

Altered Fingerprint Detection – Algorithm Performance Evaluation

INGRESS consortium – UNIL, GUC

Authors:

*Rudolf Haraksim, Alexandre Anthonioz, Christophe Champod, Martin Olsen,
John Ellingsgaard, Christoph Busch*

Keywords: altered, fingerprint, performance evaluation, detection

Abstract

The purpose of this paper is to present a comparative study on the performance of altered fingerprint detection algorithms.

Different algorithms from different institutions have been evaluated on two different datasets. Both datasets feature real alterations on fingers and the ground truth regarding the alteration is known a priori, as, in some cases, corresponding pre-altered fingerprints were also available. The performance obtained on both datasets produced by either reference state-of-the-art or custom-built algorithms is higher than the reported 10% EER from previous studies [1].

1. Introduction

As discussed as early as 1935 [2], intentional¹ fingerprint alterations mainly serve the purpose of defeating fingerprint identification measures in an attempt to hide criminal records, escape detection at borders and escape deportation measures. Fingerprint alterations are considered as one method of presentation attack detections (PAD) on fingerprint recognitions systems according to ISO/IEC 30107 [3].

The Federal Bureau of Investigation (FBI) recently confirmed that voluntary alterations either self inflicted or with surgical assistance is used to defeat identification efforts [4]. The FBI reports on the discovery of 412 fingerprint records in their AFIS system with clear indications of deliberate alterations. The most used techniques to alter the fingerprint patterns are the vertical cut (leaving a heavy scare), the removal of a vertical slice from the core of the fingerprints, the Z-cuts or the use of heat, or chemicals to burn the fingertips.

The objective of the research initiative is to offer law enforcement agencies (and border-control agencies in particular) with a mean to automatically detect fingerprints that have been altered. It falls into two distinct operational contexts: (a) being able at presentation attempts of fingerprint characteristics

¹ We can distinguish between intentional alterations when an individual deliberately applies tools (cutting, transplantations, abrasive, burning – either fire or chemical, etc.) to alter the fingerprint(s) from unintentional alterations when the fingerprint pattern is altered due to specific work activities of the individual or disease.

(on a livescan device for example) to detect the possibility of alterations and trigger further investigative actions from the law enforcement authorities in the flagged individual; (b) allowing a triage of the captured fingerprint samples, before a search in an AFIS system, between samples with signs of alterations to be treated with specific comparison algorithms and prints without signs of alterations going through the standard comparison subsystem. It is expected that the specific comparison algorithms allowing a search between altered and non-altered prints will be more computer intensive compared to a standard comparison algorithm developed for overall unaltered prints. It is clear that the requirements in terms of error rates on the detection of altered fingerprints will differ substantially between these two operational contexts. Errors rates shall seek to be much lower in (a) than in (b) because of the differing consequences of misclassification. In context (a) a failure to detect is a major security treat and false positives will impact on “innocent” citizens, whereas in context (b), the impact of false positives will only be on computing time with the direct benefits of the successful hits due to the altered-dedicated comparison algorithm.

Although the assessment of the intent associated with the alteration (either voluntary or involuntary) is decisive for the authorities, it is out of our scope and we will focus on altered structure without distinguishing between the two.

The task of detection of altered fingerprints from a biometric perspective is a classification task where a decision system, e.g. a classifier, provides a binary decision regarding the presence or absence of alteration. Recent altered fingerprint detection algorithms (referred to hereinafter as AFDA) take advantage of disruptions in the ridge flow, the detection of large number of opposing minutiae or quality measures associated with the fingerprints [1, 5-8].

In this paper, two new approaches for the detection of altered fingerprints are presented. These approaches, developed respectively by the University of Lausanne (UNIL) and the Gjøvik University College (GUC) are benchmarked to a state of the art algorithm based on the work described in [1].

2. Material and methods

2.1 Datasets

Two different datasets based on altered fingerprints, which are stemming from real operational cases, have been used to evaluate the performance of the AFDA. The first dataset (designated as UNIL_DB) is based on fingerprints selected from an AFIS database in the custody of UNIL (composed of about 1 million fingerprints originating from an operational inked card collection). The second dataset is based on the altered fingerprints, collected over the course of the past years by different forensic and border control agencies (designated as POL_DB). Both of these datasets are further described below.

2.1.1 UNIL_DB dataset

The UNIL_DB dataset of fingerprints amounts to 819 inked rolled fingerprints, recorded at 500dpi (1:1). The fingerprints are split into three different categories based on the level of alteration as assessed by a fingerprint expert: altered, slightly altered, non-altered and each of these categories is further classified into corresponding sub-categories (see table 1 below). Note that most of the discovered alterations are likely to have been acquired involuntarily (although the truth of the matter remains unknown).















Table 1 – different categories of UNIL_DB dataset

Class	Altered (158)	Slightly-Altered (217)	Non-Altered (444)
Categories	Burns	Many Lines	Perfect
	Large Scar(s)	Small Burns	Distorted
	Multiple Scars	Small Scar(s)	Contrast Issues
	Warts	Other	Small Line
	Other		Other

The dataset counts 158 altered that will be the targets of the detection against 444 non-altered and 217 slightly altered fingerprints. The purpose of the classification algorithm is to be able to detect altered fingerprints from unaltered ones (including fingerprints showing limited alterations that are of unintentional nature such as small scars).

Illustrative examples of fingerprints from the different categories are presented in table 2.

Table 2 – Fingerprints² from the different categories (altered, slightly altered and non-altered)

						
Altered burns	Altered large scar	Altered multiple scars	Altered wart	Altered other	Slightly altered many wrinkles	Slightly altered small burns
						
Slightly altered small scars	Slightly altered other	Non-altered low contrast	Non-altered distorted	Non-altered well rolled and contrasted	Non-altered limited wrinkles	Non-altered other

² For confidentiality reasons, the fingerprints from the dataset cannot be reproduced. The fingerprints presented here are from {6} and presented for illustration purposes.

2.1.2 POL_DB dataset

The images in POL_DB dataset are 935 fingerprints at 500dpi (1:1), split into two categories – altered / non-altered, based on the source of their origin. In 140 cases a non-altered version of the altered fingerprint is available. While the previous dataset contained exclusively rolled fingerprints, the POL_DB dataset is distributed between rolled and flat fingerprints (examples are shown in figure 1 below). 795 fingerprints in this dataset are altered. This dataset contains multiple cases of voluntary alterations of various types.

2.2 Performance evaluation

The performance of the different algorithms is evaluated on their ability to correctly detect and classify an altered fingerprint, but also on their false positive rate *FPR* (i.e. claiming that a fingerprint is altered when in fact it was not). Indeed, this last rate is crucial when we consider the application of the algorithm in operations. False negative rate *FNR* will be shown for the algorithms tested as well (i.e. claiming that a fingerprint is non-altered when in fact it was altered). As in literature false positive rate (and also of false negative rate) of attack detection algorithms is often confused with biometric performance testing metrics (i.e. false match rate, false non-match rate) [9], this document will align the metrics with the definitions of ISO/IEC 30107-3 Biometric Presentation Attack Detection – Part 3: Testing and Reporting [10]. According to ISO/IEC 30107-3 the accuracy of AFDA methods shall be reported in terms of **normal presentation classification error rate (NPCER)**, which is defined as *proportion of normal presentations incorrectly classified as presentation attacks* [10] for the false positive rate and on the contrary in terms of the **attack presentation classification error rate (APCER)**, which is defined as *“proportion of presentation attacks incorrectly classified as normal presentations“* for the false negative case [10].

The Decision Error Trade-off plots (DET) [11] together with the Equal Error Rate (EER) present a standard measure in biometric system performance evaluation [9] and will be used as well for presentation attack detection evaluation. The EER value extracted from the DET plot is to be taken purely as an indicative measure, as the choice of the operating points on the DET curve is arbitrary and depends on the application used.

3. The algorithms

While absolutely no control was provided over the reference state-of-the-art algorithm (supplied for the evaluation as a black-box), the approaches of UNIL and GUC differ in both – features extracted and classification algorithms used.

3.1 The reference state-of-the-art algorithm

The reference algorithm, based on the work described in [1] classifies individual images as either altered or non-altered, and alongside the binary

categorical decision provides an a-posteriori “quality score” in a fixed range (0-255).

This reference algorithm has been trained on a separate dataset described in [1] and satisfies the condition of evaluating previously unseen data, based on its single-input/single-output functionality.

3.2 The two additional developed algorithms

Both of these algorithms are operating in fully controlled conditions, meaning with an in-depth understanding of the entire process – from the image, through the feature extraction and classification algorithm training down to the a-posteriori score and binary classification. Both UNIL and GUC approaches are based on different set of features and are presented below.

3.2.1 The UNIL algorithm

UNIL algorithm uses the output of the latest generation of the Universal Latent Workstation (ULW developed by the FBI in collaboration with NOBLIS [12]), namely spatial quality information, the number of minutiae extracted from different quality areas coupled with the convex hull surface of the finger information (see figure 1 for an example).

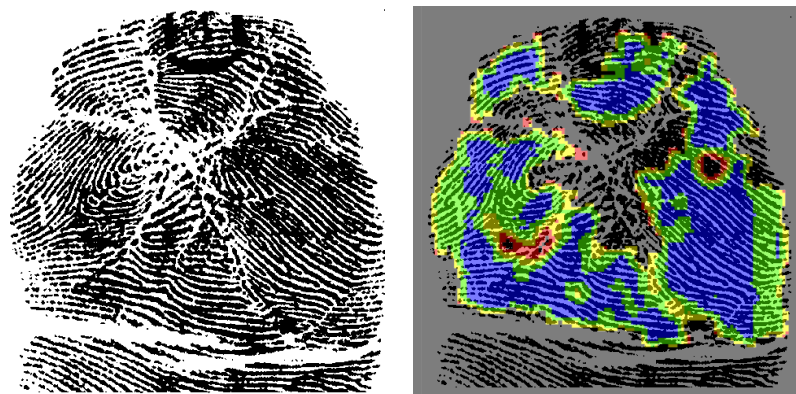


Figure 1 – ULW quality map processing stage (image originating from [13])

The basic functionality of the UNIL algorithm is shown in figure 2. For each incoming image, it performs a feature extraction using the ULW. The resulting feature vector for each fingerprint counts 15 variables (such as quality of the impression, surfaces encapsulating different quality zones, number of minutiae present within the zones of interest, overall surface of the fingerprint, etc.), which are separately normalized³ by z-score prior to using the classifier.

The features extracted are fed into the classification algorithm on the leave-one-out basis. The features in the altered / non-altered datasets serve for the algorithm training and the features extracted from the left-out fingerprint for

³ Any other score normalization method can be used instead of the standard score normalization used. The main objective of the normalization was to unify the feature space in order to boost the performance of the classification algorithms.

testing. This way over-fitting is reduced, as each “left-out” fingerprint represents previously unseen (out-of-the-bag) data points.

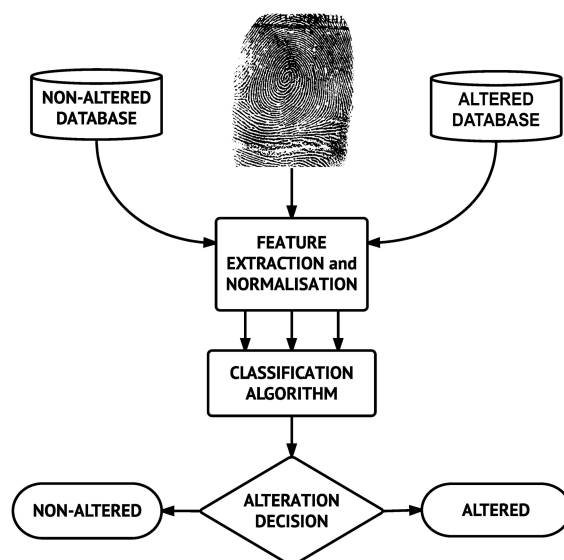


Figure 2 – UNIL altered fingerprint detection algorithm flowchart

A range of supervised classification algorithms have been evaluated for the classification task, namely random forest, neural networks, KNN, SVM. The *caret* R toolbox⁴ served as a testing platform [14]. A leave-one-out cross-validation⁵ scheme has been adopted. The results of the best performing classifiers, namely random forest and SVM will be presented in the following section.

3.2.2 GUC algorithms

Four different methods developed at the GNU have been used to extract feature vectors from the fingerprint images in both datasets. The algorithms are described in detail in the publication J. Ellingsgaard, C. Sousedik, C. Busch: "Detecting Fingerprint Alterations by Orientation Field and Minutiae Orientation Analysis", in Proceedings of the 2nd International Workshop on Biometrics and Forensics 2014 (IWBF 2014), 27-28th March 2014, Valletta, Malta, (2014), which documents the achievements that were reached in the INGRESS projects [15].

OFA: The Orientation Field Approximation (OFA) method is a re-design of the approach originally proposed in [1]. The OFA method uses a mathematical model for constructing an approximation of an estimated ridge flow of the fingerprint. The analysis identifies discontinuities based on differences of the ridge flow approximation and estimation, e.g. areas where the approximation is unable to correctly simulate the actual fingerprint image. A natural approach for extracting ridge orientation is based on computation of gradients in the fingerprint image. Block-wise averages of gradients have multiple purposes

⁴ <http://topepo.github.io/caret/index.html>

⁵ In the leave-one-out cross-validation we iteratively train the classification algorithm using all-but-one feature vectors. In other words, we retrain the classification algorithm for each feature vector.

when processing fingerprint images. Typically, the orientation (or gradients) of each pixel is first smoothed using an averaging filter from a larger area of the image before assigning block-wise orientation averages. The same averaging technique is used in both cases. Altered areas in a fingerprint, e.g. around scars and obliterated areas, can result in discontinuous or unnatural changes in the orientation field. The approximated orientation field will not be able to accurately represent these abrupt and irregular changes caused by alterations. Unaltered fingerprints of good quality will therefore only have small errors around singular points, whereas altered fingerprints can additionally have errors in scarred or mutilated areas.

MOA: The Minutia Orientation Analysis (MOA) method analyses the minutiae distribution in detail. Fingerprint alteration significantly affects the distribution of minutiae by severe skin distortion introduced during the process of alteration. Abrupt ridge endings produced by scars and unusual ridge patterns formed by mutilation will result in additional spurious minutiae. The additional spurious minutia that is caused by alterations will be located along edges of the critical areas. The MOA method conducts an additional local analysis of each detected minutia in order to identify discontinuities and changes in the orientation. The density maps from each analysis are normalized to lie in the range of $\{0,1\}$. The feature extraction will construct high-dimensional vector, which includes histograms in various bins in the range $\{0,1\}$, which are computed for each of the nine image cells (i.e. subimage).

MDA: Again the Minutiae Distribution Analysis (MDA) method is as OFA a redesign of the approach proposed in [1]. Minutiae are located at ridge endings or ridge bifurcation. In this analysis the minutiae extractor *MinDtct* in NBIS [16] is used to extract minutia from a fingerprint. The analysis is based on the observation that the minutiae distribution of altered fingerprints often differs from that of natural fingerprints [1].

SPD: The Singular Point Density Analysis (SPD) method inspects changes in the pixel-wise orientation field. It is based on the local entropy and uncertainty of orientations around scarred and mutilated areas and uses common techniques to extract core features of a fingerprint. Local areas of high curvature will be found using the Poincaré index. This is a common method for extracting singular points in which some altered regions share similar characteristics. Quality measurements of friction ridges are merged into the analysis in order to diminish the effect of uncertainties in poor quality or heavily obliterated areas. Gabor filters are used to evaluate the quality of ridges.

A feature level fusion⁶ [17] was used to boost the performance of the classification algorithm in the following configurations: SPD-MOA, SPD-MDA, OFA-MOA and OFA-SPD.

Based on the previous experiments [15] and taking into account a large number of features extracted (160 features per fingerprint) compared to relatively small dataset a SVM classifier in a leave-one-out cross-validation was used for all the GUC methods. The results obtained are presented in the following section.

4. The results

The performances of the different algorithms are presented using DET plots (figure 4) and summarized in a table 3 below.

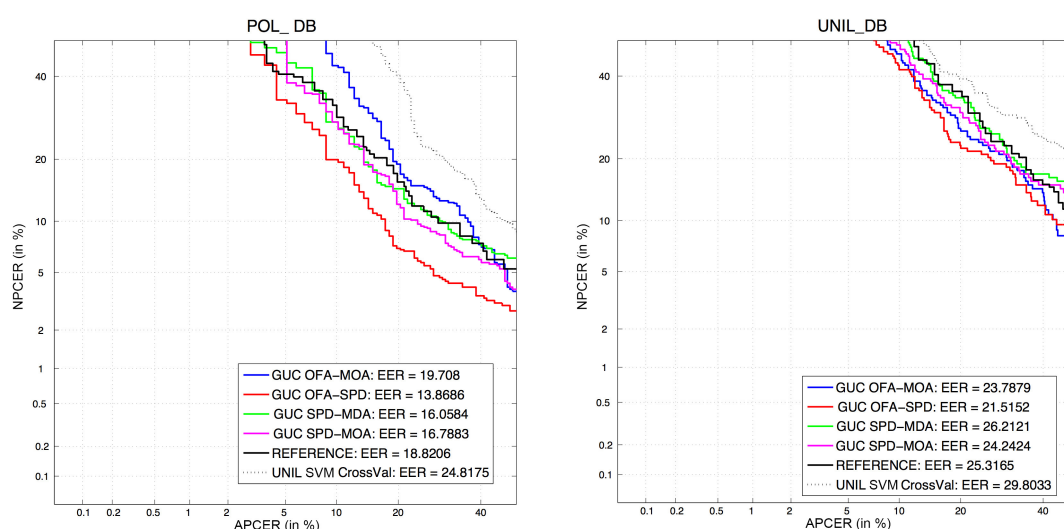


Figure 3 – Algorithm performance on the POL_DB and UNIL-DB datasets

From the 6 different altered fingerprint detection algorithms evaluated, the best overall performance on both UNIL_DB and POL_DB datasets is achieved by the GUC OFA-SPD algorithm as shown in figure 4. It achieved 21.51% EER on the UNIL_DB dataset and 13.86% EER on the POL_DB dataset (see table 3).

Table 3 – Different algorithm performance on both UNIL datasets

Algorithm	UNIL_DB EER {%}	POL_DB EER {%}
GUC OFA-MOA	23.78	19.7
GUC OFA-SPD	21.51	13.86
GUC SPD-MDA	26.21	16.05
GUC SPD-MOA	24.24	16.78
REFERENCE	25.31	18.82
UNIL SVM	29.8	24.81

⁶ By feature level fusion we mean composition of a super-vector by concatenating two or more feature vectors of each fingerprint. Different approaches to fusion in biometrics can be reviewed for example in [10].

Major differences were observed between the performance on the POL_DB and UNIL_DB datasets. While the POL_DB dataset featured predominantly altered fingerprints (797 out of the 935), the majority of the fingerprints in the UNIL_DB dataset were non-altered (444), followed by slightly altered (217) and altered (158).

Table 4 – Different algorithm performance in terms of false positive rate

Algorithm	NPCER POL_DB	APCER POL_DB	NPCER UNIL_DB	APCER UNIL_DB
GUC OFA-MOA	14.41 %	27.62 %	24.68 %	26.77 %
GUC OFA-SPD	14.34 %	23.98%	24.68 %	23.6 %
GUC SPD-MDA	14.42 %	26.32 %	20.88 %	36.3 %
GUC SPD-MOA	19.8 %	27.01 %	24.68 %	26.47 %
REFERENCE	2.87%	67.72%	7.55%	43.4%
UNIL SVM	7.02%	59.16%	9.08%	62.03%

From the results shown in table 4 we can conclude, that while the alteration detection algorithms tested achieve reasonable performance in the task of detection of altered fingerprints, they are less effective in the task of correctly classifying the non-altered fingerprints. We note also that the overall performance of all of the altered fingerprint detection algorithms evaluated is nowhere near low EER (smaller than 10%) reported in [1]. Reasons for such differences are likely to be related to differences of datasets.

Table 4 also indicates that all of the GUC supplied fingerprint alteration algorithms appear to be reasonably balanced when comparing the false positive and false negative rates compared to the reference and UNIL alteration detection algorithms.

From an operational perspective in context (a) described above, if we want to achieve a reasonable detection rate, the rate of false alarms will be such that I will hinder the overall border-crossing process and affect the impact of the technology on “innocent” citizens. In context (b), the detection algorithms can have a role to play.

5. Conclusions

The purpose of the research was to comparatively test and assess the performance of three alteration detection algorithms. A reference algorithm based on [1] against two specific developments carried in the context of this project. These algorithms have been tested against two distinct datasets showing different levels of alteration: the first dataset with alterations largely to unintentional events on the patterns, the second with alterations arising in majority from voluntary attempts to alter fingerprints. On both datasets all

algorithms led to ERR between 13% and 30%. The algorithms developed by GUC showed the best performance.

Undoubtedly, there is room for progress in this area. Given the errors rates obtained in this study, it is hard to design an operational process with an early detection of altered fingerprints of individuals passing through border controls. The reason being that in order to achieve a desired rate of positive detection (say at least above 70%), the rate of false detection will be such that the treatment of false alarm will be unmanageable. However, in the context of triage of prints to be submitted in distinct comparison algorithms in an AFIS system, these algorithms will provide benefits.

Acknowledgements

This research was conducted in the scope of the INGRESS project, funded by the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 312792.

References

1. Feng, J., A.K. Jain, and A. Ross. *Detecting Altered Fingerprints*. in *Pattern Recognition (ICPR), 2010 20th International Conference on*. 2010.
2. Cummins, H., *Attempts to Alter and Obliterate Finger-Prints*. *Journal of Criminal Law and Criminology (1931-1951)*, 1935. **25**(6): p. 982-991.
3. International Organisation for Standardisation, *ISO/IEC 30107-1 Biometric presentation attack detection - Part 1: Framework*. 2015.
4. *Altered Fingerprints: A Challenge to Law Enforcement Identification Efforts*. FBI Law Enforcement Bulletin 2015; Available from: <https://leb.fbi.gov/2015/may/forensic-spotlight-altered-fingerprints-a-challenge-to-law-enforcement-identification-efforts>.
5. Yoon, S., Z. Qijun, and A.K. Jain. *On matching altered fingerprints*. in *Biometrics (ICB), 2012 5th IAPR International Conference on*. 2012.
6. Yoon, S., J. Feng, and A.K. Jain, *Altered fingerprints: Analysis and detection*. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012. **34**(3): p. 451-464.
7. Jain, A.K. and S. Yoon, *Automatic detection of altered fingerprints*. *Computer*, 2012. **45**(1): p. 79-82.
8. Antonelli, A., et al., *Fake finger detection by skin distortion analysis*. *Information Forensics and Security, IEEE Transactions on*, 2006. **1**(3): p. 360-373.
9. International Organisation for Standardisation, *ISO/IEC 19795-1 Biometric performance testing - Part 1: Principles and framework*. 2015.

10. International Organisation for Standardisation, *ISO/IEC 30107-3 Biometric presentation attack detection - Part 3: Testing and reporting*. 2015.
11. Martin, A., et al. *The DET Curve in Assessment of Detection Task Performance*. in *EuroSpeech*. 1997. Rhodes, Greece.
12. Hicklin, R.A., J. Buscaglia, and M.A. Roberts, *Assessing the clarity of friction ridge impressions*. *Forensic Science International*, 2013. **226**(1-3): p. 106-117.
13. Samishchenko, S.S., *Atlas of the Unusual Papilla Patterns / Atlas Neobychnykh Papilliarnykh Uzorov*. 2001, Moscow: Urisprudentsiia. 307.
14. Kuhn, M. and K. Johnson, *Applied Predictive Modeling*. 2013, New-York: Springer-Verlag.
15. Ellingsgaard, J., C. Sousedik, and C. Busch. *Detecting fingerprint alterations by orientation field and minutiae orientation analysis*. in *Biometrics and Forensics (IWBF), 2014 International Workshop on*. 2014.
16. Watson, C., et al., *User's Guide to NIST Biometric Image Software (NBIS)*, NIST, Editor. p. 207.
17. Ross, A. and A.K. Jain, *Information fusion in biometrics*. *Pattern Recognition Letters*, 2003. **24**: p. 2115-2125.