is quite difficult for a machine as well as humans; however, there has been and are still ongoing efforts towards the machine estimation of human age to a high level of accuracy. Facial age estimation can be defined as the task of automatically assigning an exact age label (or age range) to an individual facial image [1] [2]. This age can be either actual age, appearance age, perceived age, or estimated age [2] [3] [4]. The goal of the automatic age estimation is to judge if age is as close to actual or appearance age as possible [5]. Usually, humans instinctively guess or predict an individual's age from his/her face and this human ability has been observed to be innate and possessed early in life. Therefore, in making the computer predict human age, the assumption is that the facial image of an individual gives sufficient ageing information about such individual. This assumption has been long established as a fact from previous age estimation algorithms which employed the facial image as the primary input.

In humans, the accuracy of a predicted age depends on (among several other factors) the experience and exposure of the individual who is predicting the age, for instance an individual who works with a criminal investigation agency might predict human ages better than a school teacher simply because of the differences in their trainings and frequent interactions and experiences. For a machine, however, the task is somewhat more difficult as ageing is affected by several intrinsic factors (gender, race, heredity, etc.) as well as extrinsic factors (weather, drugs, condition of living etc.). In addition, the temporal nature of ageing and the fact that ageing patterns are individualistic also contributing to the difficulty of age estimation as these have made it difficult to gather facial ageing dataset suitable enough for tackling the problem.

Machine Learning has been defined as an automated computing procedure based on logical or binary operation which learns a task for a series of examples. It often involves enabling the computer to automatically perform some tasks by training it with examples of such tasks. The challenging nature of age estimation has presented Machine Learning as a typical solution to the problem over the years. As it is with humans, the accuracy of automatic age estimation, therefore depends on several factors two of which are the amount of data

available for training and the performance (generalization) of the chosen learning algorithm [1] .

With the advances in data collection and storage capabilities, massive multidimensional data are being generated on a daily basis in a wide range of emerging applications and learning algorithms for knowledge extraction from these data are becoming more and more important. These massive multidimensional data are usually very high-dimensional, with a large amount of redundancy and only occupying a subspace of the input space [6] [7]. Generally, image processing applications involve relatively massive multidimensional data, this Multidimensional objects are officially called tensor data [8], here we focus on age estimation based on Multidimensional face images of human subjects using multidimensional analysis for the representation of facial images.

Thus, algorithms need to extract useful information from this data. However, the excessive dimension of the data space often brings the algorithms into the curse of dimensionality dilemma [10]. Therefore, the dimensionality reduction process is one of the most important tasks in the face analysis field [9, 11, 12]. The purpose of dimensionality reduction is to transform a high-dimensional data set into a low-dimensional space, while retaining most of the underlying structure in the data [13]. In this work, we present a novel approach to solve the supervised dimensionality reduction problem by encoding an image object as a general tensor of second or even higher order [10].

In this work, we investigate a novel framework for automatic age estimation (age estimation from face images). This framework used Tensor Exponential Discriminant Analysis (TEDA). To properly present our modest work we divide it into three main chapters:

In Chapter 1, we are giving definitions, challenges, applications and an overview of the field of facial age estimation.

In Chapter 2, we recalled the main tensor definitions, properties and the different types of tensor projections. along Tensor Exponential Discriminant Analysis (TEDA) is detailed in this chapter.

Finally, the last chapter (Implementation and Results) will detail our age estimation framework with its different blocks, the PAL database, and the results.

1.1 Introduction

As an important biological information carrier, the human face reflects many properties such as identity, age, gender, expression, and emotion [14]. The vast majority of people are able to easily recognize human traits like emotional states, where they can tell if the person is happy, sad or angry from the face. Likewise, it is easy to determine the gender of the person. However, knowing the person's age just by looking at old or recent pictures for them is often a bigger challenge. Our objective in this work is to develop a human age estimator from face images. Given a face image of the person, we label it with an estimated age as shown in Figure 1.1 [15].

Automatic age estimation is a topical topic but remains a relatively difficult task due to the diversity of the ageing process as shown in Figure 1.2. Different people of the same age may have a distinctly different appearance due to the aging process that differs from person to person [16]. Aging is a non-reversible process. Human facial characteristics change with time, which reflects major variations in appearance. The age progression, signs displayed on the faces are uncontrollable and personalized such as hair whitening, muscles dropping and wrinkles. The aging signs depend on many external factors such as life style and degree of stress. For instance, smoking causes several facial characteristic changes. A 30 years old person who smokes a box of cigarettes each day will look like a 42 year old one [15].

Age Estimation refers to the process of labelling a facial signal with an age or age group. The input signal can be 2D, 3D or image sequences. The problems that fall into this category can be divided further into two subcategories, depending on the labels of the training data: (a) real age or (b) apparent age estimation, which refers to the age that is inferred by humans based on the individual's appearance [17].

Figure 1.1 : Age estimation .

Figure 1.2 : Photos of the same person of different years reveal the changing process of aging. Each column shows the images of the same person and below the photo is the taken year [14].

1.2 Definitions

1.2.1 Definitions about human age

We want to differentiate four definitions about human age**:**

- Actual Age: The real age (cumulated years after birth) of an individual.
- Appearance Age: The age information shown on the visual appearance.
- Perceived Age: The individual age gauged by human subjects from the visual appearance.
- Estimated Age: The individual age recognized by machine from the visual appearance.

We use the actual age and estimate age defined in this work [15].

Figure 1.3: A flexible hierarchical approach for facial age estimation based on multiple features [18].
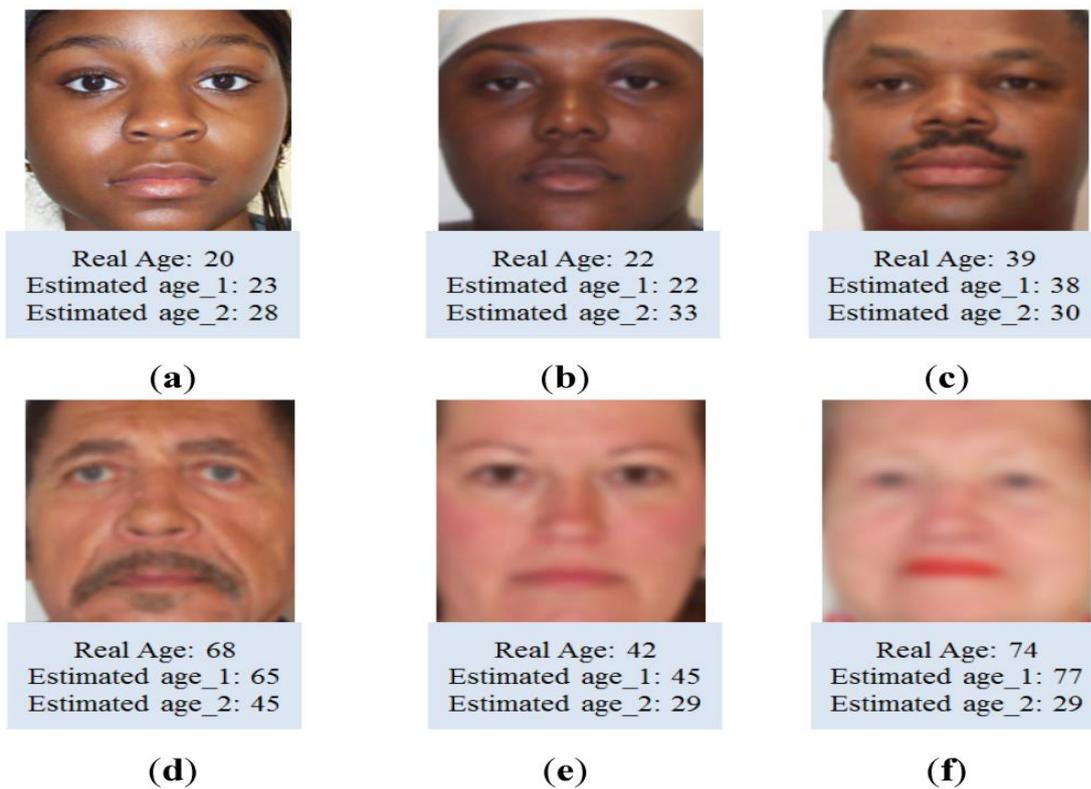
Figure 1.4: Age estimation-based soft biometrics considering optical blurring based on symmetrical sub-blocks for MLBP [18].

1.2.2 Computer vision problems

We have to distinguish between two computer vision problems. Age synthesis, which aims at simulating the aging effects on human faces [19] (i.e. simulate how the face would

look like at a certain age) with customized single or mixed facial attributes (identity, expression, gender, age, ethnicity, pose, etc.) which is the inverse procedure of age estimation. While, age estimation aims at labeling a face image automatically with the exact age (year) or the age group (year range) of the individual face [15].

- ✓ Age progression: that is, the process of transforming a facial visual input, in order to model it across different ages. The change of the age can be bidirectional, so that the facial output can appear either younger or older than the input.

- ✓ Age estimation: refers to the process of labelling a facial signal with an age or age group.

- ✓ Age-invariant facial characterization: involves the process of building a signal representation that is invariant to the facial transformations and appearance changes caused by aging [17].

1.3 Challenges

Face images can demonstrate a wide degree of variation in both shape and texture. Appearance variations are caused by individual differences, the deformation of an individual face due to changes in expression and speaking, as well as lighting variations. These issues are explained more in the following points:

- ➢ Age depends on many factors, some of them are visual and many others are non-visual such as ethnic background, living style, working environment, health condition and social life. In particular, Stone [20] stated that aging can be accelerated by smoking, genetic predisposition, emotional stress, disease processes, dramatic changes in weight, and exposure to extreme climates.

- ➢ The visual features that can help in estimating age, such as people's facial features are affected by pose, lighting and imaging conditions (variation regarding face size, image quality and occlusions).

- ➢ Males and females may have different general discriminative features displayed in images due to the different extent in using makeup, accessories as shown in Figure 1.5 and cosmetic surgeries which increase the negative influence of individual differences.

- ➢ The difficulty of acquiring large-scale databases, which covers enough age range with chronological face aging images, makes the estimation tasks more difficult to achieve [15].

Figure 1.5 : Effects of makeup and plastic surgery on perceived age [18].

1.4 Facial age estimation system

The age estimate system for the face is an individual verification system, which estimates the age of the person who belongs to the system database [16]. Age estimation systems can be represented in the following diagram:
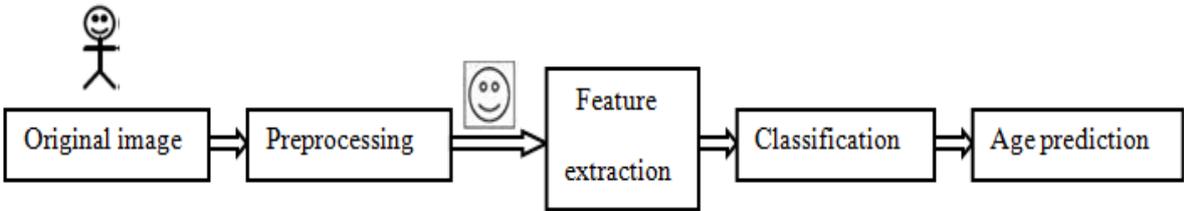
Figure 1.6 : General block scheme of an age estimation system [16].

1.5 Age estimation applications

Automatic age estimation from facial images has recently emerged as a technology with multiple interesting applications [15]. Applications where age estimation can play an important role include [16]:

Biometrics: The physical, physiological or behavioural cues based on which a person is recognized, e.g., iris, fingerprint, face are referred to as biometrics [21]. Age and kinship comprise soft biometrics [22], [23], [24] as they can be used to boost the effectiveness of recognition. Besides improving face recognition accuracy there is a need for robust towards aging and kinship. Passport checks, demand age-invariance in case of the large age gap between the passport image and the person in question. Similarly, kinship invariance can potentially boost automatic face recognition, in particular towards distinguishing between kins that look alike [17].

Electronic Customer Relationship Management (ECRM): The ECRM is a management strategy that uses latest computer vision algorithms and approaches to build interaction tools for effectively establishing different relationships with all customers and serving them individually as shown in Figure 1.7. Customers are classified to different age groups such as babies, teenagers, adults and senior adults. It is important to take their habits, preferences, responsiveness, and expectation to marketing in consideration, companies can earn more money by acknowledging this fact, responding directly to allcustomers" specific needs based on their age groups, and customizing products or services according to each customer age group [15].

Figure 1.7 : Automatically estimating the customers' age can help with efficient customer profiling and age-oriented decision making [18].

The most challenging part is to maintain enough personal information records or histories from all customers'age groups, where companies need to invest a large amount of the cost input to establish long-term customer relationships. For example, the owner of a fast food shop wants to know the most popular sandwiches or meals for each age group; the advertising companies want to target specific audiences (potential customers) for specific advertisements in terms of age groups; mobile manufacturers want to know which age group is more interested in their new product models showing in a public kiosk; clothes stores may display suitable fashions for males or females according to their age groups as shown in Figure 1.8.

Figure 1.8 : Fashions for males or females, according to their age groups [18].

Security Control and Surveillance Monitoring: Security control and surveillance monitoring systems are becoming increasingly important in our everyday life, especially with the rise the number of crimes and terrorist threats. With the help of a monitoring camera, a human age estimator application can generate a warning sound or alarm when underage drinkers are entering bars and preventing them from purchasing tobacco products from vending machines if the IDs are faked or another as shown in Figure 1.9 [15], access control, e.g., restricting the access of minors to sensible products like alcohol from vending machines or to events with adult content as shown in Figure 1.10. surveillance, e.g., automatic detection of unattended children at unusual hours and places [25].

Figure 1.9 : Cigarette vending machine [26].

Figure 1.10 : Internet safety for minors [26].

Forensics: Forensics include a set of scientific techniques that are used for crime detection. Among these techniques, forensics art demonstrates the challenging task of producing a lifelike image of a person. In some cases, forensics experts face the need to change the age of a face. Such cases include updating archive images of wanted criminals as well as images of lost children [17].

Health care systems: Age estimation applications are not limited to prevent criminals from committing crimes, but also can be used in health care systems, such as robotic nurse, intelligent intensive care unit as shown in Figure 1.11, for customized services [15].
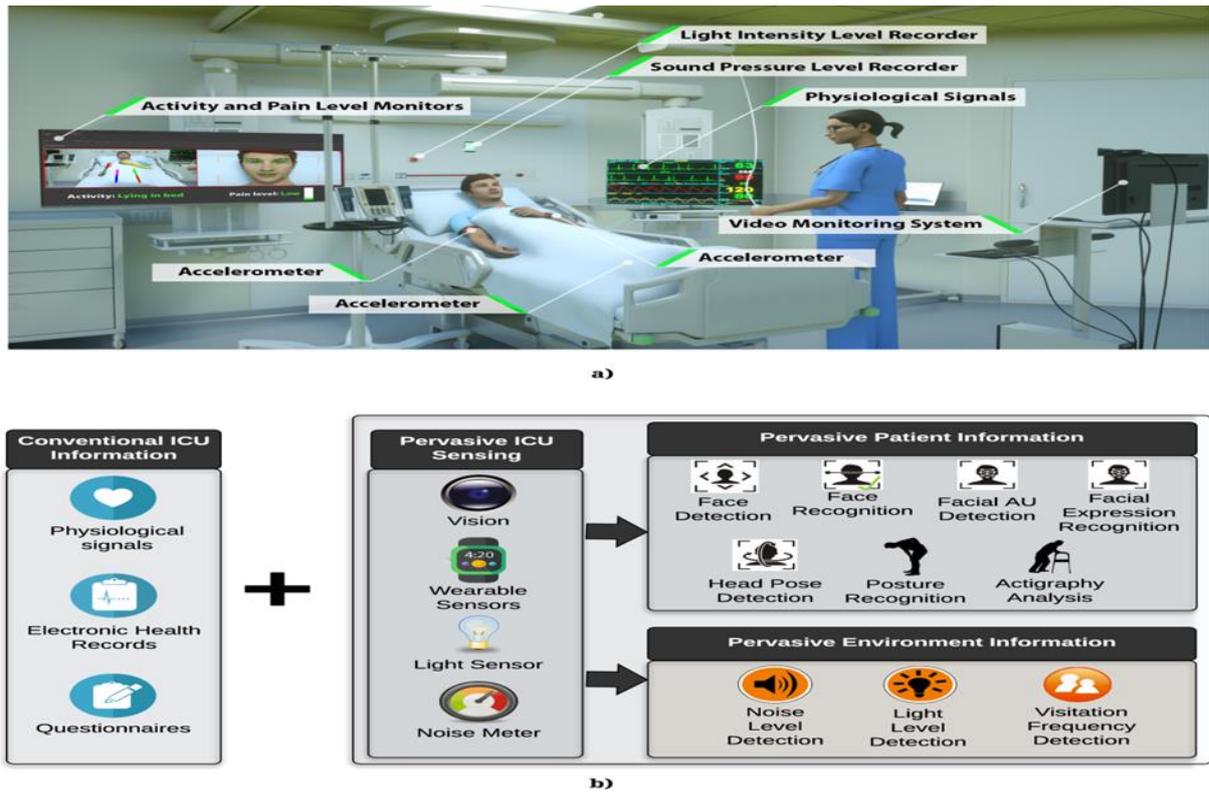
Figure 1.11: Intelligent ICU for Autonomous Patient Monitoring using Pervasive Sensing and   Deep Learning [18].

1.6 Evaluation measures

The large number of methods in age estimation has exposed the need for a common evaluation protocol. The most widely adopted measures are the Mean Absolute Error (MAE) and the Cumulative Score (CS) [17]. The MAE is calculated based on the average of the absolute errors between the estimated age and the ground truth (labelled age), which is represented as

$$\text{MAE} = \frac{1}{N}\sum_{n=1}^{N}\|l_n - y_n\|, \tag{1.1}$$

Where l_n is the ground truth label of the n$^{\text{th}}$ image and y_n represents the estimated age based on the proposed framework. N is the total number of the test samples.
The Cumulative Score (CS) is represented as:

$$\text{CS}(l) = \frac{Ne \leq l}{N} \times 100\% \tag{1.2}$$

11

Where $Ne \le l$ is the number of images with an absolute error between the estimated age and the ground truth age not greater than l years. CS is the accuracy rate for the estimation error no higher than l [14].

1.7 Datasets

The availability of labelled datasets is a cornerstone towards the development of facial models for aging [17]. The first algorithm for face recognition was introduced in mid 1960's [27] and since then numerous approaches have been proposed in the literature [28]. For decades and prior to the proliferation of deep learning (Before the advent of deep learning in 2006 [29]), various facial datasets were introduced in the literature. One common feature of all these datasets was that they contained images which were captured under controlled conditions (e.g., common background, controlled lighting setting, etc.). This restriction was imposed due to the fact that the feature extractors did not perform well on "in-the-wild" datasets (i.e., datasets that included images captured in uncontrolled conditions). Some of the most widely used databases which included images captured under controlled conditions were the XM2VTS database [30], the Multi-PIE database [31, 32], the AR face database [33], the Caltech faces database [34], the FERET database [35, 36, 37], the Yale face database [38].

During the last few years, an explosion in scientific research with respect to the development of deep learning architectures for face recognition has been witnessed. Deep learning comprises a set of methods that are able to automatically discover the patterns that may exist in raw data and, as a result, feature extractors which were utilized to transform the raw data are no longer required [39, 40]. This competitive advantage of deep learning methods against the conventional algorithms led to the introduction of several "in-the-wild" databases in the literature.

The term "in-the-wild" is used to refer to databases that contain images which have been captured under completely uncontrolled conditions (e.g., varying backgrounds, existence of noise in the pictures, existence of occlusions in the faces depicted in various images, different types of cameras used to capture the images, etc.) [28].

Early models for facial age progression and estimation date back to 1994-95[41], [42], while the problem of face recognition across ages was first investigated in 2000 [43],[17].In this section, we review recently released datasets which contain annotations for age

progression, age estimation, age-invariant facial characterization. Particular emphasis is put on data capturing naturalistic, real-world conditions, often referred as in the wild [44]. A complete catalogue of the available datasets for facial age modelling is listed in Table 1 [17].

The vast majority of the available datasets in Table 1.1 for facial age modeling contain still images and apart from the FACES, IRIP, LHI, and YGA face datasets, they are not balanced with regards to the gender and age of the subjects. In addition, while containing an abundance of different annotated faces, many of these datasets do not contain a considerable number of images of the same person at different ages, which is essential for training methods for age progression and age-invariant facial characterization [17].

Databases	#data	#subjects	Age-range	precision	In the wild	Year	modality
AgeDB[28]	16,458	568	1-101	Exact age	yes	2017	images
PAL [45]	1046	1046	18-93	Exact age	yes	2016	images
IMDB-WIKI [25]	523,051	20,284	0-100	Exact age	yes	2016	images
AFAD [46]	164,432	N/A	14-40+	Exact age	yes	2016	images
IRIP [47]	2,100	N/A	1-70	Exact age	yes	2016	images
OUI-Adience [48]	26,580	2.284	0-60+	age group	yes	2014	images
CACD [49]	163,446	2,000	N/A	Exact age	yes	2014	images
UvA-Nemo [50]	1,247	400	8-76	Exact age	no	2014	videos
VADANA [51]	2.298	43	0-78	age group	yes	2011	images
LHI Face Dataset [52]	8.000	8.000	9-89	Exact age	no	2010	images
HOIP [53]	306.600	300	15-64	age group	no	2010	images
FACES [54]	1026	171	19-80	Exact age	no	2010	images
Web Image Db [55]	219.892	N/A	1-80	Exact age	yes	2009	images
Images of Groups [56]	28.231	28.231	0-66+	age group	yes	2009	images

YGA [57]	8.000	1.600	0-93	Exact age	yes	2008	images
Iranian Face [58]	3.600	616	2-85	Exact age	yes	2007	images
Brown Sisters [59]	16	4	15-62	Exact age	yes	2007	images
Scherbaum'sDb[60]	438	438	8-18 or adult	both	no	2007	3D
MORPH2 [61]	55.134	13.618	16-77	Exact age	no	2006	images
WIT-BD [62]	26.222	5.500	3-85	age group	yes	2006	images
AI&R [63]	34	17	22-66	Exact age	no	2006	images
FRGC [64]	50.000	568	18-70	Exact age	partially	2005	Images,3D
Lifespan Database [65]	576	576	18-93	age group	yes	2004	images
FG- NET [66] , [67]	1.002	82	0-69	Exact age	yes	2004	images
PIE [68]	41.638	68	N/A	Exact age	no	2002	images
FERET [69]	14.126	1,199	N/A	Exact age	partially	1998	images
Coucasian Face Db [70]	147	147	20-62	Exact age	no	1995	images

Table 1.1 : Datasets with age labels [17] .

Data suitable for age estimation

The databases used for the task of age estimation vary greatly in terms of sample size, number of subjects, and age-range, as indicated in Table 1.1. Nevertheless, the recent success of deep learning-based models in computer vision has created a need for larger datasets with age annotations. Towards this end, the IMDB-WIKI, AFAD and OUI-Adjience datasets have been collected.

- OUI-Adience (The OUI-Adience dataset)
- UvA-NEMO (The UvA-NEMO dataset)
- AFAD (The Asian Face Age Dataset)
- IMDB-WIKI (The largest age-annotated dataset)

1.8 Previous work

The existing age estimation frameworks using face images typically consist of two main stages: age image representation and age estimation techniques.

1.8.1 Image representation

Anthropometric models :

In the past few years, much research has been conducted in human facial age estimation. The earliest paper published in the area of age classification from facial images was the work by Kwon and Lobo [71]. They proposed a human age classification method based on the cranio-facial development theory and skin wrinkle analysis, where the human faces are classified into three groups, namely, babies, young and senior adults [14].

The tasks of automatic age estimation from facial images introduced in 1994 by Kwon and Lobo [41]. Inspired by anthropometric studies [72] that describe the growth of the human head from infancy to adulthood, six facial distance ratios are used to discriminate between infants and adults. The adult faces are further classified into young adults and older adults by using snakelets (i.e., deformable curves) [73], which capture wrinkles on certain regions [17].

Active appearance models :

Lanitis et al.[68] and [74] adopted the statistical face model, Active Appearance Models (AAMs) [75], to extract the shape and texture information of facial images. In their work, the aging pattern is represented by a quadratic function called the aging function[14].

Active Appearance Model (AAM) is a generative facial model introduced in [75] by Cootes et al. AAMs employ Principal Component Analysis (PCA) to learn a linear model for shape and appearance from images and a set of landmarks. This representation was first used for age estimation in [68],[17].

Aging pattern subspace :

Geng et al. [76] and [77] proposed the Aging pattErn Subspace (AGES) algorithm based on the subspace trained on a data structure called aging pattern vector [14].

Appearance models :

Aging-related facial feature extraction is more focused by the appearance model. Both global and local features were used in existing age estimation systems. The effective texture descriptor, Local Binary Patterns (LBP) [78], has been used for appearance feature extraction in an automatic age estimation system [79] [15].

Suo et al. [80] proposed to design graphical facial features–topology, geometry, photometry, and configuration–based on their early developed multi-resolution hierarchical face model [81] [15].

Guo et al. [82] proposed to use the Biologically Inspired Features (BIF) [83] [84] for age estimation via faces [15].

1.8.2 Age estimation techniques

Given an aging feature representation, the next step is to estimate ages. Age estimation approaches fall into two categories: a) classification-based; and b) regression-based [15].The age estimation problem can be seen as a regression (Fu and Huang 2008) or as a classification problem up to a quantization error (Lanitis et al. 2004; Geng et al. 2007) [25].

Classification-based :

Among the most popular classification techniques we mention the traditional nearest neighbor (NN) and Support Vector Machines (SVMs) (Cortes and Vapnik 1995) [25].

Lanitis et al. [76] evaluated the performance of different classifiers for age estimation, including the nearest neighbor classifier, the Artificial Neural Networks (ANNs), and a quadratic function classifier [15].

Ueki et al. [78] built 11 Gaussian models in a low-dimensional 2DLDA+LDA feature space using the EM Algorithm. The age-group classification is determined by fitting the test image to each Gaussian model and comparing the likelihoods [15].

Regression-based :

Among the most popular regression techniques we mention Support Vector Regression (SVR) (Drucker et al. 1997), Partial Least Squares (PLS) (Geladi and Kowalski 1986), Canonical Correlation Analysis (CCA) (Hardoon et al. 2004) [25].

Lanitis et al. [68] investigated three formulations for the aging function: linear, quadratic, and cubic, respectively, with 50 raw model parameters.

Zhou et al. [85] presented the generalized Image Based Regression (IBR) aiming at multiple-output settings. A boosting scheme is used to select features from redundant Haar-like feature set. The proposed training algorithm can also significantly reduce the computational cost [15].

Guo et al. [86] [58] proposed a method, called Locally Adjusted Robust Regressor (LARR). They showed that a consistently better performance can be obtained by combining a classifier and a regressor [15].

Guo et al. [58] proposed a probabilistic approach to combine regression and classification results [15].

Yan et al. [87] regarded age estimation as a regression problem with nonnegative label intervals and solved the problem through semidefinite programming. They also proposed an EM algorithm to solve the regression problem and speed up the optimization process [88]. Instead of learning a specific aging pattern for each individual, a common aging trend or pattern can be learned from many individuals at different ages [14].

1.9 Conclusion

Age estimation based on the human face remains a significant problem in computer vision and pattern recognition. In order to estimate an accurate age of a facial image, We have developed a fully automatic age estimation framework in this work. We have introduced the most important applications to estimate age in biometrics. Moreover, we review evaluation protocols and metrics. A complete catalogue of publicly available datasets with manual annotations for facial age tasks is listed in section 1.7. Finally, we presented a brief review of related work on age estimation.

2.1 Introduction

Multidimensional data are generated day-to-day in many applications. This leads to a high demand for the use of multilinear algebra tools. This chapter briefly reviews some basic multilinear concepts used in this work and introduces the multilinear projection of tensor objects for the tensor dimensionality reduction. Finally, we present Tensor Exponential Discriminant Analysis (TEDA) method.

2.2 Tensor properties for multidimensional data modelling

This section introduces the mathematical notations and the background behind the subspace approach.

2.2.1 Notations and concepts

The variables and mathematical notations used in this work are defined as follows:

- Lowercase and uppercase symbols (e.g., i, j, K, L and λ) denote scalars.
- Italic lowercase symbols (e.g., *x, y* and *α*) denote vectors.
- Italic uppercase symbols (e.g., *U, N, X* and *V*) denote matrices.
- Bold uppercase symbols (e.g., **A, B, C** and **X**) denote tensors.

2.2.1.1 Definition of a tensor

The term tensor object is used here to denote a multidimensional object, the elements of which are to be addressed by more than two indices [89]. The number of indices used in the description defines the order of the tensor object and each index defines one of the so-called "modes." Many image and video data are naturally tensor objects. For example, color images are 3-D (third-order tensor) objects with column, row, and color modes [90] as well as other grayscale video sequences can be viewed as third-order tensors with column, row, and time modes. Naturally, color video sequences are fourth-order tensors with the addition of a color mode [91].

A tensor $\mathbf{A} \epsilon \ \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_m}$ is defined as a multidimensional array [91], [10]. m is the order of the tensor and **A** is called an mth-order tensor. I_K, $1 \leq k \leq m$, is the dimension of the Kth mode. Each element of the tensor **A** is denoted as $\mathbf{A}_{i_1 i_2 \ldots i_m}$, where $1 \leq i_K \ \leq \ I_K$,

$1 \leq k \leq m$ [92].
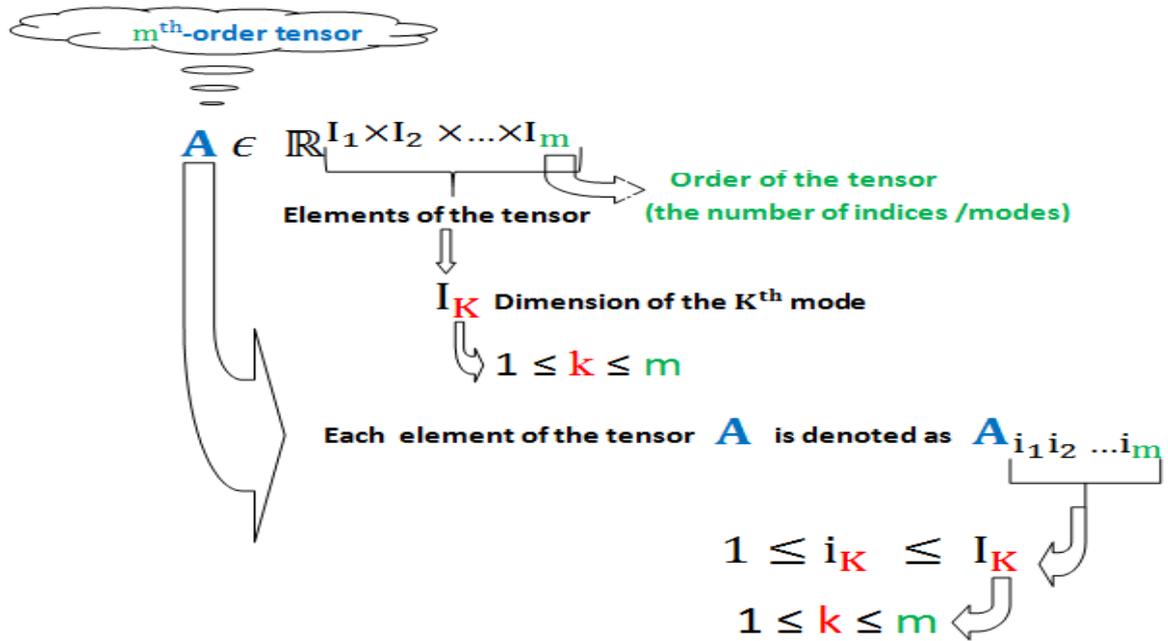
Figure 2.1 : Illustration of a higher-order tensors [93].
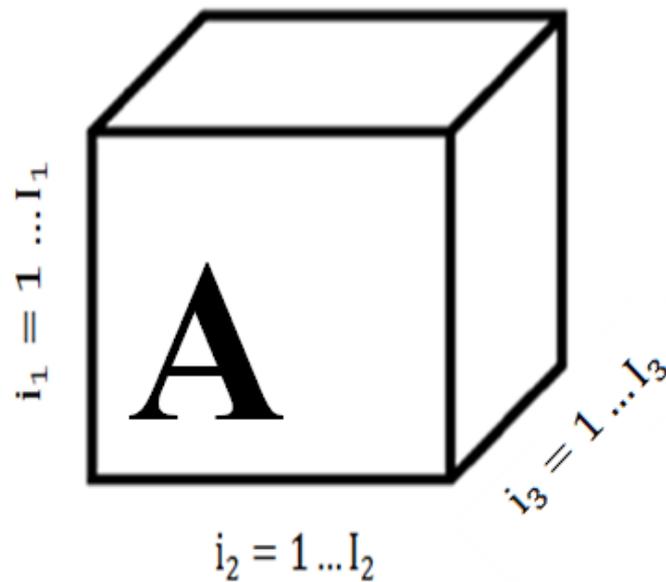
Figure 2.2 : Third-order tensor ($\mathbf{A}\epsilon$ $\mathbb{R}^{I_1 \times I_2 \times I_3}$) [9].

Figure 2.3 shows two examples of tensor data representations for a face image and a silhouette sequence:
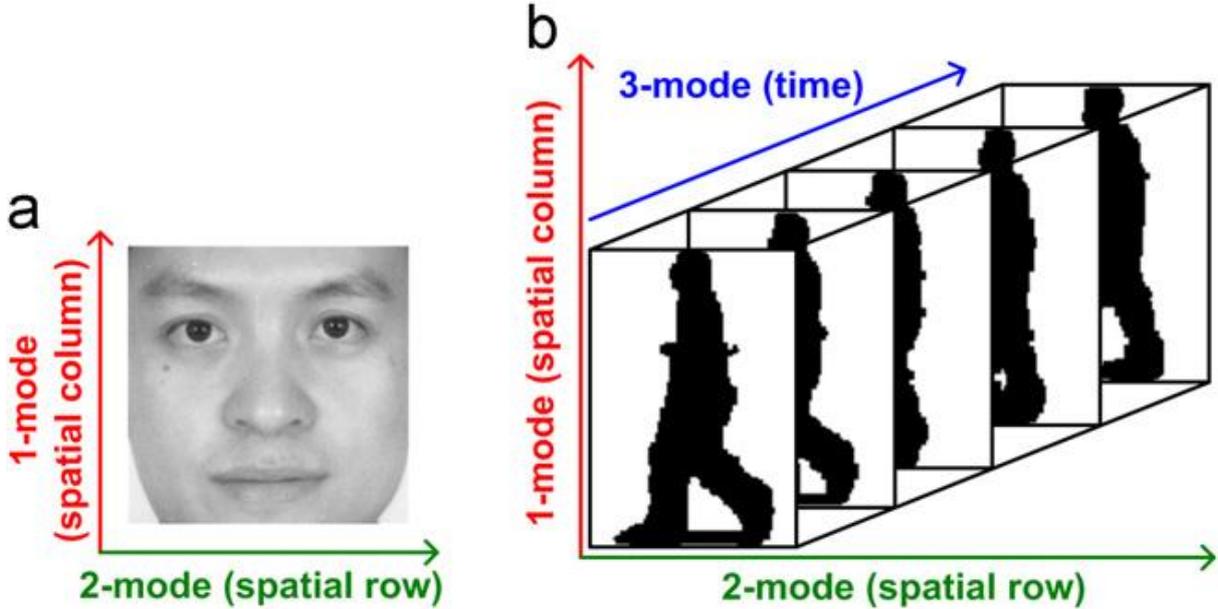
Figure 2.3 : Illustration of real-world data in their natural tensor representation: (a) a 2D face and (b) a 3D silhouette sequence [7].

Figures 2.4, 2.5 and 2.6 shows examples of a tensor of order 2, 3 and 4, respectively:
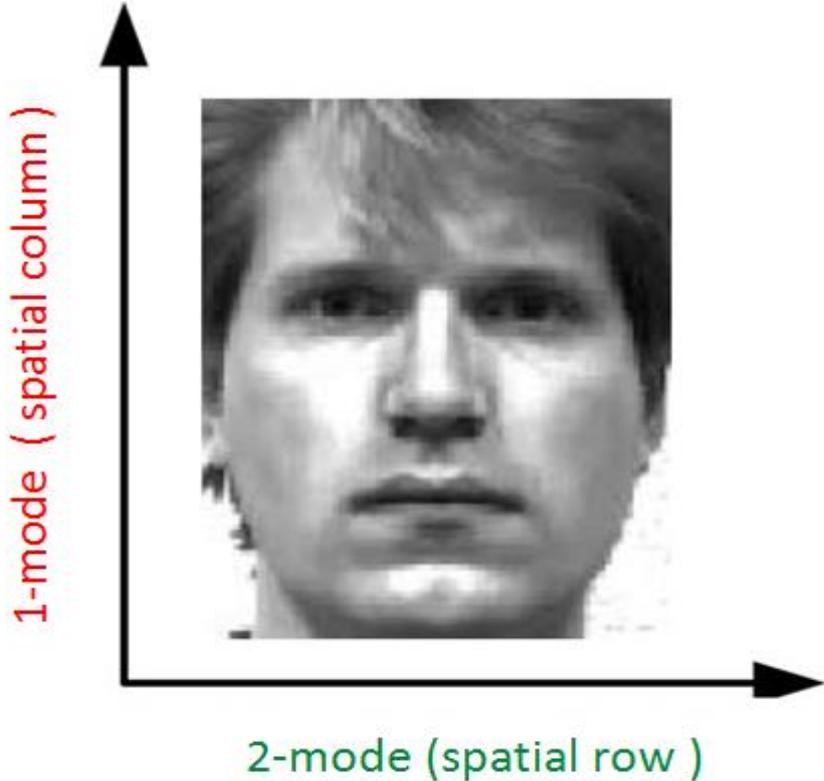
Figure 2.4 : 2nd-order tensor (grayscale image) [9].
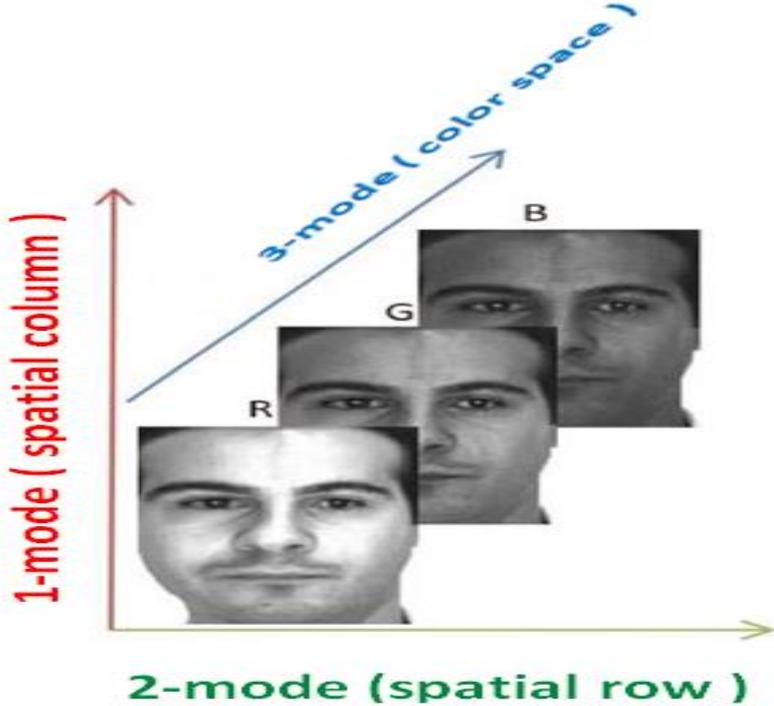
Figure 2.5 : 3rd-order tensor (coloured image) [9].
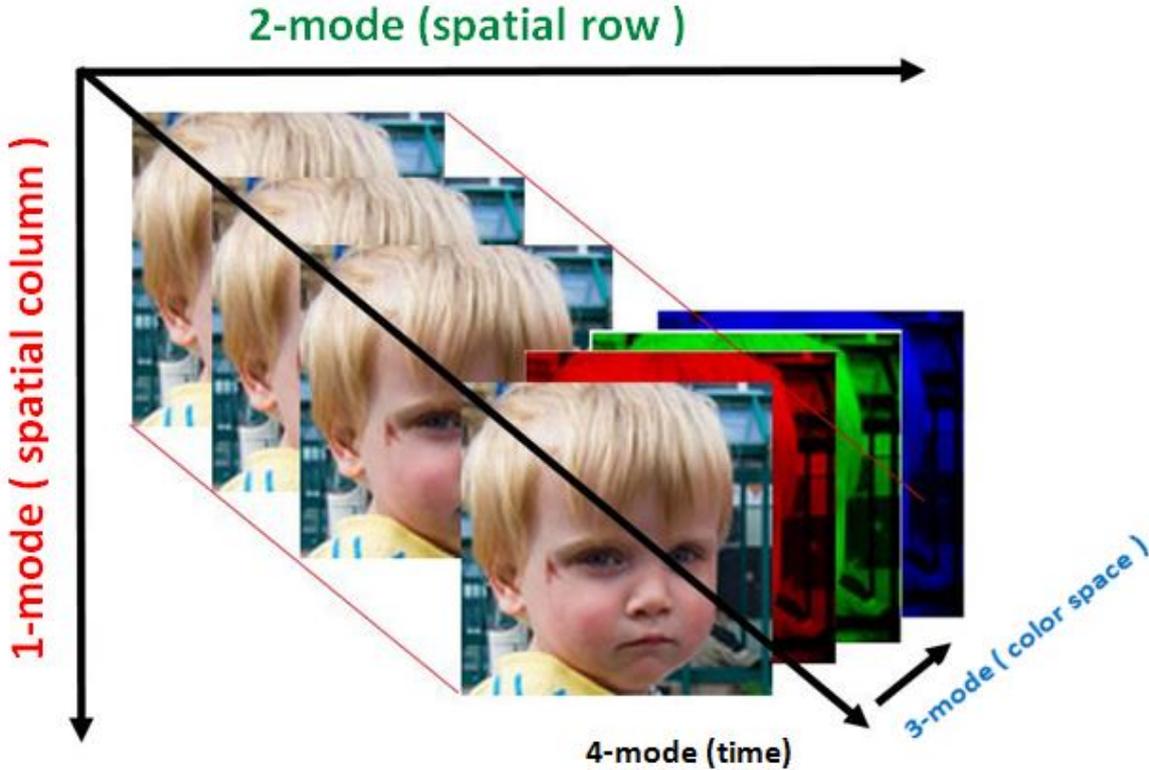
Figure 2.6 : 4th-order tensor (color video sequences) [9].

A tensor is considered a higher order extension of a vector and matrix .The order tensor space m is composed of m subspace mode. A data set including m parameters may be modelled as an input m array, in which each tensor mode is associated with a different parameter.

2.2.1.2 Basic multilinear algebra

An mth-order tensor is denoted as: $\mathbf{A} \epsilon \ \mathbb{R}^{I_1 \times I_2 \times \dots \times I_m}$, which is addressed by k indices i_k, $k = 1, \dots, m$, with each i_k in addressing the k-mode of $\mathbf{A}$ [7] .

Unfolding tensor

For all values of indices . Unfolding $\mathbf{A}$ along the k-mode is denoted as [7]:

$$A_{(K)} \in \ \mathbb{R}^{I_K \times (I_1 \times \dots \times I_{k-1} \times I_{k+1} \times \dots \times I_m)} \tag{2.1}$$

Where the column vectors of $A_{(K)}$ are the k-mode vectors of $\mathbf{A}$. Figure 2.7 (b), (c), and (d) illustrate the 1-mode, 2-mode and 3-mode vectors of a tensor $\mathbf{A}$ respectively. Figure 2.8 shows the 1-mode unfolding of the tensor $\mathbf{A}$.
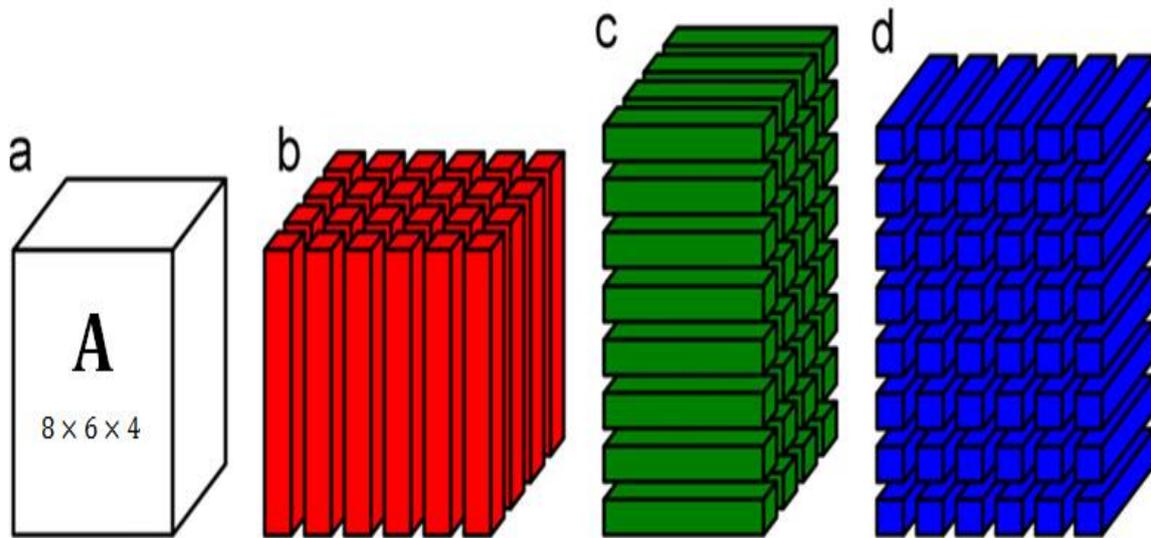
Figure 2.7 : Visual illustration of k-mode vectors : (a) a tensor $\mathbf{A} \in \mathbb{R}^{8 \times 6 \times 4}$, (b) the 1-mode (column) vectors, (c) the 2-mode (row) vectors, and (d) the 3-mode (tube) vectors [7] .
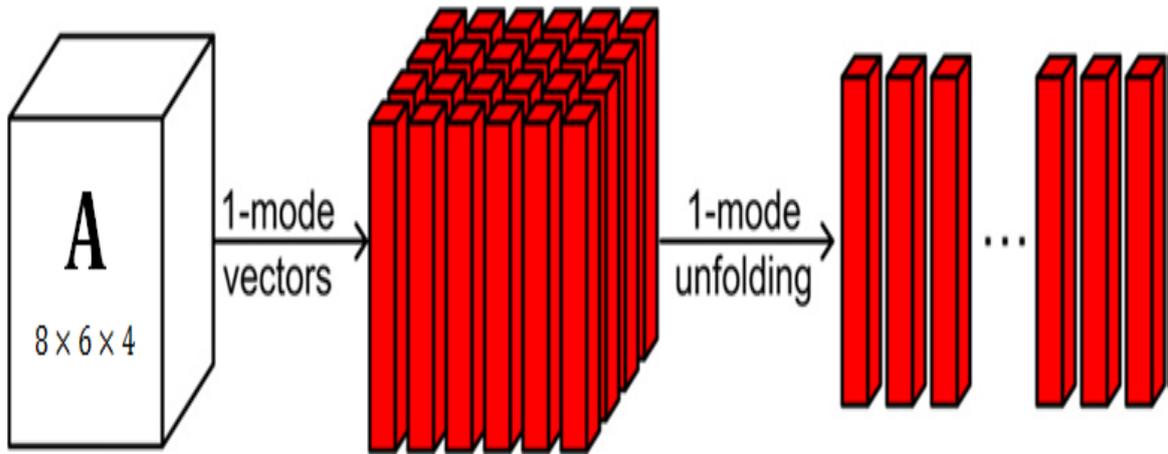
Figure 2.8 : Visual illustration of 1-mode unfolding of a third-order tensor [7].

The procedure to form k-mode matrices from an order of tensor > 2 is called unfolding tensor [9]. Figure 2.9 illustrates the unfolding of third-order tensor in the different modes.
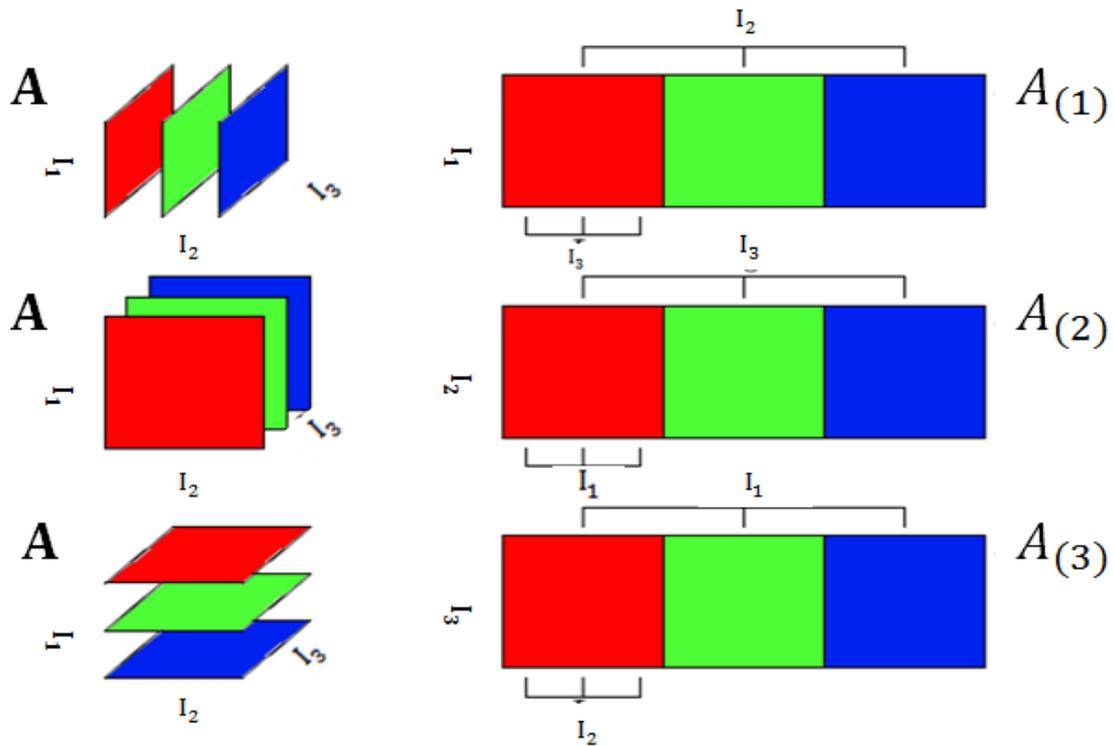
Figure 2.9: unfolding of third-order tensor in the different modes [9].

The concept is easier to understand using an example :
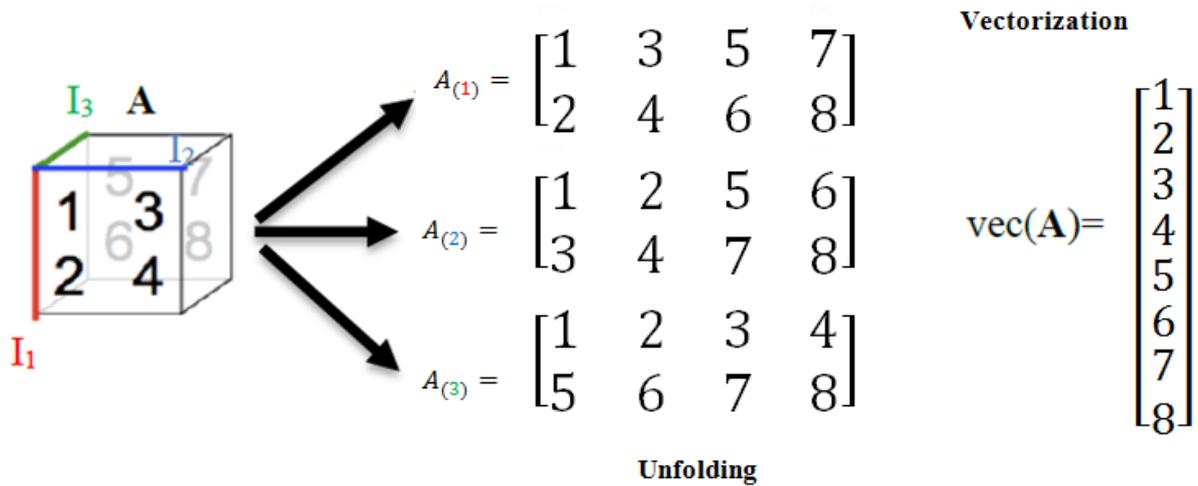
Figure 2.10 : Example of unfolding of third-order tensor in the different modes [9] .

Tensor multiplication

The k-mode (matrix) product of a tensor **A** by a matrix $U \in \mathbb{R}^{J_k \times I_k}$, denoted as $\mathbf{A} \times_k U$ is a tensor with entries [93] [7] [8] [9]:

$$\mathbf{B} = \mathbf{A} \times_k U \in \mathbb{R}^{I_1 \times \dots I_{k-1} \times J_k \times I_{k+1} \dots \times I_m}$$

$$= \sum_{i_k} \mathbf{A}_{i_1, \dots, i_m} \cdot U_{j_k, i_k}$$

$$= \sum_{i_k=1}^{I_k} a_{i_1, \dots, i_{k-1}, j_k, i_{k+1}, \dots, i_m} \cdot u_{j_k, i_k} \qquad (2.2)$$

Figure 2.11 illustrates an example of 1-mode product of the third order tensor $\mathbf{A} \in \mathbb{R}^{12 \times 7 \times 4}$ by matrix $U^{(T)} \in \mathbb{R}^{4 \times 12}$, where the result is a tensor $\mathbf{B} \in \mathbb{R}^{4 \times 7 \times 4}$.
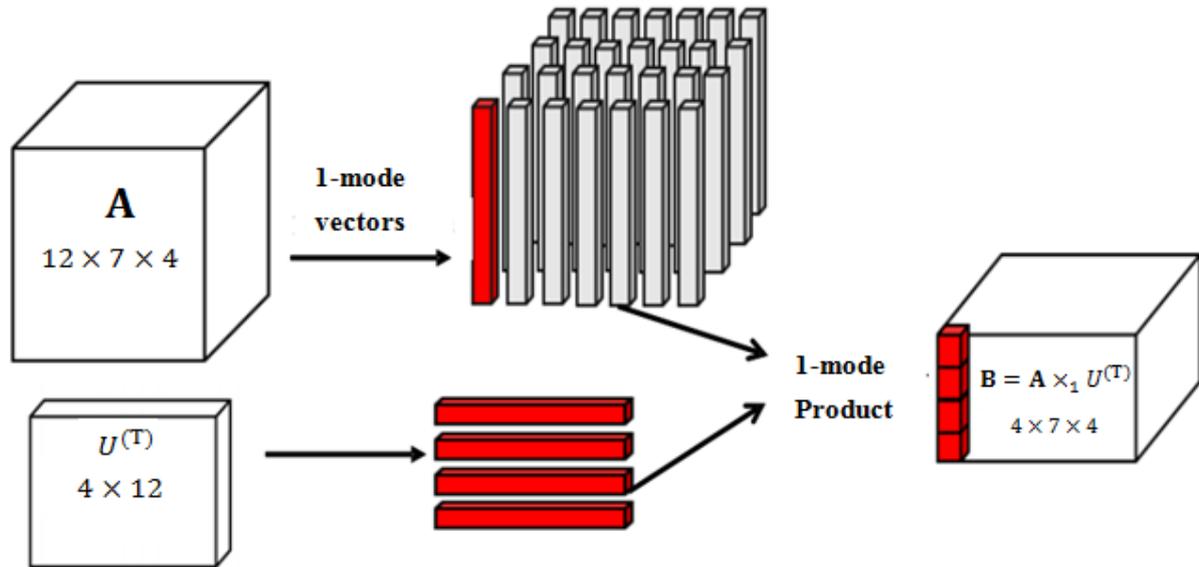
Figure 2.11: Visual illustration of 1-mode product of the third-order tensor**A** by a matrix U [9].

The concept is easier to understand using an example :

Let the frontal slices of $\mathbf{A} \epsilon \ \mathbb{R}^{3 \times 4 \times 2}$ be

$$A_1 = \begin{bmatrix} 1 & 4 & 7 & 10 \\ 2 & 5 & 8 & 11 \\ 3 & 6 & 9 & 12 \end{bmatrix}, A_2 = \begin{bmatrix} 13 & 16 & 19 & 22 \\ 14 & 17 & 20 & 23 \\ 15 & 18 & 21 & 24 \end{bmatrix}$$

and let $U \in \ \mathbb{R}^{2 \times 3}$

$$U = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}$$

Then the product $\mathbf{B} = \mathbf{A} \times_1 U \in \ \mathbb{R}^{2 \times 4 \times 2}$ is

$$B_1 = \begin{bmatrix} 22 & 49 & 76 & 103 \\ 28 & 64 & 100 & 136 \end{bmatrix}, B_2 = \begin{bmatrix} 130 & 157 & 184 & 211 \\ 172 & 208 & 244 & 280 \end{bmatrix}.$$

We can also write this result in a matrix form, which clearly shows the generalization of the matrix product:

$$B_K = U.A_k \tag{2.3}$$

Where B_K and A_k represent folding matrices of tensors $\mathbf{B}$ and $\mathbf{A}$ in k-mode.

Thus the tensor resulting from the produced k-mode between the tensor $\mathbf{A} \epsilon \ \mathbb{R}^{I_1 \times I_2 \times ... \times I_m}$ and the matrix $U^{(k)} \in \mathbb{R}^{J_k \times I_k}$, $k = 1, ..., m$ is defined by:

$$\mathbf{B} = \mathbf{A} \times_1 U^{(1)} ... \times_k U^{(k)} \in \mathbb{R}^{J_1 \times ... \times J_m} \tag{2.4}$$

Each index element $(j_1, \dots, j_m)$ of tensor **B** is given by:

$$b_{j_1,\dots,j_m} = \Sigma_{i_1=1}^{I_1} \dots \Sigma_{i_m=1}^{I_m} a_{i_1 \dots i_m} u_{j_1 i_1}^{(1)} \dots u_{j_m i_m}^{(m)} \tag{2.5}$$

With $j_k \in \{1, \dots, J_k\}$ and $J_k \in \mathbb{N}^*, \forall\, k = 1, \dots, m$.

Properties:

1. Whatever the tensor $\mathbf{A} \epsilon\ \mathbb{R}^{I_1 \times I_2 \times \dots \times I_m}$ and matrices $F \in \mathbb{R}^{J_k \times I_k}$ and $\in \mathbb{R}^{J_s \times I_s}$, $(k \neq s)$, then :

$$(\mathbf{A} \times_k F) \times_s G = (\mathbf{A} \times_s G) \times_k F = \mathbf{A} \times_k F \times_s G \tag{2.6}$$

2. Given tensor $\mathbf{A} \epsilon\ \mathbb{R}^{I_1 \times I_2 \times \dots \times I_m}$ and matrices $F \in \mathbb{R}^{J_k \times I_k}$ and $G \in \mathbb{R}^{K_k \times I_k}$ then we have :

$$(\mathbf{A} \times_k F) \times_k G = \mathbf{A} \times_k (G.F) \tag{2.7}$$

3. Whatever the tensor $\mathbf{A} \epsilon\ \mathbb{R}^{I_1 \times I_2 \times \dots \times I_m}$, if the matrices $U^{(k)} \in \mathbb{R}^{I_k \times R_k}$ are orthogonal $\left(U^{(k)}.U^{(k)\,\mathrm{T}} = I_{I_k J_k} \right) \forall\, k = 1, \dots, m$ then :

$$\mathbf{B} = \mathbf{A} \times_1 U^{(1)} \dots \times_m U^{(m)} \Leftrightarrow \mathbf{B} = \mathbf{A} \times_1 U^{(1)\,\mathrm{T}} \dots \times_m U^{(m)\,\mathrm{T}} \tag{2.8}$$

Scalar product

The scalar product of two tensors $\mathbf{A}, \mathbf{B} \epsilon\ \mathbb{R}^{I_1 \times I_2 \times \dots \times I_m}$ is defined as [93] [7] [8]:

$$\langle \mathbf{A}, \mathbf{B} \rangle = \Sigma_{i_1} \Sigma_{i_2} \dots \Sigma_{i_m} \mathbf{A}_{i_1, i_2, \dots, i_m} . \mathbf{B}_{i_1, i_2, \dots, i_m} \tag{2.9}$$

Where :

$$\langle \mathbf{A} | \mathbf{B} \rangle = \sum_{i_1, i_2, \dots, i_m} a_{i_1, i_2, \dots, i_m} . b_{i_1, i_2, \dots, i_m}$$

Rank-One Tensors

The outer product between two vectors $e^{(1)} \in E^{(1)}$ and $e^{(2)} \in E^{(2)}$ defines a matrix

$A \in \mathbb{R}^{I_1 \times I_2}$ [8]:

$$A = e^{(1)} \mathrm{o}\, e^{(2)} = e^{(1)} e^{(2)^{\mathrm{T}}} \tag{2.10}$$

The K-mode vectors of **A** are defined as the I_k-dimensional vectors obtained from **A** by varying its index I_k while keeping all the other indices fixed. A rank-one tensor **A** equals to the outer product of m vectors [7]:

$$\mathbf{A} = u^{(1)} \mathrm{o}\, u^{(2)} \mathrm{o} \dots \mathrm{o}\, u^{(m)} \tag{2.11}$$

The symbol "o " represents the vector outer product. This means that each element of the tensor is the product of the corresponding vector elements [93]:

$$\mathbf{A}_{i_1,i_2,\dots,i_m} = u^{(1)}_{i_1} \mathrm{o}\, u^{(2)}_{i_2} \mathrm{o} \dots \mathrm{o}\, u^{(m)}_{i_m} \tag{2.12}$$

Figure 2.12 illustrates $\mathbf{A} = u^{(1)} \mathrm{o}\, u^{(2)} \mathrm{o}\, u^{(3)}$, a third-order rank-one tensor.
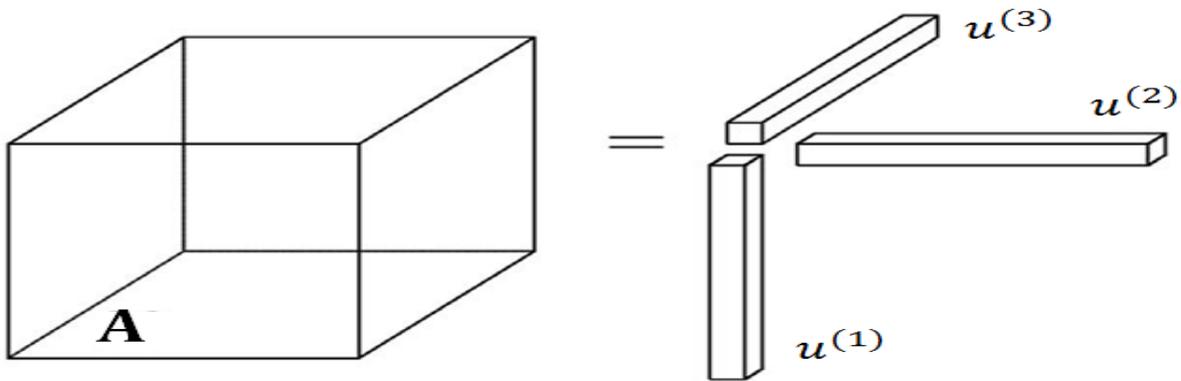
Figure 2.12 : Rank-one third-order tensor , $\mathbf{A} = u^{(1)} \mathrm{o}\, u^{(2)} \mathrm{o}\, u^{(3)}$ [93].

The Euclidean distance

The distance between tensors **A** and **B** can be measured by the Frobenius norm [7] [8]:

$$dist(\mathbf{A}, \mathbf{B}) = \|\mathbf{A} - \mathbf{B}\|_{\mathrm{F}}. \tag{2.13}$$

Although this is a tensor-based measure, it is equivalent to a distance measure of corresponding vector representations, as proven in [94]. Let vec(**A**) be the vector representation (vectoriza-tion) of **A**, then:

$$dist(\mathbf{A}, \mathbf{B}) = \|vec(\mathbf{A}) - vec(\mathbf{B})\|_2. \tag{2.14}$$

Kronecker product

The Kronecker product [93] of matrices $A \in \mathbb{R}^{I \times J}$ and $B \in \mathbb{R}^{H \times M}$ is denoted by $A \otimes B$. The result is a matrix of size $(IH) \times (JM)$ and defined by :

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1J}B \\ a_{21}B & a_{22}B & \cdots & a_{2J}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{I1}B & a_{I2}B & \cdots & a_{IJ}B \end{bmatrix}$$
$$= \begin{bmatrix} a_1 \otimes b_1 & a_1 \otimes b_2 & a_1 \otimes b_3 & \dots & a_j \otimes b_{M-1} & a_J \otimes b_M \end{bmatrix} \tag{2.15}$$

The Kronecker product [8] of matrices $A \in \mathbb{R}^{I_1 \times I_2}$ and $B \in \mathbb{R}^{J_1 \times J_2}$ is denoted by $A \otimes B$. The result is a matrix of size $(I_1 I_2 \times J_1 J_2)$ and defined by:

$$A \otimes B = \begin{pmatrix} a_{11}B & \cdots & a_{1I_2}B \\ \vdots & \ddots & \vdots \\ a_{I_1 1}B & \cdots & a_{I_1 I_2}B \end{pmatrix} \tag{2.16}$$

For example :
$$\begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & 3 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 \times 0 & 1 \times 3 & 2 \times 0 & 2 \times 3 \\ 1 \times 2 & 1 \times 1 & 2 \times 2 & 2 \times 1 \\ 3 \times 0 & 3 \times 3 & 1 \times 0 & 1 \times 3 \\ 3 \times 2 & 3 \times 1 & 1 \times 2 & 1 \times 1 \end{pmatrix} = \begin{pmatrix} 0 & 3 & 0 & 6 \\ 2 & 1 & 4 & 2 \\ 0 & 9 & 0 & 3 \\ 6 & 3 & 2 & 1 \end{pmatrix}$$

2.2.1.3 Multilinear projections

A Multilinear subspace [7] is defined through a multilinear projection that maps the input tensor data from one space to another (lower-dimensional) space. Therefore, multilinear projection is a key concept in Multilinear Subspace Learning (MSL). There are three basic multilinear projections based on the input and output of a projection: the traditional Vector-to-Vector Projection (VVP), Tensor-to-Tensor Projection (TTP), and Tensor-to-Vector Projection (TVP). In our work we use Tensor-to-Tensor Projection (TTP) because is the best [92].

Tensor-to-tensor projection

A tensor can be projected to another tensor of the same order, named as TTP. As an m th-order tensor $\mathbf{X}$ resides in the tensor space $\mathbb{R}^{I1} \otimes \mathbb{R}^{I2} ... \otimes \mathbb{R}^{Im}$, the tensor space can be viewed as the Kronecker product of m vector (linear) spaces $\mathbb{R}^{I1} \otimes \mathbb{R}^{I2} ... \otimes \mathbb{R}^{Im}$. To project a tensor $\mathbf{X}$ in a tensor space $\mathbb{R}^{I1} \otimes \mathbb{R}^{I2} ... \otimes \mathbb{R}^{Im}$ to another tensor $\mathbf{Y}$ in a lower-dimensional tensor space $\mathbb{R}^{P1} \otimes \mathbb{R}^{P2} ... \otimes \mathbb{R}^{Pm}$ where $P_k < I_k$ for all k , m projection matrices { $U^{(k)} \in \mathbb{R}^{I_K \times P_K}, k = 1, ..., K$} are used so that :

$$\mathbf{Y} = \mathbf{X} \times_1 U^{(1)^T} \times_2 U^{(2)^T} ... \times_m U^{(m)^T} \tag{2.17}$$

It can be done in m steps, where in the kth step, each k-mode vector is projected to a lower dimension P_k by $U^{(k)}$, as shown in Figure 2.13 (a) and (b) demonstrates how to project a tensor in 1-mode using a 1-mode projection matrix , which projects each 1-mode vector of the original tensor to a low-dimensional vector.
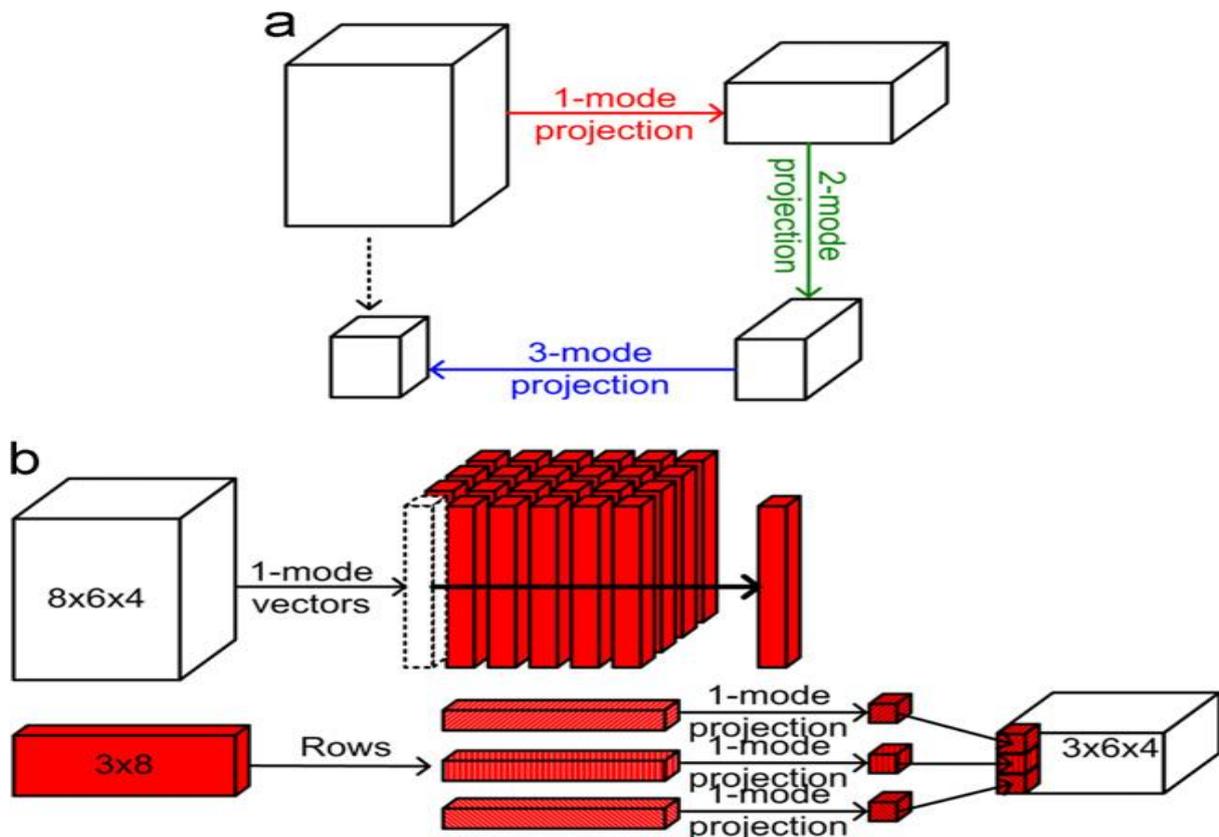
Figure 2.13 : Illustration of tensor-to-tensor projection : (a) projection of a tensor in all modes and (b) projection of a tensor in one mode [7].

2.3 **Tensor Exponential Discriminant Analysis (TEDA)**

Exponential discriminant analysis (EDA) [95] has been proposed to address Small Sample Size (SSS) problem as well as to preserve the discrimination achieved by the null space of the within-class scatter matrix in LDA. This idea has been successfully generalized to exponential dimensionality reduction approaches [96]. Inspired by these methods, Ouamane and al. [92] extend MDA [10], by introducing the exponentiation, to TEDA for tensor dimensionality reduction and discrimination improvement. We will first recall the principle of MDA, then we present its extension to TEDA.

2.3.1 **MDA**

Let the training samples represented as an mth -order tensors $\mathbf{A} \epsilon \ \mathbb{R}^{I_1 \times I_2 \times \dots \times I_m}$ belonging to L different classes and each class j contains n_j samples. MDA seeks m interrelated projection matrices U_k^* by maximizing the inter-class scatter while minimizing the intra-class scatter in each mode of the training tensor:

$$U_K^*|_{k=1}^m = \underset{U_K|_{k=1}^m}{\text{argmax}} \times \frac{\sum_{j=1}^L n_j \left\| \bar{\mathbf{A}}_j \times_1 U_1 \dots \times_m U_m - \bar{\mathbf{A}} \times_1 U_1 \dots \times_m U_m \right\|^2}{\sum_{i=1}^N \left\| \mathbf{A}_i \times_1 U_1 \dots \times_m U_m - \bar{\mathbf{A}}_{n_i} \times_1 U_1 \dots \times_m U_m \right\|^2} \qquad (2.18)$$

Where N is the number of training samples, $\bar{\mathbf{A}}_j$ is the average tensor of each class j, $\bar{\mathbf{A}}$ is the average tensor of all the training data.

Eq. (2.1) is a higher order nonlinear optimization problem with a higher order nonlinear constraint. Finding a straightforward closed solution is not obvious. Therefore, an iterative optimization approach to estimate the interrelated discriminative subspaces is proposed [10]. First, considering the optimization problem from each k-mode, we have the following objective function:

$$U_k^* = \underset{U_K}{\text{argmax}} \frac{\sum_{j=1}^L n_j \left\| \bar{\mathbf{A}}_j \times_K U_K - \bar{\mathbf{A}} \times_K U_K \right\|^2}{\sum_{i=1}^N \left\| \mathbf{A}_i \times_K U_K - \bar{\mathbf{A}}_{n_i} \times_K U_K \right\|^2} \qquad (2.19)$$

This optimization problem is reduced to a special discriminant analysis where the tensor is unfolded in the k-mode and the column vectors of the unfolded matrices are labeled

with the tensor's original label. The optimization problem in Eq. (2.19) is reformulated as a special discriminant analysis as:

$$U_k^* = \underset{U_K}{\operatorname{argmax}} \frac{\operatorname{Tr}(U_K^T S_b U_k)}{\operatorname{Tr}(U_K^T S_w U_k)} \tag{2.20}$$

Where

$$S_b^k = \sum_{j=1}^{L} n_j \left(\overline{A}_j^k - \overline{A}^k\right)\left(\overline{A}_j^k - \overline{A}^k\right)^T \tag{2.21}$$

And

$$S_w^k = \sum_{j=1}^{L} \sum_{i=1}^{n_j} \left(\overline{A}_{j,i}^k - \overline{A}_j^k\right)\left(\overline{A}_{j,i}^k - \overline{A}_j^k\right)^T \tag{2.22}$$

Are the between and within class scatter matrices, respectively. Here, $\overline{A}_{j,i}^k$ is the k-mode unfolded matrix of tensor $\mathbf{A}_i$, $\overline{A}_j^k$ the average matrix on class j and $\overline{A}^k$ is the average matrix of the whole training data. We refer the reader to [10] for the underlying mathematical details.

In each iteration of the solution of Eq. (2.18), $U_1, U_2 ..., U_m$ are supposed to be known, leading to:

$$U_K^*|_{k=1}^m = \underset{U_K}{\operatorname{argmax}} \times \frac{\sum_{j=1}^{L} n_j \left\| \overline{\mathbf{A}}_j \times_1 U_1 ... \times_m U_m - \overline{\mathbf{A}} \times_1 U_1 ... \times_m U_m \right\|^2}{\sum_{i=1}^{N} \left\| \mathbf{A}_i \times_1 U_1 ... \times_m U_m - \overline{\mathbf{A}}_{n_i} \times_1 U_1 ... \times_m U_m \right\|^2} \tag{2.23}$$

letting:

$$\mathbf{B}_i = \mathbf{A}_i \times_i U_1 ... \times_{k-1} U_{k-1} \times_{k+1} U_{k+1} ... \times_m U_m \tag{2.24}$$

We obtain

$$U_k^* = \underset{U_K}{\operatorname{argmax}} \frac{\sum_{j=1}^{L} n_j \left\| \overline{\mathbf{B}}_j \times_K U_K - \overline{\mathbf{B}} \times_K U_K \right\|^2}{\sum_{i=1}^{N} \left\| \mathbf{B}_i \times_K U_K - \overline{\mathbf{B}}_{n_i} \times_K U_K \right\|^2} \tag{2.25}$$

This is equivalent to Eq. (2.19) by changing $\mathbf{A}_i$ by $\mathbf{B}_i$. Therefore, it can be resolved using Eq. (2.20). Consequently, the projection matrices are iteratively optimized for each mode. The iterative process of MDA breaks when either a maximal number of iterations is reached or

$$\left\|U_k^{\text{iteration}} - U_k^{\text{iteration}-1}\right\| < n_k' n_k \varepsilon \text{ where } U_k^{\text{iteration}} \in \mathbb{R}^{n_k' n_k}.$$

2.3.2 Matrix Exponential

We need the following definition and properties for introducing TEDA. Let A be an $m \times m$ square matrix, its exponential, denoted by $\exp(A)$ or $e^{(A)}$, is defined as :

$\exp(A) = \sum_{i=0}^{\infty} \frac{A^i}{i!} = I + A + \frac{A^2}{2!} + \cdots + \frac{A^n}{n!} + \cdots$, where I *is* the unit matrix with similar size of A. Following are some properties of the matrix exponential :

1) $\exp(0) = I.$
2) $\exp(A)$ is a full rank matrix.
3) If $AB = BA$, then $\exp(A + B) = \exp(A)\exp(B) = \exp(B)\exp(A).$
4) $(\exp(A))^{-1} = \exp(-A).$
5) For the invertible matrix W, $\exp(W^{-1}AW) = W^{-1}\exp(A)W.$
6) If $A = diag(a_1, a_2, \dots, a_m)$ is a diagonal matrix, $\exp(A) = diag(\exp(a_1), \exp(a_2), \dots, \exp(a_m)).$
7) $|\exp(A)| = \exp(tr(A)).$
8) If $(v_1, v_2, \dots, v_m)$ are eigenvectors for matrix A and $\lambda_1, \lambda_2, \dots, \lambda_m$ the corresponding eigenvalues, then $(v_1, v_2, \dots, v_m)$ are also eigenvectors of $\exp(A)$ that correspond to the eigenvalues $\exp(\lambda_1), \exp(\lambda_2), \dots, \exp(\lambda_m).$

2.3.3 TEDA

Using the eigenvalue decomposition, the Kth mode the projection matrix, U_k^* in Eq. (2.20), is rewritten as:

$$U_k^* = \underset{U_K}{\text{argmax}} \; \frac{\text{Tr}\,(U_K^T\,(Y_b^T \Lambda_b Y_b) U_k)}{\text{Tr}\,(U_K^T (Y_w^T \Lambda_w Y_b) U_k)} \tag{2.26}$$

Where $Y_b = (v_{b_1}, v_{b_1}, \dots, v_{b_m})$ is the eigenvector matrix of S_b and
$\Lambda_b = diag\,(\lambda_{b_1}, \lambda_{b_1}, \dots, \lambda_{b_m})$ represent the corresponding eigenvalues.
$Y_w = (v_{w_1}, v_{w_1}, \dots, v_{w_m})$ is the eigenvector matrix of S_w and
$\Lambda_w = diag\,(\lambda_{w_1}, \lambda_{w_1}, \dots, \lambda_{w_m})$ represent the corresponding eigenvalues.

S_w is not full-rank matrix under the small sample size situation. In this case, the discriminant data related to the null eigenvalues of S_w has the best discriminant power [96]. However, in MDA, this data is discarded by the projection. To prevent this issue, in TEDA, the authors in [92] introduced the expectational by changing λ_{w_i} the eigenvalues of S_w by $\exp(\lambda_{w_i})$. Hence, the objective function in Eq. (2.26) become :

$$U_k^* = \underset{U_K}{\operatorname{argmax}} \frac{\operatorname{Tr}\left(U_K^T\left(\Upsilon_b^T \exp(\Lambda_b)\Upsilon_b\right)U_k\right)}{\operatorname{Tr}\left(U_K^T\left(\Upsilon_w^T \exp(\Lambda_w)\Upsilon_b\right)U_k\right)} \tag{2.27}$$

Applying the property 8) of experiential matrix, Eq. (2.26) become :

$$U_k^* = \underset{U_K}{\operatorname{argmax}} \frac{\operatorname{Tr}\left(U_K^T\left(\exp(S_b)\right)U_k\right)}{\operatorname{Tr}\left(U_K^T\left(\exp(S_w)\right)U_k\right)} \tag{2.28}$$

Based on property 2), the matrix $\exp(S_w)$ is a full-rank matrix. Consequently, the discriminant data included in the null space of S_w can be preserved by equation Eq. (2.28). The optimal projection matrix U_k^* for each k-mode and iteration comprises to the first leading $n_{(k)}$ eigenvectors of $\exp(S_b)\Upsilon = \Lambda \exp(S_w)\Upsilon$, where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{n_{(k)}}$.

2.3.4 Justifications

In pattern recognition in general and face recognition in particular, kernel methods are usually used for original data with nonlinear constructions. The objective of kernel transformation is to map the nonlinear problems in the original data to another space of higher dimension with linear problems. Similarly, to the kernel methods, exponential transformation maps the within-class $\left(S_w^k\right)$ and between-class $\left(S_b^k\right)$ scatter matrices into another nonlinear space by:

$$\Omega: \mathbb{R}^{g \times g} \rightarrow \mathbb{R}^{g \times g}$$
$$S_b^k \rightarrow \Omega\left(S_b^k\right) = \exp\left(S_b^k\right)$$
$$S_w^k \rightarrow \Omega\left(S_w^k\right) = \exp\left(S_w^k\right) \tag{2.29}$$

Then, TEDA might keep some related properties of the kernel method. Therefore, TEDA works better than MDA when dealing with problem of nonlinearity.

The objective function for discriminant MDA criterion is to simultaneously increase the between-class distance (distance$_b$) and decrease the within-class distance (distance$_w$) for each k-mode of the tensor. These two distances can be computed by trace of two scatter matrices:

$$(\text{distance}_b) = trace(S_b^k) = \sum_{j=1}^{L} n_j \left\| \overline{B}_j^k - \overline{B}^k \right\|_2^2 = \lambda_{b_1} + \lambda_{b_1} + \cdots + \lambda_{b_{n_{(k)}}}$$

$$(\text{distance}_w) = trace(S_w^k) = \sum_{j=1}^{L} \sum_{i=1}^{n_j} \left\| B_{j,i}^k - \overline{B}_j^k \right\|_2^2 = \lambda_{w_1} + \lambda_{w_1} + \cdots + \lambda_{w_{n_{(k)}}}.$$

where the matrix B^{K_i} is defined in equation Eq. (2.24) and Eq. (2.1). By applying exponential and using property 7) we get:

$$\text{distance}_b' = trace(\exp(S_b^k)) = \exp(\lambda_{b_1}) + \exp(\lambda_{b_2}) + \cdots + \exp(\lambda_{b_{n_{(k)}}})$$

and

$$\text{distance}_w' = trace(\exp(S_w^k)) = \exp(\lambda_{w_1}) + \exp(\lambda_{w_2}) + \cdots + \exp(\lambda_{w_{n_{(k)}}})$$

The eigenvalues λ_{b_i} and λ_{w_i} are mapped to $\exp(\lambda_{b_i})$ and $\exp(\lambda_{w_i})$, respectively. The biggest eigenvalue is the more important and the smallest eigenvalue is the lowest important. Figure 2.14 illustrates an example of the proportion of λ_i and $\exp(\lambda_i)$ to their respective sums. In this figure, the largest eigenvalue $\lambda_7 = 20\%$ while its corresponding exponential $\exp(\lambda_7) = 65.70\%$ and the smallest eigenvalue $\lambda_1 = 7\%$ while its corresponding exponential $\exp(\lambda_1) = 10^{-4} \%$. We have : $\frac{\lambda_{b_i}}{\lambda_{w_i}} < \frac{\exp(\lambda_{b_i})}{\exp(\lambda_{w_i})}$. Thus, the small eigenvalues are reduced and large eigenvalues are enlarged. This affects the distances between the projected tensors as the inter-samples distances within the same class are smaller than the related distances between samples from different classes.
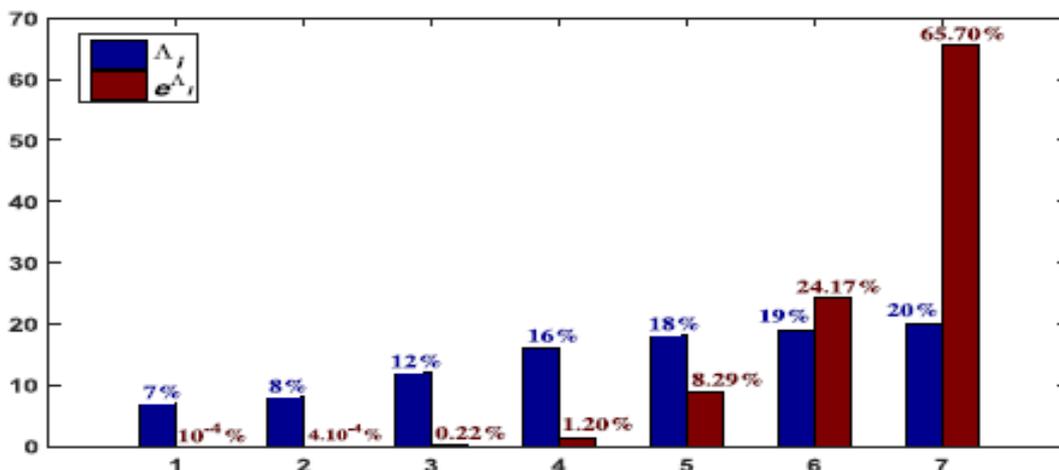
Figure 2.14 : Example of the proportions $\frac{\lambda_i}{\sum \lambda_i}$ (blue bars) and $\frac{\exp(\lambda_i)}{\sum \exp(i)}$ (red bars) [92].

2.3.5 Within Class Covariance Normalization

The within Class Covariance Normalization (WCCN) has been used mostly in the speaker recognition community. It is found in [97] that the best approach is to project the LDA reduced i-vectors to a new subspace specified by the square-root of the inverse of the within-class covariance matrix. In [92] the authors proposed a variant TEDA by integrating WCCN:

$$W_k = \sum_{j=1}^{L} \sum_{i=1}^{n_j} \frac{(U_k^*)^T B_{j,i}^k - (U_k^*)^T \overline{B_j}^k}{(U_k^*)^T B_{j,i}^k - (U_k^*)^T \overline{B_j}^k} \tag{2.30}$$

where, U_k^* is the TEDA projection matrix for k-mode. The WCCN projection matrix W_k is obtained by Cholesky decomposition of the inverse of W_k :

$W_K^{-1} = C_k C_k^T$. WCCN reduces the effect of the within class variations by minimizing the expected classification error on the training data. This is achieved by imposing upper bounds on the classification error metric [98]. The procedure of this proposed variant Tensor Exponential Discriminant Analysis integrating within class covariance normalization (TEDA+WCCN), is detailed in the next chapter.

2.4 Conclusion

In this chapter, we have introduced the fundamental principles to good understanding of multilinear algebra. Thus, we were able to define the concept of tensor which allows us to work directly with the data in the form of multidimensional tables. In the first part of this chapter, we presented, through the properties of multilinear algebra, the tools necessary to understand the tensor approach and we define what we call «tensor» with such formalism. Next, we studied the three basic types of multilinear projection. Finally, We first recalled the principle of MDA, then we presented its extension to TEDA.

3.1 Introduction

Age estimation accuracy depends on how well the input images have been represented by good general discriminative features. The choice of classification has an impact on the result of the estimated age for unknown image. In this chapter, we present a complete age estimation framework with description of each component. The output of this framework is the estimated age for the input face image.

We will citing, the different steps applied for estimating the age of faces from facial images as follows:

- Facial preprocessing, the face region should be cropped (aligned face) .
- We use deep features provided by three pre-trained Deep CNN, as facial representations.
- We use the TEDA method explained in chapter 2 to analyze different facial image representations (tensor facial representations).
- Estimation of the age using Random Forest classifier

3.2 Age estimation framework

The proposed system consists of the complete pipeline for age estimation. Given an image, we first crop the face using the preprocessing step. Then, Each aligned face is passed through a deep CNNs in features extraction step. We then each tensor facial representations is passed through a multidimensional analysis by TEDA method for reduction and separation between ages of the multidimensional features (Subspace transformation). Finally, a classification-based Random-Forest is performed on these features to estimate the facial age. The system architecture is illustrate in figure 3.1. The details of each component are presented in the following sub-sections.

In the offline (training) stage, the optimal multilinear projection matrices TEDA method are estimated, and in online (test) stage, new samples are projected by these tensors. The training 3rd order tensor $\mathbf{X} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ is constructed using the vectors   of   different   deep descriptors extracted from the preprocessed face images of the training database .The three modes of the tensor $\mathbf{X}$ are defined as follows:

1- I_1 represents weights of deep descriptors.

2- I_2 represents the vectors of deep descriptors .

3- I_3 represents face samples of all the persons of the database.

During the training phase, the input tensor $\mathbf{X}$ is reduced according to I_1 and I_2 modes and projected in another subspace based on the proposed method TEDA. Then, we obtain a reduced tensor $\mathbf{Y} \in \mathbb{R}^{L_1 \times L_2 \times L_3}$, where $L_1 \times L_2 \ll I_1 \times I_2$.

During the test phase, the test matrix X_i candidate of size $I_1 \times I_2$ is projected into the tensor subspace using TEDA in order to get a reduced subspace features Y_i of size $L_1 \times L_2$ where $L_1 \times L_2 \ll I_1 \times I_2$.
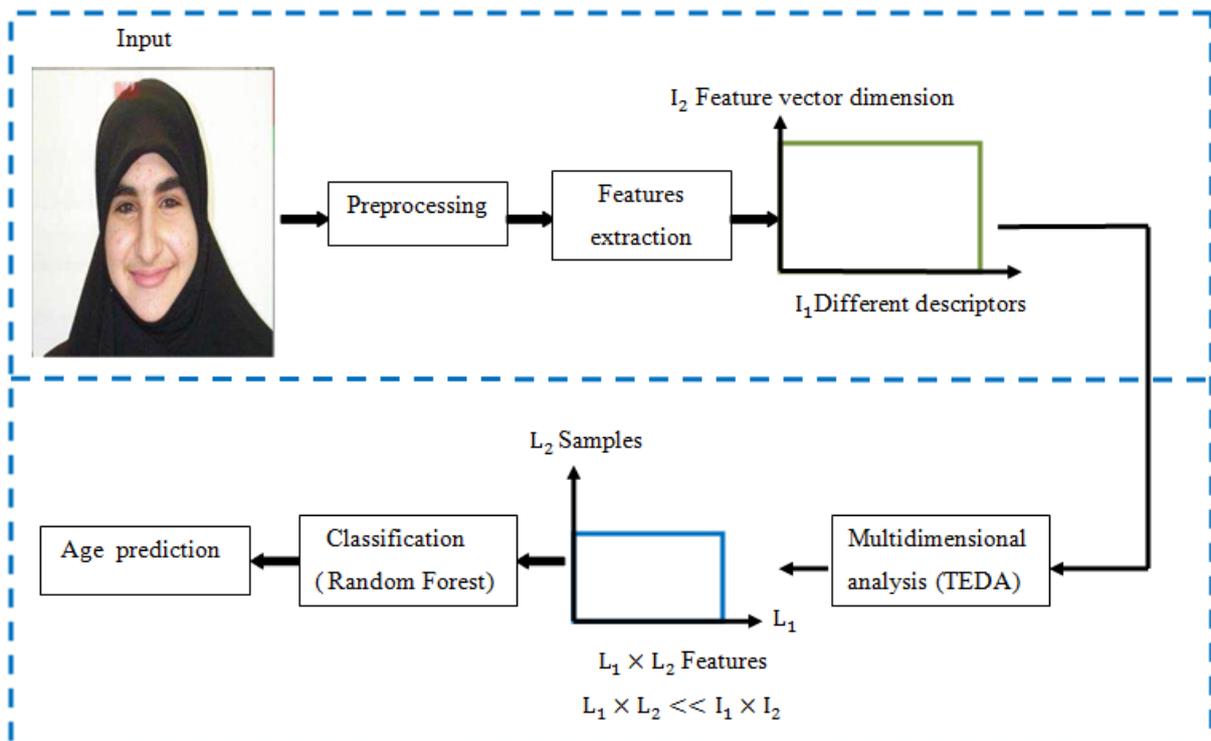
Figure 3.1: Overview of the proposed framework for facial age estimation.

3.2.1 Preprocessing

One of the most significant stages in image-based age estimation is face arrangement. In the presented work, eyes of each face are distinguished utilizing the Ensemble of Regression Trees (ERT) method [99], which is a powerful and productive calculation for facial landmark localization. When we have the 2D places of the two eyes, we use them to make up for the in-plane pivot of the face. In the wake of playing out the pivot and reusing, the face area ought to be trimmed (adjusted face) [100].

Original image

Alignment and cropped face
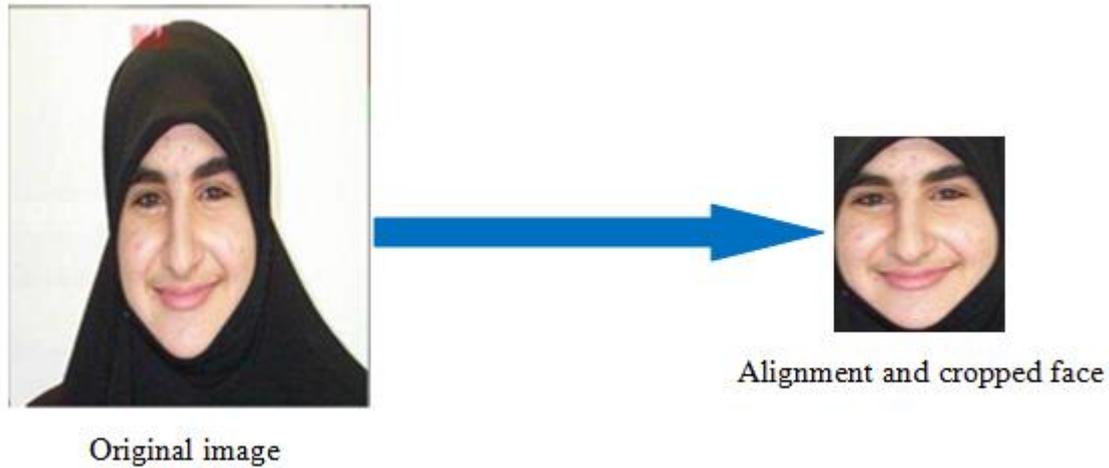
Figure 3.2: Face alignment and cropping associated with one original image in PAL database.

3.2.2 Feature Extraction

Convolutional neural network pre-trained on a large image dataset can be used as a feature extractor for other closely related datasets (PAL database in our case). Reusing a learned model in another domain or task is known as transfer learning [101]. Obviously, transfer learning only concerns the deep CNNs. For transfer learning, we use deep features provided by three pre-trained Deep CNN [102] [103]:

✓ VGG-face [104] : was trained on face images for the purpose of face identification.

✓ ImageNet [105] : was trained on images of objects for the purpose of image categorization .

✓ DEX-ChaLearn-ICCV2015 [106, 25]: was trained on face images for the purpose of age estimation. One can also notice was trained on apparent age.

All deep nets are used as feature extractors. All deep features are given by 4096 elements. Briefly, we will describe this features that used for extracting face features here :

➢ **Visual Geometry Group (VGG) Face Features** :

This CNN comprises 11 blocks, each containing a linear operator followed by one or more non-linearities such as ReLU and max pooling [104]. The CNN architecture is given in full detail in figure 3.3. The first eight such blocks are said to be convolutional as the linear operator is a bank of linear filters (linear convolution). The last three blocks are instead called Fully Connected (FC); they are the same as a convolutional layer, but the size of the filters matches the size of the input data, such that each filter senses data from the entire image. All

the convolution layers are followed by a rectification layer (ReLU). The first two FC layers output are 4096 dimensional vectors. This multi-way CNN is trained to discriminate between the 2,622 identities using about 2.6 million images. The deep features of this network are extracted by taking the 4K dimensional features and removing the last classification layer. The resulting vector is L2 normalized [102]. The 4096-dimensional vector of the second fully connected layer 'fc7 ' is used as the feature (shown in figure 3.4).
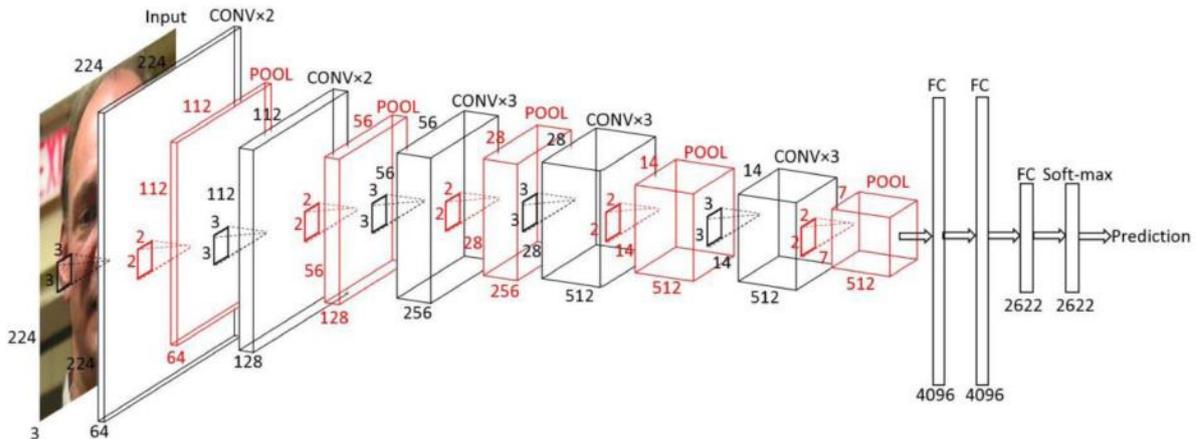
Figure 3.3: VGG-Face architecture, CONV indicates convolutional layers, POOL indicates pooling layers and FC indicates fully-connected layers. For each layer, the filter size, number of filters, size of resulted feature maps are also indicated [107].
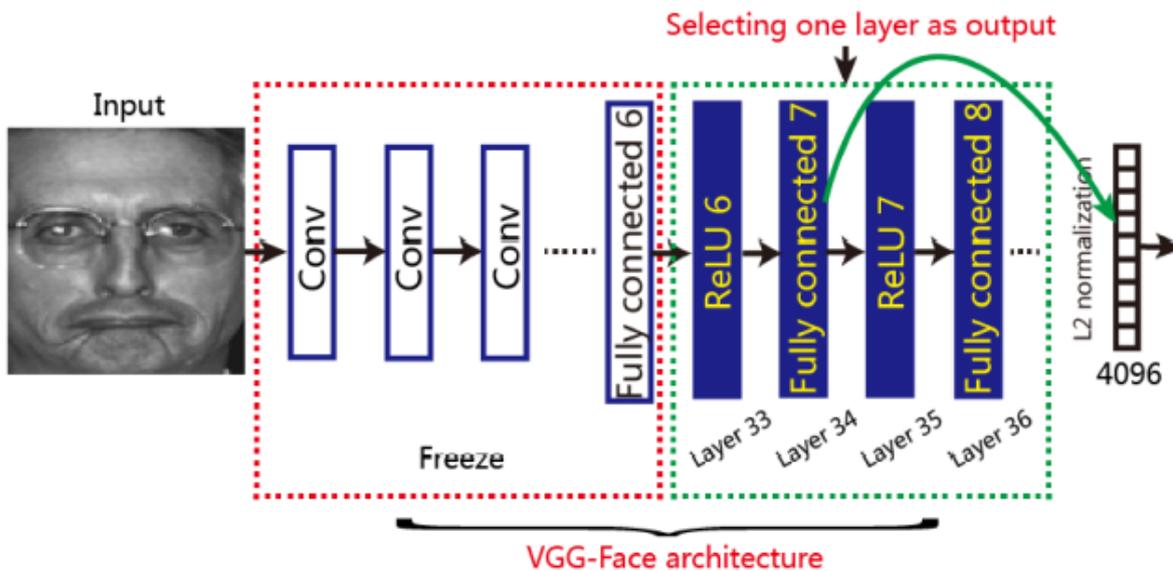
Figure 3.4 : Selecting one layer as output (using the 34th layer as output) [108].

> **ImageNet VGG-F Features** :

The Fast (VGG-F) architecture [105] (shown in Table 3.1) is similar to the one used by Krizhevsky et al. [109]. It comprises 8 learnable layers, 5 of which are convolutional, and the last 3 are fully-connected. The input image size is 224×224. Fast processing is ensured by the 4 pixel stride in the first convolutional layer. The main differences between this architecture and that of Krizhevsky are the reduced number of convolutional layers and the dense connectivity between convolutional layers (Krizhevsky used sparse connections to enable training on two GPUs). The network was trained on ILSVRC-2012 (The ImageNet Large Scale Visual Recognition Challenge) using gradient descent with momentum. The hyper-parameters are the same as used by Krizhevsky. The authors applied data augmentation in the form of random crops, horizontal flips, and RGB color jittering. We extracted the deep features from the 4K dimensional feature vector after removing the last classification layer. The resulting vector is L2 normalized. The only image pre-processing consists on resizing the input images to the network input size and subtracting the average image (provided by the authors in the network metadata) [102].

Arch.	Conv1	Conv2	Conv3	Conv4
CNN-F	64x11x11 st. 4, pad 0 LRN, x2 pool	256x5x5 st. 1, pad 2 LRN, x2 pool	256x3x3 st. 1, pad 1 -	256x3x3 st. 1, pad 1 -
	Conv5	Full6	Full7	Full8
	256x3x3 st. 1, pad 1 x2 pool	4096 drop-out	4096 drop-out	1000 soft-max

Table 3.1: ImageNet architecture [105].

Fast (CNN-F) architecture .this architecture contains 5 convolutional layers (conv 1–5) and three fully-connected layers (full 1–3). The details of the convolutional layers are given in three sub-rows: the first specifies the number of convolution filters and their receptive field size as "num x size x size"; the second indicates the convolution stride ("st.") and spatial padding ("pad"); the third indicates if Local Response Normalisation (LRN) [109] is applied, and the max-pooling down sampling factor. For full 1–3, we specify their dimensionality, which is the same for all three architectures. Full6 and full7 are regularised using dropout

[109], while the last layer acts as a multi-way soft-max classifier [105].in our work we use the second fully connected layer 'fc7 '.

> ➢ **DEX-ChaLearnFeatures** :

The Deep Expectation(DEX) on apparent age method [106, 25] uses the VGG-16 architecture (shown in figure 3.5) for its networks, which are pre-trained on ImageNet for image classification. In addition, the authors explored the benefit of fine-tuning over crawled Internet face images with available age. In total, they collected more than 500,000 images of celebrities from IMDb and Wikipedia. The networks of DEX were fine-tuned on the crawled images and then on the provided images with apparent age annotations from the ChaLearn LAP 2015 challenge on apparent age estimation. We extracted the features provided by DEX-ChaLearn-ICCV2015 network trained on apparent age using the challenge images. An ensemble of these models led to 1st place at the challenge (115 teams). The 4K features are collected from the previous to the last FC layer [102].The 4096-dimensional vector of the second fully connected layer 'fc7 ' is used as the feature.
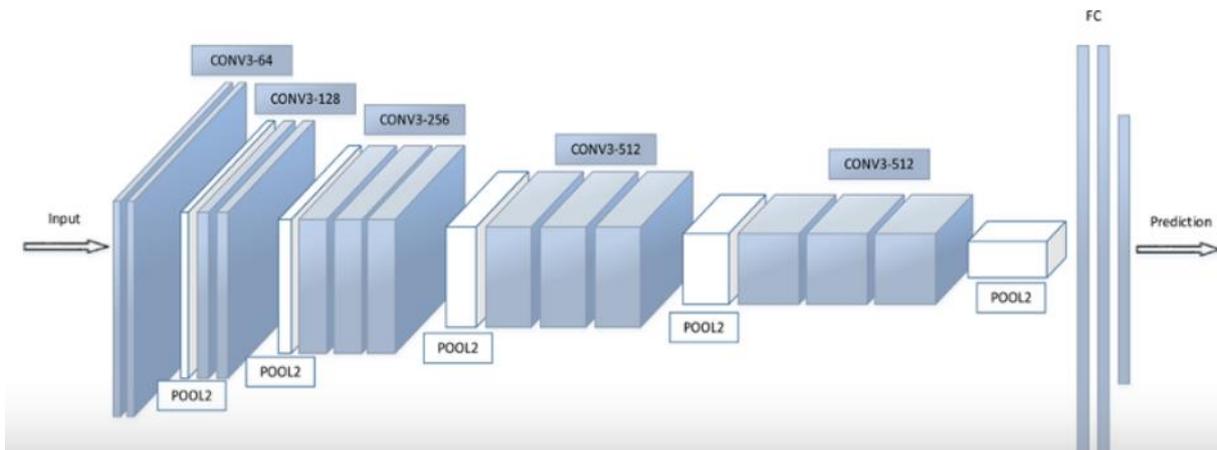
Figure 3.5 : VGG-Very-Deep-16 CNN architecture [101].

3.2.3 Multidimensional analysis using TEDA

This part is mainly based on reduction of the multidimensional features of the face. We use the TEDA+WCCN (Tensor Exponential Discriminant Analysis plus Within Class Covariance Normalization) method. The procedure of this variant, Tensor Exponential

Discriminant Analysis integrating within class covariance normalization (TEDA+WCCN), is detailed in the following algorithm [92] as follows:

Algorithm: TEDA+WCCN

Inputs :

- The tensor $\widetilde{\mathbf{A}} \in \mathbb{R}^{I_1 \times I_2 \times \dots I_m \times N}$ of the N training samples belonging to L classes, each class $\mathbf{A_j}$ contains $\mathbf{n_j}$ samples.
- itr_{max} is the maximal number of iterations .
- The final lower dimensions are $I'_1 \times I'_2 \times \dots \times I'_m$.

Outputs : The projections $U_k = U_k^t \in \mathbb{R}^{I_k \times I'_k}, k = 1, \dots, m.$

  1) **Initialization :** $U_1^0 = I_{I_1}, U_2^0 = I_{I_2}, \dots, U_m^0 = I_{I_m}$

  2) **For** $t = 1$ to itr_{max}

    a) **For** k=1 to m

       - $\boldsymbol{B}_i = \boldsymbol{A}_i \times_1 U_1^{t-1} \dots \times_{k-1} U_{k-1}^{t-1} \times_{k+1} U_{k+1}^{t-1} \dots \times_m U_m^{t-1}$

       - $B_i^k k \Leftarrow \mathbf{B}_i$

       - $S_b^k = \sum_{j=1}^{L} n_j \left(\overline{B}_j^k - \overline{B}^k \right) \left(\overline{B}_j^k - \overline{B}^k \right)^T$

       - $S_w^k = \sum_{j=1}^{L} \sum_{i=1}^{n_j} \left(B_{j,i}^k - \overline{B}_j^k \right) \left(B_{j,i}^k - \overline{B}_j^k \right)^T$

       - Compute the matrices: $\exp(S_b^k)$ and $\exp(S_w^k)$

       - $\exp(S_b^k)U_k^t = \Lambda \exp(S_w^k)U_k^t$, obtain $U_k^t \in \mathbb{R}^{I_k \times I'_k}$

       - $W_k = \sum_{j=1}^{L} \sum_{i=1}^{n_j} \left((U_k^t)^T B_{j,i}^k - (U_k^t)^T \overline{B}_j^k \right) \left((U_k^t)^T B_{j,i}^k - (U_k^t)^T \overline{B}_j^k \right)^T$

       - Compute WCCN projection matrix (C_k) : $W_k^{-1} = C_k C_k^T$

       - Compute the new U_k^t : $U_k^t = C_k^T U_k^t$

    b) **If** $t > 2$ and $\left\| U_k^t - U_k^{t-1} \right\| < I_k \times I_k \varepsilon$, $k = 1, \dots, m$,   break ;

  3) **Projection:** The feature tensor after projection is obtained by:

$$\mathbf{B}_i = \mathbf{A}_i \times_1 U_1 \times_2 U_2 \dots \times_m U_{(m)}, i = 1, \dots, m.$$

3.2.4 Classification Random Forest

Random forests are strategies for acquiring prescient models for classification and regresion. The technique executes binary decision trees, including CART trees proposed by Breiman et al. (1984).The general idea behind the method is that instead of trying to get an optimized method at once, we generate several predictors before pooling their different predictions. Use the feature to classify or regress a sample of observations described by qualitative and/or quantitative variables. In classification (variable qualitative response): the method predicts the affiliation of observations (observations, individuals) to a class of a qualitative variable, based on quantitative and/or qualitative explanatory variables. Figure 3.6 illustrate the Random Forest classifier, where each class vector is generated by counting the percentage of different classes of training examples at the leaf node where the concerned instance falls and then averaging across all trees in the same forest [100].
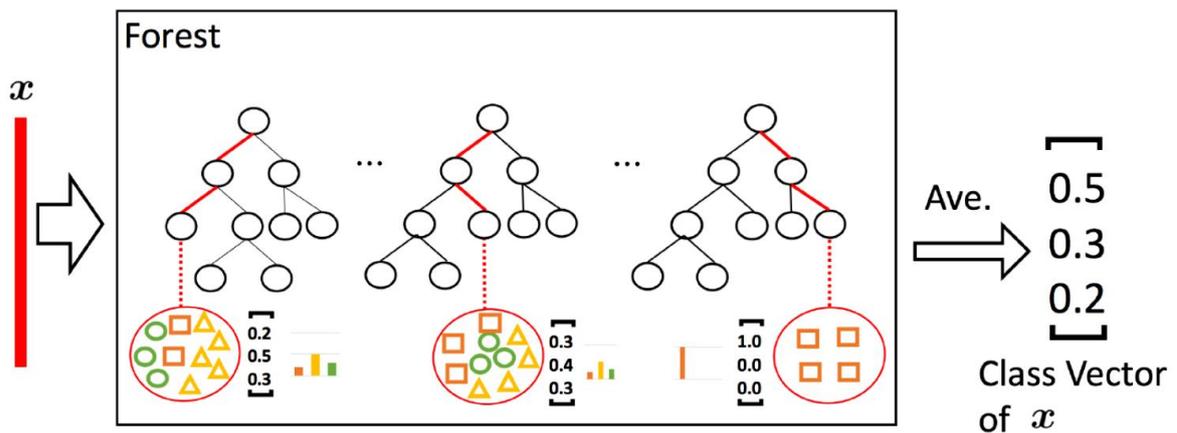
Figure 3.6: Illustration of Random Forest classifier [100].

3.3 Experiments

In this section, we present the experimental evaluation of the proposed approach and discuss the obtained results.

3.3.1 PAL database

The dataset used in the experiments is the Productive Aging Lab Face (PAL) database, from the University of Texas at Dallas [110] contains 1046 frontal face images from different

subjects (430 males and 616 females) in the age range of 18 to 93 years old. The PAL database can be divided into three main ethnicities: African-American subjects (208 images), Caucasian subjects (732 images) and other subjects (106 images). The database images contain faces having different expressions. Some samples are illustrated in figure 3.7. For the evaluation of the approach, we conduct again a 5-fold cross-validation(used for training and testing purposes) [102] [103].We perform the random partition as in [45], [103], where we randomly partition images in 80% training and the other 20% for testing. It is repeated five times. The average of the five different splits will be the final performance [100].

Figure 3.7: Sample images from PAL database [102].

3.3.2 Evaluation Protocol

The performance of the age estimation is evaluated by MAE (Mean Absolute Error), i.e., the average absolute difference between the estimated age and the real age. We used five-fold cross-validation that allows testing every test image in the considered database.

3.3.3 Results and Discussion

In the experiments, we considered three cases:

1- Maximal number of iterations
2- Calculate the final lower dimension values in mode-1 and mode-2
3- Calculate the number of trees.

3.3.3.1 Maximal number of iterations

Table 3.2 illustrates the MAE obtained on the PAL database using the subspace transformation with different maximal number of iterations.

Maximal number of iterations	Reduced subspace $I_2' = 3$		
	$I_1' = 50$	$I_1' = 100$	$I_1' = 150$
2	**2.9924**	**3.0686**	**5.5333**
3	3.1067	3.1724	5.3514
4	3.1381	3.2505	5.4152
10	3.0200	3.5343	5.5629
20	3.0496	3.3686	5.3838
50	3.2200	3.4610	5.5486
100	3.1133	3.3038	5.8276
200	3.0552	3.5448	5.4533

Table 3.2 : Mean Absolute Error (in Years) of age estimation with different final lower dimension values in mode-1 (where other modes values are fixing) on the PAL Database .

We can observe in this table that:
- Mean Absolute Error changes with the variation of Maximal number of iterations.
- Maximal number of iterations which equal 2 give the best result. So we set Maximal number of iterations to 2.

3.3.3.2 Calculate the final lower dimension values in mode-1 and mode-2

Table 3.3 shows a comparison between the MAE of the age estimation by the use of best maximal number of iterations (itr_{max}) that obtained from our previous experiments. Our facial analysis has MAE of 3.07 , 3.23 and 4.12 years for levels (I_1 =60 I_2 =3) , (I_1 =40 I_2 =2) and (I_1 =20 I_2 =1) respectively.

itr$_{max}$ = 2					
I$'_1$, I$'_2$	MAE	I$'_1$, I$'_2$	MAE	I$'_1$, I$'_2$	MAE
20 , 3	3.2829	20 , 2	3.6505	20 , 1	**4.1248**
40 , 3	3.1381	40 , 2	**3.2314**	40 , 1	4.8486
60 , 3	**3.0695**	60 , 2	3.3829	60 , 1	5.0543
80 , 3	3.2895	80 , 2	3.5762	80 , 1	4.8152
100 , 3	3.3381	100 , 2	3.8362	100 , 1	4.6333
120 , 3	3.9829	120 , 2	5.6771	120 , 1	4.4819
140 , 3	5.6390	140 , 2	5.2152	140 , 1	4.7371
160 , 3	5.4286	160 , 2	4.9143	160 , 1	4.5829
180 , 3	5.4448	180 , 2	5.3190	180 , 1	4.4629
200 , 3	5.8590	200 , 2	5.2105	200 , 1	4.5219
250 , 3	5.7105	250 , 2	4.7971	250 , 1	4.4095
300 , 3	5.4610	300 , 2	4.6590	300 , 1	4.2190
350 , 3	5.2181	350 , 2	4.4467	350 , 1	4.3238
400 , 3	5.3676	400 , 2	4.3667	400 , 1	4.6210
450 , 3	5.0562	450 , 2	4.5095	450 , 1	4.2600
500 , 3	4.7848	500 , 2	4.7914	500 , 1	4.2076

Table 3.3 : Mean Absolute Error (in Years) of Age Estimation obtained with itr$_{max}$ = 2 .

We can observe in this table that:

- Mean Absolute Error changes with the variation of final lower dimension values in mode-1 and mode-2.
- The best result which equals 3.0695 is obtained with final lower dimension values I$'_1$ = 60 and I$'_2$ = 3.

3.3.3.3 Calculate the number of trees

The following table presents the MAE results of the different number of trees at different levels of reduced subspace.

itr$_{max}$ = 2			
Number of trees	**I$_1$=60 , I$_2$=3**	**I$_1$=40 , I$_2$=2**	**I$_1$=20 , I$_2$=1**
20	5.4219	4.9019	4.8114
40	4.3152	4.2714	4.4648
60	3.7838	3.7971	4.2086
80	3.6857	3.9467	4.2200
100	3.4924	3.7667	4.0324
120	3.4943	3.4695	4.3724
140	3.5010	3.5724	4.1010
160	3.1981	3.3133	4.1514
180	3.2010	3.4476	4.1762
200	3.2981	3.1762	4.1486
250	3.3371	3.3800	4.0152
300	3.1219	3.3429	4.1495
350	3.2219	3.3448	4.0981
400	3.1086	3.1667	4.0505
500	3.1714	3.3038	3.9790
600	**2.8752**	3.2838	4.1871
700	3.2076	**3.1562**	4.2286
800	3.0143	3.3029	**4.0114**
900	2.9762	3.2295	4.1162
1000	3.2467	3.2857	4.0371
1500	3.0790	3.2886	4.1495
2000	2.9476	3.1990	4.0562
3000	2.8990	3.2371	4.0495
4000	2.9679	3.3333	4.0352

Table 3.4 : Mean Age Error (years) obtained with different number of trees.

The results obtained from Table 3.4 show that :

- Mean Absolute Error changes with the variation of number of trees.
- The number of trees = 600 gives the best MAE result = **2.88** at the I$_1$=60, I$_2$=3.
- The number of trees = 700 gives the best MAE result = 3.16 at the I$_1$=40, I$_2$=2.
- The number of trees = 800 gives the best MAE result = 4.01 at the I$_1$=2, I$_2$=1.

3.3.3.4 Comparison with state of the art

The performances of some state-of-the-art approaches are shown in Table 3.5. As can be seen, by adopting the proposed scheme, we got a significant improvement in performance. The best state-of-the art MAE was 3.23 years, whereas the best MAE obtained by our adopted scheme was 2.88 years. These results show the efficiency of our age estimation system. This system uses several CNN type as image description and multidimensional analysis (TEDA)

Publication	Approach	MAE
Gunay and Nabiyev (2016) [45]	AAM+GABOR+LBP	5.40
Nguyen et al. (2014) [111]	MLBP+GABOR+SVR	6.50
Bekhouche et al. (2014) [112]	LBP+BSIF+SVR	6.20
Choi et al. (2010) [113]	GHPF+SVR	8.40
Luu et al. (2011) [114]	CAM+SVR	6.00
Dornaika et al.(2018) [109]	Deep features+transfer	3.79
Guehairia et al.(2020) [107]	DRF	3.23
Proposed scheme	TEDA	**2.88**

Table 3.5: Mean Age Error (years) obtained with different state-of-the art approaches on PAL database [102].

3.4 Conclusion

In this chapter, we presented a age estimator system based on the (TEDA+WCCN) method. Firstly, we explored new set of deep CNN features, where we used deep features provided by these CNNs. Then, a face representation based on tensors is designed to combine all the features extracted from faces. Furthermore, to improve the discrimination of the proposed tensor face representation, we used multilinear subspace method (TEDA). Additionally, WCCN feature normalization is applied to the face tensors to reduce the classification errors. Finally, we have experimented with the dimensionality reduction and shown experimentally the superiority of the proposed contributions. Evaluated on the PAL database, our algorithm achieved high accuracy in estimating ages compared to published methods.

In human communication, the face is the first source that provides information to identify a person, as a serves a knowledge base for a lot of useful information, and as an important biological information carrier. The human face reflects many properties such as identity, age, gender, expression, and emotion. Age estimation and age classification of humans using their face images are a part of a field of biometric. Biometric systems use behavioral and physiological characteristics to recognize individuals. Soft biometric traits like age, gender, ethnicity, height, weight, in combination with hard biometric traits can be used to enhance the performance of biometric systems.

Age estimation is a type of soft biometrics that provides ancillary information of an individual's identity information. It is defined as the age of a person based on their biometric features. Due to the importance of human face aging in biometrics systems, age estimation became an attractive area, as age can enhance the performance of face recognition. For this purpose, the age estimation framework has been proposed to improve the accuracy and efficacy of these systems. Human age classification and estimation can be done in many ways, but this work is concerned with the age estimation and classification based on multi-dimensional images of people's faces.

The age estimation problem can be seen as a classification problem. Our work, confronted the challenge of age estimation as a classification problem using a multi-dimensional approach, to analyze the natural images for the purpose of facial age estimation.

Considering the importance of expanding research on this topic. In our work, we have developed a fully automatic age estimation framework, and this framework was performed based on method (TEDA) by multidimensional analysis (tensor analysis). Our age estimation framework encompassed several components: face preprocessing, feature extraction, subspace transformation (Reduction and classification of multidimensional data), and age prediction.

1- In the first step, a pre-treatment phase is necessary to keep the maximums of intrinsic variations of the face and to remove other information such as background, hair, shirt collars, ears, etc.

2- Regarding the description of the face, we used a new representation based on high-order tensors. This representation combines different deep descriptors, extracted at different scales, offering better discrimination. The proposed tensor representation is considered to be a new means of merging deep descriptors.

3- We used TEDA method, for the reduction of dimensionality and classification of tensor data when the complete data class label is missing. (TEDA) projects the face tensor of

entry into a new subspace in which the margin between samples belonging to different classes is widened while the margin in samples belonging to the same classes is reduced. In addition, TEDA reduces the size of each tensor mode.

4- In the age prediction step, This part has explored the efficacy of (TEDA) coupled with a random forest classifier for classifying the age. The dimensions of the training set was initially reduced using TEDA and then categorized based on these reduced deep features using a Random Forest Classifier.

We have provided a thorough evaluation of the performance of this approach. We have demonstrated that the performance of the age estimation can be significantly improved by adopting subspace transformation, using (TEDA) method. Multidimensional analysis still outperforms their original counterparts (the traditional linear subspace methods) by a large margin. We have shown that the performance of deep representations on the PAL database is a good indicator of their performance and that using the random forest classifier can further improve on already very strong results achieved using the combination of deep representations. Regarding results, our experiments show that the best MAE score which equals 2.88 (years) is obtained with final lower dimension values $I_1 = 60$ and $I_2 = 3$, Maximal number of iterations $= 2$ and number of trees $= 600$. Additionally, the proposed model outperforms the previous works by 0.35 (years) on PAL database.

The prospects for the evolution of this work are:

- As future work, it is interesting to investigate higher tensor orders for face representation with the proposed tensor subspace variant.

- Apply our system to another database larger than the PAL database.

- Use hand-crafted features to extract characteristics such as LBP and BSIF, with deep features.

Bibliographie

[1]   Onifade, O. F., & Akinyemi, D. J. (2015). A review on the suitability of machine learning approaches to facial age estimation. International Journal of Modern Education and Computer Science, 7(12), 17.

[2]   Fu, Y., Guo, G., & Huang, T. S. (2010). Age synthesis and estimation via faces: A survey. IEEE transactions on pattern analysis and machine intelligence, 32(11), 1955-1976.

[3]   Grd, P. (2013). Introduction to human age estimation using face images. Research Papers Faculty of Materials Science and Technology Slovak University of Technology, 21(Special-Issue), 24-30.

[4]   Angulu, R., Tapamo, J. R., & Adewumi, A. O. (2018). Age estimation via face images: a survey. EURASIP Journal on Image and Video Processing, 2018(1), 42.

[5]   Osman, O. F., & Yap, M. H. (2018). Computational intelligence in automatic face age estimation: A survey. IEEE Transactions on Emerging Topics in Computational Intelligence, 3(3), 271-285.

[6]   Shakhnarovich, G., & Moghaddam, B. (2005). Face recognition in subspaces. In Handbook of face recognition (pp. 141-168). Springer, New York, NY.

[7]   Lu, H., Plataniotis, K. N., & Venetsanopoulos, A. N. (2011). A survey of multilinear subspace learning for tensor data. Pattern Recognition, 44(7), 1540-1551.

[8]   BESSAOUDI, M. (2019). Reconnaissance de Visage basée sur l'Analyse Multidimensionnelle (Doctoral dissertation, Université Mohamed Khider–Biskra).

[9]   Ouamane, A. (2018). Nouveau descripteur Gabor locale binaire DGLB pour la vérification de visage 3D. MASTER.Télécommunications. Université Mohamed Khider Biskra.

[10]   Yan, S., Xu, D., Yang, Q., Zhang, L., Tang, X., & Zhang, H. J. (2006). Multilinear discriminant analysis for face recognition. IEEE Transactions on Image Processing, 16(1), 212-220.

[11]   Mandal, T., Wu, Q. J., & Yuan, Y. (2009). Curvelet based face recognition via dimension reduction. Signal Processing, 89(12), 2345-2353.

[12]   Yuan, S., Mao, X., & Chen, L. (2017). Multilinear spatial discriminant analysis for dimensionality reduction. IEEE Transactions on Image Processing, 26(6), 2669-2681.

[13]   Law, M. H., & Jain, A. K. (2006). Incremental nonlinear dimensionality reduction by manifold learning. IEEE transactions on pattern analysis and machine intelligence, 28(3), 377-391.

Bibliographie

[14] Hu, Z., Wen, Y., Wang, J., Wang, M., Hong, R., & Yan, S. (2016). Facial age estimation with age difference. IEEE Transactions on Image Processing, 26(7), 3087-3097.

[15] El Dib, M. Y. M. H. (2011). Automatic facial age estimation (Doctoral dissertation, Tesis de máster no publicada. Universidad del Cairo, Cairo, Egipto).

[16] Estimation de l'âge d'une personne par analyse faciale. Projet Intégratif - MASTER SCIENCES DE L'INGENIEUR Said Boufrah, RihabChetoui, Sabrina Djedi , Duc Tuan Tran, Thi Thanh Thuy Tran, Eugénie Vasson Responsable :Kevin Bailly.

[17] Georgopoulos, M., Panagakis, Y., & Pantic, M. (2018). Modeling of facial aging and kinship: A survey. Image and Vision Computing, 80, 58-79.

[18] sites internet .

[19] Suo, J., Zhu, S. C., Shan, S., & Chen, X. (2009). A compositional and dynamic model for face aging. IEEE Transactions on Pattern Analysis and Machine Intelligence, 32(3), 385-401.

[20] Stone, A. (2010). The aging process of the face & techniques of rejuvenation. http://www. aaronstonemd. com/Facial Aging Rejuvenation. shtm.

[21] Jain, A. K., Flynn, P., & Ross, A. A. (Eds.). (2007). Handbook of biometrics. Springer Science & Business Media.

[22] Jain, A. K., Dass, S. C., & Nandakumar, K. (2004, July). Soft biometric traits fo personal recognition systems. In International conference on biometric authentication (pp.731-738). Springer, Berlin, Heidelberg.

[23] Dantcheva, A., Elia, P., & Ross, A. (2015). What else does your biometric data reveal? A survey on soft biometrics. IEEE Transactions on Information Forensics and Security, 11(3), 441-467.

[24] Nixon, M. S., Correia, P. L., Nasrollahi, K., Moeslund, T. B., Hadid, A., & Tistarelli, M. (2015). On soft biometrics. Pattern Recognition Letters, 68, 218-230.

[25] Rothe, R., Timofte, R., & Van Gool, L. (2018). Deep expectation of real and apparent age from a single image without facial landmarks. International Journal of Computer Vision, 126(2-4), 144-157.

[26] Geng, X., Fu, Y., & Smith-Miles, K. (2010). Automatic facial age estimation. In 11th Pacific rim international conference on artificial intelligence (pp. 1-130).

[27] Bledsoe, W. W., & Chan, H. (1965). A man-machine facial recognition system— some preliminary results. Panoramic Research, Inc, Palo Alto, California., Technical Report PRI A, 19, 1965.

[28]   Moschoglou, S., Papaioannou, A., Sagonas, C., Deng, J., Kotsia, I., & Zafeiriou, S. (2017). Agedb: the first manually collected, in-the-wild age database. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (pp. 51-59).

[29]   Hinton, G. E. (2007). Learning multiple layers of representation. Trends in cognitive sciences, 11(10), 428-434.

[30]   Messer, K., Matas, J., Kittler, J., Luettin, J., & Maitre, G. (1999, March).XM2VTSDB: The extended M2VTS database. In Second international conference on audio and video-based biometric person authentication (Vol. 964, pp. 965-966).

[31]   Gross, R., Matthews, I., Cohn, J., Kanade, T., & Baker, S. (2007). The CMU multi-pose, illumination, and expression (Multi-PIE) face database. CMU Robotics Institute. TR-07-08, Tech. Rep.

[32]   Gross, R., Matthews, I., Cohn, J., Kanade, T., & Baker, S. (2010). Multi-pie. Image and Vision Computing, 28(5), 807-813.

[33]   Martinez, A. M. (1998). The AR face database. CVC Technical Report24.

[34]   Angelova, A., Abu-Mostafam, Y., & Perona, P. (2005, June). Pruning training sets for learning of object categories. In 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) (Vol. 1, pp. 494-501). IEEE.

[35]   Phillips, P. J., Moon, H., Rizvi, S. A., & Rauss, P. J. (2000). The FERET evaluation methodology for face-recognition algorithms. IEEE Transactions on pattern analysis and machine intelligence, 22(10), 1090-1104.

[36]   Phillips, P. J., Der, S. Z., Rauss, P. J., & Der, O. Z. (1996). FERET (face recognition technology) recognition algorithm development and test results. Adelphi, MD: Army Research Laboratory.

[37]   Rizvi, S. A., Phillips, P. J., & Moon, H. (1998, April). The FERET verification testing protocol for face recognition algorithms. In Proceedings third IEEE international conference on automatic face and gesture recognition (pp. 48-53). IEEE.

[38]   Georghiades, A., Belhumeur, P., & Kriegman, D. (1997). Yale face database. Center for computational Vision and Control at Yale University, http://cvc. yale. edu/projects/yalefaces/yalefa, 2(6), 33.

[39]   LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. nature, 521(7553), 436-444.

[40]   Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT press.

Bibliographie

[41] Kwon, Y. H. (1994). da Vitoria Lobo,". In Age classification from facial images," 1994 Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA (pp. 762-767).

[42] Burt, D. M., & Perrett, D. I. (1995). Perception of age in adult Caucasian male faces: Computer graphic manipulation of shape and colour information. Proceedings of the Royal Society of London. Series B: Biological Sciences, 259 (1355), 137-143.

[43] Lanitis, A., & Taylor, C. J. (2000, March). Towards automatic face identification robust to ageing variation. In Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580) (pp. 391-396). IEEE.

[44] Huang, G. B., Mattar, M., Berg, T., & Learned-Miller, E. (2008, October). Labeled faces in the wild: A database for studying face recognition in unconstrained environments.

[45] Günay, A., & Nabiyev, V. V. (2016). Age estimation based on hybrid features of facial images. In Information Sciences and Systems 2015 (pp. 295-304). Springer, Cham.

[46] 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).

[47] Yang, H., Huang, D., Wang, Y., Wang, H., & Tang, Y. (2016). Face aging effect simulation using hidden factor analysis joint sparse representation. IEEE Transactions on Image Processing, 25(6), 2493-2507.

[48] Eidinger, E., Enbar, R., & Hassner, T. (2014). Age and gender estimation of unfiltered faces. IEEE Transactions on Information Forensics and Security, 9(12),2170-2179.

[49] Chen, B. C., Chen, C. S., & Hsu, W. H. (2015). Face recognition and retrieval using cross-age reference coding with cross-age celebrity dataset. IEEE Transaction son Multimedia, 17(6), 804-815.

[50] Dibeklioğlu, H., Salah, A. A., & Gevers, T. (2012, October). Are you really miling at me? spontaneous versus posed enjoyment smiles. In European Conference on Computer Vision (pp. 525-538). Springer, Berlin, Heidelberg.

[51] Somanath, G., Rohith, M. V., & Kambhamettu, C. (2011, November).VADANA: A dense dataset for facial image analysis. In 2011 IEEE international conference on computer vision workshops (ICCV Workshops) (pp. 2175-2182). IEEE.

[52] LHI Image Database. 2010.

Bibliographie

[53] S. J. Foundation. Human and object interaction processing (hoip) face database. 2014.

[54] Ebner, N. C., Riediger, M., & Lindenberger, U. (2010). FACES—A database of facial expressions in young, middle-aged, and older women and men: Development and validation. Behavior research methods, 42(1), 351-362.

[55] Ni, B., Song, Z., & Yan, S. (2009, October). Web image mining towards universal age estimator. In Proceedings of the 17th ACM international conference on Multimedia (pp. 85-94).

[56] 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.

[57] Guo, G., Fu, Y., Dyer, C. R., & Huang, T. S. (2008). Image-based human age estimation by manifold learning and locally adjusted robust regression. IEEE Transactionson Image Processing, 17(7), 1178-1188.

[58] Bastanfard, A., Nik, M. A., & Dehshibi, M. M. (2007, December). Iranian face database with age, pose and expression. In 2007 International Conference on Machine Vision (pp. 50-55). IEEE.

[59] Nixon, N., & Galassi, P. (2008). The Brown Sisters: Thirty-Three Years. Museum of Modern Art.

[60] Scherbaum, K., Sunkel, M., Seidel, H. P., &Blanz, V. (2007, September). Prediction of individual non-linear aging trajectories of faces. In Computer Graphics Forum (Vol. 26, No. 3, pp. 285-294). Oxford, UK: Blackwell Publishing Ltd.

[61] Ricanek, K., & Tesafaye, T. (2006, April). Morph: A longitudinal image database of normal adult age-progression. In 7th International Conference on Automatic Face and Gesture Recognition (FGR06) (pp. 341-345). IEEE.

[62] Ueki, K., Hayashida, T., & Kobayashi, T. (2006, April). Subspace-based age-group classification using facial images under various lighting conditions. In 7th International Conference on Automatic Face and Gesture Recognition (FGR06) (pp. 6-pp).IEEE.

[63] Fu, Y., & Zheng, N. (2006). M-face: An appearance-based photorealistic model for multiple facial attributes rendering. IEEE Transactions on Circuits and Systems for Video technology, 16(7), 830-842.

[64] Phillips, P. J., Flynn, P. J., Scruggs, T., Bowyer, K. W., Chang, J., Hoffman, K., ... & Worek, W. (2005, June). Overview of the face recognition grand challenge. In 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05) (Vol. 1, pp. 947-954). IEEE.

Bibliographie

[65] Minear, M., & Park, D. C. (2004). A lifespan database of adult facial stimuli. Behavior Research Methods, Instruments, & Computers, 36(4), 630-633.

[66] Lanitis, A. The FG-NET aging database, 2002. Available: www-prima. Inrialpes. fr/FGnet/html/benchmarks. html, 2.

[67] Lanitis, A., Taylor, C. J., & Cootes, T. F. (2002). Toward automatic simulation of aging effects on face images. IEEE Transactions on pattern Analysis and machine Intelligence, 24(4), 442-455.

[68] Sim, T., Baker, S., & Bsat, M. (2002, May). The CMU pose, illumination, and expression (PIE) database. In Proceedings of Fifth IEEE International Conference on Automatic Face Gesture Recognition (pp. 53-58). IEEE.

[69] Phillips, P. J., Wechsler, H., Huang, J., & Rauss, P. J. (1998). The FERET database and evaluation procedure for face-recognition algorithms. Image and vision computing, 16(5), 295-306.

[70] Burt, D. M., & Perrett, D. I. (1995). Perception of age in adult Caucasian male faces: Computer graphic manipulation of shape and colour information. Proceedings of the Royal Society of London. Series B: Biological Sciences, 259 (1355), 137-143.

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

[72] Berry, D. S., Zebrowitz-MeArthur, L., & Alley, T. R. (1988). Social and applied aspects of perceiving faces.

[73] Kass, M., Witkin, A., & Terzopoulos, D. (1988). Snakes: Active contour models. International journal of computer vision, 1(4), 321-331.

[74] Lanitis, A., Draganova, C., & Christodoulou, C. (2004). Comparing different classifiers for automatic age estimation. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 34(1), 621-628.

[75] Cootes, T. F., Edwards, G. J., & Taylor, C. J. (2001). Active appearance models. IEEE Transactions on pattern analysis and machine intelligence, 23(6), 681-685.

[76] Geng, X., Zhou, Z. H., Zhang, Y., Li, G., & Dai, H. (2006, October). Learning from facial aging patterns for automatic age estimation. In Proceedings of the 14th ACM international conference on Multimedia (pp. 307-316).

[77] 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.

[78] Ahonen, T., Hadid, A., & Pietikainen, M. (2006). Face description with local

binary patterns: Application to face recognition. IEEE transactions on pattern analysis and machine intelligence, 28(12), 2037-2041.

[79]   Hill, C. M., Solomon, C. J., & Gibson, S. J. (2005). Aging the human face-a statistically rigorous approach.

[80]   Suo, J., Wu, T., Zhu, S., Shan, S., Chen, X., & Gao, W. (2008, September).Design sparse features for age estimation using hierarchical face model. In 2008 8th IEEE International Conference on Automatic Face & Gesture Recognition (pp. 1-6). IEEE.

[81]   Suo, J., Min, F., Zhu, S., Shan, S., & Chen, X. (2007, June). A multi-resolution dynamic model for face aging simulation. In 2007 IEEE Conference on Computer Vision and Pattern Recognition (pp. 1-8). IEEE.

[82]   Guo, G., Mu, G., Fu, Y., & Huang, T. S. (2009, June). Human age estimation using bio-inspired features. In 2009 IEEE Conference on Computer Vision and Pattern Recognition (pp. 112-119). IEEE.

[83]   Serre, T., Wolf, L., Bileschi, S., Riesenhuber, M., & Poggio, T. (2007). Robust object recognition with cortex-like mechanisms. IEEE transactions on pattern analysis and machine intelligence, 29(3), 411-426.

[84]   Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. Nature neuroscience, 2(11), 1019-1025.

[85]   Zhou, S. K., Georgescu, B., Zhou, X. S., & Comaniciu, D. (2005, October).Image based regression using boosting method. In Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1 (Vol. 1, pp. 541-548). IEEE.

[86]   Guo, G., Fu, Y., Huang, T. S., & Dyer, C. R. (2008, January). Locally adjusted robust regression for human age estimation. In 2008 IEEE Workshop on Applications of Computer Vision (pp. 1-6). IEEE.

[87]   Yan, S., Wang, H., Tang, X., & Huang, T. S. (2007, October). Learning auto-Structured regressor from uncertain nonnegative labels. In 2007 IEEE 11th international conference on computer vision (pp. 1-8). IEEE.

[88]   Yan, S., Wang, H., Huang, T. S., Yang, Q., & Tang, X. (2007, July). Ranking with uncertain labels. In 2007 IEEE International Conference on Multimedia and Expo (pp. 96-99). IEEE.

[89]   De Lathauwer, L., De Moor, B., & Vandewalle, J. (2000). On the best rank-1 and rank-(r 1, r 2,..., rn) approximation of higher-order tensors. SIAM journal on Matrix Analysis and Applications, 21(4), 1324-1342.

Bibliographie

[90]   Plataniotis, K. N., & Venetsanopoulos, A. N. (2013). Color image processing and applications. Springer Science & Business Media.

[91]   Lu, H., Plataniotis, K. N., & Venetsanopoulos, A. N. (2008). MPCA: Multilinear principal component analysis of tensor objects. IEEE transactions on Neural Networks, 19(1), 18-39.

[92]   Ouamane, A., Chouchane, A., Boutellaa, E., Belahcene, M., Bourennane, S., & Hadid, A. (2017). Efficient tensor-based 2d+ 3d face verification. IEEE Transactions on Information Forensics and Security, 12(11), 2751-2762.

[93]   Kolda, T. G., & Bader, B. W. (2009). Tensor decompositions and applications. SIAM review, 51(3), 455-500.

[94]   Lu, H. (2008). Multilinear subspace learning for face and gait recognition. University of Toront.

[95]   Wang, S. J., Yan, S., Yang, J., Zhou, C. G., & Fu, X. (2014). A general exponential framework for dimensionality reduction. IEEE Transactions on Image Processing, 23(2), 920-930.

[96]   Zhang, T., Fang, B., Tang, Y. Y., Shang, Z., & Xu, B. (2009). Generalized discriminant analysis: A matrix exponential approach. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 40(1), 186-197.

[97]   Dehak, N., Kenny, P. J., Dehak, R., Dumouchel, P., & Ouellet, P. (2010). Front-end factor analysis for speaker verification. IEEE Transactions on Audio, Speech, and Language Processing, 19(4), 788-798.

[98]   Barkan, O., Weill, J., Wolf, L., & Aronowitz, H. (2013). Fast high dimensional vector multiplication face recognition. In Proceedings of the IEEE international conference on computer vision (pp. 1960-1967).

[99]   Kazemi, V., & Sullivan, J. (2014). One millisecond face alignment with an ensemble of regression trees. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1867-1874).

[100]   Oussama, G., Ouamane, A., Dornaika, F., Taleb-Ahmed, A. (2020). Deep Random Forest for Facial Age Estimation Based on Face Images. the First International Conference on Communications, Control Systems and Signal Prossessing.

[101]   El Khiyari, H., & Wechsler, H. (2016). Face recognition across time lapse using convolutional neural networks. Journal of Information Security, 7(3), 141-151.

[102]   Belver, C., Arganda-Carreras, I., & Dornaika, F. (2016). Comparative study of human age estimation based on hand-crafted and deep face features. In Video Analytics. Face

and Facial Expression Recognition and Audience Measurement (pp. 98-112). Springer, Cham.

[103] Dornaika, F., Arganda-Carreras, I., & Belver, C. (2019). Age estimation in facial images through transfer learning. Machine Vision and Applications, 30(1), 177-187.

[104] Parkhi, O. M., Vedaldi, A., & Zisserman, A. (2015). Deep face recognition.

[105] Chatfield, K., Simonyan, K., Vedaldi, A., & Zisserman, A. (2014). Return of the devil in the details: Delving deep into convolutional nets. arXiv preprint arXiv:1405.3531.

[106] Rothe, R., Timofte, R., & Van Gool, L. (2015). Dex: Deep expectation of apparent age from a single image. In Proceedings of the IEEE international conference on computer vision workshops (pp. 10-15).

[107] Lu, Z., Jiang, X., & Kot, A. (2017, June). Enhance deep learning performance in face recognition. In 2017 2nd International Conference on Image, Vision and Computing (ICIVC) (pp. 244-248). IEEE.

[108] Wan, L., Liu, N., Huo, H., & Fang, T. (2017, June). Face recognition with convolutional neural networks and subspace learning. In 2017 2nd international conference on image, vision and computing (ICIVC) (pp. 228-233). IEEE.

[109] 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).

[110] Minear, M., & Park, D. C. (2004). A lifespan database of adult facial stimuli. Behavior Research Methods, Instruments, & Computers, 36(4), 630-633.

[111] Nguyen, D. T., Cho, S. R., Shin, K. Y., Bang, J. W., & Park, K. R. (2014). Comparative study of human age estimation with or without preclassification of gender and facial expression. The Scientific World Journal, 2014.

[112] Bekhouche, S. E., Ouafi, A., Taleb-Ahmed, A., Hadid, A., & Benlamoudi, A. (2016). Facial age estimation using bsif and lbp. arXiv preprint arXiv:1601.01876.

[113] Choi, S. E., Lee, Y. J., Lee, S. J., Park, K. R., & Kim, J. (2010, December). A comparative study of local feature extraction for age estimation. In 2010 11th International Conference on Control Automation Robotics & Vision (pp. 1280-1284). IEEE.

[114] Luu, K., Seshadri, K., Savvides, M., Bui, T. D., & Suen, C. Y. (2011, October). Contourlet appearance model for facial age estimation. In 2011 international joint conference on biometrics (IJCB) (pp. 1-8). IEEE.