# A Study of Age and Ageing in Fingerprint Biometrics

Javier Galbally, Rudolf Haraksim, and Laurent Beslay

***Abstract*— Thanks to Mr. James Bond, we are aware that diamonds are forever but, are fingerprints? It is well known that biometrics brings to the security field a new paradigm; unlike traditional systems, individuals are not identified by something that they have or they know, but by what they are. While such an approach entails some clear advantages, an important question remains: is what we are today the same as what we will be tomorrow? This paper addresses such a key problem in the fingerprint modality based on a database of over 400K impressions coming from more than 250K different fingers. The database was acquired under real operational conditions and contains fingerprints from subjects aged 0–25 and 65–98 years. Fingerprint pairs were collected with a time difference that ranges between 0 and 7 years. Such a unique set of data has allowed us to analyze both the age and ageing effects, shedding some new light into issues, such as fingerprint permanence and fingerprint quality.**

***Index Terms*— Biometrics, fingerprint recognition, ageing, children, elderly, fingerprint quality.**

## I. Introduction

"*Every single cell in the human body replaces itself over a period of seven years. That means there's not even the smallest part of you now that was part of you seven years ago.*" - Steven Hall, The Raw Shark Texts.

*YOU are your own key*. Behind this catchy principle biometrics have become an attractive alternative to traditional identification methods such as tokens or passwords. However, what would happen if this new natural in-built key changed over time? Would it still open the door it was designed for?

To answer these legitimate questions, there is the need to analyse the way in which time affects biometric characteristics and the effect that such changes have on the performance of automatic biometric recognition systems. In particular, the present paper focuses on the study of fingerprints and time.

In order to understand how fingerprint recognition systems are affected by time, it is important to notice that each user is acquired at two separate points, during the enrolment of its reference template and during the collection of the probe sample. This double interaction with the system produces two different (but linked) time effects: the age effect and the ageing effect.

- **Age effect**. This effect accounts for the variations in accuracy between different user groups according to their age, such as, for example, children, adults and elderly. In this case, *assuming a short time difference between the reference and probe acquisitions*, the question being addressed is: How does the performance of fingerprints vary through life? Can we expect the same performance from fingerprint recognition systems for 3-year old children, than for 25-year old adults, than for 90-year old elders? This effect is mainly related to the *collectability* of fingerprints which has a direct impact on their *quality*.
- **Ageing effect**. This effect accounts for the variations in accuracy due to the increase of the time difference between the reference sample and the probe sample [1]. Accordingly, the question being addressed in this case is: Can we expect the same Genuine Matching Score (GMS) distribution when the time difference between the reference and probe samples is 1 year, 5 years or 10 years? Furthermore, is this effect dependent on the age of the reference sample? Ageing is mainly related to the *permanence* of fingerprints, or rather, to the lack of it.
  Note that, in the present article, *fingerprint permanence* (also *fingerprint persistence*) does not refer to the ridge structure anatomy of the fingertip, but to the ability to reliably acquire and recognise in the digital domain this ridge structure over time. That is, it does not refer to the physical world, but to changes in the digital images captured with current live-scan touch-based technology. Theoretically, fingertips may withstand the passing of time. However, in practice, if the images acquired for recognition change, eventually, the fingertips may be unusable to automatically discern individuals apart.
  In summary, the present article does not address the theoretical immutability of physical fingertips, but their practical usability for recognition purposes over time.
  The same "digital" interpretation of fingerprint permanence has been used in previous studies from the literature [2].

Compared to the other factors such as the universality or uniqueness of fingerprints [3]–[5], a clear gap exists in terms of research effort with respect to the previous two time effects. How does time affect fingerprints performance? Do these changes affect automatic fingerprint recognition systems? Are there any age limits for the use of fingerprints with current technology? When is the time difference

Manuscript received June 18, 2018; revised October 8, 2018; accepted October 12, 2018. Date of publication October 25, 2018; date of current version January 30, 2019. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Adams W. K. Kong.
*(Corresponding author: Javier Galbally.)*

The authors are with the European Commission, Joint Research Center (DG-JRC), 21027 Ispra, Italy (e-mail: javier.galbally@ec.europa.eu; rudolf.haraksim@ec.europa.eu; laurent.beslay@ec.europa.eu).

This paper has supplementary downloadable material at http://ieeexplore.ieee.org, provided by the authors. The supplementary material consists of a pdf file containing Annex A which reports further experimental results.

Digital Object Identifier 10.1109/TIFS.2018.2878160

between two samples of the same user too large to trust their matching? How does fingerprint quality evolve through a lifetime? Although some valuable efforts have been conducted to address these and other similar issues (see Sect. II for a review of the state of the art), there is still not enough consistent evidence to be able to provide reliable answers to the questions above.

The lack of large comprehensive studies addressing the effect of time on fingerprint-based technology is mostly explained by two factors: 1) On the one hand, the acquisition of fingerprints with the new generation of live-scan devices, instead of the traditional ink-and-paper method, is quite recent. As such, long-term data with which to carry out such studies is scarce. 2) On the other hand, large datasets of fingerprint images acquired in real operational conditions are, rightly so, secured under data protection regulations that severely restrict the access to these data, even for research purposes.

This situation of data scarcity has forced researchers to conduct their time-related studies on limited sets of fingerprints, in many cases acquired under controlled laboratory-like conditions for the purpose of the experiments [6]–[8]. While the results of these works are certainly valid to point out general trends and to formulate hypotheses, further analysis is required on larger, more comprehensive and realistic sets of data in order to confirm those results and to provide more consistent evidence that supports the findings.

The present research study is an attempt to bridge this existing gap and to shed some further light into the problem of fingerprints and time. To that end, the study has been conducted on a database of fingerprints captured under real operational conditions for the issuing of passports. The database contains almost half a million fingerprints of ages between 0 and 98, with a time separation between samples of the same finger of 0 to 7 years. Based on this unique set of data, the main contributions of the work are:

- First comprehensive study of fingerprint quality for the whole age range of human life (from 0 to 25 and from 65 to 98 years of age).
- First comprehensive study of fingerprint matching through the whole age range of human life (from 0 to 25 and from 65 to 98 years of age).
- First comparative study of the effect of ageing for different age-groups (e.g., children, adults, elderly).
- First comprehensive study of the possible limitations on the interaction of elders with fingerprint recognition systems.

The rest of the work is structured as follows. A review of the main works dealing with age-related factors in fingerprint recognition is given in Sect. II. Sect. III describes the key characteristics of the database used in the work. The experimental protocol, divided into age- and ageing-related experiments, is presented in Sect. IV. The results obtained following this protocol, as well as some partial findings, are given in Sect. V. Final conclusions are drawn in Sect. VI.

## II. Related Works

The effect of time on biometric technology has been lately the focus of books [9], surveys [1] and specific publications in biometric characteristics such as fingerprint [2], face [10], iris [11], [12], hand [13], or signature [14]. In the present section we will only consider the most relevant works published in the field of fingerprint-based technology.

Two of the pioneers in the development of fingerprint recognition, Sir William Herschel and Francis Galton, in two of their first articles, already considered the problem of fingerprint permanence [15], [16]. Both research studies were very limited: three fingerprints of one person (his son) taken at 7, 17 and 40 years of age in the case of Herschel and six subjects in the case of Galton, with time gaps between the two collections of 11 to 31 years. However, these early works already showed the importance of the permanence issue and they set the basis for other larger studies that came afterwards.

In more recent times, several small-scale studies performed on live-scanned data have shown that the ageing effect on fingerprints can be perceived for a time difference as short as three to four years [17]–[19]. While these works were valuable to alert on the potential problems posed by ageing, they were carried out on limited sets of data that prevented from extracting conclusive findings.

The most comprehensive study to date focused on the permanence of fingerprints was published in 2015 by researchers from Michigan State University [2]. While the works by Galton and Herschel focused on analysing the variability (or invariability) of the physical ridge structure, MSU's study concentrates on the impact that changes in the digital representation of the ridge structure (i.e., fingerprint images) may have on the genuine matching scores of automatic recognition systems working with current live-scan imaging technology (i.e., ageing). This "digital" interpretation of fingerprints permanence is the same considered in the present article.

The database used in [2] was acquired in real operational conditions for law-enforcement purposes and contains an average of 10-print cards of 15K subjects. The average time difference between the first and last acquisition for each individual is 9 years. The main limitation of the database is that the vast majority of individuals are adults, belonging to the age-range 15-40. This way, while the dataset is well suited to perform ageing-effect experiments, it does not allow analysing the age-effect derived from possible differences between age groups, e.g., differences between children, adults and elders.

The age effect which was not covered in [2] due to the lack of a suitable database, has been addressed in different previous articles [6]–[8], [20]–[24]. The main limitation of all these works is the significantly low-scale datasets used, which did not exceed 5,000 samples. This way, while the general methodology followed is correct and some interesting trends can be observed regarding the differences between age-groups, further tests are required on more comprehensive sets of data to confirm the conclusions drawn in those studies.

Furthermore, the literature on the age effect focuses almost completely on the differences between adults and children fingerprints. However, very little research has been carried out on such an important demographic group as the elderly. Unlike children, especially of young ages, elders have a significantly higher degree of autonomy both from a legal and an economic perspective. This freedom entails that they are more prone to interact on a daily basis with biometric systems in order to access the ever growing range of activities, applications and benefits secured by this technology. It is also important to highlight that we live in a society where the elderly are the fastest growing demographic [25]. Therefore, it should become a top priority for the biometric community to understand what are the challenges faced by biometrics when dealing with data coming from this segment of the population, in order to prevent potential situations of age-based discrimination [26].

## III. The Dataset

The dataset used in the experiments was provided under strict security and data protection measures by the Portuguese authorities. It contains real fingerprint operational data acquired for the issuing of passports. The data was acquired at multiple locations but in all cases optical live-scan readers working at 500 dpi were used. The acquisition process was supervised by governmental civil servants (not law-enforcers) with general knowledge about biometrics.

In total, the database contains fingerprint impressions from 265,321 different fingers which have produced a total 421,388 images. These data can be divided in three main groups according to the age of the fingers at the time of the first acquisition: children (ages 0-17), adults (ages 18-25) and elderly (ages 65-98). The dataset contains no fingerprints in the age range 26-64. The fingerprints distribution per age is shown in Fig. 1. Following the experimental protocol that will be explained in Sect. IV, the children group has been further divided into three sub-groups: children1 0-4, children2 5-12 and children3 13-17; and the elderly group into four sub-groups: elderly1 65-69, elderly2 70-74, elderly3 75-79, elderly4 80-98.

As can be seen in Table I, the fingers in the database present one or two samples. For the 156,067 fingers with two acquisitions, the separation between samples is 0 to 7 years. Table II shows the number of fingerprint pairs for each time separation and for each age-group. For a detailed year-by-year distribution of the fingerprint pairs in the database we refer the reader to Annex A, provided as accompanying material of the present article.

As a side note to the database description, the reader should be aware that, due to reasons beyond the authors' control, the database generously provided by the Portuguese authorities under an agreement with DG JRC, unfortunately did not contain any fingerprint data in the age range 26-64 years. Therefore, as will be explained in the experimental sections of this article, only estimations of the behaviour of fingerprints in these ages could be made. However, we do believe that the experiments performed on the data received (0-25 and 65-98) strongly support the estimations made.

## IV. Experimental Protocol

As mentioned in the introduction, the way time affects the accuracy of fingerprint recognition systems can be seen from two different angles, depending on whether the focus is 1) on the age of the individual at the time of the acquisition of the reference template (age effect) or whether it is 2) on the time difference between the reference template and the probe template (ageing effect). A diagram summarizing these two effects in the experimental database is depicted in Fig. 2.

It should be noted that, although ultimately it is the variability of systems accuracy that we are interested in, age and ageing are two phenomenons mostly related to genuine matching scores (i.e., matching scores between samples of the same finger). This is why, in some cases, typical accuracy

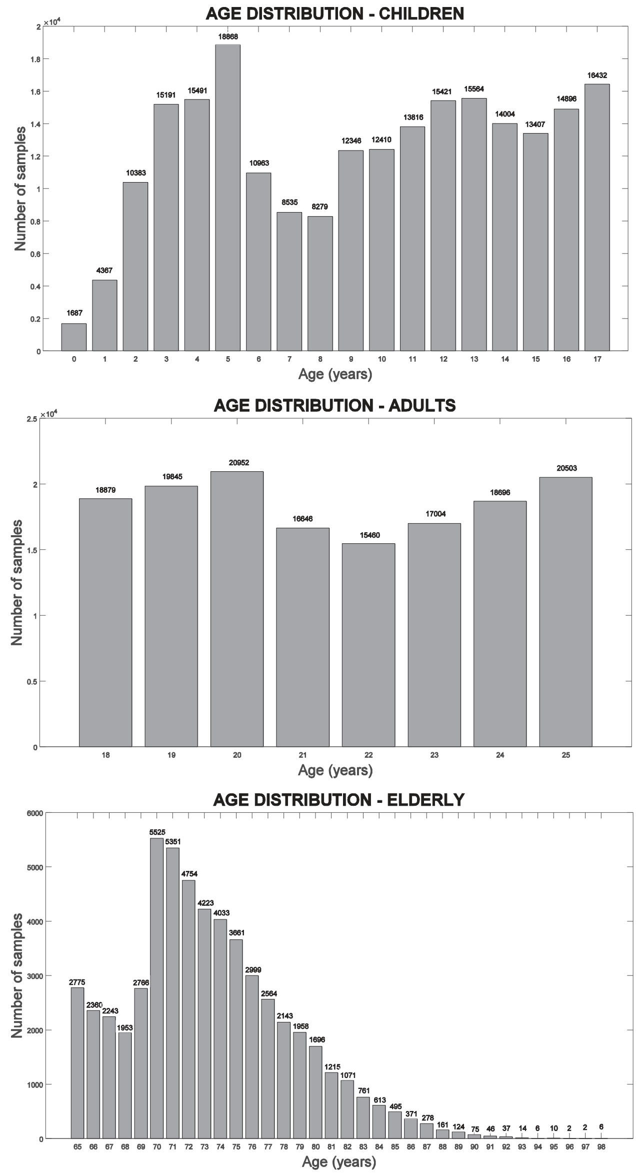


Fig. 1. Fingerprint distributions in the database according to the age.

TABLE I

Number of Fingers in the Database With 1 and 2 Samples, Divided by Age-Groups: Children (0-17), Adults (18-25) and Elderly (65-98)

| # samples | # fingers (0-17) | # fingers (18-25) | # fingers (65-98) | TOTAL |
|---|---|---|---|---|
| 1 | 45,158 | 41,992 | 22,104 | 109,254 |
| 2 | 102,600 | 36,528 | 16,939 | 156,067 |

TABLE II

NUMBER OF FINGERPRINT PAIRS IN THE EXPERIMENTAL DATASET ACCORDING TO: $i$) THE TIME DIFFERENCE BETWEEN THE FIRST AND SECOND SAMPLES (COLUMNS) AND $ii$) THE AGE-GROUP (ROWS). THE YEAR-BY-YEAR DISTRIBUTION OF THE FINGERPRINT PAIRS CAN BE CONSULTED IN ANNEX A, PROVIDED AS ACCOMPANYING MATERIAL OF THE PRESENT ARTICLE

| | | Years between first and second samples | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 0-1 | 1-2 | 2-3 | 3-4 | 4-5 | 5-6 | 6-7 | 7-8 |
| Age first acquisit. | 0-4 | 69 | 1,072 | 11,881 | 6,587 | 4,725 | 5,378 | 1,677 | 280 |
| | 5-12 | 249 | 381 | 444 | 609 | 5,323 | 24,874 | 6,466 | 607 |
| | 13-17 | 247 | 389 | 566 | 646 | 5,178 | 19,464 | 4,947 | 539 |
| | 18-25 | 1,613 | 3,525 | 2,842 | 2,193 | 7,101 | 16,012 | 2,657 | 146 |
| | 65-69 | 299 | 449 | 319 | 330 | 1661 | 5,698 | 515 | |
| | 70-74 | 147 | 181 | 177 | 124 | 788 | 3,134 | 261 | |
| | 75-79 | 91 | 99 | 78 | 57 | 419 | 1,217 | 134 | |
| | 80-98 | 47 | 77 | 41 | 26 | 135 | 349 | 42 | |

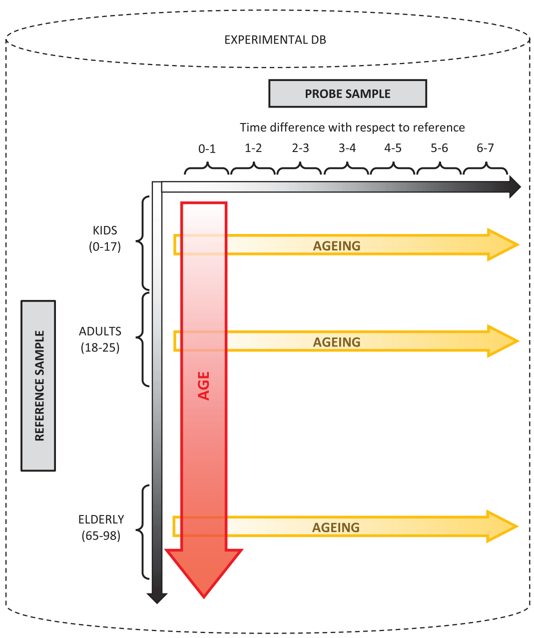


Fig. 2. Diagram depicting the age-effect and the ageing-effect in the experimental database.

metrics such as the FAR and the FRR (False Acceptance and False Rejection Rates) may not tell the whole story about these effects. Therefore, following previous related works [14], [2], in the present article in addition to the traditional DET curves (Detection Error Trade-off), we also analyse the changes suffered directly by the genuine score distribution over time.

It is also important to highlight that, the effects of age and ageing are not independent. For example, ageing may impact differently children, adults and elderly. However, studying both effects in experiments as decoupled from each other as possible can help to better understand the way time affects fingerprint-based systems and, eventually, to minimize its effects whenever possible.

Following this rationale, the experimental protocol has been divided in two main sets of experiments: 1) experiments directed to analyse the age effect (described in Sect. IV-A); and 2) experiments focused on analysing the ageing effect (described in Sect. IV-B).

In these two sets of experiments two different publicly available software tools, one matcher and one quality metric, have been used:

- **Matcher: VeriFinger**. VeriFinger (Version 10.0 of the Neurotechnology feature extraction and matching algorithms), based on the MegaMatcher identification engine and compliant with NIST MINEX [27]. This system obtained state-of-the-art results in the Fingerprint Vendor Technology Evaluation (FpVTE) organised by NIST in 2012 and has been regularly updated since [28]. The feature extraction and matching algorithms use minutiae points and other non specified proprietary algorithmic solutions, which enhance the performance and reliability of the system. The system is available under different charged licensing possibilities through the Neurotechnology webpage.[1]
- **Quality metric: NFIQ2**. The development of NFIQ2 was driven by the advances in fingerprint quality estimation since the original version of NFIQ was published in 2004 [29]. It was initiated in 2011 by the US NIST, who leaded a team of different partners coming from law-enforcement and research. The major differences in comparison with the original NFIQ are: 1) modular design; 2) possibility to be retrained to adapt to specific contexts (e.g., latent fingerprints); 3) increased speed; 4) increased accuracy in the estimation of fingerprint quality; 5) increased sensitivity range to 0-100. Furthermore, NFIQ2 quality features are being formally standardized as part of ISO/IEC 29794-4 Biometric Sample Quality [30]. Alike the original NFIQ, NFIQ2 is also supplied as an open-source platform through the NIST portal,[2]

  It should be noted that the distributable version of NIFQ2 has been pre-trained using solely: adult fingerprint data, acquired with live-scan optical sensors at 500 dpi. As such, if the test data differs significantly from these characteristics, results can be inaccurate.

The experimental protocol described in the following subsections for VeriFinger and NFIQ2, was replicated for another matcher (NIST NBIS) and two other quality metrics (NFIQ1 and VERIQ). Due to limitations of space, the results obtained with those tools, as well as their description, can be consulted in Annex A, provided as accompanying material of the present article. In brief, those results provide further confirmation of the findings and conclusions extracted in the main text.

### A. Experimental Protocol: Age Effect

The objective of this set of experiments is to determine if the age of the individual can play a role in the performance of biometric systems. The age effect is studied from two linked points of view: 1) influence on the fingerprint image quality and 2) influence on the system accuracy.

Since biometric quality and biometric accuracy are closely interdependent, the goal of the matching experiments is to

[1] http://www.neurotechnology.com/

[2] https://www.nist.gov/services-resources/software/development-nfiq-20

determine to what extent quality metrics are capable of reflecting the variations in the accuracy of fingerprint recognition systems due to age.

*1) Age Effect: Quality Experiments:* The quality scores of all the 421,388 samples present in the experimental dataset are extracted using the NFIQ2 metric.

Quality distributions for each of the three main groups (i.e., children, adults and elderly) are computed as well as for each of the children and elderly sub-groups.

The mean quality value per age of acquisition is also computed (i.e, mean quality value for ages 0-25 and 65-98).

*2) Age Effect: Matching Experiments:* The accuracy of the systems will be evaluated based on the DET (Detection Error Trade-off) curves for the VeriFinger matcher. In order to extract these curves two sets of matching scores are needed, commonly referred to in the biometric literature as genuine and impostor[3]:

- *Genuine scores*. To compute these scores, only those fingers with two samples in the database are considered (see Table I). The age of the pair is determined by the first sample. Genuine scores are produced by matching the reference sample of each finger (first sample) to its respective probe sample (second sample).
  In order to analyse only the age effect, dissociating it to the largest extent possible from the ageing effect, it is preferable to use fingerprint pairs with the smallest time difference between the reference and probe samples. However, taking only pairs (reference-probe) that were acquired, for instance, on the same year, would reduce drastically the available data and reduce the statistical relevance of the results. As such, a compromise had to be reached between: 1) the temporal proximity of the reference and probe samples and 2) the amount of available data. Following this necessary compromise, 87,011 fingerprint pairs (i.e., genuine matching scores) were selected for the experiments (47,782 pairs coming from children, 33,725 from adults and 5,504 from elders).
- *Impostor scores*. The population of impostors is taken from those fingers that have just one sample in the database (see Table I). One impostor fingerprint, the one with the highest quality value, is selected for each age between 0-25 and between 65-90, which results in a total of 52 impostor fingerprints. Impostor scores are computed matching the 52 impostor fingerprints to the each of the reference samples used to compute the genuine scores. This way, the number of impostor scores is 52 times the number of genuine scores.
  Please note that impostor scores are needed in order to properly evaluate the accuracy of fingerprint recognition systems. However, both the age and ageing effects are intrinsically linked to genuine scores (as explained at the beginning of Sect. IV). Therefore, in order to avoid that the two analyzed effects are concealed due to uncontrolled changes in the impostor scores, the protocol has been designed in order to minimize eventual variability factors that can affect the impostor score distribution. That is why, in the present study: 1) the same 52 impostor samples are used in all scenarios; 2) impostors of all ages are equally represented (one fingerprint per age); 3) the highest quality fingerprints for each age are selected in order to reduce the potential effect that quality may have on the impostor score distribution, as a certain correlation (much lower than in the case of genuine scores) has been reported in some works [31].

[3]Defined as *mated* and *non-mated* matching scores in the Harmonized Biometric Vocabulary contained in the standard ISO/IEC 2382-37

Finally, the mean genuine matching score value per age of acquisition is also computed (i.e, mean genuine score value for ages 0-25 and 65-98).

The full age-effect protocol, including both the quality and matching experiments, is depicted in Fig. 3. Results are presented in Sect. V-A.

### *B. Experimental Protocol: Ageing Effect*

The objective of these experiments is threefold: 1) determine if ageing has an effect on the accuracy (DET curves) of fingerprint recognition systems for the time gap represented in the experimental database between the reference and probe samples (i.e., 7 years); 2) estimate the variation of the genuine matching scores (GMS) distribution when the time difference between the reference and the probe fingerprints increases; and 3) determine whether the variation in accuracy and in the GMS differs depending on the age of the individual at the enrolment of the reference sample.

The experiments are carried out on all 156,067 fingers with two samples in the dataset. The fingerprint impression acquired at a younger age is used as the reference sample and the one captured at an older age as the probe sample. Genuine scores are produced by matching the reference sample of each finger to its respective probe sample.

As in the case of the age-effect experiments, eight different age-groups are considered and each finger is assigned to one of the groups according to the age at which the first fingerprint sample was enrolled to the system: children1 0-4, children2 5-12, children3 13-17, adults 18-25, elderly1 65-69, elderly2 70-74, elderly3 75-79 and elderly4 80-98.

Each of these eight age-groups is then further divided into eight sub-groups according to the time difference between the reference and the probe sample: 0-1 years, 1-2 years, 2-3 years, 3-4 years, 4-5 years, 5-6 years, 6-7 years and 7-8 years. Accordingly, a total of $8 \times 8 = 64$ sub-sets are considered.

The result of the complete set of experiments is constituted by 156,067 genuine scores. These scores are divided among each of the 64 sub-sets as specified in Table II. Impostor scores for the 64 sub-sets are computed as described for the age experiments in Sect. IV-A.

A diagram summarizing the experimental protocol for the ageing experiments is shown in Fig. 4. The main results are presented in Sect. V-B.

## V. RESULTS

This section presents the results that have been obtained following the protocol defined in Sect. IV. Throughout this section the reader will find a number of FINDINGS and HYPOTHESES that are derived from the experiments.

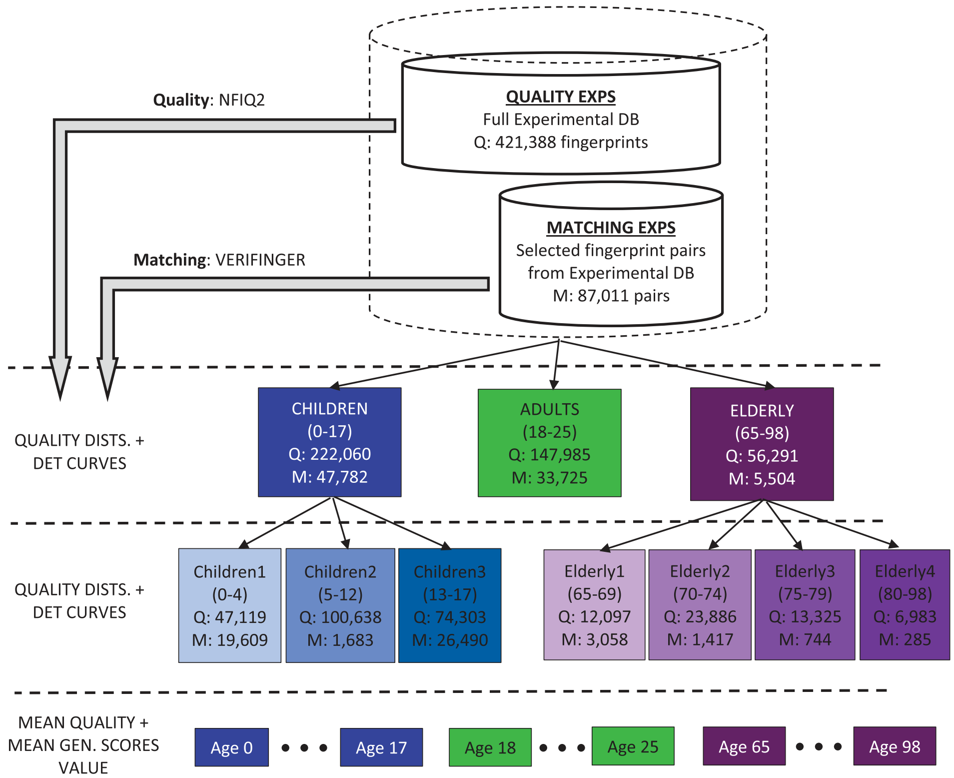


Fig. 3. Diagram depicting the protocol followed to analyse the age effect. The evaluation metrics used for the analysis of the age effect are specified to the left of the age groups (i.e., quality distributions, DET curves, mean value of quality scores genuine matching scores). The figures after 'Q:' and 'M:' indicate respectively, the number of quality scores and the number of genuine matching scores computed for each age-group. This same protocol was replicated also for the NFIQ1 and VERIQ quality metrics and results are presented in Annex A (provided as accompanying material of this article).

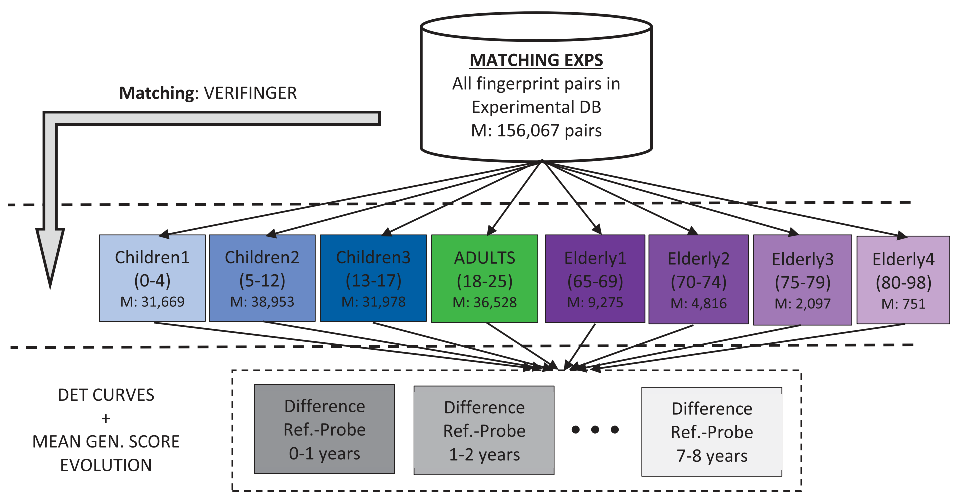


Fig. 4. Diagram depicting the protocol followed to analyse the ageing effect. The evaluation metrics used for the analysis of the ageing effect are specified to the left of the age groups (i.e., DET curves, mean value of genuine scores). The figures after 'M:' indicate the number of fingerprint pairs (i.e., genuine scores) for each age group. See Table II for the number of fingerprint pairs available for each time separation (0, 1, 2,…7 years) for the different age-groups. This same protocol was replicated also for the NIST matcher and results are presented in Annex A (provided as accompanying material of this article).

By FINDINGS we refer to observations for which some level of support is provided by the results obtained, even if, in some cases, further experimentation is required in order to fully confirm them. On the other hand, we use the term HYPOTHESIS to refer to a reasonable conjecture based on the results presented in this article, but for which no experiments have been carried out either to confirm or invalidate it.

### *A. Results: Age Effect*

The results presented in this section have been obtained following the experimental protocol described in Sect. IV-A (see Fig. 3).

*1) Results Age Effect: Quality Experiments:* Fig. 5 shows in the top chart the comparison of the NFIQ2 fingerprint quality distributions corresponding to the three overall age groups in

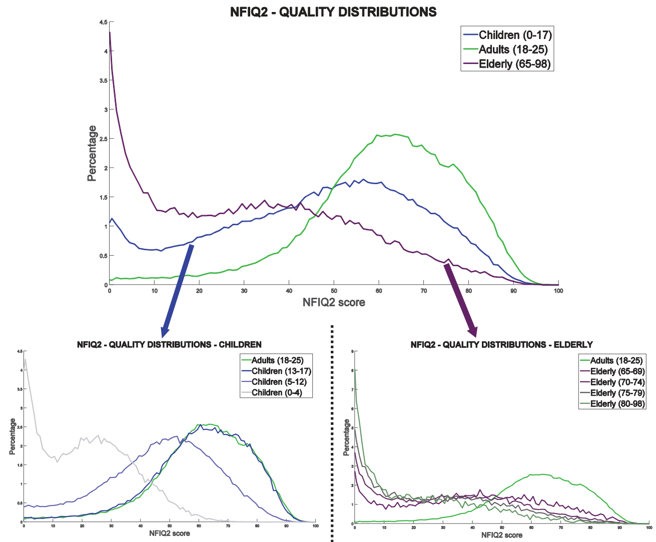


Fig. 5. (Top) NFIQ2 quality distributions of the fingerprints belonging to each of the three main age-groups represented in the experimental dataset: children (0-17), adults (18-25) and elderly (65-98). (Bottom left) Quality distributions of the three children sub-groups: Children1 (0-4), children2 (5-12), children3 (13-17). (Bottom right) Quality distributions corresponding to the four elderly sub-groups: elderly1 65-69, elderly2 70-74, elderly3 75-79 and elderly4 80-98. In the two bottom charts the adults quality distribution is also given for reference. These same distributions for the NFIQ1 and VERIQ quality metrics are shown in Figs. 1 and 2 of Annex A (provided as accompanying material of the present article).

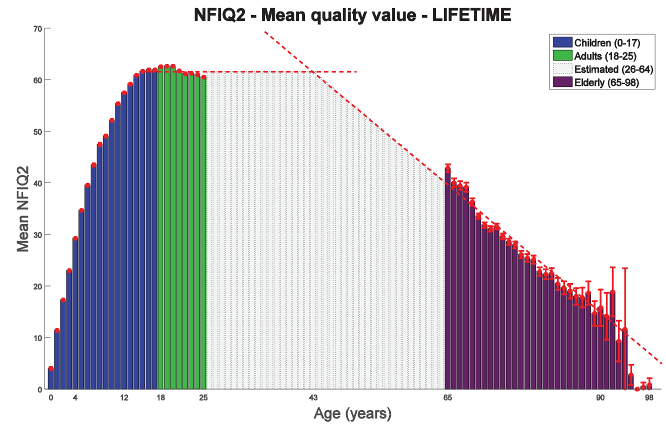


Fig. 6. Lifetime evolution of the NFIQ2 mean quality value. The 90% confidence intervals are shown in vertical red lines. Values for ages 26-64 have been estimated using two linear fits (shown with dashed red lines). Analogous plots for the NFIQ1 and VERIQ quality metrics are shown in Figs. 3 and 4 of Annex A (provided as accompanying material of the present article).

the experimental dataset: children, adults and elderly. The children and elderly distributions have been further subdivided in the bottom charts. The left chart shows the quality distributions corresponding to the three children sub-groups, while the right chart shows the quality distribution of the four elderly sub-groups. In the two bottom charts the adults distribution is also given for reference (in green).

Given that the fingerprints in the experimental dataset are not uniformly distributed age-wise (see Fig. 1), the quality distributions shown in Fig. 5 should not be taken as a perfect reflection of reality. However, given the amount of data considered, these distributions do reflect the general trend that can be expected from fingerprint data in these large three age-groups (children, adults and elderly). As such, we believe it is safe to extract the next conclusion from the results shown in Fig. 5:

- ***FINDING 1***. In terms of general fingerprint quality (see Fig. 5 top), the most challenging age-group is the elderly (65 years of age and above), which presents an overall quality significantly lower than that of children (0-17 years of age). As could be expected, adults clearly present the highest fingerprint quality.
- ***FINDING 2***. For children (see Fig. 5 bottom left), clearly the most problematic group is 0-4. For ages 5-12 fingerprint quality is already acceptable, while for 13-17 it is equivalent to that of adults.
- ***FINDING 3***. For the elderly (see Fig. 5 bottom right), there is a gradual degradation of fingerprint quality from group 65-69 to group 81-98. However, unlike children, where a big difference in fingerprint quality could be seen among groups, in the case of the elderly, for all four groups the quality level is similarly low (in between that of children 0-4 and 5-12).

Fig. 6 shows the year-by-year evolution of the mean fingerprint image quality in the experimental dataset. The 90% confidence intervals for the mean values are shown as vertical red lines.

For those ages not present in the dataset, that is, ages between 26 and 64 (plotted in light grey in Fig. 6), the mean fingerprint quality has been estimated using values for ages 18-25 and 65-90 as described below. In the description, $X$ represents the theoretical age at which fingerprint quality starts decreasing from adulthood to old-age:

- **Mean quality estimation: ages 26-$X$**. The estimation has been done following the hypothesis that during adult life, fingerprint quality does not vary significantly. Given that only eight mean quality values are available for adults (ages 18-25), the mean quality for ages 26-$X$ has been estimated as the average of the mean quality values for 18-25. This estimate corresponds to the horizontal red dashed line in Fig. 6.
- **Mean quality estimation: ages $X$-26**. For ages $X$-64, the mean quality values have been estimated with a linear regression fit using the mean quality values from ages 65-90. Mean quality values corresponding to ages 91-98 have not been considered due to the low amount of fingerprint impressions available in the database. This estimate corresponds to the diagonal red dashed line in Fig. 6.

The age $X$ is defined by the intersection of both linear fits intersect (i.e., horizontal dashed red line and diagonal dashed red line). As mentioned above, it represents an estimation of the point in the fingerprint lifetime at which its quality starts degrading after peaking during adulthood.

The age-wise evolution of fingerprint quality shown in Fig. 6 allows us to conclude that:

- ***FINDING 4***. Quality of children fingerprint impressions increases between 0 and 12 years of age. This increase is very fast between 0 and 4 years of age while it reduces its rate between 5 and 12. From 12 years old until 17, fingerprint quality stabilizes and can be considered equal to that of adults (18-25).

- ***FINDING 5***. For adults, fingerprint quality is quite stable, with an almost negligible decreasing slope between 18 and 25 years. Given the limited amount of data available for adults from an age-wise perspective, covering only ages 18-25, this invariable behaviour of fingerprint quality should still be confirmed.
- ***FINDING 6***. For elders in the range 65-90, fingerprint image quality decreases linearly with age. According to the estimation made in the study, this linear decrease starts at around $X = 40$-45 years of age. It is interesting to underline that for subjects 70 years old, fingerprint quality is equivalent to that of 4-5 years old children.

*2) Results Age Effect: Matching Experiments:* As mentioned in the description of the experimental protocol in Sect. IV-A, the matching tests were performed to confirm, or to complement if necessary, the observations made in the quality-related results presented above.

Matching results have been obtained on approximately one fifth of the data of the quality results as explained in the general experimental protocol in Sect. IV-A: 421,388 fingerprint samples for the quality experiments with respect to 87,011 fingerprint pairs for the matching experiments (see Fig. 3). This means that, from a statistical perspective, matching results are somewhat less reliable (as will be shown in the results by the larger 90% confidence intervals). However, we believe that the amount of data remains significant and offers the possibility to extract valid conclusions.

Fig. 7 (top) shows the comparison of the DET curves for the three main age groups in the experimental dataset: children, adults and elderly (please see Sect. IV-A for a description of the computation of the genuine and impostor score distributions). As in the case of quality, given the non-uniformly age distribution of the experimental dataset, these DET curves should be taken as a general indication of matching performance and not as a perfect representation of reality.

The FINDING 1 extracted from the quality experiments is not fully confirmed by the matching error rates. While children presented a better overall quality than the elderly, results presented in Fig. 7 show that:

- ***FINDING 7***. Fingerprint impressions of the elderly perform, in general (Fig. 7 top), better than those of children. As such, better quality in this case does not directly translate into better accuracy.
- ***FINDING 8***. The FINDING 2 of the quality experiments is confirmed by the accuracy results shown for the different children sub-groups in Fig. 7 (bottom left). The worst overall performing age-group (including the elderly) are children between 0 and 4. Children 5-12 present acceptable error rates, while children 13-17 can be considered as adults in terms of fingerprint accuracy.
- ***FINDING 9***. Regarding the accuracy of the four elderly sub-groups (shown in Fig. 7, bottom right), it can be seen that, as happened with quality (FINDING 3), there is a gradual deteriorations from group 65-69 to group 80-98. Interestingly, the performance of elderly 65-69 is almost equal to children 5-12. The accuracy of the other three elderly groups (70-74, 75-79 and 80-98), is in between children 5-12 and children 0-4.

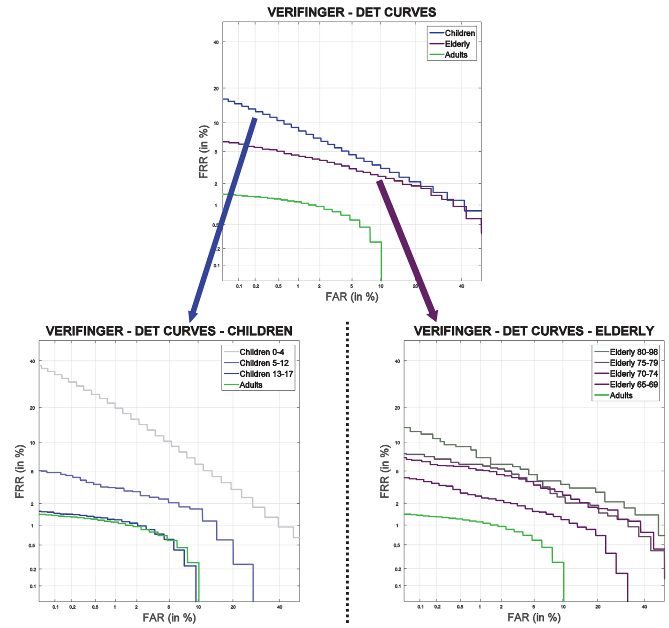


Fig. 7. (Top) DET curves for the three main age-groups represented in the experimental dataset: children (0-17), adults (18-25) and elderly (65-98). (Bottom left) DET curves for the three children sub-groups: Children1 (0-4), children2 (5-12), children3 (13-17). (Bottom right) DET curves corresponding to the four elderly sub-groups: elderly1 65-69, elderly2 70-74, elderly3 75-79 and elderly4 80-98. In the two bottom charts the adults DET is also given for reference (in green). An analogous plot for the NIST matcher is presented in Fig. 5 of Annex A.

This apparent incongruity between FINDING 1 (quality) and FINDING 7 (matching) may have two possible explanations derived from the type of data commonly used to train and test fingerprint matching algorithms and quality metrics:

- As explained in Sect. IV, the NFIQ2 quality metric used in this study was exclusively trained on adults data. This is the case for the vast majority of quality metrics proposed in the literature. Accordingly quality metrics designed for adult fingerprints may be inaccurate when predicting the matching performance of children data. Depending on the age range of the adults fingerprints used for their training, the discrepancy between quality scores and matching scores could also be applicable to elderly fingerprints (e.g., if training data does not take into consideration fingerprints above, for instance, 50 years of age).
- Similarly to quality metrics, fingerprint matching algorithms are typically trained and tested on adults data. As such, they may be inefficient at exploiting the discriminative information conveyed by children's fingerprints even if these are of sufficient quality.

Following the quality experiments, Fig. 8 shows the year-by-year evolution of the mean genuine matching score value. The 90% confidence intervals for each of the mean values are

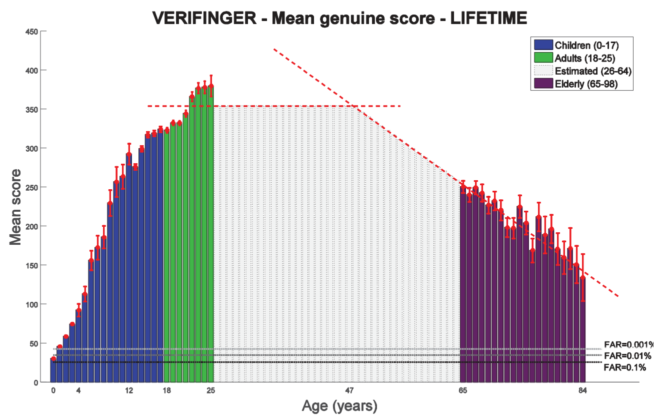


Fig. 8. Lifetime evolution of the mean GMS value. The 90% confidence intervals are shown as vertical red lines. Values for ages 26-64 have been estimated using two linear fits (shown with dashed red lines). The thresholds for FAR = [0.1%, 0.01%, 0.001%] computed on the adults data are given for reference as dotted horizontal lines. An analogous plot for the NIST matcher is presented in Fig. 6 of Annex A.

shown as vertical red lines. For some ages, these confidence intervals are noticeable since they have been obtained on around a fifth of the data used in the quality-related protocol (see the specific numbers under 'Q:' and 'M:' in Fig. 3). Even if their statistical reliability is smaller, they do help to show the overall trends of fingerprint matching performance with respect to age and they complement what was observed in the quality-based experiments. The thresholds for FAR = [0.1%, 0.01%, 0.001%] computed on the adults data are given for reference as dotted horizontal lines.

For those ages not present in the experimental dataset, that is, ages between 26 and 64 (plotted in light grey in Fig. 8), the mean genuine matching scores (GMS) have been estimated following the same process as in the quality experiments. That is: A) the horizontal red dashed line represents the linear estimate for ages 26-$X$ and it has been computed as the average of the mean GMS values for ages 18-25; B) the diagonal red dashed line represent the linear estimate for ages $X$-64 and it has been computed as the linear regression fit of the mean GMS values for ages 65-85 (mean GMS values corresponding to ages 85-98 have not been considered due to the insufficient quantity of GMS available for those ages). In the description above, $X$ represents the estimated age at which GMS start degrading from adulthood to old-age. For further details on the rationale to use these linear fits please see Sect. V-A.1.

The matching results shown in Fig. 8 are consistent with the equivalent quality-related results presented in Fig. 6. This way, the conclusions drawn from the quality experiments are confirmed with small variations:

- ***FINDING 10***. Genuine matching scores of children increase between 0 and 18 years of age. This increase is linear and very rapid between 0 and 12 years of age while it considerably reduces its rate between 12 and 17.
- ***FINDING 11***. For adults, although a certain increasing trend can be observed between 18 and 25 years of age, considering the range of the 90% confidence intervals, it is not possible to confirm such improvement. Rather, based on FINDING 5 of the quality experiments, it is more reasonable to assume that matching scores during adulthood should be fairly constant. However, given the limited amount of data available for adults from an age-wise perspective, covering ages 18-25 (i.e., eight points), this assumption regarding the stable behaviour of fingerprint genuine matching performance for adults should still be confirmed on a set of data covering the age range 25-64.
- ***FINDING 12***. For elderly in the range 65-84, fingerprint genuine matching scores decrease linearly with age. According to the estimation made in the study, this linear decrease starts at around 40-45 years of age (which is consistent with the estimation made in the quality experiments). The mean value of genuine matching scores of 70-year olds is similar to that of children close to 5 years old (as was already observed in FINDING 6 of the quality experiments).

The results presented in this section, summarized in FINDINGS 1-12, have shown the big challenge posed to fingerprint recognition systems by very young children (0-4) and, to a lesser extent, also by the elderly (especially above 70). Based on these findings and on previous experience gained in the field of fingerprint biometrics, we present here plausible explanations for this poor performance and we put forward two hypotheses on how to improve the interaction of these problematic age-groups with fingerprint-based technology.

The size of fingerprints and the frequency (width) of ridges and valleys are two of the major parameters that are taken into account in the development of quality metrics and feature extraction algorithms. These parameters are typically adapted to the average size and ridge width of adults fingerprints. As such, the small overall size and narrow ridge structure of fingerprints belonging to very young children (0-4 years of age) is likely to be one of the main reasons for their low quality and poor matching performance.

- ***HYPOTHESIS 1***. Developing specific quality metrics and matching algorithms adapted to the reduced size of children fingerprints could significantly improve both their image quality scores and their overall accuracy.

Following the previous hypothesis, some vendors already include a *juvenile option* in their recognition systems in order to adapt certain parameters of the embedded algorithms to the specific size particularities of children fingerprints. However, still further research is required to fully assess with experimental results the improvement offered by these children-tailored solutions.

For the elderly, as for adults, fingerprints size and ridge width remains basically invariable with age. However, the skin condition changes as we grow older, gradually losing its elasticity, firmness and becoming drier, mostly due to the decrease of collagen [32]. These variations, together with other possible medical sufferings typical of old age such as arthritis, hinder the acquisition of fingerprints with current live-scan touch-based scanners, which entails a decrease in their overall image quality.

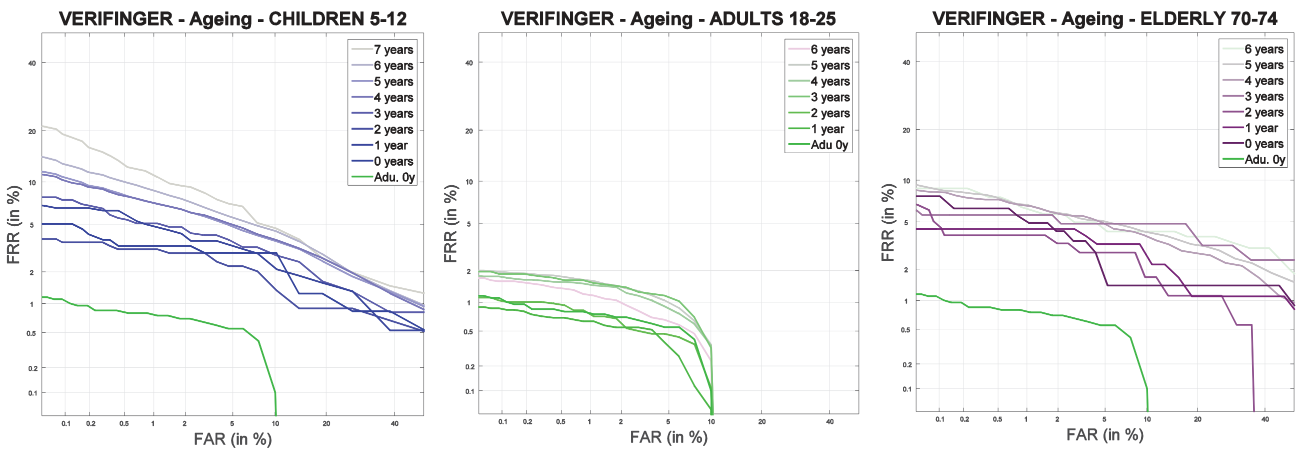


Fig. 9. Evolution of the DET curves (i.e., system accuracy) as the time difference between the reference and the probe samples increases from 0 years to 7 years. The plots are given for: children 5-12 (left), adults 18-25 (center) and elderly 70-74 (right). Adults DET curve for a time difference of 0 years between reference and probe (Adu. 0y) is provided as reference in the children and elderly plots (in green). Find a similar plot but for the NIST matcher in Fig. 7 of Annex A.

- ***HYPOTHESIS 2***. From a technological perspective, new touchless acquisition devices could improve the quality and, therefore, the matching performance of elderly fingerprints.

From a pure procedural perspective, with current touch-based technology, moisturizing the fingertip skin prior to the acquisition can also help to obtain images with better quality (improving this way the matching scores).

### *B. Results: Ageing Effect*

The results presented in this section have been obtained following the experimental protocol described in Sect. IV-B (see Fig. 4).

Fig. 9 shows the evolution of the system accuracy in terms of the DET curves when the time difference between the reference and probe samples increases from 0 to 7 years. As illustrative examples, results are given for age groups: children2 (5-12), adults (18-25) and elderly2 (70-74).

From these plots it seems that the largest ageing effect happens for the children group, where the DET curves present a gradual degradation from 0 years (darkest shade of blue) to 7 years (lightest shade of blue). For adults and elderly, the effect is not as clear, although it does appear that lighter shades of the DET curves tend to be higher up in the plot (larger error rates).

The DET plots presented in Fig. 9 are enough to show that the accuracy of systems can suffer a certain degradation, especially for the case of children, in the 7-year time-gap considered. However, based on these accuracy metrics it is difficult to quantify, either visually or numerically, the ageing effect. In addition, the DET curves do not allow for a clear comparison of ageing among age-groups.

For the reasons expressed above, DET curves are not the best suited tools for the analysis of ageing. As explained in the introduction of Sect. IV, ageing is a phenomenon mostly related to genuine matching scores. Therefore, the fairly consistent False Acceptance and False Rejection error rates (FAR and FRR) shown in Fig. 9 for adults and elderly, does not necessarily imply that there is no ageing, as the genuine score distribution may have started to drift towards the impostor score distribution which, eventually, will result in a decrease of the system accuracy.

According to this rationale, also applied in previous studies [2], in the following ageing is further analysed based on the variation of the mean value of the genuine matching score distributions.

Fig. 10 shows the evolution of the mean genuine scores when the time difference between the reference and probe samples increases from 0 to 7 years. Results are given for age group categories: 1) Top row: children1 (0-4), children2 (5-12), children3 (13-17) 2) elderly1 (65-69), elderly2 (70-74), elderly3 (75-79), elderly4 (80-98). Adults (18-25), is plotted in green in both cases as reference.

The left plots in Fig. 10 show the evolution of the mean absolute values. The right plots show the normalized mean values. The normalization is such that, for all age groups, the mean genuine score for a time difference of 0-1 years represents 100%. This way it is possible to visualize the variation in percentage of the mean value, where a steeper slope implies a larger ageing effect. For each point, the 90% confidence interval is given as a vertical bar.

The graphs given in Fig. 10 allow for a better analysis of ageing than the DET curves of Fig. 9. The main conclusions that may be extracted from the children results (top row) are:

- ***FINDING 13***. Looking at the absolute values of the genuine matching scores (i.e., top row left in Fig. 10), the results obtained in the age-effect experiments presented in Sect. V-A are confirmed: children fingerprints in the range 0-4 show lower genuine matching scores (GMS) than children fingerprints in the range 5-12, while GMS for children 13-17 and adults are very similar.
- ***FINDING 14***. For all groups, a larger time difference between the reference and probe samples implies a decrease in the performance of the genuine matching scores. The ageing effect is therefore confirmed. This loss of matching performance is (see Fig. 10, top row, right plot): 1) around 15% in the case of

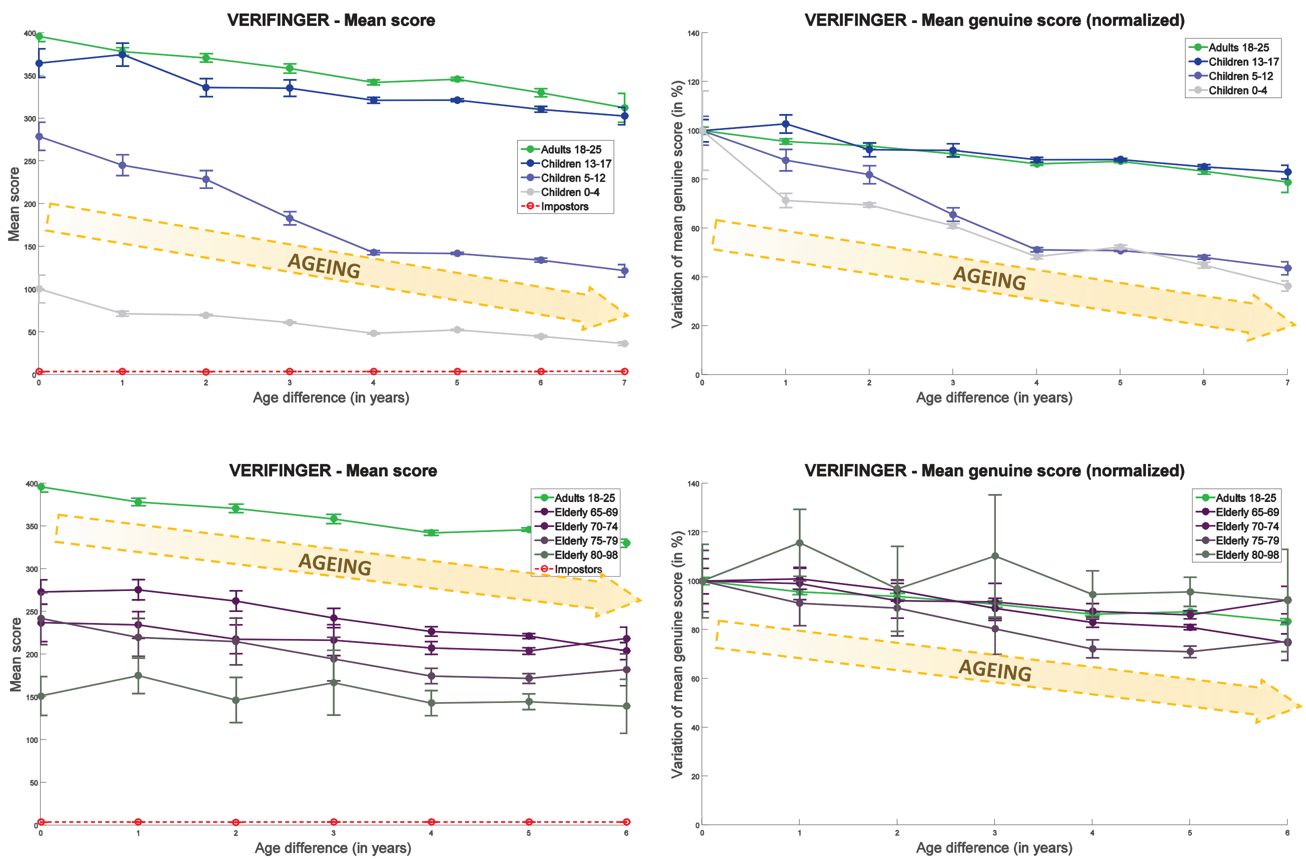


Fig. 10. Evolution of the mean genuine score as the time difference between the reference and the probe samples increases from 0 years to 7 years. The right column shows the same results as the left, but normalized so that the first point represents in all cases 100%. The plots are given for: 1) the three children sub-groups (top row); 2) the four elderly sub-groups (bottom row). Adults are given as reference (in green). The evolution of the mean impostor score is also plotted in red as reference (left column). Fig. 8 in Annex A shows the same results obtained with the NIST matcher.

adults and children 13-17; 2) around 50% in the case of children 5-12 and 0-4. Note that, given the 90% confidence intervals, this loss has a tolerance of around $\pm 3\%$.

- ***FINDING 15***. For adults and children 13-17, the total 15% decrease in the GMS is almost linear between 0 and 7-years difference, i.e., there is around a 2% loss with each additional year between the reference and probe samples. On the contrary, for children 0-4 and 5-12, the biggest ageing effect occurs when the time difference between the reference and probe samples increases from 2 to 4 years (steepest slope in the right plots). In this 2-year gap there is a 30%-40% loss in the GMS (out of the total 50% over 7 years). Note that, given the 90% confidence intervals, such performance loss has a tolerance of around $\pm 5 - \pm 16\%$, depending on the age group.

The following set of conclusions may be drawn from the results presented in Fig. 10 for the elderly (bottom row):

- ***FINDING 16***. Looking at the absolute GMS values, i.e., left plot, the results obtained in the age-effect experiments presented in Sect. V-A are confirmed: elderly fingerprints perform worse as the age of the reference template increases, that is, age group 65-69 performs better than 70-74, which performs better than 75-79, which performs better than 80-98.
- ***FINDING 17***. For all groups, a larger time difference between the reference and probe samples implies a loss in the GMS mean value. Therefore, ageing is confirmed. This decrease of the genuine matching scores is very similar for all groups around 15% (similar to that of adults).
- ***FINDING 18***. For all groups, the total 15% GMS decrease is almost linear between 0 and 6-years difference, i.e., there is around a 2-3% loss with each additional year difference between the reference and probe samples (similar to adults).

In summary, it can be stated that ageing between 13 and 98 years of age happens in a very similar way, while this effect is significantly larger between 0 and 12 years.

As a coarse way of quantifying together age and ageing, Fig. 11 shows the estimation of the time difference required in order for the mean value of the genuine scores to become equal to the mean value of the impostor scores, assuming a linear degradation of the genuine scores for all age groups. The thresholds for FAR = [0.1%, 0.01%, 0.001%] computed on adults data are also given for reference. This figure can be interpreted as a conjecture of how the curves given in the left column of Fig. 10 may evolve with time. As can be seen, for children 0-12, in around 10 years time from the acquisition of the reference sample, the system would become

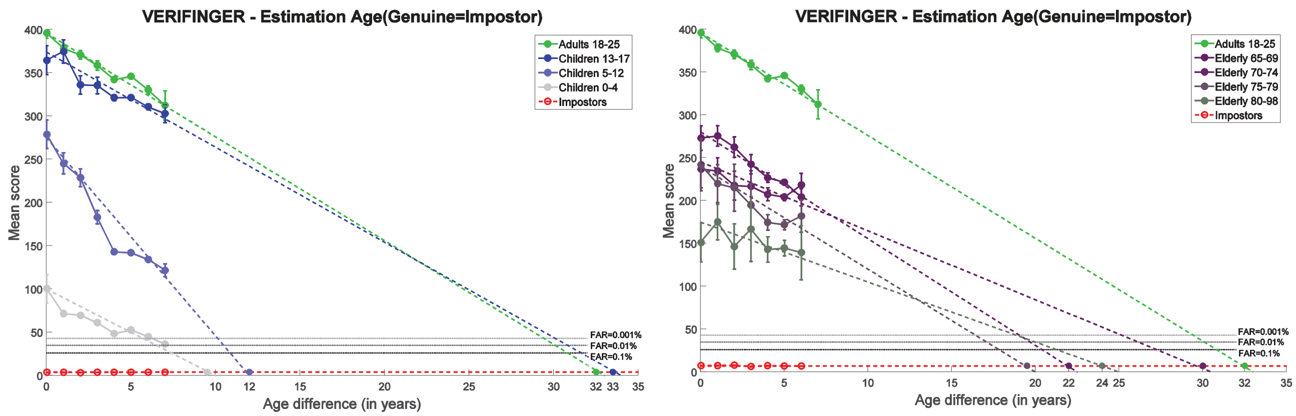


Fig. 11. Estimation of the time difference in which the mean genuine score value eventually becomes equal to the mean impostor score value, assuming linear ageing over time for all age groups. The thresholds for FAR = [0.1%, 0.01%, 0.001%] computed on adults data are given for reference. An analogous plot is shown in Fig. 9 of Annex A.

completely unusable (i.e., overlapped genuine and impostor scores distributions). This time gap increases to 20-30 years for the elderly groups and goes beyond 30 years for subjects of ages between 13 and 25 (adults). These figures can be useful in order to define a policy for template update depending on the age of the user (e.g., validity period for passports).

As mentioned above, the previous findings have shown that the population segment most affected by ageing are children between 0 and 12 years of age. This is the age range where individuals grow at the fastest rate. From our perspective, the displacement of the minutiae points due to this rapid growth is the most probable cause for the larger ageing in this group. Based on this rationale, we can state that:

- ***HYPOTHESIS 3***. From an algorithmic perspective, the development of a reliable growth model for fingerprints between 0 and 12 years could help to predict the new position of minutiae points and other discriminative features at a certain point in the future, with respect to the reference template, helping this way to reduce the ageing effect in young children.

An early work has already provided some initial experimental support to the previous hypothesis, under the assumption of the isotropic displacement of minutiae points [33].

From a pure procedural perspective, ageing can also be prevented by reducing the validity of the reference templates (e.g., in the case of travel documents this would entail a shorter expiry period).

## VI. Conclusions

According to folk wisdom "nothing is immutable" (except for death and taxes).[4] If that is so, can we trust biometrics as a mean for personal authentication?

The present article has addressed this difficult issue in the field of fingerprint recognition, presenting some new insights into the way time affects fingerprint-based technology. The main goal has been to produce valuable results that can help researchers, vendors and users to further understand the level of reliability of automatic fingerprint recognition systems depending on the age of the subject and the time difference between the reference and probe samples.

To reach this objective, we have used a unique database of over 400K fingerprints which contains fingers ranging between 0-25 years and between 65-98 years, with a time difference between samples of the same finger of 0 to 7 years. This dataset has allowed us to study the effect of time on fingerprint recognition systems from two linked perspectives: age and ageing. These two effects have been evaluated considering both fingerprint quality and fingerprint matching.

The analysis of the results has generated a number of findings highlighted throughout the text. These findings are summarised in the next set of wrap-up conclusions which either: 1) confirm similar results reached in previous works (usually over significantly less amount of data); 2) challenge conclusions reached in previous works; 3) constitute new knowledge in the field.

- ***CONCLUSION 1***. From a quality point of view, children fingerprint impressions show better quality than those of the elderly. However, from an accuracy perspective, elderly fingerprint images show somewhat lower error rates than those of children. Both from a quality and a matching perspective adults fingerprints are clearly those that present the best behaviour.
  This comparison of the three main age-groups (i.e., children, adults and elderly) from a quality and matching perspective is a new contribution from this work. However, some similar trends had already been pointed out in previous low-scale works considering only children [7], [23], [24].
- ***CONCLUSION 2***. Fingerprints quality and genuine matching scores: 1) increase very rapidly between 0 and 12 years of age, where they stabilise; 2) both remain fairly constant during adulthood until 40-45 years of age; 3) at 40-45 both start to decrease linearly. Please recall that, both the stable behaviour during adulthood and the age at which the scores begin to decrease linearly, are

[4] Paraphrase of a famous quote usually attributed to Benjamin Franklin: "*In this world, nothing can be said to be certain except death and taxes*"

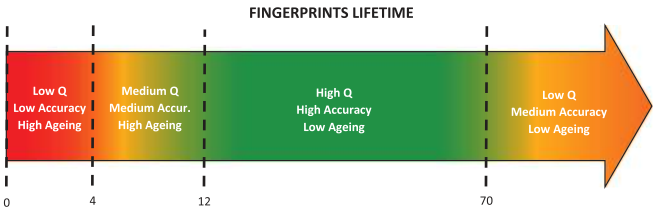


Fig. 12. Diagram showing the different age zones in which the fingerprints lifetime can be divided according to their quality/matching and the ageing effect (following the results presented in Sect. V). Numbers indicate age in years.

estimations that still need to be confirmed on a database with fingerprint images in the age range 26-64.

This result challenges previous conclusions presented in [2] where it was claimed that both genuine matching scores and fingerprint quality continuously decrease between 0 and 80 years of age.

- ***CONCLUSION 3***. Ageing occurs for all age groups: the larger the time difference between the reference and the probe samples, the larger the matching performance loss of the genuine scores. This happens for a time difference as small as 1-2 years.

  This result confirms the conclusions reached in the previous large-scale ageing study [2].
- ***CONCLUSION 4***. Ageing is larger for children whose fingerprint reference sample has been enrolled to the system at 0 to 12 years old. In this age range, for a time difference of 7 years the genuine matching scores decrease by around 50%. For the age range between 13 and 98 years of age, ageing is very similar. It occurs linearly with a drop in genuine matching performance of around 1.5%-3% every increase of 1 year between the reference and probe samples.

  Ageing had been studied for adults in [2], where similar conclusions were reached. However, the present work adds information regarding the comparison of ageing through different age-groups.

The findings reached in the work, summarised in the four previous conclusions, can be used to identify four different age zones for fingerprints, depending on the level of the age effect and the ageing effect. This four age zones are depicted in Fig. 12. Please note that, the ages given as limits between zones, should not be taken as precise and definitive markers but as general guidelines with some tolerance, that can help to comprise the evolution of fingerprints through life.

The results have also led us to put forward a number of hypotheses (highlighted in the text), which give probable explanations to the effects observed in each of these four age zones, at the same time that possible solutions are proposed to reduce these effects. The hypotheses need to be confirmed/refuted through further development and experimentation, opening paths for future research.

The four fingerprint age zones that can be identified thanks to the conclusions of the work are:

- **Very young children, 0-4**. This age-group is the most challenging of all the analysed ones. It is characterized by: 1) poor fingerprint image quality; 2) poor accuracy; 3) a pronounced ageing effect. Specific fingerprint algorithms/procedures could be conceived for this segment of the population.

  As expressed in HYPOTHESIS 1, new quality and feature extraction algorithms may be developed, specifically adapted to the small size of these fingerprints and to their narrow ridges and valleys.

  In addition, following HYPOTHESIS 3, the development of a reliable growth model for the displacement of minutiae points through childhood could be a powerful tool to counteract the effect of ageing.

  From a procedural perspective, shorter validity periods for the reference templates could also be an advisable measure to put in place for this age group.
- **Children, 5-12**. For this age group, while quality and matching performance clearly improve with respect to children 0-4 and get closer to adults, the ageing effect is still significantly higher. Therefore, analogue measures to those described in HYPOTHESIS 3 for very young children (0-4) could be followed to minimise this effect.
- **Teenagers, adults and young-elders, 13-69**. For this population segment it can be safely stated that fingerprint recognition systems work, approximately, as evaluated on adults.

  It is true that 40-45 years has been estimated as the age at which both fingerprint quality and genuine matching scores start to linearly degrade. While this degradation will eventually affect the overall accuracy of fingerprint systems, based on the results for elders, we believe that, until approximately 70 years of age, this performance loss will not be significant enough.

  It is also important to notice that, to fully validate the previous statements, further experiments on real data are required in order to accurately model the age and ageing effects for the age range 26-64 (missing in the dataset). Those results could show some variations with respect to the estimations made in the article. However, all the evidence presented in this work indicates that it is unlikely that the behaviour of fingerprints between 26-64 may differ significantly from that of either adults (18-25) or the first elderly group (65-69).
- **Elders, 70+**. The quality degradation of the fingerprint impressions for this part of the population is quite significant, to the point that their quality is the lowest of all age groups considered, including children 0-4. This very low quality is not fully reflected on the accuracy of the systems which is comparable to that of children 5-12. As stated in HYPOTHESIS 2, new touchless acquisition technology could help to improve the fingerprint quality for these users and, therefore, also the final accuracy of systems.

  From a procedural perspective, practical acquisition measures such as moisturizing the skin prior to the scanning can also help to obtain better quality fingerprints for this age group.

Lastly, we would like to highlight that, as pointed out in some of the findings of the study, the elderly can pose a significant challenge to fingerprint recognition systems, comparable, or even bigger, than children in the age range 5-12. This fact can have big practical implications. We should not forget that Europe has stated a commitment to "*the rights of the elderly to lead a life of dignity and independence and to participate in social and cultural life*" [34]. This implies to take the necessary measures to ensure the inclusion of elders in every day life and to guarantee their access to services available to the general population. The results presented in the article have shown that, given the quality deterioration of fingerprints at advanced points in life, there is a potential risk of age-based discrimination against elders due to increased rates of failure-to-capture or failure-to-enrol. We believe that this should be an important issue to be considered in the design of fingerprint recognition systems in order to avoid possible inter-generational inequality [26].

A good practical illustrative example of the situation described above can be found in the field of border management and travel control. Elders are, unlike children, fully autonomous to cross borders and, in general, have the economic resources to do so. Therefore, all automatic systems put in place to supervise and regulate the flow of travelers, such as the ePassport or ABC gates, should take into account the biometric particularities of elders. For instance, the results presented in the article can help to define important policies like: 1) setting different validity periods for travel documents depending on the age of the holder or 2) setting different quality thresholds for fingerprint samples according to the age-group.

## Acknowledgment

The research reported in this paper has been carried out thanks to the invaluable support of the Portuguese Government, and in particular its Passport and Immigration Service (SEF), who generously allowed Unit E.3 from DG JRC to have access to the fingerprint data on which the experiments were performed. The authors gratefully acknowledge the cooperative spirit of all people involved from SEF and thank them for their continuous support and help.

## References

[1] A. Lanitis, "A survey of the effects of aging on biometric identity verification," *Int. J. Biometrics*, vol. 2, no. 1, pp. 34–52, Dec. 2010.

[2] S. Yoon and A. K. Jain, "Longitudinal study of fingerprint recognition," *Proc. Nat. Acad. Sci. USA*, vol. 112, no. 28, pp. 8555–8560, 2015.

[3] S. Pankanti, S. Prabhakar, and A. K. Jain, "On the individuality of fingerprints," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 8, pp. 1010–1025, Aug. 2002.

[4] Y. Zhu, S. C. Dass, and A. K. Jain, "Statistical models for assessing the individuality of fingerprints," *IEEE Trans. Inf. Forensics Security*, vol. 2, no. 3, pp. 391–401, Sep. 2007.

[5] C. Su and S. N. Srihari, "Generative models for fingerprint individuality using ridge models," in *Proc. 19th Int. Conf. Pattern Recognit. (ICPR)*, Dec. 2008, pp. 1–4.

[6] A. Uhl and P. Wild, "Comparing verification performance of kids and adults for fingerprint, palmprint, hand-geometry and digitprint biometrics," in *Proc. IEEE 3rd Int. Conf. Biometrics, Theory, Appl., Syst.*, Sep. 2009, pp. 1–6.

[7] D. Jrc, "Fingerprint recognition for children," Eur. Commission, Brussels, Belgium, Tech. Rep. EUR 26193 EN, 2013.

[8] A. K. Jain, K. Cao, and S. S. Arora, "Recognizing infants and toddlers using fingerprints: Increasing the vaccination coverage," in *Proc. Int. Joint Conf. Biometrics (IJCB)*, 2014, pp. 1–8.

[9] M. Fairhurst, Ed., *Age Factors in Biometric Processing*. Stevenage, U.K.: IET, 2013.

[10] G. Panis, A. Lanitis, N. Tsapatsoulis, and T. F. Cootes, "Overview of research on facial ageing using the FG-NET ageing database," *IET Biometrics*, vol. 5, no. 2, pp. 37–46, May 2016.

[11] P. Grother, J. Matey, E. Tabassi, G. Quinn, and M. Chumakov, "IREX VI: Temporal stability of iris recognition accuracy," U.S. Nat. Inst. Standards Technol., Gaithersburg, MD, USA, Interag. Rep. 7948, 2013.

[12] H. Mehrotra, M. Vatsa, R. Singh, and B. Majhi, "Does iris change over time?" *PLoS ONE*, vol. 8, no. 11, p. e78333, 2013.

[13] A. Uhl and P. Wild, "Experimental evidence of ageing in hand biometrics," in *Proc. Int. Conf. BIOSIG Spec. Interest Group (BIOSIG)*, Sep. 2013, pp. 1–6.

[14] J. Galbally, M. Martinez-Diaz, and J. Fierrez, "Aging in biometrics: An experimental analysis on on-line signature," *PLoS ONE*, vol. 8, no. 7, p. e69897, 2013.

[15] F. Galton, *Finger Prints*. London, U.K.: Macmillan, 1892.

[16] W. J. Herschel, *The Origins of Fingerprinting*. London, U.K.: Oxford Univ. Press, 1916.

[17] M. Arnold, C. Busch, and H. Ihmor, "Investigating performance and impacts on fingerprint recognition systems," in *Proc. 6th Annu. IEEE SMC Inf. Assurance Workshop*, Jun. 2005, pp. 1–7.

[18] S. Kirchgasser and A. Uhl, "Fingerprint template ageing vs. template changes revisited," in *Proc. Int. Conf. Biometrics Spec. Interest Group (BIOSIG)*, Sep. 2017, pp. 1–7.

[19] S. Kirchgasser and A. Uhl, "Template ageing and quality analysis in time-span separated fingerprint data," in *Proc. IEEE Int. Conf. Identity, Secur. Behav. Anal. (ISBA)*, Feb. 2017, pp.1–8.

[20] Dutch Government, "Evaluation report biometrics trial 2b or not 2b," Dutch Ministry Interior Kingdom Relations, The Hague, The Netherlands, Tech. Rep., 2005.

[21] N. C. Sickler and S. J. Elliott, "An evaluation of fingerprint image quality across an elderly population vis-a-vis an 18–25 year old population," in *Proc. 39th Annu. Int. Carnahan Conf. Secur. Technol.*, Oct. 2005, pp. 68–73.

[22] S. K. Modi and S. J. Elliott, "Impact of image quality on performance: Comparison of young and elderly fingerprints," in *Proc. 6th Int. Conf. Recent Adv. Soft Comput. (ICRASC)*, 2006, pp. 445–449.

[23] S. K. Modi, S. J. Elliott, J. Whetsone, and H. Kim, "Impact of age groups on fingerprint recognition performance," in *Proc. IEEE Workshop Automat. Identificat. Adv. Technol.*, Jun. 2007, pp. 19–23.

[24] A. K. Jain, S. S. Arora, K. Cao, L. Best-Rowden, and A. Bhatnagar, "Fingerprint recognition of young children," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 7, pp. 1501–1514, Jul. 2017.

[25] "*The Demographic Future of Europe—from Challenge to Opportunity*," Eur. Commission, Brussels, Belgium, Oct. 2006.

[26] A. P. Rebera and B. Guihen, "Biometrics for an ageing society societal and ethical factors in biometrics and ageing," in *Proc. Int. Conf. Biometrics Spec. Interest Group (BIOSIG)*, Sep. 2012, pp. 1–4.

[27] P. Grother *et al.*, "MINEX: Performance and interoperability of the INCITS 378 fingerprint template," U.S. Nat. Inst. Standards Technol., Gaithersburg, MD, USA, Tech. Rep. NISTIR 7296, 2006.

[28] C. Watson, G. Fiumara, E. Tabassi, S. L. Cheng, P. Flanagan, and W. Salamon, "Fingerprint vendor technology evaluation," NIST, Gaithersburg, MD, USA, Tech. Rep. NISTIR 8034, 2014.

[29] E. Tabassi, C. Wilson, and C. Watson, "Fingerprint image quality," U.S. Nat. Inst. Standards Technol., Gaithersburg, MD, USA, NIST Internal Rep. 7151, 2004.

[30] *Information Technology—Biometric Sample Quality—Part 4: Finger Image Data*, ISO/IEC Standard 29794-4, 2017.

[31] P. Grother and E. Tabassi, "Performance of biometric quality measures," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 4, pp. 531–543, Apr. 2007.

[32] E. Carmeli, H. Patish, and R. Coleman, "The aging hand," *J. Gerontology, Sci.*, vol. 58A, no. 2 pp. M146–M152, Feb. 2003.

[33] C. Gottschlich, T. Hotz, R. Lorenz, S. Bernhardt, M. Hantschel, and A. Munk, "Modeling the growth of fingerprints improves matching for adolescents," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 1165–1169, Sep. 2011.

[34] European Parliament, "Charter of fundamental rights of the European Union," *Off. J. Eur. Communities*, vol. C 364, pp. 364-1–364-22, Dec. 2000.

**Javier Galbally** received the M.Sc. degree in electrical engineering from the Universidad de Cantabria, Spain, in 2005, and the Ph.D. degree in electrical engineering from the Universidad Autónoma de Madrid, Spain, in 2009. He was an Assistant Professor with the Universidad Autónoma de Madrid until 2012. In 2013, he joined the European Commission, DG Joint Research Centre, where he is currently a Scientific/Technical Officer. He is actively involved in different European projects dealing with biometrics and their use in large IT systems. He has authored over 100 publications mainly focused on the biometrics field. His research interests are mainly focused on the security and performance evaluation of biometric systems and pattern and biometric recognition. He was a recipient of a number of distinctions, including the IBM Best Student Paper Award at ICPR 2008, a Finalist of the EBF European Biometric Research Award 2009, the Best Ph.D. Thesis Award by the Universidad Autónoma de Madrid 2010, the Best Poster Award at ICB 2013, the Best Paper Award at ICB 2015, the Best Poster Award at IJCB 2017, and the Best Paper Award at BIOSIG 2018. He is the Chair of the EAB Research Project Conference.

**Rudolf Haraksim** received the M.Sc. degree from the Faculty of Electrical Engineering, Manchester Metropolitan University, U.K., in 2007, and the Ph.D. degree from the Services, Cybersecurity and Safety Group, Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, Enschede, The Netherlands, in 2014, with the support of the Netherlands Forensic Institute (Netherlands Ministry of Justice), The Hague, The Netherlands. In 2014, he joined the Institute of Scientific Police, University of Lausanne (CH), where he was a Post-Doctoral Research Assistant on the topics related to fingerprint biometrics. In 2015, he moved to the Speech Processing and Biometrics Group, Swiss Federal Institute of Technology (EPFL, CH), where he was a Research Assistant, involved in the topics related to face, fingerprint, and voice biometrics. In 2017, he joined the DG Joint Research Centre, European Commission, where he was a Scientific Project Officer on topics related to biometrics and their applications in the fields of safety, security, and migration. Throughout his scientific career, he authored numerous scientific articles published in peer reviewed journals and some of his most recent distinctions include the BIOSIG 2018 Best Paper Award.

**Laurent Beslay** has been a European Commission Official with the DG Joint Research Centre, Directorate E on Space, Security and Migration, since 2011, where he is currently a Scientific Project Leader of law enforcement technologies and the citizen. He manages research activities related to fight against cybercrime and organized crime, biometric systems, and privacy safeguards. From 2004 to 2011, he was a Coordinator on Security and Technology and a Scientific Advisor of the European Data Protection Supervisor, Brussels. He was with the JRC—Institute for Prospective Technological Studies, for six years, as a Ph.D. candidate of electronic surveillance: benefits and risks for the EU, and as a Project Officer in the field of cyber-security.