Article

LQMetric: A Latent Fingerprint Quality Metric for Predicting AFIS Performance and Assessing the Value of Latent Fingerprints

Nathan D. Kalka
Michael Beachler
R. Austin Hicklin

Noblis
Reston, VA

Abstract: We describe LQMetric, an automated tool for measuring the image quality of latent fingerprints. The value returned by LQMetric is an estimate of the probability that an image-only search of the Federal Bureau of Investigation's (FBI) Next Generation Identification (NGI) automated fingerprint identification system (AFIS) would hit at rank 1 if the subject's exemplar (rolled) fingerprints are enrolled in the gallery. LQMetric can also be used in assessing the value of latent fingerprints in non-AFIS casework. LQMetric is incorporated into the FBI's Universal Latent Workstation (ULW) software and has been used operationally since 2014. The development of an automated latent fingerprint quality metric was driven by practical use cases including predicting the likelihood of successful AFIS matching; helping examiners determine whether an image-only or human-markup search is more appropriate for a particular latent fingerprint; supporting a quality-directed workflow whereby a backlog is prioritized based on quality or lower quality latent prints are directed to highly experienced examiners; or providing an objective difficulty measure for quality assurance purposes such as flagging complex prints for special handling or additional verification. We describe how LQMetric was developed and trained, how well it predicts NGI AFIS search results, and we also discuss human examiner latent fingerprint value assessments.

Received April 14, 2020; accepted May 28, 2020

Introduction

Latent fingerprints are often key forensic evidence. The value or utility of latent prints is significantly affected by their quality, which affects both the likelihood that they can be used successfully in searching large-scale automated fingerprint identification systems (AFIS) and the ability of expert latent print examiners to reach conclusions. Here we describe LQMetric, an objective, automated algorithm to measure the quality of latent prints, designed both to predict AFIS performance and to augment or replace the informal subjective assessments of quality used by latent print examiners.

LQMetric builds directly on the FBI Laboratory's Latent Quality Study [1, 2], in which 86 latent print examiners each assessed the quality of approximately 70 fingerprint images (out of a pool of 1,090 latent and exemplar fingerprint images), resulting in a total of 5,245 quality assessments. In that study, examiners provided an overall assessment of the quality of each image and indicated within each impression that examiner's degree of confidence in discerning the features in the image. That study also resulted in the development of prototype software tools for the manual or automated definition of local clarity[1] maps, automated overall clarity metrics, and calculation of metrics assessing corresponding clarity for comparisons. As part of the LQMetric project, much of the LQAS functionality was enhanced and incorporated for operational use in the FBI's Universal Latent Workstation (ULW) software[2] [2], which is the software used by latent print examiners in federal, state, and local law enforcement agencies to search the FBI's Next Generation Identification (NGI) system (by way of their respective state or federal conduits or automated biometric systems). Most of the work described here was conducted in 2012 to 2013. LQMetric has been used operationally since 2014 as part of the FBI's ULW software, but no detailed description has previously been published. Since LQMetric was developed, a variety of intriguing fingerprint quality metrics relevant to latent print quality assessment have been proposed that may offer improvements on LQMetric [3–9].

---

[1] **Note on Terminology** – We use "clarity" to refer to the fidelity with which anatomical details are represented in an impression. "Quality" is a broader term, defined in ISO/IEC 29794-1 [10] to include not just fidelity (clarity), but also character (quantity or distinctiveness of the physical features) and utility (appropriateness for use, either in human examinations or by automated systems). A fingerprint quality metric is therefore a means of quantifying the expected utility of a given fingerprint.

[2] FBI ULW software is available at www.fbibiospecs.cjis.gov/Latent/PrintServices.

Latent Quality Metric (LQMetric)

LQMetric assesses quality in two stages (Figure 1). First, a local clarity map is generated using image processing and machine learning algorithms, resulting in a representation of the clarity at each pixel location in an image, defined using the ANSI/NIST-ITL standard [11]. Second, the overall quality is calculated based on clarity map and latent print attributes and is trained and calibrated using AFIS matcher scores to predict the estimated probability of a rank-1 hit of a search of the FBI's NGI AFIS, if the subject is enrolled in the gallery (database). Note that LQMetric was developed and trained on latent fingerprints; for it to be of use for latent palmprints, both the local clarity and overall quality algorithms would have to be retrained.

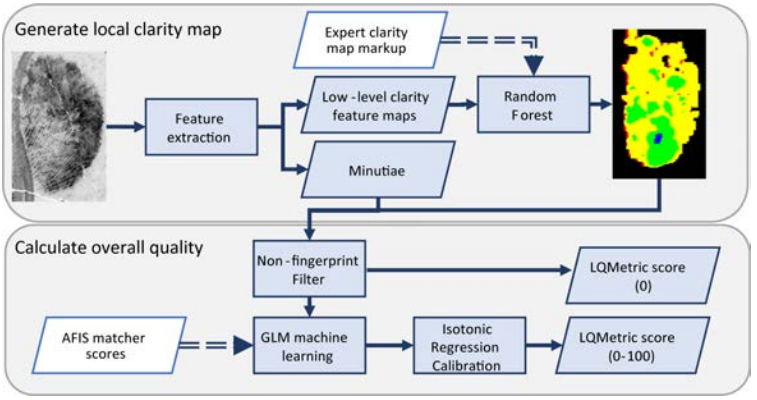

Figure 1

LQMetric process for generating the local clarity map and calculating the overall score. Dashed lines indicate an input that was only used during the training process.

Local Clarity Map Generation

Clarity maps provide a latent print examiner (or algorithm) reviewing the image a standard means of defining the level of confidence in the presence or absence of features in each location in a fingerprint image. LQMetric automatically generates local clarity maps that comply with ANSI/NIST-ITL [11], which is the international standard for interchange of forensic fingerprint data (as well as other biometric data). These local clarity maps are defined as a categorical scale, with a standardized color-coding scheme, with each level defined in terms of the fidelity of reproduction of different types of features at each location in an image grid: level 0 (black) is nonfriction ridge area; level 1

(red) has debatable ridge flow; level 2 (yellow) has clear ridge flow but debatable minutiae; level 3 (green) has clear presence or absence of minutiae but debatable ridge edges; and level 4 (blue) has clear ridge edges. Level 5 clarity (aqua), defined as clear pore detail, was rarely used by human examiners and therefore is not included in LQMetric clarity maps. For example, a green area in a clarity map indicates high confidence in the marked minutiae in that area and that there are no missing minutiae in that area; a yellow area may include missing or false minutiae. Each map has an effective resolution of 125 pixels per inch[3] (ppi, 4.9 pixels per mm), which can be seen as a grid of 4 x 4 pixel cells at 500 ppi (19.7 pixels per mm).

Automatic generation of a local clarity map is based on a variety of algorithms that assess various image attributes related to clarity. The study identified pre-existing tools that could be leveraged and developed some new image analysis tools. The algorithms that were used came from these sources:

- Custom revisions to the NIST MINDTCT (MINutiae DeTeCTor) feature extractor [12] to extract intermediate feature values and representations of data. MINDTCT was derived from the U.K. Home Office HO-39 feature extractor, which was used as the basis for the minutiae detector in the FBI's ULW software (hereafter "Baseline features", described in Appendix A).
- Custom revisions to the FBI's remote fingerprint editing software (RFES), which was developed by Lockheed Martin in 1998 to 2000. Training with features derived from RFES provided classification accuracy similar to the Baseline features but required significantly more computation time. Because LQMetric was designed for operational use, the minimal increase in performance did not justify the increased processing time, and so RFES-derived features were not included in LQMetric.
- Our implementation of Fourier-based features derived from the Latent Fingerprint Image Quality (LFIQ) algorithm [10] (hereafter "FFT features").

For each latent print, over 40 different pixel-wise feature maps were extracted in total using a combination of these algorithms. These features measure a combination of general image attributes along with information related to the frequency and consistency of the local latent fingerprint ridge structure.

[3] American units are used here because FBI image capture specifications require that fingerprints are scanned at 500 or 1000 pixels per inch (19.7 or 39.4 pixels per mm) and the metric equivalents would be rounded.

The training and cross-validation of LQMetric used 1035 latent fingerprint images with local clarity maps marked by (human) latent fingerprint examiners, from the NIST ELFT-EFS#2 [13, 14] evaluation. The training process followed a stratified 10-fold cross-validation strategy. The training data was stratified by randomly selecting samples from the pool of training samples for each clarity level so that each clarity level contained the same number of training samples as clarity level 4 (resulting in 179,757 samples for each clarity level). Clarity level 5 was significantly under-represented within this dataset and was therefore merged with clarity level 4. After stratification, the data was split into 10 folds for cross-validation. Training commenced on nine folds and the last fold was utilized for testing. This process was repeated until every fold served as the test fold. The set of pixel-wise feature maps from each extractor was paired with the pixel-wise ground-truth human-marked clarity information and used as input to an OpenCV CvRTrees random forest machine learning algorithm [15]. This process results in a model that takes the feature maps as input and outputs a predicted clarity map. The predicted clarity maps are post-processed to only include the largest contiguous area of ridge flow (yellow or better) to eliminate isolated cells and to approximate a region of interest as marked by latent print examiners. Figure 2 shows examples of latent prints from the training and validation set, with examiner-marked clarity maps, and output of the CvRTrees model when training with the Baseline and FFT features.

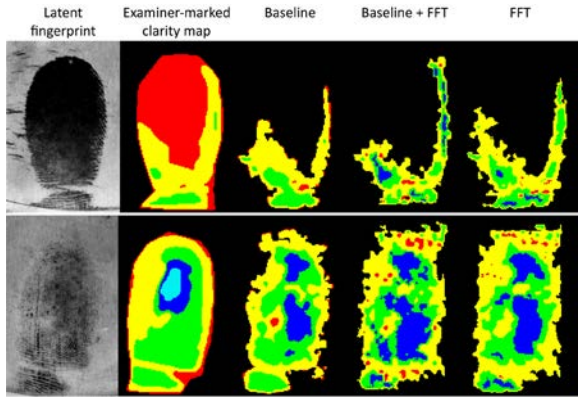

Figure 2

Examples of latent fingerprints and corresponding human-marked clarity maps, as well as the predicted clarity maps utilizing the different feature maps as input. Note that the local clarity map algorithm is designed to crop the clarity maps to the largest contiguous area of ridge flow. LQMetric Overall Latent Quality assessment score: upper image = 40; lower image = 78.

Overall Latent Quality Assessment

LQMetric's automated assessment of Overall Latent Quality uses metrics that are derived from the local clarity maps and from automatically extracted minutiae, trained against matcher scores from searches of those latents against large-scale AFISs.

The latent fingerprint images that were used for training and evaluation for predicting AFIS performance were selected from a subset of the NIST ELFT-EFS Evaluation #2 dataset (EE2, 997 117 latent prints) and the NIST ELFT-EFS Public Challenge dataset (PC2, 245 latent prints) [13, 14]. Latent prints were used only if good-quality exemplars were available and if human latent print examiners determined that identification was possible. Therefore, 69 images from the EE2 dataset were omitted that had only low-quality exemplars (NFIQ=5 for both plain and rolled exemplars) or an examiner-determined status of inconclusive (e.g., no overlap between the latent and exemplars).

The overall quality assessment includes a filter to flag nonfingerprint images, which are assigned an overall quality of zero. To train the nonfingerprint filter, we used LQMetric on more than 8000 nonfingerprint images from a variety of sources. Based on these results, partition trees were used to set thresholds for images with extremely small areas of potential ridge detail or extremely large counts of debatable "minutiae", which are assumed to be nonfingerprint images. Additional details are described in Appendix B.

The matcher scores used in training were from searches of two AFIS Systems. The operational FBI NGI system had a gallery of 69 million subjects (at the time) but could only be searched using the PC2 dataset. The NGI Testbed was a clone of the NGI latent AFIS algorithms, but with a gallery of only 500,000 subjects, searchable using both the EE2 and PC2 datasets. Each subject in the gallery had two associated exemplar sets: one set of 10 rolled fingerprint images and one set of 10 plain fingerprint images. All searches were conducted against all fingers, so that the effective gallery size was 690 million for the operational NGI and 5 million for the NGI Testbed.

Regression models were trained to predict the NGI Testbed score from the local clarity map and other fingerprint attributes. The data was fit using a generalized linear modeling technique via penalized maximum likelihood [5] (lasso from the R package glmnet [16]). The predicted NGI Testbed matcher scores were calibrated using isotonic regression so that an LQMetric score indicates the probability that a latent fingerprint image-only

search (i.e., without manually marked minutiae) would hit at rank-1 on the NGI operational system if the subject is enrolled in the gallery.

LQMetric's Overall Latent Quality is an estimate of the accuracy of searches conducted against NGI; this performance is not necessarily generalizable to other AFIS algorithms or systems. However, experience with ELFT-EFS [13, 14] indicates that AFISs from different vendors are generally consistent regarding which latents result in the highest matcher scores, indicating that latents with very high LQMetric values are likely to be searched successfully across AFIS algorithms or systems, but there may be more disparity among latents with low LQMetric values.

Feature Selection

Selection of the features implemented in the operational LQMetric algorithm was based primarily on accuracy in predicting AFIS matching performance and secondarily on accuracy in generating local clarity maps. Algorithm speed was used to select among features when performance was otherwise equivalent. Cross-validated (10-fold) performance measures for each candidate model were used in selection of the final algorithm [including mean squared error (MSE); root mean squared error (RMSE); mean absolute percentage error (MAPE); Spearman's Rho; Kendall's Tau; percent deviance, R^2; and area under the receiver operating characteristic curve (AUC)]. With respect to accuracy in generating local clarity maps, the CvRTrees models individually trained with the Baseline, RFES, and FFT features all provided roughly similar classification accuracy, and, therefore, feature selection was based on accuracy in predicting NGI matching performance.

With respect to accuracy in predicting NGI matching performance, the cross-validated results described in terms of the R^2 metric are shown in Table 1. For comparison, Table 1 also shows the effectiveness of the clarity maps marked by latent print examiners in predicting AFIS performance, which serves as a rough upper bound on performance. Overall, the CvRTrees model trained with the Baseline features provided the highest average R^2 metric whereas the CvRTrees model trained with both Baseline and FFT features provided the highest R^2. The CvRTrees model trained with the Baseline features was selected for LQMetric because it provided the highest average R-squared (R^2).

	Min R^2	Mean R^2	Max R^2
Examiner-Marked Clarity Maps	0.18	0.43	0.60
Model Features			
Baseline	0.08	0.37	0.53
FFT	0.01	0.30	0.48
Baseline+FFT	0.07	0.36	0.58

Table 1

Cross-validated average R^2 summaries for examiner-marked clarity, clarity generated from Baseline features, FFT features, and their combination.

The final algorithm implemented in LQMetric uses the following features, each of which is calculated based on the local clarity map (and can be returned using LQMetric's "verbose" mode):

- The area (in square millimeters) of clear ridge flow and minutiae (green and better clarity)
- The "good flow" area (in square millimeters) of clear ridge flow and minutiae, after using erosion then dilation morphological operations to eliminate minor gaps or protrusions
- The ratio of clear versus unclear areas of the image (area of green and better clarity over area of yellow clarity)
- The count of very high confidence minutiae (minutiae in blue areas)
- The count of debatable minutiae (minutiae in yellow areas)

For additional details on these features, see Appendix C. For examples of images with local clarity maps and overall quality calculated by LQMetric, see Appendix D.

Evaluations of Performance

The performance of LQMetric can be evaluated in three ways: the classification accuracy of the local clarity maps, the accuracy of predicting AFIS performance, and the accuracy of predicting latent print examiners' assessments of value.

Evaluation of Local Clarity Maps

Table 2 shows the accuracy of classification for the LQMetric-predicted clarity levels versus the actual clarity levels marked by latent print examiners. Overall, the clarity as marked by the human latent print examiners was correctly predicted for 61.4% of samples; the prediction was within ± 1 clarity level for 94.2% of samples. This is slightly better than the reproducibility of clarity markup among human latent print examiners as reported by Ulery et al. [17], in which examiners agreed on clarity levels in 46.3% of locations and agreed within ± 1 clarity level for 89.1% of locations.[4] The extremes (clarity levels 0 and 4) resulted in the most accurate predictions.

		Human-Marked Clarity Level				
		0 (Black)	1 (Red)	2 (Yellow)	3 (Green)	4 (Blue)
LQMetric Predicted Clarity Level	0	71.2%	22.3%	5.5%	0.9%	0.1%
	1	22.4%	51.1%	23.8%	2.3%	0.4%
	2	9.4%	23.8%	47.8%	16.1%	2.9%
	3	1.7%	3.4%	17.0%	60.1%	17.8%
	4	0.2%	0.2%	1.0%	19.7%	78.9%

Table 2

Confusion matrix describing average classification accuracy and error rates for the local clarity maps generated using CvRTrees model training with Baseline features across 10-fold cross-validation. (n=179,757 samples for each clarity level.)

[4] In Ulery et al. [17], Appendix A, Table 11 compares the reproducibility of clarity marked by multiple examiners in 44,941 locations in which at least one of the examiners marked a minutia; because the data is conditioned on the presence of minutiae, the lowest clarity levels are under-represented.

Evaluation of Predictions of AFIS Performance

Figure 3 shows the effectiveness of LQMetric in predicting AFIS hits for both the Testbed and operational NGI searches. For both, LQMetric is monotonic with respect to hit rate: higher LQMetric scores indicate there is a higher likelihood to hit at rank 1. The measured rates are lower for operational NGI than the Testbed, as would be expected given the much larger gallery size. With respect to the Testbed, LQMetric accurately predicts the hit rate for most of its range but is less accurate for LQMetric values less than 20. The operational NGI results are similar, but slightly less accurate.

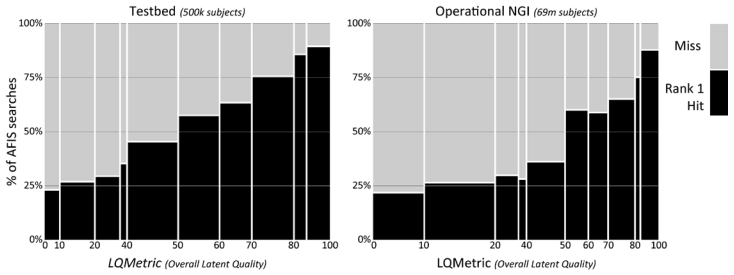

Figure 3

Proportion of image-only searches that resulted in rank-1 hits, limited to searches for which the subject was known to be in the database. Left: NGI Testbed, 1241 latent prints searched against a gallery of 5 million distinct fingers (500k subjects). Right: Operational NGI, 245 latent prints searched against a gallery of 690 million distinct fingers (69m subjects).

Human Examiner Evaluation

We evaluated LQMetric with respect to latent print examiners' assessments of the value of the latent prints, as well as in terms of latent print examiners' assessments of relative differences in quality. Figures 4a and 4b show the association between LQMetric and human examiners' value assessments: 15 latent print examiners were each presented 240 latent prints and asked to rate the value of each latent print (out of a pool of 1152 latents from the ELFT-EFS #2 dataset; 3 to 6 examiners reviewed each latent). The LQMetric software provides an option to return "verbose" details in addition to the default (which in verbose mode is referred to as Overall Latent Quality). Here we also show the predicted VID value, which rescales the response to estimate the probability that a latent print examiner would assess the latent as of value for identification. (LQMetric's verbose mode also returns the various specialized assessments of quality and clarity described by Hicklin et al. [1])

There is a strong association between LQMetric scores and examiners' assessments of value, especially when the LQMetric Overall Latent Quality score is 50 or larger. High LQMetric scores are overwhelmingly assessed as high quality by examiners. However, when LQMetric indicates that a latent is low quality, examiners often consider latents to be of value. Note that because LQMetric is predicting a successful AFIS search and examiners were asked to rate based on utility for individualization, it is not surprising that examiners assessed some images with low LQMetric scores as "of value"; examiners may be able to use a latent print in casework even if it is unlikely to hit in an AFIS search.

Figure 4c shows the association between LQMetric and a different quality scale often used by human examiners, the informal "good, bad, ugly" scale [17]. Again, LQMetric is strongly associated with the human quality assessments, but is not a perfect predictor of the human assessments. Note the differences in "good" in Figure 4a versus 4b, underscoring the variability of informal assessments and different examiners.

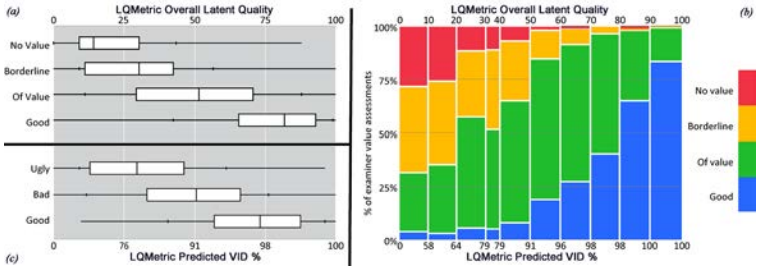

Figure 4

Comparisons of human assessments of value with LQMetric Overall Latent Quality and predicted VID measures. (a) LQMetric vs distributions of examiners' value assessments [(n=3600 examiner assessments of 1152 latents) (medians: 14,30,51,81)]. (b) LQMetric vs consensus of examiners' value assessments [n=1089 latents from (a) for which a majority of examiners agreed on a value assessment]. (c) LQMetric vs informal "Good, Bad, Ugly" assessments [(n=1252 latents, 1 examiner per latent; different examiners from (a) and (b) (medians: 30,51,73)]. Crossbars in boxplots indicate deciles (10%, 90%).

To assess LQMetric in terms of latent print examiners' assessments of relative differences in quality, the 15 examiners were shown each latent print side-by-side with a randomly selected latent (from a different subject), and the examiners were asked to rate whether one of the latent prints was of higher quality or whether they were of equal quality. Figure 5 shows the extent to which the examiners' assessments of higher quality were associated with the higher LQMetric value. Each column contains a 10-point range of LQMetric score differences between the two latents in an image pair, and the colors indicate whether the latents that examiners assessed as higher quality had the higher LQMetric score. For example, for latents that had a difference of LQMetric scores between 30 and 40, 68% of examiners rated the latent with the higher LQMetric score as higher quality, 16% rated the latent with the higher LQMetric score as lower quality, and the remainder rated the latents as equal quality. In general, examiners' assessments of the relative quality of two latent prints correspond to LQMetric, and this association is strongest as the difference in LQMetric scores between the two images increases.

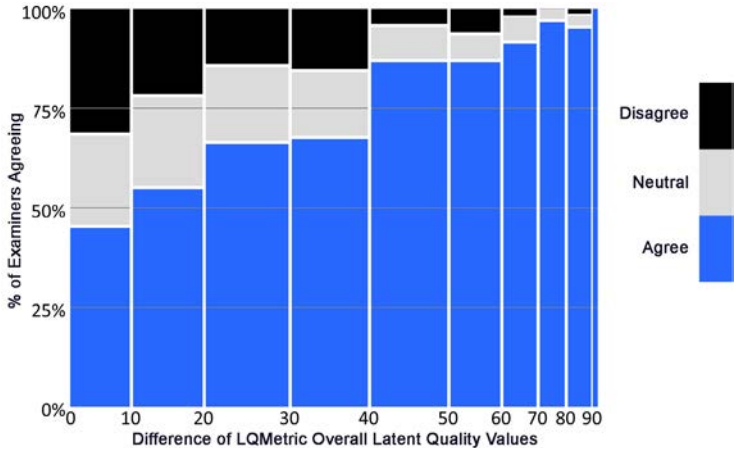

Figure 5

Comparisons of human assessments of relative quality with LQMetric. [n=1800 examiner assessments of the relative quality of pairs of latents (1309 agreements, 234 disagreements)].

Conclusions

The primary purpose of LQMetric is to predict the probability of a successful AFIS search. However, a tool that can reliably and objectively measure the quality of latent prints may be valuable for a variety of purposes, such as providing automated value assessments for use in quality assurance; flagging specific latent prints for additional verification; providing assistance to examiners in determining which latents are appropriate for image-only searching (as opposed to requiring human-marked minutiae); replacing the informal "good", "bad", and "ugly" categories for assessments of the quality of latents in a case or in a dataset; enabling quality-directed workflow by directing poor-quality latents to more expert examiners, or prioritizing backlog based on quality; and in research evaluating the efficacy of new latent print processing or development methods. Clarity maps are used operationally by some latent print examiners in documenting their analyses of the latent prints for archiving or casework exchange. NIST's Latent Interoperability Transmission Specification [18] includes local clarity maps in defining transactions for latent AFIS data exchange and for interchange of latent print annotation among examiners as part of non-AFIS casework.

In this paper, we describe LQMetric, operational software that has been widely used since its release in 2014 and referenced in multiple publications [19–24]. Here we describe the underlying algorithms used in LQMetric and how it was developed and trained. LQMetric is incorporated for operational use in ULW, which is the software used by latent print examiners in federal, state, and local law enforcement agencies to search the FBI's NGI system (by way of their respective state or federal conduits or automated biometric systems). Our results indicate that LQMetric is an effective predictor of AFIS performance as measured by searches of NGI and that there is general agreement between latent print examiners and the overall LQMetric score. Future work may further enhance its functionality, notably with respect to adding palmprint functionality.

Acknowledgments

This is publication number 20-56 of the Laboratory Division of the FBI. The views expressed are those of the authors and do not necessarily reflect the official policy or position of the FBI or the U.S. Government. Development of LQMetric in 2012–2014 was funded by FBI Criminal Justice Information Services

Division and the FBI Biometric Center of Excellence; funding for this publication was provided by the FBI Laboratory Division. Fingerprints shown in illustrations are from pre-existing publicly distributed datasets: latent prints in Figure 2 are from the Noblis Multi-Latent Dataset, which was collected under the approval of the Noblis Institutional Review Board; latent prints in appendix D are from the NIST ELFT-EFS Public Challenge Dataset. We would like to thank our colleagues who aided in developing, evaluating, and sponsoring LQMetric: Matthew Schwarz, Ron Smith, Erik Stanford, Brad Ulery, Bethany Retton, JoAnn Buscaglia, and Maria Antonia Roberts.

For further information, please contact:

R. Austin Hicklin
Noblis
2002 Edmund Halley Dr.
Reston, VA 20191
hicklin@noblis.org

References

1. Hicklin, R. A.; Buscaglia, J.; Roberts, M. A. Assessing the Clarity of Friction Ridge Impressions. *For. Sci. Int*. **2013**, *226* (1–3), 106–117.
2. Hicklin, R. A.; Buscaglia, J.; Roberts, M. A.; Meagher, S. B.; Fellner, W.; Burge, M. J.; Monaco, M.; Vera, D; Pantzer, L. R.; Yeung, C. C.; Unnikumaran, T. N. Latent Fingerprint Quality: A Survey of Examiners. *J. For. Ident*. **2011**, *61* (4), 385–419.
3. Cao, K.; Nguyen, D.-L.; Tymoszek, C.; Jain, A. K. End-to-End Latent Fingerprint Search, Michigan State University Technical Report MSU-CSE-18-3, 2018.
4. Chugh, T.; Cao, K.; Zhou, E.; Tabassi, E.; Jain, A. K. Latent Fingerprint Value Prediction: Crowd-Based Learning, *IEEE Trans. Info. For. and Sec*. **2018**, *13*, (1), 20–34.
5. Friedman, J.; Hastie, T.; Tibshirani, R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J. Stat. Software* **2010**, *33* (1), 1–22.
6. Guan, H.; Dienstfrey, A. M.; Theofanos, M. F. A New Metric for Latent Fingerprint Image Preprocessing. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR) Workshops* **2013**, 84–91.

7. Sankaran, A.; Vatsa, M.; Singh, R. Automated Clarity and Quality Assessment for Latent Fingerprints. *IEEE Sixth International Conference on Biometrics: Theory, Applications, and Systems (BTAS),* Arlington, VA, 2013; pp 1–6; doi:10.1109/BTAS.2013.6712716.
8. Yoon, S.; Cao, K.; Liu, E.; Jain, A. K. LFIQ: Latent Fingerprint Image Quality. *2013 IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, Arlington, VA; 1–8; doi: 10.1109/BTAS.2013.6712750.
9. *NFIQ 2.0 Draft for Comment.* National Institute of Standards and Technology: Gaithersburg, MD. www.nist.gov/system/files/documents/2018/11/29/nfiq2_report.pdf, April 2016.
10. ISO (International Organization for Standardization) *Information Technology - Biometric Sample Quality - Part 1: Framework*; ISO/IEC 29794-1:2016; ISO: Geneva, Switzerland, 2016.
11. Mangold, K. C., Ed. *American National Standard for Information Systems- Data Format for the Interchange of Fingerprint, Facial & Other Biometric Information; ANSI/NIST-ITL 1-2011 NIST Special Publication 500-290,* Edition 3 (ANSI/ NIST-ITL 1-2011:Update 2015). U.S. Department of Commerce, National Institute of Standards and Technology: Gaithersburg, MD, 2015.
12. Ko, K.; Salamon, W. J. NIST Biometric Image Software (NBIS). U.S. Department of Commerce, National Institute of Standards and Technology: Gaithersburg, MD, 2010. Available at http://fingerprint.nist.gov/NBIS.
13. Indovina, M.; Hicklin, R.; Kiebuzinski, G. *NIST ELFT-EFS Evaluation of Latent Fingerprint Technologies: Extended Feature Sets*. NIST Interagency/Internal Report (NISTIR) 7775, 2011. U.S. Department of Commerce, National Institute of Standards and Technology: Gaithersburg, MD.
14. Indovina, M. D.; Dvornychenko, V. N.; Hicklin, R. A.; Kiebuzinski, G. I. *ELFT-EFS Evaluation of Latent Fingerprint Technologies: Extended Feature Sets [Evaluation #2].* NIST Interagency/Internal Report (NISTIR) 7859, 2012. U.S. Department of Commerce, National Institute of Standards and Technology: Gaithersburg, MD.
15. OpenCV (Open Source Computer Vision Library). "OpenCV Random Trees", https://docs.opencv.org/2.4/modules/ml/doc/random_trees.html. [accessed March 2020].
16. Friedman, J.; Hastie, T.; Tibshirani, R.; Narasimhan, B.; Tay, K.; Simon, N.; Qian, J. glmnet: Lasso and Elastic-Net Regularized Generalized Linear Models. https://cran.r-project.org/web/packages/glmnet (accessed 26 March 2020).

17. Ulery, B. T.; Hicklin, R. A.: Roberts, M. A.; Buscaglia, J. Interexaminer Variation of Minutia Markup on Latent Fingerprints. *For. Sci. Int*. **2016**, *264*, 89–99.
18. Chapman, W.; Hicklin, A.; Keibuzinski, G.; Komarinski, P.; Mayer-Splain, J.; Taylor, M.; Wallner, R. *Latent Interoperability Transmission Specification*. NIST Special Publication 1153. U.S. Department of Commerce, National Institute of Standards and Technology: Gaithersburg, MD, 2013.
19. Haraksim, R.; Galbally, J.; Beslay, L. *Study on Fingermark and Palmmark Identification Technologies for their Implementation in the Schengen Information System*. Joint Research Centre (European Commission): Ispra, Italy, 2019.
20. Koertner, A. J.; Swofford, H. J. Comparison of Latent Print Proficiency Tests with Latent Prints Obtained in Routine Casework Using Automated and Objective Quality Metrics. *J. For. Ident*. **2018**, *68* (3), 379–388.
21. Swofford, H. J.; Koertner, A. J.; Zemp, F.; Ausdemore, M.; Liu, A.; Salyards, M. J. A Method for the Statistical Intepretation of Friction Ridge Skin Impression Evidence: Method Development and Validation. *For. Sci. Int*. **2018**, *287*, 113–126.
22. Kelly, S.; Gardner, B.; Murrie, D.; Pan, K.; Kafadar, K. How Do Latent Print Examiners Perceive Proficiency Testing? An Analysis of Examiner Perceptions, Performance, and Print Quality. *Sci. & Just*. **2020**, *60* (2), 120–127.
23. FBI, *Next Generation Identification (NGI) Latent Best Practices*. 2015. https://www.fbibiospecs.cjis.gov/Document/Get?fileName=20151109%20Latent%20Best%20Practices.pdf.
24. FBI, *Next Generation Identification (NGI) Latent Fingerprint Search Strategies*. 2015. https://www.fbibiospecs.cjis.gov/Document/Get?fileName=20151105%20Latent%20Search%20Strategies.docx.

Appendix A

Local Clarity Map Generation

To generate a local clarity map, LQMetric uses 20 low-level clarity feature maps as input to a random forest machine learning algorithm, trained on local clarity maps manually marked by human examiners (1035 latents).

Each feature map is based on sampling points (SPs) with an effective resolution of 125 pixels per inch (ppi, 4.9 pixels per mm). This can be seen as a grid of cells, with each SP representing one 4x4 pixel cell at 500 ppi (pixels per mm).

Each of the following 10 base clarity feature maps is accompanied by an absolute deviation feature map, calculated by taking the absolute value of the difference of the center SP with each of the neighboring SPs (i.e., the 7x7 area of SPs centered on the current SP) and returning the mean. These are used to assess the consistency or noise of the base feature map at that point. If the absolute deviation value is 0, then every sampling point in the 7x7 square in the base feature map has an identical value.

The algorithm is also passed a flag indicating SPs at the extreme edges of images (less than 1/8").

Direction-Based Clarity Feature Maps:

- Curvature: maximum difference in direction between a cell and its eight immediate neighbors, used to flag valid points of inflection as well as ill-defined areas
- Direction change: difference between the direction in a cell and the mean direction of the neighboring cells

Grayscale-Based Clarity Feature Maps:

- Grayscale count: number of distinct grayscale values present in a cell
- Grayscale median: median grayscale value in a cell
- Grayscale range: difference between the 10th and 90th% values in grayscale histogram

Discrete Fourier Transform (DFT)-Based Clarity Feature Maps:

- Low-frequency magnitude: magnitude of low-frequency results from the DFTs used to determine ridge flow direction

- Maximum magnitude: maximum magnitude of DFT direction flow, limited to cells that are within acceptable bounds of normalized magnitude and low-frequency magnitude
- Normalized magnitude: normalized magnitude of DFT frequency

Other Clarity Feature Maps:

- Valid neighbors: number of valid neighboring cells
- Quality: pre-existing metric developed for ULW, combining ridge flow, grayscale, and high curvature data

All of the low-level clarity feature maps are derived from a revised version of the NIST MINDTCT code.

Appendix B

Flagging Nonfingerprint Images

To flag nonfingerprint images, we used LQMetric on 8471 nonfingerprint images from a variety of sources and, based on these results, partition trees were used to set thresholds for images with extremely small areas of potential ridge detail or extremely large counts of debatable "minutiae", which are assumed to be nonfingerprint images. The following thresholds were used:

- The largest contiguous area of possible friction ridge area (i.e., red or better) is less than 30 square millimeters.
- The largest contiguous area of good-quality friction ridge area (green or better) is less than or equal to 0.5 square millimeters.
- There are more than 795 automatically extracted minutia.

Appendix C

Calculation of Overall Quality

In LQMetric, the overall quality value is calculated based on the following measures. Each of the measures is calculated using the local clarity map.

Area3plus: area of green and better clarity in sq mm (i.e., area of clear ridge flow and minutiae)

- GFA3plus: good flow area of green and better clarity in sq mm, after using erosion then dilation morphological operations to eliminate minor gaps or protrusions less than 2 SPs (1.0 mm, approx. 1.8 ridges)
- Area3plusDiv2: ratio of area of green and better clarity over area of yellow clarity (i.e., ratio of clear to unclear area of the image)
- Min2: number of minutiae in yellow areas (i.e., number of debatable minutiae; larger values indicate lower quality)
- Min4: number of minutiae in blue areas (i.e., number of very high confidence minutiae, in areas of clear ridge edges)

Appendix D

LQMetric Examples

Six examples of latent prints with the LQMetric-generated local clarity map and overall quality for each.

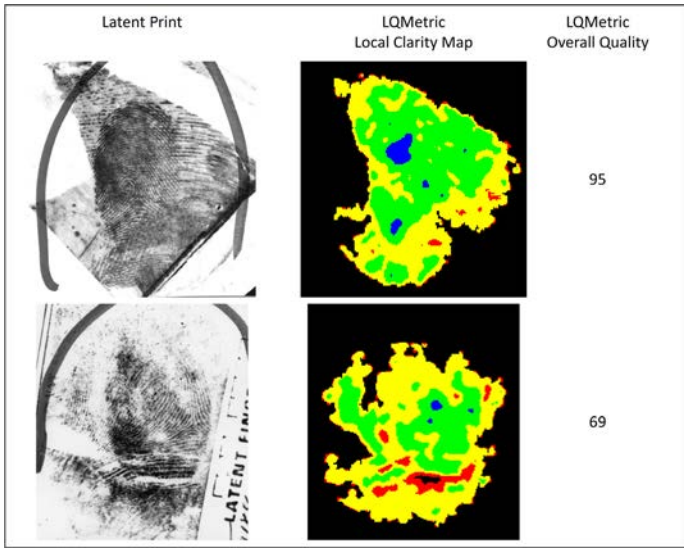

Figure 6

Examples of latent prints with the LQMetric-generated local clarity map and overall quality for each (Images 1–2 of 6).

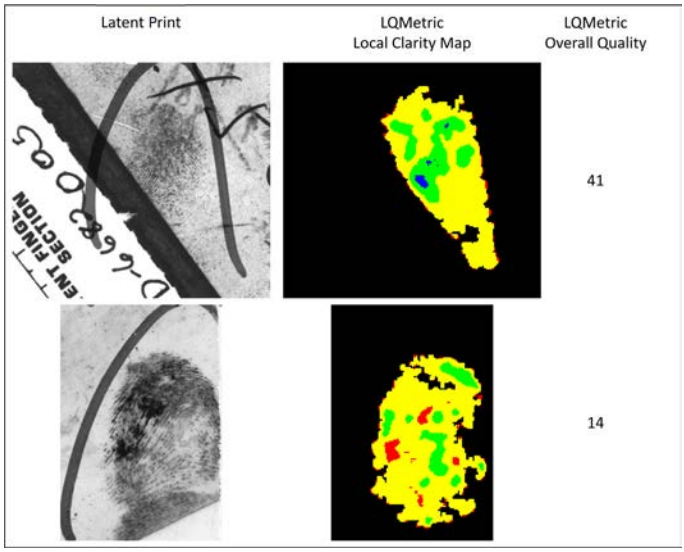

Figure 7

Examples of latent prints with the LQMetric-generated local clarity map and overall quality for each (Images 3–4 of 6).

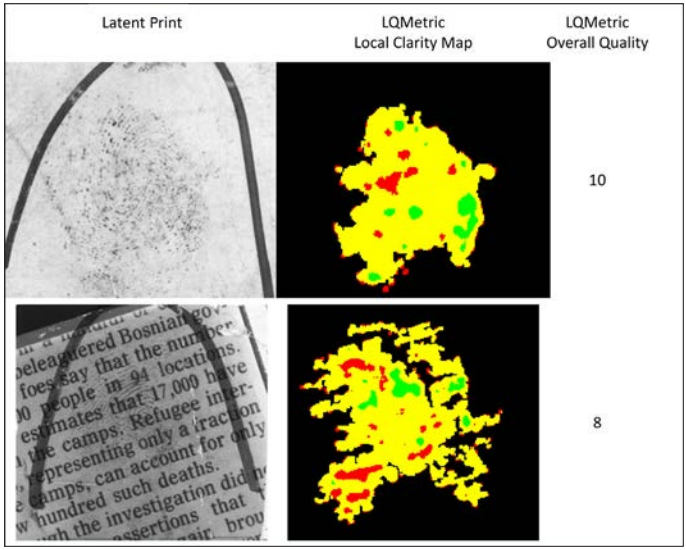

Figure 8

Examples of latent prints with the LQMetric-generated local clarity map and overall quality for each (Images 5–6 of 6).