UNIL | Université de Lausanne
Ecole des sciences criminelles
bâtiment Batochime
CH-1015 Lausanne

Maîtrise universitaire ès Sciences en science forensique
Orientation identification physique

Mémoire de maîtrise

De l'utilisation d'images d'empreintes digitales de synthèse pour tester les performances d'un système AFIS

Alessandro Costa

Sous la direction du Professeur Sébastien Marcel

Juin 2022

Abstract

An automatic biometric recognition system needs large-scale datasets to be trained and benchmarked which involves certain limitations in terms of time, money and privacy. Recent developments in the field of Artificial Intelligence (AI) and, more in details, the successes achieved by the Generative Adversarial Networks (Goodfellow et al., 2014) in the generation of synthetic images offer numerous possibilities trying to solve these constraints. In this study, the *CFG* fully synthetic fingerprint database of Bahmani et al. (2021) has been the subject of the experimentations to test the following hypothesis: (1) the results derived from the evaluation on generated synthetic fingerprints datasets are similar to a real one; (2) the intra- and inter-class variability of a real and a synthetic database are similar. Moreover, further works will develop more in detail the hypothesis (3), according to which, a fully synthetic fingerprint database could be used to train a biometric system (AFIS) instead of using a real fingerprint database.

Keywords: Fingerprints, Biometrics, Generative Adversarial Networks (GAN), Automated Fingerprint Identification System (AFIS), Artificial Intelligence

Résumé

Un système de reconnaissance biométrique automatique a besoin de données à grande échelle pour pouvoir être entraîné et évalué. Ceci implique certaines limitations en termes de temps, d'argent et de confidentialité (privacy). Les récents développements dans le domaine de l'Intelligence Artificielle (IA) et, plus en détail, les succès obtenus par les Generative Adversarial Networks (Goodfellow et al., 2014) dans la génération d'images synthétiques offrent de nombreuses possibilités pour tenter de résoudre ces contraintes. Dans cette étude, la base de données *CFG* d'empreintes digitales entièrement synthétiques, de Bahmani et al. (2021) a fait l'objet d'expérimentations afin de tester les hypothèses suivantes : (1) les résultats d'évaluation des jeux de données d'empreintes digitales synthétiques sont similaires à ceux d'une base de données réelle ; (2) l'intra- et inter-variabilité d'une base de données réelle et d'une base de données synthétique est similaire de sorte que, (3) une base de données d'empreintes digitales entièrement synthétique pourrait être utilisée pour entraîner un système biométrique (AFIS) au lieu d'utiliser une base de données d'empreintes digitales réelle.

Mots clés: Empreintes digitales, Biométrie, Generative Adversarial Networks (GAN), Automated Fingerprint Identification System (AFIS), Intelligence Artificielle

École des Sciences Criminelles
Université de Lausanne
1015 - Lausanne Dorigny

Table of contents

École des Sciences Criminelles
Université de Lausanne
1015 - Lausanne Dorigny

1. Introduction

The recent progresses accomplished in the field of Artificial Intelligence (AI) allow to create fully synthetic images of many different types. This means that nowadays it is possible to generate images of animals, cities, faces, fingerprints, etc. that do not exist and could be tricky to be distinguished from the reality. For instance, images (and art) could also be created just from a description in natural language (Ramesh et al., 2022). Image 1, here below, represents an example of an image generated just from a description, using Artificial Intelligence:

***Image 1**: example of an image created with DALL-E 2[1] AI system from the description "a close up of a hand palm with leaves growing from it" (Illustration: Ramesh et al., 2022)*

Now, the synthesis of images could also be used to solve one big challenge in the biometrics field. In fact, an automatic biometric recognition system needs large-scale datasets to be trained and benchmarked. However, to create an accurate dataset, which can reach the scale of hundreds of thousands of identities to be used in the field of biometrics research is challenging. If the number of resources that needs to be employed to collect and sort all pertinent biometrics features costing time and money was the one and only limitation, it could have been easily exceeded. However, there are also limitations concerning legislations (Colbois, Freitas Pereira, et al., 2021). In fact, as it's explained in their article, the General Data Protection Regulation (GDPR) considers that "processing of personal data revealing […] biometric data for the purpose of uniquely identifying a natural person, […] shall be prohibited" if the informed consents are not obtained from data subjects[2].

[1] OpenAI DALL-E 2: https://openai.com/dall-e-2/ [accessed online Wednesday 25 May 2022]

[2] European Commission. General data protection regulation - processing of special categories of personal data, 2018. https://gdpr-info.eu/art-9-gdpr/. [accessed online Wednesday 25 May 2022]

Consequently, the generation of fully synthetic datasets could potentially make the collection and distribution of biometrics datasets much easier, not needing the informed consent and all other regulations. On the other hand, Colbois, Freitas Pereira, & al. (2021) highlight that the synthetic datasets should satisfy three important requirements (Zhang & Jain, 2006) in order to substitute a real one:

1. **Precision:** the results obtained from the evaluation of a synthetic biometric dataset should be equal to the ones derived from a real biometric dataset.
2. **Universality**: the precision requirement should be satisfied for all evaluated authentication algorithms.
3. **Privacy**: each biometric data in the synthetic dataset should not represent any real person.

In this work, the *privacy requirement* will be just indirectly discussed as, how you will discover more in details later on, an existing fully synthetic fingerprint database was used, and this requirement has already been discussed in the correlated article (Bahmani et al., 2021).

Moreover, the evaluation of the second requirement (*universality*) demands much more resources in terms of time and costs than are available for this Master's thesis. We cover it partially in this work by considering two authentication algorithms, but it will be more extensively addressed in future work.

As we will see in the next section (1.1. Objectives and key points), the aim of this thesis is to evaluate the *precision requirement* of the synthetic fingerprint dataset, generated in the framework of this thesis. The latter, is based on the synthetic fingerprint images issued from the study of Bahmani, Plesh, et al. (2021). Subsequently, these fingerprints have been warped, thanks to a warping model implemented during the studies of Marco De Donno, which are still in progress under the supervision of Prof. Champod of the University of Lausanne. For this reason, only the results of the warping model will be exposed in this thesis. In fact, it contains confidential data that cannot be published and must remain publicly unavailable.

1.1. Objectives, hypothesis and key points

As aforementioned, the aim of this Master thesis will be to study if the evaluation results of generated fingerprint datasets are similar to the one from the real dataset (first hypothesis). This, observing also if a real and a synthetic database could have similar intra- and inter-class distributions (second hypothesis) so that to use a synthetic database to train biometric systems (third hypothesis). To do this, here are the key points that will be treated in this thesis:

- An introduction to the **Biometrics basics**, where a general description of biometrics features and their principles will allow to better understand what will be discussed later;
- A section will be dedicated to the **synthesis of images**, to approach the Generative Adversarial Network (GAN) and understand what it is, how it works and what it is possible to do with GAN. In this section there will also be place for a literature review with of course some examples of the images that can be generated. After that, we will focus on the **synthetic fingerprint generator** that has been used in the framework of this thesis with the goal of understanding how it works and discuss the *privacy* requirement, briefly approached previously;
- Moreover, as anticipated before a discussion on the **synthetic fingerprint deformation** will also take place. Therefore, the details of the used code cannot be given as it contains confidential data;
- The section dedicated to the **practical part** of the work will first describe what is an Automated Fingerprint Identification System (AFIS), how it works, and which one/s has been used for the experimentations. Secondly, we'll take a look on the AFIS performances when using real or synthetic fingerprint datasets, describing the methodology and the submission's protocol;
- The evaluation results comparison of the AFIS's performances between the real and the synthetic fingerprint with a discussion of the possible benefits and limitations;
- The conclusion of the thesis, which will consist in a summary of the practical results obtained with the evaluation of the achieved goals and the further works that might be carried out in the near future.

2. Synthetic Fingerprint Images

2.1. The basics of Biometrics

To be able to assess the value of a fingerprint datasets as we aim to do in the framework of this thesis and to understand how to approach this evaluation, it could be helpful to take a moment for perceive the basics of this very vast branch.

Each of us, just like most of the animals that populate our planet, accomplishes almost every day a sort of identity recognition process. For instance, recent studies show that, for a dog, just the voice of its owner could be enough to recognize him (Gábor et al., 2022). Moreover, humans can recognize each other from their body language and physical appearance.

Therefore, continuous technology improvements in the Computer Science field have made the automation of many process possible, including in the biometrics field, which aims to automatically identify individuals from their physical, chemical and/or behavioral attributes (Jain and Ross, 2008).

- **Physical trait**: corresponds to the anatomical human's features. For instance:

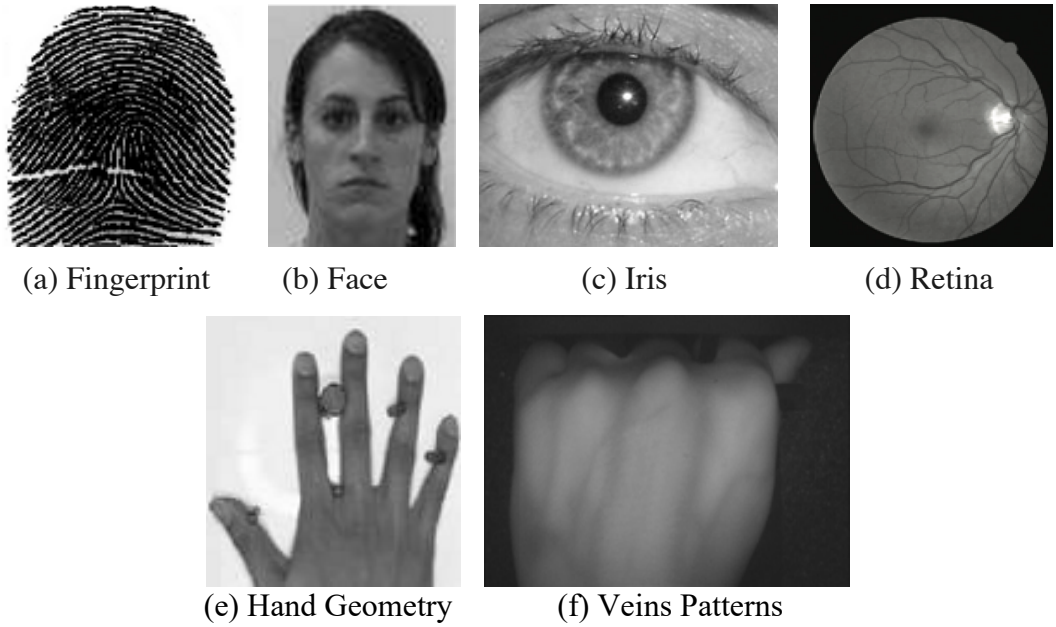
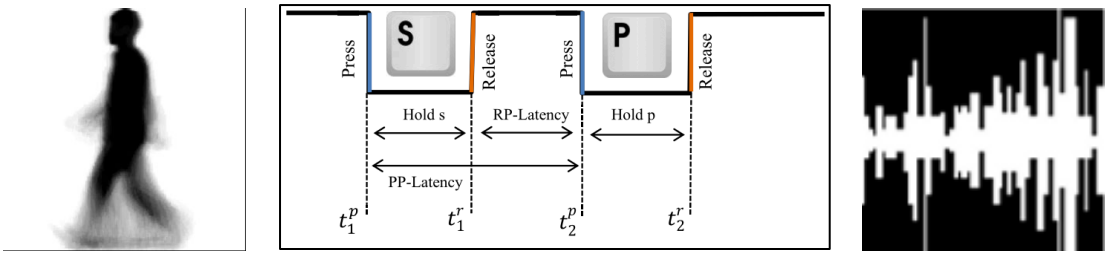


***Image 2**: (a) Fingerprint (Illustration: Jain & Ross, 2008, pg. 4) , (b) Face (Illustration: Jain & Ross, 2008, pg. 4), (c) Iris (Illustration: Erturk, 2006, pg. 413), (d) Retina (Illustration: Farzin et al., 2008, pg. 5), (e) Hand Geometry (Illustration: Jain & Ross, 2008, pg. 4) , (f) Veins Patterns (Illustration: Soni et al., 2010, pg. 504)*

- **Behavioural trait**: refers to the manner in which a person act. This could include:



(a) Gait (b) Keystroke dynamics (c) Voice

***Image 3**: (a) Gait recognition (Illustration: Rida et al., 2016, pg. 465), (b) Keystroke dynamics (Illustration: Morales et al., 2016, pg. 7739), (c) Voice recognition (Illustration: Jain & Ross, 2008, pg. 4)*

- **Chemical trait**: relates to the chemical composition of a person. For example: DNA and body odour.

Moreover, a biometric trait, may this be physical, behavioral or chemical should meet the seven attributes that, Jain et al. (1999), identified to determine "the suitability of a […] trait to be used in a biometric application" (Jain and Ross, 2008) which are the following: (1) Universality, (2) Uniqueness, (3) Permanence, (4) Measurability, (5) Per-formance, (6) Acceptability, (7) Circumvention (the imitation of the trait using artifacts should not be easy) (Jain et al., 1999, quoted in Jain & Ross, 2008). Hence, if we apply these attributes to the main subject of this thesis, fingerprints are admissible as part of the physical biometric characteristics. This means that they can be used to prove someone's identity through different applications (i.e., commercial, government, forensic applications), and as such it is considered a personal data (GDPR).

Previously, in the first section of this document (1. Introduction), the subject of privacy has been briefly mentioned and it's important to identify which of these privacy issues could be, as it is critical today as it was in the past (Davies, 1994; quoted in Jain & Ross, 2008). More precisely forging attacks, like spoofing (Rebera et al., 2014), could be carried out, putting the identity and sensitive information (i.e., medical and health conditions) of peoples at risk of abuse or unintended use (Marcel et al., 2014).

Finally, in the framework of this thesis it could also be useful to know how the biometrics features can be employed. Generally a biometric system is used to acquire the desired trait and extract features from it (Jain and Ross, 2008). Afterwards, the pattern recognition system operates the comparison between the acquired trait's features and the stored ones (Jain and Ross, 2008).

Moreover, depending on the application, the comparison can be performed in two different modes (Jain and Ross, 2008):

- **Verification** mode: the individual claims an identity (i.e., via a smart card, password, etc.) and submits the required biometric trait to the system which compares it with the (previously) stored features corresponding to the claimed identity (Jain and Ross, 2008). It consists in a one-to-one comparison and a typical example question: "Does the presented iris features belong to the claimed identity?".
- **Identification** mode: consists in a one-to-many comparison between the probe (i.e., the presented biometric trait) and multiples references (i.e., the enrolled features in the database). The aim of the identification mode is to determine who a person is.

In both the verification and identification mode, the biometric system will tell how close the probe and the reference features are, by the *score*. Most of the times the score is a measure of the similarity between two features, so, the higher the score, the more closely the features match. After that, a decision-making process needs to be performed to determine the *threshold* where, if the score is grater or equal, it is a *Match,* and, if it is smaller, it is a *No Match*. This process is often quite tricky and requires a trade-off between the possibility of accepting impostor users and the possibility of rejecting genuine users because of the multiple factors that can affect an individual biometric trait (i.e., sensing conditions, alteration of the characteristic, ambient conditions, interaction with the sensor) (Jain et al., 2008). Furthermore, in their chapter, Jain & Ross (2008) clarify how the performances of biometric system can be evaluating observing the False Match Rate (FMR) and the False Non-Match Rate (FNMR). The first one (FMR) tells the percentage of the times where the biometric system concludes the comparison as a match, but the features come from two different identities. On the other side, the FNMR tells the percentage of times where the system concludes the comparison as a non-match, but the features come from the same identity (Jain and Ross, 2008).

In conclusion, as said before, this section only briefly presents the basics of biometrics. Later in this document certain important notions (i.e., operations of a biometric systems, performances, etc.) will be extended and discussed.

2.2. Generative Adversarial Network (GAN) for images synthesis

Recent progress in the field of artificial intelligence (AI) has made machine learning more and more popular. This particular field was defined for the first time in 1959 by Arthur Samuel as "the field of study that gives computers the ability to learn without being explicitly programmed" (Samuel, 1959) and, for instance, nowadays it can be used "[…] to identify objects in images, transcribe speech into text […]" and so on (LeCun et al., 2015).

Hence, deep learning, which is a branch of machine learning (Chassagnon et al., 2020) refers to a network of neurons organized in multiple successive layers, each one performing simple operations and sending the results to others neurons (LeCun et al., 2015). This allows to model very elaborated functions using millions of parameters, with great results. Essentially, deep learning can be: *generative* (i.e., where the aim is the synthesis of an image), *discriminative* (i.e., where the aim is the classification of an image) or *hybrid*, which combines the generative and the discriminative architectures (Kim et al., 2018).

Goodfellow et al., (2014) advanced a hybrid deep learning architecture: Generative Adversarial Network. In this framework, the generative and the discriminative architectures are in competition and they are both trained at the same time (Kim et al., 2018). To take an example, it can be considered that in a particular application the aim of the discriminator (D) is to correctly recognize if a face image is real or fake, and, on the other hand, the generator (G) aims to deceive the discriminator. Therefore, the generator (G) is trained thanks to the feedback of the discriminator (i.e., if D classified the synthetic face image as "Fake", G would try to synthetize a more realistic face image) and the discriminator learns from weather or not the classification was correct (figure 3).

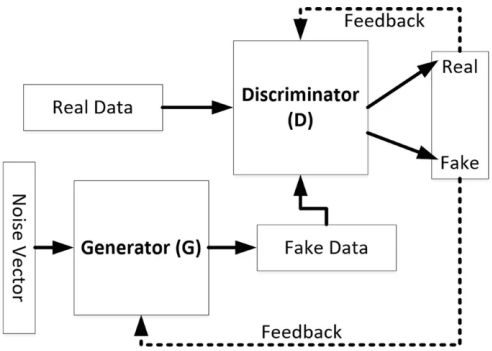


***Image 4**: typical architecture of Generative Adversarial Network (GAN) (Illustration: Kim et al., 2018)*

GAN's technology has been subject of many studies and research during the last years and has also been applied in various fields of the biometrics, mostly in face recognition, where impressive results are being achieved. One of the main challenges over the past few years regards the synthesis process not being trivial to synthetize more images of the same identity. For instance, concerning real faces datasets it is possible to take several pictures of the same person, from different angles, with different illumination and expressions. Thus, as shown in Image 5, Colbois et al. (2021) recent studies demonstrate that it is possible to identify the parameters responsible of the angle, illumination and expression and change them in order to generate intra-class variability (different images of the same identity) with satisfying results.

However, in the next section we will focus on the synthesis of fingerprint images, which is the main subject of the thesis, and it will be shown that even if the general process of generating synthetic fingerprint and face images it is very similar, some differences were observed between the two field's state-of-the-art.

***Image 5**: Same synthetic identity with variation of expression (first row), pose (second row) and illumination (third row). The image top left (red contour) represent reference. Illustration from: Colbois et al., 2021*

2.2.1. Synthetic fingerprint generators

As discussed in the introduction of this document, establishing a large fingerprint database that can be used to evaluate a fingerprint recognition system could involve high costs in terms of money and time, as well as certain privacy's related risks (Maltoni, 2009). Therefore, the synthesis of fingerprint has a quite long history and the GANs architectures, which represent the current state-of-the-art in the synthesis of images, became popular just in the last decade (Goodfellow et al., 2014). As a matter of fact, Cappelli, Maio and Maltoni (2000a, b, 2004) (quoted by Maltoni, 2009), developed a software (SFinGe) which allows the automatic generation of a "large database of fingerprints […]" (Maltoni, 2009).

Essentially, to generate synthetic fingerprints, SFinGe execute the three steps of a typical fingerprint features extractor, but in a different order: from a fingerprint area, a frequency image and an orientation image produce a binary ridge pattern image (Maltoni, 2009). Afterwards the obtained images are rendered more realistic by "adding fingerprint-specific noise" (Maltoni, 2009) and by changing parameters such as a displacement, rotation, noise, etc., it is possible to synthetize more impressions of the same fingerprint identity (intra-class variability).

Now, as it has been discussed previously, Generative Adversarial Networks does not require labelling and annotating data thanks to its unsupervised learning nature and can generate a high number of high-quality synthetic data looking more and more realistic[3]. Moreover, the reason why the focus of this thesis is on GANs is that deep learning models have a more flexible generation process than mathematical ones (i.e., SFinGe). In other terms, SFinGE (mathematical model) generates a synthetic fingerprint from a given set of three maps (fingerprint area, frequency map and directional map) while the input of a GAN based model is a *random noise* (Riazi et al., 2020).

Since Goodfellow et al. (2014) designed this machine-learning architecture, the synthesis of fingerprint images gained in popularity. As a consequence, many researchers tried to implement their algorithms based on GANs with always greater results in terms of quality and realism during the years. Roy et al. (2017) generated synthetic fingerprint templates that can be used to authenticate an impostor as a genuine user.

This particular research motivated Bontrager et al. (2018) to design a model able to generate fingerprint images (instead of fingerprint templates) similar to real fingerprint: *DeepMasterPrints*. To do this, Bontrager et al. (2018) generated fingerprint images using WGAN algorithm, where, during the training of the Discriminator, the distance between the real and the generated distribution are measured by the Wasserstein distance function instead of the Jensen-Shannon divergence metric, increasing the stability of the training (Arjovsky et al., 2017).

[3] Analytics Vidhya: Generative Adversarial Networks | GANs for Image Data : https://www.analyticsvidhya.com /blog/2021/03/why-are-generative-adversarial-networksgans-so-famous-and-how-will-gans-be-in-the-future/ [accessed online Tuesday 7 June 2022]

Minaee & Abdolrashidi (2018) proposed a *deep convolutional GAN* (DC-GAN) architecture trained with two fingerprint databases (FVC 2006 Fingerprint Database[4] and PolyU Fingerprint Databases). As they explain in their paper, the lines composing a fingerprint image form a connected component that GAN can learn better, thanks to the term added to the loss function of the proposed DC-GAN architecture (Minaee and Abdolrashidi, 2018).

Furthermore, to the best of my knowledge, Mistry et al. (2020) generated the largest synthetic fingerprint image dataset consisting of 100 millions synthetic fingerprints. To do that, they used an Improved-WGAN (I-WGAN) architecture initialized by the outputs of the Convolutional Autoencoder. As shown by Image 7 (d), the visual quality of the synthetic rolled fingerprints seems fine, unlike the synthetic plain fingerprint images, seeming to have very wavy ridges. However, the evaluation results shown in their study demonstrate that the proposed database raised the bar in terms of realism and quantity.

In parallel, during the same year, (Fahim and Jung, 2020) presented a different GAN architecture for fingerprint generation able to speed-up the synthesis process and to generate "[…] whole and cropped fingerprint patches with 128 by 128 and 256 by 256" (Fahim and Jung, 2020).

For their part, Riazi et al. (2020) developed a two-phases method (SYNFI) with which they first generate a low-quality fingerprint image form a random latent variable (GAN model) and then use a *Super-Resolution* (SR)[5] model in order to add realism to the image and turn it into an high-quality one. SYNFI was developed to meet real fingerprint expectations, which (Riazi et al., 2020) summarized essentially in four goals: (1) the features used for AFIS should be preserved, i.e., ridge structure and minutiae. (2) It generates full-finger images, (3) that should be indistinguishable from real fingerprints impressions and (4) the system should do this automatically (Riazi et al., 2020).

[4] University of Bologna FVC 2006; Fingerprint Verification Competition: http://bias.csr.unibo.it/fvc2006/databases.asp [accessed online Tuesday 7 June 2022]

[5] Wikipedia, *Super Resolution Imaging:* https://en.wikipedia.org/wiki/Super-resolution_imaging [accessed online Tuesday 8 June 2022]

Moreover, they computationally tested the *realism* of the generated fingerprints by using different classifiers and concluded that the best one that was used "could distinguish synthetic fingerprints from real ones only 0.43% better than a random guess" (Riazi et al., 2020) which demonstrate that they met the fixed goals of their research.

With the aim of generate a fully synthetic 512 x 512 pixels at 500 dpi, plain impression fingerprint images dataset at scale, (Bahmani et al., 2021) took a step forward in this field and worked on the *Clarkson Fingerprint Generator* (CFG). With this model, the authors from the Clarkson University (Potsdam, NY, USA) and from Precise Biometrics (Potsdam, NY, USA) improved the quality of fingerprint images based on previous models and demonstrated that the generated fingerprints did not disclose the identity of the real fingerprint utilized for the training of the *CFG* model. In addition, they made the pre-trained *Clarkson Fingerprint Generator* model publicly available as well as the synthetic fingerprint dataset[6] generated by them with the *CFG* model[7]. As we discussed in the previous section, and as it will be discussed more in detail later, one of the biggest challenges in working with Generative Adversarial Networks is that it is not trivial to implement a method capable of synthetize more images of a same identity.

As a matter of fact, contrary to the mathematical method SFinGe (Cappelli et al., 2002), the proposed StyleGAN-based *Clarkson Fingerprint Generator* (Bahmani et al., 2021) only generate a single illustration for each identity. For this reason, with the current version of their publicly available dataset, only inter-class variability analysis are possible while intra-class variability distribution cannot be evaluated.

On the other hand, the current state-of-the-art in synthetic fingerprint generation is *PrintsGAN* (Engelsma et al., 2022), which overcame several limitations of previous works in this field: most of them lack quality in terms of realism or, as previously shown, they cannot create intra-class variability for a same given identity (Engelsma et al., 2022).

[6] Clarkson Fingerprint Generator (CFG) dataset - 50k Synthetically Generated Fingerprints: https://drive.google.com/file/d/1KQUjnol9JjYQtx6D0eVN6mfUs91eWcS3/view?usp=sharing [Downloaded online Thursday 10.03.2022 by IDIAP – non-commercial license]. Contact: Bahmank@Clarkson.edu

[7] Clarkson Fingerprint Generator (CFG): https://github.com/keivanB/Clarkson_Finger_Gen [accessed last time Tuesday 8 June 2022]

The proposed method, *PrintsGAN* (Engelsma et al., 2022), is also based on Generative Adversarial Networks, but, thanks to the combined style transfer and warping module, it made it possible to synthetize fingerprint images with a high realism quality.

Furthermore, it also provides intra-class variability, in terms of distortion, moisture and pressure, since the method can generate a considerable number of images of the same finger identity. Finally, Engelsma et al. (2022), as well as (Bahmani et al., 2021), demonstrated the fact that no synthetic fingerprint reveal the identity of the real fingerprint identity used to train their deep-learning architecture. *PrintsGAN* generation process essentially requires three steps, summarized in Image 6 here belove.

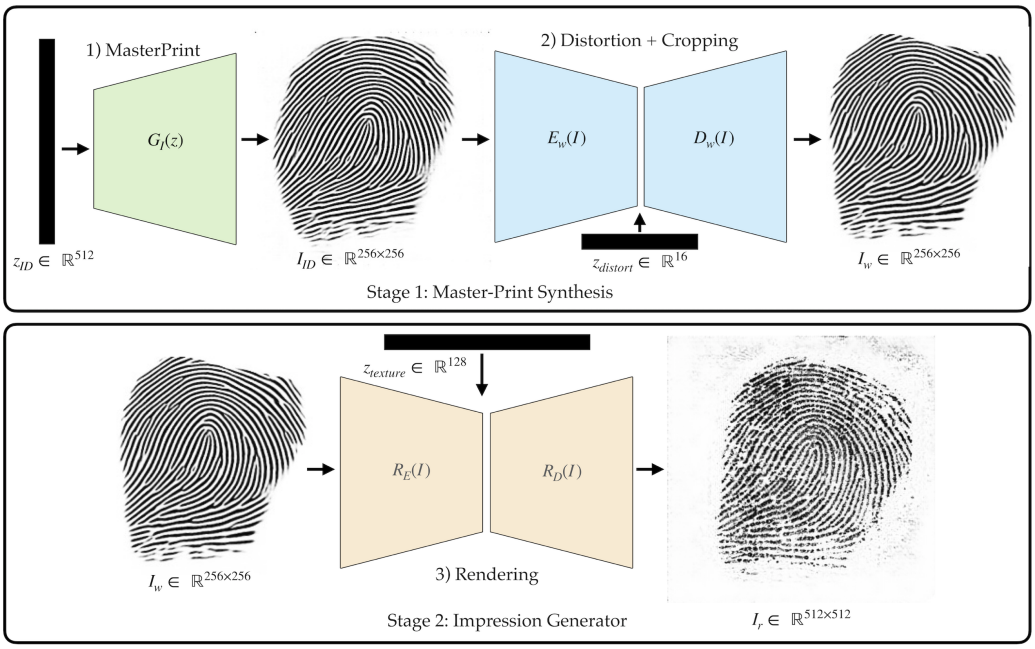


***Image 6**: Scheme of the three steps required by PrintsGAN architecture. 1) MasterPrint generation; (2) Warping and Cropping operations; (3) Textural details addition. (Engelsma et al., 2022)*

First, form a random noise a fingerprint identity is synthetized. The generated MasterPrint is warped and cropped to generate intra-class variability. Even more intra-variability is added by the last step that consists in adding texture to the image. Moreover, in order to demonstrate that the method can generate high-quality 512 x 512 pixels, rolled fingerprints providing high realism (in terms of minutiae quantity, type and quality distributions) and that it can emulate real fingerprint intra- and inter-class variability, they performed several experiments (Engelsma et al., 2022).

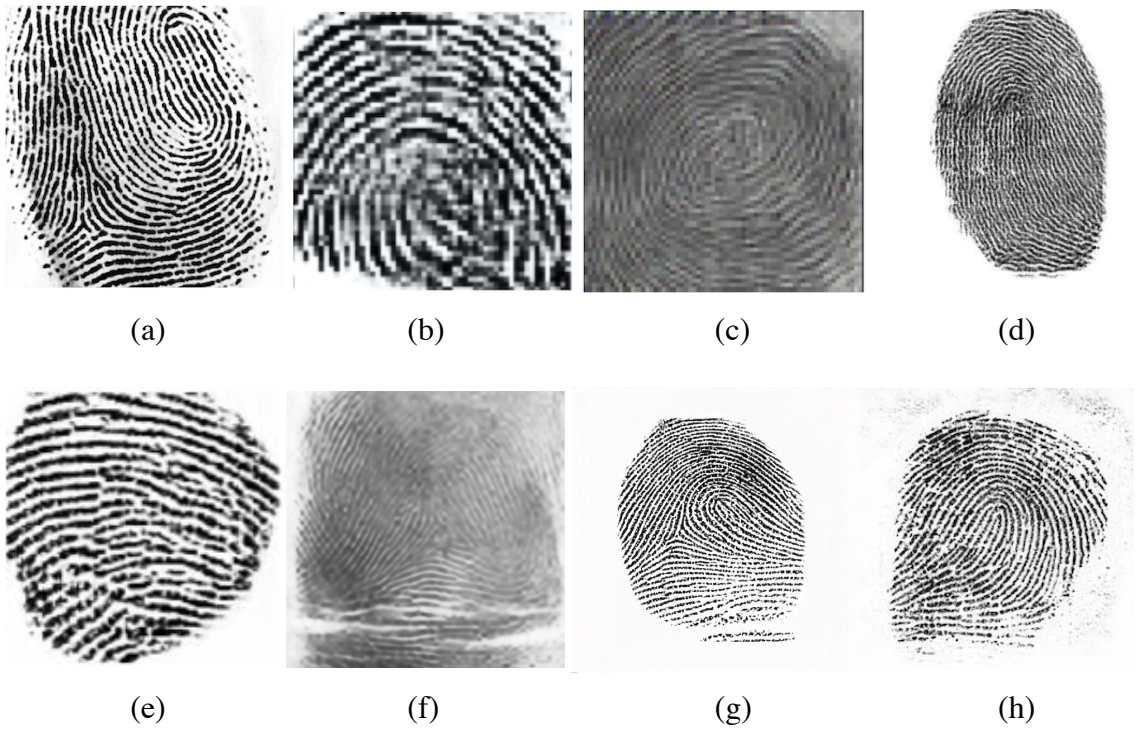


***Image 7**: Example of synthetic fingerprint issued form the related previous works discussed in this section: (a) SFinGe model (Cappelli et al., 2002), (b) DeepMasterPrints (Bontrager et al., 2018), (c) Finger-GAN (Minaee and Abdolrashidi, 2018), (d) Fingerprint Synthesis (Mistry et al., 2020), (e) Large scale FP generator (Fahim and Jung, 2020), (f) SynFi (Riazi et al., 2020), (g) Clarkson Fingerprint Generator (Bahmani et al., 2021), (h) PrintsGAN (Engelsma et al., 2022)*

Image 7 here above shows an example of synthetic fingerprint for every study discussed in this section concerning the synthesis of fingerprint images. Based on the examples shown above, the synthetic fingerprints from (a) to (f) show considerable gaps in terms of realism quality. In fact, it is possible to see that certain features like minutiae quantity and distribution as well as ridge density and shape lack realism. Moreover, it seems that the resolution of (b) and (c) is not optimal for the aim of this thesis. However, as previously stated, (Bahmani et al., 2021) reached great results in terms of fingerprint quality (g) with their *Clarkson Fingerprint Generator*, still not implementing a method to generate intra-class variability. Finally, fingerprint images (h) represent an example of what *PrintsGAN*, (Engelsma et al., 2022) can generate, thus affirming itself the current state-of-the-art in synthetic fingerprint generation.[8]

[8] The considerations made about realism quality are made form a personal point of view and it has not been validated. A crowdsourcing experiment should be performed like (Engelsma et al., 2022) have done during their research. This will allow to have solid and scientific results to base the considerations.

In conclusion, it is also important to specify that there are also other similar works, which tried to create a model to generate synthetic fingerprints and are not covered by this document, i.e., Zhao et al. (2012), (Johnson et al., 2013), (Attia et al., 2019), (Cao and Jain, 2018), (Wyzykowski et al., 2020). However, in their study, Engelsma et al. (2022) classified these works in a similar way to those shown throughout this document (Table 1):

Method (Ref.)	Image Size	Nb. of Images	Open Source	Open Data
SFinGe (Cappelli et al., 2002)	Variable	100k per 24h	No	Yes, some of them (FVC)
DeepMasterPrints (Bontrager et al., 2018)	128 x 128	Not found	Not official code found	Not official Data found
Finger-GAN (Minaee and Abdolrashidi, 2018)	512 x 512	Not found	Not official code found	Not official Data found
Fingerprint Synthesis (Mistry et al., 2020)	512 x 512	100 million	Not official code found	Not official Data found
Large scale FP generator (Fahim and Jung, 2020)	128 x 128 or 256 x 256	Not found	No	No
SYNFI (Riazi et al., 2020)	256 x 256	Not found	Yes[9]	No
CFG (Bahmani et al., 2021)	512 x 512	50k	Yes	Yes
PrintsGAN (Engelsma et al., 2022)	512 x 512	525k	No	No

Table 1: *Summary of the GAN-based models discussed previously during this study*

This section (2.2.1.) provided a general overview of the previous related works in the field of synthetic fingerprint generation by also comparing and discussing the obtained results and the current state-of-the-art. In the next section, the practical part will be introduced, by initially considering which fingerprint generator system was chosen in the framework of this thesis, why and how it works in detail and what can be done to overcome the limitations of this generator.

[9] SYNFI code : https://github.com/MohammadChavosh/synthetic-fingerprint-generation [accessed last time Tuesday 21 June 2022]

2.3. Synthetic fingerprint generator of choice

First, it is important to remember that in the aim of this work is not to create a new GAN-based fingerprint generator model. In fact, doing this would require more time than allowed for conducting this study. For this reason, it was decided to choose one of the existing synthetic fingerprint generators, to conduct experiments on them to reach the set goal, which is to determine if the evaluation results of generated fingerprint datasets are similar to the ones from the real dataset (§ 1.1. Objectives and key points).

Since, as shown in the previous section (§ 2.2.1 Synthetic fingerprint generators), *PrintsGAN* (Engelsma et al., 2022) is the current state-of-the-art model in fingerprint generation, it was initially the model of choice for developing the practical part of this study. Furthermore, based on their abstract they "plan to release […] database of synthetic fingerprints to the public" (Engelsma et al., 2022). However, after contacting the authors of the research, it was been possible to obtain neither the generated database of the synthetic fingerprints issued from *PrintsGAN,* nor the code developed by them to eventually recreate another synthetic database.

Hence, since it is impossible to use *PrintsGAN* database and/or code, it was **decided to utilize *Clarkson Fingerprint Generator* model** (Bahmani et al., 2021), which was made publicly available as well as the synthetic fingerprint dataset[10] generated by them with the *CFG* model[11]. As previously stated, *CFG* provides fully synthetic 512 x 512 pixels at 500 dpi, plain impression fingerprint images, each of which represent a different identity form the used training real fingerprint dataset.

More in detail, *Clarkson Fingerprint Generator* is the first model based on StyleGAN architecture (Karras et al., 2019) and was trained in an unsupervised manner using a dataset composed by 72'000 real fingerprint images (512 x 512 pixels) captured form 250 unique identities using a fingerprint scanner: Crossmatch Guardian scanner (Bahmani et al., 2021). Thereafter, the *CFG* has been utilized to generate a 50'000 fully synthetic fingerprint database which is publicly available[10].

[10] Clarkson Fingerprint Generator (CFG) dataset - 50k Synthetically Generated Fingerprints: https://drive.google.com/file/d/1KQUjnol9JjYQtx6D0eVN6mfUs91eWcS3/view?usp=sharing [Downloaded online Thursday 10.03.2022 by IDIAP – non-commercial license]. Contact: Bahmank@Clarkson.edu

[11] Clarkson Fingerprint Generator (CFG): https://github.com/keivanB/Clarkson_Finger_Gen [accessed last time online Tuesday 8 June 2022]

The evaluation of generated fingerprint database essentially relies on two measures:

- Fréchet Inception Distance[12] (FID): is a more consistent metric than the Inception Score and it is considered the standard to evaluate the quality of Generative Adversarial Networks. Introduced by Heusel et al. (2018), it has the advantage of using real images distribution and compare it to the synthetic samples distribution. The lower the FID, the more similar the distributions of real and generated images (Heusel et al., 2018);
- Structural Similarity[13] (SSIM): is an objective index utilized to evaluate perceptual quality measures of an image based on the degradation of structural information (Wang et al., 2004).

In the framework of their research, Bahmani et al. (2021) evaluated the quality of their generated dataset only with Fréchet Inception Distance (FID) and achieved a result of 24 for this specific metric. Hence, it is possible to state that this corresponds to a considerable improvement over the results obtained by *Finger-GAN* model (Minaee and Abdolrashidi, 2018) which achieved FID of 70. However, only a few models considered FID to evaluate the quality of their synthetic fingerprint images distribution, and, because of this, it was possible to directly compare them on the base of this metric. Anyway, the main reason why other researches did not provide a Fréchet Inception Distance could be that it works better for natural images then biometric ones (Huang et al., 2006).

Therefore, Bahmani et al. (2021) also evaluated the *privacy requirement* (Zhang and Jain, 2006) throughout the impostor distribution of the synthetic fingerprint images, by submitting them to BOZORTH3 fingerprint matcher (Watson et al., 2007). Finally, other quality evaluation tests were performed. The results from NIST NFIQ 2.0 (Tabassi et al., 2021) utilization allow to state that the quality of the generated fingerprint with CFG is satisfying.

[12] Wikipedia, Fréchet inception distance: https://en.wikipedia.org/wiki/Fréchet_inception_distance [accessed online Thursday 9 June 2022]

[13] Wikipedia, Structural similarity: https://en.wikipedia.org/wiki/Structural_similarity [accessed online Thursday 09 June 2022]

On the other hand, NIST NBIS software (Watson et al., 2007) supports the thesis that the minutiae configuration is similar to the one observed on real fingerprints. Image 8, here below, represents some examples of synthetic fingerprint images taken from the publicly available database[14] generated with the *Clarkson Fingerprint Generator* (Bahmani et al., 2021) in comparison with real fingerprint images. These latter are taken from the database "DB1_B" originally used by the participants of the 2004 Fingerprint Verification Competition and now freely available online[15].

Image 8: *examples of synthetic fingerprint images taken from the synthetic database generated by the CFG model (a, b and c) (Bahmani et al., 2021) and the publicly available DB1_B[16] form the Fingerprint Verification Competition 2004 (FVC2004) (Maltoni, 2009).*

In conclusion, all the reasons listed in this section made it possible to choose the CFG (Bahmani et al., 2021) Style-GAN-based (Karras et al., 2019) model for the execution of the practical part of the work. The next section will be dedicated to the distortion model used to overcome the intra-class variability issue previously discussed.

[14] Clarkson Fingerprint Generator (CFG) dataset - 50k Synthetically Generated Fingerprints: https://drive.google.com/file/d/1KQUjnol9JjYQtx6D0eVN6mfUs91eWcS3/view?usp=sharing [Downloaded online Thursday 10.03.2022 by IDIAP – non-commercial license]. Contact: Bahmank@Clarkson.edu

[15] Fingerprint Verification Competition 2004 (FVC 2004): http://bias.csr.unibo.it/fvc2004/download.asp [accessed online last time Friday 10 June 2022]

[16] Real fingerprint database DB1_B from the FVC2004004: http://bias.csr.unibo.it/fvc2004/Downloads/DB1_B.zip [accessed online last time Friday 10 June 2022]

2.4. Synthetic fingerprint deformation

As pointed out several times in this work, as well as in the study of (Engelsma et al., 2022), one of the main limitations of the *Clarkson Fingerprint Generator* (Bahmani et al., 2021) and, more generally, of most of the GAN-based models, is that it cannot provide more than one illustration for fingerprint identity. This causes a lack of intra-class variability, not allowing to evaluate and compare the intra-class distribution of synthetic fingerprints to the real fingerprint one. For this reason, during this work, it has been decided to apply deformations to the fingerprint images issued by the *CFG* model.

As a matter of fact, it is indeed rare that fingerprints of the same identity are left with an identical pressure or angle. Moreover, it is also possible that a fingerprint is left from a slip on the surface, or the finger twisted during the deposition. However, all this factors can cause fingerprint deformations (Maceo, 2009), and, thus, enlarges the intra-class variability.

Now, warping the fingerprint images in an automated way is a non-trivial operation and, for this, it is necessary to thank Marco De Donno and Prof. Christophe Champod who provided me with the distortion model they are working on, and, on which they are writing an article. This model is based on the study of Bookstein (1989) where the subject were the "*Principal Warps: Thin-Plate Splines and Decomposition of Deformations*" (Bookstein, 1989) and is an implementation of the *Thin-Plate Splines* (TPS) theory from the aforementioned article. However, since (1) they are writing an article on the distortion model which also (2) contains confidential data, it was decided not to provide precise details and this will remain a more general discourse on how the model works.

First of all, the warping parameters must be within a certain range, so that the results can be consistent with the deformations observed in the real world. To do this, couples of real fingerprint minutiae configurations have been observed in order to obtain realistic parameters of deformations, and this precisely explains where the confidentiality of this model come from. Once the parameters are calculated, it is possible to call a function (implemented by Marco De Donno) where the inputs are the image that has to be warped, and the deformation parameters previously calculated. The result is a new fingerprint impression (warped) of the same fingerprint identity.

Image 9 provides some examples of the results of the application of the deformation model, supplied by Marco De Donno and Prof. Christophe Champod, on some synthetic fingerprint images taken from the *CFG* database[17].

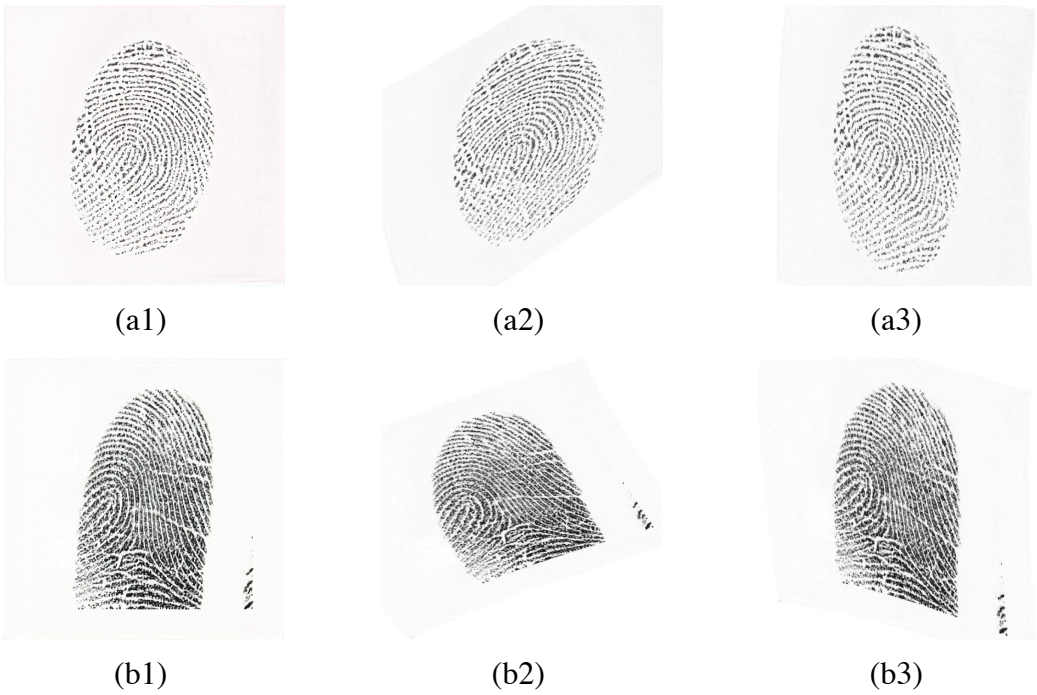


Image 9*: example of two different fingerprint from the synthetic CFG database (a1, b1) warped two times (a2, a3 and b2, b3) with the TPS deformation model provided by Marco De Donno and Prof. Christophe Champod. Images a1, a2, a3 represent three different illustrations of the same fingerprint identity, as well as b1, b2, b3.*

The provided deformation model could also allow to directly perform the warping functions on the extracted minutiae of the fingerprints. However, it was preferred to deform the fingerprint images rather than those features (1) because not all fingerprint recognition systems uses minutiae for their operations (Bontrager et al., 2018) and (2) for providing a more complete experience (extraction and rather than only comparison).

Moreover, to the best of my knowledge, this deformation model, based on TPS, is similar to the one implemented in *PrintsGAN* (Engelsma et al., 2022), but, on the other hand, they performed deformations into the Generation process in such a way that the model could also be trained to provide deformations of higher quality (Engelsma et al., 2022).

The next section will be devoted to the methodological part, the results and discussion of the conducted experiences.

[17] Clarkson Fingerprint Generator (CFG) dataset - 50k Synthetically Generated Fingerprints: https://drive.google.com/file/d/1KQUjnol9JjYQtx6D0eVN6mfUs91eWcS3/view?usp=sharing [Downloaded online Thursday 10.03.2022 by IDIAP – non-commercial license]. Contact: Bahmank@Clarkson.edu

3. Synthetic Fingerprint images to test AFIS

3.1. Methodology aspects

This section, which introduces the practical part of the thesis, summarizes the methodological aspects of the research. In fact, it could be helpful to know what was done in the framework of this study to reach the fixed objectives, allowing, in this way, anyone else to have a critical look at this work.

First of all, the literature research carried out thanks to the resources provided for UNIL students[18] and the frequent meetings with the director of the thesis, Prof. Marcel, allowed the identification of the steps needed to achieve the fixed goals.

Thereafter, as previously mentioned (§2.3. Synthetic fingerprint generator of choice), *Clarkson Fingerprint Generator* (Bahmani et al., 2021) database[19] was chosen to perform the practical experimentations. It contains 50'000 synthetic fingerprints images generated with CFG model and it is publicly available. Subsequently, these images were warped using the TPS deformation model provided by Marco De Donno et Prof. Christophe Champod (§2.4. Synthetic fingerprint deformation).

More specifically, for every synthetic fingerprint from the synthetic CFG-database, five random deformations were performed. This leads to obtaining six fingerprint impressions for every fingerprint identity for a total of 300'000 fingerprint images of 50'000 different identities.

Unfortunately, not all deformations led to satisfactory results: a considerable number of warped fingerprint images have been found to be completely blank (with no exploitable information) or not realistic (deformation was too strong and decreased realism). To solve this, it was decided to manually select 5'050 fingerprint images of 1'010 fingerprint identities, out of the 250'000 images produced by the TPS deformation model. More information about the synthetic databases will be lately provided in the next subsection (§3.1.1 Comparison Protocol).

[18] Ezproxy; selection of electronic resources of the EPFL Library: http://bib-ezproxy.epfl.ch/ [accessed last time 13 June 2022]

[19] Clarkson Fingerprint Generator (CFG) dataset - 50k Synthetically Generated Fingerprints: https://drive.google.com/file/d/1KQUjnol9JjYQtx6D0eVN6mfUs91eWcS3/view?usp=sharing [Downloaded online Thursday 10.03.2022 by IDIAP – non-commercial license]. Contact: Bahmank@Clarkson.edu

The third database (DB3_B) contains 320 fingerprint images of 40 different identities. This is the result of a merge of four fingerprint databases (DB1_B to DB4_B) freely available at the following link: http://bias.csr.unibo.it/fvc2004/download.asp. These databases were originally made available for the participants of the 2004 Fingerprint Verification Competition (FVC2004)[20] and it is important to know that DB4_B contains 80 synthetic fingerprint generated by SFinGe (Cappelli et al., 2002) but were considered as real fingerprints. More details about FVC2004 databases are available on http://bias.csr.unibo.it/fvc2004/databases.asp (scanners, participants, collecting method, etc.).

As the aim is to study if the evaluation results of generated fingerprint datasets are similar to the one from the real dataset, an Automated Fingerprint Identification System (AFIS)[21] had to be chosen. The very first intention was to use such a system from the "Ecole de Sciences Criminelles" of the University of Lausanne, but, unfortunately, it was not possible. For this reason, it has been decided to use the packages from NIST's open-source Biometric Image Software (NBIS)[22] implemented in IDIAP by Vedrana Krivokuća Hahn into "lab-fingerprint" code. The latter was modified to adapt it to the needs of this thesis. It is essentially a Python source-code running on the web-based environment Jupyter Notebook[23]. Moreover, due to some python's packages update, it was necessary to use a specific Miniconda[24] environment to use previous versions of some python's packages (details provided by Annexe A).Essentially this AFIS performs two different operations: (1) feature extraction of the fingerprint with the package NIBIS "MINDTCT" (2) fingerprint comparison (minutiae level) with "BOZORTH3" (Watson et al., 2007). Finally, the following sub-section will provide the comparison protocol, to explain in detail how the fingerprint images were submitted to the AFIS.

[20] Fingerprint Verification Competition 2004 (FVC 2004): http://bias.csr.unibo.it/fvc2004/download.asp [accessed online last time 10 June 2022]

[21] Wikipedia: Automated Fingerprint Identification System: https://en.wikipedia.org/wiki/Automated_fingerprint_identification [accessed last time 13 June 2022]

[22] NBIS packages available at: https://www.nist.gov/itl/iad/image-group/products-and-services/image-group-open-source-server-nigos#Releases [accessed online last time 13 June 2022]

[23] Jupyter Notebook: https://jupyter.org [accessed last time 13 June 2022]

[24] Conda environments: https://docs.conda.io/projects/conda/en/latest/user-guide/tasks/manage-environments.html [accessed last time 13 June 2022]

3.1.1. Comparison Protocol

To recapitulate, for this important step (fingerprint features comparison) three different databases were used. Two of them (DB1_B and DB2_B) contains synthetic (and warped) fingerprint. The third database (DB3_B) only presents real fingerprints[25].

DB_1 and DB_2 contain 3'030 synthetic fingerprint images each. For every synthetic identity (1'010 in DB1_B + DB2_B) 5 different deformations were executed, which means that there are 6 different images of the same fingerprint identity: five of them (ex. 1_2.tif, 1_3.tif, 1_4.tif, 1_5.tif, 1_6.tif) are obtained after a warping process of the same synthetic fingerprint (ex. 1_1.tif). In conclusion 6'060 synthetic fingerprint images of 1010 different identities were equally split between DB1_B and DB2_B.

On the other hand, DB3_B contains 320 fingerprint images of 40 different identities (8 impression for every fingerprint identity), as discussed previously in section 3.1. Finally, it is also important to specify that the same image does not appear in two different databases, but only in one of them. Hence, to be able to evaluate the AFIS performances, a number of **genuine** and **impostor** fingerprint comparisons were conducted in three steps: DEV_SET, EVAL_SET_1 and EVAL_SET_2.

- DEV_SET: contains the images from the first synthetic fingerprint database DB1_B. This step was used to calculate the **threshold** (θ) which better minimize both the False Match Rate (FMR) and the False Non-Match Rate (FNMR). The references are all the fingerprints with samples IDs 1 (i.e., 1_1, 2_1, 3_1, … 505_1). The probes are all the fingerprints with sample IDs 2, 3, 4, 5 and 6 (i.e., 1_2, 1_3, 2_4, 3_5, … 505_6). Image 10 below show the described arrangement.

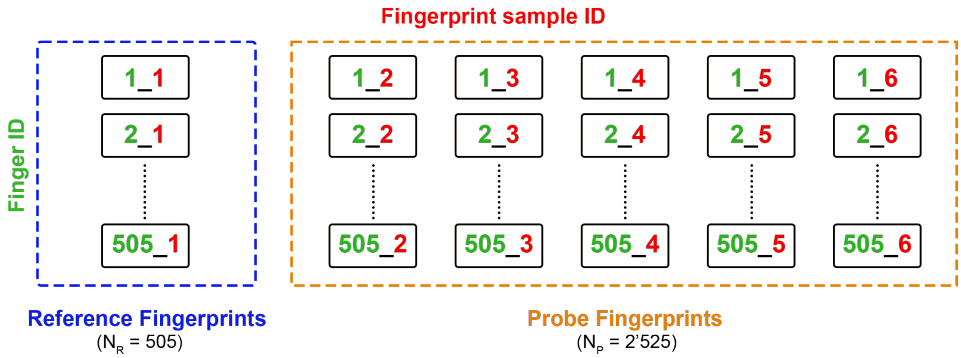


***Image 10**: arrangement of the DEV_SET used to calculate the threshold (θ).*

[25] It contains 80 synthetic fingerprint generated by SFinGe (Cappelli et al., 2002) but have been considered as real fingerprints.

- EVAL_SET_1: it contains the images from the second synthetic fingerprint database DB2_B. These comparisons allow to observe the intra- and inter-class variability distributions and evaluate the error rates using the threshold θ obtained from the previous step on a *synthetic* fingerprint dataset.
  The arrangement of the fingerprints is very similar to the "DEV_SET" one, as it is shown by the Image 11 here below:

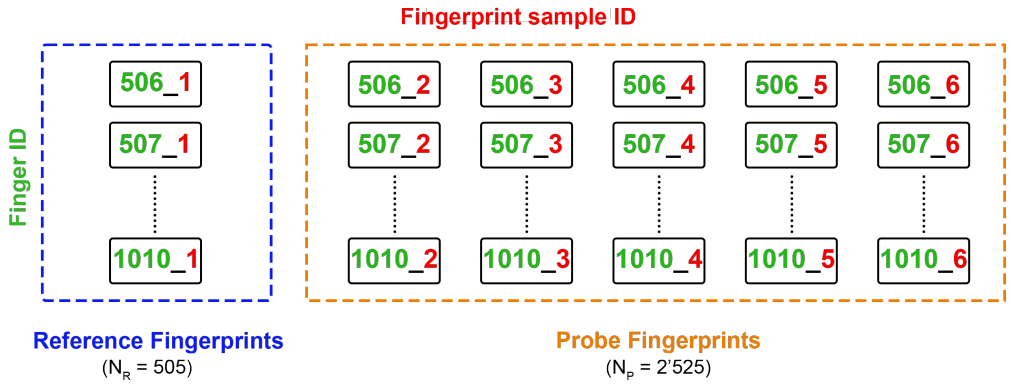


***Image 11**: EVAL_SET_1 arrangement; used to evaluate the results on a synthetic fingerprint dataset*

- EVAL_SET_2: contains the images from the third fingerprint database DB3_B. These comparisons will be performed using the threshold θ obtained from the first step and will allow to observe the intra- and inter- variability distribution and evaluate the error rates using the threshold θ obtained from the previous step on a *real* fingerprint dataset. This dataset is a little different from the others two previously observed. Here, the references are all the fingerprints with samples IDs 1 (i.e., 101_1, 102_1, 103_1, … 140_1). The probes are all the fingerprints with sample IDs 2, 3, 4, 5 and 6 (i.e., 101_2, 101_3, 102_4, 103_5, … 140_6). Image 12 shows this arrangement.

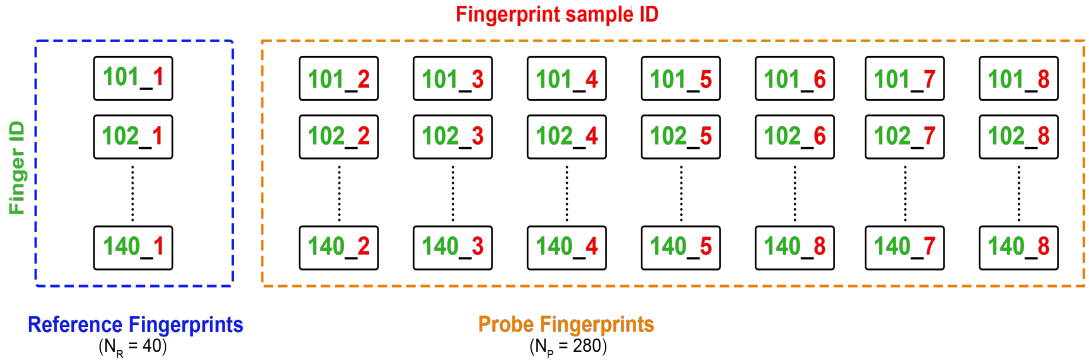


***Image 12**: EVAL_SET_1 arrangement; used to evaluate the results on a real fingerprint dataset*

In conclusion, regarding the various types of comparisons: a *genuine* comparison is made between two fingerprint impressions of the same fingerprint identity, while an *impostor* comparison is a comparison between two fingerprint images of two different fingerprint identities. To provide an easier understanding of the number of the performed comparison, the following Table 2 summarizes them.

	Genuine **Comparisons**	*Impostor* **Comparisons**
DEV SET	**2'525** (= 5 x 505)	**1'272'600** (= 505 x [2'525 - 5])
EVAL SET 1	**2'525** (= 5 x 505)	**1'272'600** (= 505 x [2'525 - 5])
EVAL SET 2[26]	**280** (= 7 x 40)	**10'920** (= 40 x [280 -7])

***Table 2**: Number of the performed comparisons for both genuine and impostor type, and for each comparison SET.*

The next section will provide the results obtained by following the described methodology aspects. Moreover, it will also cover a discussion of the results, with a critical overview of the benefits and the limits of the experimentations.

[26] The features extraction of a certain number of fingerprint images of the DB3_B did not work correctly. In fact, fingerprint 131_1 (DB_3) failed to enroll while 5 other images (134_7, 134_4, 136_8, 137_6 and 137_5) failed to acquire. This caused a total of 4% invalid both genuine and impostor comparisons. Details will be discussed in §3.2.

3.2. Results and discussion

In section 2.1. "The basics of Biometrics" it was briefly mentioned how a biometric system can be evaluated (Jain et al., 2008). As many of the existent AFIS, the output of the comparator system "BOZORTH3" is a *similarity score*. Essentially, it is a number that tells how similar the compared fingerprints are, so, the higher the score, the more closely the compared features match. This score must be interpreted through a decision-making process to determine the threshold θ where, if the score is grater or equal, it is a *Match,* and, if it is smaller, it is a *No Match*. Now, the False Match Rate (FMR) and the False Non-Match Rate (FNMR) (Jain and Ross, 2008) are closely linked to the determination of the threshold θ, which strictly depends on the biometric system application. For instance, if the system should have a very low possibility of accepting impostor users (FMR ~ 0), increasing this way the possibility of rejecting a genuine user (high FNMR), then a higher threshold θ should be used and vice-versa.

In the present case, considering that no special applications of the system were envisaged, it was decided to use as the threshold θ, the score (threshold θ) where the two error rates (FMR, FNMR) are approximately equal. This point is known as the Equal Error Rate (EER) and can be obtained in two steps: (1) calculate the FMR and FNMR at every possible value of the match threshold θ and (2) plot the two error rates (FMR, FNMR) versus the threshold θ. The intersection of the two curves gives an approximation of the EER. Image 13 shows the obtained curves of the FMR and FNMR of the DEV_SET, used to determinate the threshold θ (at the EER).

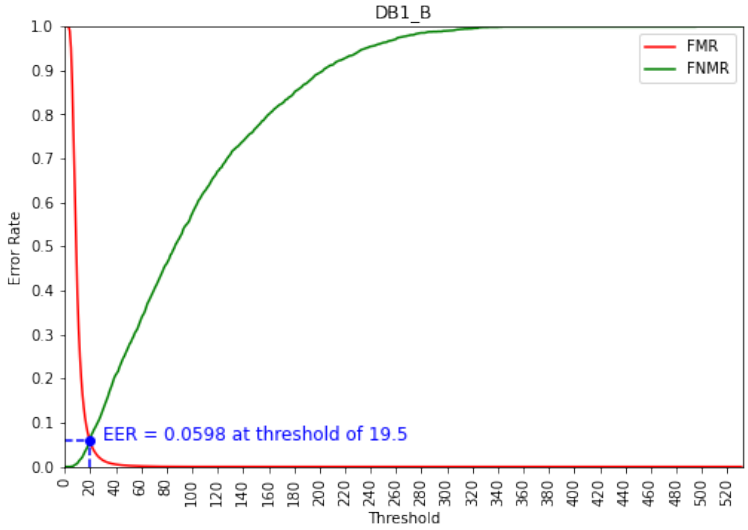


***Image 13**: plot of the error rate as a function of the threshold for the evaluation of the EER of the fully synthetic DEV_SET (DB1_B). Red curve correspond to the FMR, green to the FNMR as a function of the threshold.*

The plot represented on Image 13 shows the error rate as a function of the given threshold, and the red curve represents the False Match Rate (FMR) while the green one indicates the False Non-Match Rate (FNMR). As denoted by on the plot, the two error curves of the DEV_SET (DB1_B) intersect at an Equal Error Rate of 0.0598 at threshold θ = 19.5. Therefore, as discussed in §3.1.1. "Comparison Protocol", it corresponds to the score (θ = 19.5), used for the EVAL_SET_1 and EVAL_SET_2 evaluation to assess whether or not the comparison is a Match (comparison score ≥ 20) or a Non-Match (comparison score ≤ 19).

Applying the threshold (θ = 19.5), obtained by the DEV_SET, to the EVAL_SET_1 and EVAL_SET_2 made it possible to evaluate the performances of the biometric system in an operational context, where a certain number of references are enrolled and it has to decide whether or not the probe matches the reference, based on the threshold θ of 19.5. The obtained results of the error rate evaluation suggest that the performances are similar, and are summarized in the Table 3:

	FMR *(θ = 19.5)*	**FNMR** *(θ = 19.5)*
EVAL_SET_1	0.0572 (5.72%)	0.1983 (19.83%)
EVAL_SET_2	0.0428 (4.28%)	0.1500 (15.00%)

***Table 3**: False Match Rate (FMR) and False Non-Match Rate (FNMR) obtained for the two Evaluation datasets EVAL_SET_1 and EVAL_SET_2 at a threshold of 19.5.*

Moreover, another method allowing the visualization of the performances of the biometric system is represented by the comparison of the Receiver Operating Characteristics (ROC) graphs (Fawcett, 2006), shown on the following Image 14:

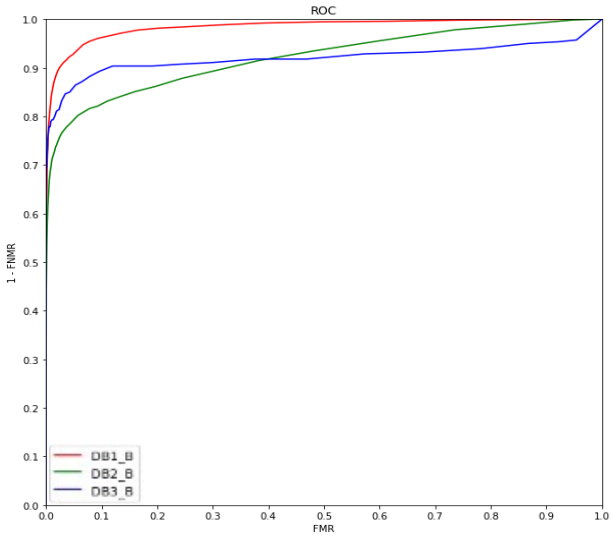


***Image 14**: comparison of the ROC curves (1-FNMR as a function of FMR) obtained from the four different datasets: in red the performances of the DEV_SET, in green the ones of the EVAL_SET_1 and blue the EVAL_SET_2.*

Both the results observed in Table 3 (FMR and FNMR at θ = 19.5) and the ROC curves comparison (Image 14) suggest that the performances of the two evaluation datasets are similar. In fact, the ROC graphs suggest that the performances of the real fingerprint dataset EVAL_SET_2 (in blue) are slightly better than the ones observed for the EVAL_SET_1. This hypothesis finds also support with the error rates obtained at a threshold of 19.5, determined with the synthetic dataset DEV_SET (DB1_B).

Furthermore, we also present histograms of the resulting score distribution (Image 15: a1, b1 and c1) for the three datasets. Additionally, to take a closer look to the zone of interest, the logarithm in base 10 of the scores was calculated to plot more detailed histograms (Image 15: a2, b2 and c2), with the red line representing the logarithm of the threshold.

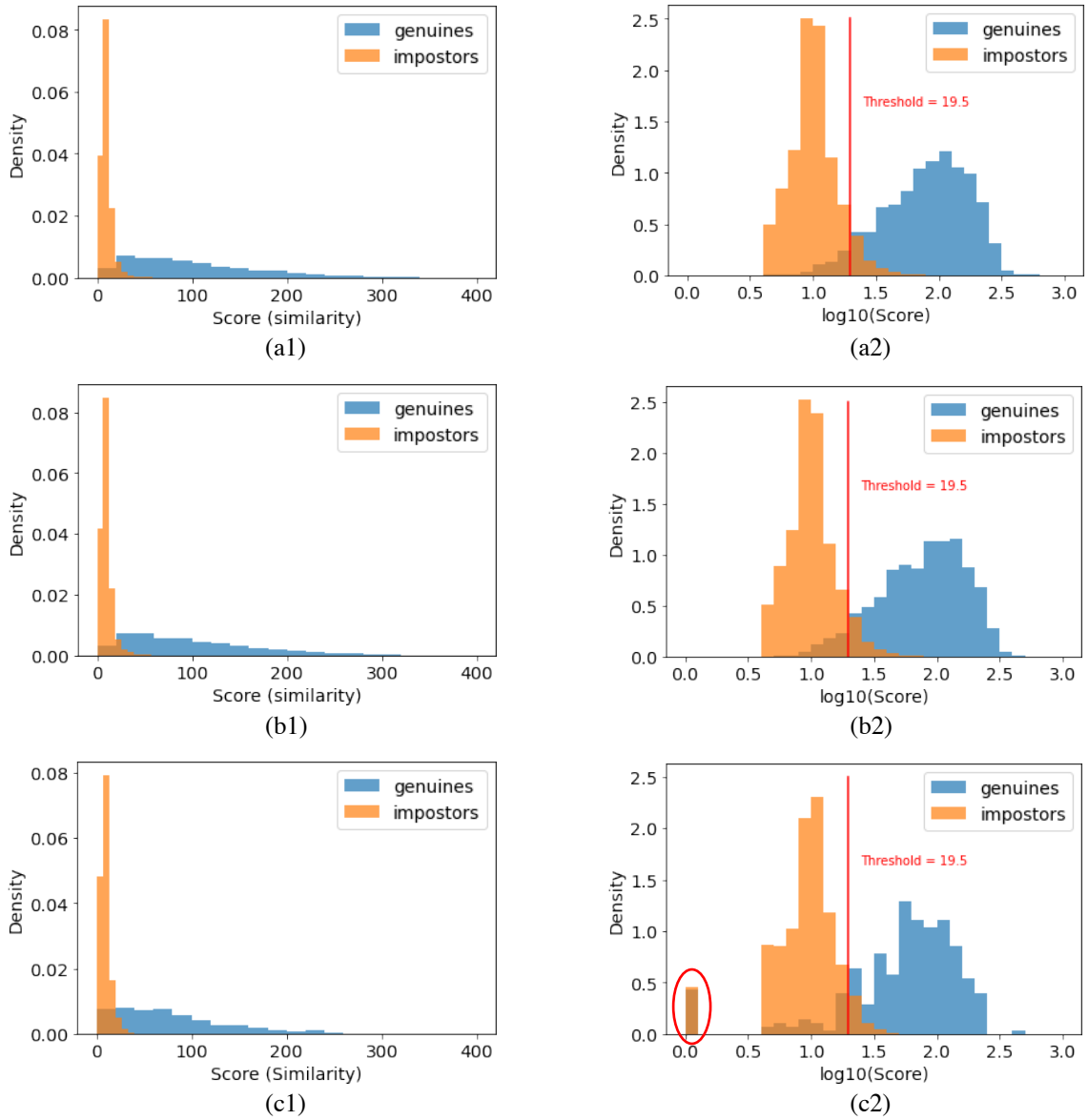


***Image 15**: histograms of the score distribution (a1, b1, c1) and of the log(score) distribution (a2, c2, b2) for the three evaluated datasets DEV_SET (a1, a2), EVAL_SET_1 (b1, b2) and EVAL_SET_3 (c1, c2).*

Image 15 provides a comparison of the different histograms of the score and the log(score) distribution of the three given datasets. The comparison shows an overall similarity of the distributions throughout the three datasets and, this way, it supports the hypothesis that the intra-class and the inter-class variability of the synthetic database is similar to the one of the given real databases. The resulting distribution of the genuine scores (in blue on the Image 15) suggests that the intra-variability of a synthetic database is similar to the one observed in a real database. In the same way, the distribution of the impostor scores (in orange on the Image 15) suggests that the inter-class variability generated by the employed model is comparable to the variation of fingerprint identity observed in a real fingerprint database.

However, the (c2) histogram (EVAL_SET_2) of Image 15 presents a not-negligible density of genuine (in blue) and impostor (in orange) comparisons resulting in null score (score = 0). Further investigations on these data showed that the features extraction of a certain number of fingerprint images of the DB3_B did not work correctly. Indeed, taking a closer look to the extracted feature files permitted to understand that, for seven fingerprint images, no minutiae were extracted during this step. More in detail, fingerprint 131_1, contained in DB3_B, failed to enroll while 5 other images (134_7, 134_4, 136_8, 137_6 and 137_5) failed to acquire, causing a total of 4% invalid both genuine and impostor comparisons. It is important to clarify, that this issue was considered for the calculation of the error rates previously discussed. If it had been the case, the error rates of the EVAL_SET_2 would have probably been slightly lower than calculated and would have produced more accurate results.

Another limitation of the presented experimentations is that, unlike the *PrintGAN* model (Engelsma et al., 2022), the TPS-warping model is not integrated directly into the GAN-based model, so that the distortion of the fingerprint cannot be trained. For this reason, a considerable number of warped images could not be used for the experimentations, while the ones composing DB1_B and DB2_B were manually selected, thus increasing the risk of human errors and biases. The integration of a warping function into the GAN architecture, as well as the consideration of the invalid comparisons could be the subject of further works.

Finally, the experimental results support all three initial hypotheses. In fact, that the first hypothesis, according to which the evaluation results of generated fingerprint datasets (DB1_B and DB2_B) are similar to the ones from the real dataset (DB3_B), is supported by the error rates (Table 3) and the ROC curves comparison (Image 14).

The results of the score distributions shown on Image 15, supports the second formulated hypothesis, according to which the intra- and inter-class variability of the real and the synthetic datasets are similar.

In conclusion, the whole methodology and the set of results obtained during practical experimentations eventually support the third hypothesis, according to which a fully synthetic fingerprint database could be used to train a biometric system (AFIS) instead of using a real fingerprint database, which involves limitations in terms of quantity, time, money, and privacy. Further works will allow this hypothesis to be developed in details.

4. Conclusions

An automatic biometric recogniser needs large-scale datasets in order to be trained and benchmarked which involves certain limitations in terms of time, money and privacy. Recent developments in the field of Artificial Intelligence (AI) and, more in detail, the successes achieved by the Generative Adversarial Networks (Goodfellow et al., 2014) in the generation of synthetic images offer numerous possibilities to try to solve these constraints.

In the context of this study, a fully synthetic and publicly available fingerprint database (Bahmani et al., 2021) has been the subject of the experimentations. To solve the issue of the intra-variability (not provided by the *Clarkson Fingerprint Generator (Bahmani et al., 2021)*) the synthetic fingerprints images were warped with a TPS-deformation model (provided by Marco De Donno and Prof. Champod). Although, unlike the *PrintsGAN* model, which represent the current state-of-the-art in fingerprint images generation (Engelsma et al., 2022), the warping function was implemented in the GAN architecture, making it impossible to be trained.

However, three new different fingerprint datasets were created. The DEV_SET, containing 3'030 synthetic fingerprint images of 505 different identities, was used to calculate a threshold of 19.5 to be applied to the others for the evaluation of the performances of a synthetic (EVAL_SET_1) and a real (EVAL_SET_2) datasets.

Finally, the experimental results support all three initial hypotheses. In fact, that the first hypothesis, according to which the evaluation results of generated fingerprint datasets (DB1_B and DB2_B) are similar to the ones from the real dataset (DB3_B), is supported by the error rates (Table 3) and the ROC curves comparison (Image 14).

The results of the score distributions shown on Image 15, supports the second formulated hypothesis, according to which the intra- and inter-class variability of the real and the synthetic datasets are similar.

In conclusion, the whole methodology and the set of results obtained during practical experimentations eventually support the third hypothesis, according to which a fully synthetic fingerprint database could be used to train a biometric system (AFIS) instead of using a real fingerprint database, which involves limitations in terms of time, money, and privacy. Further works will allow this hypothesis to be developed more in detail.

5. Bibliography

Arjovsky M, Chintala S and Bottou L (2017) Wasserstein GAN. arXiv:1701.07875. arXiv. Available at: http://arxiv.org/abs/1701.07875 (accessed 7 June 2022).

Attia M, Attia MH, Iskander J, et al. (2019) Fingerprint Synthesis Via Latent Space Representation. In: *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, Bari, Italy, October 2019, pp. 1855–1861. IEEE. DOI: 10.1109/SMC.2019.8914499.

Bahmani K, Plesh R, Johnson P, et al. (2021) High Fidelity Fingerprint Generation: Quality, Uniqueness, and Privacy. *arXiv:2105.10403 [cs, eess]*. Available at: http://arxiv.org/abs/2105.10403 (accessed 11 March 2022).

Bontrager P, Roy A, Togelius J, et al. (2018) DeepMasterPrints: Generating MasterPrints for Dictionary Attacks via Latent Variable Evolution. arXiv:1705.07386. arXiv. Available at: http://arxiv.org/abs/1705.07386 (accessed 7 June 2022).

Bookstein FL (1989) Principal warps: thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11(6): 567–585. DOI: 10.1109/34.24792.

Cao K and Jain A (2018) Fingerprint Synthesis: Evaluating Fingerprint Search at Scale. In: *2018 International Conference on Biometrics (ICB)*, Gold Coast, QLD, February 2018, pp. 31–38. IEEE. DOI: 10.1109/ICB2018.2018.00016.

Cappelli R, Maio D and Maltoni D (2002) Synthetic fingerprint-database generation. In: *Object recognition supported by user interaction for service robots*, Quebec City, Que., Canada, 2002, pp. 744–747. IEEE Comput. Soc. DOI: 10.1109/ICPR.2002.1048096.

Chassagnon G, Vakalopolou M, Paragios N, et al. (2020) Deep learning: definition and perspectives for thoracic imaging. *European Radiology* 30(4): 2021–2030. DOI: 10.1007/s00330-019-06564-3.

Colbois L, Freitas Pereira T de and Marcel S (2021) On the use of automatically generated synthetic image datasets for benchmarking face recognition. In: *2021 IEEE International Joint Conference on Biometrics (IJCB)*, Shenzhen, China, 4 August 2021, pp. 1–8. IEEE. DOI: 10.1109/IJCB52358.2021.9484363.

Davies SG (1994) Touching Big Brother: How Biometric Technology Will Fuse Flesh and Machine. *Information Technology & People* 7(4): 38–47. DOI: 10.1108/09593849410076807.

Engelsma JJ, Grosz SA and Jain AK (2022) PrintsGAN: Synthetic Fingerprint Generator. *arXiv:2201.03674 [cs]*. Available at: http://arxiv.org/abs/2201.03674 (accessed 22 February 2022).

Erturk S (2006) Nonintrusive Iris Image Extraction for Iris Recognition-Based Biometric Identification. *Circuits, Systems & Signal Processing* 25(3): 405–419. DOI: 10.1007/s00034-005-0305-6.

Fahim MA-NI and Jung HY (2020) A lightweight GAN network for large scale fingerprint generation. *IEEE Access*: 1–1. DOI: 10.1109/ACCESS.2020.2994371.

Farzin H, Abrishami-Moghaddam H and Moin M-S (2008) A Novel Retinal Identification System. *EURASIP Journal on Advances in Signal Processing* 2008(1): 280635. DOI: 10.1155/2008/280635.

Fawcett T (2006) An introduction to ROC analysis. *Pattern Recognition Letters* 27(8): 861–874. DOI: 10.1016/j.patrec.2005.10.010.

Gábor A, Kaszás N, Faragó T, et al. (2022) The acoustic bases of human voice identity processing in dogs. *Animal Cognition*. DOI: 10.1007/s10071-022-01601-z.

Goodfellow I, Pouget-Abadie J, Mirza M, et al. (2014) Generative Adversarial Nets. In: *Advances in Neural Information Processing Systems* (eds Z Ghahramani, M Welling, C Cortes, et al.), 2014. Curran Associates, Inc. Available at: https://proceedings.neurips.cc/paper/2014/file/5ca3e9b122f61f8f06494c97b1afccf3-Paper.pdf.

Heusel M, Ramsauer H, Unterthiner T, et al. (2018) GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium. arXiv:1706.08500. arXiv. Available at: http://arxiv.org/abs/1706.08500 (accessed 9 June 2022).

Huang D, Allen TT, Notz WI, et al. (2006) Global Optimization of Stochastic Black-Box Systems via Sequential Kriging Meta-Models. *Journal of Global Optimization* 34(3): 441–466. DOI: 10.1007/s10898-005-2454-3.

Jain AK and Ross A (2008) Introduction to Biometrics. In: Jain AK, Flynn P, and Ross AA (eds) *Handbook of Biometrics*. Boston, MA: Springer US, pp. 1–22. DOI: 10.1007/978-0-387-71041-9_1.

Jain AK, Bolle R and Pankanti S (eds) (1999) *Biometrics: Personal Identification in Networked Society*. The Kluwer international series in engineering and computer science SECS 479. Boston: Kluwer.

Jain AK, Flynn P and Ross AA (eds) (2008) *Handbook of Biometrics*. New York: Springer.

Johnson P, Hua F and Schuckers S (2013) Texture Modeling for Synthetic Fingerprint Generation. In: *2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops*, OR, USA, June 2013, pp. 154–159. IEEE. DOI: 10.1109/CVPRW.2013.30.

Karras T, Laine S and Aila T (2019) A Style-Based Generator Architecture for Generative Adversarial Networks. *arXiv:1812.04948 [cs, stat]*. Available at: http://arxiv.org/abs/1812.04948 (accessed 11 March 2022).

Kim K, Aminanto ME and Tanuwidjaja HC (2018) Deep Learning. In: *Network Intrusion Detection Using Deep Learning*. SpringerBriefs on Cyber Security Systems and Networks. Singapore: Springer Singapore, pp. 27–34. DOI: 10.1007/978-981-13-1444-5_4.

LeCun Y, Bengio Y and Hinton G (2015) Deep learning. *Nature* 521(7553): 436–444. DOI: 10.1038/nature14539.

Maceo AV (2009) Qualitative Assessment of Skin Deformation: A Pilot Study. *Journal of Forensic Identification* Vol. 59: 390–440.

Maltoni D (ed.) (2009) *Handbook of Fingerprint Recognition*. 2nd ed. London: Springer.

Marcel S, Nixon MS and Li SZ (eds) (2014) *Handbook of Biometric Anti-Spoofing*. Advances in Computer Vision and Pattern Recognition. London: Springer London. DOI: 10.1007/978-1-4471-6524-8.

Minaee S and Abdolrashidi A (2018) Finger-GAN: Generating Realistic Fingerprint Images Using Connectivity Imposed GAN. *arXiv:1812.10482 [cs]*. Available at: http://arxiv.org/abs/1812.10482 (accessed 22 February 2022).

Mistry V, Engelsma JJ and Jain AK (2020) Fingerprint Synthesis: Search with 100 Million Prints. *arXiv:1912.07195 [cs]*. Available at: http://arxiv.org/abs/1912.07195 (accessed 11 March 2022).

Morales A, Fierrez J, Tolosana R, et al. (2016) Keystroke Biometrics Ongoing Competition. *IEEE Access* 4: 7736–7746. DOI: 10.1109/ACCESS.2016.2626718.

Ramesh A, Dhariwal P, Nichol A, et al. (2022) Hierarchical Text-Conditional Image Generation with CLIP Latents. arXiv:2204.06125. arXiv. Available at: http://arxiv.org/abs/2204.06125 (accessed 25 May 2022).

Rebera AP, Bonfanti ME and Venier S (2014) Societal and Ethical Implications of Anti-Spoofing Technologies in Biometrics. *Science and Engineering Ethics* 20(1): 155–169. DOI: 10.1007/s11948-013-9440-9.

Riazi MS, Chavoshian SM and Koushanfar F (2020) SynFi: Automatic Synthetic Fingerprint Generation. *arXiv:2002.08900 [cs, eess]*. Available at: http://arxiv.org/abs/2002.08900 (accessed 11 March 2022).

Rida I, Almaadeed S and Bouridane A (2016) Gait recognition based on modified phase-only correlation. *Signal, Image and Video Processing* 10(3): 463–470. DOI: 10.1007/s11760-015-0766-4.

Roy A, Memon N and Ross A (2017) MasterPrint: Exploring the Vulnerability of Partial Fingerprint-Based Authentication Systems. *IEEE Transactions on Information Forensics and Security* 12(9): 2013–2025. DOI: 10.1109/TIFS.2017.2691658.

Samuel AL (1959) Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development* 3(3): 210–229. DOI: 10.1147/rd.33.0210.

Soni M, Gupta S and Gupta P (2010) A New Approach for Vein Pattern-Based Recognition. In: Huang D-S, Zhao Z, Bevilacqua V, et al. (eds) *Advanced Intelligent Computing Theories and Applications*. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 482–489. DOI: 10.1007/978-3-642-14922-1_60.

Tabassi E, Olsen M, Bausinger O, et al. (2021) *NFIQ 2 NIST Fingerprint Image Quality*. 13 July. National Institute of Standards and Technology. DOI: 10.6028/NIST.IR.8382.

Wang Z, Bovik AC, Sheikh HR, et al. (2004) Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Transactions on Image Processing* 13(4): 600–612. DOI: 10.1109/TIP.2003.819861.

Watson CI, Garris MD, Tabassi E, et al. (2007) *User's guide to export controlled distribution of NIST biometric image software (NBIS-EC)*. 0 ed. NIST IR 7391. Gaithersburg, MD: National Institute of Standards and Technology. DOI: 10.6028/NIST.IR.7391.

Wyzykowski ABV, Segundo MP and Lemes R de P (2020) Level Three Synthetic Fingerprint Generation. arXiv:2002.03809. arXiv. Available at: http://arxiv.org/abs/2002.03809 (accessed 8 June 2022).

Zhang D and Jain AK (eds) (2006) *Advances in Biometrics: International Conference, ICB 2006, Hong Kong, China, January 5-7, 2006: Proceedings*. Lecture notes in computer science 3832. Berlin ; New York: Springer-Verlag.

Zhao Q, Jain AK, Paulter NG, et al. (2012) Fingerprint image synthesis based on statistical feature models. In: *2012 IEEE Fifth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, Arlington, VA, USA, September 2012, pp. 23–30. IEEE. DOI: 10.1109/BTAS.2012.6374554.

6. Acknowledgements

First, I would like to thank the director of my thesis, Professor Sébastien Marcel, for guiding me on this path through our regular meetings during these months. I wish to extend my special thanks to Laurent Colbois for helping me whenever I had a technical issue and for integrating me into IDIAP. His role was crucial throughout my thesis period.

I would also like to show my deep appreciation to Vedrana Krivokuća Hahn for providing me with technical support on the fingerprint-lab code, which she kindly shared with me. A big thank you also goes to the rest of the Biometrics Security and Privacy Group at IDIAP Research Institute in Martigny for the welcome during these months.

I am also extremely grateful to Cristina Bertani for proofreading my work and correcting orthographic, lexical and grammatical mistakes.

Last but surely not least, I would like to thank my family for the constant support not only during my thesis, but also throughout my university career.

7. Annexes

A. Miniconda Environment

The same conda environment can be created using a Python3 environment and installing the following packages summarized in Table 4:

Package	Version
jupyter	1.0.0
matplotlib	3.5.1
numpy	1.19.2
pandas	1.3.5
Pillow	9.0.1
pip	21.2.2
scipy	1.1.0

***Table 4**: Conda environnement's packages*

More information on how to create a conda environment can be found on the following link: https://docs.conda.io/projects/conda/en/latest/user-guide/tasks/manage-environments.html [accessed 13 June 2022]

B. Data location

All data used for this work have been submitted, including the used codes, the source version of this document (Word) and a README.md file providing information about the code. The location of these files is summarized in the following Table 5:

Data	Location
Readme.md	*https://gitlab.idiap.ch/alcosta/costa_memoire/-/blob/master/README.md*
Costa_mémoire.docx	*https://gitlab.idiap.ch/alcosta/costa_memoire/-/blob/master/Costa_Mémoire.docx*
Codes	*https://gitlab.idiap.ch/alcosta/costa_memoire/-/tree/master/Costa_Mémoire_Code*
Images	*https://gitlab.idiap.ch/alcosta/costa_memoire/-/tree/master/Costa_Mémoire_Images*
DB1_B, DB2_B, DB3_B (databases)	*/idiap/project/biometricscenter/students/alcosta/costa_mémoire/databases*
Scores	*/idiap/project/biometricscenter/students/alcosta/costa_mémoire/scores*
Results (extraction, comparison)	*/idiap/project/biometricscenter/students/alcosta/costa_mémoire/results*
Fingerprint Recognition Lab	*/idiap/project/biometricscenter/students/alcosta/lab-fingerprint-recognition*

Table 5: *Data Location*