Analysis, Comparison, and Assessment of Latent Fingerprint Image Preprocessing

Haiying Guan*, Paul Lee*, Andrew Dienstfrey*, Mary Theofanos*, Curtis Lamp*, Brian Stanton*, and Matthew T. Schwarz+
*National Institute of Standards and Technology Gaithersburg, MD, USA
*Email: {haiying.guan, paul.lee, andrew.dienstfrey, mary.theofanos, curtis.lamp, brian.stanton}@nist.gov
+ Schwarz Forensic Enterprises, Inc., +Email: matt@schwarzforensic.com

Abstract

Latent fingerprints obtained from crime scenes are rarely immediately suitable for identification purposes. Instead, most latent fingerprint images must be preprocessed to enhance the fingerprint information held within the digital image, while suppressing interference arising from noise and otherwise unwanted image features. In the following we present results of our ongoing research to assess this critical step in the forensic workflow. Previously we discussed the creation of a new database of latent fingerprint images to support such research. The new contributions of this paper are twofold. First, we implement a study in which a group of trained Latent Print Examiners provide Extended Feature Set markups of all images. We discuss the experimental design of this study, and its execution. Next, we propose metrics for measuring the increase of fingerprint information provided by latent fingerprint image preprocessing, and we present preliminary analysis of these metrics when applied to the images in our database[1]. We consider formally defined quality scales (Good, Bad, Ugly), and minutiae identifications of latent fingerprint images before and after preprocessing. All analyses show that latent fingerprint image preprocessing results in a statistically significant increase in fingerprint information and quality.

I. INTRODUCTION

Latent fingerprints are friction ridge impressions left unintentionally on the surface of an object. Images of latent fingerprints can be obtained (i.e., "lifted" or "developed") through numerous methods ranging from precision photography to complex physical and chemical processing techniques [1]. Latent fingerprint evidence plays an important role in forensic science and is routinely used as evidence to convict offenders of crimes. From the unintentional deposition and complexity of acquisition, it follows that the initial latent fingerprint images collected directly from a crime scene may be incomplete or hard to visualize, leading to images of very poor quality. Lighting, pressure, and underlying surface qualities such as texture and color are just a few factors that may affect the quality of a fingerprint digital image [2].

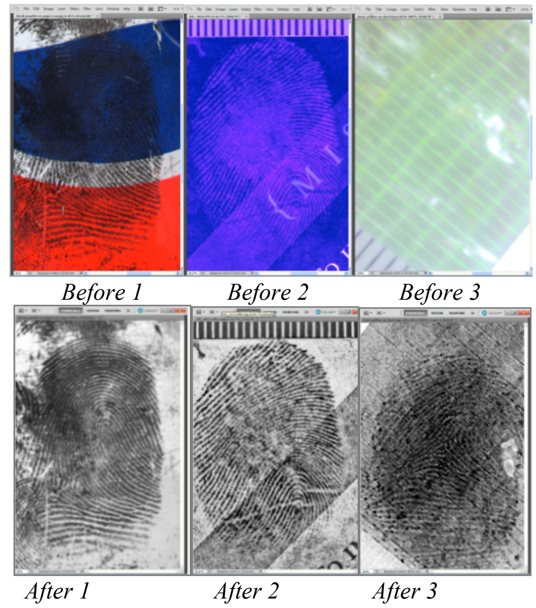


Figure 1: Latent Fingerprint Before and After Preprocessing Examples.

Consider the top row of latent fingerprint examples shown in Figure 1. Due to the low signal quality of the fingerprint in relation to other systematic image features, such as color, pattern, text, etc., the initial fingerprint image quality may be of only marginal value for identification. In some extreme cases, latent prints are identified as "no value." In this context, "no value" is a formal determination that the print is of such poor quality that no identification—neither individualization nor exclusion—is possible. This is true regardless of the score of a potential match between the latent to other prints held in a database [3]. Thus, potentially usable latent images are classified as unsuitable for feature markup, entry into databases, or input into fingerprint identification software to search for matches.

To mitigate this issue, current practice allows for a Latent Print Examiner (LPE) to perform image preprocessing prior to markup and feature analysis. The forensics community currently uses a variety of image analysis and preprocessing tools to significantly improve the quality of these images and enhance fingerprint features. The bottom row of latent fingerprint examples in Figure 1 are the result of preprocessing the images in the top row, which show that the changes can be extraordinary. For example, the ridge patterns are significantly more visible in the first and second images. In the third, a grid-like background has been removed to

1 The latent fingerprint images are from a training data set provided by the course from FORAY technologies and Schwarz Forensic Enterprises, Inc.

reveal fingerprint information "underneath". In short, latent fingerprint image preprocessing can transform raw images with little or no value into ones suitable for evidentiary analysis.

Some of the key components of evidentiary analysis such as automatic fingerprint feature extraction, matching, and print type identification are well studied, regulated, and implemented in existing systems. However, the preprocessing step is currently overlooked. Preprocessing is the first step of the analysis workflow, and can be critical to the accuracy of subsequent analysis [4]. For instance, an image with extraneous noise introduced during the preprocessing phase may lead to incorrect feature extraction, which may have a negative effect during the matching and identification stages. Despite the importance of this step, there exist few databases for controlled experimentation and scientific study, and even fewer standards. Detrimental consequences for reproducibility, traceability, and quantification of accuracy naturally follow. Our research hopes to shed some light on this topic.

Previously Guan et al. [4] presented the results of the collaboration with forensic scientists to design and collect a database of latent fingerprint images consisting of: original latent fingerprint images ("Before"), their pre-processed counterparts ("After"), and documentation of the image transformation procedures executed during the preprocessing stage. The paper also proposed a new latent print quality measurement metric. Here we extend this previous work in two significant ways. First, we designed a round-robin experiment contracting an independent set of LPEs certified by the International Association for Identification to provide Extended Feature Set (EFS) [5] markups for all images. Such markup information significantly increases the value of this database. Next, we conducted experiments analyzing the fingerprint data quality in the several image classes within the database. More specifically, we use three metrics—value determination, minutiae count, and quality confidence score—to compare changes in image quality and fingerprint information that result from preprocessing. We find that examiners mark more minutiae on the After latent image than the Before. Additionally, examiners categorize the preprocessed images higher on a quality scale, resulting in an improved quality confidence scores as compared to the Before images. Finally, our analysis shows that LPEs identify more minutiae in color images than in grayscale (i.e., "Before Color" as compared to "Before Gray"). This suggests that there may be value in having color images available for input in Automated Fingerprint Identification Systems (AFIS).

In summary, we intend that these results will provide foundational elements for a systematic and scientific basis for latent fingerprint analysis. Furthermore, we hope that they may serve as a test case for the development of comparable analysis for other image-based methods in forensic science in the future.

II. METHODOLOGY

2.1 Study Objective

The objective of this study is to determine the quantitative value of latent fingerprint images before and after preprocessing, focusing on notable changes in detectable fingerprint minutiae. Additionally, we are also looking at quality scale changes and quality map changes in color vs. grayscale images. Note that unlike other studies [3][6], this study does not compare the markups among examiners nor evaluate examiners' performance. No identifying information is kept or linked to the images. The chief goal is to determine whether and to what extent latent fingerprint preprocessing improves the ability to gain information in the identification of latent impressions, as well as to what extent it transforms latent images with no comparison value into images that can be used for analysis.

2.2 Initial Dataset

Previously we created a database of latent fingerprint images isolating several steps within the preprocessing workflow. This database includes 89 latent fingerprint image pairs that were developed using a cross-section of forensic field work techniques including: ninhydrin, silver magnesium powder, white powder, bi-chromatic powder, bi-chromatic mag powder, and black ink. The original images were scanned by high-resolution flatbed scanners and subsequently preprocessed within Adobe Photoshop, the primary image analysis tool used by latent examiners practicing today.[2] The image transformations in the preprocessing workflow were recorded in Adobe Photoshop and saved in an accompanying metadata file as per existing best-practice guidelines [7]. The result was a collection of triplets consisting of original image, processed image, and metadata file. This database has proven to be an invaluable source of controlled data for developing scientific analyses of forensic image preprocessing.

2.3 Experimental Design

A team of 9 independent LPEs was assembled. Images were distributed and presented to the LPEs in a pre-determined order per the following assignment criteria:

- Examiners receive and mark up one image at a time. No information regarding whether the image has undergone preprocessing is given to the examiner.
- Generally, an examiner will not mark up the After image corresponding to any previously seen Before image.

[2] Any mention of commercial products or reference to commercial organizations is for information only. It does not imply recommendation nor endorsement by NIST, nor does it imply that the products mentioned are the best available for the purpose.

- Each examiner receives at least one good quality image and one bad quality image. The remainder will be a mix of good, bad, and ugly images.
- Finger source distribution is randomized, ensuring approximately the same distribution amongst examiners. The study has three phases:

Phase I: Examiners mark Before Grayscale images.
Phase II: Examiners mark the corresponding Before Color images. It is acceptable for an examiner to view the Grayscale image to assist in this markup process.
Phase III: Examiners mark After Grayscale images.

In each phase, the examiners were given a list of images to analyze. The results of the previous phase were collected prior to release of the next phase's image set. We implemented a sorting algorithm to assign the images to the examiners in different phases, attempting to satisfy the above design criteria as best as possible.

When performing the markup, LPEs assume the images provided are the only images available, and that physical evidence, lift cards, fingerprint cards, additional exemplars, and different images of these prints are not available. For consistency, LPEs use Universal Latent Workstation Latent Editor software, ULW-EFS 6.4.0 or newer (https://www.fbibiospecs.cjis.gov/Latent/PrintServices), to do markup. Each received a standardized instruction document on how to proceed at all stages of the study to guide their work.

Upon receiving an image, the LPEs were required to perform the following steps: (1) Paint the quality (clarity) of the latent (throughout the entire region of interest), (2) Annotate EFS features within the image, and (3) Record the final value impression determination of each print using the Good, Bad or Ugly (GBU) quality scale [2].

The EFS was developed by Noblis (http://www.noblis.org) in collaboration with the Federal Bureau of Investigations and standardizes the diverse fingerprint image metadata considered useful for identification analysis. The EFS augments the ridge-flow information contained within a fingerprint image by inserting standardized indications of features including: ridge quality maps, incipient ridges, minutiae, cores, deltas, and others. LPEs followed instructions of the ACE-V (Analysis, Comparison, Evaluation, and Verification) methodology to assess images for the presence of: friction ridges, fingerprint information available, the confidence of such information etc. Enhancement tools present in the ULW Latent Editor software or in any other software that the examiner might have available were strictly forbidden. Under our study, three versions of each latent image were marked by examiners and the EFS information are held within the database: the original Before Color image, the original Before Grayscale image, and the preprocessed After image. Each latent image was marked by two different LPEs.

In addition, we collected rolled print images and performed EFS markups for every finger source. The markup of these prints allows us to furnish "ground truth" EFS data, which serves as a basis of comparison between Before and After images. Prior to comparing markups of a latent image to its corresponding rolled image, the two images must be aligned. In forensic practice, such image registration is accomplished as a sub-task of EFS feature comparison. In the present study, we sought to eliminate this source of variability. For each latent record an independent examiner identified a number (>3) of benchmark minutiae that could be found on both the latent and its associated ground-truth. Corresponding minutiae were indicated by color. Ideally these features are as separated as possible throughout the region of interest. A color point detection algorithm identified the locations of corresponding features, and a least-squares algorithm was used to estimate the rigid transformation parameters (rotation and translation) to transform ground-truth orientation to that of the latent.

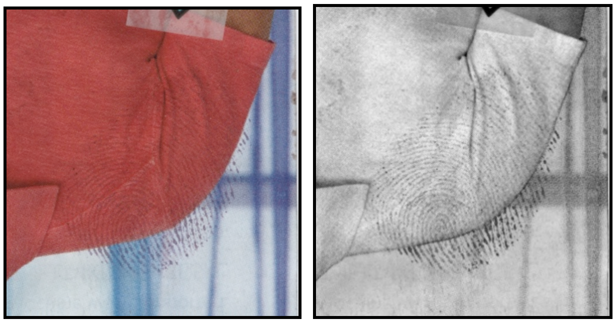
III. LATENT PREPROCESSING DATABASE

The latent preprocessing database contains 89 fingerprint records. Structurally, each record is a directory containing: several image files, their EFS markups saved in the Latent Friction Features Search format (defined by the Electronic Biometric Transmission Specification described in https://www.fbibiospecs.cjis.gov/ebts/), the source finger's card image and its EFS markup, and metadata files. The various metadata files include: examiner ID, source finger ID, GBU value determinations, image resolutions, specific latent lifting techniques, etc. In total, there are 28 files for each record. We describe these in more detail below.

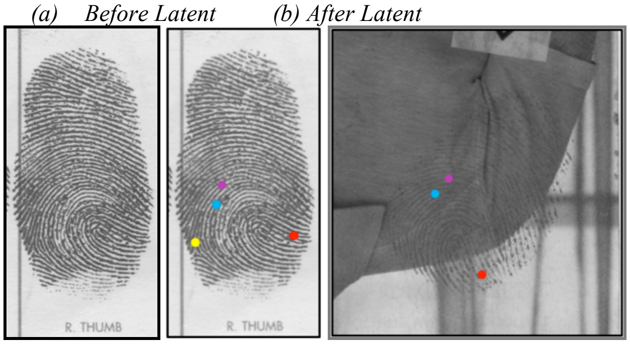
3.1 Image files

The Before latent image is the latent fingerprint scan that has yet to undergo preprocessing. Note that while preprocessing is performed in Adobe photoshop on high-resolution images scanned at 1200ppi, the ULW-EFS 6.4.0 requires images to be 1000ppi. Thus there are three Before image files in our dataset: the original Before latent color image in its native scanned resolution of 1200ppi, its downscaled color version of 1000ppi, and its downscaled grayscale version of 1000ppi. All down sampling was done using OpenCV (http://www.opencv.org/) bicubic interpolation.

Once the original Before Color image scan at 1200ppi undergoes preprocessing, the After Grayscale image is obtained. Unlike the Before category with its grayscale and color versions, After images are only in grayscale. These high contrast versions of latent images are commonly used in AFIS [8] search or matching; the grayscale property is required by this system. In the database, there are two After files: one latent image at the native scanned resolution (1200ppi) and one scaled down to 1000ppi. Once again, the 1000ppi image is required by the ULW software. Figure 2 (a) and (b) show a sample pair of a Before Color latent image and its After preprocessed image.



(a) Before Latent (b) After Latent



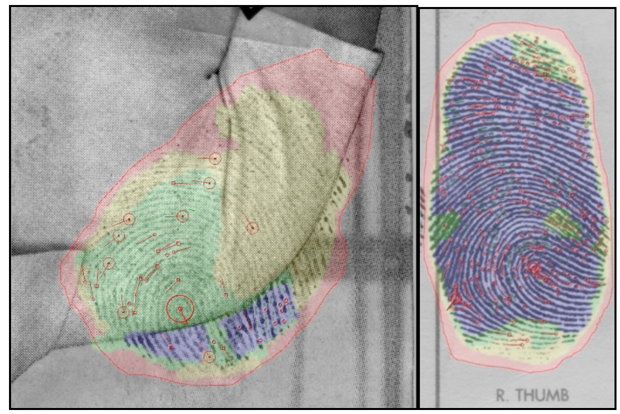
(c) Card (d) Card Minutiae Annotation (e) Latent Minutiae Annotation

Figure 2: Images in latent preprocessing database

We also collect the source finger's card image and its EFS markup file as the reference ground-truth for the latent images. Figure 2 (c) is an example of the original, unmarked finger source's card image. To align the minutiae in latent image with the minutiae in card image, the latent examiner uses colored dots to annotate at least three minutiae in the card image and in the After image (Figure 2 (d) and (e)). Note that the color dot radius is enlarged for illustration purposes. The actual dot radius within the image is 5 pixels at 1200ppi.

3.2 Markup Files

Following the experimental design, in the first round two latent print examiners mark up the EFS Before Color and Before Grayscale images. Examiners inspect and mark miscellaneous minutiae, bifurcations, incipient ridges, ridge endings, dots, the region of interest, and distinctive quality areas. Bifurcations are marked with squares, incipient ridges by green lines, ridge endings by small circles with trailing tails, and deltas by large circles with bisected centers. Minutiae too obscure to classify are represented by lone circles of two sizes to represent the uncertainty, with higher quality unknown minutiae corresponding to the smaller of the two. When the first round is complete, After Grayscale images are released. Two examiners repeat the process for the preprocessed images, marking all the same feature categories. In total, seven EFS markup files in the EBTS format (.lffs) are collected: a markup of the card image, and two independent markup files for the Before Grayscale, Before Color, and After Grayscale latent images. Each of these markup files are accompanied by a corresponding text document containing data such as: minutiae coordinates, minutiae types, the quality map matrix, image metadata, etc. The After Grayscale markup of Figure 2(b) is shown below in Figure 3(a). The markup of the card(source) image is also collected as shown in Figure 3(b).



(a)After Grayscale Markup (b) Card(Source)Markup

Figure 3: Full Markup Images

Latent Quality Mapping is used to document the level of confidence in the marked features. Image quality is documented by painting over the image using standard color definitions for latent region quality markup. The color scale range includes cyan, blue, green, yellow, red, and black, in order of the largest level of confidence to the smallest. Teal indicates that there are clear definitive ridge edges plus, dots, pores and level three detail throughout the area, blue indicates that there are clear ridges, and green indicates that it is certain that every minutia in the area is marked. Note that green (or better) means that the examiner is certain of the presence of all minutiae they've marked in that region AND they are certain that there are no unmarked minutiae. Yellow indicates that the examiner is not confident in the presence or location of marked minutiae and there may be minutiae in the area that they did not mark. Finally, red indicates any discontinuities (e.g., smears), and black indicates the lack of ridge data in a particular area of the image. The Latent Quality Markup of Figure 2 can be seen above in Figure 3. Additional information about the ridge quality map and feature markup on this study can be found in the ANSI/NIST standard [9] and Markup Instructions for Extended Friction Ridge Features [7].

3.3 Metadata

Alongside the various images, six metadata files are also included. The first information spreadsheet contains most experimental design related details, including the source fingerprint, the examiner IDs, the latent acquisition procedure used, the GBU classification of the Before and After files, and the various image resolutions. The second information spreadsheet holds relational database details. The action history of the preprocessing editing session are recorded in a word document, including file creation, color channel selection, color scheme conversion, use of the burn tool, etc. The last three files are all single item files that hold the translational matrix for the latent image to card image shift, the manually annotated rectangular region of interest coordinates, and the manually annotated polygon coordinates respectively.

IV. ANALYSIS

Data points represented here include the 89 sample records. The significance of preprocessing was determined by analysis of changes in three quality metrics: image value determination, minutiae count, and quality confidence score.

4.1 GBU Value Determination Comparison

Upon receiving the latent image, examiners determine the overall latent quality using the Good, Bad, and Ugly scale [2]. After pre-processing, the images were examined again and re-categorized. Table 1 shows the number of images classified as Good, Bad and Ugly for the Before and After datasets as well as the change in quality scale determination after preprocessing.

Table 1: Value Determination and Re-Categorization

After / Before	Good	Bad	Ugly	Total
Good	25	0	0	25
Bad	23	12	0	35
Ugly	5	16	8	29
Total	53	28	8	89

Across each Before row (Before Good, Before Bad, and Before Ugly), the number of images initially classified as such are noted. Of the 89 records, 25 of the Before images were classified as Good, 35 were classified as Bad, and 29 were classified as Ugly. Down each After column (After Good, After Bad, and After Ugly), the number of images re-classified into these categories can be seen. Of the 64 images previously categorized as either Bad or Ugly, 28 were assessed to be of Good quality after preprocessing (Bad: 23, Ugly: 5). The remaining 36 images were split between Ugly images rising to Bad quality (16), and images in both categories staying in their initial determination category (Bad: 12, Ugly: 8). Each of the 25 initial Good quality latent images remained in the Good category. No instances of quality deterioration were found.

These tables show that across quality determination categories, 49.43% of latent fingerprint images showed marked improvement after preprocessing. Excluding the images initially rated Good (as these cannot be improved), we find that 68.75% are improved by preprocessing.

4.2 Minutiae analysis

In our collection, we have three types of latent images: Before Color, Before Grayscale, and After Grayscale. Each latent image was reviewed and marked by two examiners. The original .lffs files were fed through our minutiae reader tool to analyze minutiae markup data. After accounting for differences in resolution, horizontal and vertical offsets, and verifying miscellaneous examiner markups, the result were 6 sets of minutiae coordinates aligned in the same coordinate system: two for Before Color, two for Before Grayscale, and another two for After Grayscale.

To compensate for possible differences between examiners, we designed and implemented an algorithm to identify corresponding minutiae between two markups of the same image. Minutiae correspondence was determined by procedure involving a combination of: minutiae proximity, distance hierarchy assignments, minutiae type, and final manual verification examinations to guarantee the correctness. For a pair of markup files of the same image, this intersection set represents a consensus understanding of an image's feature set, with a singular representation of each minutiae. Unmatched minutiae are retained in the database but are not included as part of the following analysis.

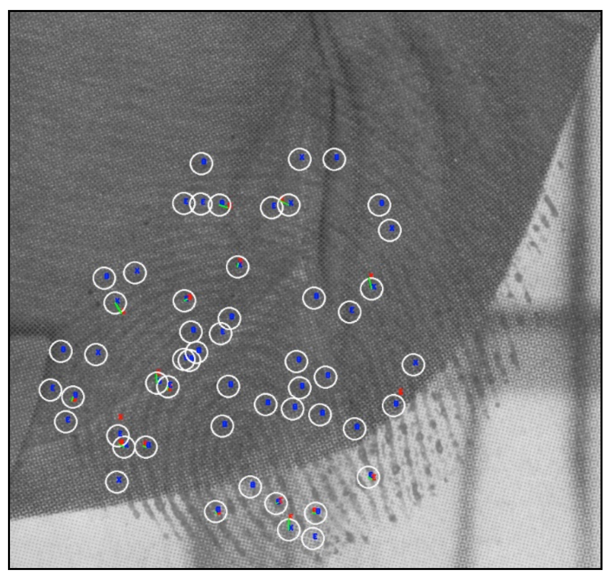


Figure 4: Before Gray and After Grayscale Intersection Minutiae

In Figure 4, both intersection minutiae sets of Before Gray and After Grayscale of Figure 2 can be seen. The Before Gray intersection set of Figure 2 is shown in red, while the After-Grayscale intersection set is shown in blue. The white circle around each After Grayscale minutiae represents the "search area" used to seek out matching Before minutiae, with green lines marking a successful match. Different types of minutiae are marked with different letters, with 'E' used for endpoints, 'B' used for bifurcations, and 'X' used for unclassified minutiae.

We present the results based off the intersection data set below.

4.2.1 Percentage Gain

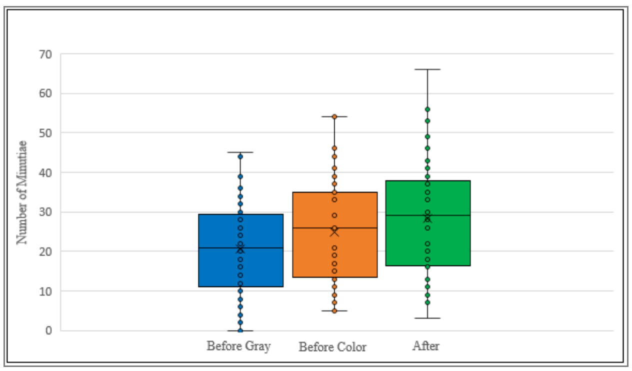


Figure 6: Minutiae Count

Overall minutiae feature count across image types is shown in Figure 6. There was an average of 20.74 minutiae in the Before Grayscale images, 24.82 minutiae in the Before Color images, and 28.44 minutiae in the After images.

The corresponding median minutiae gain percentages are shown in Table 2. With the most commonly used AFIS systems and matching tools limited to or being heavily reliant on grayscale images, only the After latent image in grayscale is currently available to be studied. In the future, based on our experiments, we suggest that After (preprocessed) images in color may also help preserve useful feature information.

Table 2: Minutiae Gain Percentage

Image Comparison	Median	Mean
BG to BC	14.84%	30.55%
BG to AG	34.59%	70.39%
BC to AG	8.82%	30.92%

For each latent image pair, we calculated the increase in minutiae count from the Before image to the After image, and divided by the Before image minutiae count. We then derived the mean and median gain percentages across the entire dataset using this series of percentages. Comparing Before Grayscale to Before Color latent images results in a 14.84% increase in minutiae found. Before Grayscale to After measures in at a 34.59% increase, while Before Color to After results 8.82% increase. Mean percentage gain comes in much higher across the board, with increases of 30.55% for Before Grayscale to Before Color, 70.39% for Before Grayscale to After, and 30.92% for Before Color to After.

The larger outliers in the gain percentage distribution contribute to the differences between median and mean gain, as seen in Figure 7. The figure on the right presents the same graph, but with a reduced scale for easier viewing. The figure only covers two out of the four comparison categories, but the distribution is similar for all four: majority clusters from about -40% to 100%, then decreasing distribution until about 200%, with a few outliers of more than five times the original minutiae count.

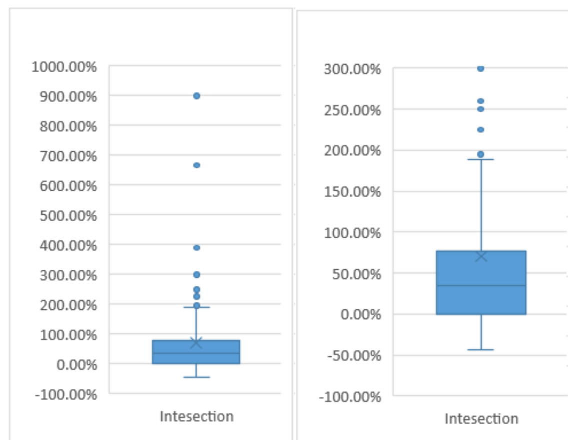


Figure 7: Minutiae Gain Percentage Distribution

4.2.2 Signed Ranks Test

The Wilcoxon Signed Rank Test [10] was used to test the significance of differences in minutiae counts between treatment groups, for example, Before Grayscale to After. Given a paired list of minutiae counts, this nonparametric test computes a score by: 1. rank ordering the absolute value of all differences, 2. reassigning the sign of the difference to the ranked list, and 3. evaluating the signed rank sum (W). Under the null hypothesis that the minutiae count distribution is the same between the two groups, W will be close to zero. Considering this comparison with zero, as all sample sizes are greater than 10, W may be approximated by a normal random variable. We calculate the z-value by dividing the critical value (W) by the standard deviation of its sampling distribution (σw). The standard deviation is derived by taking the square root of $(N_r(N_r+1)(2N_r+1))/6$, where N_r is the sample size.

$$z = \frac{W}{\sigma w}, \sigma w = \sqrt{\frac{N_r(N_r+1)(2N_r+1)}{6}} \quad (1)$$

Based on the z-value, we can determine the two-tailed probability score P (<0.05). This description is brief. For more details see [10].

Like the minutiae gain analysis, all significance tests compared the relationship between Before Grayscale and Before Color latent images, Before Grayscale latent image and After images, and Before Color and After images, using the 89 previously used records. Note the different values of N_r are the result of image pairs with no change in the number of minutiae features identified. In these cases, the sample is not included in the significance analysis. The results can be seen below in Table 3.

The first comparison was done to measure any marked improvement between image qualities of the starting sample. The Before Grayscale to Before Color comparison results in W=-2507 and a Z value of -6.13, leading to a P=<.0001. The second comparison was performed to measure any improvement in minutiae detection due to pre-processing. Comparing Before Grayscale to After, there are 87 samples with W=-2769, resulting in a Z value of -5.86 and P=<.001. The final comparison measures the ability of the Before Color image to preserve feature information. With 85 samples, Before Color to After has W=-1451, Z=-3.18, and P=.0015.

Table 3: Wilcoxon Signed Ranks Test - Minutiae Count

Comparison	N	Test Statistic(W)	Z	P
BG to BC	79	-2507	-6.13	<.0001
BG to AG	87	-2769	-5.86	<.0001
BC to AG	85	-1451	-3.18	0.0015

With α=.05, the Wilcoxon Signed Rank Test indicates that differences between all treatments groups in Table 3 are statistically significant. In other words, more minutiae are identified in the Before Color than their grayscale counterparts (BG to BC), and also, the preprocessed images contain more minutiae than either of the before images (BG to AG, and BC to AG). We note that we also tested for the significance of minutiae count differences using a Random Matched Sample analysis which confirmed the results shown here.

4.3 Quality Confidence Score

In addition to the improvement in image quality that can be derived from GBU value determination and minutiae count gain, we also assess image quality gains by looking at the differences in the quality confidence score of image treatments. The quality confidence score of each markup image is derived by cross referencing the coordinate position of each marked minutiae with the Latent Quality Mapping of the image as detailed in section 3.2. The quality confidence score of a given latent fingerprint is defined as follows: given each marked minutia in the latent image, locate its position in the quality map, and obtain its quality map value given that position. Then we sum up all minutia quality map values and obtain a final single quality confidence score for the image. The quality confidence score measures how thoroughly a LPE could mark the features of the latent image, as well as how confident they are in specific minutiae locations. More minutiae or larger areas of higher quality results in a higher overall quality score.

Each record's quality confidence score gain percentage was collected and averaged, resulting in a mean score gain of 22.81% when comparing Before Grayscale to Before Color, and 23.55% when comparing Before Grayscale to After. Median score gains were slightly more detached, with a 20.00% gain for Before Grayscale to Before Color and 29.38% for Before Grayscale to After. To measure the significance of the result, we again used the Wilcoxon Signed Ranks Test.

Of the 89 records previously used in quality comparisons, 86 were used in conjunction with the Wilcoxon Signed Ranks Test to determine quality change significance between Before Grayscale and Before Color images. Of the three unused records, two were removed due to missing minutiae markup, while the third had the same quality confidence score for both the Before Grayscale and Before Color images. For the Before Grayscale to After comparison, 87 out of the 89 records were used. The two unused records were the same two removed in the Before Grayscale to Before Color comparison due to missing minutiae markup. This holds true for the Before Color and After comparison as well. The results of the test can be seen in Table 5.

Table 5: Wilcoxon Signed Ranks Test - Quality Confidence Score

Comparison	N	Test Statistic(W)	Z	P
BG to BC	86	-3456	-7.440	<.0001
BG to AG	87	-2886	-6.107	<.0001
BC to AG	87	-1014	-2.05	.0324

The first comparison, Before Grayscale to Before Color, has the test statistic (W) = -3456 and the Z value = -7.440, which leads to P = <.0001. The second comparison, Before Grayscale to After, has the test statistic (W) = -2886 and the Z value = -6.107, which leads to P = <.0001. The third, Before Color to After, has the test statistic (W) = -1014 and the Z value = -2.05, which leads to P = .0324. With α=.05, the Wilcoxon Signed Ranks test indicates that the differences in quality score between the Before Grayscale and Before Color, the Before Grayscale and After, and Before Color to After are all statistically significant.

V. CONCLUSION

Currently many prints that could be preprocessed through software are not analyzed or compared because they are deemed "no value" [11]. Furthermore, due to the lack of quantitative techniques for image preprocessing many forensic laboratories currently do not employ or allow image preprocessing software. We hope that the dataset discussed in this paper, complete with images and EFS markups, will provide forensic analysts a testbed for future preprocessing studies.

Along these lines, we designed a series of comparison experiments to examine the effectiveness of preprocessing in relation to feature marks. The quantitative results show that the After latent image is significantly improved by preprocessing. While the scope of this paper is limited to latent fingerprints preprocessing, the design approach and analyses methods are applicable to other biometrics comparative disciplines including handwriting, footwear, tool marks, tread marks, firearm compressions, bite marks, bruising, and so on.

Future work involving the database will include another series of comparisons using the quality map feature. We will also continue to provide techniques and processes enabling latent fingerprint examiners to analyze and compare evidence more effectively, as well as build foundations for future academic research and standards formulation.

REFERENCES

[1] H. C. Lee and R. E. Gaensslen, Eds., *Advances in fingerprint technology*, 2nd ed. Boca Raton, Fla: CRC Press, 2001.

[2] R. Rajkumar and K. Hemachandran, “Latent Fingerprint Enhancement in Preprocessing Stage,” in *International Conference on Computer Science and Intelligent Systems (CSIS’11)*, 2011.

[3] B. T. Ulery, R. A. Hicklin, G. I. Kiebuzinski, M. A. Roberts, and J. Buscaglia, “Understanding the sufficiency of information for latent fingerprint value determinations,” *Forensic Sci. Int.*, vol. 230, no. 1–3, pp. 99–106, Jul. 2013.

[4] H. Guan, A. M. Dienstfrey, and M. F. Theofanos, “A New Metric for Latent Fingerprint Image Preprocessing,” in *2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2013, pp. 84–91.

[5] M. Taylor *et al.*, “Extended Feature Set Profile Specification,” National Institute of Standards and Technology, Gaithersburg, MD, NIST SP 1134, Dec. 2012.

[6] B. T. Ulery, R. A. Hicklin, M. A. Roberts, and J. Buscaglia, “Measuring What Latent Fingerprint Examiners Consider Sufficient Information for Individualization Determinations,” *PLOS ONE*, vol. 9, no. 11, p. e110179, Nov. 2014.

[7] P. Peterson *et al.*, “Latent Prints: A Perspective on the State of the Science,” *FBI Forensic Sci. Commun.*, vol. 11, no. 4.

[8] L. M. I. Dexter and A. J. Pouratian, “Automated fingerprint identification system,” US5869822 A, 09-Feb-1999.

[9] K. C. Mangold, “Data Format for the Interchange of Fingerprint, Facial & Other Biometric Information ANSI/NIST-ITL 1-2011.” NIST Special Publication 500-290 Edition 3, 22-Aug-2016.

[10] R. Lowry, “Concepts and Applications of Inferential Statistics, by Richard Lowry,” *Online Book*, 2015. [Online]. Available: http://onlinebooks.library.upenn.edu/webbin/book/lookupid?key=olbp66608. [Accessed: 03-Mar-2017].

[11] M. Theofanos, B. Stanton, D. Witzke, and M. Schwarz, “Characterizing the Latent Fingerprint Pre-Processing Procedures,” National Institute of Standards and Technology, NIST Internal or Interagency Reports (NISTIR), 2017.