2016

University of North Carolina Wilmington

Master of Science in

Computer Science and Information Systems

Proceedings

https://csbapp.uncw.edu/mscsis

CHILD FACE RECOGNITION
ESTABLISHING BASELINE PERFORMANCE METRICS

Shivani Bhardwaj

A Capstone Project Submitted to the
University of North Carolina Wilmington in Partial Fulfillment
of the Requirements for the Degree of
Master of Science

Department of Computer Science
Department of Information Systems and Operations Management

University of North Carolina Wilmington

2016

Approved by

Advisory Committee

Dr. Jeffrey W. Cummings _____

Dr. Toni B. Pence _____

_____
Dr. Karl Jr. Ricanek, Chair

Accepted By

_____
Dean, Graduate School

TABLE OF CONTENTS

ABSTRACT

Automatic face recognition is a challenging task that has made significant advancements over the last decade against the problems of pose, illumination, and expression (PIE). There has also been an improvement in the systems against the challenges of time displacement, also known as aging. Aging results in face variation, which affects the performance of a face recognition system. However, this work uncovers a problem of aging that has not, as of yet, received attention from the research community. This body of work explores the challenges of face recognition, and by inference soft biometrics or facial analytics, for sub-adults, the population of faces that exists between ages 0 and 18 years. The velocity of craniofacial morphology in the sub-adult population can be quite aggressive as compared to the changes in the adult population, and such rapid changes in both the hard (bony) tissue and the soft tissue can cause demonstrable degradation in face recognition as well as facial analytics. The objective of this work is to highlight the challenges with the issues of longitudinal face recognition and the soft biometrics on sub-adults. Further, this work establishes the difficulty of this problem by establishing a seminal baseline as well as the critical biological underpinnings for the cause of the problem of child face recognition. This work institutes the baseline against the largest longitudinal sub-adult corpus created to date on a set of commercial matchers. The impact of variation caused because of aging on face recognition technology was explored with commercial face recognition systems, Cognitec's FaceVacs v8.50, Rank One Computing v1.20, and Neurotechnology's Verilook v6.0. The performance of these commercial systems reveals the challenges associated with this problem: average performance across the systems tested for identification of a person from its younger image: 25.06% (Rank-1) and 47.03% (Rank-20) with a dataset of 501 subjects.

DEDICATION

This work is dedicated to my husband, Rajiv, who motivated me every single day and supported me unconditionally. To my parents, Mr. Umakant Bhardwaj and Mrs. Madhu Bhardwaj who have always instilled the value of education while letting me find my own path. To my grandparents and in-laws, for their understanding and support.

ACKNOWLEDGMENT

My sincere gratitude to my advisor Dr. Karl Ricanek Jr. for his continuous support, patience and motivation. His guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor. Other members of my capstone committee: Dr. Jeff Cummings and Dr. Toni B. Pence for their encouragement and insightful comments that incented me to widen my research from various perspectives.

Very special thanks to Dr. A. Midori Albert and Amrutha Prasad for always focusing my direction, and for creating an enjoyable academic environment. Thanks to my scholastic colleagues, Harry, Chad and Charissa, for providing knowledge and stability.

My friend Pooja Mishra for pushing me when I wanted to give up. Finally, big appreciation to my family, friends and Face Aging Lab members for their constant support and encouragement, they provided throughout this work.

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

CHAPTER 1. INTRODUCTION

Humans are prewired for face recognition from birth, thanks to centuries of evolution. In fact, faces are considered to be among the most significant visual stimuli [1] and the ability to recognize a face is perhaps the most developed visual perceptual skill among humans [2]. Humans are able to recognize faces accurately, independent of their age, poses, expressions, and illumination conditions. The face is not only important in identity recognition, but it also provides a diverse pool of information. The wealth of information about an individual that can be extracted from a single glance at their face, includes their identity, ethnicity (race and culture), sex, age, health state, emotional state (happiness, sadness, anger, frustration, attraction, etc.), and the direction of their attention, to name just a few [1]. Using the human face as a key to identity, biometric technology has received significant attention in the past decade due to its potential markets especially in security domain. The security market is seeking a "lights out" automatic face recognition system capable of running fully autonomously with very few errors[1]. Due to the increasing popularity of social networks and the availability of image recording devices, law enforcement agencies have access to a variety of crime scene images, including CCTV recordings and other media outlets. Law enforcement agencies adopt facial recognition technology when analyzing these crime scene images against mug shot databases [3], which is significantly higher than the amount of latent fingerprints or DNA samples. Likewise, due to the rise in fraud in the Banking, Financial services, and Insurance (BFSI) sector, facial recognition systems are being increasingly used in these industries to ensure

[1] Some law enforcement agencies estimate that up to a quarter of complaint cases contain facial images of the suspect or an accomplice [3], which is significantly higher than the number of latent fingerprints or DNA samples.

security and reduce crimes. Furthermore, facial analytics systems are employed in the retail industry to determine the age, gender, and facial expressions of the customers to target them for marketing. However, these market domains have been explored only for adult users. There is promising potential for the growth of this technology for sub-adults as well which has not been fully explored. The objective of this paper is to evaluate the performance of automated face recognition system for sub-adults, i.e. people under 18 years of age.

1.1 Applications Scenarios

Automatic facial recognition is commonly used in monitoring/surveillance systems. Government agencies employ the use of automatic facial recognition in mass surveillance systems to monitor terrorism, secure border areas, and catch criminals like drug smugglers. The quality that makes face recognition technology an effective monitoring system is its covert and passive nature. Covert refers to the instance in which biometric samples are being collected at a location that is not known to by bystanders. An example of a covert environment might involve an airport checkpoint where face images of passengers are captured and compared without their knowledge [4] for possible terrorist or criminals. Another quality that makes facial recognition technology more acceptable is that it can use existing hardware infrastructure. Any place that is equipped with any image capture device like cameras could end up employing face recognition technology. Therefore, any time you are in public whether you are attending a protest rally or entering a shopping mall or visiting churches or a bar, face recognition systems (FRS) are capable of identifying and tracking your movements. FRS can help in identifying social disturbances, child endangerment and crime, which may ultimately lead to the prevention

of such instances. Casinos are using it to track their high rollers and keep out card cheats [5] while retailers are using it to catch shoplifters and get to know their shoppers [6].

Researchers are racing to develop reliable methods of handling vast amounts of data, making facial recognition systems more robust to perform operations like sorting and searching image data, according to the picture content. It is a challenging task to identify a person from the immense quantity of image data; however, this is the biggest business sector for the technology, some of the example includes photo tagging and identity fraud detection. It can also be deployed to fight child pornography; a system can be deployed with capability to scan the vast amount of internet data and flags all the alleged pornographic video or image with a sub-adult in it [7].

The application of facial recognition technology is not only limited to law enforcement and security agencies, but it has also been adopted by commercial markets. For example, social media uses FRS for identifying friends in images that pass through out social media, e.g. Facebook, Instagram, Pinterest, Twitter or Google Plus+. Who is that person in your social media photo? Face recognition can provide the answer. This feature is called auto-tagging and is an accepted feature for all social media and photo uploading sites like iCloud. An FRS has matched anyone who has ever been tagged in a photo. With the advancement in consumer the technology, most of the handheld devices are now capable of obtaining an image such as smart phones and digital camera, and are also equipped with FR technology. These digital devices now detect a face in the frame, and when it finds one, its camera can automatically adjust both the focus and the exposure to provide the best portrait possible. Additionally, searching and organizing digital data by its content is another feature that can leverage face recognition technology. Search engines like Google already provides a way to search person of interest by photographs.

Visual content is taking over the social media and the digital marketing strategy. According to a marketing theory, images are better for promotion as they grab the reader's attention quickly and the brain capability to process an image is a lot faster than text [8]. Facial analytics, which uses many of the same principles of face recognition, can help in gauging demographic data analysis, and soft biometric are useful for targeted digital marketing. It allows retailers to customize their service and products as well as their advertising. Besides, it is beneficial to the customer because they do not have to waste time essentially relaying that information [9]. Systems equipped with advanced facial recognition technology allow humans to interact with computers in novel ways. This capability can provide an innovative way for human-computer interaction with a personalized experience, as soon as you walk in front of a machine, it can recognize you and interact with you accordingly.

1.2 Motivation and Contribution

The greatest motivation for improving on child facial recognition is its potential applications in an extensive range of access control and monitoring systems, primarily to safeguard children. Access control systems are the systems designed to regulate who or what can view or use resources (physical or logical) in a computing environment. Physical access control limits access to campuses, buildings, rooms and physical assets. Logical access limits connections to computer networks, system files, and data [10]. Monitoring systems let parents remotely check-up on their kids anytime and provide a method of assurance that their children are safe. A tracking system with facial recognition technology could also be configured to send you alerts if certain conditions are met.

Automatic facial recognition technology can provide an entirely new smart technology to protect and support latched-key kids[2]. A child can easily lose a key and hiding a key under a mat is not so secure. Also, there are no means to check if a child reaches home safely after school when parents are not home. Face recognition for smart environments includes smart homes, which can identify and provide access child. In addition, system can be configured, to not only trigger an alarm if the child does not arrive home by a certain time, but also if any unknown person is in your house. The system can also capture the photographs of all the people visited your house, and those pictures can be sent to your mobile phone when your child is alone. The same technology can also be used in a school or daycare for taking attendance or to locate the child in a building [11], [12], [13], [14], [15], [16]. It can aid in monitoring age-restricted areas and provides access control for various internet-of-things across different age groups. It can also assist to protect the sub-adult from predators and inappropriate web content. A system can be developed to flag a child predator if the person is near a school or children's play area. Volvo uses face recognition to help tired drivers; the same technique can be used to monitor your teen or first-time drivers. If, your teenage child jumps behind the wheel, the inside camera sends a photo to your phone for authentication, where you can approve the car's ignition sequence and even set a driving time limit, max audio sound and speed limit. Besides, if you want, you can monitor your new driver.

Face recognition can help/aid in the identification of missing and exploited children. A "Missing Persons Website/App" could be developed that allows the public to

[2]A latchkey kid or latchkey child is a child who returns from school to an empty home because their parent or parents are away at work or a child who is often left at home with little parental supervision. The house key is often strung around the child's neck or left hidden under a mat (or some other object) at the rear door to the property [11]

upload a picture of a person for matching against known missing persons. If a substantial match is recorded, law enforcement could be alerted and put in contact with the individual who uploaded the image. Another example of how a real child-based face recognition system could be used is for Amber Alerts. During Amber Alerts, a virtual watch list for the missing child could be developed that would allow anyone to upload or text, a picture of a potential match from a mobile device. For consumer applications, a better child face recognition system would make auto-tagging much more accurate. Current systems are not perfect at auto-tagging for the sub-adult population of people, according to the work of Michael Sodomsky [17].

There is an enormous market for the applications of facial recognition for sub-adults. In addition, it can aid in reducing crime against children and making cities a lot safer and secure for children. However, face recognition has been built for the adult use-case, this work identifies this flaw and will help to eradicate by providing researchers with the data required to develop novel solutions for child FR as well as overall FR.

The changes because of growth and development of a child make face recognition a more challenging problem. For all the systems, older individuals are both easier to recognize as themselves, and easy to tell apart from each other. The opposite is true in children and infants. As the variation in face because of aging in adult is different from sub-adults. Temporal changes after maturation is dominated by morphological and soft tissue changes, i.e. skin texture, wrinkles; while during growth and development phase, human craniofacial complex changes in proportion and the typically wide appearing child face, gains height and shape later in age. Over the time, the variation in the face belongs to the same individual increase, that makes it difficult to recognize the person and affects facial recognition system performance.

Current research focuses on the temporal changes (aging) in the adult face, and even commercial biometric systems lack basic performance goals with sub-adult faces. Facebook's Deepface algorithm claims accuracy of 97.35% on Label Faces in Wild dataset [18]. Klare et al., showed a true accept rate of 72.4% at a fixed false accept rate of 1.0% in adults with 10 years of lapse [19]. Michael Sodomsky conducted a preliminary evaluation of FRS for sub-adult faces. His study includes In-The-Wild Child Celebrity (ITWCC) image dataset, which focuses on the growth and development of human face and six different face recognition systems. The best-performing commercial system, Cognetic's FaceVacs v8.50 achieved true accept rate of only 37% at 1% false accept rate [17]. Moreover, the best non-commercial algorithm, OpenBR's SF4 algorithm attained true accept rate of 25% at 1% false accept rate [17]. As a matter of note, the developers of this algorithm are the founders of Rank One Computing, and both Cognitec's FaceVacs and Rank One Computing's FRS are evaluated in this work. The results of Sodomsky's evaluation showed that this area needs attention by the research community [17]. The focus of my current research addresses the performance gap of the automatic face recognition system between adults and sub-adults on commercial matchers only. The performance is compared to suitable performance against a congruent data corpus of adult faces, in this work the congruent data corpus is the Labeled Faces in the Wild.

This study will provide an independent and open assessment of modern commercial face recognition systems for sub-adults, in different operating conditions, like access control and photo tagging. The commercial systems included in this study include Cognitec's FaceVACS v8.50 [20], Rank One Computing's ROC v1.20 [21] and Neurotechnology's Verilook v6.0 trial [22]. I will compare and contrast the recognition accuracy of commercial systems with adult and sub-adult subjects. As no competent dataset

is available to explore this domain, I have extended the original dataset created by Sodomsky [17] by 30,793 number of images. This data set is called In-the-Wild Child Celebrity-2 (ITWCC-2) dataset.

This work is organized in different chapters as follows: Chapter 2 provides the essentials of an automatic facial recognition system and explore the previous work done in the field of face recognition and aging. Chapter 3 explains the methodology to evaluate the performance of facial recognition systems. Chapter 4 contains the experimental setup. Chapter 5 details the results of experiments performed. The results obtained from the analysis are concluded and summarized in chapter 6. It also describes the direction for the future work. A glossary is also included as Appendix A, which lists the terms or words relevant to facial recognition and biometric systems.

CHAPTER 2. BACKGROUND

The biometric system empowers recognition of individuals based on their behavioral or biological characteristics [23]. Standard biological biometric characteristics include fingerprints, retina, iris, facial images, hand geometry. Whereas, standard behavioral biometric characteristics include signature, gait, voice recordings, and keystroke rhythms. In biometric literature, these characteristics are referred to as traits, modalities, indicators or identifiers [24]. Biometric systems are widely accepted as an identity management and access control systems and primarily used to validate the identity of an individual seeking access to a restricted area or to establish the identity of a person.

Automatic face recognition is one of the many modalities of a biometric system and uses the digital image of the visible physical structure of an individual's face for recognition purposes [23]. Among all the biometric traits, the characteristics that make face recognition a unique biometric modality is as follows [24]:

- Universality: Face is the trait possessed by every individual.
- Semi–permanence: permanence is necessary for longitudinal work, and it is important that the biometric feature is sufficiently consistent over a period, or how the biometric system can incorporate the changes in the trait. With the advancement in facial recognition technology, it is possible to recognize an adult over more than ten years of time lapse [26].
- Collectability: The required feature (face) can be collected via a non-intrusive method that offer an added advantage over other biometric systems. Instead of requiring a person to provide their fingerprint or asking them to submit to iris scanning or retinal scanning, facial characteristics can be videoed or captured easily without requiring any contact at all and at a

distance. For many purposes, such as crime deterrent or security purposes, the ability to collect facial biometric information without the subject knowing is paramount.

- Acceptability: Society as well as the courts allow for the ubiquitous collection of face images, the example includes CCTV, cell phones & selfies.
- Performance/Accuracy: Face recognition is accurate as humans. In FRVT 2006, the experiment, comparing human and algorithm performance, the best-performing face recognition algorithms like Tsinghua (Ts2-norm) were more accurate than human [27] and again it can be used in wide variety of applications.

The identity of a person is not the only usable information that can be extracted from the facial image. Facial features can be further analyzed to determine other traits that can be useful to describe an individual like gender, race, and age. The size and geometry of the chin, lips, nose, eyebrows, skin color and other face components can be used to distinguish gender, race, and ethnicity, while other features such as creases, lines, sagging and wrinkles can reveal clues about age [24], refer to figure 1. These traits are known as soft biometric traits. Jain [29] defined soft biometrics as characteristics that provide some information about an individual but lack the distinctiveness and permanence to differentiate any two persons sufficiently.

Figure 1: Examples of Soft Biometric Traits. Notice the difference in facial features, and skin color and texture among different ethnicity, gender and race [88].

2.1 Face Recognition System

Like any biometric system, facial recognition is essentially a pattern recognition technique that, (1) acquires a digital image of the face from an individual, (2) extracts the salient feature set from the acquired image, (3) compares the extracted feature set against stored feature sets in a database and (4) provides the match result. When an individual uses a facial recognition system for the first time, their facial images (identifying features) are enrolled in the system as a reference for future comparison. This reference image(s) is known as an enrolled image(s) or gallery image(s) and may be stored in a central database as per the application requirement. When identity recognition is required, the individual's facial image is captured again. This time, however, the captured image is compared with the stored reference image to determine if there is a close match. Figure 2 is a block diagram of facial recognition system according to [24]. In brief, the basic operation of facial biometric system is explained below:

- Sensor module: It is designed to acquire the required biometric trait from the user. For face recognition system, the sensor can be any device with the capability of capturing images, like a camera.

- Feature Extraction module: For efficiency reasons, rather than using a recorded facial image directly, it usually to extract identifying features from the sample image and encodes these features in a matrix form that facilitates storage and comparison. It is crucial to precisely locate the features in an image to efficiently extract the features. Feature detection is often considered as the most difficult step in the process, whereas feature encoding, turning the area of interest of the face into features used for matching, is a relatively simple process defined by a set of algorithms used as a transformation.
- Match module: The matching module compares the matrix (probe) with the set (or subset) of the enrolled images in the system's database and provides a similarity score. The higher the score value the more similar the images.
- Decision module: The next step is to decide if the comparison is genuine or an impostor using a decision threshold (DT). If the score is higher than the threshold, the algorithm determines it is a genuine comparison (match); on the contrary, it is classified as an impostor comparison (non-match). Frequently, the decision threshold is a parameter that can be tuned by algorithm's users.

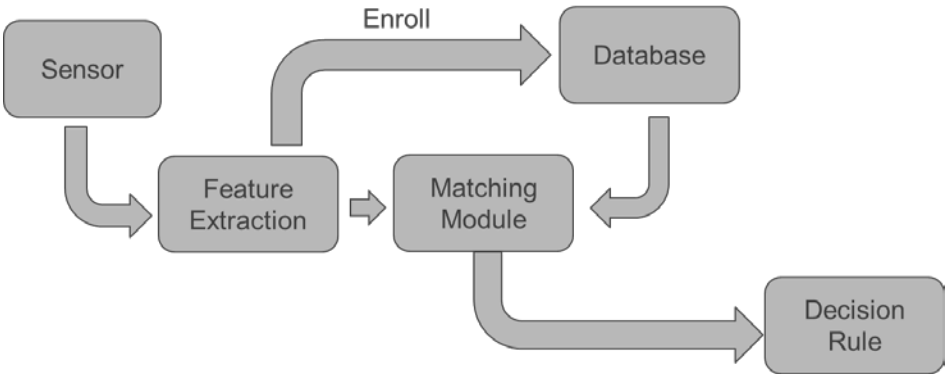


Figure 2: The basic components of a facial biometric system.

2.2 Face Recognition Challenges

While facial recognition systems impart several advantages over other biometric systems, it is imperative that the vulnerabilities of these systems are considered. The robustness of facial recognition systems are restricted by large intra-class variations caused by the environmental conditions including illumination, camera variation, image quality, pose variations, facial expressions, occlusion and also the temporal variation caused as a result of aging [25].

Over the last two decades, face recognition systems have advanced considerably. Performance measures such as recognition accuracy have improved substantially. As poor image quality and variation in the acquisition parameters were the main reason for matching errors in facial biometric systems; ISO/IEC 19794-5 image standard was proposed to improve face recognition accuracy and to support a variety of devices for facial recognition, including those with limited resources. Additional information on the ISO/IEC 19794-5 is located in the standard [31]. Also, all modern biometric passport photos comply with this standard. An illustration of the rule that governs the use of face images for official documents, e.g. national id's, passports, etc. is shown in figure 3.

Figure 3: ISO/IEC 19794-5 image sample [89]

There are many applications of face recognition technology, like visa and passport where one can specify and control the parameters of image acquisition such as pose variations, illumination variations, occlusions, expression variations, out-of-focus blur, and image resolution. However, there are also many applications like a surveillance system, where it is not possible to control such parameters. A lot of work has been done to support facial recognition in an unconstrained environment. Some of this work includes enhancement in sensors, correction for varying illumination across images, betterment in normalization techniques and pose corrections, advancement in algorithms from partially automatic to fully automatic [25], [32], [33], [34]. Images with the variation seen in everyday life are an example of unconstrained images also referred as "in-the-wild" images. Figure 4 shows some sample images acquired in unconstrained environmental condition.

Figure 4: Example of unconstrained images.

Facial recognition is highly researched not only because of its potential applications but also because of its challenging and complex nature. Researchers have reported a variety of challenges in the facial recognition domain; also, many solutions have been proposed, and advancement have been made. However, this improvement is best with images captured in constrained environments [25], [26], [27], [28], [29], [30]. Some of the reported recognition work also includes plastic surgery [35], similarity of twins [36], sketch-to-photo matching [37], [38], aging [39] , [40], [19], [41] and disguise [42].

The majority of the research and commercialization of the automatic face recognition has been dedicated on adult faces, whether it is to compensate for the intra-class variation or temporal changes caused because of aging. It is important to understand that the change caused by aging in growth and development phase of human life is different from deviation because of adult aging. Child face recognition is a challenging domain and has not been addressed adequately.

2.3 Advances in Face Recognition Performance

According to National Institute of Standards and Technology (NIST) 2014 report, since the initial commercialization of face recognition technology, error rates have declined massively in the last two decades [43]. Figure 5 represents a reduction in error rate for face recognition from 1993 to 2010 [43]. NIST has tracked the improvement in facial

recognition system from 1993 to present through a series of Face Recognition Challenges (http://frvt.org) and has fostered improvements in the technology [25]. The NIST report includes performance figures for prototype algorithms from the research laboratories of many of the major commercial suppliers of face recognition technologies including NEC, Morpho, Toshiba, Cognitec, 3M and Neurotechnology.

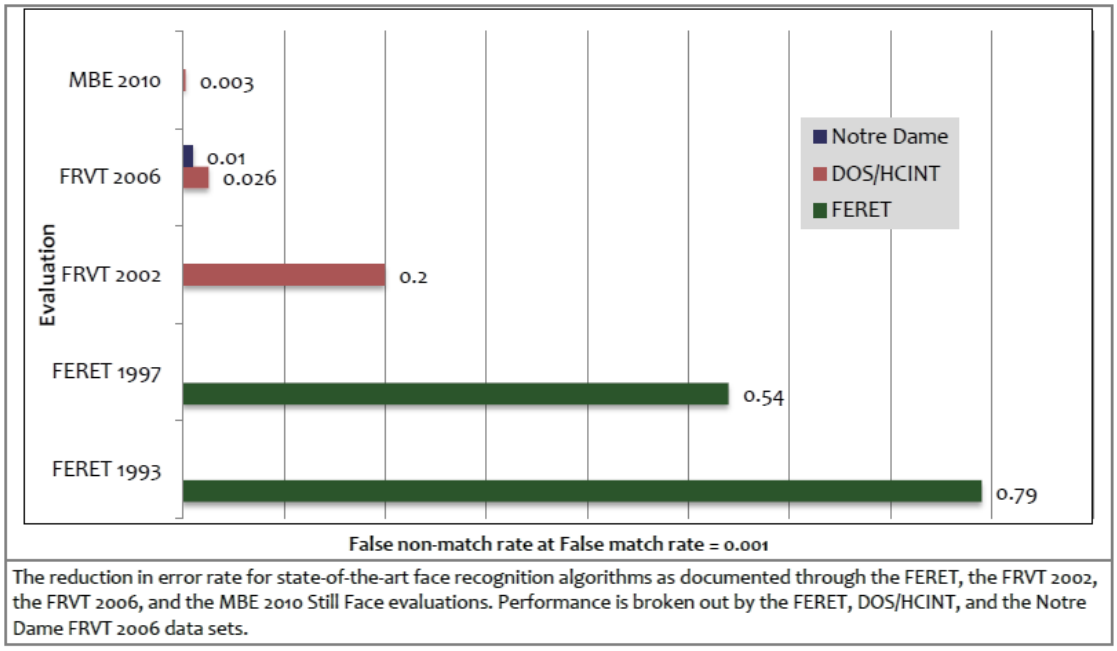


Figure 5: NIST 2014 - Reduction in error rate for face recognition from 1993 to 2010 [29]

According to the NIST 2014 report, the accuracy of facial recognition systems has improved in all the evaluated scenarios from 2010 to 2013. Some of the evaluation scenarios include one-to-one verification comparisons to determine if two samples originate from the same person or not, one-to-many identification searches to ascertain the identity of the individual and it also include cases to determine the sex or age of a person in one or more input images. Four research groups, Cognitec, Neurotechnology, Morpho, and NEC, are enrolled in both the 2013 and the previous 2010 test, allowing NIST researchers to compare performance improvements over time. From 2010 to 2013, rank

one miss rates have reduced by about 10% for Cognitec, Neurotechnology, and Morpho, and by about 28% for NEC (from 8.9% to 6.4%). Details of the latest report from NIST can be found in [43]. Whereas a preliminary evaluation of the facial recognition system with sub-adults dataset reveals that, the rank one miss rate is 47% [44].

2.3 Effect of Aging

Accuracy of automatic face recognition system is strongly dependent on subject age. Many attempts have been made to understand the effects aging has on facial recognition systems [45], [11], [12], [15], [16]. Brendan Klare and Anil K. Jain evaluated multiple facial recognition algorithms on a time lapsed adult database [19]. Their work showed that the best performing algorithm had a 96.3%, 94.3%, 88.6%, and 80.5% true accept rate at 1.0% false accept rate in the following time-lapse periods, respectively: 0 to 1, 1 to 5, 5 to 10, 10+ [19]. These results suggest that performance does decrease as the time between probe and gallery increases.

For all techniques/systems, older individuals are both easier to recognize as themselves, and easy to tell apart from each other. The opposite is true for children: both false negative and false positive rates are much higher, with infants being very hard to identify. Moreover, the trends are progressive throughout adulthood, with young adults [11] being identified with worse accuracy than older [43], [17]. In addition, it is difficult to determine the soft biometric trait like gender, as there are no significant sex differences in shape among non-adults [46]. According to the NIST report, the majority of the algorithms estimate age more accurately, for the adult age group, i.e., 18 years - 55 years [43]. The majority of the work done has been focused on adults and deals with the dynamics of matured faces and performance as a subject of growth and development phases of childhood, has just begun to be fully explored by researchers.

The most challenging issue is the vast amount of data that is required to fully understand and investigate the human face and its maturation process. The human face keeps changing throughout the life of an individual. This dynamic nature of the face makes it difficult to recognize the person over time. Because humans are very accurate at recognizing and estimating the age of a person from their face [9]; human visual perception skills can provide a benchmark and is often studied to understand the fundamentals about the age group of a person.

The alteration in the human face during the growth and development age is not the same as the changes in the face after maturation. Anthropologists and forensic studies have contributed significantly to this field and can be referred to understand these differences. Human aging can be explored as a two-staged process: stage one involves the growth and development phase and the second stage deals with the effect of maturity as age progresses [47], [48]. Because of the rapid structural changes in childhood, temporal displacement has a more profound impact on automatic facial recognition systems than that of adult aging.

2.4 Aging Process of Human Face

The human face is a complex structure that varies with age. It is a known fact that the shape of the bone changes with time because of remodeling throughout the entire human life [49]. The variation in adults is dominated by morphological and soft tissue changes, whereas, in sub-adults, the variation is because of craniofacial growth and development. Craniofacial growth and development involve the construction of facial structure, which results in the change of shape and the increase in size, in all three directions: vertical height, transversal width and anteroposterior depth [49], [50]. Figure 6 shows the facial changes in growth and development age.

Figure 6: Variations in Growth and Development Age

In broad terms, growth can be defined as a change in magnitude or increase in physical size, whereas development leads to maturation and is concerned with the details and cause of growth. Bastir et al. defines growth as the process, which results in the change in size over time, whereas development outcomes in the change in shape over time [49]. The fully-grown cranium represents the summation of its different components with differential growth. Growth and development are a continuous process that is sporadic and non-uniform in nature. However, it does not occur randomly. For most of the parts, the growth of any dimension or part of the body occurs according to a pattern. Most body dimensions follow trends that involve the rapid growth separated by a period of relatively slower or steady growth [51]. However, the growth and development of craniofacial skeleton is not isolated but it is related to other parts, and it develops as a unit [52].

At birth, the human skull consists of at least 45 different bones, which undergoes fusion during the growth and development process and results in the adult facial structure, which consists of approximately 22 bones [53]. Apart from fusion, the two basic growth movements involve the process of remodeling and displacement [54], [52], refer to figure 7. Maturation is achieved in males between the ages of 12 and 15 years while the same applies for females between 10 years and 13 years [50]. After maturation, the underlying structure of the face will continue to grow, however, not as rapidly.

Figure 7: Growth and Development in Sub-Adults [86]

Adult aging is dominated by morphological and soft tissue changes i.e. skin texture, wrinkles, etc., but some skeletal changes continue to occur [55]. Early adulthood shows the first signs of soft tissue stressing. Hyper-dynamic expressions will start to show wrinkles on the face. Fine facial lines will appear horizontally on the forehead, vertical lines between the forehead and thin lines around the outside corners of the eyes will appear. From ages, 40 to 50 there are noticeable changes in skin texture while minimal changes are found in younger years. The aging rate of adults differs massively from the individual, which is not the case for sub-adults [56] . These differences can be attributed to genetics

and external features. Biological changes in adults alter the shape and texture of a face. As the skull continues to change with age, the eyes appear smaller as they sink deep into their orbits. As the skins, elasticity begins to degrade wrinkles form, most notably in the eyelids, and the corners of the mouth [56]. These features start to sag and change in size, thus changing the relationships of the characteristics of the face, as in figure 8. Details of the changes in the facial feature concerning age can be found in Appendix F.

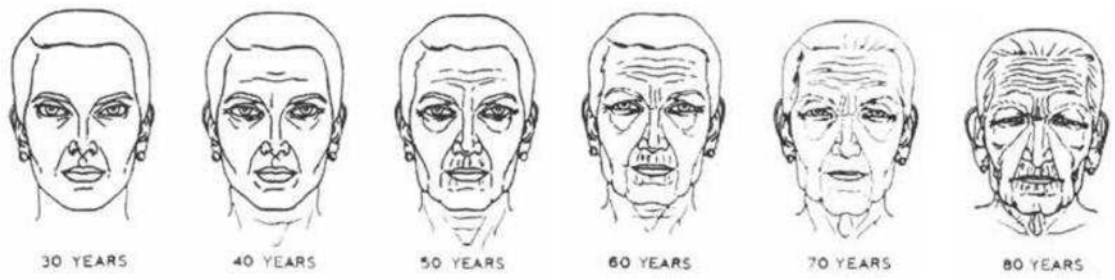


Figure 8: Soft Tissue Changes with Adult Aging [57]

2.5 Image Dataset

Facial images have been analyzed and researched for a variety of applications. The contribution to facial image datasets is not confined to computer vision (face recognition, verification, age progression); it has been widely used in several research areas including anthropology, neuroscience and psychology [1].

Data is the key to enabling design, analysis, comparison or benchmarking facial recognition systems. Along with the development of face recognition algorithms, numerous face databases have been collected for face recognition evaluation. However, as data is contextual in nature, many of these databases are tailored to specific testing conditions. Some of these data sets include:

- "Label Faces in Wild" [58], a database of human face images designed to aid in studying the problem of unconstrained face recognition,

- "SCface - Surveillance Cameras Face Database" [59], emphasizing different law enforcement and surveillance use case scenarios,
- "YouTube Makeup Database" [60] and "Makeup In the Wild" [61] to analyze the effect of makeup on automatic facial recognition,
- "Look-alike Face Database" [62] to address the cases where the intra-class similarity is higher than the inter-class variation for two individuals (look-alike).
- Longitudinal databases like MORPH [63], FGNet [64] and VADANA [65] were introduced to address issues related to aging, verification with age progression and age estimation.

For the exploration of sub-adult facial recognition, we needed an image database with child faces, faces that have not reached biological or physical maturation, which for this work are faces between ages 0 and 18. The public databases that include child faces are FG-NET [64] and the Adience dataset [66]. However, FG-NET did not offer the sufficient number of subjects to evaluate the face recognition systems for children. In addition, most of the images in the dataset are scanned from photographs, which tend to lose anthropometric measures of faces as well as introduces scanning artifacts that are difficult to decouple. That is the reason only a few subjects, around 82 are usable from the FG-NET dataset. Adience dataset contains non-adult subjects and labels its subjects for 8 different age groups (0-2, 4-6, 8-13, 15-20, 25-32, 38-43, 48-53, 60-) [66]; however it is a cross-sectional in nature and does not provide any longitudinal information of subjects. Table 1 outlines all the available aging datasets.

Table 1: Outline of all the available aging datasets

Database	# Subjects	# Images	Images per subject	Age range	Image Quality	Label for Age	Nature
VADNA [65]	43	2,298	3-300	0-78	24-bit colored, 30 scanned	Yes	Longitudinal
FGNET [64]	82	1,002	6 -18	0-69	Mostly scanned images	Yes	Longitudinal
MORPH (Album1) [63]	631	1,690	1-6	16-69	Digitally scanned at 300dpi, gray scaled	Yes	Longitudinal
MORPH (Album 2) [63]	13,673	55,608	1-53	16 – 99	8-bit color 200x240 JPEG or 400x480 JPEG	Yes	Longitudinal
Cross-Age Celebrity Dataset [67]	2,000	163,446	--	16 – 62	24-bit colored images	Yes	Longitudinal
Adience Dataset [66]	2,284	26,580	--	0-65	Flicker album images.	Labeled for age groups	Cross Sectional

The original In-the-Wild Child Celebrity or ITWCC dataset is the only dataset introduced to support the preliminary evaluation of craniofacial morphological changes due to natural aging of sub-adults against face recognition systems [17]. ITWCC focuses on having large sets of individuals, where the subject growth and development can be observed. ITWCC is comprised of 304 subjects with 1,718 images [17]. While ITWCC dataset contains the faces of non-mature subjects over time, it does not provide enough data to carry out some of the experiments and the depiction of real life scenarios. This body of work has extended the original ITWCC dataset to increase the number of subjects and total image count, 501 subjects and 32,515 images.

CHAPTER 3. METHODOLOGY

This work explores the difficulties of using facial recognition frameworks on juvenile faces. While face images have traditionally been used in identification documents such as passports, driver's licenses, voter ID, etc., in recent years, face images are being increasingly utilized in a wide variety of applications that require identification and analytics of young (juvenile) faces such as photo-tagging. Another example, in the security domain is what is the reliability of an e-passport face verification system of a 12 or 13-year-old against his or hers 16 to 18-year-old self. This body of work has examined this type of scenario and has uncovered startling poor results. The culprit as identified in earlier sections is the natural changes that occur during the growth and development phase, which significantly alters the face. Hence, it is important to understand the challenges of juvenile age progression in affecting the performance of face recognition is important.

Depending on the application context, a facial recognition system can be configured for verification, a system confirms the claimed identity of a face presented to it, and identification, a system identifies an unknown face by matching to a set (gallery) of known faces [68]. This work involves investigating the current state-of-the-art via open assessment of modern commercial face recognition technology in both the verification and identification scenarios. The In-The-Wild Child Celebrity-2 (ITWCC-2) dataset is introduced to examine the difficulty of sub-adult aging.

3.1 Commercial Systems Included in Study

The commercial systems involved in the study include Cognitec [69], Rank One Computing [21], and Neurotechnology [22]. These commercial systems were selected to represent the many commercial face matchers available primarily for their ease of access, i.e. the systems are available to the Face Aging Group research laboratory.

- Cognitec: Cognitec's FaceVacs SDK v8.50 is used in this study. It is one of the leading systems used by many governments and leading organizations for image quality check, verification for document issuance, and verification for access control. It also integrates an upgraded algorithm for age estimation and gender detection in newer release [69]. According to the NIST 2014 evaluation report, it is also one of the best performers with the highest accuracy in all the age groups [70].
- Rank One Computing: ROC SDK v1.2 is used in this study. Rank One Computing's roots lie in academia and open source software development with over 30 peer-reviewed papers on the topic of automated face recognition. They have provided contract support in a lead capacity to multiple U.S. government agencies, licensed the technology to industry partners in the biometric space, and released open source software, OpenBR that has since gained a large user base [71]. ROC SDK has claimed the capability of fast and real time face recognition with enrollment speed of around 50 faces, per CPU core, per second and matching speed of 25 million comparisons, per CPU core, per second [21].
- Neurotechnology: Neurotechnology facial identification technology VeriLook SDK, 6.0 trial version is used in this study. Neurotechnology provides computer-based vision and object recognition products to security companies, system integrators, and hardware manufacturers. It provides an algorithm and software development product for a variety of biometric modalities including fingerprint, face, iris, voice, and palm print. With millions of customer installations worldwide, Neurotechnology's product is

used for both civil and forensic applications, including border crossings, criminal investigations, systems for voter registration, verification and duplication checking, passport issuance and other national-scale projects. The VeriLook technology assures system performance and reliability with live face detection, simultaneous multiple face recognition and fast face matching in 1-to-1 and 1-to-many modes. It claims to match up to 40,000 faces per second on a PC, can enroll a face in 0.6 seconds, and can use as little as 4 kilobytes for a face template in a database [22].

3.2 In-The-Wild Child Celebrity-2 Dataset

The original ITWCC dataset was the largest longitudinal collection of adolescent faces at its inception; however, it is not sufficiently large enough to evaluate the full spectrum of the technology. To determine the accuracy and robustness of automatic facial recognition for adolescent images, In-the-Wild Child Celebrity, or ITWCC, was extended. ITWCC-2 is currently the largest dataset with juvenile faces over time. ITWCC-2 focuses on having large sets of individuals, where the subject growth and development can be observed. As the dataset's name 'In-the-wild' suggests, the images are collected with the unrestricted face, and the data corpus is designed to emulate real-life captures. Like ITWCC, images were obtained by exploiting the fame of the subjects and gathered through open Internet sources that are free to use. The data was collected until September of 2015. Table 2 shows the difference between ITWICC and ITWICC-2. The requirements used to augment ITWCC-2 dataset are as follows:

- The subject must have at least three images to qualify.
- The subject must have at least two images less than 16 years of age, the age of sexual maturity.

- The date or at least the year that the photo was taken must be available.
- Subject's year of birth should be known.

Table 2: ITWCC Vs ITWCC-2

	ITWCC	**ITWCC-2**
Number of Subjects	304	501
Number of Images	1,718	32,511
Min Image Age	4 months	4 months
Max Image Age	32 years	32 years
Mean Image Age	13.4 years	11.65 years
Standard Deviation	3.4 years	3.10 years
Median Age	13 year	11 year

The ITWCC-2 dataset is comprised of 501 subjects and 32,511 images. The subject's age ranges from 4 months to 32 years. The mean age of all images is 11.65 years with a standard deviation of 3.10 years. The average age of the first capture is 9.07 years; furthermore, the mean age of final capture is 16.50 years. Figure 9 shows the age ranges of subject and figure 10 shows the number of images for each age in the ITWCC-2 dataset.

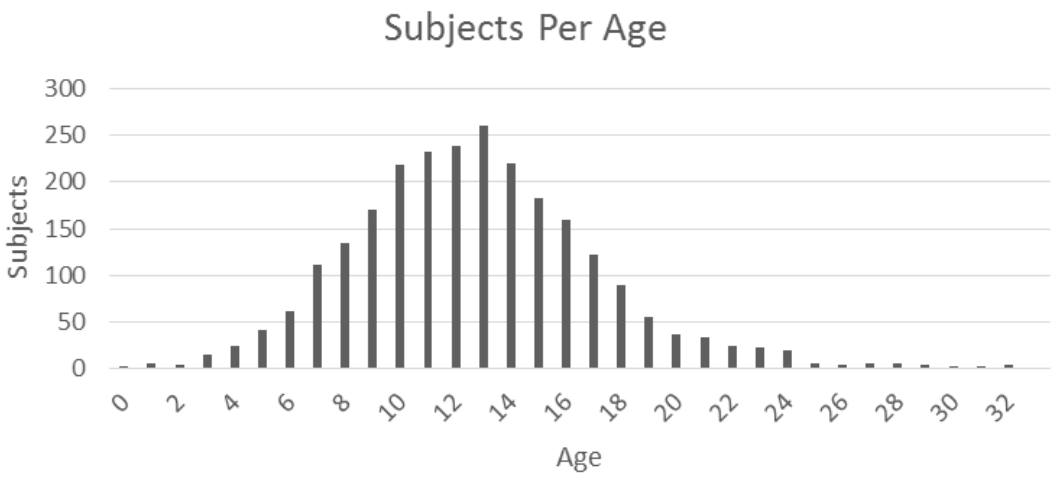


Figure 9: Histogram of Subjects per Age

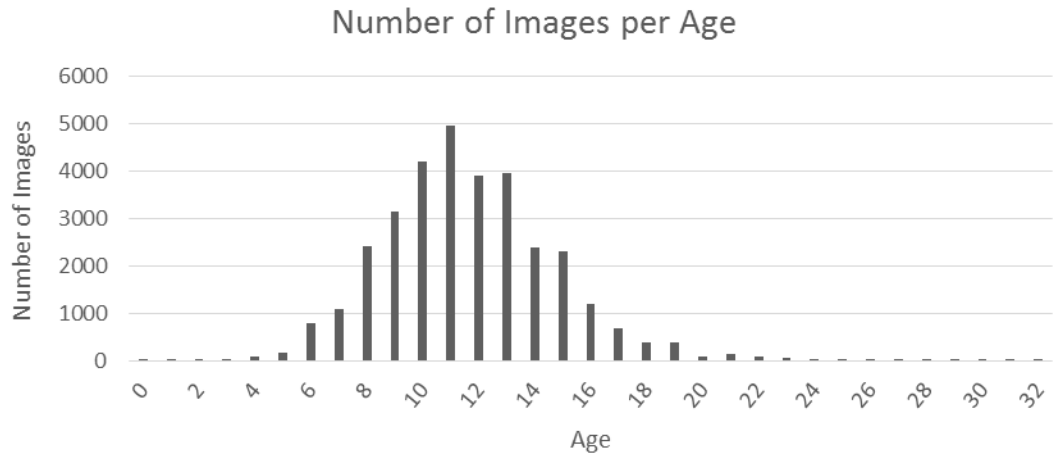


Figure 10: Histogram of Number of Images per Age

In addition to the age information, other meta-data are also captured, including name, date of birth, race, gender, date of capture (photograph taken), a unique image identifier, makeup indicator, glass marker, eye coordinates, the distance between the eyes and URL. This information can further illuminate the differences in gender-specific aging variations and occlusion's effects on facial recognition systems. Table 3 provides the sample of metadata ITWCC-2 dataset. If a month or day from the image capture date is missing, they default to June 15. Dlib's version 18.17 [72], eye detection algorithm is used to acquire the eye coordinates and the distance between eyes. Out of 32,511 images, in 584 images eye detection algorithm fails to extract the eye coordinate. For these 584 images, the eye coordinates were manually annotated. Figure 11 shows a subset of the ITWCC-2 dataset.

Figure 11: In-the-Wild Child Celebrity-2 Dataset

Table 3: Sample metadata collected for the ITWCC-2 dataset.

Name	ID	Count	Image	Age	Ext.	DOB	DOA	Gender	Race	Makeup	Glasses	URL	eye x1	eye y1	eye x2	eye y2	Distance	x1	y1	x2	y2
KyleMassey	CS0013	0	CS0013_00m11	11	.jpg	8/28/1991	1/11/2003	m	a	0	0	http://www.imdb.com/media/rm2245236736/nm10818	96.8	91.7	150.0	103.3	54.4	0	0	292	496
KyleMassey	CS0013	2	CS0013_02m14	14	.jpg	8/28/1991	2/22/2006	m	a	0	0	http://www.imdb.com/media/rm1060412160/nm10818	188.3	205.0	297.0	212.3	108.9	0	0	504	599
KyleMassey	CS0013	3	CS0013_03m18	18	.jpg	8/28/1991	12/5/2009	m	a	0	0	http://www.imdb.com/media/rm4001205760/nm10818	113.5	121.0	146.8	124.2	33.5	0	0	219	366
KyleMassey	CS0013	1	CS0013_01m19	19	.jpg	8/28/1991	10/1/2010	m	a	0	0	http://ia.media-imdb.com/images/M/MV5	99.6	139.2	170.0	139.8	70.3	0	0	284	400
AshleyTisdale	CS0002	1	CS0002_01f01	1	.jpg	7/2/1985	1/1/1987	f	c	0	0	http://www.blog-city.info/en/img6/6589_7_p	126.2	125.2	167.7	109.8	44.2	0	0	282	245
AshleyTisdale	CS0002	2	CS0002_02f05	5	.jpg	7/2/1985	1/1/1991	f	c	0	0	http://i4.ytimg.com/vi/7bgsnW_1J98/hqdefault.jpg	208.2	166.0	271.7	166.8	63.0	0	0	480	360
AshleyTisdale	CS0002	4	CS0002_04f16	16	.jpg	7/2/1985	1/1/2002	f	c	0	0	http://ashleytisdale.org/photos/albums/photoshoot/Offi	141.7	120.2	248.7	146.7	109.8	0	0	391	500
AshleyTisdale	CS0002	6	CS0002_06f19	19	.jpg	7/2/1985	1/1/2005	f	c	1	0	http://ashleytisdale.org/photos/albums/photoshoot/And	319.5	99.0	372.0	100.8	52.5	0	0	500	332
ChristianBale	CS0170	0	CS0170_00m13	13	.jpg	1/30/1974	3/3/1987	m	c	0	0	http://24.media.tumblr.com/tumblr_m9f2y2cFcF1rov8l	181.2	112.2	227.8	112.5	46.7	0	0	400	400
ChristianBale	CS0170	3	CS0170_03m14	14	.jpg	1/30/1974	3/3/1988	M	C	0	0	http://25.media.tumblr.com/tumblr_lgor1roKpD1qef7d	132.7	148.5	195.5	154.0	63.1	0	0	308	400
ChristianBale	CS0170	4	CS0170_04m18	18	.jpg	1/30/1974	2/2/1992	m	c	0	0	http://images2.fanpop.com/image/photos/11100000/Ne	137.7	125.7	186.0	132.0	49.3	0	0	402	287
ChristianBale	CS0170	1	CS0170_01m24	24	.jpg	1/30/1974	10/26/1998	m	c	0	0	http://i2.cdnds.net/12/27/618x473/movies_christian_ba	289.7	144.5	349.0	142.7	59.4	0	0	618	473
MichelleTrachtenberg	CS0606	11	CS0606_11f08	8	.jpg	10/11/1985	6/15/1994	f	c	1	0	http://www.childstarlets.com/captures/videocaps/mtra	194.2	183.8	334.0	175.8	140.1	0	0	627	474
MichelleTrachtenberg	CS0606	91	CS0606_91f10	10	.jpg	10/11/1985	6/15/1996	f	c	0	0	http://www.childstarlets.com/captures/videocaps/mtra	366.3	206.8	491.2	207.2	124.8	0	0	938	520
MichelleTrachtenberg	CS0606	1	CS0606_01f15	15	.jpg	10/11/1985	1/6/2001	f	c	0	0	http://www.imdb.com/media/rm4030896640/nm00055	109.7	118.0	146.5	115.2	36.9	0	0	272	400
MichelleTrachtenberg	CS0606	4	CS0606_04f18	18	.jpg	10/11/1985	11/20/2003	f	c	1	0	http://www.imdb.com/media/rm2536741376/nm00055	103.8	129.8	153.0	121.3	49.9	0	0	266	400

3.3 Operating Environments/Mode

Face recognition technology can be deployed in different environments to attack different operating scenarios. However, depending on the application context, a facial recognition system may operate in two different modes: verification or identification. This section will provide the overview of the working environment for the experiments, and a definition of the types of experiments conducted.

Verification systems are the biometric system that seek to confirm, whether the individual is whom they said, they are [68]. It is a system designed to answer the question, "Is this person who they say they are?" Under such facial recognition system, an individual present's himself as a particular person. The system checks their facial image (acquired as the probe or query) against a picture profile that already exists in the database linked to that person's file (gallery or enrolled) to find a match. The aim of verification systems is to authenticate the claimed identity.

Face verification systems are described as a one-to-one matching system because the system tries to compare an image presented by the individual against a specific enrolled image. As per the comparison score, it provides a genuine or imposter decision. Verification systems are typically used for positive recognition with the aim to prevent multiple user from accessing the system. Verification is used to establish and confirm identity that assists in securing networks and digital assets, identify threats and persons of interest. Because verification systems are one-to-one matching systems, they can generate results more quickly and are more accurate than identification systems, even when the size of the database increases, i.e. the performance of verification is independent on the number

of enrolled as far as the access time of the enrolled image is not accounted. Figure 12 represents a simple face verification process.

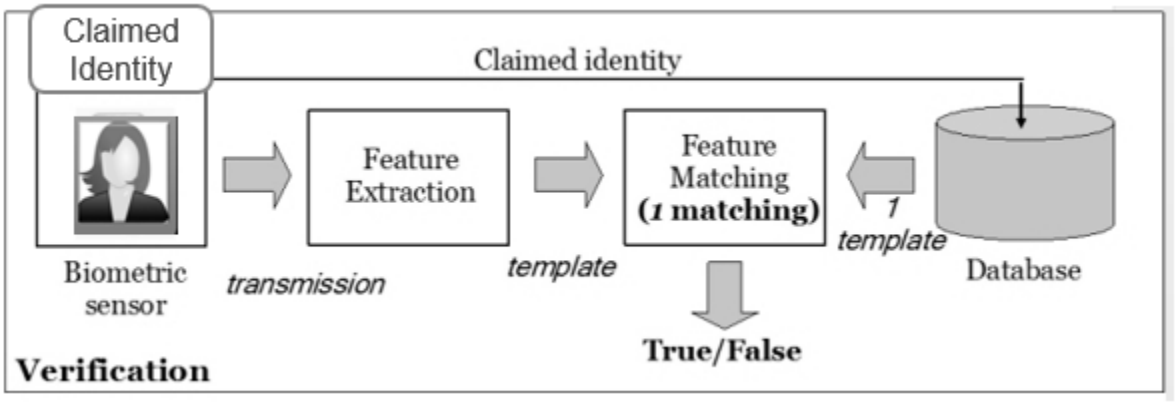


Figure 12: Verification Process

Face identification systems seek to identify an unknown person. These systems try to answer the questions "Who is this person in the presented image?" [68]. It is more complicated than the verification process as it includes a comparison of the acquired and processed facial trait of an individual with all the enrolled templates in the database. That is also the reason, its accuracy and speed depends on the number of registered images.

Identification systems are described as one-to-many match systems and gives a ranked list of matches. It is a critical component in negative recognition applications with a purpose to prevent the same person from using multiple identities. This is also known as de-duplication as used in driver's license scenario. The one-to-many identification is the largest market for face recognition technology. Grother said, "These algorithms are used around the world to detect duplicates in databases, fraudulent applications for passports and driving licenses, in tokenless access control, surveillance, social media tagging, lookalike discovery and criminal investigations" [43]. Figure 13 represents a simple identification process.

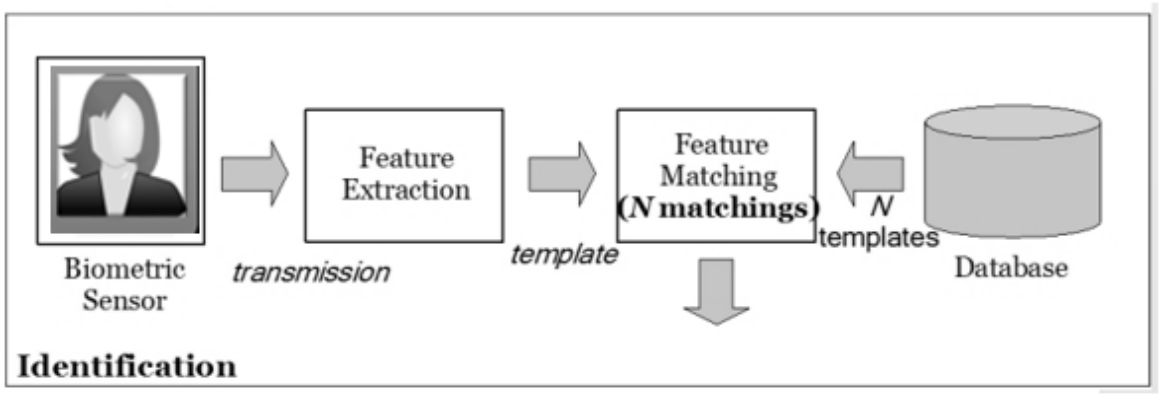


Figure 13: Identification Process

This work includes the evaluation of facial recognition technology in both, verification and identification modes.

3.4 Evaluation Scenarios

The environments included in this study to evaluate the recognition accuracy for sub-adults, were designed to mimic the real operating scenarios of automatic facial recognition technology. In-the-Wild Child Celebrity-2 dataset is used to comprehend the effect of sub-adult aging. The first scenario is the access control scenario, which is a verification scenario, where the presented facial image of a user is matched against an enrolled image to verify the user's identity and grant the access. This scene includes all 501 ITWCC-2 subjects where all the subject images match against all other images of the subject to confirm the identity of the person. This situation will identify problems that may occur due to verification matching against temporally displaced images, i.e. subject's young image is enrolled and match is made against all older images of the subject. If a match does not occur then one can presuppose that the issue is the time displacement. The reality is that in this dataset a match may not be made because of PIE errors, i.e. variation because of pose, illumination, or expression.

Next scenarios depict the photo-tagging scenario to verify the performance of automatic facial recognition technology with temporally displaced images. One is to gauge the identification accuracy of the system from a younger image. Here suppose an adult, 22-year-old person, has a Facebook account with all their current pictures. Now the user uploads an image from his/her childhood, will the system be able to recognize the person correctly and apply the right auto-tag? A parent sets up an online photo organizations app like Picas and uploads childhood images (the earliest) images of their children; will the system be able to auto-tag all older images presented to the system? What is the accuracy rate of this auto-tagging system?

The face images are highly unconstrained and hence, degradation in performance could be a result of the PIE attributes of the image. So, to understand the effects of PIE and to divorce these effects from this work a baseline is established against a similar adult dataset. The adult dataset used in Labeled Faces in the Wild. This dataset has been used in many evaluations of face recognition system. LFW database exhibits "natural" variability in the pose, lighting, focus, resolution, facial expression, age, gender, race, accessories, make-up, occlusions, background, and photographic quality [58]. LFW contains only adult subjects, and the acquisition and diversity of images are similar to ITWCC-2. LFW and ITWCC-2 dataset mainly differ in subject age range, which makes them perfect for measuring the challenges and differences in face recognition technology because of adult and sub-adult aging.

3.5 Performance Evaluation

A traditional token-based authentication approach like passwords can guarantee 100% accuracy in match, i.e. either the password is correct or it is not, however, in biometric systems authentication match is not 100%. The match in a face recognition

system designates the degree of similarity between the presented biometric trait (sample template) and the biometric trait stored in a database (reference template). Typically, the presented sample inevitably varies somewhat from the reference template, due to subtle changes over time, imperfect and varying sensing conditions, variation in interaction with the sensor and errors in the process of feature extraction. That is the reason the distance between the two templates originated from the same biometric trait of a user is typically non-zero. In fact, in a biometric system such as face recognition, a perfect match may flag a potential fraud; presenting an exact match may imply that the reference template database has been compromised [73]. The variability observed between the templates is referred as intra-class variation and the variability between the templates originated from two different individuals is regarded as an inter-class variation.

Fundamentally, each system will match at least two biometric templates, one being the stored template or gallery/enrolled image and the other being probe, to produce a match score that will determine acceptance or rejection. This match score is a standardized number that shows the likeness between the two templates. A genuine user is a person (template) who should match with the reference template, whereas an imposter is a person (template) that should not match with the reference template. Ideally, all legitimate users should be accepted while all impostors should be rejected. The performance of a face recognition system is traditionally characterized by following statistics [68], also shown in figure 14:

- True accept rate (TAR), the ratio of genuine users who have been accepted [68].
- True reject rate (TRR), the ratio of impostors who have been correctly rejected [68].

- False accepts rate (FAR), a system incorrectly accepts an identity [68].
- False rejects rate (FRR) is the measure of the likelihood that the biometric system will incorrectly reject an access attempt by an authorized user [68].

		Gallery	
		Genuine	Imposter
Probe	Genuine	TAR	FRR
	Imposter	FAR	TRR

Figure 14: Statistics measures to characterize the performance of a face recognition system

To evaluate the performance of all three commercial systems, Open Source Biometric Recognition (OpenBR) evaluation toolkit version 0.5.0 is used. OpenBR is a collaborative tool that provides a method for researchers to compare algorithms in a controlled environment [71]. Klontz et al. stated, "OpenBR provides tools to design and evaluate new biometric algorithms and an interface to incorporate biometric technology into end-user applications" [71]. OpenBr leverages the extensive support from the research community and provides a simple command line interface that can be used to create and evaluate the algorithms, as well as an API for creating plug-ins. OpenBr toolkit evaluates the similarity score information and then plots the information in a standardized format. With the mentioned statistics, following plots are also studied to measure the performance of face recognition system:

- Receiver operating characteristic (ROC) is based on aggregate statistics of match scores corresponding to all biometric samples. It shows the accuracy of a biometric system, by comparing the false acceptance rate versus the verification rate. For ROC, higher values indicate greater accuracy [4].

- Detection tasks can be viewed as involving a tradeoff between two error types: missed detections and false alarms. Detection error trade-off (DET) Curve has distinct advantages over the standard ROC type curve for presenting performance results where tradeoffs between these error types are involved. DET plots the false reject rate against the false accept rate to determine the error rates for each algorithm [4].
- Score histograms plot the frequency of verification scores. Values are listed from 0 - 1 and both genuine and imposters are shown on this chart. An ideal plot would have genuine matches nearing 1 while imposters nearing 0 with a clear division between them [74].
- Cumulative match characteristic (CMC) shows the accuracy of a closed-set identification process. During the identification process, all images are scored and ranked in an ascending order based on their match score. The CMC graph shows the percentage that the legitimate user is in a rank less than or equal to the current rank being plotted [4].

CHAPTER 4. EXPERIMENT DESIGN

This work involves exploring the difficulty of temporally displaced sub-adult data i.e. children from 0-18 years. To accomplish this, the recognition accuracy of three different commercial systems were examined in three different scenarios, where temporally displaced image data would be often used in real life. Cognitec's FaceVace v8.50 [20], Rank One Computing's ROC v1.20 [21] and Neurotechnology's Verilook v6.0 [22] are the commercial face recognition systems included in the study. These systems are widely deployed and hence, are a good representation of the field of commercial systems.

As discussed in section 2.1, face recognition is a multistep process and the performance of a system as a whole depends on the accuracy of its sub-modules like normalization, feature extraction, or match module. Face alignment is an important part of face recognition process and has a great impact on recognition accuracy of a system [75]. Face alignment involves spatially scaling and rotating a facial image to match with face pictures in the database. In addition, many algorithms use eye positions, for face alignment. Therefore, it is essential to have a robust algorithm to detect and extract eye coordinates. For most of the face recognition methods, eye positions are manually given. However, for a real-world application of face recognition, manually detecting eye positions is not realistic. Therefore, an automatic eye detection algorithm is needed for a fully automatic face recognition system.

The systems included in the study are evaluated independently for sub-adults. Each system provides its API for eye detection and comparison. These API's are used to process and compare the images. The match scores from the comparison API is then fed to OpenBR's evaluation toolkit to evaluate the accuracy of the system and to plot the results.

4.1. Failure to Acquire/Enroll (FTA/FTE)

Failure to acquire can be defined as a failure of a biometric system to capture and/or extract usable information from a biometric sample [4]. FTA is one of the significant performance ratio and measures the efficiency of feature extractor/face detection module of a biometric system. As mentioned in section 3.1, out of 32,511 ITWCC-2 images, Dlib's eye detection algorithm [72] failed to extract the eye coordinates for 584 images. It means Dlib's fails to find eye positions for 584 out of 32,511 images resulting in a FTA rate 1.80%.

Table 4 shows the FTA rate of all three systems included in the study with ITWCC-2 dataset and LFW dataset, with a total of 32,511 images in ITWCC-2 and 13,233 images in LFW dataset. As the system can only perform the comparison if it can extract eye coordinates from the presented image, all the images are ignored in the experiment for the system, where it is not able to detect eye. FTA is highest for Verilook with both adults (LFW) and sub-adult (ITWCC-2) data, which indicates that Verilook is the poorest performing system to detect a face from unconstrained image data.

Table 4: Failure to Acquire

SDK	**ITWCC-2**	**LFW**
Verilook (6.0 trial)	33.92%	79.51%
FaceVacs (8.50)	19.22%	1.16%
Rank One Computing (1.2)	9.01%	3.92%

4.2 Access Control/All-to-All Verification

The first experimental scenario considered is the access control scenario. The purpose is to determine the effectiveness of current technology while comparing temporally displaced sub-adult image in verification mode. This experiment verifies all images within the ITWCC-2 dataset against all other images. As shown in figure 15, an

image of an individual is compared with all the images belonging to the same person and all other enrolled images. The Access Control Scenario was conducted to understand how efficiently the current commercial systems could verify adolescent faces regardless of which image was enrolled. All the images of the ITWCC-2 dataset are used in this scenario. Table 5 lists the number of gallery and probe images of the different commercial systems and table 6 describes the match matrix of all-to-all comparison. Match matrix lists the number of genuine score, imposter score, and total number of comparisons for the experiment. It also lists the ignored match, these are the number of images that are compared to themselves and are masked from evaluation.

Table 5: All-to-All Verification Experiment Data Usage (ITWCC-2)

SDK	Probe	Gallery
Verilook (6.0 trial)	21,485	21,485
FaceVacs (8.50)	26,266	26,266
Rank One Computing (1.2)	29,568	29,568

Table 6: All-to-All Verification Experiment Match Comparison (ITWCC-2)

SDK	Num. of Genuine Matches	Num. of Imposter Matches	Num. of Ignore Matches	Num. of Total Matches
Verilook (6.0 trial)	2,978,380	458,605,360	21,485	461,605,225
FaceVacs (8.50)	4,043,980	685,832,510	26,266	689,902,756
Rank One Computing (1.2)	5,595,924	868,641,132	29,568	874,266,624

The same experiment is repeated with the labeled faces in-the-wild (LFW) dataset. The purpose is to determine the effectiveness of face verification process with adult subjects. It also provides a benchmark to verify the gap in performance of state-of-the-art for adult and sub-adult subjects. All the images of LFW dataset are used in this scenario. Table 7 lists the number of gallery and probe images of the different commercial systems,

all the images where the system fails to detect eye coordinates are ignored for the system. Table 8 contains the match matrix of all-to-all comparison.

Table 7: All-to-All Verification Experiment Data Usage (LFW)

SDK	Probe	Gallery
Verilook (6.0 trial)	12,714	12,714
FaceVacs (8.50)	2,712	2,712
Rank One Computing (1.2)	13,077	13,077

Table 8: All-to-All Verification Experiment Match Comparison (LFW)

SDK	Num. of Genuine Matches	Num. of Imposter Matches	Num. of Ignore Matches	Num. of Total Matches
Verilook (6.0 trial)	451,960	161,181,122	12,714	161,645,796
FaceVacs (8.50)	37,322	7,314,910	2,712	7,354,944
Rank One Computing (1.2)	477,186	170,517,666	13,077	171,007,929

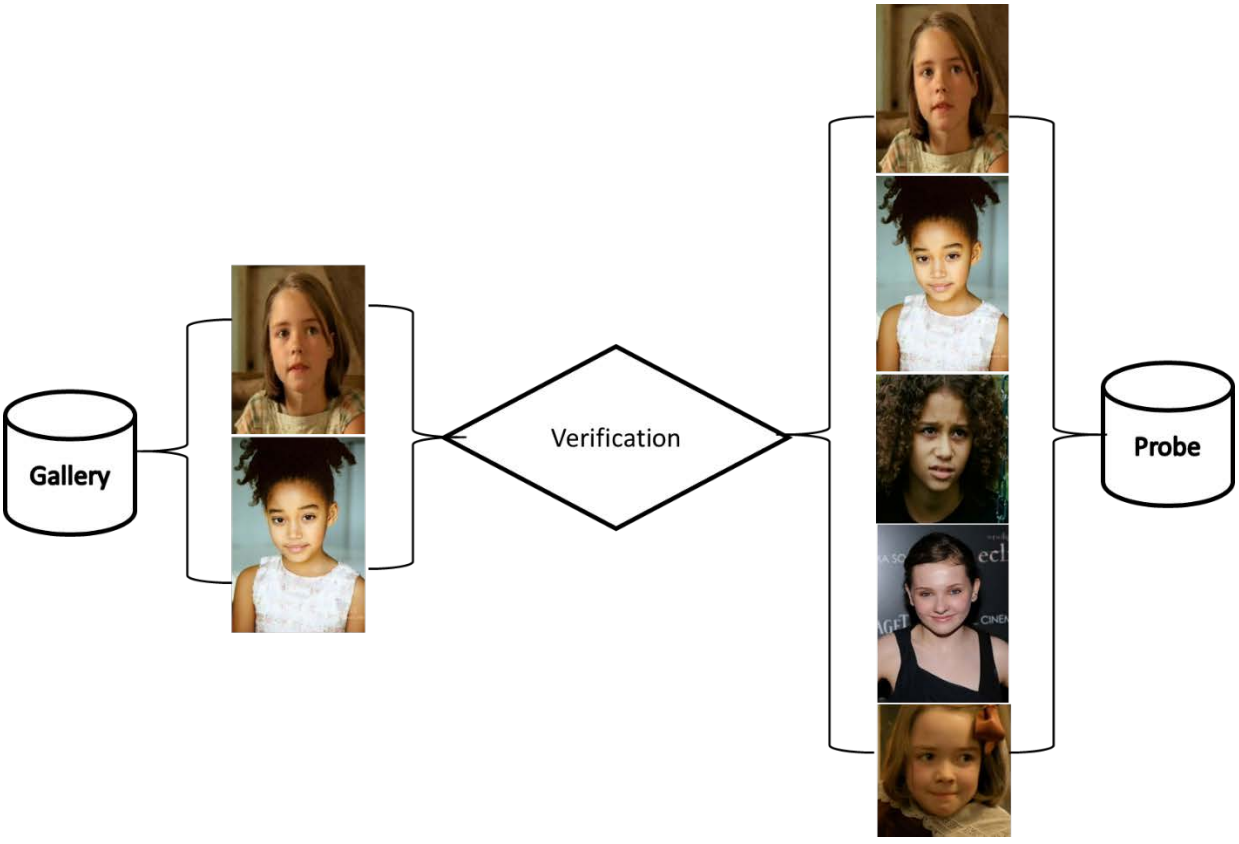


Figure 15: Access Control Design

4.3 Sub-Adult Aging Identification

This scenario is intended to explore how adolescent aging will affect the performance of face recognition system in identification scenario. This experiment attempts to setup a scenario in which an end user of a photo-tagging tool, such as Facebook, Picasa, etc., would begin adding images over a span of time.

4.3.1 Young-to-Old Identification

The presented experiment is to evaluate the performance of a face identification system when the system tries to recognize a person from its younger image. Similar to experiment 1, the ITWCC-2 dataset is used to measure the performance of face recognition technology with adolescence data. Only the youngest image of each person is enrolled in the gallery, and all other images for the individual are used as probe image. Figure 16 depicts the experiment scene where the youngest image is matched to all of its older images. The approximate mean age of the gallery is 9.5 years with a standard deviation of 2.4 years. The remaining images were then placed in the probe set. The mean age of the probe set is approximately 12.2 years with a standard deviation of 3.1 years. These mean represent the data as a whole and do not reflect the individual age differences between each subject. Table 9 describes the experiment data usage, and table 10 lists the match matrix of young-to-old comparison. Note: the difference in the probe and gallery data and as a result the mean (average) age is due to failure to acquire/enroll. Further, in this experiment there was not any overlapping images from the probe-gallery as seen in Table 10.

Table 9: Young-to-Old Identification Experiment Data Usage (ITWCC-2)

SDK	**Probe**	**Gallery**	**Average Age**		**Standard Deviation**	
			Probe	**Gallery**	**Probe**	**Gallery**
Verilook (6.0 trial)	25,160	4,400	12.1	9.4	3.1	2.3
FaceVacs (8.50)	17,831	3,622	12.4	9.6	3.1	2.4
Rank One Computing (1.2)	22,078	4,184	12.1	9.4	3.1	2.5

Table 10: Young-to-Old Identification Experiment Match Comparison (ITWCC-2)

SDK	Num. of Genuine Matches	Num. of Imposter Matches	Num. of Ignore Matches	Num. of Total Matches
Verilook (6.0 trial)	370,397	110,333,603	0	110,704,000
FaceVacs (8.50)	241,845	64,342,037	0	64,583,882
Rank One Computing (1.2)	286,963	92,087,389	0	92,374,352

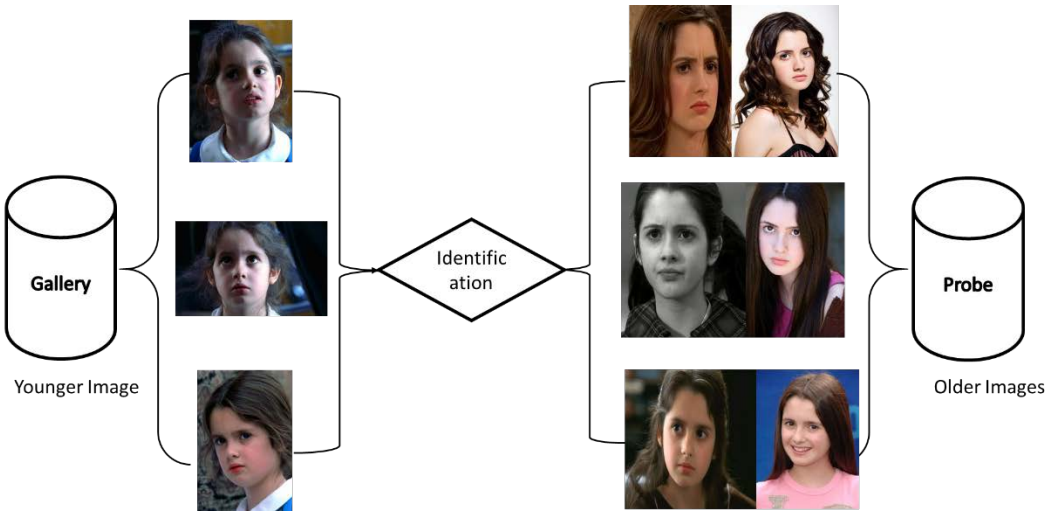


Figure 16: Young-to-Old Experiment Design

4.3.2 Old-to-Young Identification

The presented experiment is to evaluate the performance of a face identification system when the system tries to recognize a person from its older image under unconstrained conditions. The ITWCC-2 dataset is used to measure the performance of face recognition technology with adolescent data. Only the oldest image of each individual is enrolled in the gallery, and all other images for the individual are used as probe image. Figure 17 depicts the experiment setup where the oldest image is matched against all of its younger pictures. The mean age of the gallery is approximately 14.7 years with a standard deviation of 3.7 years. The remaining images were placed in the probe set. The mean age of the probe set is 11.4 years with a standard deviation of 2.9 years. Here, mean represent

the data as a whole and do not reflect the individual age differences between each subject. Table 11 describes the experiment data usage, and table 12 lists the match matrix of Old-to-Young comparison.

Table 11: Old-to-Young Identification Experiment Data Usage (ITWCC-2)

SDK	Probe	Gallery	Average Age		Standard Deviation	
			Probe	Gallery	Probe	Gallery
Verilook (6.0 trial)	25,795	3,765	11.4	14.5	2.9	3.7
FaceVacs (8.50)	18,500	2,953	11.5	14.6	2.9	3.8
Rank One Computing (1.2)	23,105	3,157	11.3	14.4	2.9	3.7

Table 12: Old-to-Young Identification Experiment Match Comparison (ITWCC-2)

SDK	Num. of Genuine Matches	Num. of Imposter Matches	Num. of Ignore Matches	Total Matches
Verilook (6.0 trial)	405,283	96,712,892	0	97,118,175
FaceVacs (8.50)	216,956	54,413,544	0	54,630,500
Rank One Computing (1.2)	255,436	72,687,049	0	72,942,485

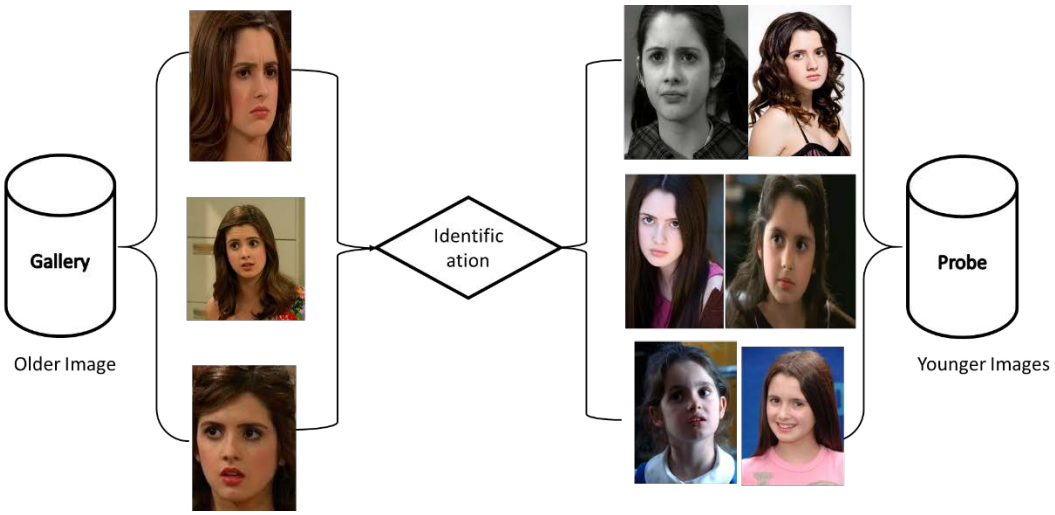


Figure 17: Old-to-Young Experiment Design

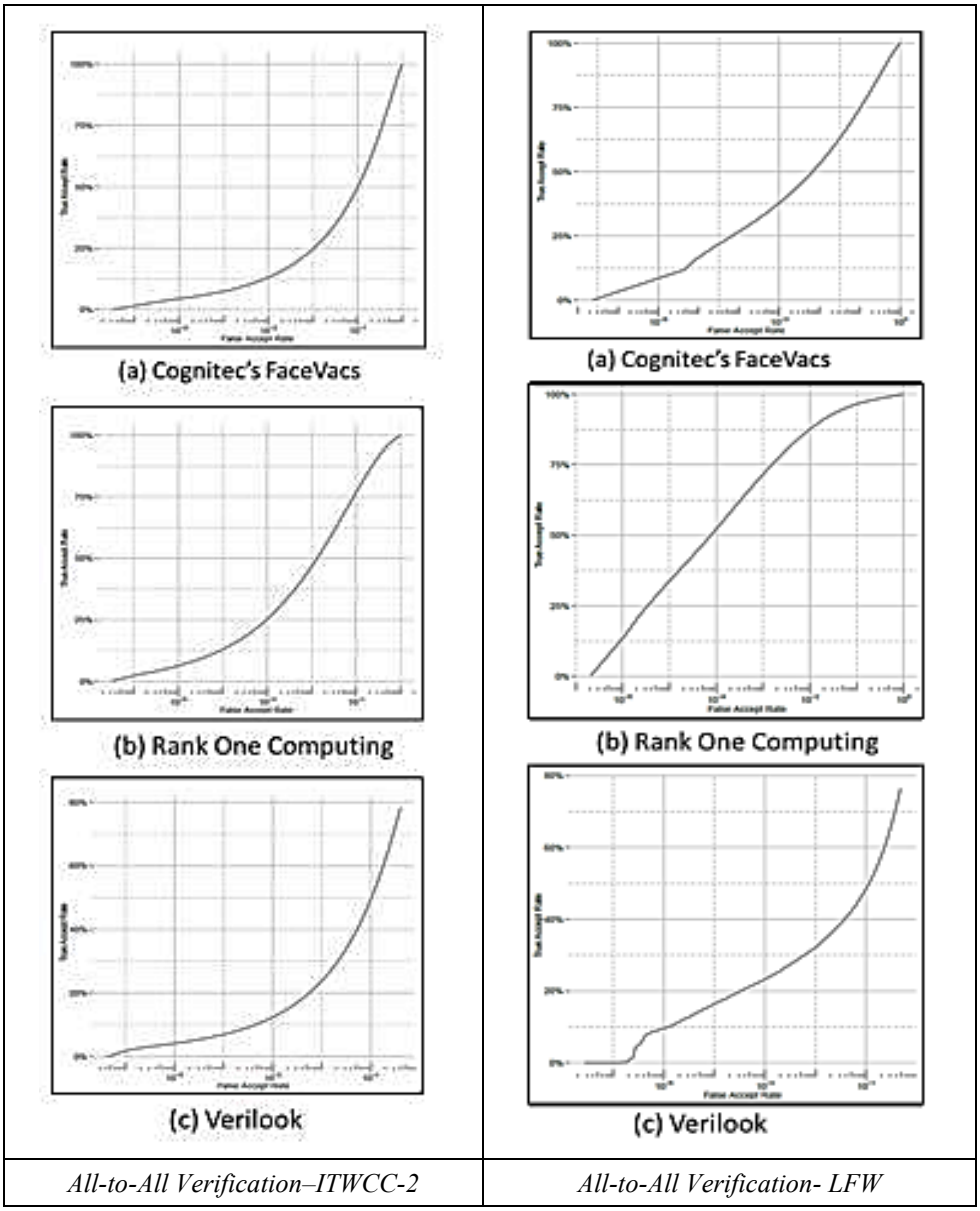
CHAPTER 5. RESULT

This work provides two scenarios to understand the challenges of facial recognition on temporally distanced sub-adult faces: access control and photo tagging. The considered scenarios, and crafted experiments facilitates the evaluation of FRS for sub-adult verification and identification. To access the performance, this work evaluated three commercial algorithms using standard performance measures, e.g. true accept rate (TAR), false accept rate (FAR), match score histograms, detection error trade-off (DET), and receiver operating characteristic (ROC). All measures collected for this work against the commercial matchers can be found in Appendix B-E

The ROC is the favoured metric for evaluating FRS against one another on a common data corpus. The ROC plots the TAR against the FAR to give an estimate of how often an imposter will gain access to the system versus how often the system accepts a legitimate user. A 1% FAR means that one imposter out of 100 will gain access and FAR of .1% (1e-03) means, that one imposter out of 1,000 who attempt to access the system will indeed breach access. This is compared to the TAR, which is the percentage of genuine users that gain access because they are who they claim to be. This plot is often used to tune FR systems for their task. For example, a high security complex, like a biohazard lab, may want to ensure that it is rare that an imposter will gain access; the system may set FAR 1e-06, which means that one in a million imposter attempts on average will be successful. Now this means that the TAR may suffer which means that some genuine users may have to present their face images to the system multiple times before a match is made and they gain access. Table 13 lists the TAR of the All-to-All Verification experiment with different FAR and provides an estimate of the accuracy of the commercial systems included in the study with sub-adult images of ITWCC-2. The Access control scenario shows that Rank

Table 13: All-to-All ROC Values

True Accept Rate at different False Accept Rate			
FAR	**Cognitec v8.50**	**Verilook v6.0**	**Rank One Computing v1.20**
$1e^{-06}$	1.8%	1.04%	2.4%
$1e^{-05}$	4.5%	04.1%	6.4%
$1e^{-04}$	7.5%	7.0%	13%
$1e^{-03}$	13.1%	12.3%	25.4%
$1e^{-02}$	***24.7%***	***23.7%***	***46.8%***
$1e^{-01}$	49.8%	33.0%	76.9%



All-to-All Verification–ITWCC-2 | *All-to-All Verification- LFW*

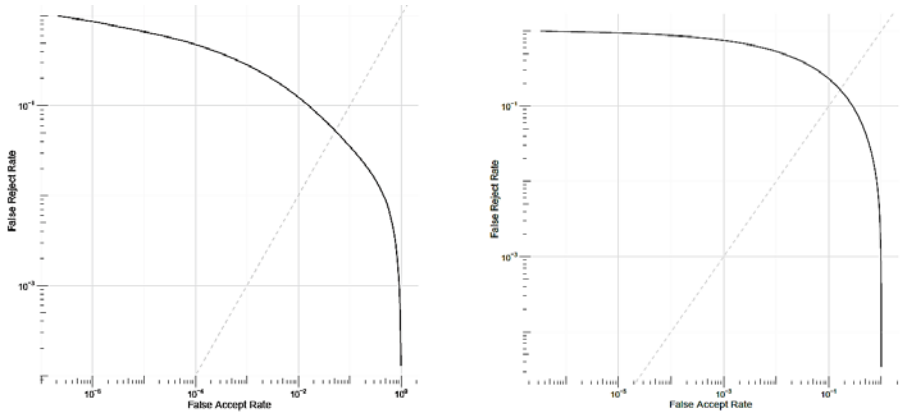
Figure 18: Receiver Operating Characteristics Curves for All-to-All Verification Experiments

One Computing v1.20 performed significantly better than the other two systems with a 46.8% TAR at 1% FAR, followed by FaceVacs v8.50 with 24.7%, and then Verilook v6.0 with 23.7%.

The all-to-all verification experiment conducted with adult faces of LFW provides a way to compare and contrast the performance in terms of recognition for the current-state-of-the-art as it relates to the system's ability to deal with unconstrained face matching. Again, the hypothesis is that if an adult unconstrained face corpus performances markedly better than the sub-adult then the nature of the changes of the sub-adult face is the cause of the performance degradation. From Table 14 and figure 18, one can observe that the LFW (adult unconstrained corpus) performed demonstrably better at every FAR. Hence, there is something else driving the degradation in performance other than the unconstrained nature of the images, i.e. it is not the pose, illumination, or expression, it must be the aging, craniofacial morphology associated with maturation. From the table and the other graphs included in Appendix B and E, confirms that the recognition accuracy is significantly better for adult faces. The DET graphs shown in figure 19 revels that both false negative and false positive rates are much higher for ITWCC-2 sub-adult faces, which indicate that it is more difficult to identify children.

Table 14: TAR comparisons for adult and sub-adult faces

	Rank One Computing v1.20		**Cognitec v8.50**		**Verilook v6.0**	
FAR	**ITWCC-2 (sub-adult)**	**LFW (adult)**	**ITWCC-2 (sub-adult)**	**LFW (adult)**	**ITWCC-2 (sub-adult)**	**LFW (adult)**
$1e^{-06}$	2.4%	12.2%	1.8%	2%	1.04%	0%
$1e^{-05}$	6.4%	33.8%	4.5%	16.5%	04.1%	9.5%
$1e^{-04}$	13%	52.2%	7.5%	26.7%	7.0%	16.4%
$1e^{-03}$	25.4%	71.6%	13.1%	37.9%	12.3%	23.3%
$1e^{-02}$	***46.8%***	**87.6%**	***24.7%***	**53.1%**	**23.7%**	**32.1%**
$1e^{-01}$	76.9%	96.5%	49.8%	74.7%	33.0%	48.4%



(a) DET Plot for LWF Faces (b) DET Plot for ITWCC-2 Faces

Figure 19: DET Plots for Rank One Computing 1.20 SDK

Table 15 lists the true accept rates for the Young-to-Old Identification experiment, and Table 16 describes the true accept rate for the Old-to-Young Identification test, with different false acceptance rates. Once again, in both the experiments Young-to-Old Identification and Old-to-Young Identification, Rank One Computing v1.20 outperforms other two commercial systems included in the study by a relatively significant margin. Figure 20 and figure 21, displays the ROC plots for Old-to-Young and Young-to-Old experiments.

Table 15: Young-to-Old True Accept Values

True Accept Rate			
FAR	**Cognitec v8.50**	**Verilook v6.0**	**Rank One Computing v1.20**
$1e^{-06}$	0.2%	0.2%	0.7%
$1e^{-05}$	2.3%	1.8%	3.7%
$1e^{-04}$	4.7%	4.1%	8.6%
$1e^{-03}$	9.5%	8.4%	18.3%
$1e^{-02}$	***19.8%***	***18.3%***	***37.1%***
$1e^{-01}$	43.8%	42.5%	68.2%

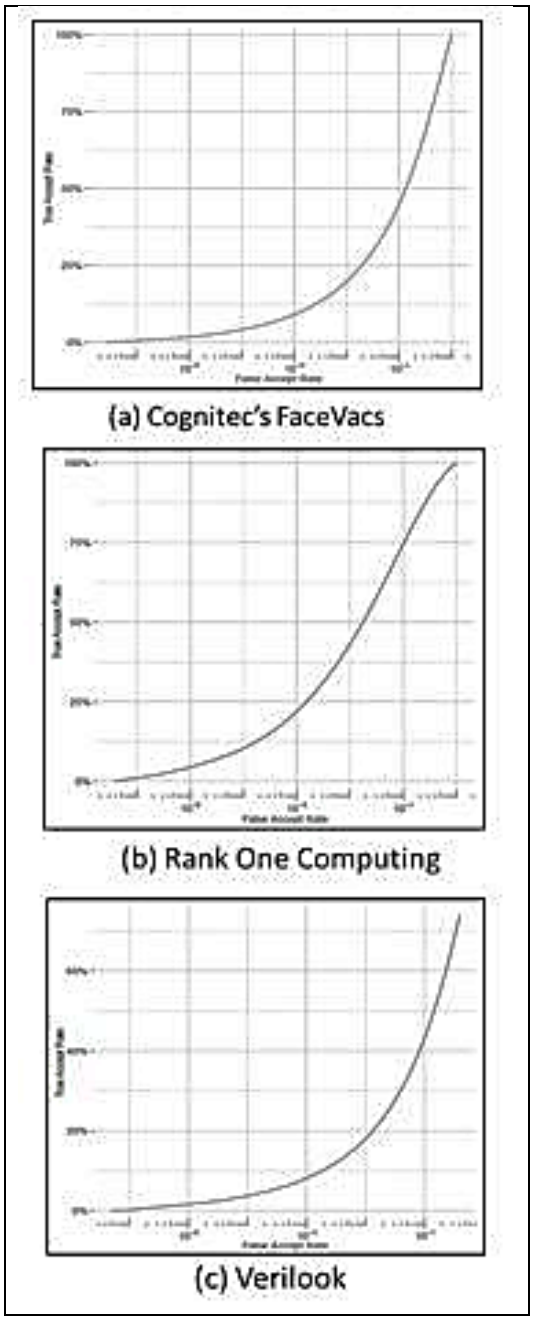


Figure 20: Receiver Operating Characteristics – Young-to-Old Experiment

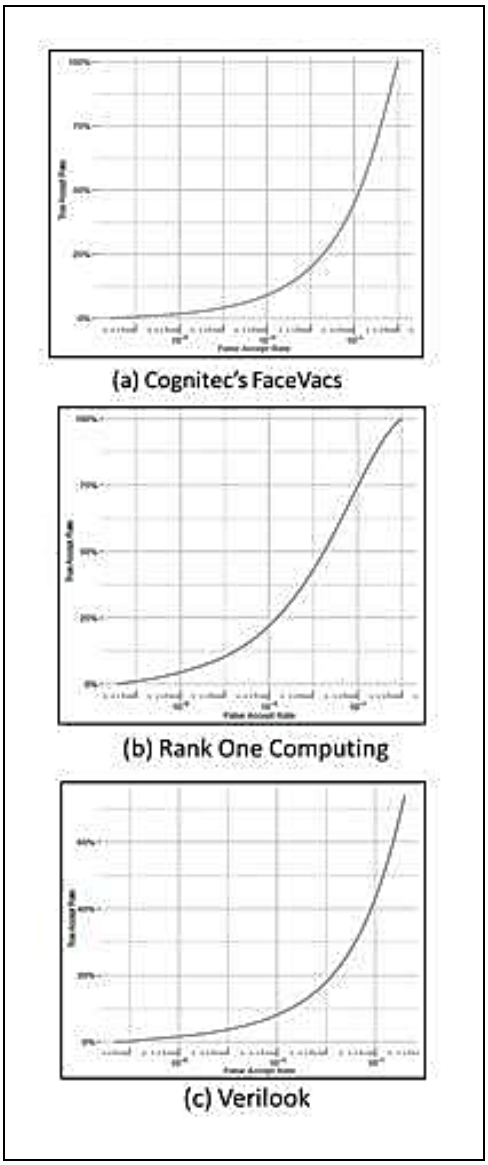


Figure 21: Receiver Operating Characteristics - Old-to-Young Experiment

Table 16: Old-to-Young True Accept Values

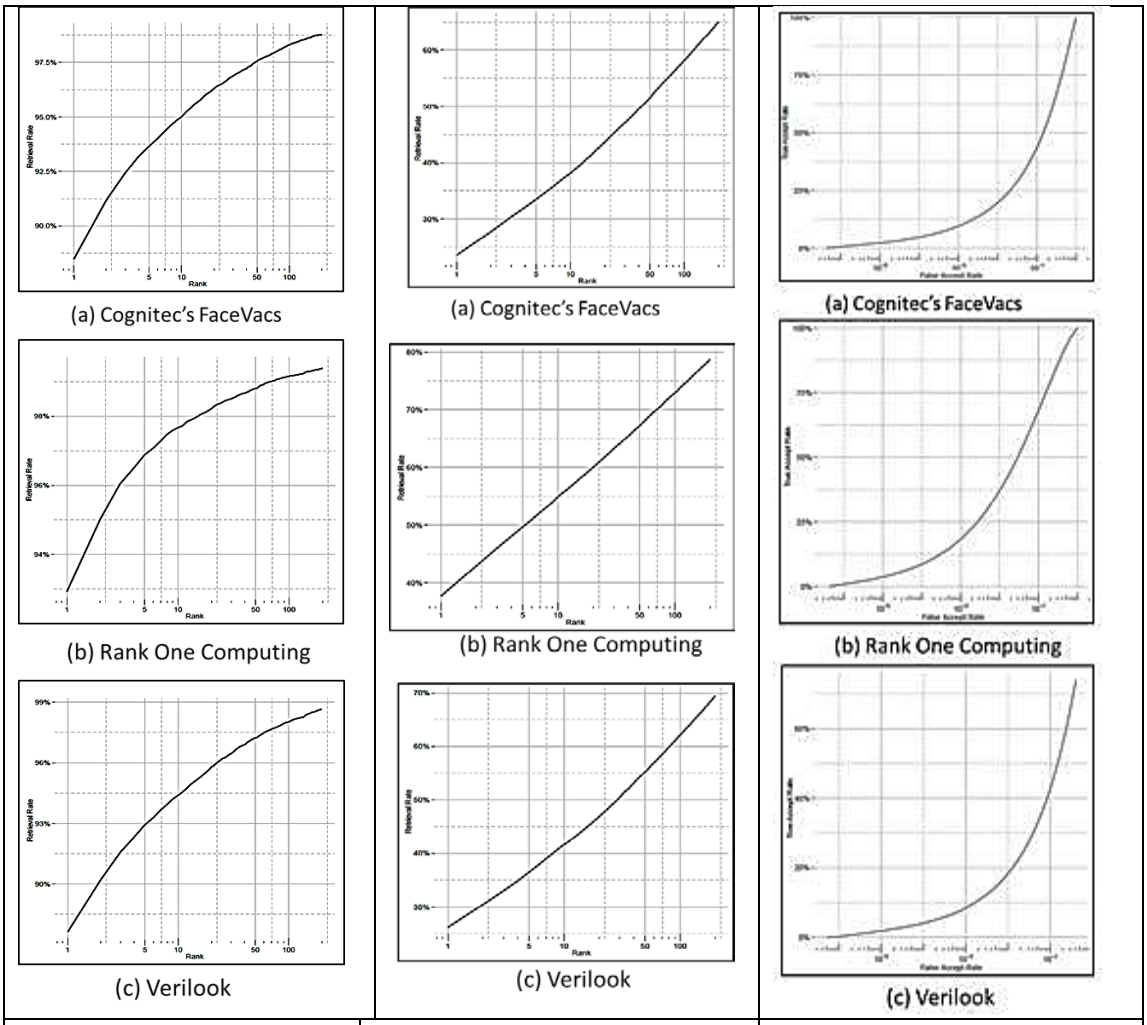
True Accept Rate			
FAR	**Cognitec v8.50**	**Verilook v6.0**	**Rank One Computing v1.20**
$1e^{-06}$	0.2%	0.1%	0.5%
$1e^{-05}$	1.8%	1.6%	4.5%
$1e^{-04}$	4.1%	3.7%	10.4%
$1e^{-03}$	9.1%	8.1%	21.9%
$1e^{-02}$	**19.9%**	**18.0%**	***43.0%***
$1e^{-01}$	45.2%	43.2%	74.5%

During the identification process, all images are scored and ranked in an ascending order based on their match score. Table 17 and figure 22 provides the rank retrieval rate for all the experiments conducted with the ITWCC-2 data set. It can be derived from table 16 data that it is much more challenging to identify a person from its temporal displaced photograph (younger or older) because there is a significant decrease in the rate of retrieving a legitimate user as rank-one in both Young-to-Old and Old-to-Young experiments.

Table 17: Rank-1 Retrieval Rate

SDK	**All-to All**	**Young-to-Old**	**Old-to-Young**
Cognitec v8.50	88.5%	20.5%	23.5%
Rank One Computing v1.20	***92.9%***	***34.6%***	***37.7%***
Verilook v6.0	87.6%	20.1%	26.2%

Moreover, it also shows that, Rank One Computing v1.20 not only performs better in the verification scenario but also outpaces other systems in identification setting.



All-to-All Verification– ITWCC-2	Old-to-Young Identification– ITWCC-2	Young-to-Old Identification– ITWCC-2

Figure 22: CMC plots

The different graphs, like receiver operating characteristic (ROC), detection error trade-off (DET), equal error rate (EER) and score histograms that are used to evaluate the performance of a facial recognition system are generated using the OpenBR Biometric 0.5.0 Toolkit [71] . The score histogram, DET curve and EER included in the study shows that Rank One Computing v1.20 [21] is not only better in terms of separating genuine and imposter (score histogram) but also it is efficient in classifying the user compared to Cognitec's FaceVacs v8.50 [20] and Verilook v6.0 trial [22].

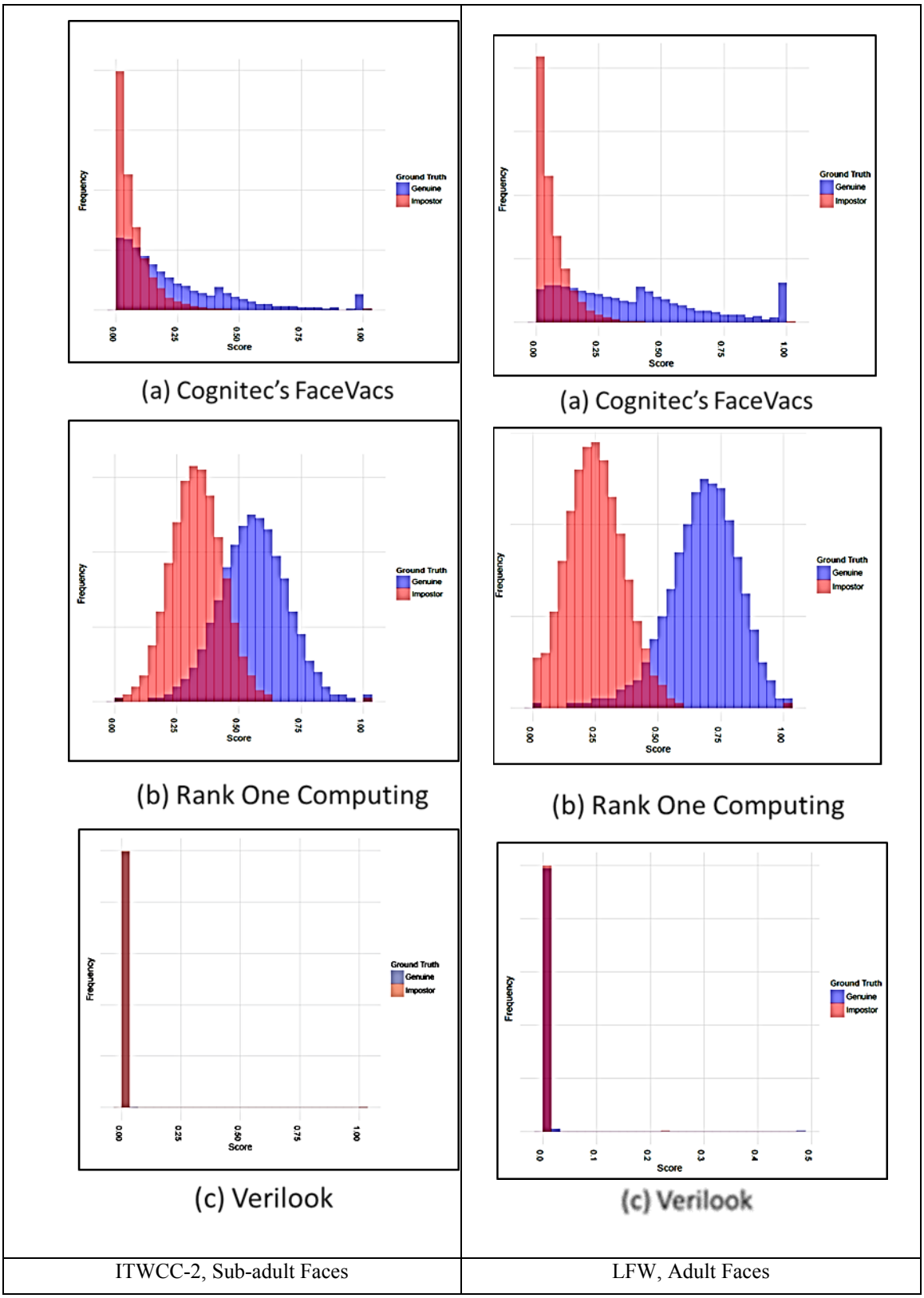
CHAPTER 6. CONCLUSIONS AND FUTURE WORK

This work presents the second dedicated research focused on child only longitudinal dataset; its objective is to determine if commercial face recognition systems developed with adult faces in mind can be used for child recognition. The results presented in the previous chapter clearly indicate that significant research into solving the problem of sub-adult face recognition is needed; the three commercial systems used in this work did not perform at the same rate on child face as they did on adult face. ROC outperformed both Cognitec and Neurotechnology handily. Further, this work provides deep insight into the growth and development phase and how this maturation process can denigrate face matchers. The experimental scenarios: Access control and Photo-tagging, were designed to explore the difficulty of sub-adult aging. Three major commercial systems were used to test the hypothesis, including Cognitec's FaceVacs v8.50 [20], Rank One Computing v1.20 [21], and Neurotechnology Verilook v6.0 [22].

Experiments concluded that Rank One Computing 1.20 recognition accuracy leaves behind both Cognitec's FaceVacs 8.50 and Neurotechnology Verilook 6.0 in both verification and identification scenarios. Results on the ITWCC-2 dataset shows that the recognition accuracy of facial recognition algorithms decreases with sub-adult subjects. The most accurate algorithm for the verification task has a true accept rate of only 46.8% at 1.0% false accept rate. However, in a similar experiment conducted on adult data set (LFW), a true accept rate of 87.6% is achieved at 1.0% false accept rate.

Also in the photo tagging scenario where the performance across the systems were tested for identification of a person from its temporally displaced image, the most accurate system had true accept rate of 37.1% for Young-to-Old verification and 43.0% for Old-to-Young verification at 1.0% false accept rate. A similar study conducted by Klare, et al.,

showed, a true accept rate of 72.4% at a fixed false accept rate of 1.0% in adults with 10 years of lapse [19]. Which demonstrates that the problem of longevity for adults has made significant advancements and that the solutions developed to tackle longitudinal displacement for adults is not sufficient for sub-adults (children). In addition, the rank-1 retrieval rate for the best performing system for identification of a person from its younger image is only 34.6%, whereas if person is identified from its older image, rank-1 retrieval rate is 37.7%. The score histogram as shown in figure 23, revealed that it is also difficult to separate genuine and imposter sub-adult users as compared to adult.



ITWCC-2, Sub-adult Faces	LFW, Adult Faces

Figure 23: Score Histogram for All-to-All Verification Experiment

The results conclude that sub-adult domain is a challenge for facial recognition systems. With the provided dataset, the researchers can start exploring the problem space and developing true face matchers for the case of children.

The future work includes the evaluation of sub-adult data in an identification scenario also known as Watch List. The watch list is a list of person of interests (POI), with individuals identified by biometric trait. It is an open-set recognition task. Open set corresponds to the watchlist, face recognition task, where the face recognition engine must detect or reject the probe, as it operates under the assumption that not all the probes have mates in the gallery. This is what makes it the most challenging among the biometric tasks and it requires the availability of a reject option. Watchlist design is very useful in flagging listed person while passive recognition. A shopping mall can have different list example VIPs, big spenders, undesirables, registered disabled or staff. Cameras always monitor the flow of people, and when a match is detected, a notification can be triggered by suitable follow-up actions. Suppose if a shoplifter is identified, the most senior security personnel could be notified for a follow-up action.

References

[1] D. Leopold and R. G., "Comparative View of Face Perception.," Journal of comparative psychology, pp. 233-251, 2010.

[2] "Face Recognition Homepage," [Online]. Available: http://www.face-rec.org/general-info/. [Accessed 17 12 2015].

[3] M. Petrov, Law Enforcement Applications of Forensic, Advanced Solutions, 2012.

[4] N. Subcommittee, "Introduction to Biometrics," 14 September 2006. [Online]. Available: http://www.biometrics.gov/ReferenceRoom/Introduction.aspx.

[5] O. C. King, "Casino ID- Facial Recognition Technology Used in Casinos," Online Casino King , [Online]. Available: http://www.onlinecasinoking.com/insights/casino-id-facial-recognition-technology-used-in-casinos/. [Accessed 17 12 2015].

[6] P. Rubens, "Facial recognition: Shop where everybody knows your name," BBC News Services, 9 December 2014. [Online]. Available: http://www.bbc.com/news/business-30219820. [Accessed 17 12 2015].

[7] K. Ricanek Jr and C. Boehnen, "Facial Analytics: From Big Data to Law Enforcement.," IEEE Computer Society Identity Sciences, pp. 95-97, 2012.

[8] M. Parkinson, "The power of visual communication," Billion Dollar Graphics, 2012.

[9] A. Jain and J. Huang, "Integrating independent components and linear discriminant analysis for gender classification," Sixth IEEE International Conference on Automatic Face and Gesture Recognition, p. 159 – 163, 2004.

[10] T. Target, "access control definition," Tech Target, June 2014. [Online]. Available: http://searchsecurity.techtarget.com/definition/access-control. [Accessed 17 12 2015].

[11] E. Kemal, Hazim, J. Stallkamp, H. Gao, M. Fischer and R. and Stiefelhagen, "Face recognition for smart interactions.," IEEE International Conference on InMultimedia and Expo, pp. 1007-1010, 2007.

[12] S. Johannes, H. K. Ekenel and R. Stiefelhagen, "Video-based Face Recognition on Real-World Data," IEEE on Computer Vision (ICCV'07), pp. 1-8, 2007.

[13] H. Ekenel, F. M. and S. R., "Face Recognition in Smart Rooms," 4th MLMI, Brno, Czech Republic, 2007.

[14] E. H.K., J. Q and F. M, "ISL Person Identification Systems in CLEAR 2007," CLEAR Evaluation Workshop, Baltimore, US, 2007.

[15] R. Stiefelhagen, K. Bernardin, H. K. Ekenel, M. J., K. Nickel, M. Voit and M. Wölfel, "Audio-visual perception of a lecturer in a smart seminar room," Signal Processing 86, 2006.

[16] E. Kemal, Hazim and A. Pnevmatikakis, "Video-based face recognition evaluation in the chil project-run 1," 7th International Conference on Automatic Face and Gesture Recognition, p. 6, 2006.

[17] M. K. Sodomsky, "An Evaluation of Longitudnal Face Recognition Performance Throughout the Growth and Development Stages," University of North Carolina Wilmington, Wilmington, 2014.

[18] T. Yaniv, M. Yang, M. Ranzato and L. Wolf, "Deepface: Closing the gap to human-level performance in face verification," 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1701-1708, 2014.

[19] B. Klare and A. K. Jain, "Face recognition across time lapse: On learning feature subspaces.," 2011 International Joint Conference on Biometrics (IJCB) , pp. 1-8, 2011.

[20] C. S. GmbH, Facevacs software developer kit, 2014.

[21] "ROC SDK," Rank One Computing, [Online]. Available: http://www.rankone.io/sdk/. [Accessed 17 12 2015].

[22] "Verilook SDK," Neurotechnology, [Online]. Available: http://www.neurotechnology.com/verilook.html. [Accessed 17 12 2015].

[23] N. S. &. T. Council, "Glossary," International Biometrics & Identification Association (IBIA), [Online]. Available: https://www.ibia.org/biometrics/glossary/.

[24] K. Ricanek and B. Barbour, "What Are Soft Biometrics and How Can They Be Used?," IEEE Computer, vol. 44, no. 9, pp. 106-108, 2011.

[25] P. J. Phillips, W. T. Scruggs, A. J. O'Toole, P. J. Flynn, K. W. Bowyer, C. L. Schott and M. Sharpe, "FRVT 2006 and ICE 2006 large-scale results.," National Institute of Standards and Technology, NISTIR, 2007.

[26] P. J. Phillips, H. Moon, S. Rizvi and P. J. Rauss, "The FERET evaluation methodology for face-recognition algorithms.," Pattern Analysis and Machine Intelligence, IEEE Transactions on , vol. 22, no. 10, pp. 1090-1104., 2000.

[27] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min and W. Worek, "Overview of the face recognition grand challenge," in Computer vision and pattern recognition. CVPR 2005. , 2005.

[28] P. J. Phillips, P. J. Flynn, J. R. Beveridge, W. T. Scruggs, A. J. O'toole, D. Bolme and B. K. W., "Overview of the multiple biometrics grand challenge.," in In Advances in Biometrics, , 2009.

[29] P. J. Grother, G. W. Quinn and P. J. Phillips, "Report on the evaluation of 2D still-image face recognition algorithms.," NIST interagency report, 2010.

[30] M. Turk and A. Pentland, "Eigenfaces for recognition," Journal of cognitive neuroscience, vol. 3, no. 1, pp. 71-86., 1991.

[31] "Biometric Data Interchange Formats Part 5: Face Image Data," ISO/EC 19794-5, 2004.

[32] P. J. Phillips, J. R. Beveridge, B. A. Draper, G. Givens, A. J. O'Toole, D. S. Bolme, J. Dunlop, Y. M. Lui, H. Sahibzada and S. Weimer, "An introduction to the good, the bad, & the ugly face recognition challenge problem.," in In Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference, 2011.

[33] T. Wu, P. Turaga and R. Chellappa, "Age estimation and face verification across aging using landmarks.," Information Forensics and Security, vol. 7, no. 6, pp. 1780-1788, 2012.

[34] Y. M. Lui, D. Bolme, P. J. Phillips, J. R. Beveridge and B. A. Draper, "Preliminary studies on the good, the bad, and the ugly face recognition challenge problem.," in In Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference, 2012.

[35] R. Singh, M. Vatsa, H. S. Bhatt, S. Bharadwaj, A. Noore and S. S. Nooreyezdan, "Plastic surgery: A new dimension to face recognition.," IEEE Transactions on Information Forensics and Security, vol. 5, no. 3, pp. 441-448., 2010.

[36] M. G. and Ricanek., "Investigating the effects of gender and age group based differences in identical twins," Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), pp. 1 -4, 2013.

[37] B. F. Klare, Z. Li and A. K. Jain, "Matching forensic sketches to mug shot photos," IEEE Transactions on Pattern Analysis and Machine Intelligence, , vol. 33, no. 3, pp. 639-646., 2011.

[38] S. Klum, H. Han, A. K. Jain and B. Klare, "Sketch based face recognition: Forensic vs. composite sketches.," International Conference on Biometrics (ICB), pp. 1-8, 2013.

[39] N. Ramanathan and R. Chellappa, ""Face verification across age progression." .," IEEE Transactions on Image Processing, , vol. 15, no. 11, pp. 3349-3361, 2006.

[40] G. Guodong, G. Mu and K. Karl Ricanek, "Cross-age face recognition on a very large database: the performance versus age intervals and improvement using soft biometric traits."," in 20th International Conference on Pattern Recognition (ICPR),, 2010.

[41] U. Park, Y. Tong and A. K. Jain, "Age-invariant face recognition.," Pattern Analysis and Machine Intelligence, IEEE Transactions on 32, pp. 947-954, 2010.

[42] R. Singh, M. Vatsa and A. Noore, " Face recognition with disguise and single gallery images.," Image and Vision Computing , vol. 27, no. 3, pp. 245-257, 2009.

[43] P. Grother and N. M., "Face recognition vendor test (frvt) performance of face identification algorithms.," NIST Interagency Report 8009 , 2014.

[44] K. Ricanek, S. Bhardwaj and S. M., "A Review of Face Recognition Against Longitudinal Child Faces," in 14th International Conference of the Biometrics Special Interest Group (BioSig 2015), Darmstadt, Germany, 2015.

[45] R. Stiefelhagen, H. K. Ekenel, C. Fügen, P. Gieselmann, H. Holzapfel, F. Kraft, K. Nickel, M. Voit and A. Waibel, "Enabling Multimodal Human-Robot Interaction fort the Karlsruhe Humanoid Robot," IEEE Trans. on Robotics, vol. 23, no. 5, 2007.

[46] K. J., B. J., C. V, K. V. and V. J, "Development of facial sexual dimorphism in children aged between 12 and 15 years: a three-dimensional longitudinal study."," Orthodontics & craniofacial research ., 2015.

[47] L. Mark, J. Pittenger and H. Hines, "Wrinkling and head shape as coordinated sources of age-Ievel information," Perception and Psychophysics, vol. 27, pp. 117-124, 1980.

[48] K. Ricanek, A. Sethuram, E. K. Patterson, A. M. Albert and E. J. Boone, "Craniofacial aging," Wiley Handbook of Science and Technology for Homeland Security, 2009.

[49] M. Bastir, A. Rosas and O. Paul, "Craniofacial levels and the morphological maturation of the human skull.," Journal of Anatomy , pp. 637-654, 2006.

[50] L. G. Farkas, J. C. Posnick and T. M. Hreczko, "Growth patterns of the face: a morphometric study.," The Cleft Palate-Craniofacial Journal , pp. 308-315, 1992.

[51] D. F. Huelke, "An overview of anatomical considerations of infants and children in the adult world of automobile safety design.," Annual Proceedings/Association for the Advancement of Automotive Medicine, vol. 42, 1998.

[52] D. S. Gill and F. B. Naini, "An Introduction to Human Craniofacial Growth and Development.," Orthodontics: Principles and Practice, pp. 1-16.

[53] C. W. Cummings, B. H. Haughey, J. R. Thomas, L. A. Harker, K. T. Robbins, D. E. Schuller, P. W. Flint and M. A. Richardson, Cummings Otolaryngology-Head and Neck Surgery, 2005.

[54] M. Hellman, "An introduction to growth of the human face from infancy to adulthood.," International Journal of Orthodontia, Oral Surgery and Radiography , pp. 777-798, 1932.

[55] K. T. Taylor, Forensic art and illustration., CRC Press, 2000.

[56] A. A. Midori, K. Ricanek and E. Patterson, "A review of the literature on the aging adult skull and face: Implications for forensic science research and applications.," Forensic Science International 172, pp. 1-9, 2007.

[57] Y. Fu, G. Guo and T. S. Huang, "Age synthesis and estimation via faces: A survey.," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 11, pp. 1955-1976, 2010.

[58] G. B. Huang, M. Ramesh, T. Berg and E. Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," Technical Report 07-49, University of Massachusetts, Amherst, 2007.

[59] M. Grgic, K. Delac and S. Grgic, "SCface–surveillance cameras face database.," Multimedia tools and applications , vol. 51, no. 3, pp. 863-879, 2011.

[60] L. Wolf, T. Hassner and I. Maoz, "Face recognition in unconstrained videos with matched background similarity.," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 529-534, 2011.

[61] C. Chen, A. Dantcheva and A. Ross, "Automatic facial makeup detection with application in face recognition.," International Conference on Biometrics (ICB), pp. 1-8, 2013.

[62] H. Lamba, A. Sarkar, M. Vatsa, R. Singh and A. Noore, "Face recognition for look-alikes: A preliminary study.," International Joint Conference on Biometrics (IJCB) , pp. 1-6, 2011.

[63] A. M. Albert and J. Karl Ricanek, "The MORPH Database: Investigating the Effects of Adult Craniofacial Aging on Automated Face-Recognition Technology," Forensic Science Communications, vol. 10, no. 2, 2008.

[64] F.-N. Consortium, "The FG-NET aging database," 2007-11-12. http://sting. cycollege. ac. cy/-alanitis/fgnetaging/index. htm, 2012.

[65] G. Somanath, M. Rohith and C. Kambhamettu, "VADANA: A dense dataset for facial image analysis," Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference , pp. 2175 - 2182 , 2011.

[66] E. Eidinger, R. Enbar and T. Tal Hassner, "Age and Gender Estimation of Unfiltered Faces," Transactions on Information Forensics and Security (IEEE-TIFS), special issue on Facial Biometrics in the Wild, vol. 9, no. 12, pp. 2170 - 2179, 2014.

[67] B.-C. Chen, C.-S. Chen and W. H. Hsu, "Cross-age reference coding for age-invariant face recognition and retrieval.," Computer Vision–ECCV, pp. 768-783, 2014.

[68] A. A. Ross, K. Nandakumar and A. K. Jain, Handbook of multibiometrics, Springer Science & Business Media, 2006.

[69] C. Systems, "FaceVACS Technology," Cognitec, [Online]. Available: http://www.cognitec.com/technology.html. [Accessed 17 12 2015].

[70] P. Grother, G. W. Quinn and N. M., "FRVT Still Face Image and Video Concept, Evaluation Plan and API Version 1.4," NIST, 2013.

[71] J. C. Klontz, B. F. Klare, S. Klum, A. K. Jain and M. J. Burge, "Open source biometric recognition," IEEE Sixth International Conference on Biometrics: Theory Applications and Systems (BTAS), pp. 1-8, 2013.

[72] D. E. King, " Dlib-ml: A Machine Learning Toolkit.," Journal of Machine Learning Research , pp. 1755-1758, 2009 .

[73] J. L. Wayman, "Fundamentals of biometric authentication technologies.," International Journal of Image and Graphics , vol. 1, pp. 93-113, 2001.

[74] B. Ulery, W. Fellner, P. Hallinan, A. Hicklin and C. Watson, "Modeling Biometric Score Distributions," NIST, 2006.

[75] P. Wang, M. B. Green, Q. Ji and J. Wayman, "Automatic eye detection and its validation.," Computer Vision and Pattern Recognition-Workshops (CVPR) , 2005.

[76] G. P. A and H. G, "Factors Influencing the Accuracy of age Estimates of unfamiliar Faces," Perception, vol. 24, pp. 1059-1073, 1995.

[77] P. N. Belhumeur, J. P. Hespanha and D. J. Kriegman., "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection.," Pattern Analysis and Machine Intelligence, IEEE Transactions, vol. 19, no. 7, p. 711–720, 1997.

[78] M. Turk and A. Pentland., "Face recognition using eigenfaces," in In Computer Vision and Pattern Recognition, 1991.

[79] G. Bradski, "The OpenCV Library," Dr. Dobb's Journal of Software Tools, 2000.

[80] A. K. Jain, B. Klare and U. Park, " Face recognition: Some challenges in forensics," IEEE International Conference on Automatic Face and Gesture Recognition, pp. 726-733, 2011.

[81] "Latchkey Kid," Wikipedia, 9 December 2015. [Online]. Available: https://en.wikipedia.org/wiki/Latchkey_kid. [Accessed 17 12 2015].

[82] M. Tistarelli, S. E. Barrett and A. J. O'Toole, "Facial Recognition, Facial Expression and Intention Detection.," In Second Generation Biometrics: The Ethical, Legal and Social Context, pp. 229-255, 2012.

[83] A. J. Toole, P. J. Phillips and A. Narvekar, "Humans versus algorithms: Comparisons from the face recognition vendor test 2006.," in 8th IEEE International Conference on Automatic Face & Gesture Recognition, 2008. FG'08. , 2008.

[84] A. Adler and M. E. Schuckers, "Comparing human and automatic face recognition performance," in Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on, 2007.

[85] K. Ricanek and T. Tesafaye, "Morph: A longitudinal image database of normal adult age-progression," Automatic Face and Gesture Recognition, 2006. FGR 2006. 7th International Conference, pp. 341--345}, 2006.

[86] "An interview with Mark G. Hans," Dental Press Journal of Orthodontics, [Online]. Available: http://www.scielo.br/scielo.php?pid=S2176-94512014000300026&script=sci_arttext. [Accessed 17 12 2015].

[87] A. K. Jain, S. C. Dass and K. Nandakumar, "Soft biometric traits for personal recognition systems.," in Biometric Authentication, Springer Berlin Heidelberg, 2004.

[88] A. K. Jain, S. C. Dass and K. Nandakumar, "Can soft biometric traits assist user recognition?,," in International Society for Optics and Photonics In Defense and Security, 2004.

[89] "ISO/IEC 19794-5," Correlance, [Online]. Available: http://www.correlance.com/cms/en/iso19794-5. [Accessed 2 1 2016].

APPENDIX A

Biometric Glossary

BIOMETRICS GLOSSARY

The set of terms included in this section was developed by the National Science & Technology Council's (NSTC) [4]. The statements here are intended to further the understanding of a general audience and are not intended to replace or compete with sources that may be more technically descriptive/prescriptive.

Algorithm

A limited sequence of instructions or steps that tells a computer system, how to solve a particular problem. A biometric system will have multiple algorithms, for example: image processing, template generation, comparisons, etc.

Application Programming Interface (API)

Formatting instructions or tools used by an application developer to link and build hardware or software applications.

Authentication

- The process of establishing confidence in the truth of some claim. The claim could be any declarative statement, for example: "This individual's name is 'Joseph K.' " or "This child is more than 5 feet tall."
- In biometrics, "authentication" is sometimes used as a generic synonym for verification.

Automated Biometric Identification System (ABIS)

Generic term sometimes used in the biometrics community to discuss a biometric system.

Behavioral Biometric Characteristic

A biometric characteristic that is learned and acquired over time rather than one based primarily on biology. All biometric characteristics depend somewhat upon both

behavioral and biological characteristics. Examples of biometric modalities for which behavioral characteristics may dominate include signature recognition and keystroke dynamics.

Benchmarking

The process of comparing measured performance against a standard, openly available, reference.

Biological Biometric Characteristic

A biometric characteristic based primarily on an anatomical or physiological characteristic, rather than a learned behavior. All biometric characteristics depend somewhat upon both behavioral and biological characteristics. Examples of biometric modalities for which biological characteristics may dominate include fingerprint and hand geometry.

Biometrics

A general term used alternatively to describe a characteristic or a process.

- As a characteristic: A measurable biological (anatomical and physiological) and behavioral characteristic that can be used for automated recognition.
- As a process: Automated methods of recognizing an individual based on measurable biological (anatomical and physiological) and behavioral characteristics.

Biometric Data

A catch-all phrase for computer data created during a biometric process. It encompasses raw sensor observations, biometric samples, models, templates and/or similarity scores. Biometric data is used to describe the information collected during an

enrollment, verification, or identification process, but does not apply to end user information such as user name, demographic information and authorizations.

Biometric Sample

Information or computer data obtained from a biometric sensor device. Examples are images of a face or fingerprint.

Biometric System

Multiple individual components (such as sensor, matching algorithm, and result display) that combine to make a fully operational system. A biometric system is an automated system capable of:

1. Capturing a biometric sample from an end user.
2. Extracting and processing the biometric data from that sample.
3. Storing the extracted information in a database.
4. Comparing the biometric data with data contained in one or more reference.
5. Deciding how well they match and indicating whether an identification or verification of identity has been achieved.

A biometric system may be a component of a larger system.

Capture

The process of collecting a biometric sample from an individual via a sensor. See also submission.

Challenge Response

A method used to confirm the presence of a person by eliciting direct responses from the individual. Responses can be either voluntary or involuntary. In a voluntary response, the end user will consciously react to something that the system presents. In an

involuntary response, the end user's body automatically responds to a stimulus. A challenge response can be used to protect the system against attacks. See also liveness detection.

Claim of identity

A statement that a person is or is not the source of a reference in a database. Claims can be positive (I am in the database), negative (I am not in the database) or specific (I am end user 123 in the database).

Closed-set Identification

A biometric task where an unidentified individual is known to be in the database and the system attempts to determine his/her identity. Performance is measured by the frequency with which the individual appears in the system's top rank (or top 5, 10, etc.).

Comparison

The process of comparing a biometric reference to a previously stored reference or references in order to make an identification or verification decision.

Cooperative User

An individual that willingly provides his/her biometric to the biometric system for capture. Example: A worker submits his/her biometric to clock in and out of work.

Covert

An instance in which biometric samples are being collected at a location that is not known to bystanders. An example of a covert environment might involve an airport checkpoint where face images of passengers are captured and compared to a watch list without their knowledge.

Cumulative Match Characteristic (CMC)

A method of showing measured accuracy performance of a biometric system operating in the closed-set identification task Templates are compared and ranked based on their similarity.

The CMC shows how often the individuals' template appears in the ranks (1, 5, 10, 100, etc.), based on the match rate. A CMC compares the rank (1, 5, 10, 100, etc.) versus identification rate as illustrated below.

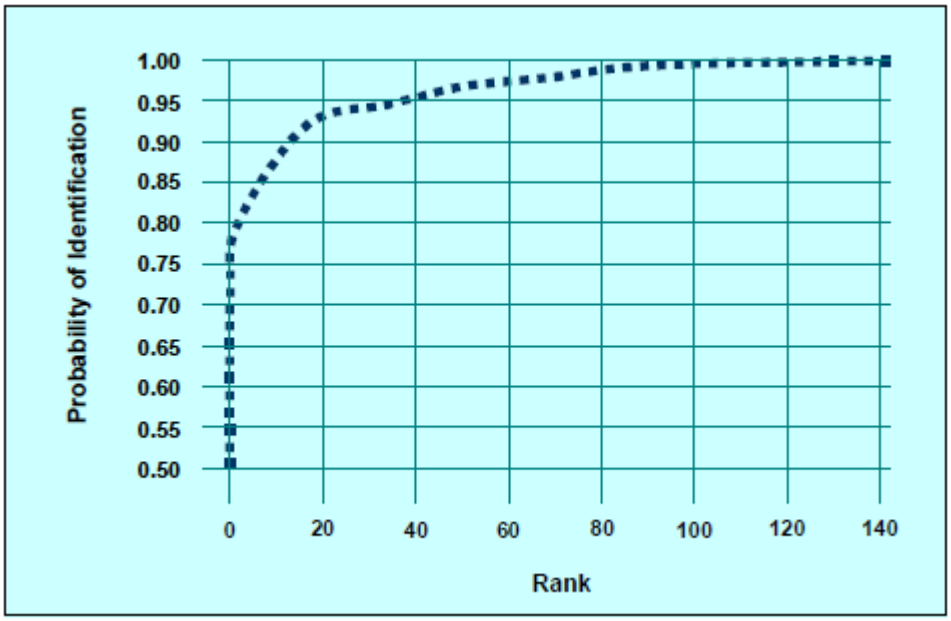


Figure 24: Cumulative Match Characteristics

Database

A collection of one or more computer files. For biometric systems, these files could consist of biometric sensor readings, templates, match results, related end user information, etc.

Decision

The resultant action taken (either automated or manual) based on a comparison of a similarity score (or similar measure) and the system's threshold.

Detection and Identification Rate

The rate at which individuals, who are in a database, are properly identified in an open-set identification (watch-list) application.

Detection Error Trade-off (DET) Curve

A graphical plot of measured error rates, as illustrated below. DET curves typically plot matching error rates (false non-match rate vs. false match rate) or decision error rates (false reject rate vs. false accept rate).

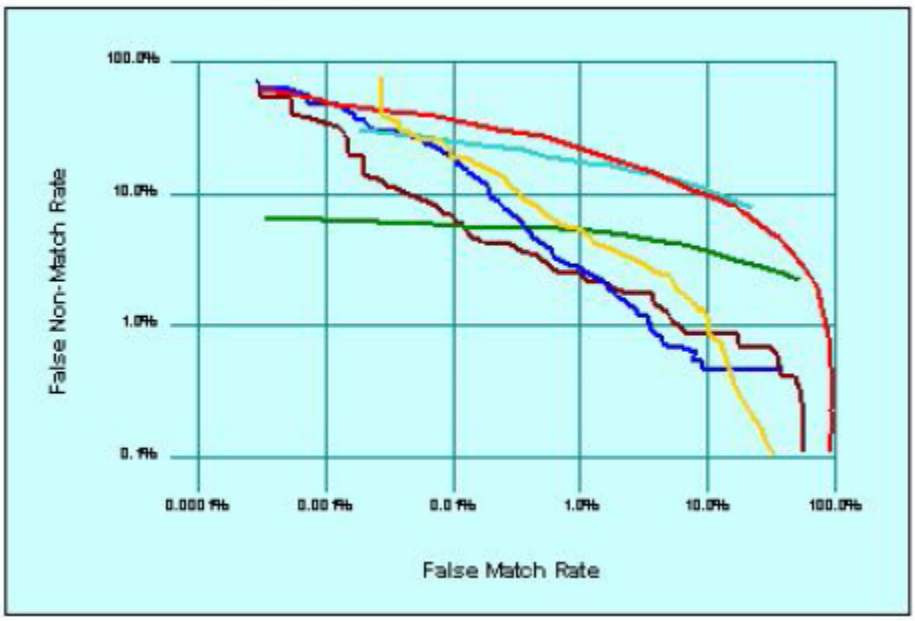


Figure 25: Detection Error Trade-off Curve

Difference Score

A value returned by a biometric algorithm that indicates the degree of difference between a biometric sample and a reference.

End User

The individual who will interact with the system to enroll, to verify, or to identify. See also cooperative user, indifferent user, non-cooperative user, uncooperative user, and user.

Enrollment

The process of collecting a biometric sample from an end user, converting it into a biometric reference, and storing it in the biometric system's database for later comparison.

Equal Error Rate (EER)

A statistic used to show biometric performance, typically when operating in the verification task. The EER is the location on a ROC or DET curve where the false accept rate and false reject rate (or one minus the verification rate {1-VR}) are equal, as illustrated below. In general, the lower the equal error rate value, the higher the accuracy of the biometric system. Note, however, that most operating systems are not set to operate at the "equal error rate" so the measure's true usefulness is limited to comparing biometric system performance. The EER is sometimes referred to as the "Crossover Error Rate."

Extraction

The process of converting a captured biometric sample into biometric data so that it can be compared to a reference.

Face Recognition

A biometric modality that uses an image of the visible physical structure of an individual's face for recognition purposes.

Failure to Acquire (FTA)

Failure of a biometric system to capture and/or extract usable information from a biometric sample.

Failure to Enroll (FTE)

Failure of a biometric system to form a proper enrollment reference for an end user. Common failures include end users who are not properly trained to provide their

biometrics, the sensor not capturing information correctly, or captured sensor data of insufficient quality to develop a template.

False Acceptance Rate (FAR)

A statistic used to measure biometric performance when operating in the verification task. The percentage of times a system produces a false accept, which occurs when an individual is incorrectly matched to another individual's existing biometric. Example: Frank claims to be John and the system verifies the claim.

False Alarm Rate

A statistic used to measure biometric performance when operating in the open-set identification (sometimes referred to as watch-list) task. This is the percentage of times an alarm is incorrectly sounded on an individual who is not in the biometric system's database (the system alarms on Frank when Frank isn't in the database), or an alarm is sounded but the wrong person is identified (the system alarms on John when John is in the database, but the system thinks John is Steve).

False Match Rate

A statistic used to measure biometric performance when. Similar to the False Acceptance Rate (FAR).

False Non-Match Rate

A statistic used to measure biometric performance. Similar to the False Reject Rate (FRR), except the FRR includes the Failure to acquire error rate and the False Non-Match Rate does not.

False Rejection Rate (FRR)

A statistic used to measure biometric performance when operating in the verification task. The percentage of times the system produces a false reject. A false reject

occurs when an individual is not matched to his/her own existing biometric template. Example: John claims to be John, but the system incorrectly denies the claim.

Feature(s)

Distinctive mathematical characteristic(s) derived from a biometric sample; used to generate a reference.

Feature Extraction

See extraction.

FERET - FacE REcognition Technology program

A face recognition development and evaluation program sponsored by the U.S. Government from 1993 through 1997. For more information, visit www.frvt.org/FERET/default.htm.

FRVT - Face Recognition Vendor Test

A series of large-scale independent technology evaluations of face recognition systems. The evaluations have occurred in 2000, 2002, and 2005. For more information, visit www.frvt.org/FRVT2005/default.aspx.

Gallery

The biometric system's database or set of known individuals, for a specific implementation or evaluation experiment.

Identification

A task where the biometric system searches a database for a reference matching a submitted biometric sample and if found, returns a corresponding identity. A biometric is collected and compared to all the references in a database. Identification is "closed-set" if the person is known to exist in the database. In "open-set" identification, sometimes

referred to as a "watch-list," the person is not guaranteed to exist in the database. The system must determine whether the person is in the database, then return the identity.

Identification Rate

The rate at which an individual in a database is correctly identified.

Identity Governance

The combination of policies and actions taken to ensure enterprise-wide consistency, privacy protection and appropriate interoperability between individual identity management systems.

Identity Management

The combination of systems, rules and procedures that defines an agreement between an individual and organization(s) regarding ownership, utilization and safeguard of personal identity information.

Impostor

A person who submits a biometric sample in either an intentional or an inadvertent attempt to claim the identity of another person to a biometric system. See also attempt.

Indifferent User

An individual who knows his/her biometric sample is being collected and does not attempt to help or hinder the collection of the sample. For example, an individual, aware that a camera is being used for face recognition, looks in the general direction of the sensor, neither avoiding nor directly looking at it.

ISO - International Organization for Standardization

A non-governmental network of the national standards institutes from 151 countries. The ISO acts as a bridging organization in which a consensus can be reached on solutions that meet both the requirements of business and the broader needs of society,

such as the needs of stakeholder groups like consumers and users. For more information, visit www.iso.org.

Liveness Detection

A technique used to ensure that the biometric sample is submitted from an end user. A liveness detection method can help protect the system against some types of spoofing attacks.

Match

A decision that a biometric sample and a stored template comes from the same human source, based on their high level of similarity (difference or hamming distance).

Matching

The process of comparing a biometric sample against a previously stored template and scoring the level of similarity (difference or hamming distance). Systems then make decisions based on this score and its relationship (above or below) a predetermined threshold.

Modality

A type or class of biometric system. For example: face recognition, fingerprint recognition, iris recognition, etc.

Model

A representation used to characterize an individual. Behavioral based biometric systems, because of the inherently dynamic characteristics, use models rather than static templates.

Multimodal Biometric System

A biometric system in which two or more of the modality components (biometric characteristic, sensor type or feature extraction algorithm) occurs in multiple.

NIST - National Institute of Standards and Technology

A non-regulatory federal agency within the U.S. Department of Commerce that develops and promotes measurement, standards, and technology to enhance productivity, facilitate trade, and improve the quality of life. NIST's measurement and standards work promoting the well-being of the nation and helps improve, among many other things, the nation's homeland security. For more information, visit www.nist.gov.

Noise

Unwanted components in a signal that degrade the quality of data or interfere with the desired signals processed by a system.

Non-cooperative User

An individual who is not aware that his/her biometric sample is being collected. Example: A traveler passing through a security line at an airport is unaware that a camera is capturing his/her face image. See also cooperative user, indifferent user, and uncooperative user.

One-to-many

A phrase used in the biometrics community to describe a system that compares one reference to many enrolled references to make a decision. The phrase typically refers to the identification or watch list tasks.

One-to-one

A phrase used in the biometrics community to describe a system that compares one reference to one enrolled reference to make a decision. The phrase typically refers to the verification task (though not all verification tasks are truly one-to-one) and the identification task can be accomplished by a series of one-to-one comparisons.

Open-set Identification

Biometric task that more closely follows operational biometric system conditions to 1) determine if someone is in a database and 2) find the record of the individual in the database. This is sometimes referred to as the "watch list" task to differentiate it from the more commonly referenced closed-set identification.

Operational Evaluation

One of the three types of performance evaluations. The primary goal of an operational evaluation is to determine the workflow impact seen by the addition of a biometric system.

Overt

Biometric sample collection where end users know they are being collected and at what location. An example of an overt environment is the US-VISIT program where non-U.S. citizens entering the United States submit their fingerprint data.

Performance

A catch-all phrase for describing a measurement of the characteristics, such as accuracy or speed, of a biometric algorithm or system. See also accuracy, crossover error rate, cumulative match characteristics, d-prime, detection error tradeoff, equal error rate, false accept rate, false alarm rate, false match rate, false reject rate, identification rate, operational evaluation, receiver operating characteristics, scenario evaluation, technology evaluation, true accept rate, true reject rate, verification rate.

Pixel

A picture element. This is the smallest element of a display that can be assigned a color value. See also pixels per inch (PPI), resolution.

Pixels-Per-Inch (PPI)

A measure of the resolution of a digital image. The higher the PPI, the more information is included in the image, and the larger the file size.

Population

The set of potential end users for an application.

Probe

The biometric sample that is submitted to the biometric system to compare against one or more references in the gallery.

Receiver Operating Characteristics (ROC)

A method of showing measured accuracy performance of a biometric system. A verification ROC compares false accept rate vs. verification rate. An open-set identification (watch-list) ROC compares false alarm rates vs. detection and identification rate.

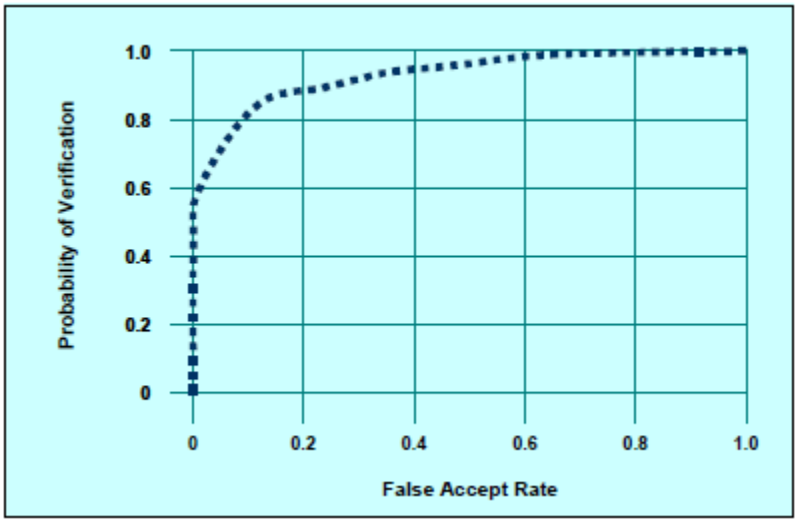


Figure 26: Receiver Operating Characteristics

Recognition

A generic term used in the description of biometric systems (e.g. face recognition or iris recognition) relating to their fundamental function. The term "recognition" does not

inherently imply verification, closed-set identification or open-set identification (watch-list).

Record

The template and other information about the end user (e.g. name, access permissions).

Reference

The biometric data stored for an individual for use in future recognition. A reference can be one or more templates, models or raw images.

Resolution

The number of pixels per unit distance in the image. Describes the sharpness and clarity of an image. See also pixel, pixels per inch (PPI).

Scenario Evaluation

One of the three types of performance evaluations. The primary goal of a scenario evaluation is to measure performance of a biometric system operating in a specific application.

Sensor

Hardware found on a biometric device that converts biometric input into a digital signal and conveys this information to the processing device.

Sensor Aging

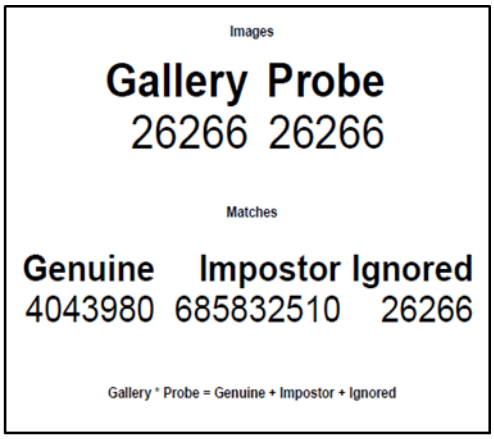
The gradual degradation in performance of a sensor over time.

Similarity Score

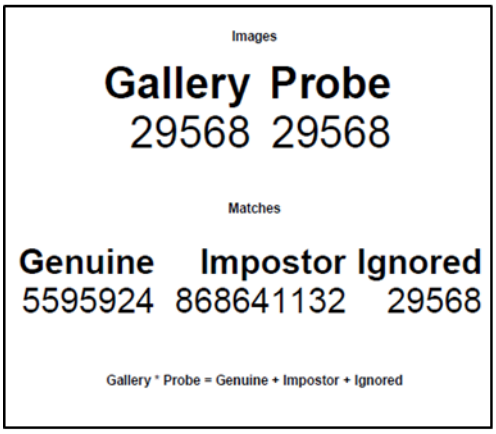
A value returned by a biometric algorithm that indicates the degree of similarity or correlation between a biometric sample and a reference. See also difference score, hamming distance.

APPENDIX B

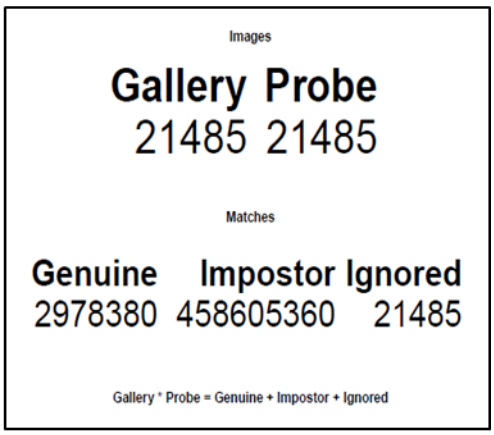
All-to-All Verification Experiment Results



(a) Cognitec's FaceVacs



(b) Rank One Computing



(c) Verilook

Figure 27: All-to-All Verification-TWCC-2-Data and Match Metrics

Table of True Accept Rates at various False Accept Rates

	eval
FAR = 1e−06	0.018
FAR = 1e−05	0.045
FAR = 1e−04	0.075
FAR = 1e−03	0.131
FAR = 1e−02	0.247
FAR = 1e−01	0.498

Table of retrieval rate at various ranks

	eval
Rank 1	0.885
Rank 5	0.937
Rank 10	0.95
Rank 20	0.963
Rank 50	0.976
Rank 100	0.983

(a) Cognitec's FaceVacs

Table of True Accept Rates at various False Accept Rates

	eval
FAR = 1e−06	0.024
FAR = 1e−05	0.064
FAR = 1e−04	0.13
FAR = 1e−03	0.254
FAR = 1e−02	0.468
FAR = 1e−01	0.769

Table of retrieval rate at various ranks

	eval
Rank 1	0.929
Rank 5	0.969
Rank 10	0.977
Rank 20	0.982
Rank 50	0.988
Rank 100	0.992

(b) Rank One Computing

Table of True Accept Rates at various False Accept Rates

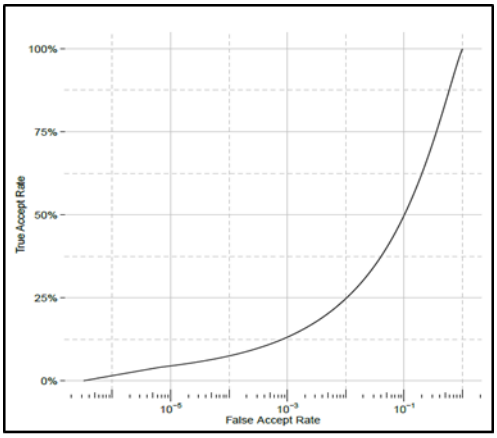
	eval
FAR = 1e−06	0.018
FAR = 1e−05	0.041
FAR = 1e−04	0.07
FAR = 1e−03	0.123
FAR = 1e−02	0.237
FAR = 1e−01	0.492

Table of retrieval rate at various ranks

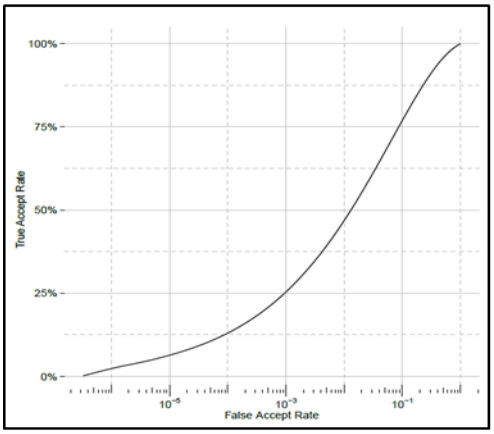
	eval
Rank 1	0.876
Rank 5	0.929
Rank 10	0.944
Rank 20	0.958
Rank 50	0.972
Rank 100	0.98

(c) Verilook

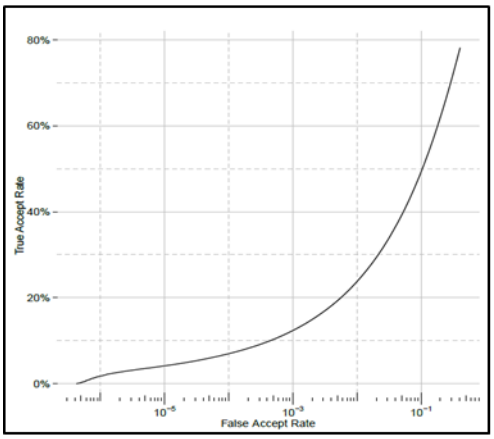
Figure 28: All-to-All Verification – ITWCC-2 TAR and Rank Values



(a) Cognitec's FaceVacs

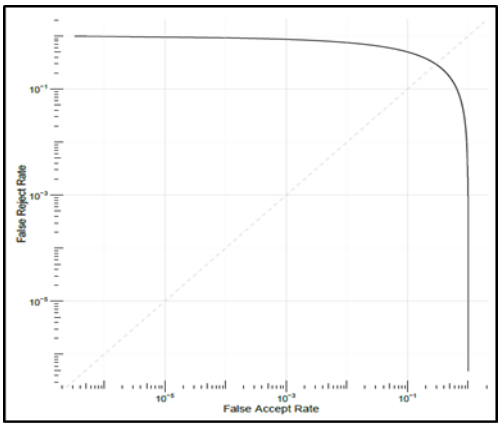


(b) Rank One Computing

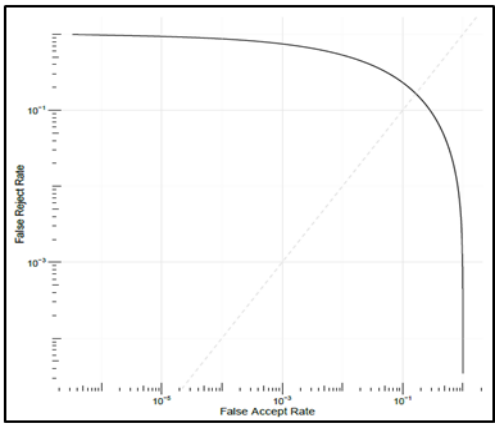


(c) Verilook

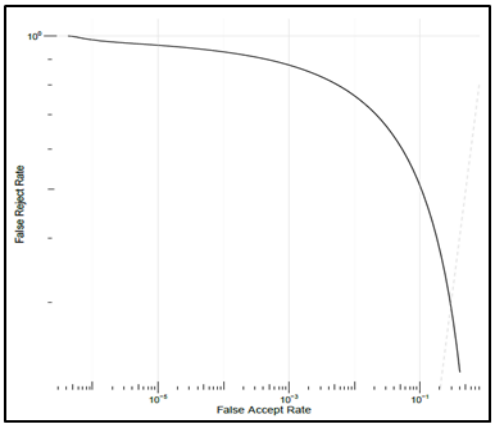
Figure 29: All-to-All Verification-ITWCC-2-ROC



(a) Cognitec's FaceVacs

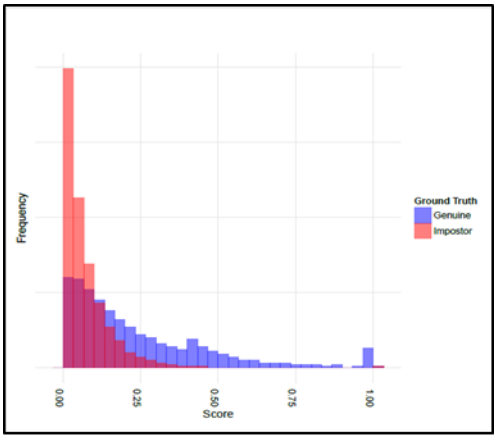


(b) Rank One Computing

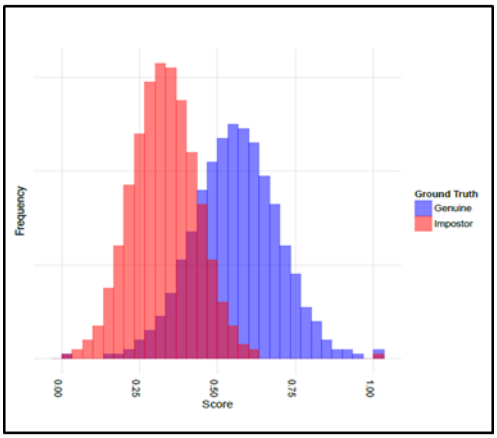


(c) Verilook

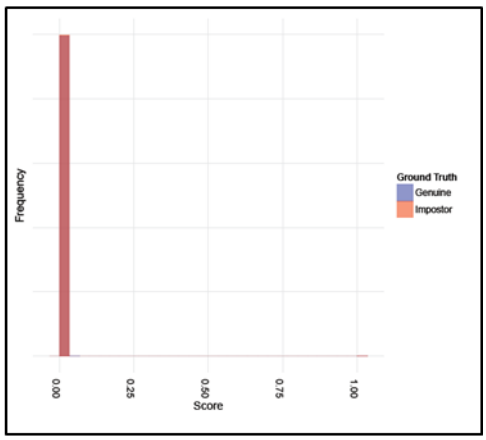
Figure 30: All-to-All Verification –ITWCC-2- DET



(a) Cognitec's FaceVacs

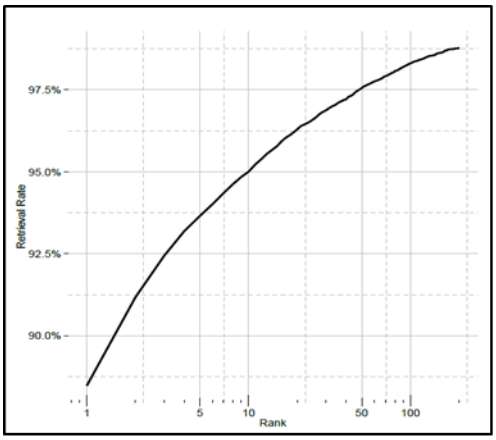


(b) Rank One Computing

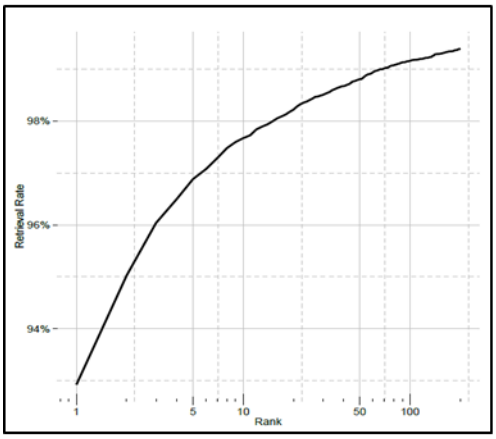


(c) Verilook

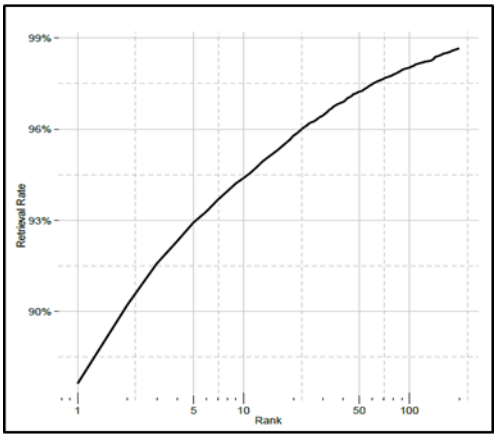
Figure 31: All-to-All Verification –ITWCC-2- Score Histogram



(a) Cognitec's FaceVacs

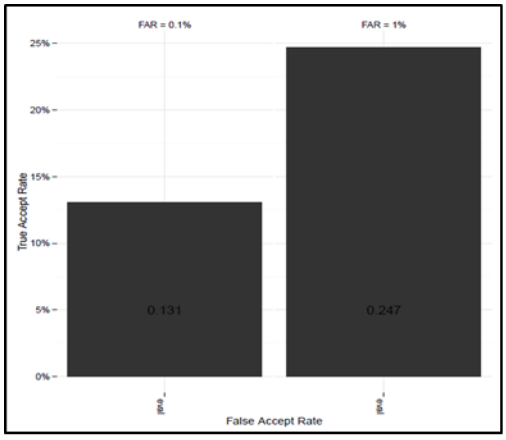


(b) Rank One Computing

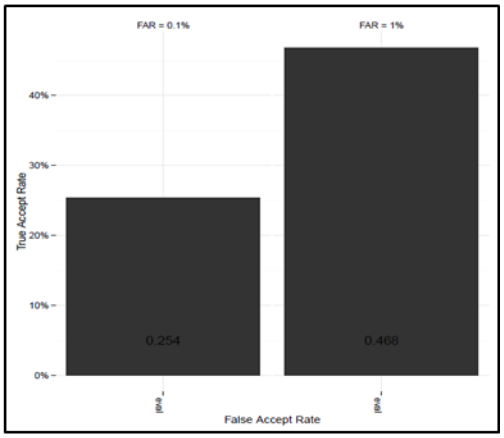


(c) Verilook

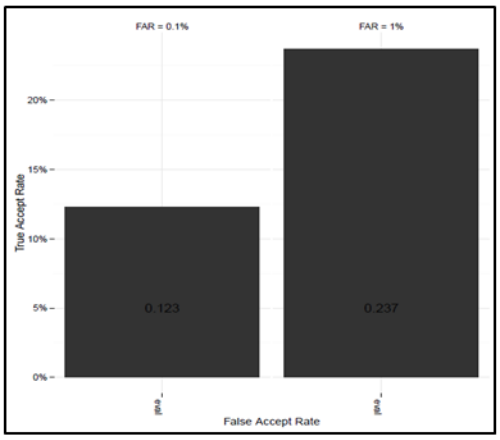
Figure 33: All-to-All Verification –ITWCC-2- CMC



(a) Cognitec's FaceVacs

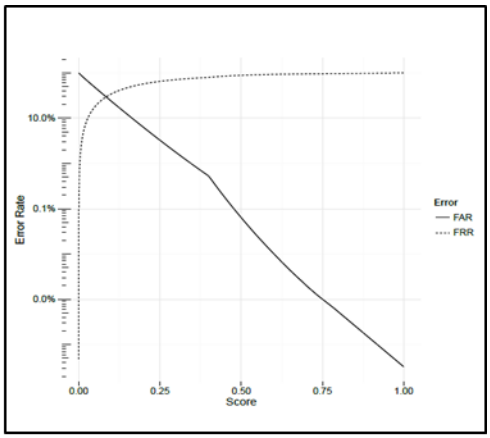


(b) Rank One Computing

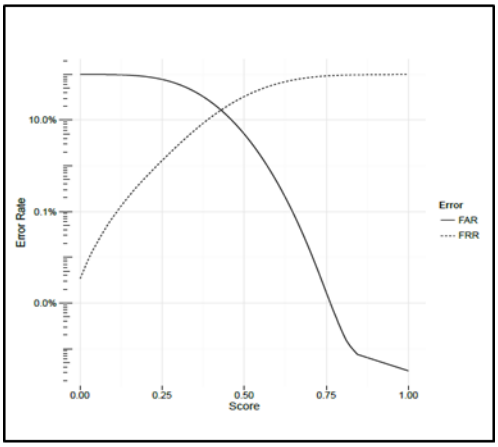


(c) Verilook

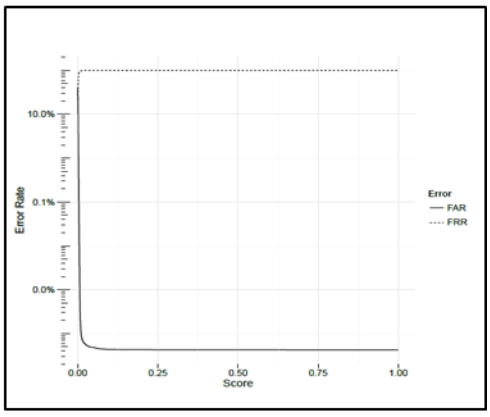
Figure 34: All-to-All Verification –ITWCC-2- ROC Scores



(a) Cognitec's FaceVacs



(b) Rank One Computing

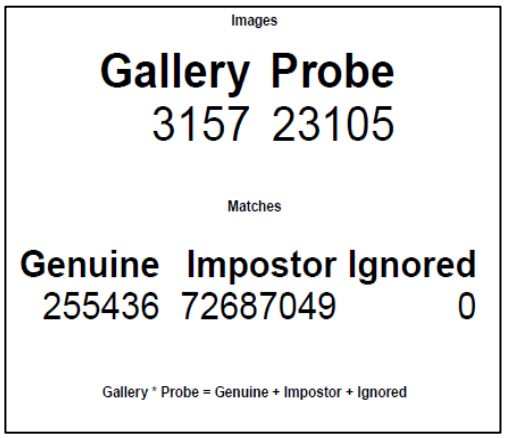


(c) Verilook

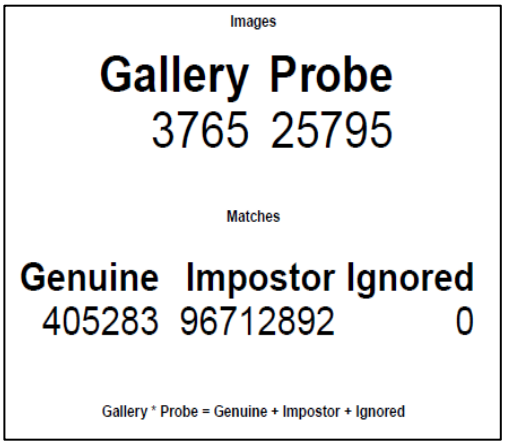
Figure 34: All-to-All Verification –ITWCC-2- EER

Appendix C

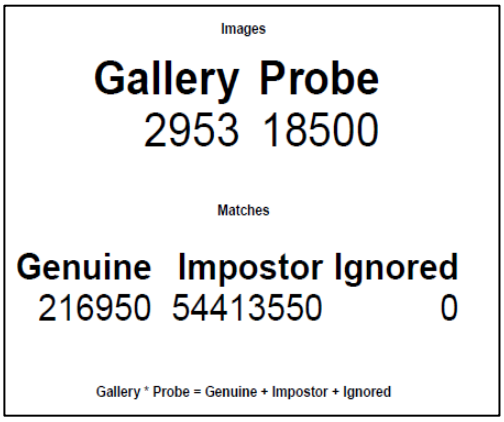
Old-To-Young Experiment Results



(a) Cognitec's FaceVacs



(b) Rank One Computing



(c) Verilook

Figure 36: Old-to-Young– ITWCC-2-Data and Match Metrics

Table of True Accept Rates at various False Accept Rates

	eval
FAR = 1e−06	0.002
FAR = 1e−05	0.018
FAR = 1e−04	0.041
FAR = 1e−03	0.091
FAR = 1e−02	0.199
FAR = 1e−01	0.452

Table of retrieval rate at various ranks

	eval
Rank 1	0.235
Rank 5	0.335
Rank 10	0.381
Rank 20	0.436
Rank 50	0.515
Rank 100	0.581

(a) Cognitec's FaceVacs

Table of True Accept Rates at various False Accept Rates

	eval
FAR = 1e−06	0.005
FAR = 1e−05	0.045
FAR = 1e−04	0.104
FAR = 1e−03	0.219
FAR = 1e−02	0.43
FAR = 1e−01	0.745

Table of retrieval rate at various ranks

	eval
Rank 1	0.377
Rank 5	0.497
Rank 10	0.549
Rank 20	0.6
Rank 50	0.673
Rank 100	0.729

(b) Rank One Computing

Table of True Accept Rates at various False Accept Rates

	eval
FAR = 1e−06	0.001
FAR = 1e−05	0.015
FAR = 1e−04	0.037
FAR = 1e−03	0.081
FAR = 1e−02	0.18
FAR = 1e−01	0.432

Table of retrieval rate at various ranks

	eval
Rank 1	0.262
Rank 5	0.365
Rank 10	0.416
Rank 20	0.468
Rank 50	0.552
Rank 100	0.621

(c) Verilook

Figure 37: Old-to-Young – ITWCC-2-TAR and Rank Values

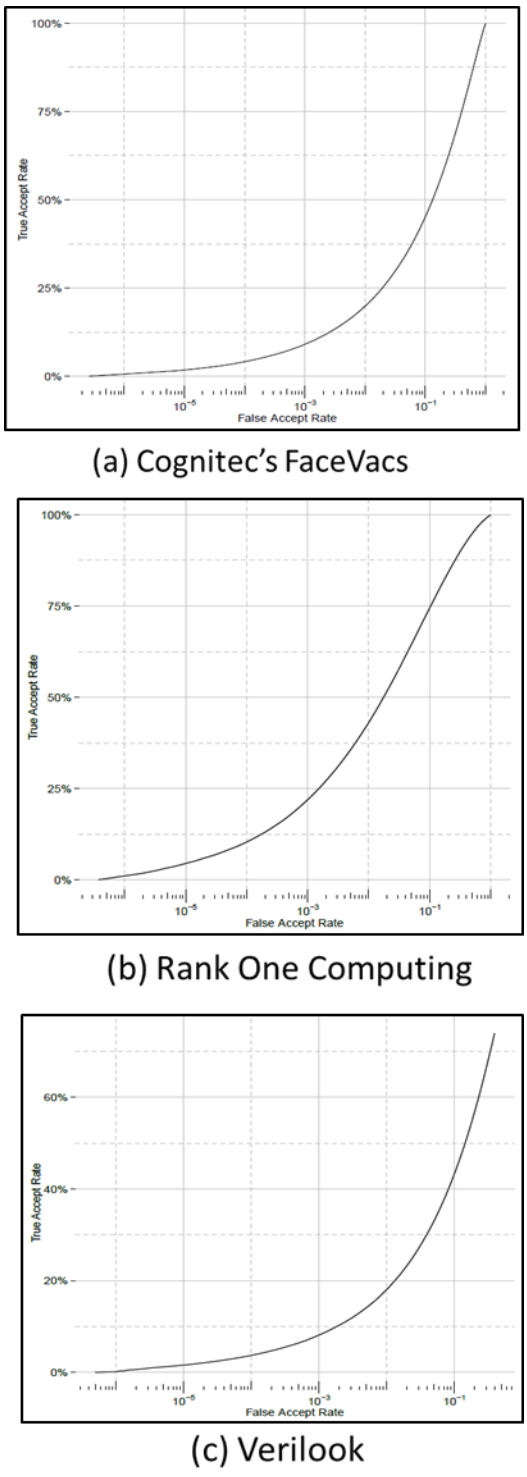
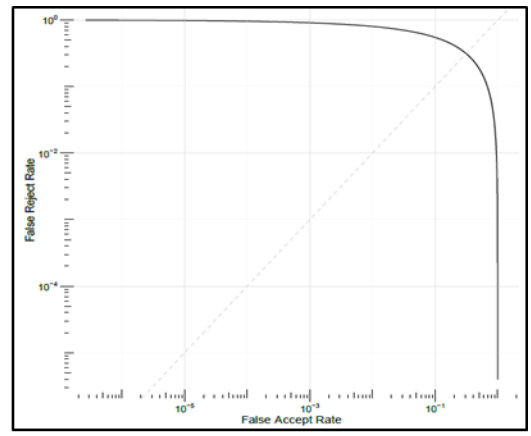
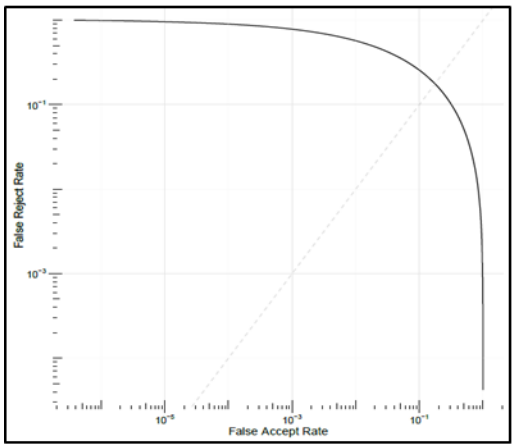


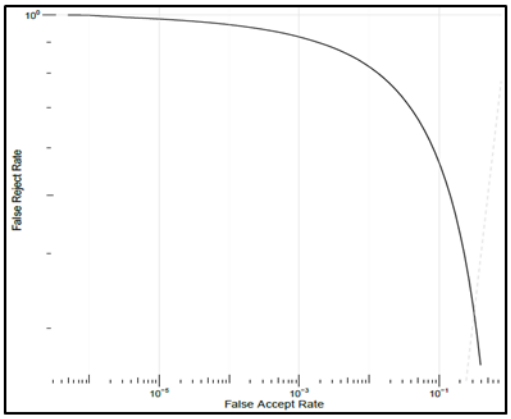
Figure 38: Old-to-Young –ITWCC-2- ROC



(a) Cognitec's FaceVacs

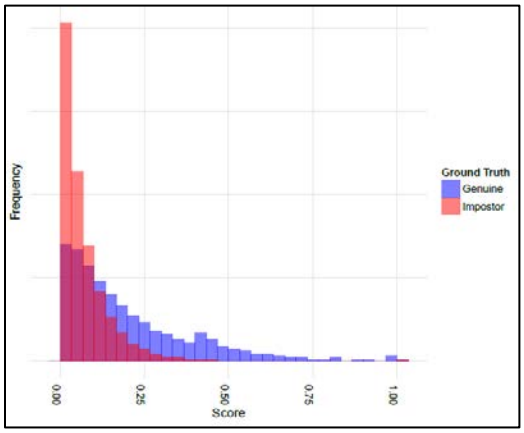


(b) Rank One Computing

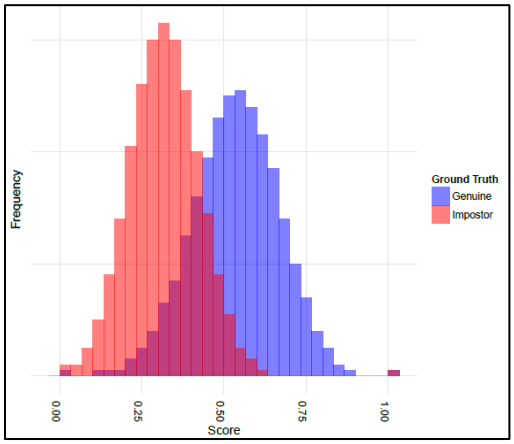


(c) Verilook

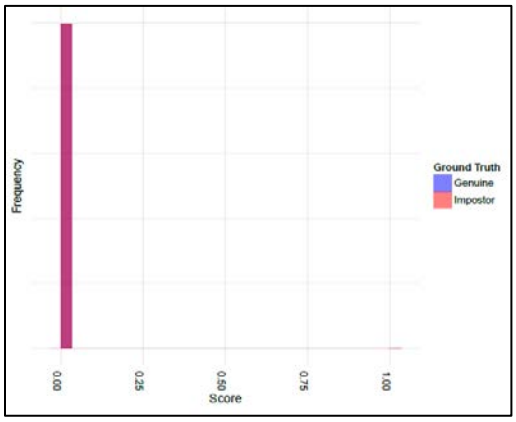
Figure 39: Old-to-Young –ITWCC-2- DET



(a) Cognitec's FaceVacs

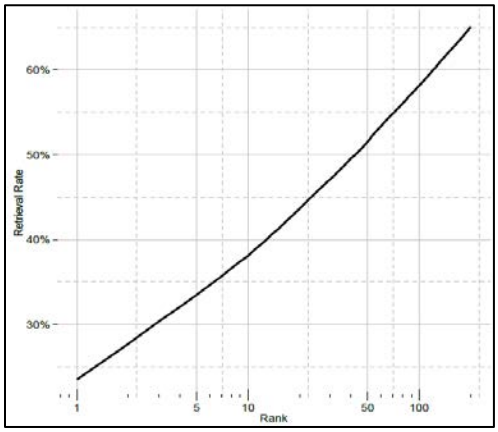


(b) Rank One Computing

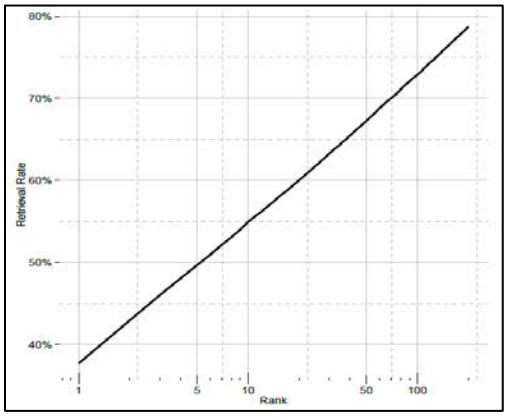


(c) Verilook

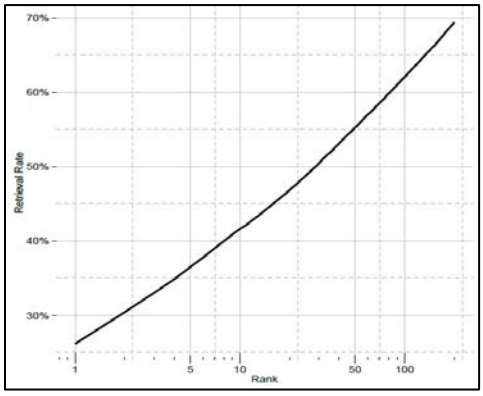
Figure 40: Old-to-Young –ITWCC-2- Score Histogram



(a) Cognitec's FaceVacs

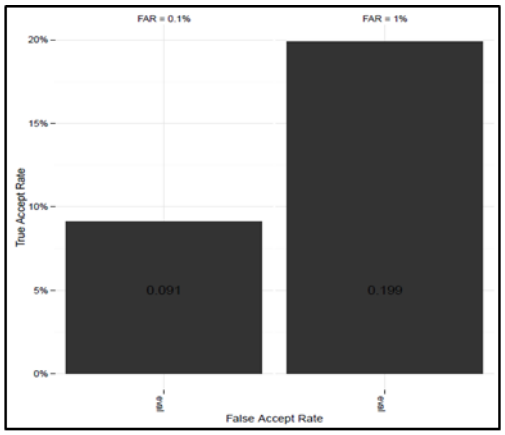


(b) Rank One Computing

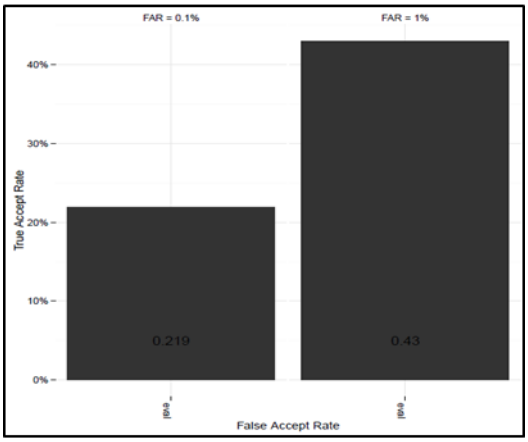


(c) Verilook

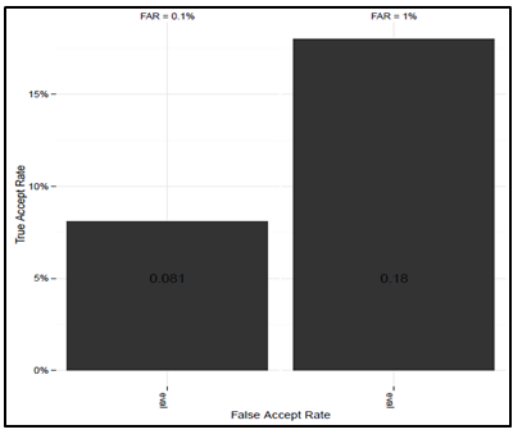
Figure 42: Old-to-Young –ITWCC-2- CMC



(a) Cognitec's FaceVacs

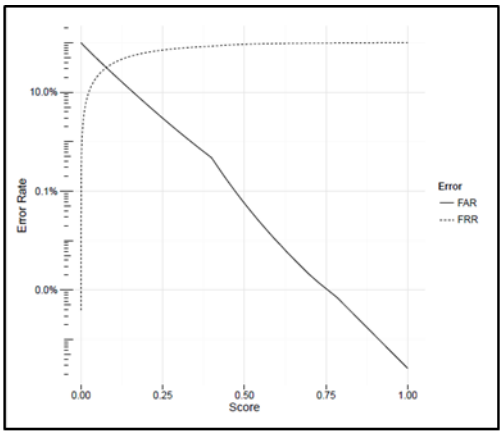


(b) Rank One Computing

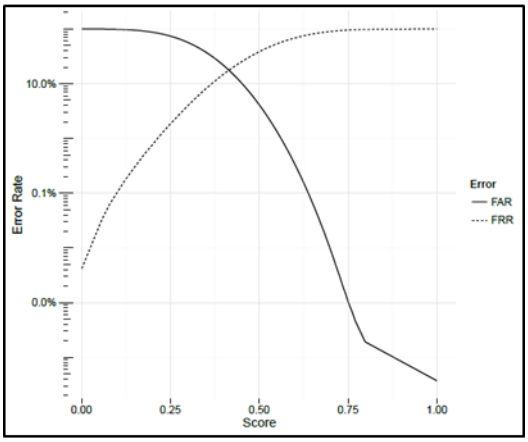


(c) Verilook

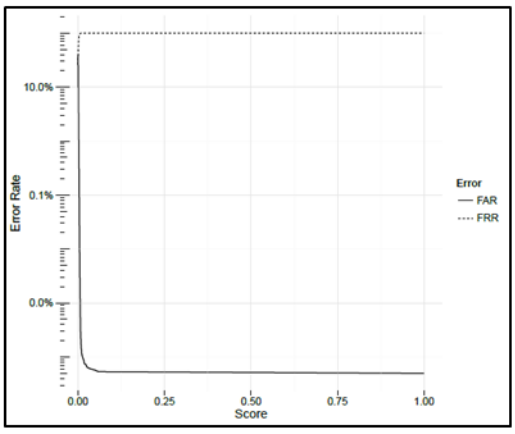
Figure 43: Old-to-Young –ITWCC-2- ROC Scores



(a) Cognitec's FaceVacs



(b) Rank One Computing

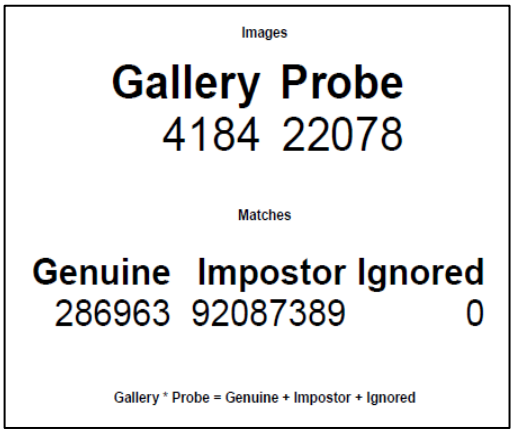


(c) Verilook

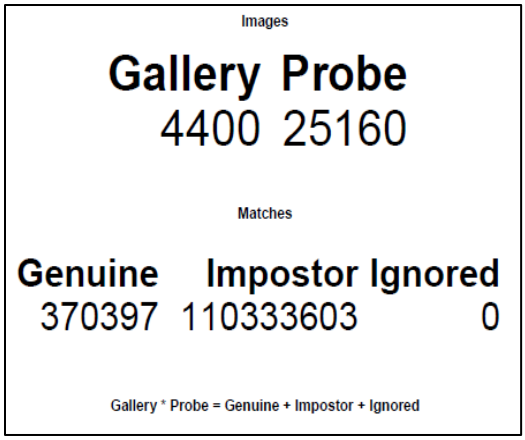
Figure 42: Old-to-Young –ITWCC-2- EER

Appendix D

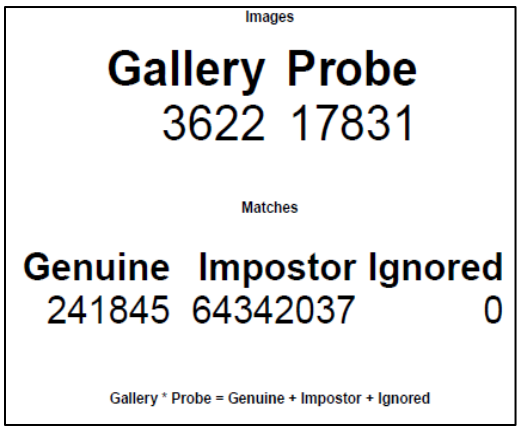
Young-to-Old Experiment Results



(a) Cognitec's FaceVacs



(b) Rank One Computing



(c) Verilook

Figure 44: Young-to-Old– ITWCC-2-Data and Match Metrics

Table of True Accept Rates at various False Accept Rates

	eval
FAR = 1e−06	0.002
FAR = 1e−05	0.023
FAR = 1e−04	0.047
FAR = 1e−03	0.095
FAR = 1e−02	0.198
FAR = 1e−01	0.438

Table of retrieval rate at various ranks

	eval
Rank 1	0.205
Rank 5	0.3
Rank 10	0.35
Rank 20	0.407
Rank 50	0.492
Rank 100	0.567

(a) Cognitec's FaceVacs

Table of True Accept Rates at various False Accept Rates

	eval_error
FAR = 1e−06	0.007
FAR = 1e−05	0.037
FAR = 1e−04	0.086
FAR = 1e−03	0.183
FAR = 1e−02	0.371
FAR = 1e−01	0.682

Table of retrieval rate at various ranks

	eval_error
Rank 1	0.346
Rank 5	0.462
Rank 10	0.541
Rank 20	0.601
Rank 50	0.685
Rank 100	0.75

(b) Rank One Computing

Table of True Accept Rates at various False Accept Rates

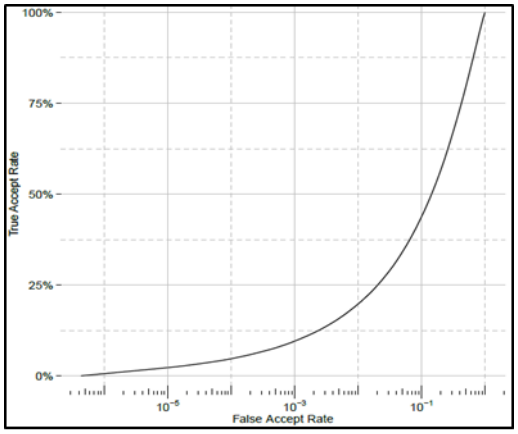
	eval
FAR = 1e−06	0.002
FAR = 1e−05	0.018
FAR = 1e−04	0.041
FAR = 1e−03	0.084
FAR = 1e−02	0.183
FAR = 1e−01	0.425

Table of retrieval rate at various ranks

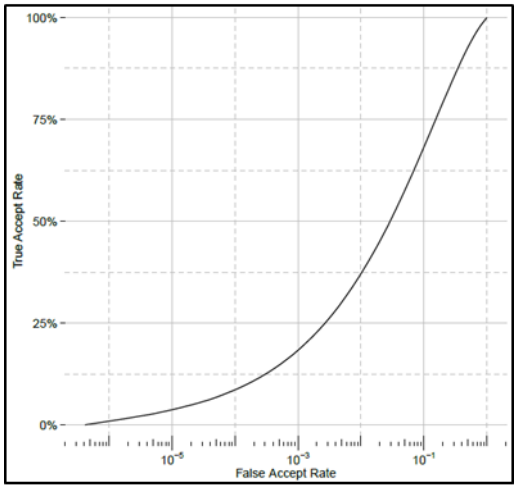
	eval
Rank 1	0.201
Rank 5	0.289
Rank 10	0.342
Rank 20	0.403
Rank 50	0.495
Rank 100	0.574

(c) Verilook

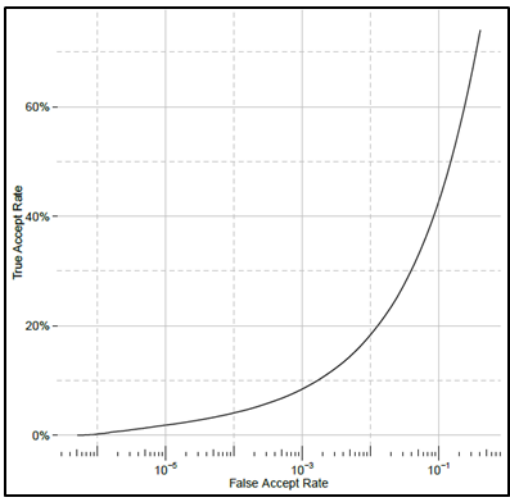
Figure 44: Young-to-Old – ITWCC-2-TAR and Rank Values



(a) Cognitec's FaceVacs

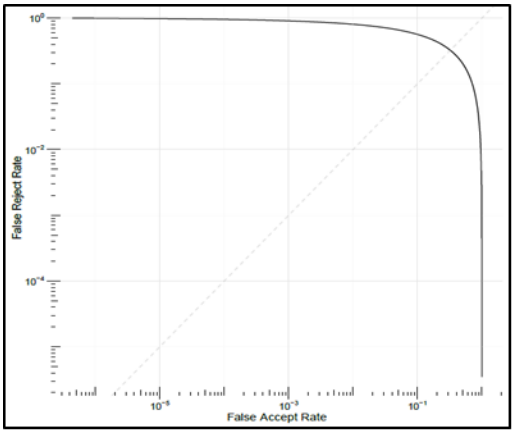


(b) Rank One Computing

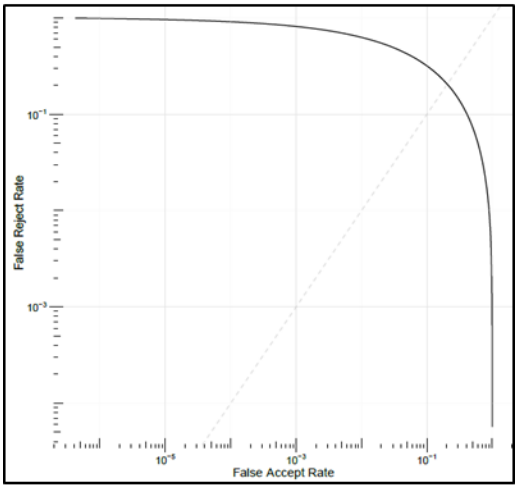


(c) Verilook

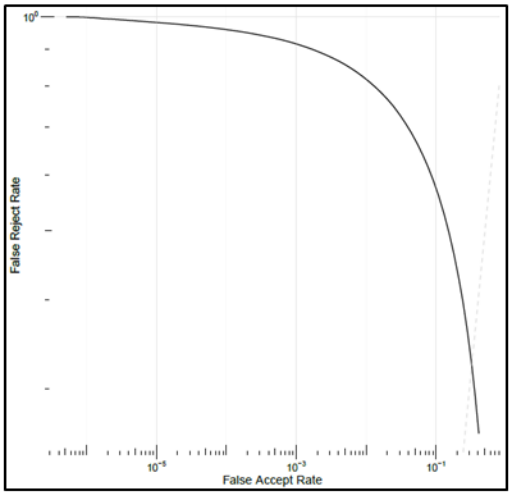
Figure 46: Young-to-Old –ITWCC-2- ROC



(a) Cognitec's FaceVacs

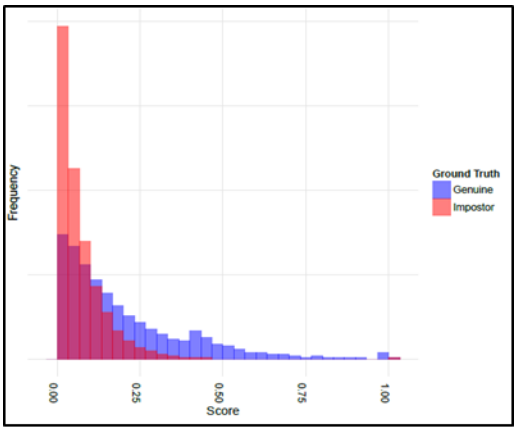


(b) Rank One Computing

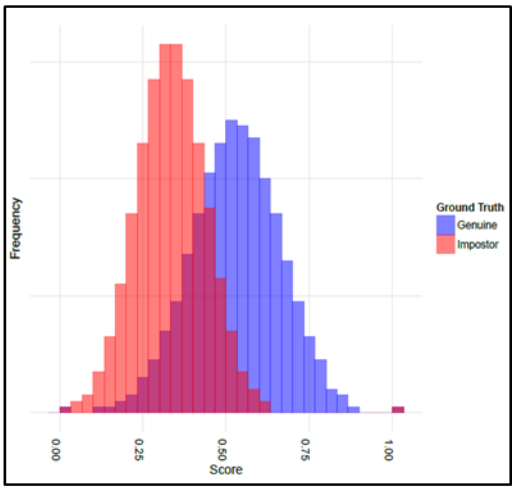


(c) Verilook

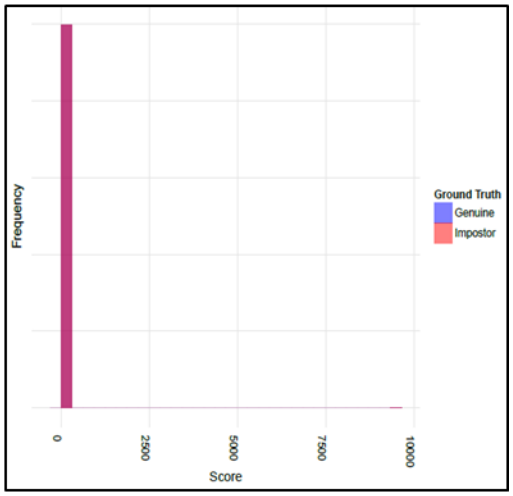
Figure 47: Young-to-Old –ITWCC-2- DET



(a) Cognitec's FaceVacs

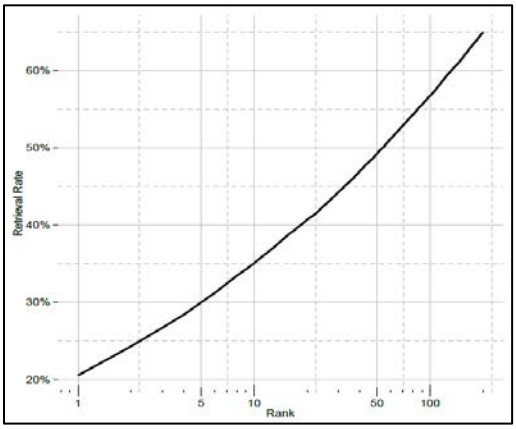


(b) Rank One Computing

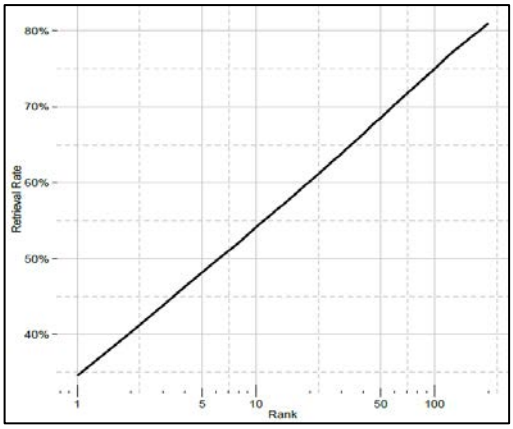


(c) Verilook

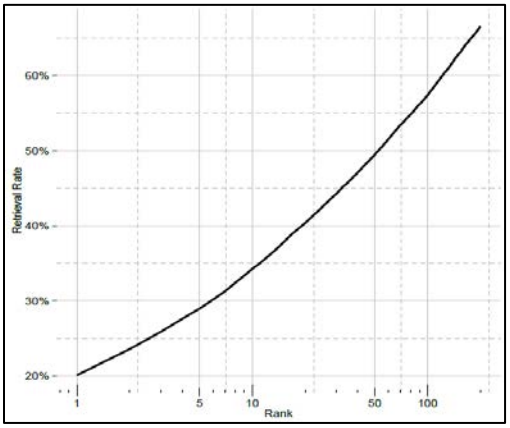
Figure 48: Young-to-Old –ITWCC-2- Score Histogram



(a) Cognitec's FaceVacs

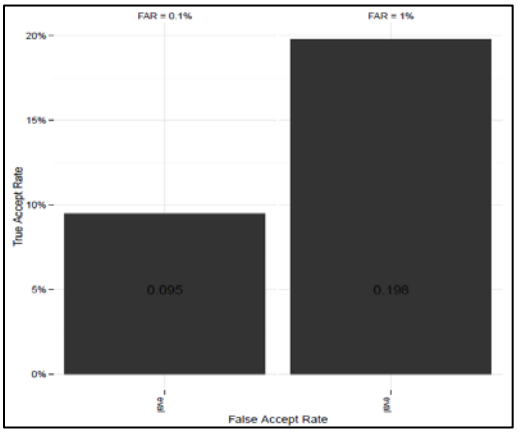


(b) Rank One Computing

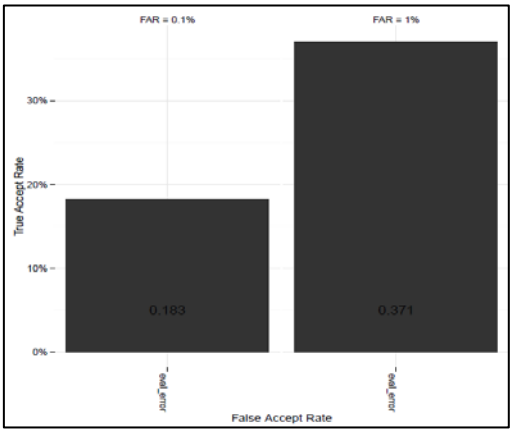


(c) Verilook

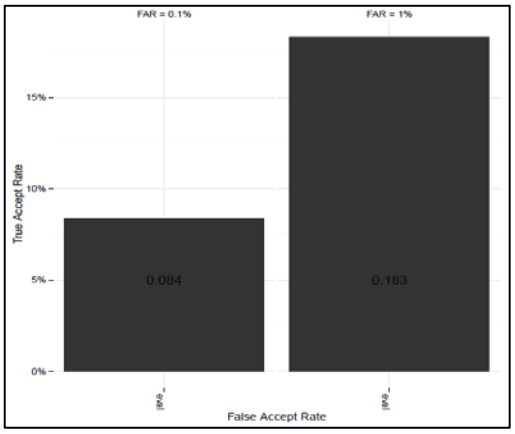
Figure 49: Young-to-Old –ITWCC-2- CMC



(a) Cognitec's FaceVacs

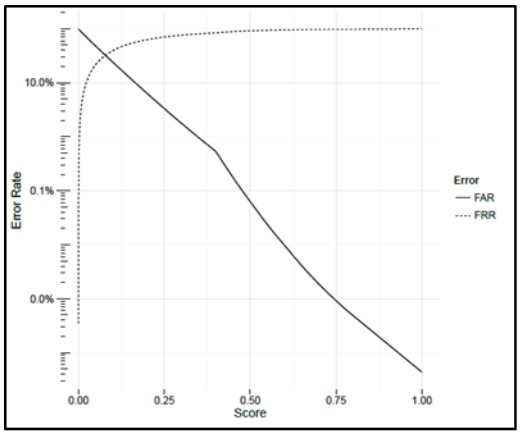


(b) Rank One Computing

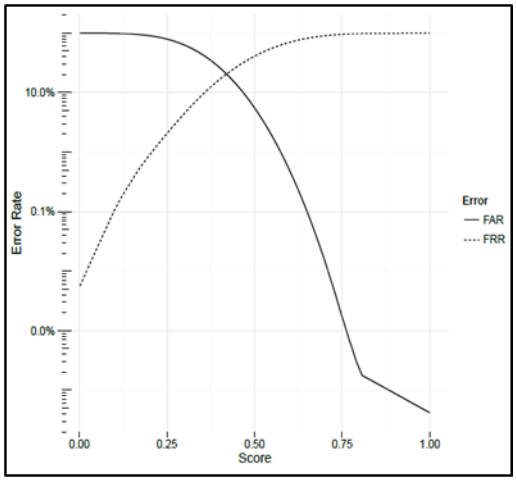


(c) Verilook

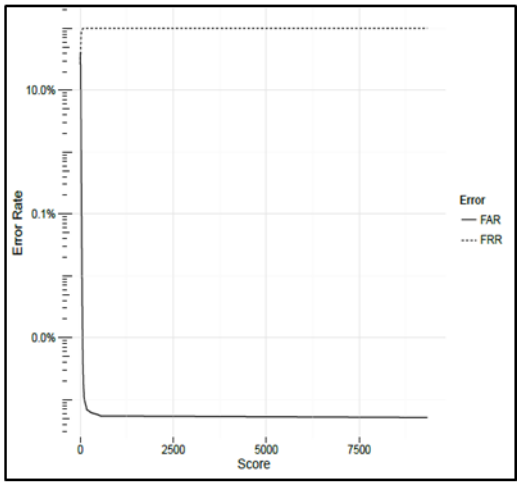
Figure 50: Young-to-Old –ITWCC-2- ROC Scores



(a) Cognitec's FaceVacs



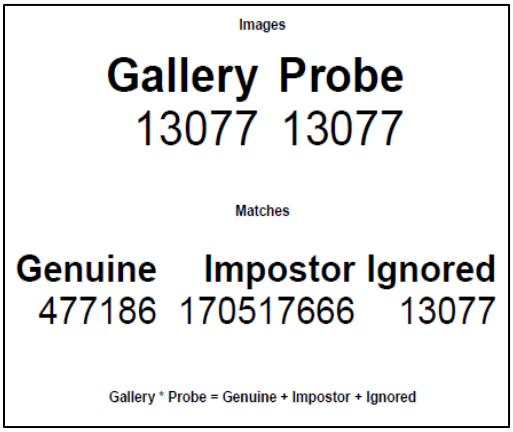
(b) Rank One Computing



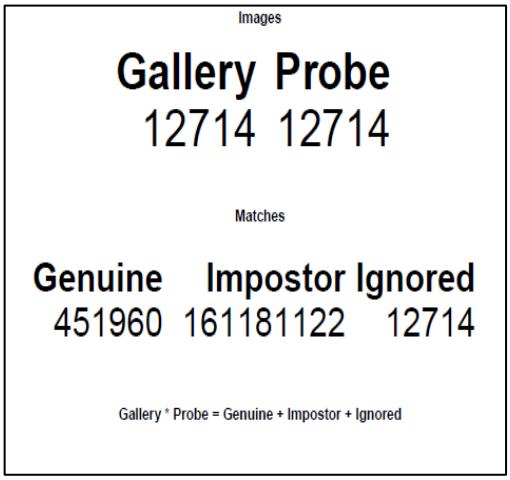
(c) Verilook

Figure 51: Young-to-Old –ITWCC-2- EER

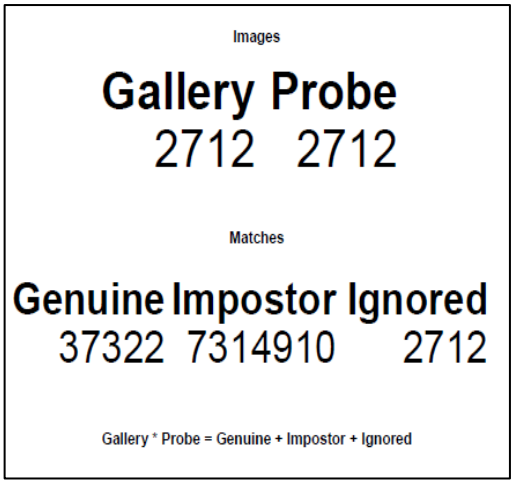
Appendix E
LFW All-to-All Verification Experiment Results



(a) Cognitec's FaceVacs



(b) Rank One Computing



(c) Verilook

Figure 52: All-to-All Verification – LFW-Data and Match Metrics

Table of True Accept Rates at various False Accept Rates

	eval
FAR = 1e−06	0.002
FAR = 1e−05	0.165
FAR = 1e−04	0.267
FAR = 1e−03	0.379
FAR = 1e−02	0.531
FAR = 1e−01	0.747

Table of retrieval rate at various ranks

	eval
Rank 1	0.691
Rank 5	0.752
Rank 10	0.777
Rank 20	0.801
Rank 50	0.829
Rank 100	0.852

(a) Cognitec's FaceVacs

Table of True Accept Rates at various False Accept Rates

	eval
FAR = 1e−06	0.122
FAR = 1e−05	0.338
FAR = 1e−04	0.522
FAR = 1e−03	0.716
FAR = 1e−02	0.876
FAR = 1e−01	0.965

Table of retrieval rate at various ranks

	eval
Rank 1	0.845
Rank 5	0.901
Rank 10	0.917
Rank 20	0.934
Rank 50	0.951
Rank 100	0.962

(b) Rank One Computing

Table of True Accept Rates at various False Accept Rates

	eval
FAR = 1e−06	0
FAR = 1e−05	0.095
FAR = 1e−04	0.164
FAR = 1e−03	0.233
FAR = 1e−02	0.321
FAR = 1e−01	0.484

Table of retrieval rate at various ranks

	eval
Rank 1	0.457
Rank 5	0.505
Rank 10	0.537
Rank 20	0.572
Rank 50	0.63
Rank 100	0.689

(c) Verilook

Figure 52: All-to-All Verification – LFW -TAR and Rank Values

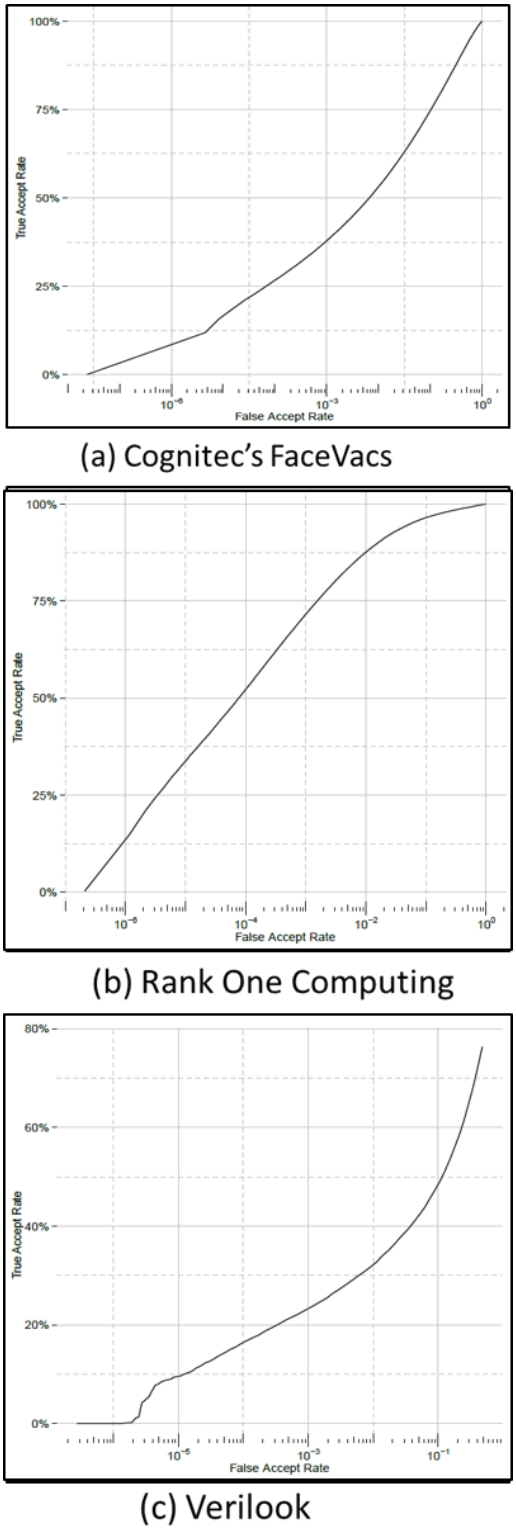
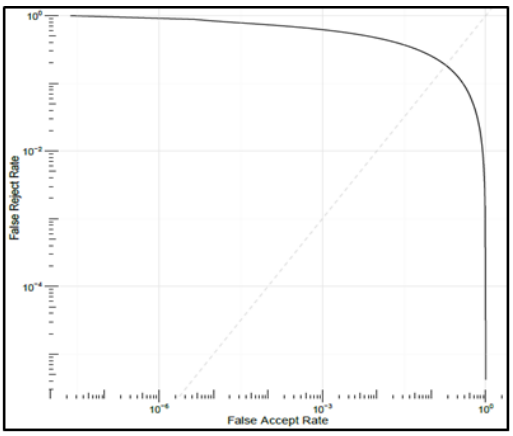
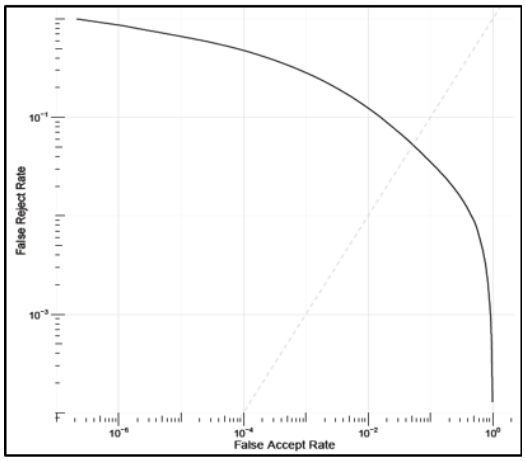


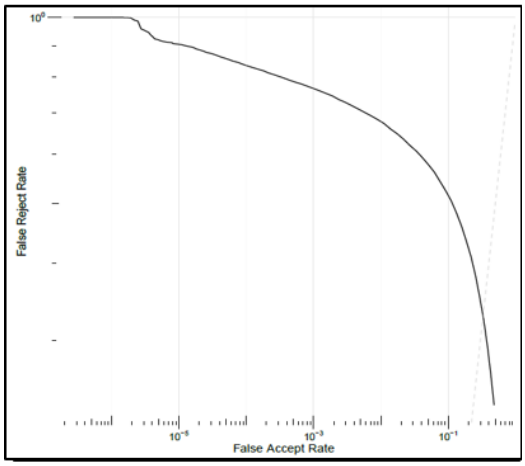
Figure 53: All-to-All Verification – LFW - ROC



(a) Cognitec's FaceVacs

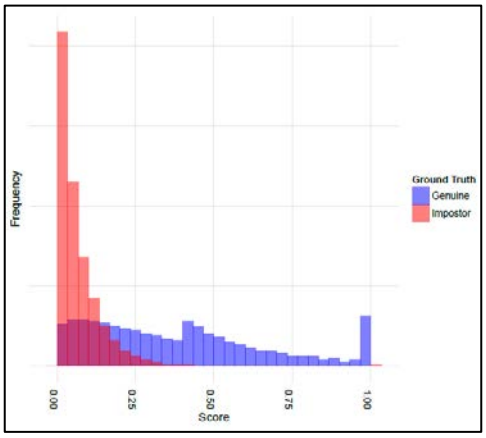


(b) Rank One Computing

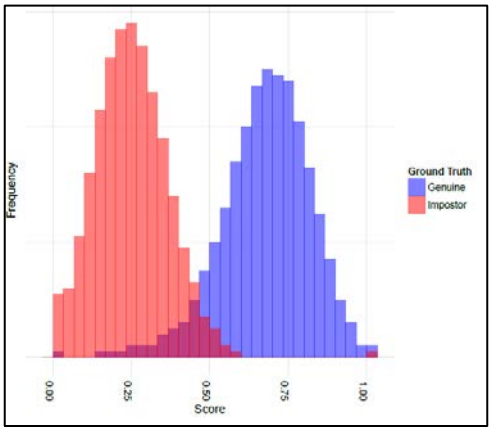


(c) Verilook

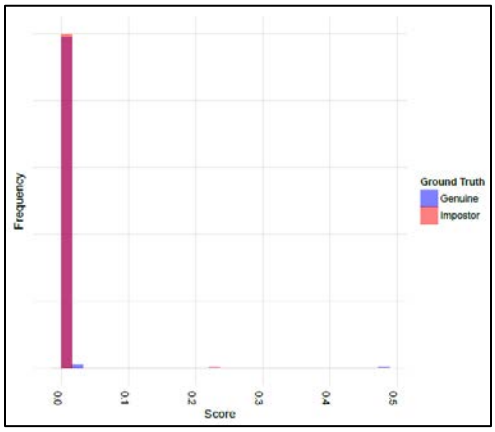
Figure 54: All-to-All Verification – LFW - DET



(a) Cognitec's FaceVacs

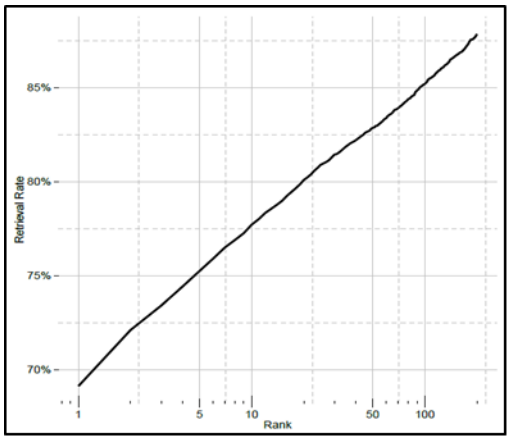


(b) Rank One Computing

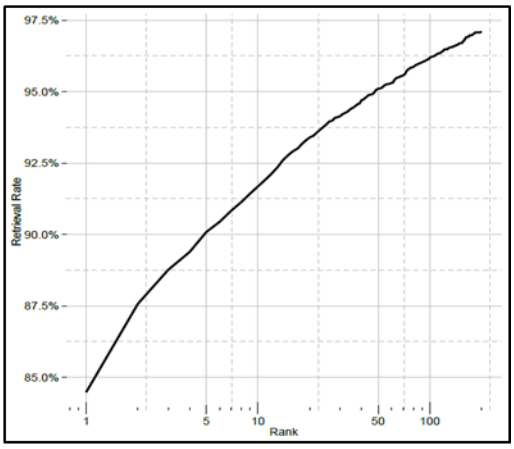


(c) Verilook

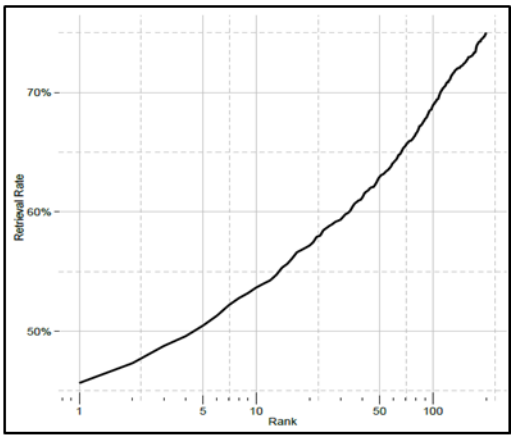
Figure 55: All-to-All Verification – LFW - Score Histogram



(a) Cognitec's FaceVacs

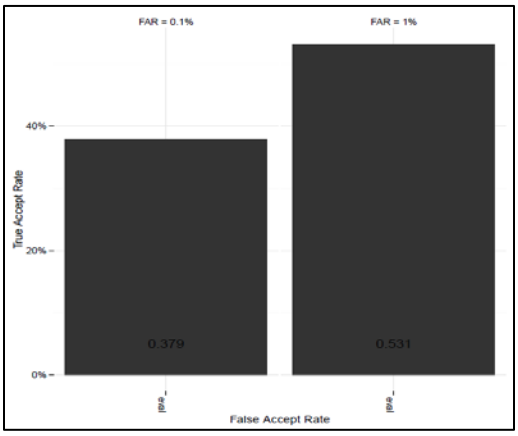


(b) Rank One Computing

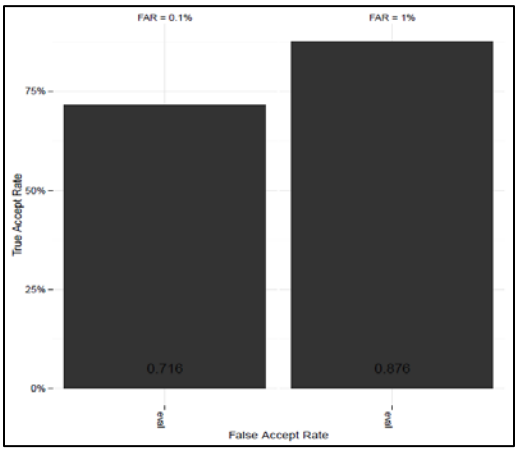


(c) Verilook

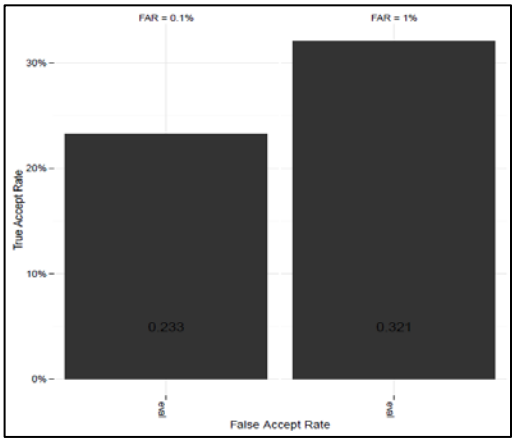
Figure 56: All-to-All Verification – LFW - CMC



(a) Cognitec's FaceVacs

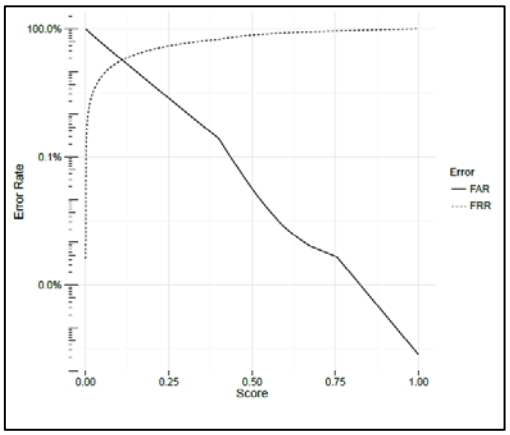


(b) Rank One Computing

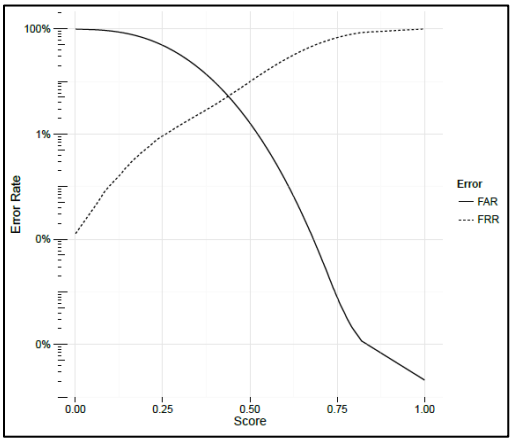


(c) Verilook

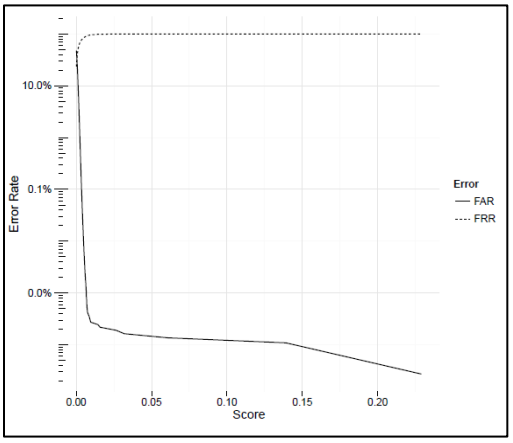
Figure 57: All-to-All Verification – LFW - ROC Scores



(a) Cognitec's FaceVacs



(b) Rank One Computing



(c) Verilook

Figure 58: All-to-All Verification – LFW - EER

Appendix F

Aging in Sub-Adults and Adults

Karen Taylors [55] work, describes the facial deformations of an adolescent male as follows:

Age 1½

- the head is very rounded and the facial contours are full and rounded
- the face occupies the lower portion of the head and there is a lot of cranium
- the bridge of the nose is flattened
- there is the appearance of epicanthic folds and the eyes look large and rounded
- the hair is soft and fine

Age 3

- the cranium expands to accommodate the growing brain
- the eyes are slightly more elongated and less rounded
- the maxilla and mandible have enlarged and widened to allow room for the deciduous dentition
- the interorbital distance is almost established

Age 4

- the nose is still small and buttonlike with small nostrils, but the bridge has begun to form
- the interorbital distance is basically established
- the anterior "baby teeth" are visible
- the ears seem very large and very low on the head
- the chin has taken some shape

Age 5

- the bridge of the nose continues to rise up, lifting some of the excess skin from the medial corners of the eyes
- the face continues to elongate as the nose length and the chin length increase
- growth pattern of the hair seems firmly established, it is less fine, and the color darkens

Age 6½

- the forehead has become less prominent and bulbous looking
- the bridge of the nose continues to rise up and the nostril size and nose width increase slightly
- the nose continues to grow in length, as does the chin
- the forms of the lower cartilages of the nose become apparent and the tip takes shape
- the central and lateral deciduous incisors have been lost and replaced by the permanent maxillary central incisors, thus the dentition is "mixed"
- the mouth has to grow to accommodate the permanent teeth

Age 8

- there is not much proportional change
- the permanent maxillary lateral incisors have appeared but they are not fully down

Age 9½

- the bridge of the nose continues to rise up and the nostril size and nose width still increase slightly
- the face elongates slightly
- the squarish form of the chin becomes obvious
- the permanent maxillary lateral incisors are fully down
- the teeth seem big for the face

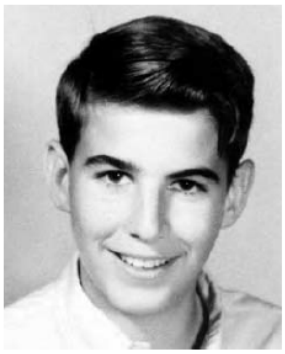
Age 11

- the "childlike" face is looking more "juvenile" as some of the facial forms become more apparent due to less "baby fat"
- the bridge of the nose continues to rise up
- the upper lip has remained about the same for several years
- the form of the chin becomes more masculine and square
- the ears still seem large for the face although they do not appear so low on the head

Age 12

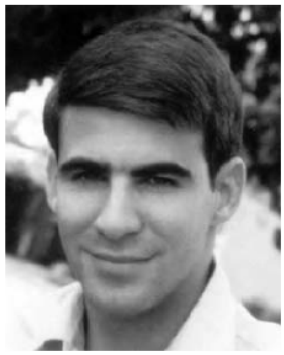
- the nose continues to grow, both in the bridge and the nostril size
- the teeth still seem big for the face
- the permanent maxillary canines or cuspids are in
- the mandible and chin continue to grow
- the neck musculature is slightly more masculine

Age 15



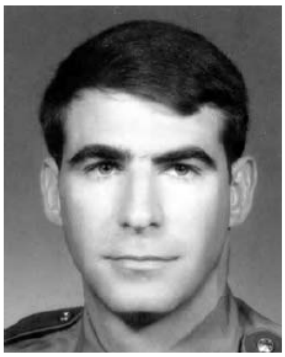
- the "juvenile" face is starting to look more "teen-aged" (and would probably be even more noticeable in a female)
- the teeth don't seem so oversized for the face as the mandible continues to grow and become more masculine
- the cheekbones have become relatively more prominent
- the ears don't seem so oversized for the face as it has elongated and they don't seem too low
- the eyebrows have become more masculine
- oops! … a zit or two!

Age 20



- the nose has grown even more, revealing the nasal bones at the bridge
- the forehead shape has remained consistent though it has grown and risen up at glabella
- the face now looks mature, particularly due to the appearance of facial hair
- the eyebrows are even fuller
- the angle of the mandible is much more squared and masculine
- the neck musculature is mature

Age 25



- the eyebrows are heavier
- the mandible has continued to square and form, looking more masculine and mature
- the neck musculature is more defined and the Adam's apple is visible

Study conducted by A. Midori Albert et al. describes the hard and soft tissue changes in adults as follows [56].

Adult hard and soft-tissue age-related changes

Approximate age range (years)	Likely bony change	Probable soft tissue or facial appearance effect
20–30	• Slight craniofacial skeletal growth. • Slight anterior (mostly lower) face height increase. • Mandibular length increase.	• Upper eyelid drooping begins. • Eyes appear smaller. • Nasolabial lines begin to form. • Lateral orbital lines begin to form. • Upper lip retrusion begins in females.
30–40	• Dentoalveolar regression suggesting eruptive movement of teeth. • Maxillary retrusion progressing, contributing to nasolabial folds. • Mandibular length increase.	• Circumoral striae begin to form. • Lines begin to form from lateral edges of nose to lateral edges of mouth. • Upper lip thickness decreasing.
40–50	• Craniofacial skeletal remodeling progresses. • Dental alveolar regression and dental eruption progressing. • Maxillary and mandibular dental arch lengths decreasing.	• Facial lines and folds continue to increase in depth. • Nose and chin positioning affected as dental arch lengths decrease. • Most profound morphological changes of the head, face, and neck are evident.
50–60	• Craniofacial remodeling continues. • Cranial thickness likely unchanging. • Alveolar bone remodeling. • Possible dental attrition affecting vertical face height.	• Facial lines and folds continue to increase in depth. • Protuberance of nose and ears due to greater craniofacial convexity.
>60	• Decrease in craniofacial size. • Greater craniofacial convexity (excluding maxilla and mandible). • Possible temporomandibular joint arthritis and joint flattening. • Alveolar bone remodeling continues.	• Protuberance of nose and ears continues. • Concave appearance in cheek hollows due to alveolar bone remodeling. • Diminished jaws.