Multisensor Optical and Latent Fingerprint Database

Anush Sankaran, *Student Member, IEEE,* Mayank Vatsa, *Senior Member, IEEE,* and Richa Singh, *Senior Member, IEEE*

***Abstract*—Large-scale fingerprint recognition involves capturing ridge patterns at different time intervals using various methods, such as live-scan and paper-ink approaches, introducing intra-class variations in the fingerprint. The performance of existing algorithms is significantly affected when fingerprints are captured with diverse acquisition settings such as multi-session, multi-spectral, multi-resolution, with slap and with latent fingerprints. One of the primary challenges in developing a generic and robust fingerprint matching algorithm is the limited availability of large datasets that capture such intra-class diversity. In this paper, we present the Multisensor Optical and Latent Fingerprint (MOLF) database of more than 19000 fingerprint images with different intra-class variations during fingerprint capture. We also showcase the baseline results of various matching experiments on this database. The database is aimed to drive research in building robust algorithms towards solving the problem of latent fingerprint matching and handling intra-class variations in fingerprint capture. Some potential applications for this database are identified and the research challenges that can be addressed using this database are also discussed.**

I. Introduction

AFTER decades of research, fingerprint recognition has become one of the most reliable and commonly used biometric modalities. In 2012, the market for automated fingerprint identification systems and fingerprint technologies contributed the largest share of the global biometrics market and is to continue to be the primary source of overall market revenues [1]. This can be ascertained by the growing number of deployed applications over the last decade using fingerprint biometrics. Some notable large scale applications are:

- The Office of Biometric Identity Management (OBIM), previously called the US-VISIT program [2], provides biometric identification services by collecting fingerprints and other biometric modalities from all the visitors applying for U.S. visas. A fingerprint database of over 90 million identities is currently accessible to federal and state government agencies.
- Aadhaar [3], the brand name of Unique Identification of Authority of India (UIDAI), is one of the largest biometrics projects, providing civil and commercial applications for Indian residents. It uses a combination of fingerprint and iris biometrics for de-duplication and authentication for over 800 million population.
- FBI IAFIS [4] is the U.S. national fingerprint and criminal history system. It houses one of the largest fingerprint databases, recording more than 70 million suspects, along with more than 34 million civil prints.

On the basis of capture type, fingerprints can be classified as (i) inked fingerprints, (ii) live-scan fingerprints, and (iii) latent fingerprints. Using inked methods or using a live-scan device (e.g. optical sensors, capacitive sensors), different fingerprint information can be captured such as flat-dap (single finger flat capture), slap (four finger flat capture), or rolled fingerprints (nail-to-nail information). Extensive research has been undertaken for recognizing fingerprints captured using these methods [5], [6], [7]. Latent fingerprints, on the contrary, are impressions that are deposited when the sweat, amino acids, proteins, and natural secretions present in the skin surface comes in contact with an external surface [8]. These fingerprints are not directly visible to human eyes and after using special (chemical) procedures, the latent prints can be lifted or photographed for further processing. In the same context, simultaneous latent fingerprints are defined as two or more latent fingerprints of the same hand deposited concurrently on the same surface [9]. Research in automated latent fingerprint recognition and simultaneous latent fingerprint recognition is still in development stage.

The evolution of fingerprint authentication has resulted in a broad spectrum of uses including personal authentication, e-commerce, security, and forensic applications. This widespread usage has also led to emergence of different challenges in fingerprint recognition. Some of these challenges are:

- *Interoperability across multiple fingerprint sensors*: Wide range of intra-class variations can occur based on the method or the sensor by which the fingerprint is captured [10]. Fig. 1 shows sample images of the right index fingerprint of a subject captured using different capturing methods, concurrently. It can be observed that these images visually differ with variations in capture process or the acquisition sensor. The report by the National Research Council [11] also discusses this important challenge and suggests the availability of a large database with fingerprint impressions from multiple fingerprint devices can help in improving the performance of algorithms (Recommendation 12).
- *Matching latent prints to slap or rolled fingerprints*: Forensic experts in law enforcement agencies lift latent fingerprints from crime scenes and match them with enrolled databases containing slap or rolled fingerprints. Since the information content and quality of latent fingerprints is significantly different from slap

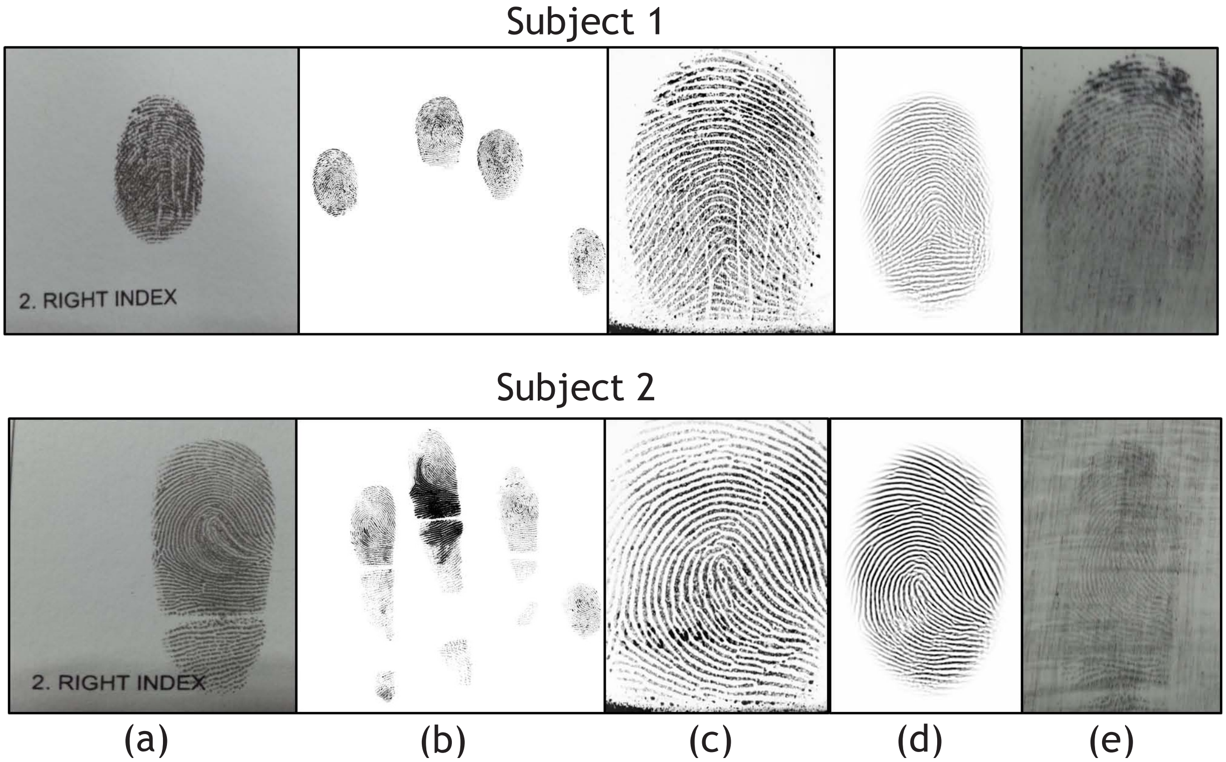


Fig. 1: Right index fingerprint captured from two different subjects showing high inter-class and intra-class variations. Variations are introduced due to different capture methods: (a) inked fingerprint, (b)-(d) live scan fingerprints: (b) CrossMatch sensor, (c) Secugen Hamster-IV sensor, (d) Lumidigm multi-spectral sensor, and (e) latent fingerprint using black powder dusting process. (Figure best viewed under zoom).

and rolled fingerprints, there is significant research required to improve the performance of current systems [8].

- *Matching fingerprint images of different resolutions and spectrums*: Fingerprint capture technology was primarily driven by optical and capacitive sensors. However, with growing usage of fingerprint in e-commerce applications and advent of smart mobile phones, matching fingerprints across different resolutions is also gaining importance. Further, there are fingerprint sensors, such as Lumidigm Venus that utilize information from multiple spectrums for fingerprint capture. Matching such images with the ones obtained from optical or capacitive sensors requires additional research.

Similar to other data driven research areas, advancements in fingerprint recognition, especially in the academic community, are dependent on the availability of large databases. Some of the large publicly available fingerprint databases include card ink-print databases, live scan fingerprint databases, multisensor fingerprint databases, multi-resolution fingerprint databases, latent with corresponding full fingerprint databases, and other special databases. A comparative analysis of all the existing public fingerprint databases is provided in Table I, which also provides a listing of the types of research challenges that can be addressed using each database.

Existing databases primarily have two limitations:

- they generally only contain image variations corresponding to a few challenges, and
- some challenges such as latent fingerprint recognition and cross spectral matching have small databases associated with them.

Some of these challenges are being researched using non-public databases and therefore, it becomes challenging to understand the progression in the state-of-the-art in fingerprints and to reproduce the results. It is our assertion that the availability of a large fingerprint database containing images with variations such as multisensor, multi-spectral, and latent vs. live-scan fingerprint images can significantly instigate research in the academic community and help visualize improvements in the literature. Therefore, we have created a new fingerprint database, termed as Multisensor Optical and Latent Fingerprint (MOLF) database. The MOLF database contains 19,200 multisensor, multi-spectral, dap and slap fingerprint images of 100 subjects obtained from three different sensors along with mated latent and simultaneous latent fingerprints. Moreover, the latent and simultaneous latent fingerprints have manually annotated features. This database provides a scope for development, evaluation, and performance assessment of fingerprint matching algorithms based on single-variate matching as well as cross-variate matching in several applications. The next section presents the details of the database.

TABLE I: Comparison of publicly available fingerprint databases in terms of capture methodology, database size, and the research challenges that can be addressed.

Capture type	Database	Classes	Images	Research Challenges: Multi-session	Multisensor	Multi-spectral	Multi-resolution	Slap fingerprint	Rolled fingerprint	Latent fingerprint	Characteristics
Inked	NIST SD-30 [12]	360	1008	✓		✓	✓		✓		Card fingerprint database of 500ppi and 1000ppi images.
	NIST SD-29 [13]	2160	3024	✓				✓	✓		Card fingerprint database. Scanned at 500 ppi.
	NIST SD-4 [14]	2000	4000	✓					✓		Card fingerprint database. Classified into 5 different L1 class.
	NIST SD-10 [15]	5520	5520						✓		Card fingerprint database of rare fingerprint level 1 patterns.
	NIST SD-14 v2 [16]	27000	54000	✓					✓		Card fingerprint database. WSQ compression.
Live-scan	UCSD WWF [17]	300	300	✓		✓					Wet and Wrinkled fingerprint matching.
	ATVS-FFp [18]	68×2	1632×2	✓	✓						Fake fingerprint matching. Captured using 3 sensors.
	FVC 2000 [19]	110×4	880×4	✓	✓	✓					Low-cost optical, Low-cost capacitive, optical, and synthetic fingerprint subsets.
	FVC 2002 [20]	110×4	880×4	✓	✓	✓					Optical, capacitive, and synthetic fingerprint subsets.
	FVC 2004 [21]	120×4	1440×4	✓	✓	✓					Optical, thermal sweep, and synthetic fingerprint subsets.
	FVC 2006 [22]	150×4	1800×4	✓	✓	✓	✓				Electric field, optical, thermal sweep, and synthetic fingerprint subsets.
	WVU multimodal [23]	272	7219	✓							Captured using CrossMatch, Precise Biometrics, SecuGen sensors at 500 dpi.
	CASIA v5.0 [24]	4000	20000	✓							Captured using URU4000 fingerprint sensor.
	MCYT bimodal [25]	1000	24000	✓	✓						Digital Persona UareU, Precise Biometrics SC-100 sensors.
Camera	NIST SD-24 [26]	100	100 (video)	✓			✓				MPEG-2 Compressed digital video of live-scan fingerprint data. 10 seconds of fingerprints at various rotated angles.
	HKPU low-resolution [27]	306	3080	✓			✓				Fingerprints captured directly using a web camera.
	PolyU HRF [28]	148	3170	✓			✓				Fingerprints captured directly using a high-resolution camera.
Latent	Tsinghua OLF [29]	12	100							✓	Overlapped latent fingerprint segmentation and matching.
	NIST SD-27A [30]	258	258				✓		✓	✓	Latent to 500 ppi and 1000 ppi exemplars matching. Manual annotation of features for latent prints available.
	IIIT-D SLF [9]	180	420	✓				✓		✓	Simultaneous latent fingerprints with 500 ppi slap prints. Manual annotation of features available.
	IIIT-D Latent Fingerprint [31]	150	1241	✓	✓		✓			✓	Latent to latent with 500 ppi slap fingerprints. Latent images directly captured using a high-resolution camera.
	IIIT-D MOLF Database (proposed)	1000	19200	✓	✓	✓		✓		✓	Dap, slap, latent and simultaneous latent fingerprints. Manual annotation of features available.

TABLE II: Different subsets of the MOLF database along with fingerprint type, capture protocol, and its properties.

Subset	Fingerprint type	No. of Images	Image Size	Capture protocol	Comment
$DB1$	Multi-spectral live-scan dap	4000	352 × 544	100 users × 10 fingers × 2 sessions × 2 instances	Lumidigm Venus IP65 Shell
$DB2$	Live-scan dap	4000	258 × 336	100 users × 10 fingers × 2 sessions × 2 instances	Secugen Hamster-IV
$DB3$	Live-scan slap	1200	1600 × 1500	100 users × 3 slap prints × 2 sessions × 2 instances	CrossMatch L-Scan Patrol
$DB3\_A$	Live-scan dap	4000	variable	100 users × 10 fingers × 2 sessions × 2 instances	Cropped prints from DB3
$DB4$	Latent	4400	variable	100 users × 2 hands × 2 sessions × 11 instances	Latent fingerprints, cropped from simultaneous prints
$DB5$	Simultaneous latent	1600	1924 × 1232	100 users × 2 hands × 2 sessions × 4 instances	Simultaneous impression with annotated ROI, core points and minutiae

II. Multisensor Optical and Latent Fingerprint Database

The MOLF database contains large number of fingerprint images with variations in terms of sensor, resolution, and capture spectrum, with slap, latent, and simultaneous latent fingerprint images. Therefore, it provides the opportunity to develop and evaluate algorithms for preprocessing, feature extraction, and matching in different scenarios including latent fingerprint matching. As shown in Table II, the database contains 19,200 fingerprint samples from all 10 fingers of 100 individuals (1000 classes, treating each finger as a class) captured in two independent sessions with an average time difference of 15 days. There are 68 male participants and 32 female participants and the overall age range of the participants is between 18 and 52. The database is captured in an indoor environment under controlled illumination. During each session, each individual provides the following information:

1) two independent instances of all 10 fingerprints captured using Lumidigm Venus sensor,
2) two independent instances of all 10 fingerprints captured using Secugen Hamster-IV sensor,
3) two independent instances of slap fingerprints (4+4+2) captured using CrossMatch L-Scan Patrol sensor, and
4) four independent simultaneous latent impressions ($2+2+3+4$ latent fingerprints) of left and right hand fingers, separately.

A sample fingerprint instance captured from all the sensors is shown in Fig. 2. Depending on the type of problems that can be addressed, the database is partitioned into six subsets: $DB1$ contains the flat dap (all 10) fingerprints collected using Lumidigm Venus sensor and $DB2$ contains the same fingerprints collected using Secugen Hamster-IV sensor. $DB3$ contains the slap fingerprints ($4+4+2$ configuration) collected using CrossMatch L-Scan patrol sensor while $DB3\_A$ contains the dap fingerprints cropped from $DB3$ using *NFSEG* tool [32]. $DB4$ contains the latent fingerprints and $DB5$ contains the simultaneous latent fingerprints. Latent fingerprints are obtained by manually cropping the simultaneous latent fingerprints. Table II provides details about the different subsets of the database.

A. Fingerprint Data Collected with Optical Sensors

The MOLF database has fingerprints taken using three optical sensors: (i) Lumidigm Venus IP65 Shell, (ii) Secugen Hamster-IV, and (iii) CrossMatch L-Scan Patrol. The three sensors comply with FBI's Image Quality Specifications (IQS). The resolution of images captured from Lumidigm, Secugen, and CrossMatch sensors are 500ppi each while the image sizes are 352×544, 258×336, and 1600×1500, pixels respectively.

For 100 individuals, each of the 10 fingerprints is captured in two sessions and in each session, two independent instances are captured. For each sensor, there are 4000 images ($DB1$, $DB2$, $DB3\_A$) with 1000 fingerprint classes. During the first session, the whole process of collection is explained to all the volunteers (subjects) and they are assisted in cleaning their fingers using dry or wet tissues, depending on the requirement. During the second session, the volunteers are allowed to act upon their own and without forced cleaning. The capture is not controlled by the expert and no constraints are applied on the finger's condition. The key motive behind this procedure is to mimic the practical situation of an intentionally registered gallery fingerprint (session I) and an unconstrained probe fingerprint (session II).

B. Latent Fingerprint Collection

The latent and simultaneous latent fingerprints are captured using a black powder dusting process [6]. The usual method of lifting dusted fingerprints using forensic tapes introduces non-linear distortion in the fingerprint ridge information. Therefore, instead of lifting the dusted fingerprints using tapes, a camera setup is created to directly capture the simultaneous latent fingerprint. The camera setup is an improved version of the setup created during the capture of the IIIT-D SLF database [9]. The setup consists of a USB programmable UEye camera with a capture size of 3840×2748 pixels. It has a 0.5-inch CMOS sensor and captures at a maximum rate of three frames per second. A manual C-Mount CCTV lens having a focal length of 8mm is mounted on the camera with finer focus for capturing the latent fingerprint. An illumination ring is attached around the camera to enhance the capture quality. The camera setup is mounted on a flexible Manfrotto magic arm - an elbow arm, clamped to the camera on one end using a Manfrotto super clamp and clamped to a table or to any support (near the dusted fingerprint) on the other end. Fig. 3(a) shows the camera setup used for capturing latent fingerprints.

The volunteers deposit their simultaneous latent fingerprints on a ceramic tile. Though the data collection happens in a closed environment, the participants are completely unconstrained, introducing a large amount of variation and challenges in the deposited latent fingerprint. Two different slabs of the same tile are used to capture the left and right hands of the user during a single session. Four impressions of both hands of the user are captured during each session as follows:

1) thumb and index finger,
2) index and middle finger,
3) index, middle, and ring finger,
4) index, middle, ring, and little finger.

Fingerprints are then directly captured using the camera apparatus. Thus, 16 instances of simultaneous latent fingerprints are captured from each individual in two different sessions. A total of 1600 simultaneous latent impressions are captured constituting $DB5$. The simultaneous latent fingerprints are manually cropped to get the individual latent fingerprints, thus forming $DB4$. As shown in Table II, there are a total of 4400 latent fingerprints from 100 subjects with 1000 classes. $DB4$ contains two latent print instances of every thumb and little finger, four instances of ring finger, six instances of middle finger, and eight instances of index finger.

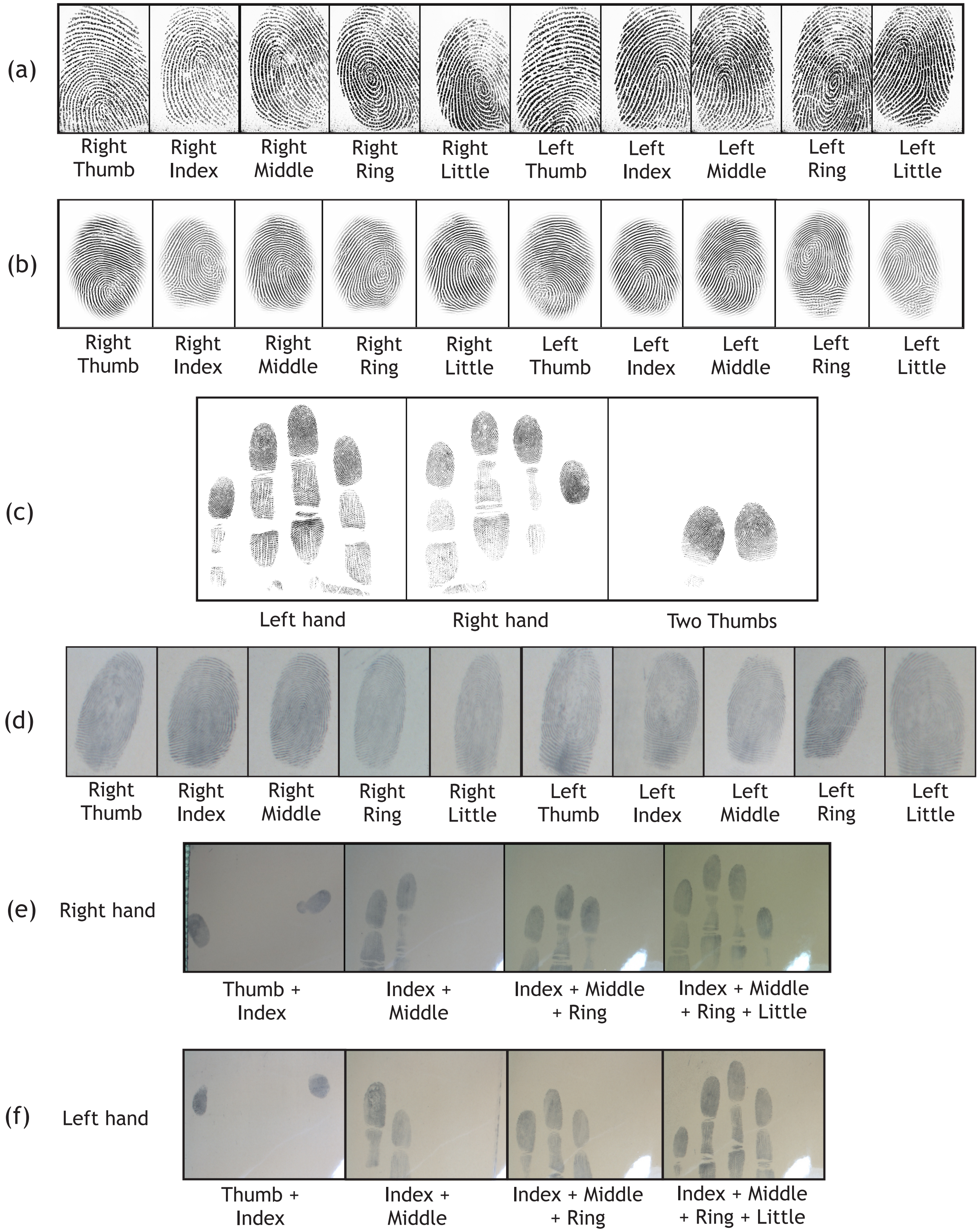
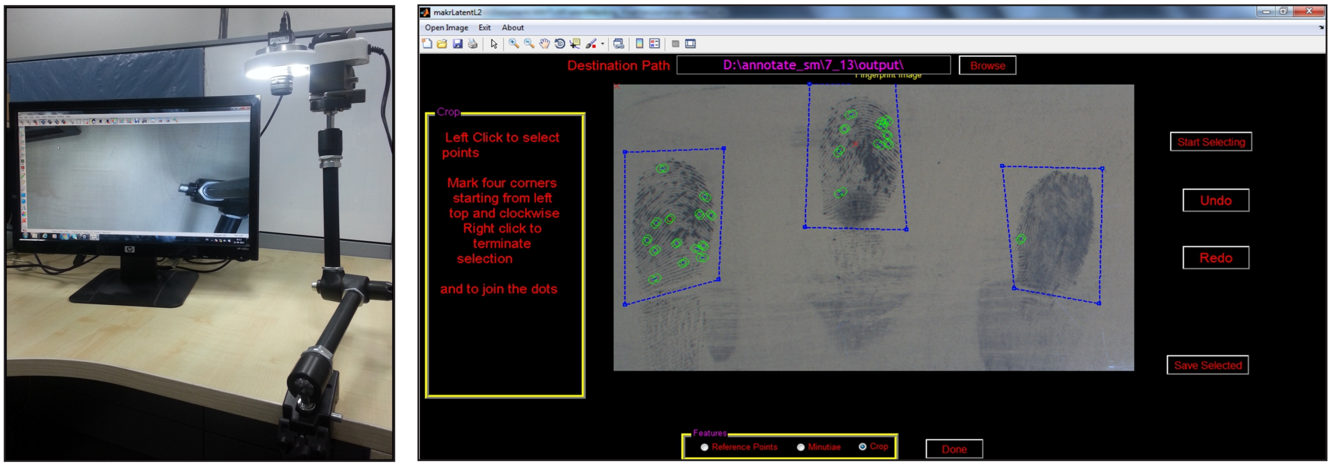


Fig. 2: Sample fingerprints captured of a subject representing capture variations in the MOLF database: (a) 500ppi fingerprint set from Secugen live scan sensor, (b) multi-spectral fingerprint set from Lumidigm live scan sensor, (c) slap fingerprint set from CrossMatch L-Scan Patrol live scan sensor, (d) latent fingerprint set, (e) simultaneous latent fingerprint set of subject's right hand, and (f) simultaneous latent fingerprint set of subject's left hand. The simultaneous impressions are captured with black powder dusting method and are directly captured using a camera setup created. The latent fingerprints are manually cropped from the simultaneous impressions.



(a) (b)

Fig. 3: (a) The latent fingerprint capture setup utilizing a 0.5-inch CMOS sensor with a 8mm focal length CCTV lens mounted on a Manfrotto magic arm that yields an image of size 3840×2748 and (b) a sample screen shot of the GUI based software tool developed for fingerprint feature annotation.

C. Latent Fingerprint Annotation

Automatic feature extraction from latent and simultaneous latent fingerprints is an important research challenge [33]. One of the major goals of FBI's Next Generation Identification (NGI) system is to develop a "lights-out" (fully automatic) fingerprint matching algorithm. To evaluate automated algorithms, a large latent fingerprint database, with manually annotated feature points, is essential. To facilitate the evaluation of such systems, manually marked ground truth feature points are provided for latent and simultaneous latent fingerprints in $DB4$ and $DB5$, respectively. For every simultaneous latent impression from $DB5$, three different features are marked: (i) Region Of Interest (ROI) boundary around every finger impression, (ii) singular points - core and delta (only those found within the available impression) on each finger, and (iii) minutiae on all fingers. Two different annotators[1] independently marked the features, each annotating equal number of images from $DB5$. The annotators marked these features at the rate of $2-3$ subjects per day and completed the annotation task in 22 days. The annotators worked for about 8 hours a day with regular breaks to avoid stress. Using the manually marked ROI, individual fingerprints are cropped from the simultaneous impressions and provided as latent fingerprints in $DB4$. The corresponding features for individual latent fingerprints are also separated and provided along with $DB4$.

To enable simultaneous latent fingerprint annotation and to ease the process, we have also developed a manual annotation tool in Matlab. A screenshot of the tool is shown in Fig. 3(b). The GUI based tool allows the annotator to mark the singular (reference) points, minutiae, and ROI. Along with the database and manually annotated feature points, the tool for manual annotation will also be made available to the research community. As the manually annotated features are provided publicly, their accuracy could be improved by further verification from experts.

[1]The annotators are not certified latent experts. The authors request the researchers in the biometrics and forensics community to improve the annotation and make it publicly available.

D. Availability of Database

All the fingerprints are available in compressed WSQ (Wavelet Scalar Quantization) format and uncompressed BMP format. Table III shows the naming convention of images in different subsets of the MOLF database. *subjectID* defines the subject number (1-100) while *captureID* defines the capture session instance number (1-4) where 1 and 2 belong to the first session, while 3 and 4 belong to the second session. *fingerID* defines the captured finger number (1-10) with 1-5 from right thumb to right little finger and 6-10 from left thumb to left little finger. *handID* defines the slap fingerprint capture ID where 1 denotes the right four fingers, 2 denotes the left four finger, and 3 denotes the two thumbs. *handCode* defines which hand the simultaneous latent is captured from (L, R), and *instanceID* is the particular instance of capture of the impression where 1-4 belongs to first session and 5-8 belongs to second session. The total size of the database in WSQ format is 600 MB and in uncompressed BMP format is 18.2 GB. The database is made available for research purpose at: http://research.iiitd.edu.in/groups/iab/molf.html

TABLE III: The nomenclature followed for the five subsets of the MOLF database.

Database	Image Nomenclature
$DB1$	*subjectID_captureID_fingerID*
$DB2$	*subjectID_captureID_fingerID*
$DB3$	*subjectID_captureID_handID*
$DB3\_A$	*subjectID_captureID_fingerID*
$DB4$	*subjectID_handCode_instanceID_fingerID*
$DB5$	*subjectID_handID_instanceID*

III. Research applications of the database

The MOLF database provides an opportunity to study multiple challenging problems related to fingerprint recognition. Major applications and new research challenges that can be addressed using the database are discussed as follows:

- *Intersensor fingerprint matching:* $DB1$, $DB2$, and $DB3\_A$ contain images captured from three different live-scan fingerprint sensors. By having one of the subsets as a gallery and any other as a probe, the performance of a fingerprint matcher can be evaluated for sensor interoperability. This also represents a practical scenario where the gallery and probe images are not captured using the same sensor.
- *Latent fingerprint feature extraction and matching*: Forensic applications require matching latent fingerprint with live-scan fingerprints [33]. Extracting reliable features from latent fingerprints is a challenging task [34]. Given the ground truth minutiae annotations, the performance of a minutiae extraction algorithm can be evaluated with good confidence. Also, with an exemplar gallery set (any one of $DB1$, $DB2$, or $DB3\_A$) and latent probe set ($DB4$), the performance of a latent fingerprint matching system can be analyzed.
- *Latent to latent fingerprint matching*: The $DB4$ subset can be used for evaluating the performance of a latent to latent fingerprint matcher for crime scene linking applications [31]. Since the latent prints in $DB4$ consist of multiple instances of the same finger, both gallery and probe can be formed using latent prints in $DB4$.
- *Simultaneous latent fingerprint matching*: The $DB5$ subset can be used for matching simultaneous latent fingerprints [9], [35]. Simultaneous latent fingerprints in $DB5$ can be matched with live-scan dap fingerprints in $DB1$, $DB2$, or $DB3\_A$, and slap fingerprints in $DB3$ to evaluate the performance of the matcher.
- *Simultaneous latent fingerprint segmentation:* As the manual segmentation results for simultaneous latent fingerprints in $DB5$ are provided, the ground truth can be used to assess the proficiency of automatic segmentation algorithms.

IV. Experimental Evaluation for Baseline Results

To establish the baseline performance on the MOLF database, several experiments are performed. These experiments are designed to demonstrate the challenges associated with the proposed database and to highlight its usage. The baseline results for livescan fingerprint experiments are computed using two fingerprint matching algorithms: NBIS (NIST Biometric Imaging Software) [32] and VeriFinger [36]. NBIS is an open source minutiae based matching algorithm developed by NIST whereas VeriFinger is a low cost proprietary software by Neurotechnology.

Latent fingerprint matching is an open research problem that the community is attempting to address. It is important to note that there is no standard latent fingerprint matching Software Development Kit (SDK) or commercial system available in the public domain, using which baseline performance can be established. In literature, we have observed that local Minutiae Cylinder Code (MCC) [37], [38] description for manually marked minutiae provides state-of-the-art results [39]. Therefore, MCC descriptors are utilized for establishing baseline results on the latent fingerprint dataset.

First, a NFIQ-based [40] analysis is performed to understand the quality distributions of different subsets of the databases. Thereafter, three different sets of experiments are performed to establish the baseline in different application scenarios. The first experiment (*Experiment I*) evaluates the performance of optical scanner fingerprints while the other two experiments (*Experiment II* and *Experiment III*) pertain to latent fingerprint matching. For *Experiment I*, both identification and verification experiments are performed and the results are reported using the Cumulative Match Characteristics (CMC) curve and the Receiver Operating Characteristics (ROC) curve, respectively. For *Experiment I* and *Experiment II*, identification experiments are performed and the results are reported in terms of the CMC curve.

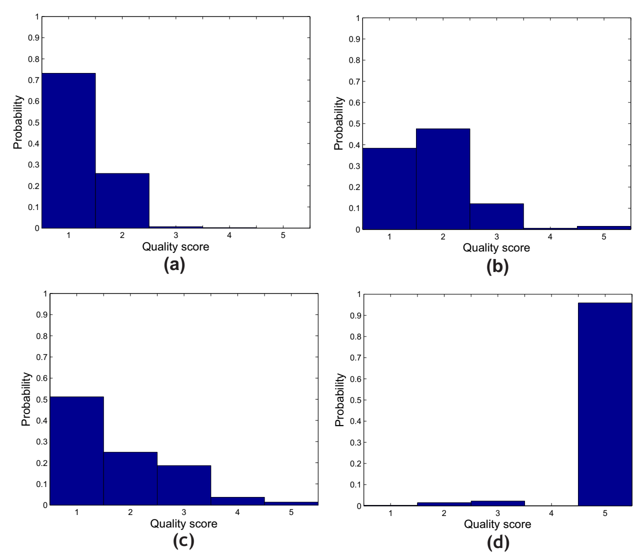


Fig. 4: NFIQ quality score distribution of (a) $DB1$ (Lumidigm) images, (b) $DB2$ (Secugen) images, (c) $DB3\_A$ (CrossMatch) images, and (d) $DB4$ (latent) images. In NFIQ measure, 1 denotes the best quality score while 5 denotes the worst.

A. Quality Analysis

Quality of all the fingerprints captured is analyzed using NFIQ (NBIS Fingerprint Image Quality) [40]. It is an open source minutiae-based quality extraction algorithm that provides a quality value $\{1, 2, 3, 4, 5\}$, with 1 representing the best quality and 5 the worst. NFIQ quality distribution

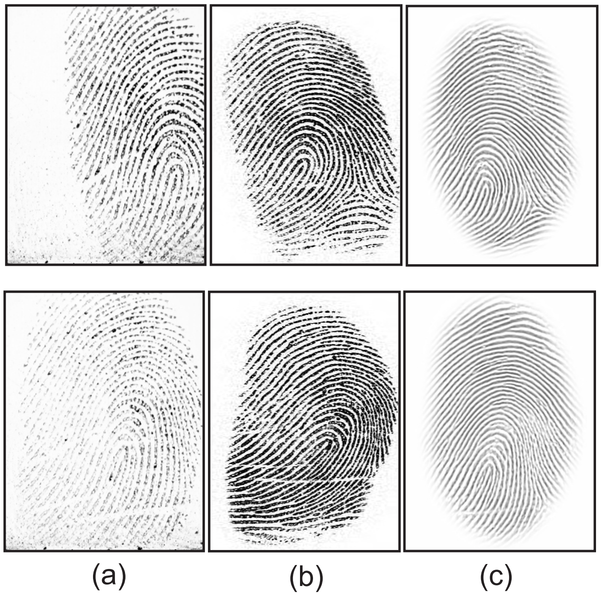


Fig. 5: Sample images showing quality variations across the three sensors (a) Secugen Hamster-IV, (b) CrossMatch L-Scan Patrol, and (c) Lumidigm Venus. It can be observed that some of the images captured using Secugen Hamster-IV have poor capture quality because of its unconstrained capture mode.

of $DB1$, $DB2$, $DB3\_A$, and $DB4$ are shown in Fig. 4. In live-scan fingerprints, it can be observed that the images from $DB1$ (Lumidigm) have the best quality images highlighting the robustness of multi-spectral images. Lumidigm Venus sensor captures the fingerprint in multiple spectrums and while fusing them, enhances the image quality. Also, CrossMatch L-Scan Patrol has an in-built quality control mechanism and captures only those fingerprints that pass the quality threshold. However, no such quality constraint is imposed on Secugen Hamster-IV scanner, thus some of the fingerprints in $DB2$ have relatively lower quality scores, as shown in Fig. 5. As expected, latent fingerprints in $DB4$ are poor quality fingerprints with almost 96% of them having a quality score of 5. However, NFIQ is not designed to evaluate the quality of latent fingerprints and a standard (open source) latent fingerprint specific assessment algorithm is still a research challenge [41]. Similarly, there is no exclusive quality measure for simultaneous latent fingerprints ($DB5$) as well. Therefore, this is a high impact research challenge which could be addressed using this database.

B. Sensor Interoperability Analysis

This experiment (termed as *Experiment I*) is performed to establish the baseline accuracy with fingerprints captured in different sessions using multiple sensors. In all three subsets, the first two instances captured during the first session are taken as gallery and the fingerprints captured during the second session are used as probe. Thus, the gallery and probe both contain 2000 images pertaining to 1000 (100×10) classes. Datasets $DB1$, $DB2$, and $DB3\_A$ are used. Since $DB3$ contains slap fingerprints, it is not used for this experiment. NBIS [32] and VeriFinger SDK [36] are then used for feature extraction and matching. Both identification and verification experiments are performed and the results are reported in Table IV. The corresponding CMC curves are shown Fig. 6, Fig. 7 and the ROC curves in Fig. 8, Fig. 9. The major observations made are as follows:

- In Experiment I, VeriFinger is observed to yield higher accuracies compared to NBIS on all three subsets of the database. VeriFinger provides same-sensor rank-1 matching accuracy in the range of 96%-98% whereas NBIS is at least 7% lower in performance.
- From Experiment I, it can be observed that matching performance is high when the gallery and probe fingerprints are captured using the same sensor. However, when the gallery and probe fingerprints are captured using different sensors, performance is reduced significantly for both NBIS and VeriFinger. This highlights that cross-sensor fingerprint matching, especially when one sensor is a multi-spectral sensor, is a research challenge.
- Verification experiments performed using NBIS show the clear impact of cross-sensor matching, having about 40% more errors than same-sensor matching. However, VeriFinger reduces the effect of cross-sensor matching to great extent showing a difference of only about 3%. Nonetheless, in large scale applications such as India's Aadhaar project, 3% is a significant error and might have a greater impact.

C. Latent Fingerprint Matching

This experiment is performed to establish the baseline accuracy of latent fingerprint matching. There are two different experiments performed on latent fingerprint matching: (i) latent fingerprint matching using manually annotated minutiae (termed as *Experiment II*), and (ii) latent fingerprint matching using automatically extracted minutiae (termed as *Experiment III*). In Experiment II, 4400 latent images in $DB4$ are used as probes and matched against three different galleries of $DB1$, $DB2$, and $DB3\_A$. The results are computed with two different approaches (a) MCC descriptor and (b) Bozorth3 (an open source matcher) available as a part of NBIS. The results are reported in Table V and the CMC curves are shown in Fig. 10. In Experiment III, both MINDTCT (NBIS) and VeriFinger are used for feature extraction and matching. Latent fingerprints in $DB4$ are matched with live-scan fingerprints in $DB1$, $DB2$, and $DB3\_A$, individually. The gallery-probe splits used are the same as in Experiment II. Two sets of experiments are performed: (a) using all probe images in $DB4$, and (b) after removing the *Failed To Process (FTP)* latent fingerprints from $DB4$. The results of all latent fingerprint matching using automatically extracted minutiae are reported in Experiment IIIa. During automatic minutiae extraction in Experiment IIIa experiments, the minutiae extractor (MINDTCT or VeriFinger) failed to extract even

TABLE IV: Rank-1 identification accuracy and equal error rate (for verification) pertaining to experiment I (sensor interoperability analysis). Equal Error Rate (EER) is the value where false accept rate and false reject rate are equal.

Experiment	Algorithm	Gallery Images	Probe Images	Accuracy (%)	EER (%)
I: Live-scan fingerprints	NBIS	$DB1$ (Lumidigm)	$DB1$ (Lumidigm)	84.90	8.57
			$DB2$ (Secugen)	42.50	10.11
			$DB3\_A$ (CrossMatch)	43.50	49.67
		$DB2$ (Secugen)	$DB1$ (Lumidigm)	44.75	10.05
			$DB2$ (Secugen)	91.70	7.85
			$DB3\_A$ (CrossMatch)	44.70	49.77
		$DB3\_A$ (CrossMatch)	$DB1$ (Lumidigm)	42.45	46.74
			$DB2$ (Secugen)	43.95	46.67
			$DB3\_A$ (CrossMatch)	84.90	08.88
	Verifinger	$DB1$ (Lumidigm)	$DB1$ (Lumidigm)	96.75	3.16
			$DB2$ (Secugen)	47.40	6.46
			$DB3\_A$ (CrossMatch)	46.90	6.42
		$DB2$ (Secugen)	$DB1$ (Lumidigm)	47.35	6.47
			$DB2$ (Secugen)	98.10	3.20
			$DB3\_A$ (CrossMatch)	46.20	3.94
		$DB3\_A$ (CrossMatch)	$DB1$ (Lumidigm)	47.80	6.42
			$DB2$ (Secugen)	43.25	3.94
			$DB3\_A$ (CrossMatch)	97.05	3.51

TABLE V: Rank-50 identification accuracy of experiment II (latent matching with manually marked minutiae) and experiment III (latent matching with automatically extracted minutiae).

No.	Experiment	Algorithm	Gallery Images	Probe Images	Accuracy (%)
IIa	Latent fingerprints (manually annotated minutiae)	MCC	$DB1$ (Lumidigm)	$DB4$ (Latent)	7.84
			$DB2$ (Secugen)		7.28
			$DB3\_A$ (CrossMatch)		5.88
IIb	Latent fingerprints (manually annotated minutiae)	$Bozorth3$	$DB1$ (Lumidigm)	$DB4$ (Latent)	31.86
			$DB2$ (Secugen)		31.49
			$DB3\_A$ (CrossMatch)		33.38
IIIa	Latent fingerprints (automatically extracted minutiae - without FTP)	NBIS	$DB1$ (Lumidigm)	$DB4$ (Latent)	6.06
			$DB2$ (Secugen)		9.09
			$DB3\_A$ (CrossMatch)		10.60
		VeriFinger	$DB1$ (Lumidigm)	$DB4$ (Latent)	6.80
			$DB2$ (Secugen)		6.37
			$DB3\_A$ (CrossMatch)		6.51
IIIb	Latent fingerprints (automatically extracted minutiae - with FTP)	NBIS	$DB1$ (Lumidigm)	$DB4$ (Latent)	53.03
			$DB2$ (Secugen)		42.42
			$DB3\_A$ (CrossMatch)		46.97
		VeriFinger	$DB1$ (Lumidigm)	$DB4$ (Latent)	55.60
			$DB2$ (Secugen)		49.27
			$DB3\_A$ (CrossMatch)		56.09

one minutia from several latent probes. In Experiment IIIb, these images are excluded from the probe set and considered as *Failed To Process* error [7]. The identification results are reported in Fig. 10 and Table V. The following key observations are made:

- Experiment IIa exhibits that state-of-the-art MCC descriptor provides very low rank-50 identification accuracy of about 5%-7%, showcasing the challenging nature of latent fingerprints in this database.
- Experiment IIb shows that with manually annotated minutiae, rank-50 matching accuracy of latent fingerprints is in the range of 31%-34%. This indicates that even after manual annotation of minutiae, latent fingerprint matching has a scope for designing robust algorithms for minutiae matching in partial fingerprints.
- For Experiment III with $DB4$ subset, MINDTCT (NBIS) extracts an average of four minutiae per latent fingerprint, while VeriFinger extracted almost 42 minutiae per latent fingerprint. On the other hand, an average of 11 minutiae per latent fingerprint are marked during manual annotation. This indicates that MINDTCT produces too few minutiae while VeriFinger extracts too many spurious minutiae for latent fingerprints.
- Experiment IIIa shows the results of matching latent and live-scan prints using an automated feature extractor and matcher. The results obtained are in the range of 6%-11%, which shows that automated feature extraction requires a significant amount of research. Similar to Experiment II, the best matching performance is obtained for NBIS matcher while using $DB3\_A$ (CrossMatch) as gallery.
- After removing the FTP latent fingerprints from $DB4$, the performance improves and the accuracy of Experiment IIIb is found to be in the range of 42%-56%. It is interesting to note that NBIS shows a very high FTP rate of almost 78% while the FTP rate for VeriFinger is approximately 17%. However, we would like to emphasize that VeriFinger and NBIS are not meant for matching latent fingerprints.

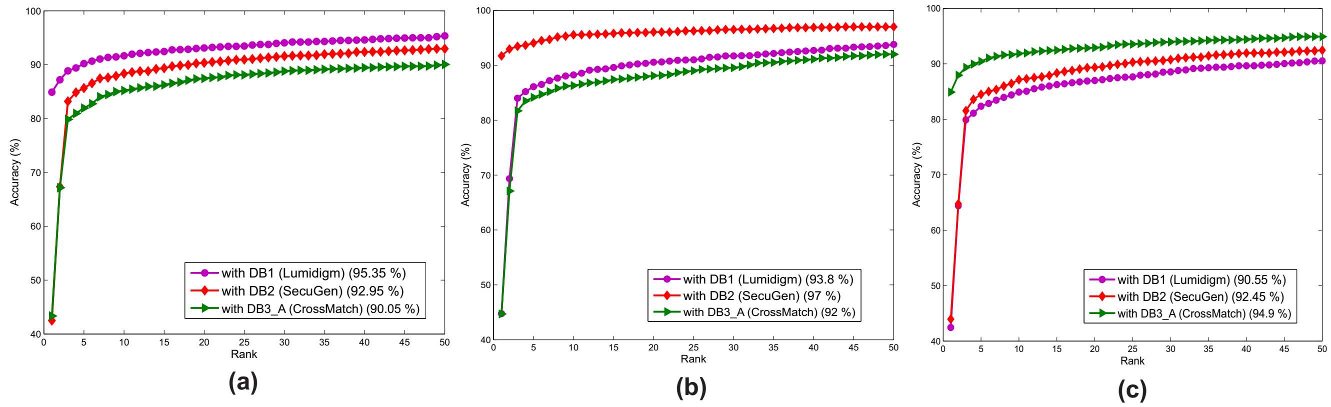


Fig. 6: CMC curves using NBIS for experiment I. (a) $DB1$ (Lumidigm) as gallery, (b) $DB2$ (Secugen) as gallery, and (c) $DB3\_A$ (CrossMatch) as gallery. For all three cases, probe is also varied to study the effect of interoperability.

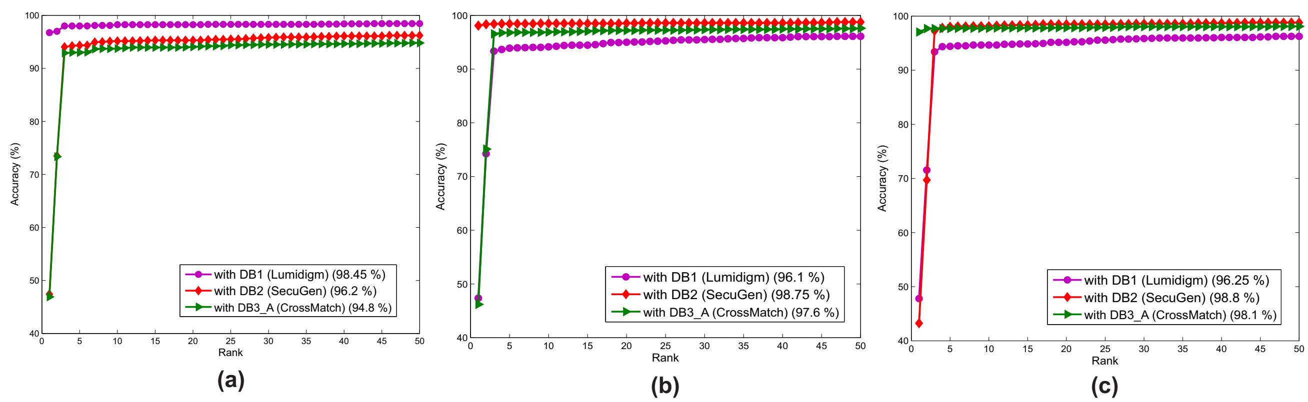


Fig. 7: CMC curves using VeriFinger for experiment I. (a) $DB1$ (Lumidigm) as gallery, (b) $DB2$ (Secugen) as gallery, and (c) $DB3\_A$ (CrossMatch) as gallery. For all three cases, probe is also varied to study the effect of interoperability

- NFSEG in NBIS is used to crop slap fingerprints captured using CrossMatch sensor. A segmentation accuracy of 98.4% is obtained for segmenting 1200 slap fingerprints into 4000 individual fingerprints, failing to segment 64 fingerprint images. These images are further manually cropped for our experiments. However, NFSEG fails to perform segmentation in simultaneous fingerprints, segmenting only 134 latent fingerprints from a total of 4400 prints (with $\sim 3\%$ accuracy).

Since there is no automatic algorithm for establishing simultaneity or automatic simultaneous latent fingerprint matching, baseline results are not computed for $DB5$.

V. Conclusion

Fingerprint matching with live-scan fingerprints is a well studied research problem. However, the academic research in latent fingerprints is in nascent stages. The primary reason for this is the lack of a large publicly available database. In this research work, we have developed a new fingerprint database, the Multisensor Optical and Latent Fingerprint (MOLF) database, that addresses this limitation. This database also acts as a very important resource to address diverse challenges in fingerprint recognition including interoperability between optical and multi-spectral sensors, latent to slap fingerprint matching, latent to latent fingerprint matching, and simultaneous latent fingerprint matching. It is our assertion that the availability of such a database will promote further research in the community and improve the state-of-the-art in these challenging and important research problems.

VI. Acknowledgement

The authors would like to thank Brian Powell and Ishan Nigam for their feedback on this paper. Mayank Vatsa is partly supported through the DST FAST grant from the Department of Science and Technology, India. Anush Sankaran is partly supported by the TCS PhD research fellowship.

References

[1] "Biometrics Research Group Inc.." http://www.biometricupdate.com.
[2] "United States Visitor and Immigrant Status Indicator Technology program (US-VISIT)." http://www.dhs.gov/obim.
[3] "Aadhaar, Unique Identification Authority of India (UIDAI)." Planning Commission, Government of India, http://uidai.gov.in/.
[4] "Integrated Automated Fingerprint Identification System." http://www.fbi.gov/about-us/cjis/fingerprints_biometrics/iafis/iafis.

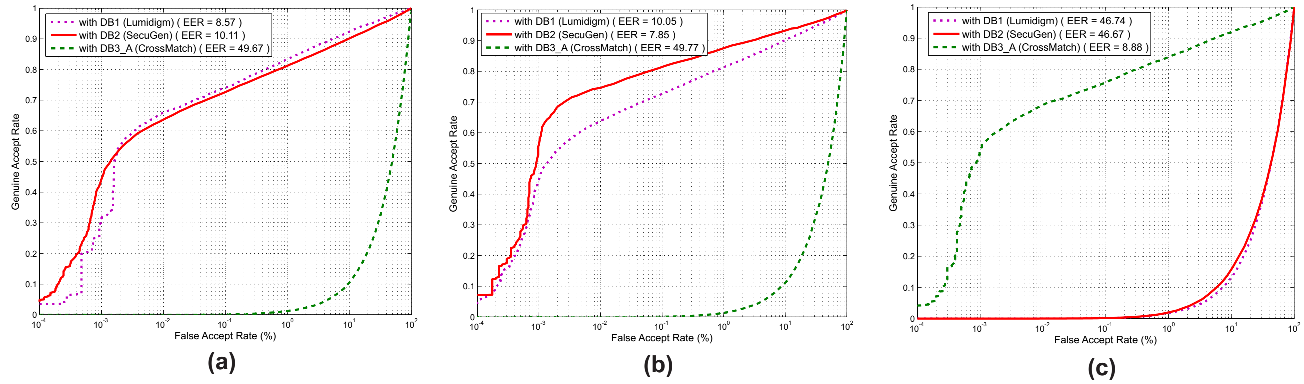


Fig. 8: ROC curves using NBIS for experiment I. (a) $DB1$ (Lumidigm) as gallery, (b) $DB2$ (Secugen) as gallery, and (c) $DB3\_A$ (CrossMatch) as gallery. For all three cases, probe is also varied to study the effect of interoperability.

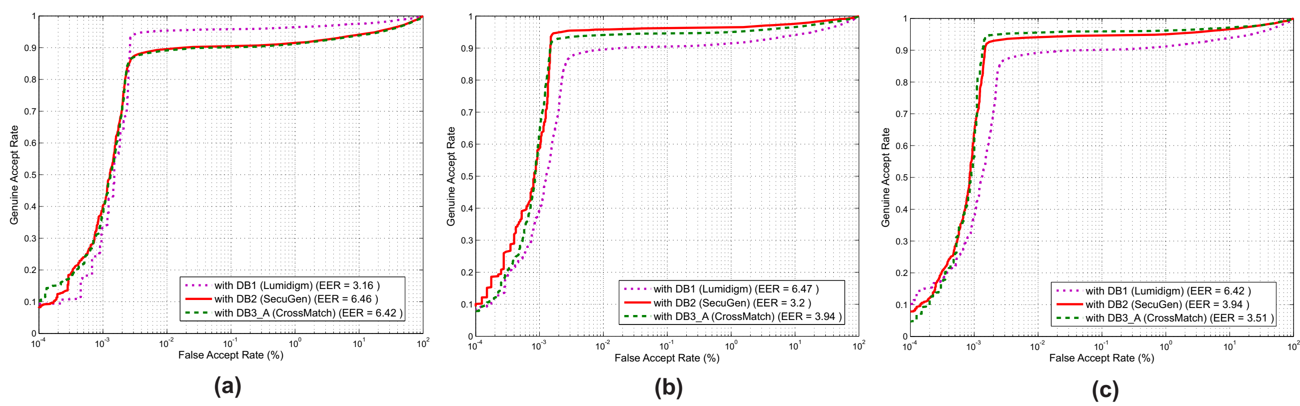


Fig. 9: ROC curves using VeriFinger for experiment I. (a) $DB1$ (Lumidigm) as gallery, (b) $DB2$ (Secugen) as gallery, and (c) $DB3\_A$ (CrossMatch) as gallery. For all three cases, probe is also varied to study the effect of interoperability

[5] P. Komarinski and Ed., "Automated Fingerprint Identification Systems (AFIS)," *Elsevier Academic Press*, 2001.

[6] H. C. Lee, R. Ramotowski, and R. Gaensslen, *Advances in fingerprint technology*. CRC press, 2001.

[7] D. Maltoni, D. Maio, A. K. Jain, and S. Prabhakar, *Handbook of fingerprint recognition, 2nd edition*. Springer-Verlag, 2009.

[8] A. Sankaran, M. Vatsa, and R. Singh, "Latent fingerprint matching: A survey," *IEEE Access*, vol. 2, pp. 982–1004, 2014.

[9] A. Sankaran, M. Vatsa, and R. Singh, "Hierarchical fusion for matching simultaneous latent fingerprint," *International Conference on Biometrics: Theory, Applications and Systems*, pp. 377–382, 2012.

[10] F. Alonso-Fernandez, R. Veldhuis, A. Bazen, J. Fierrez-Aguilar, and J. Ortega-Garcia, "Sensor interoperability and fusion in fingerprint verification: A case study using minutiae-and ridge-based matchers," in *International Conference on Control, Automation, Robotics and Vision*, pp. 1–6, 2006.

[11] "Strengthening forensic science in the United States: a path forward, National Research Council," 2009.

[12] "Dual resolution images from paired fingerprint cards - NIST special database 30." http://www.nist.gov/srd/nistsd30.cfm, 2010.

[13] "Plain and rolled images from paired fingerprint cards - NIST special database 29." http://www.nist.gov/srd/nistsd29.htm, 2010.

[14] "Fingerprint Minutiae from Latent and Matching Tenprint Images - NIST special database 4." http://www.nist.gov/srd/nistsd4.htm, 2010.

[15] "NIST Supplemental Fingerprint Card Data (SFCD - NIST special database 10." http://www.nist.gov/srd/nistsd10.cfm, 2010.

[16] "NIST 8-bit gray scale images of Fingerprint Image Groups (FIGS) - NIST special database 14 version 2," 2010.

[17] P. Krishnasamy, S. Belongie, and D. Kriegman, "Wet fingerprint recognition: Challenges and opportunities," in *International Joint Conference on Biometrics*, pp. 1–7, 2011.

[18] J. Galbally, J. Fierrez, F. Alonso-Fernandez, and M. Martinez-Diaz, "Evaluation of direct attacks to fingerprint verification systems," *Telecommunication Systems*, vol. 47, no. 3, pp. 243–254, 2011.

[19] D. Maio, D. Maltoni, R. Cappelli, J. Wayman, and A. Jain, "FVC2000: Fingerprint Verification Competition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 3, pp. 402 –412, 2002.

[20] D. Maio, D. Maltoni, R. Cappelli, J. L. Wayman, and A. K. Jain, "FVC2002: second Fingerprint Verification Competition," in *International Conference on Pattern Recognition*, pp. 811–814, 2002.

[21] D. Maio, D. Maltoni, R. Cappelli, J. Wayman, and A. Jain, "FVC2004: third Fingerprint Verification Competition," *Biometric Authentication*, pp. 1–5, 2004.

[22] R. Cappelli, M. Ferrara, A. Franco, and D. Maltoni, "Fingerprint Verification Competition 2006," *Biometric Technology Today*, vol. 15, no. 7, pp. 7–9, 2007.

[23] L. Hornak, A. Ross, S. G. Crihalmeanu, and S. A. Schuckers, "A protocol for multibiometric data acquisition storage and dissemination," tech. rep., West Virginia University, https://eidr. wvu. edu/esra/documentdata. eSRA, 2007.

[24] "CASIA fingerprint image database version 5.0." http://biometrics.idealtest.org/.

[25] J. Ortega-Garcia, J. Fierrez-Aguilar, D. Simon, J. Gonzalez, M. Faundez-Zanuy, V. Espinosa, A. Satue, I. Hernaez, J.-J. Igarza, C. Vivaracho, *et al.*, "MCYT multimodal baseline corpus: a bimodal biometric database," *IEE Proceedings on Vision, Image and Signal Processing*, vol. 150, no. 6, pp. 395–401, 2003.

[26] "NIST Digital Video of Live-scan Fingerprint Database - NIST special database 24." http://www.nist.gov/srd/nistsd24.htm, 2010.

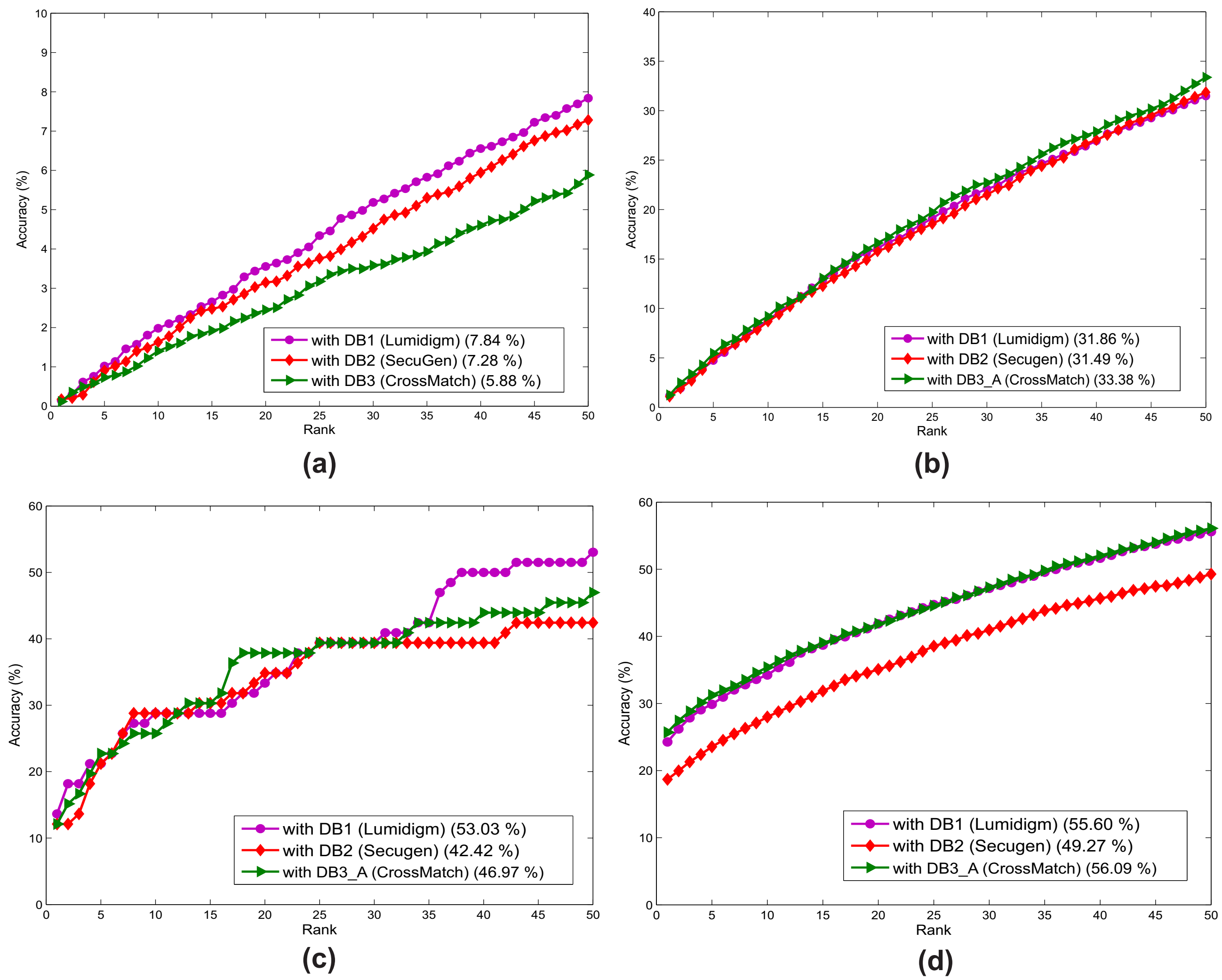


Fig. 10: CMC curves for experiments IIa, IIb and IIIb. The results are computed with (a) manually marked minutiae matched using Minutiae Cylinder Code, (b) manually marked features matched using BOZORTH3, (c) NBIS, and (d) VeriFinger.

[27] A. Kumar and Y. Zhou, "Contactless fingerprint identification using level zero features," in *Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 114–119, 2011.

[28] "The Hong Kong Polytechnic University (PolyU) High-Resolution-Fingerprint (HRF) database." http://www.comp.polyu.edu.hk/ biometrics/HRF/HRF.htm.

[29] J. Feng, Y. Shi, and J. Zhou, "Robust and efficient algorithms for separating latent overlapped fingerprints," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 5, pp. 1498 –1510, 2012.

[30] "Fingerprint Minutiae from Latent and Matching Tenprint Images - NIST special database 27 (a)." http://www.nist.gov/srd/nistsd27.htm, 2010.

[31] A. Sankaran, T. I. Dhamecha, M. Vatsa, and R. Singh, "On matching latent to latent fingerprints," *International Joint Conference on Biometrics*, pp. 1–6, 2010.

[32] "NBIS (NIST Biometric Image Software)." http://www.nist.gov/itl/iad/ig/nbis.cfm.

[33] A. Jain and J. Feng, "Latent fingerprint matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 1, pp. 88–100, 2011.

[34] A. Sankaran, P. Pandey, M. Vatsa, and R. Singh, "On latent fingerprint minutiae extraction using stacked denoising sparse autoencoders," in *International Joint Conference on Biometrics*, pp. 1–7, 2014.

[35] M. Vatsa, R. Singh, A. Noore, and K. Morris, "Simultaneous latent fingerprint recognition," *Applied Soft Computing*, vol. 11, no. 7, pp. 4260–4266, 2011.

[36] "VeriFinger by NeuroTechnology." www.neurotechnology.com/verifinger.html.

[37] R. Cappelli, M. Ferrara, and D. Maltoni, "Minutia cylinder-code: A new representation and matching technique for fingerprint recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 12, pp. 2128–2141, 2010.

[38] M. Ferrara, D. Maltoni, and R. Cappelli, "Noninvertible minutia cylinder-code representation," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 6, pp. 1727–1737, 2012.

[39] A. A. Paulino, J. Feng, and A. K. Jain, "Latent fingerprint matching using descriptor-based Hough transform," *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 1, pp. 31–45, 2013.

[40] E. Tabassi, C. Wilson, and C. Watson, "NIST Fingerprint Image Quality (NFIQ)," *NIST Report NISTIR7151*, 2004.

[41] A. Sankaran, M. Vatsa, and R. Singh, "Automated clarity and quality assessment for latent fingerprints," in *International Conference on Biometrics: Theory, Applications and Systems*, pp. 1–6, 2013.

Anush Sankaran received the B.Tech. degree in computer science from the Coimbatore Institute of Technology, Coimbatore, India, in 2010. He is currently pursuing the Ph.D. degree with the Indraprastha Institute of Information Technology, New Delhi, India. His research interests include deep learning, image processing, and their applications in biometrics. He was a recipient of the TCS Ph.D. Research Fellowship from 2010 to 2015, and the Best Poster Awards in the IEEE BTAS 2013 and the IEEE IJCB 2014.

Mayank Vatsa received the M.S. and Ph.D. degrees in computer science in 2005 and 2008, respectively from the West Virginia University, Morgantown, USA. He is currently an Associate Professor and AR Krishnaswamy Faculty Research Fellow at the Indraprastha Institute of Information Technology (IIIT) Delhi, India. He has more than 150 publications in refereed journals, book chapters, and conferences. His research has been funded by the UIDAI and DeitY, Government of India. He is a recipient of the FAST award by DST, India. His areas of interest are biometrics, image processing, computer vision, and information fusion. Dr. Vatsa is a member of the IEEE, Computer Society and the Association for Computing Machinery. He is the recipient of several best paper and best poster awards in international conferences. He is also an associate editor of IEEE Access, area editor of Information Fusion, Elsevier, and IEEE Biometric Compendium, and served as the PC Co-Chair of ICB 2013 and IJCB 2014.

Richa Singh received the M.S. and Ph.D. degrees in computer science in 2005 and 2008, respectively from the West Virginia University, Morgantown, USA. She is currently an Associate Professor and recipient of the Kusum and Mohandas Pai Faculty Research Fellowship at the Indraprastha Institute of Information Technology (IIIT) Delhi, India. Her research has been funded by the UIDAI and DeitY, Government of India. She is a recipient of the FAST award by DST, India. Her areas of interest are biometrics, pattern recognition, and machine learning. She has more than 150 publications in refereed journals, book chapters, and conferences. She is also an editorial board member of Information Fusion, Elsevier and EURASIP Journal on Image and Video Processing, Springer. Dr. Singh is a member of the CDEFFS, IEEE, Computer Society and the Association for Computing Machinery. She is a recipient of several best paper and best poster awards in international conferences.