Channel Refinement of Fingerprint Preprocessing Models

Indu Joshi[1,2]*, Ayush Utkarsh[3], Tashvik Dhamija[4], Pravendra Singh[5], Antitza Dantcheva[1], Sumantra Dutta Roy[2] and Prem Kumar Kalra[2]

[1] *Inria Sophia Antipolis, France*

[2] *IIT Delhi, India*

[3] *Independent Researcher, India*

[4] *Delhi Technological University, India*

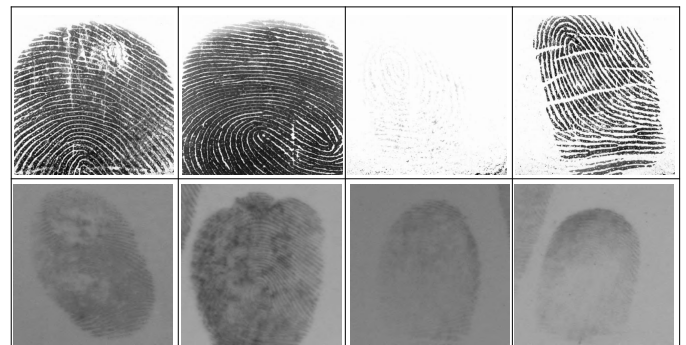[5] *IIT Roorkee, India*

* E-mail: indu.joshi@inria.fr

Abstract: Deep models are the state-of-the-art models for fingerprint preprocessing. However, these models have very high number of parameters, usually in millions. As a result, redundancy is observed among the features learnt by deep learning based fingerprint preprocessing models. A popular technique to help deep models learn distinct and informative features is channel refinement. A recent study has illustrated the capability of channel refinement to improve generalization of fingerprint enhancement models. Motivated by the above-mentioned study, this chapter delves into presenting a detailed study illustrating the usefulness of channel refinement in reducing redundancy and imparting generalization ability to fingerprint enhancement models. Furthermore, we extend this study to assess whether channel refinement generalizes on fingerprint ROI segmentation. Extensive experiments on fourteen challenging publicly available fingerprint databases and a private database of fingerprints of the rural Indian population are conducted to assess the potential of channel refinement on fingerprint preprocessing models.
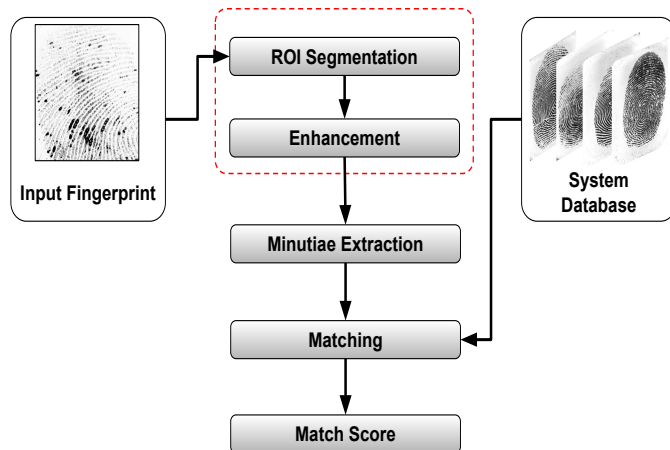
Fig. 1: Schematic diagram of an AFRS. This chapter discusses the impact of channel refinement on deep learning based fingerprint preprocessing models. For understanding, fingerprint preprocessing blocks constituting of *fingerprint ROI segmentation* and *fingerprint enhancement* are marked using the red box.


Fig. 2: Sample poor quality fingerprints used to study the impact of channel refinement on fingerprint preprocessing models.

1 Introduction

The robustness of fingerprints has long established it as a tool for facilitating person identification. Subsequently, Consequently, fingerprints have emerged as one of the most widely used biometrics modalities. These are used in a wide range of applications, including law enforcement, border security and civilian applications. An automated fingerprint recognition (AFRS) constitutes various steps which can be broadly categorized as: *fingerprint acquisition*, *preprocessing*, *feature extraction* and *matching*. A fingerprint preprocessing model is a term used to jointly refer to fingerprint region of interest (ROI) segmentation and fingerprint enhancement (see Figure 1).

Fingerprint preprocessing initiates with the fingerprint ROI segmentation step. This step separates the background and the foreground fingerprint region in a given fingerprint image [1]. Formally, the foreground fingerprint region can be described as the fingerprint image region constituting of fingerprint ridges and valleys. On the other hand, background region in a fingerprint image typically constitutes unnecessary textured patterns such as overlapping text, or noise observed to the presence of oil and dirt on sensor surface or wearing of fingerprint sensing device. Subsequent to fingerprint ROI segmentation, what follows is the fingerprint enhancement step. This steps improves the quality of fingerprint images. In particular, fingerprint enhancement refines the ridge structure of a fingerprint image by improving ridge-valley contrast, removing noise observed in background and predicting the ridge details in unclear regions [2]. Both the fingerprint preprocessing steps are crucial for appropriately extracting features and achieve satisfactory comparison results. The need for preprocessing models is even more profound for poor quality fingerprints with visibly low contrast and missing ridge details (see Figure 2).
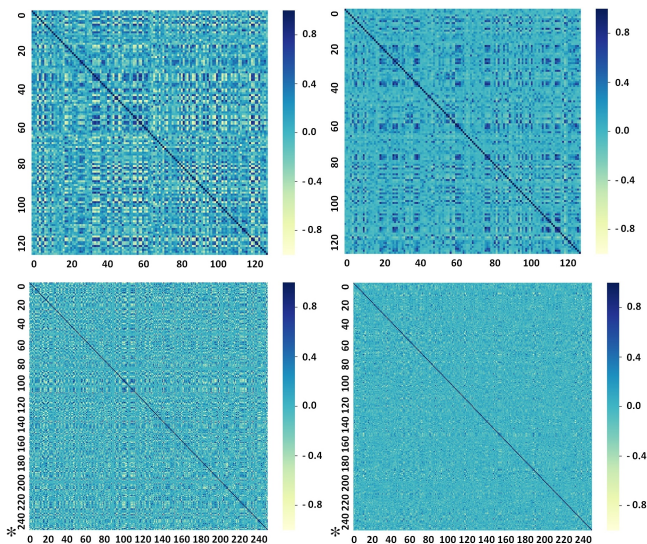
Fig. 3: Left- correlation matrix demonstrating high redundancy among features learnt by cutting-edge CNN based fingerprint pre-processing model. Right- redundancy among features is significantly reduced after incorporating proposed channel refinement.

Deep convolutional neural networks (CNNs) dominate modern fingerprint preprocessing models for fingerprint recognition. These networks have a very large number of layers and model parameters. Although, an effective deep model is expected to learn distinct and useful features, however, with such a high model capacity, it is often observed that the learnt features have high correlation. A cutting-edge method to lessen redundancy among channel weights (features) learnt by CNNs is through channel refinement. Recently, Joshi et al. [3] demonstrate that channel weights learnt by fingerprint enhancement models have redundant information and propose channel refinement of fingerprint enhancement models and demonstrate its effectiveness to reduce redundancy among channel weights (see Figure 3).

Research Contributions

Following research contributions are made in this chapter:

● A recently conducted by Joshi et al. [3] demonstrates the ability of channel refinement to improve the generalization ability of fingerprint enhancement models. This chapter delves into evaluating whether channel refinement generalizes for fingerprint ROI segmentation.
● The performance of channel refinement on several varying fingerprint preprocessing models is studied.
● Additionally, study is done on how channel refining applies to various cutting-edge deep learning models.
● Visualization of feature correlation matrix is provided to understand redundancy among features.
● Relevant ablation studies and comparisons with cutting-edge techniques of channel attention are provided.
● Rigorous experimentation on fifteen fingerprint databases is conducted to understand the generalization ability of channel refinement.

2 Related Work

2.1 Fingerprint Enhancement

Several factors make fingerprint enhancement a highly challenging problem. Some of these factors include presence of smudged ridge patterns, poor contrast between ridges and valleys due to wet or dry

fingertips, presence of sensor noise, oil and dirt on sensing surface or structured background noise due to overlapping text or fingerprint patterns. Several algorithms have been proposed to address these challenges. However, the design of a generalizable fingerprint enhancement algorithm is still an open challenge. We proceed to describe the literature on fingerprint enhancement.

2.1.1 Classical Image Processing Techniques for Fingerprint Enhancement: *Filtering* in spatial or Fourier domain constitutes the most widely used classical image processing based fingerprint enhancement methods. These methods utilize ridge information, e.g., ridge continuity or ridge orientations to predict missing ridge information in poor quality regions [4–11]. The contribution of Hong et al. [4] is one of the introductory and seminal contributions in the field. To approximate the frequency and orientation of the ridge, the authors simulate a sinusoid wave in a direction perpendicular to the ridge orientation. The enhanced fingerprint image is computed via using Gabor filters that are set in accordance with the approximated frequency and orientation. Yoon et al. [12] approximate fingerprint ridge orientation using zero-pole method based on singular points as well as fingerprint skin distortion and rotation model. Gottschlich and Schönlieb [5] exploit anisotropic filtering to propose a locally adaptive fingerprint enhancement algorithm. Filtering is applied with respect to the ridge orientations in the local fingerprint region. Turroni et al. [6] as well utilize context. The context in their approach is approximated through local fingerprint quality, frequency and orientation. The local context is used for adapting the filter accordingly. The authors suggest iterative application of contextual filtering. Filtering begins from regions with high fingerprint quality to the ones with low fingerprint quality. Ramos et al. [7] propose adaptive Gabor filtering that is tuned to the signal frequency. Wang et al. [8] overcome the limited bandwidth challenge of Gabor filter by proposing log-Gabor filters. For filtering, the authors utilize curvature, ridge frequency and the orientation information. Similarly, Gottschlich [9] exploit curvature information by proposing filtering using curved Gabor filters. The contribution of Chikkerur et al. [10] is among the first few approaches that that exploit frequency domain to study fingerprints. The authors approximate ridge frequency and orientation using Short Time Fourier Transform (STFT). Ghafoor et al. [11] propose fingerprint enhancement through filtering in both spatial and frequency domains.

Hsieh et al. [13] propose utilizing both local and global contextual information. The authors execute wavelet decomposition to extract global information. The enhanced fingerprint is generated using wavelet reconstruction. Jirachaweng and Areekul [14] note the limitations of Gabor filters in generating fingerprints with blocking artifacts and poor ridge continuity around high curvature fingerprint regions. The authors exploit filtering on the Discrete Cosine Transform (DCT) domain to address the limitations of Gabor filters. We note that since the classical image processing based fingerprint enhancement methods directly exploit contextual information, ridges in poor quality fingerprint regions with unreliable contextual information are often incorrectly enhanced. Learning based fingerprint enhancement methods are proposed to address these limitations of classical image processing based methods.

2.1.2 Learning Based Algorithms for Fingerprint Enhancement: Several notable works in fingerprint enhancement utilize *learnable dictionaries* to estimate fingerprint ridge orientations [15–19]. Feng et al. [15] propose a dictionary based fingerprint enhancement method that exploits compatibility between orientations in neighbouring fingerprint image patches to estimate fingerprint ridge orientations. Yang et al. [16] observe that only certain orientations exist in a specific fingerprint region. The authors exploit this information by proposing location-specific dictionaries for quicker dictionary look-ups and lower orientation estimation errors. Chen et al. [17] observe that dictionary with larger patches are better suited for poor quality fingerprint regions. Motivated by this observation, the authors propose multi-scale dictionaries to account for varying amount of noise in different fingerprint regions. Liu et al. [18] learn efficient dictionaries for fingerprint enhancement by sparse coding the fingerprint ridge orientation dictionaries. Chaidee et al.

[19] exploit information from both Gabor and curved filters. The response of both these filters are combined to construct a frequency domain dictionary which is subsequently to estimate orientations and generate the enhanced image.

A standard practice for dictionary based fingerprint enhancement approaches is that the orientation dictionaries are computed from good quality fingerprint regions. Thus, dictionary based methods often perform poorly on low quality fingerprint regions. This shortcoming of dictionary based approaches is addressed by *orientation prediction networks* [20, 21]. Cao and Jain [20] formulate fingerprint ridge orientation approximation to be a classification task. The authors propose a CNN to perform the orientation classification. Qu et al. [21] pose orientation angle prediction as a regression problem and exploit a deep model estimate the direction of the fingerprint ridge. Driven by the accomplishments of deep models in orientation estimation, researchers argued to *directly generate the enhanced fingerprint* as opposed to predicting orientation [2, 22–27].

Sahasrabudhe and Namboodiri [22] suggest enhancing fingerprint using a deep belief network. Schuch *et al.* [23] exploit a deconvolutional autoencoder (DeConvNet) for enhancing fingerprints. Svoboda *et al.* [25] introduce domain knowledge into the fingerprint enhancement domain by introducing an gradient and orientation optimizing autoencoder model. Qian *et al.* [26] propose to enhance fingerprint patches using DenseUnet based fingerprint enhancement model. Li *et al.* [2] and Wong and Lai [27] demonstrate that multitasking using an orientation correction task facilitates improved fingerprint enhancement performance. Joshi *et al.* [28, 29] demonstrate that generative adversarial network is an effective model for fingerprint enhancement. Later, the authors demonstrate that Monte Carlo Dropout [30] imparts interpretability to fingerprint enhancement models while estimating data uncertainty facilitates noise-aware fingerprint enhancement [31]. The most recent works in fingerprint enhancement domain include cross-domain consistency [32] and context-aware enhancement by solving jigsaw puzzles [33]. For a comprehensive survey on fingerprint enhancement methods, the readers are referred to the survey by Schuch *et al.* [34]. To summarize, we note that existing state-of-the-art fingerprint approaches are autoencoder or generative adversarial network based approaches. However, designing a generalizable fingerprint enhancement model is still an open challenge in the domain.

2.2 Fingerprint ROI Segmentation

Similar intensities of foreground and background pixels around fingerprint boundaries make Fingerprint ROI segmentation a challenging problem. Many algorithms are proposed to address the challenges posed by fingerprint ROI segmentation. However, generalization on disparate sensing technologies is an existing challenge. We now describe the literature on fingerprint ROI segmentation which can be broadly categorized as follows:

2.2.1 Classical Image Processing Techniques for Fingerprint ROI Segmentation:
Hu *et al.* [35] and Hai *et al.* [36] suggest *frequency domain filtering* to differentiate between foreground and background. A fusion-based segmentation technique is recommended by Hu *et al.* [35] that combines information from the frequency domain and domain knowledge of orientations. At first, they apply log-Gabor filtering on the fingerprint image followed by adaptive thresholding to obtain the first segmentation mask. The second mask is obtained using the orientation reliability metric defined by the authors. Both these masks are fused followed by postprocessing to obtain the segmented ROI. Thai *et al.* [36] argue that the frequencies of ridge patterns observed in the foreground fingerprint image region lie only in a specific band of the Fourier spectrum. They suggest Fourier domain filtering using factorized directional bandpass filter. The aforementioned technique attenuates frequencies occurring due to artifacts and only preserve the relevant frequencies pertaining to the true fingerprint region. The reconstructed image is then followed by morphological operations to obtain the segmented fingerprint image.

Some approaches use only *morphological operations* for fingerprint ROI segmentation [37, 38]. To separate the noise and

fingerprint region, Thai and Gottsclich [37] suggest a three-part decomposition approach. To create the segmented image, morphological techniques are applied to the binarized fingerprint image. For the input fingerprint image, Fahmy and Thabet [38] compute the range image. Subsequently, adaptive thresholding is used to transform the range image into a binary image. To create the segmented image, morphological techniques and contour smoothing are used to the binary image. Another prominent direction explored by researchers to approach fingerprint ROI segmentation is through making use of *ridge orientation information* [39, 40]. Teixeira and Leite [39] suggest a multi-scale pyramidal structuring element for monotonic filtering of image extrema while a multi-scale directional operator is used to determine the orientation of each pixel. The computed directional field is used to estimate the background. The segmented image is obtained using after subtracting the estimated background followed by postprocessing. Raimundo *et al.* [40] point out that the ridges and valleys can be identified as pixels parallel and normal to the ridge orientation, respectively. Further, the authors use orientation information to assess the quality of ridge and valleys and compute directional images. Clustering on the directional image followed by postprocessing is performed to find the segmented image. Different from the above-mentioned approaches, Wu *et al.* [41] observe that the strength of Harris corner points in the background is lower compared to the foreground. The high corner strength possessed by ridge boundaries helps to distinguish between foreground and background.

A significant limitation of traditional image processing based fingerprint ROI segmentation approaches is that these approaches do not perform well when the intensity of background pixels is similar to the intensity of foreground fingerprint pixels. Furthermore, the segmentation performance of these approaches is heavily dependent on the postprocessing step. Therefore, to improve the segmentation performance, learning based approaches are proposed, which are described next.

2.2.2 Learning Based Algorithms for Fingerprint ROI Segmentation:
Initial approaches to learning based ROI segmentation algorithms propose to *cluster image pixels* to identify foreground and background [42–45]. Yang *et al.* [42] calculate coherence, mean and variance features from non-overlapping blocks of the fingerprint image. These features are then clustered through K-means clustering. The clusters are then classified into foreground or background using voting of neighbours. Morphological postprocessing is performed to obtain the final segmented image. Ferreira *et al.* [43] apply block-wise range filter to enhance the ridges of the fingerprint image. The output range image that results is clustered and binarized. Later, the ROI segmentation mask is attained by conducting postprocessing. Lei and Lin *et al.* [44] propose to apply a range filter to enhance fingerprint ridge boundaries. The clusters created by the range filter are then combined. To create the ROI mask, morphological operations are used. Ferreira *et al.* [45] apply filtering through the range, entropy and variance filters. Three different clustering methods are evaluated. The decision is re-evaluated through different classifiers. The final image is obtained after postprocessing of the binarized image. However, a common shortcoming of clustering based approaches is that these require prior knowledge of the number of clusters and their performance is heavily dependent on this parameter. This adversely affects the generalization ability of clustering based approaches.

Some learning based methods work on a patch-level and *classify a given patch as foreground or background* [1, 46–51]. Stojanović *et al.* [1] propose to divide each image into patches and classify each patch using off-the-shelf CNN model, AlexNet. Predictions for each block are combined for achieving ROI mask. This mask is smoother by postprocessing to obtain the segmented ROI mask. Liu *et al.* [46] extract handcrafted texture and intensity features from patches of a fingerprint image. The foreground or background label for each patch is obtained using an Adaboost classifier. Later, postprocessing is applied to obtain a smooth ROI mask. Zhu *et al.* [47] extract multi-sized overlapping patches and classify them using three different neural networks. Predictions from the three networks are

combined to generate a prediction score for each patch. The prediction scores are thresholded, followed by postprocessing to generate the segmentation mask. Ezeobiejesi and Bhanu [48] propose a two-phased method for ROI segmentation where in the first phase model learns representative fingerprint features by training a hierarchy of Restricted Boltzmann Machines (RBMs) to learn an identity mapping of image patches. In the second phase, the network is trained to predict foreground or background class for each patch. Predictions of all the patches are merged to obtain the ROI mask. Serafim et al. [49] suggest classifying image patches through AlexNet and apply smoothing and hole filling postprocessing techniques on the combined predictions for achieving ROI mask. Sankaran et al. [50] extract saliency, intensity, gradient, ridge and quality based features for each patch and perform feature selection to find the most discriminative features. Each patch is classified as either foreground or background using a random forest classifier trained on the chosen features. Khan and Wani [51] exploit CNN for classifying a patch as background or foreground. Furthermore, they take a majority of neighbours to decide whether a patch is misclassified and change the label of a misclassified patch.

The patch based classification approaches for ROI segmentation suffer from many limitations. The first being the high computational time since the network has to make a prediction for each patch. Second, the segmented ROI obtained through these approaches suffer from block-effect around boundaries and needs postprocessing. Joshi et al. [30, 31] demonstrate that recurrent Unet (RUnet) [52] is a promising CNN architecture for fingerprint ROI segmentation. However, designing a generalizable fingerprint ROI segmentation is still an open challenge in fingerprint ROI segmentation domain.

2.3 Attention Mechanisms

The attention mechanism is a technique for directing the viewer's attention on an image's most crucial components and ignoring unimportant aspects. The attention mechanism of the human visual system is used to examine [53, 54] and comprehend complicated pictures effectively and efficiently. This has motivated researchers to enhance computer vision systems' functionality by adding attention methods. A visual system's attention mechanism can be conceptualized as a dynamic process for selection that is implemented through adaptively weighing features in line with the significance of the input. We can classify existing attention methods into six categories: channel attention, spatial attention, temporal attention, branch attention, channel & spatial attention, and spatial & temporal attention. Channel attention [55–61] creates a channel-wide attention mask and uses it to identify the most crucial channels. Spatial attention [62–67] creates attention masks across spatial domains and employs them to either directly predict or pick out important spatial locations. By creating an attention mask in real-time, temporal attention [68, 69] chooses key frames. Branch attention [70–73] creates an attention mask over all of the branches and uses it to identify the most significant ones. In order to choose key characteristics, channel & spatial attention [74–78] either directly produce a joint 3-D channel, width, and height attention mask, or individually forecast channel and spatial attention masks. In order to focus attention on informative locations, spatial & temporal [79–82] attention computes temporal and spatial attention masks individually or generates a joint spatiotemporal attention mask.

2.3.1 Channel Attention:
In deep CNNs, distinct channels of various feature maps typically signify distinct objects [83]. Channel attention is comparable to a mechanism of choosing an object to decide what to focus on by adaptively recalibrating [84] the weight of each channel. Hu et al. [55] introduced SENet and originally put out the idea of channel attention. A squeeze-and-excitation (SE) block, which is the core element of SENet, is utilised to gathering global data, record associations between channels, and enhance representational capacity. Squeeze modules and excitation modules are the two components that make up SE blocks. Global average pooling is used in the squeeze module to collect spatial data on a global scale. Utilizing fully connected layers, the excitation module collects channel-wise relationships and generates an attention vector.

The relevant component of the attention vector is multiplied to scale each channel of the input feature [85].

In order to enhance the squeeze module, Gao et al. [56] suggested utilizing a global second-order pooling block for representing high-order statistics while accumulating global information. It is not practical to employ a SE block post each convolution block because of the computation complexity and a large number of parameters in the fully connected layer in the excitation module. Additionally, an implicit approach for modeling channel relationships is to use fully connected layers. Yang et al. [57] suggested using gated channel transformation to efficiently capture information while explicitly describing channel-wise interactions in order to solve the aforementioned issues. SENet decreases the number of channels to avoid having a complex model. However, this approach fails to accurately simulate the relationship between the weight vectors and the inputs, which lowers the quality of the output. Wang et al. [58] suggested the efficient channel attention block to get around this problem by determining the interaction between channels using a 1D convolution rather than dimensionality reduction. The squeeze module can only use global average pooling, which has representational limitations. Qin et al. [86] exploit frequency domain information for applying global average pooling operation so as to get a more potent representation capability. In order to describe the link between probabilities of object categories and scene context, Zhang et al. [87] suggested the semantic encoding loss pertaining to a context encoding module. For semantic segmentation, this approach makes advantage of contextual information present in a global scene.

2.3.2 Spatial Attention:
One way to think of spatial attention is as an adaptive mechanism for choosing which spatial region to focus on. Particularly for big inputs, convolutional neural networks require a significant computing cost. Mnih et al. [62] introduced the recurrent attention model , which uses reinforcement learning [88] and recurrent neural networks (RNNs) [89] to instruct the network where to focus its attention for concentrating scarce computational assets on critical areas. Ba et al. [63] suggested a deep recurrent network, similar to [62], that can process an image glimpse for tasks involving several objects. Here glimpse refers to an image croppings at several different resolutions. To be specific, the suggested model uses a glimpse as an input to update its hidden state, and at each step predicts both the location of the subsequent glimpse and a new object. The network's computational efficiency is enhanced by the fact that the glimpse is typically substantially smaller compared to entire image. Xu et al. [64] suggested to exploit both soft and hard attention to help an image caption generation model visualise where and what it ought to concentrate. By enabling the user to discern the model's emphasis, the application of suggested attention model enhances the interpretability of the procedure of creating image captions. Additionally, it aids in enhancing the network's capacity for better representation. Hu et al. [65] developed GENet, which draws inspiration from SENet, with the goal of offering a spatial domain recalibration function for gathering long-range dependencies present in spatial domain. Part gathering and excitation procedures are combined in GENet. It combines input features over broad neighbourhoods in the first stage and simulates the relationship between various spatial locations. The second stage begins by utilising interpolation to create an attention map whose dimensions are similar to the dimensions of feature that is provided as the input. The next step is to scale the input feature map by multiplying each point by the matching element in the attention map.

2.3.3 Temporal Attention:
Video processing typically makes use of temporal attention, which may be regarded of as a dynamic means for choosing the time when the model must be attentive. In research on video representation learning, temporal pooling and RNNs has been frequently utilised to capture interaction among different frames. However, such techniques suffer from limited ability to model temporal relationship. In order to get around these, Li et al. [68] suggested to learn global-local temporal representation to take advantage of many scales of temporal information present in a video clip. The suggested attention model consisted of a a

temporal self attention module that helps to capture temporal dependencies on a global scale. Additionally, the model proposed to learn temporal dependencies on a local scale through employing dilated convolutions that span a variety of temporal ranges with progressively increasing dilatation rates. Different outputs are concatenated to merge multi-scale information. Liu et al. [69] introduced a temporal adaptive module to efficiently and adaptably capture complicated temporal interactions. In order to collect context at a global scale, however, in lesser time than [68], it utilised an adaptive kernel as opposed to self-attention.

2.3.4 *Branch Attention:*
When employed with a multi-branch structure, branch attention may be regarded of as a dynamic means for choosing the branch on which the module must focus. In order to solve the issue of training very deep networks, Srivastava et al. [70] suggested highway networks, which use adaptive gating techniques to facilitate information flows across layers. Using straightforward gradient descent techniques, very deep highway networks may be immediately trained. This is made possible by the gating mechanism and skip-connection structure. Information is routed through layers due to the gating mechanism, which, in contrast to fixed skip-connections, adapts to the input. According to studies conducted in the field of neuroscience, in response to the input signal, visual cortical neurons adaptively modify the dimensions pertaining to receptive fields [90]. Li et al. [71] were motivated by this to propose the selective kernel (SK) convolution using the automatic selection method. Three operations are used to implement SK convolution: split, fuse, and select. The feature map is subjected to transformations with various kernel sizes during split to produce various sizes RFs. Then, to compute the gate vector, element-wise summation is used to merge information available from the different branches. By doing this, the information flow from the various branches is regulated. The final feature map is then created via combining all branch feature maps under the guidance of the gate vector.

Convolution kernel uniformity is a fundamental presumption in CNNs. In light of this, increasing a network's depth or width is typically the best approach to improve its ability to represent information, but doing so comes with a high additional cost to computation. Yang et al. [72] suggested a multi-branch convolution referred as CondConv for boosting the capability of convolutional neural networks. Leveraging a computationally less expensive means for branch attention, CondConv fully exploits the benefits of the multi-branch structure. It offers an innovative way to effectively improve networks' capacity. The model's width and depth are constrained by lightweight CNNs' incredibly low computational cost, thus lowering the representational power of the networks. In order to solve the aforementioned issue, dynamic convolution was suggested by Chen et al. [73], that in addition to CondConv [72], boosts representational power at a small additional computational cost without altering the network's width or depth.

2.3.5 *Channel & Spatial Attention:*
The benefits of both channel and spatial attention are combined in channel & spatial attention. It chooses important objects and regions in an adaptive manner. Residual attention network [74] highlighted the significance of relevant elements in both the dimensions: spatial and channel. Subsequently, the authors [74] established the channel spatial attention domain. It uses numerous convolutions in a bottom-up framework to build a 3D (channel, breadth, and height) attention map. Woo et al. [75] suggested to serially stack channel and spatial attention and referred it as a convolutional block attention module. The proposed method improved both informative channels and important regions. Decoupling channel and spatial attention enables increased computational efficiency. Global pooling is then used to make use of global spatial information. In addition to [75], Park et al. [76] also put out the bottleneck attention module (BAM), which sought to effectively increase networks' representational capacity. The spatial attention sub-receptive module's field is expanded using dilated convolution, and to reduce computational costs, a bottleneck structure is built in accordance with ResNet's suggestions. The feature map is subjected to global pooling by a SE block, which combines global spatial data. However, it disregards the spatial information at

the pixel level, which is crucial for dense prediction problems. Roy et al. [77] thus suggested spatial and channel SE blocks. In a manner similar to BAM, spatial SE blocks are utilised in addition to SE blocks for enabling concentration on salient input regions through weights pertaining to spatial dimensions. Both convolution block attention module [75] and BAM [76] individually estimated separately in CBAM and BAM, neglecting the connections between these two domains. Triplet attention, a straightforward but effective attention method to model cross-domain interaction, was first described by Misra et al. [78].

2.3.6 *Spatial & Temporal Attention:*
The benefits of both spatial attention and temporal attention are combined in spatial & temporal attention, which adaptively chooses both salient regions and crucial frames. Each type of action in human action recognition often only depends on a small number of distinct kinematic joints. Multiple actions might also be carried out across time. These findings led Song et al. [79] to propose a hybrid LSTM model [89] for spatial & temporal attention. Suggested model helps to adaptively identify keyframes and discriminative features. A spatial attention sub-network that selects significant areas and a temporal attention sub-network that selects critical frames are its key modules for learning attention weights. In order to choose important aspects globally and adaptively in order to capture spatial and temporal information that exists in video frames, Du et al. [80] introduced spatiotemporal attention. The sequential application of spatial attention component followed by a temporal attention component constitutes the framework for spatiotemporal attention in [80]. Prior attention models to re-identify individuals in videos merely gave each frame a weighted attention value; they were unable to recognise joint spatial and temporal relationships. Fu et al. [81] presented a novel spatiotemporal attention method to address this problem, which rates attention in every spatial location in various frames with the use of no additional parameters. In order to understand the spatial dependencies inside a frame and the temporal links among different frames, Yang et al. [82] introduced a spatiotemporal graph convolutional network. The network is trained to learn discriminative features from a video. Pairwise similarity is used to build a patch graph. Subsequently, combines data using graph convolution.

To summarize, depending upon the nature of application, different kinds of attention mechanisms can be introduced in a deep neural network. For a deep fingerprint preprocessing model, we hypothesize that its huge model capacity makes it prone to poor generalization. In order to ensure learning of distinct features that generalize over different databases, for fingerprint preprocessing models, we propose to exploit channel-level attention.

3 Proposed Method

This subject of this research is to study the significance of channel-level dependencies in a deep learning based model fingerprint preprocessing. We propose a method to effectively refine the features learnt by a fingerprint preprocessing model such that redundancy among features is reduced and improved generalization is obtained. We now proceed to describe channel refinement unit (CRU) to refine the features of fingerprint preprocessing models.

3.1 *Channel Refinement Unit*

A channel refinement unit (CRU) is designed to transform the features such that redundancy among channels is reduced. For an input feature $X \in R^{L \times H \times C}$, CRU transforms it into $X_{new} \in R^{L \times H \times C}$ such that X_{new} has lesser channel level redundancy and more informative features are learnt by the fingerprint preprocessing model (see Figure 4). Convolution is a local operation due to its limited receptive field. As a result, standard convolution operation cannot successfully extract the global information from fingerprint images. CRU exploits global information derived using the activation maps to evaluate the significance of each channel towards fingerprint processing. Therefore, as its very first step, CRU computes the global representation corresponding to each channel of a given convolution
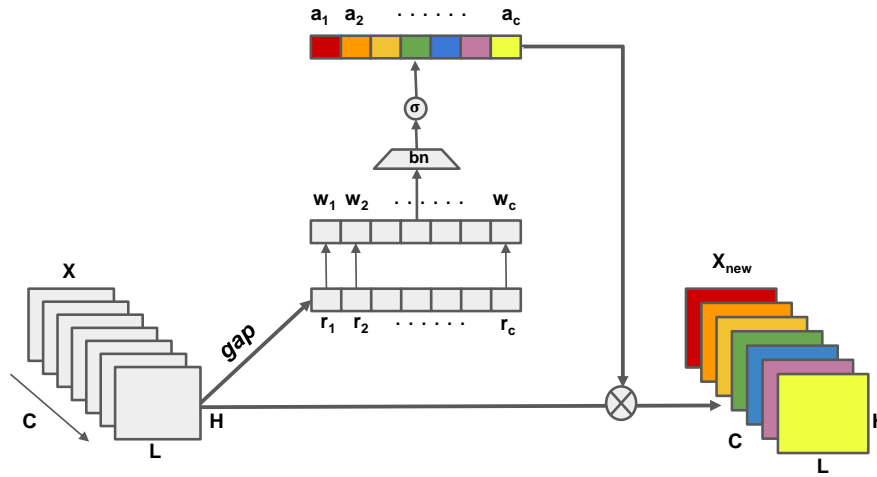
Fig. 4: CRU transforms X to X_{new} such that redundancy among channels is reduced and the fingerprint preprocessing model is enabled to learn more informative features, leading to improved fingerprint preprocessing performance.
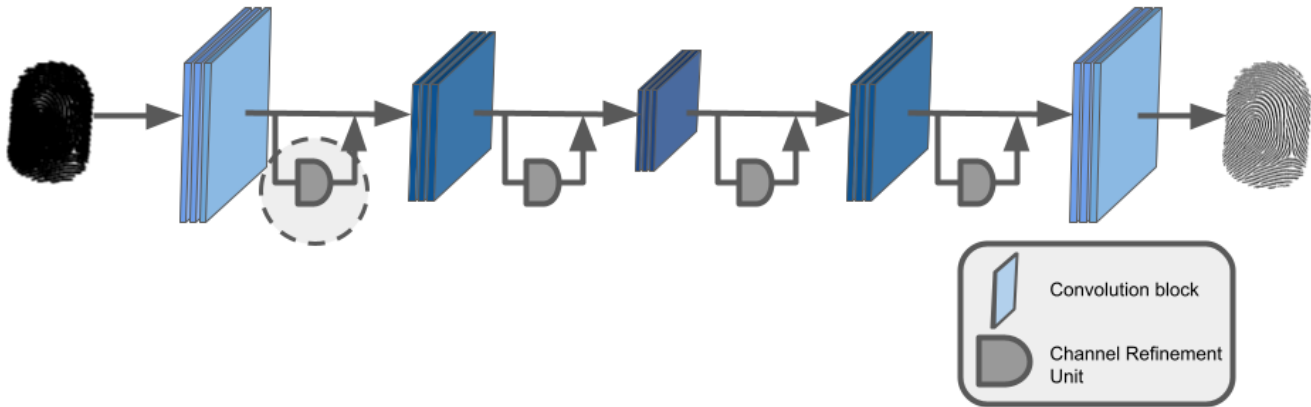
Fig. 5: Flowchart depicting introduction of CRU into an existing fingerprint preprocessing model. CRU transforms the features learnt by the baseline fingerprint preprocessing model such that redundancy among channels is reduced. Thus, CRU helps to learn more informative features, facilitating improved fingerprint preprocessing performance.

layer. To compute the global representation, CRU leverages a global average pooling (gap) layer. This (gap) layer averages activations across channels, to output a C dimensional vector, corresponding to each input channel.

Next operation in CRU is aimed towards understanding the significance of each of the C channels. To achieve this, CRU executes C depth wise convolution (dwc) (1×1). This provides an output vector $W=[w_1, w_2, ...w_C]$, where $w_i \in R$. The operations that follow on Vector W are batch normalization (bn) and sigmoid activation (σ). These operations output a refinement vector (A) that represents how much each channel must be refined to obtain the optimal fingerprint preprocessing performance. $A=[a_1, a_2, ...a_C]$ $(a_i \in R)$, this vector contains a refinement weight for each of the input channel. The refined feature X_{new} is obtained through conducting element wise product between input feature X and the corresponding refinement weight pertaining to its channels.

$$A = \sigma(bn(dwc(gap(X))))$$
$$X_{new} = [x_1 \cdot a_1, x_2 \cdot a_2, ...x_C \cdot a_C]$$

Instead of X, after exploiting CRU the refined feature X_{new} is used for computation at the following layers to generate the improved preprocessed fingerprints.

3.2 Introducing Channel Refinement Unit into a Fingerprint Preprocessing Model

The incorporation of CRU does not need much change in the network design of the backbone deep model for fingerprint preprocessing. CRU is included after each convolution layer. As a result, refined features after each convolution layer and these refined features are forwarded to the next convolution layer for generating the preprocessed fingerprint image at the output layer (see Figure 5). The refined features have lesser redundancy and therefore capture distinct and more informative characteristics from input fingerprints. Learning of more informative features facilitates improved fingerprint preprocessing performance.

4 Experimental Setup

We rigorously assess the impact of channel refinement by introducing CRU into different fingerprint preprocessing models conducting different fingerprint preprocessing tasks (fingerprint ROI segmentation and fingerprint enhancement).

4.1 Databases

For evaluating this research, we use fourteen challenging and publicly available fingerprints databases, and a challenging private

Database	Resolution	Sensor Name	Sensing Tech.
2000 DB1	300×300	S.D. Scanner	Optical
2000 DB2	256×364	TouchChip	Capacitive
2000 DB3	448×478	DF-90	Optical
2000 DB4	240×320	Synthetic Generator	NA
2002 DB1	388×374	TouchView II	Optical
2002 DB2	296×560	FX2000	Optical
2002 DB3	300×300	100 SC	Capacitive
2002 DB4	288×384	Synthetic Generator	NA
2004 DB1	640×480	V300	Optical
2004 DB2	328×364	U.are.U 4000	Optical
2004 DB3	300×480	FingerChip	Thermal
2004 DB4	288×384	Synthetic Generator	NA

Table 1 Details about FVC databases exploited to assess fingerprint ROI segmentation performance.

Fig. 6: Examples presenting latent fingerprints [93] enhanced by CR-GAN and baseline FP-E-GAN.

Enhancement Algorithm	Quality Score
Raw Image	4.96
FP-E-GAN [28]	1.91
CR-GAN	**1.77**

Table 2 An analysis of the latent fingerprints' [93] average quality scores computed through NFIQ.

dataset of poor quality fingerprints, resulting in a total of fifteen databases for experimental evaluation of the proposed CRU.

4.1.1 Fingerprint Enhancement: as As suggested in the literature [28, 29, 91], fingerprint enhancement models are trained in a supervised manner using synthetic data. For evaluating the fingerprint processing performance on the enhancement task, we use the following databases:

1. *Multisensor Optical and Latent Fingerprint Database (MOLF)* [92]: MOLF is the largest publicly available latent fingerprint database. This databases constitutes 4400 poor quality latent fingerprints acquired from 100 volunteers. The volunteers press their fingertips on ceramic tiles. The latent fingerprints are captured via exploiting the standard black powder dusting process.
2. *Rural Indian Fingerprint Database* [93]: This is a challenging open access dataset of poor quality fingerprints. These fingerprints are collected from rual Indian volunteers rigorously involved with physical activity requiring extensive use of fingers, such as farming. In total, this dataset has 1631 poor quality fingerprints obtained from an optical sensor.
3. A private rural Indian fingerprint database [94] with samples from aging and the subjects engaged with rigorous manual work. This dataset has 1000 poor quality fingerprint samples obtained from an optical sensor.

4.1.2 Fingerprint ROI Segmentation: To evaluate the fingerprint processing performance on the ROI segmentation task, we use the *Fingerprint Verification Challenge (FVC) databases*. We evaluate the fingerprint ROI segmentation on different series of FVC competitions, FVC2000, FVC2002 and FVC2004. The fingerprints in these databases are acquired from different fingerprint sensing technologies, i.e. optical, capacitive, thermal and even synthetic fingerprints [95]. Each series has 4 databases, each with 80 training samples and 800 testing samples. Therefore, in total, we train the fingerprint ROI segmentation model on 960 images and evaluate the fingerprint ROI segmentation performance on 9600 fingerprints. The authors of [37] prepared the ground truth annotations of ROI. The information about the FVC databases is presented in Table 1.

4.2 Assessment Criteria

We proceed to describe the different assessment criteria used to judge the fingerprint preprocessing performance.

4.2.1 Fingerprint Enhancement: We exploit the following metrics to evaluate the fingerprint enhancement performance:

1. *Ridge Structure Preservation Ability*: Ridge details in a fingerprint image entail the identity information. Therefore, it is crucial that the fingerprint details are preserved while enhancing the fingerprint images. To measure the fingerprint enhancement model's capacity to preserve ridge structure, we calculate common measures like SSIM [96], Jaccard similarity score [97], PSNR [98], and
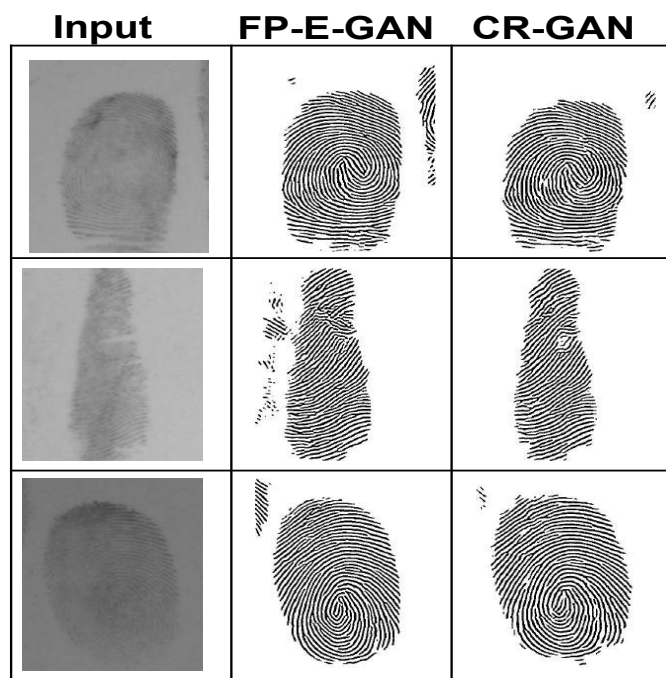
Dice score [99] between ground truth binarization and the enhanced fingerprint.
2. *Fingerprint Quality Assessment*: A fingerprint enhancement algorithm is meant to improve the quality of the input fingerprint image. Using the publicly available tool *nfiq* [100], we compute fingerprint quality scores for measuring the improvement in quality of fingerprints. This tool assess a given input fingerprint on various parameters such as ridge-valley clarity, number of minutiae and ridge smoothness.
3. *Fingerprint Matching Performance*: We also quantify the improved matching performance on enhanced images. Matching performance on latent fingerprints is analyzed in identification mode of fingerprint recognition. As a result, for latent fingerprints, the matching performance is analyzed by calculating rank-50 accuracy and plotting the corresponding cumulative matching curve (CMC). On the other hand, the matching performance on the rural Indian fingerprints is analyzed in verification mode of fingerprint recognition. Subsequently, for rural Indian fingerprints, the matching performance is analyzed by presenting the detection error trade-off (DET) curve that corresponds to the average equal error rate (EER).

4.2.2 Fingerprint ROI Segmentation: We evaluate the fingerprint ROI segmentation performance by computing the standard metrics for assessing segmentation performance: *Dice score* [99] and *Jaccard score* [97].
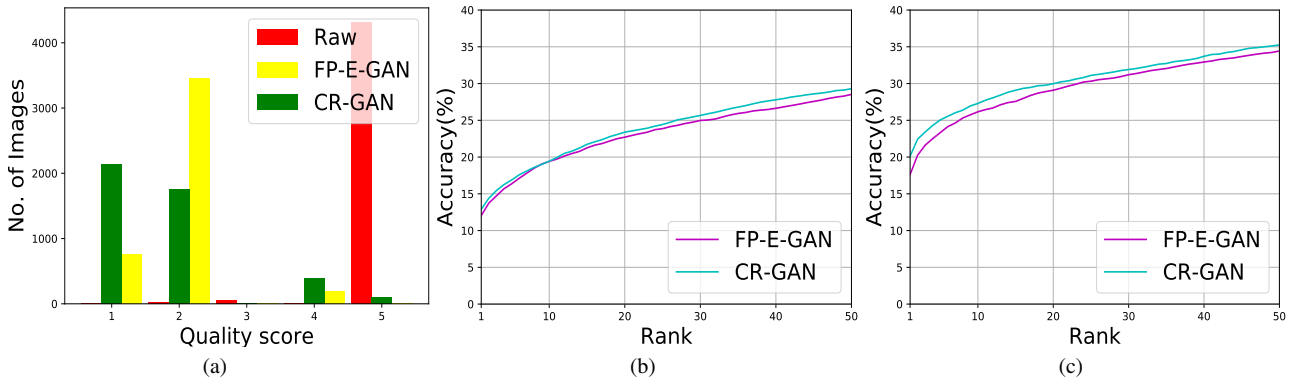
Fig. 7: Enhancement performance before (FP-E-GAN [28]) and after introduction of CRU (CR-GAN). (a) histogram presenting the distribution of quality scores of fingerprints, CMC curves obtained using (b) Bozorth and (c) MCC.
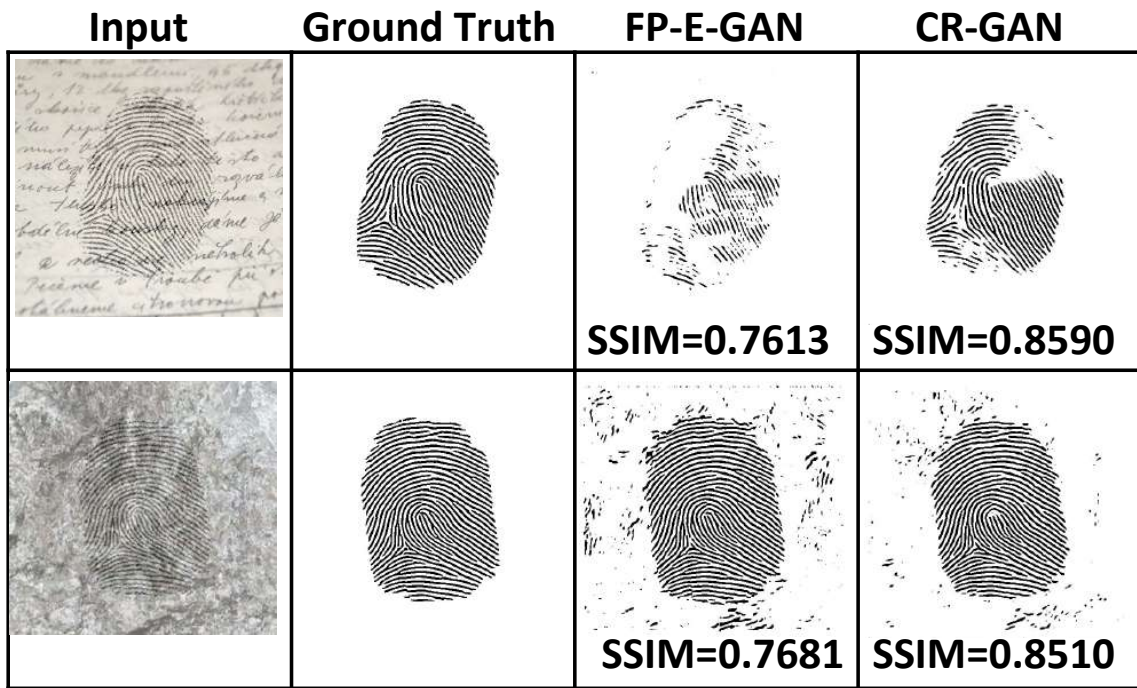
Fig. 8: Examples demonstrating improved ridge preservation ability after introduction of CRU.
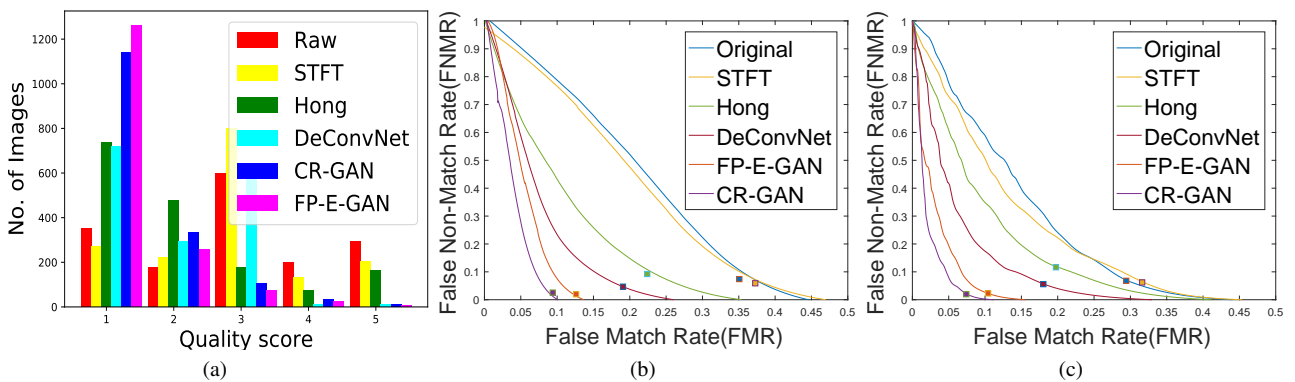
Fig. 9: State-of-the-art methods for fingerprint enhancement are compared on [92] using (a) histogram presenting the distribution of quality scores of fingerprints, DET curves acquired from (b) Bozorth, and (c) MCC.

Enhancement Algorithm	Bozorth	MCC
Raw Image	5.45	6.06
Svoboda *et al.* [25]	NA	22.36
FP-E-GAN [28]	28.52	34.43
CR-GAN	**29.30**	**35.25**

Table 3 An analysis of the latent fingerprints' [93] identification results when compared against gallery of Lumidigm sensor.

Enhancement Algorithm	Bozorth	MCC
Raw Image	16.36	13.23
STFT [10]	18.13	14.52
Hong *et al.* [4]	11.01	11.46
DeConvNet [23]	10.93	10.86
FP-E-GAN [28]	7.30	5.96
CR-GAN	**5.72**	**4.45**

Table 4 An analysis of verification results obtained on [92] quantified by average equal error rate.

5 Results: Fingerprint Enhancement

5.1 Enhancement of Latent Fingerprints

We begin the analysis of impact of introducing CRU into a fingerprint enhancement model by assessing its performance on latent fingerprints. Figure 6 depicts samples enhanced before (FP-E-GAN) and after introducing CRU (CR-GAN). We find that introduction of CRU enables improved enhancement performance quantified by improved fingerprint quality scores and improved rank-50 accuracy on both Bozorth and MCC fingerprint matching tools (see Table 2 and Table 3). The corresponding histogram comparing fingerprint quality scores and CMC curves corresponding to both the matching tools are presented in Figure 7. We also find that introducing CRU enables improved ridge preservation ability to FP-E-GAN [28], quantified by higher SSIM scores after introducing CRU (CR-GAN) compared to FP-E-GAN (see Figure 8). Having successfully demonstrated improved fingerprint enhancement of latent fingerprints, we proceed with analysis of fingerprint enhancement performance on rural Indian fingerprints.

5.2 Enhancement of Rural Indian Fingerprints

Subsequently, we assess the enhancement results obtained by CR-GAN on two rural Indian fingerprint databases. We contrast the enhancement results obtained by CR-GAN with the cutting-edge models for fingerprint enhancement: STFT [10], DeConvNet [23], FP-E-GAN [28], and Hong [4]. Table 4 and Table 5 reflect on the improved performance of fingerprint matching as measured by average equal error rate (EER) and the corresponding DET curves are presented Figure 9 (b) and Figure 9 (c). Regardless of the fingerprint matching algorithm chosen, the average EER is much lower in both rural fingerprint databases. This demonstrates how the suggested CRU enhances FP-E-GAN's performance [28]. As a result, on both datasets, CR-GAN performs better than state-of-the-art. These findings support the assertion that FP-E-GAN does indeed learn some redundant channel weights during training, and that channel weight refinement aids in enhancing FP-E-GAN's performance. Then, we contrast the CR-GAN fingerprint quality ratings with cutting-edge algorithms for fingerprint enhancement. We display distribution of NFIQ values through a histogram in Figure 9 (a), while Table 6 displays the average NFIQ score. The results demonstrate that while matching performance is greatly increased, the quality of enhanced fingerprints that CR-GAN produces is comparable to that of FP-E-GAN.

The example restored fingerprint pictures in Figure 10 are created using both the proposed CR-GAN and the most recent cutting-edge algorithms for fingerprint enhancement. When compared to the currently available fingerprint enhancement algorithms, CR-GAN

Enhancement Algorithm	Bozorth	MCC
DeconvNet [23]	28.75	26.80
FP-E-GAN [28]	17.06	15.85
CR-GAN	**13.23**	**11.52**

Table 5 An analysis of verification results obtained on the private fingerprint database quantified by average equal error rate.

Enhancement Algorithm	Quality Score
Raw Image	2.94
STFT [10]	2.86
Hong *et al.* [4]	2.05
DeconvNet [23]	1.95
CR-GAN	1.42
FP-E-GAN [28]	**1.31**

Table 6 An analysis of the openly accessible rural Indian fingerprints' [92] average quality scores computed through NFIQ.

Network	Generator	Discriminator	Total
FP-E-GAN [28]	11376129	2765505	14141634
CR-GAN	11383041	2768193	14151234
SE-GAN [55]	12072081	3165177	15237258

Table 7 Comparison of the CRU's introduced model parameters with model parameters introduced by SE-block [55].

Enhancement Algorithm	Quality Score
SE-GAN [55]	1.76
CR-GAN	**1.42**

**Table 8** An analysis of average quality scores (computed through NFIQ) achieved for proposed CRU and SE block.

Enhancement Algorithm	Bozorth	MCC
SE-GAN [55]	12.34	10.50
CR-GAN	**5.72**	**4.45**

Table 9 Comparison of verification results obtained after introducing CRU and SE-block quantified by average equal error rate.

produces the most continuous ridge patterns for all of the sample inputs. The scenario of high pressure when taking a fingerprint impression is shown in the first row. As a result, valleys are obscured and very thick ridges are produced. The fingerprint image produced by CR-GAN has the highest ridge-valley clarity and performs the best at predicting ridges and valleys. The issue of low ridge clarity caused by uneven pressure is seen in the second row. While many of the most cutting-edge enhancement algorithms produce fictitious ridge features, CR-GAN properly predicts any ridge details that are obscure in the input fingerprint image. In addition, in the third and fourth rows, where there are creases, the ridge features that are missing are more accurately predicted by CR-GAN than by other techniques.

### 5.3 Contrasting CRU with Squeeze and Excitation Block

Now we contrast the suggested CRU's performance to that of the squeeze and excitation (SE) block [55], a cutting-edge channel attention model. Inside CR-GAN's design, the SE block is used to substitute the suggested CRU, and the resulting model is named SE-GAN. Table 7 shows a comparison of parameter counts, and we find that CR-GAN has fewer parameters than SE-GAN. This observation demonstrates that the SE-block introduces more parameters into the architecture of a fingerprint preprocessing model than the suggested CRU.
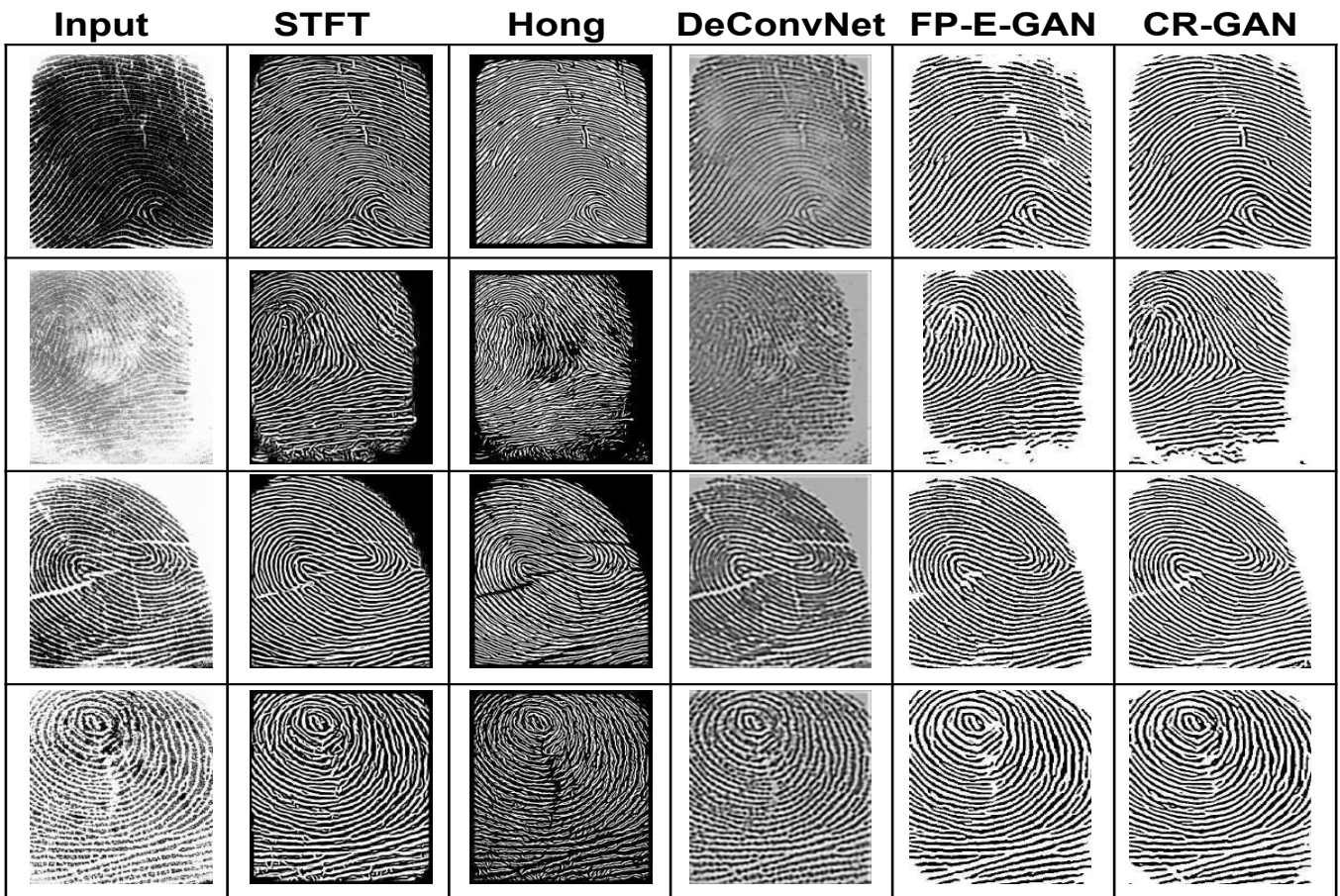
Fig. 10: Examples presenting effective fingerprint enhancement obtained through CR-GAN and comparisons with cutting-edge algorithms.
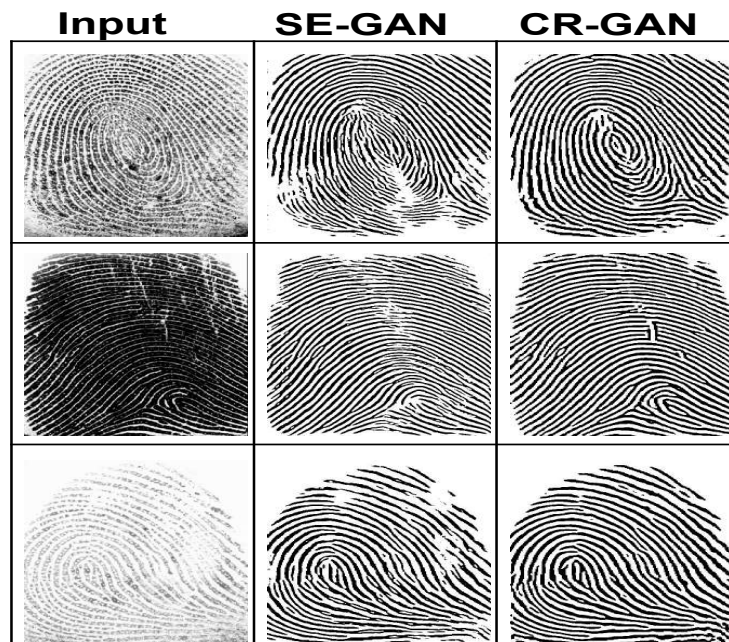
Fig. 11: Comparison of enhanced fingerprints generated from the fingerprint enhancement model with CRU (CR-GAN) and the fingerprint enhancement model with SE block (SE-GAN).
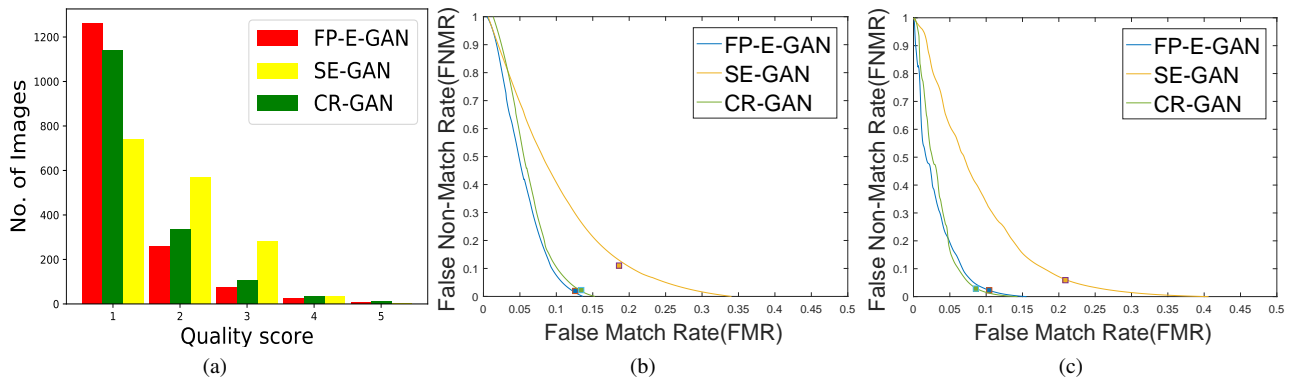
Fig. 12: (a) Histogram presenting the distribution of quality scores of fingerprints, DET curves produced using (b) Bozorth, and (c) MCC. Comparisons between CR-GAN, FP-E-GAN, and SE-GAN are provided on [92].
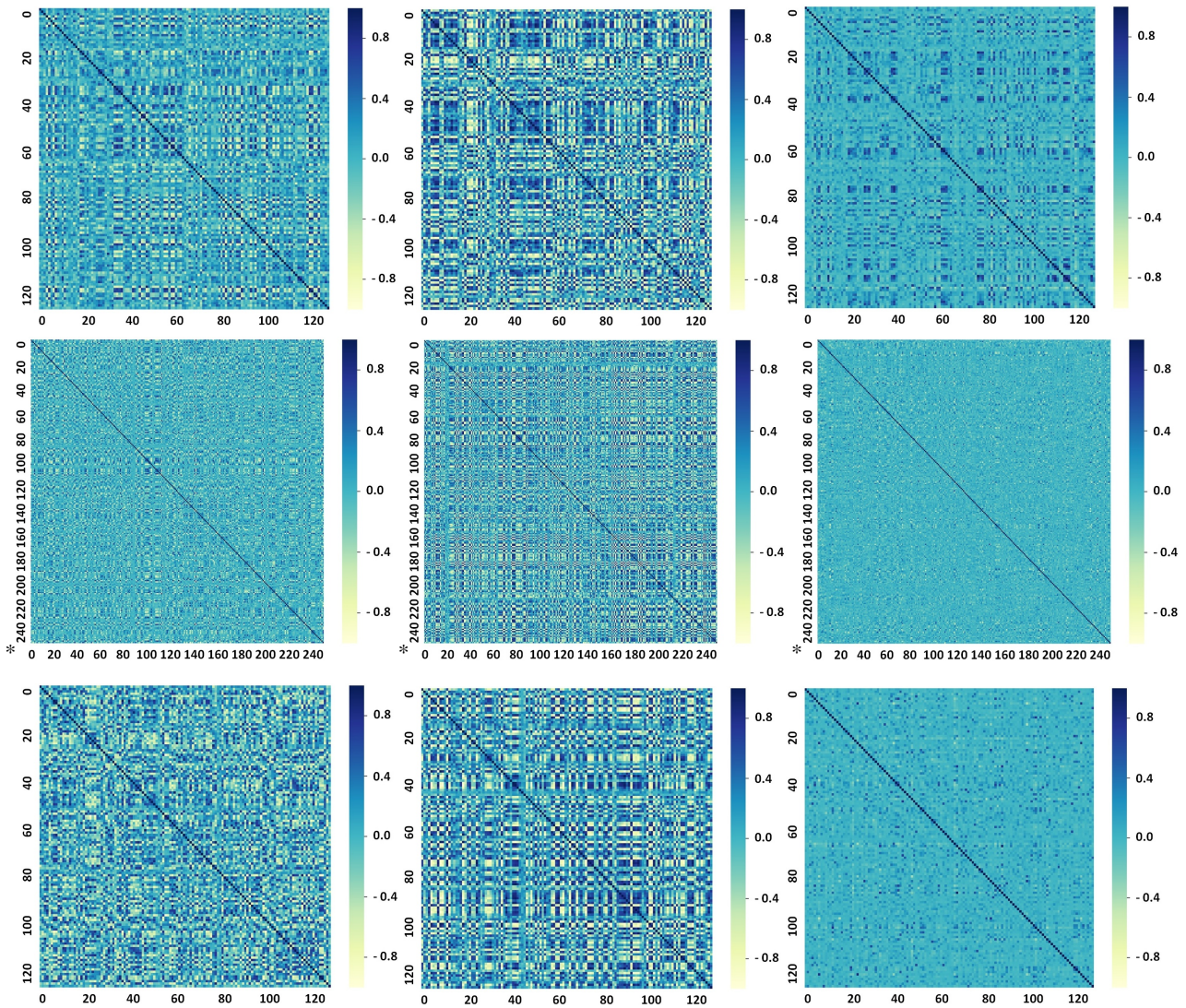
Fig. 13: FP-E-GAN, SE-GAN, and CR-GAN channel weights correlation matrices (from left to right). The first, second, and third rows in all three columns correspond to layers 3, 16, and 21, respectively of the corresponding fingerprint enhancement model.

Enhancement Algorithm	Quality Score
DeconvNet [23]	1.95
CR-DeConvNet	*1.78*
Unet [101]	1.45
Att-Unet [102]	1.51
CR-Unet	*1.45*
CR-GAN	**1.42**

Table 10 An analysis of average quality scores (computed through NFIQ) achieved by cutting-edge deep models.

Enhancement Algorithm	Bozorth	MCC
Unet [101]	11.35	10.58
DeConvNet [23]	10.93	10.86
Att-Unet [102]	9.50	9.08
FP-E-GAN [28]	7.30	5.96
CR-DeConvNet	*6.53*	*5.45*
CR-Unet	*5.99*	*5.55*
CR-GAN	**5.72**	**4.45**

Table 11 An analysis of verification results obtained by several cutting-edge deep models on [92]. Performance is quantified by average equal error rate.

Quality scores of the enhanced fingerprints generated by CR-GAN and SE-GAN are compared in Table 8. The matching distribution of NFIQ scores represented through a histogram is shown in Figure 12 (a). Results reveal that when compared to CR-GAN, SE-GAN generates inferior quality enhanced fingerprints. The average EER is presented in Table 9 in order to compare matching performance. In Figure 12 (b) and Figure 12 (c), the relevant DET curves are provided. In comparison to SE-GAN, the verification error for fingerprints enhanced using CR-GAN is much lower for both Bozorth and MCC matchers. Due to the inferior verification results of SE-GAN in comparison to the FP-E-GAN baseline, these findings demonstrate that SE-block is inappropriate for fingerprint enhancement. The performance of CR-GAN, on the other hand, with the suggested CRU, surpasses that of SE-GAN and FP-E-GAN, showing that it is a good candidate to introduce channel attention into a preprocessing model enhancing fingerprints. These findings further show that the CR-GAN's enhanced performance cannot be only explained by an increase in model capacity as a result of CR-GAN's higher parameter space than FP-E-GAN. SE-GAN performs substantially worse than FP-E-GAN despite having larger parameter space than CR-GAN. In Figure 13, we present the matrices of correlations between different layers' learnt channel weights. We study matrices of correlation pertaining to different enhancement models: FP-E-GAN, SE-GAN, and CR-GAN, to examine the impact of introduction of their respective channel attention on the features the model learns in contrast to the features learnt by the backbone model (low correlation values signify less redundant features).

In theory, a channel level attention mechanism is introduced to lessen redundancy in the backbone network's channel weights. In contrast to the FP-E-GAN backbone, we discover that SE-GAN exhibits stronger channel correlation rather than decreased correlation among channel weights. This demonstrates how the SE-block introduces redundant features. Conversely, the correlation values for CR-GAN are the lowest. This shows that the suggested CRU lowers feature redundancy and aids in learning robust features. CR-GAN outperforms both FP-E-GAN and SE-GAN due to its learning of more robust features than both these enhancement models. The representative cases presenting the recovered fingerprint images produced using the proposed CR-GAN and SE-GAN are provided in Figure 11. CR-GAN performs better than SE-GAN in every situation. CR-GAN outputs fingerprints with greater clarity between ridges and valleys, smoother ridges, and less erroneous ridge details, when contrasted to SE-GAN.

5.4 Application of CRU to Various Deep Models for Fingerprint Enhancement

As of right now, we can see that the suggested CRU enhances the reconstruction ability of a fingerprint enhancement model leveraging a GAN framework at its core. We then examine the proposed CRU's generalizability for various network architecture options for fingerprint enhancement. As the foundational network designs for our experiment, we use Unet [101] and DeConvNet [23], a autoencoder model for enhancing low quality fingerprints. We create new architectures on both of these designs named *CR-Unet* and *CR-DeConvNet* by adding CRU after each convolution block, like done for CR-GAN. To see if the suggested CRU improves the performance of these networks, we contrast the verification error obtained using CR-Unet with that of Unet and likewise of CR-DeConvNet to

DeConvNet. Additionally, we contrast CR-Unet's verification error to that of the Attention Unet (Att-Unet), which is based on spatial attention. Baseline Unet and Att-Unet implementation is taken from the source provided in [3].

First, we evaluate the enhanced fingerprint images produced by various models over fingerprint quality ratings. The histogram presenting the distribution of quality scores of fingerprints is presented in Figure 15 (a) while Table 10 reports the average NFIQ score. The results back up the assertion that the proposed CRU enhances performance. When reconstructing fingerprints, CR-DeConvNet outperforms DeConvNet, in terms of quality scores. Similar observation holds for CR-Unet and Unet. Table 11 reports the average verification error obtained by all cutting-edge deep models, while Figure 15 (b) and Figure 15 (c) illustrate the related DET curves. Average EER is dramatically decreased for both fingerprint matchers. These findings show how cutting-edge deep models trained to learn fingerprint enhancement may learn redundant features. The suggested CRU helps learn robust features by reducing feature redundancy. Subsequently, the enhancement performance of CR-DeConvNet turns out significantly better than DeConvNet. Likewise results are reported for CR-Unet and Unet. The claim that the suggested CRU generalizes to several cutting-edge deep models is verified by all of these results.

Additionally, we see that CR-Unet performs noticeably better than Att-Unet [102]. While CR-Unet makes use of channel-level attention, Att-Unet makes use of spatial attention. Because CR-Unet surpasses Att-Unet both in terms of matching performance and fingerprint quality score, we conclude that spatial attention is less useful for fingerprint enhancement than channel attention. It is also interesting to note that both the deep models: CR-DeConvNet and CR-Unet, that were suggested in this subsection perform better than FP-E-GAN. However, FP-E-GAN performs much better than Unet and DeConvNet at baseline compared to those two networks. The suggested CR-GAN hence performs better than CR-Unet and CR-DeConvNet. We contrast the sample reconstructed fingerprints produced by Unet, Att-Unet, and CR-Unet in Figure 14 (a). We discover that CR-Unet's produced fingerprints have superior clarity between ridges and valleys and fewer false ridge characteristics than those created by Unet and Att-Unet. When CR-DeConvNet and DeConvNet are compared, similar findings are found (see Figure 14 (b)).

6 Contrasting CR-GAN with Cutting-edge Fingerprint Enhancement Methods Leveraging GAN Framework

This section contrasts the Cycle-GAN [103] and DU-GAN [31], two cutting-edge generative adversarial network-based fingerprint enhancement models, in terms of performance with the proposed CR-GAN. Figure 16 (a) and Table 12 compare the average fingerprint quality scores obtained on enhanced fingerprints obtained by Cycle-GAN, DU-GAN, and the proposed CR-GAN. In comparison to Cycle-GAN and DU-GAN, we discover that the fingerprint quality of photos produced by CR-GAN is substantially higher. The average EER derived by each model is shown in Figure 16 (b) and (c) and Table 13. We discover that Cycle-GAN performs poorly
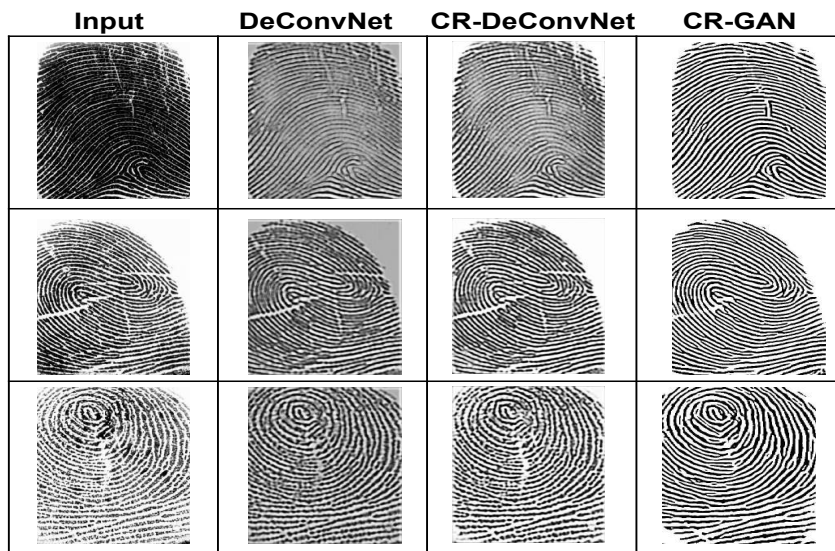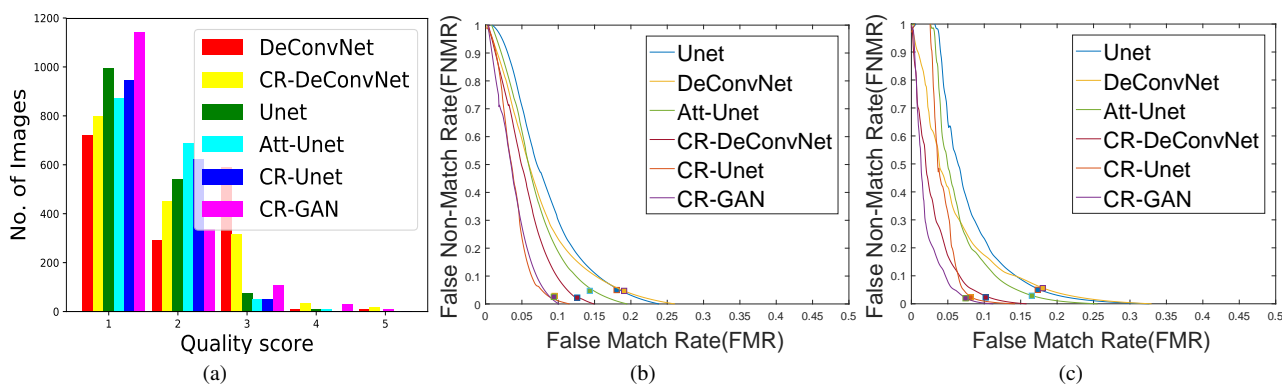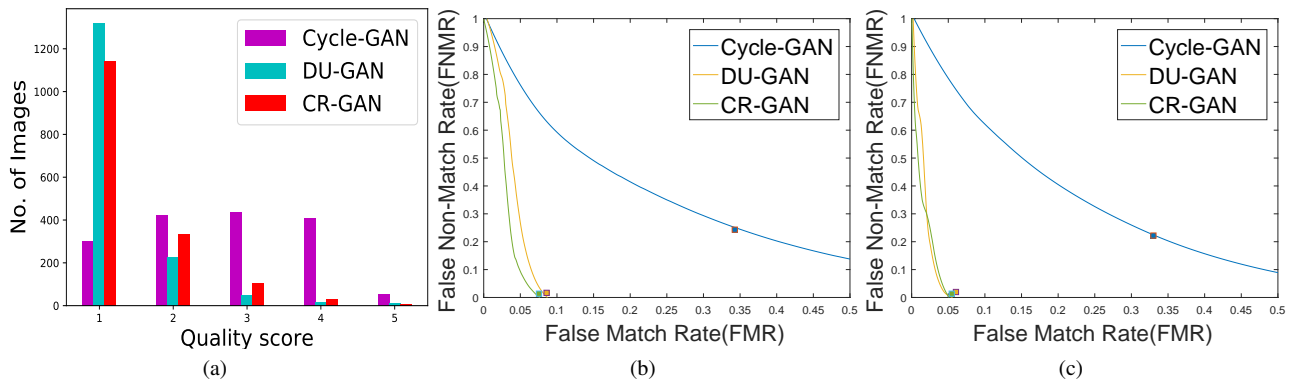
(a)
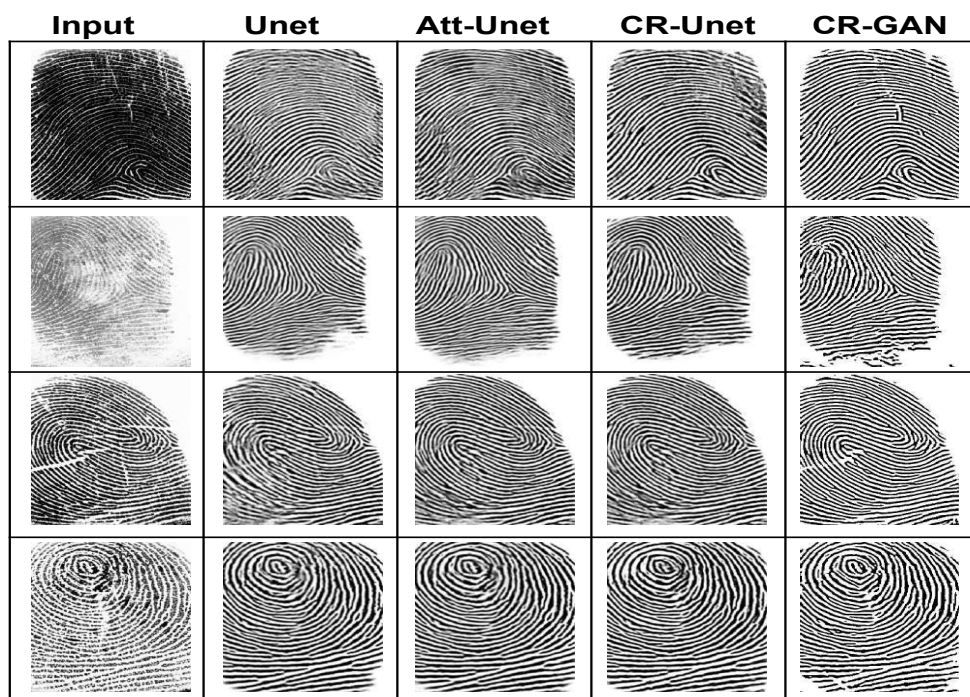
(b)

Fig. 14: Examples showing the suggested CRU's generalizability on (a) DeConvNet and (b) Unet designs, while contrasting with the suggested CR-GAN.

Fig. 15: Evaluation of the suggested CRU's generalizability over cutting-edge deep architectures: (a) histogram presenting the distribution of quality scores of fingerprints; DET curves generated using (b) Bozorth; (c) MCC.

Fig. 16: On [92], the suggested CR-GAN is compared against recent fingerprint enhancement models that employ a GAN framework: Cycle-GAN and DU-GAN: (a) Histogram presenting the distribution of quality scores of fingerprints; DET curves derived using (b) Bozorth (c) MCC.

Input	Ground Truth	Discriminator	Generator	Both
		Dice: 96.87 Jaccard: 94.60 SSIM: 85.16 PSNR: 12.63	Dice:97.99 Jaccard: 96.55 SSIM: 89.57 PSNR: 14.61	Dice:98.38 Jaccard: 97.30 SSIM: 90.96 PSNR: 15.50
		Dice:95.05 Jaccard: 90.61 SSIM: 75.52 PSNR: 12.44	Dice:95.95 Jaccard: 92.28 SSIM: 80.04 PSNR: 13.43	Dice:96.08 Jaccard: 92.51 SSIM: 81.06 PSNR: 13.55
		Dice:97.63 Jaccard: 95.92 SSIM: 88.08 PSNR: 13.85	Dice:98.33 Jaccard: 97.19 SSIM: 91.22 PSNR: 15.42	Dice:98.39 Jaccard: 97.28 SSIM: 91.51 PSNR: 15.58

Fig. 17: Example scenarios that show the ridge preservation capabilities of the CR-GAN and its variations (analyzed during ablation) to keep the ridge details intact.
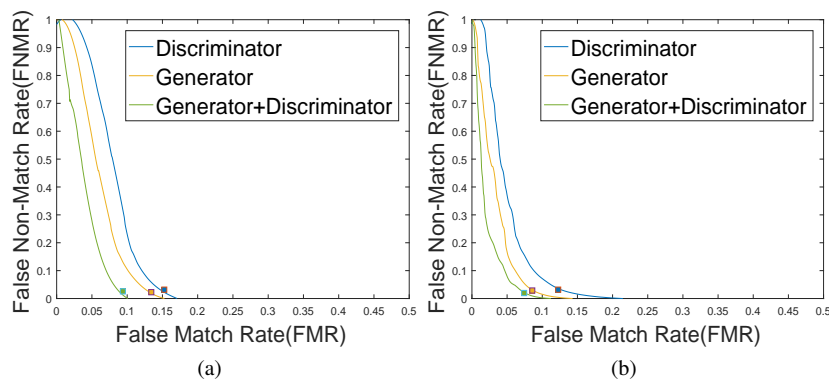
Fig. 18: DET curves utilizing (a) Bozorth and (b) MCC to illustrate the importance of the suggested channel refinement.

Enhancement Algorithm	Quality Score
Cycle-GAN [103]	1.76
DU-GAN [31]	**1.26**
CR-GAN	1.42

Table 12 An analysis of average quality scores (computed through NFIQ) achieved by several enhancement models leveraging GAN framework.

Enhancement Algorithm	Bozorth	MCC
Cycle-GAN [103]	29.52	27.96
DU-GAN [31]	7.13	5.13
CR-GAN	**5.72**	**4.45**

Table 13 An analysis of verification results obtained on [92] by several models leveraging GAN framework for enhancement. Performance is quantified by average equal error rate.

Refinement	Bozorth	MCC
Discriminator	7.68	5.81
Generator	6.79	4.73
Both	**5.72**	**4.45**

Table 14 An analysis of verification results obtained on [92] by GAN variants curated by introducing the suggested CRU into its different sub-networks.

Model	Total
RUnet	3104178
CR-RUnet	3110586
SE-RUnet	3643599

Table 15 Comparison of the CRU's introduced model parameters with model parameters introduced by SE-block [55].

Database	Jaccard Similarity (↑)		Dice Score (↑)	
	RUnet	CR-RUnet	RUnet	CR-RUnet
2000DB1	**88.15**	86.97	**93.34**	92.71
2000DB2	**86.40**	84.87	**92.39**	91.55
2000DB3	**93.74**	93.04	**96.50**	96.16
2000DB4	**94.28**	88.68	**97.04**	93.94
2002DB1	**96.95**	96.65	**98.44**	98.29
2002DB2	**94.88**	93.93	**97.28**	96.73
2002DB3	91.83	**92.75**	95.53	**96.11**
2002DB4	**91.17**	88.67	**95.32**	93.93
2004DB1	**98.78**	98.64	**99.38**	99.31
2004DB2	93.94	**95.46**	96.69	**97.65**
2004DB3	94.62	**94.90**	97.17	**97.35**
2004DB4	94.73	**95.17**	97.21	**97.48**

Table 16 Fingerprint ROI segmentation performance degrades after incorporating CRU into RUnet architecture.

because it is unable to maintain ridge structure throughout improvement. In contrast to the other two enhancement models leveraging GAN framework, suggested CR-GAN achieves the most satisfactory enhanced fingerprints, as evidenced by the lowest verification error rate.

7 Impact of CRU on Fingerprint Enhancement

7.1 Ridge Structure Preservation

A poor input fingerprint may be accompanied by many types of noise patterns. In order to measure the suggested CR-GAN's capacity to preserve ridges over a variety of noise patterns observed in real poor quality fingerprints, we compute SSIM similarity scores between the ground truth binarized image and the output corresponding to the input. A high similarity score means the model keeps ridge details, i.e., while enhancing them, the fingerprint picture input preserves the type of fingerprint pattern, the direction of the ridges, and minutiae details. Sample degraded fingerprints are shown in the first column from the left in Figure 17 along with enhanced pictures produced by CR-GAN (rightmost column). The related ground truth binarized fingerprint images are shown in the second column. High similarity scores between CR-GAN's output and ground truth are attained, proving that the suggested method preserves the input fingerprint's ridge structure while enhancing it.

7.2 Ablation Study

In order to more precisely measure the effects on the generator and discriminator networks by incorporating the suggested CRU, we analyze CR-GAN's enhancement results on different variants of CR-GAN. When using the suggested CRU, we investigate three alternative variants: using just the generator, just the discriminator, and using both the generator and the discriminator. Table 14 reports the verification error rate for all three versions, and Figure 18(a) and Figure 18 (b) illustrate the related DET curves. Figure 17 and 19 display sample reconstructions produced by each of the three variations.

7.3 Successful Scenarios

Few samples on which the suggested CR-GAN achieves satisfactory enhancement results are illustrated in Figure 20. The two left-most columns are latent fingerprints that are acquired using dusting with a chemical powder which sometimes lead to non-uniform amount of powder at different fingerprint regions. As a result, latent fingerprints may possess unclear ridge structure. However, interestingly, with decent ridge-valley clarity, CR-GAN successfully reconstructs fingerprints for both the samples. The scenario of indistinct valleys caused by thick ridges arising due to moist finger or excessive pressure is depicted in the third column. Once more, the suggested CR-GAN properly predicts details about ridges and valleys and produces a fingerprint with significantly improved ridge information. The lost ridge details owing to warts or creases is shown in the the two rightmost columns. CR-GAN accurately approximates the otherwise missing ridge features at fingerprint image pixels with creases and cuts.

7.4 Challenging Scenarios

The effectiveness of CR-GAN on few difficult instances is shown in Figure 21. CR-GAN produces erroneous ridge patterns close to the distorted region, in the top row. The input fingerprint's ridges in the middle row, close to the area of excessive pressure, are either invisible or very dark. Subsequently, CR-GAN produces ridges that are not smooth or erroneous ridge features. Nonetheless, CR-GAN consistently beats the backbone FP-E-GAN (in addition to several cutting-edge models for fingerprint enhancement). This confirms the idea that refining channel weights enhances model performance in terms of both quality and subsequent match scores.

7.5 Fingerprint ROI Segmentation

Lastly, we study the generalization ability of the proposed CRU on fingerprint preprocessing tasks by evaluating the fingerprint ROI segmentation task. For this experiment, we take RUnet [52], a state-of-art deep model for fingerprint preprocessing [30, 104–106] as the backbone network. Subsequently, we introduce CRU into RUnet and design *CR-RUnet*. Later, CR-RUnet is trained in a fully supervised manner to learn to segment foreground and background fingerprint regions. As any channel level attention model increases the model parameters (see Table 15), which under limited availability of training data, can deteriorate the model performance. Table 16 compares the fingerprint ROI segmentation performance with (CR-RUnet) and without (RUnet) the introduction of CRU. We find that fingerprint ROI segmentation ability of the model after introducing CRU is competitive to RUnet, with RUnet performing better on the majority
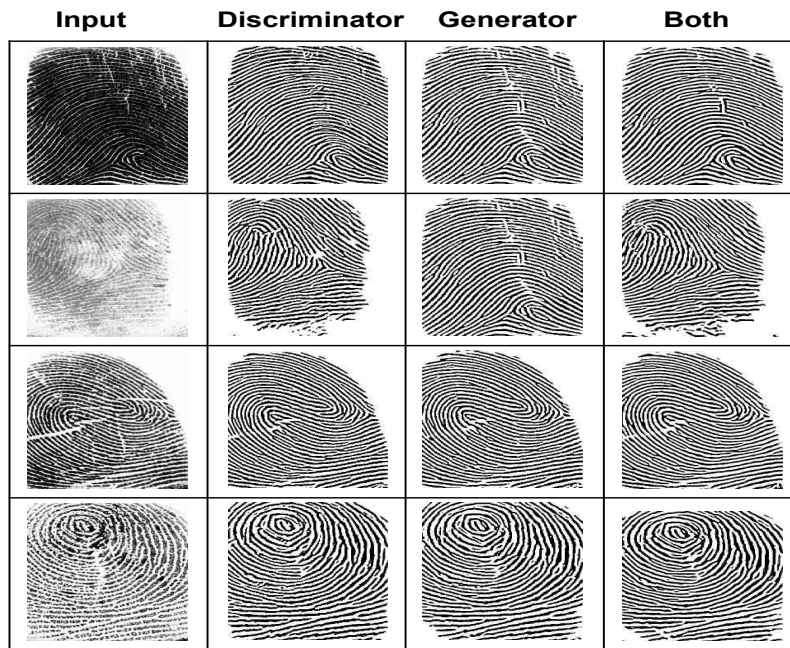
Fig. 19: Examples from [92] showing the effect of introducing the suggested CRU into FP-E-GAN's [28] different sub-networks.
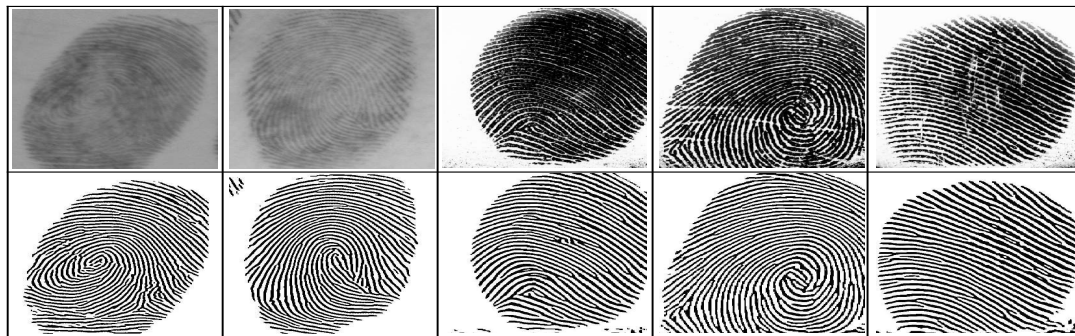
Fig. 20: Examples of effective enhancement results obtained using the suggested CR-GAN.
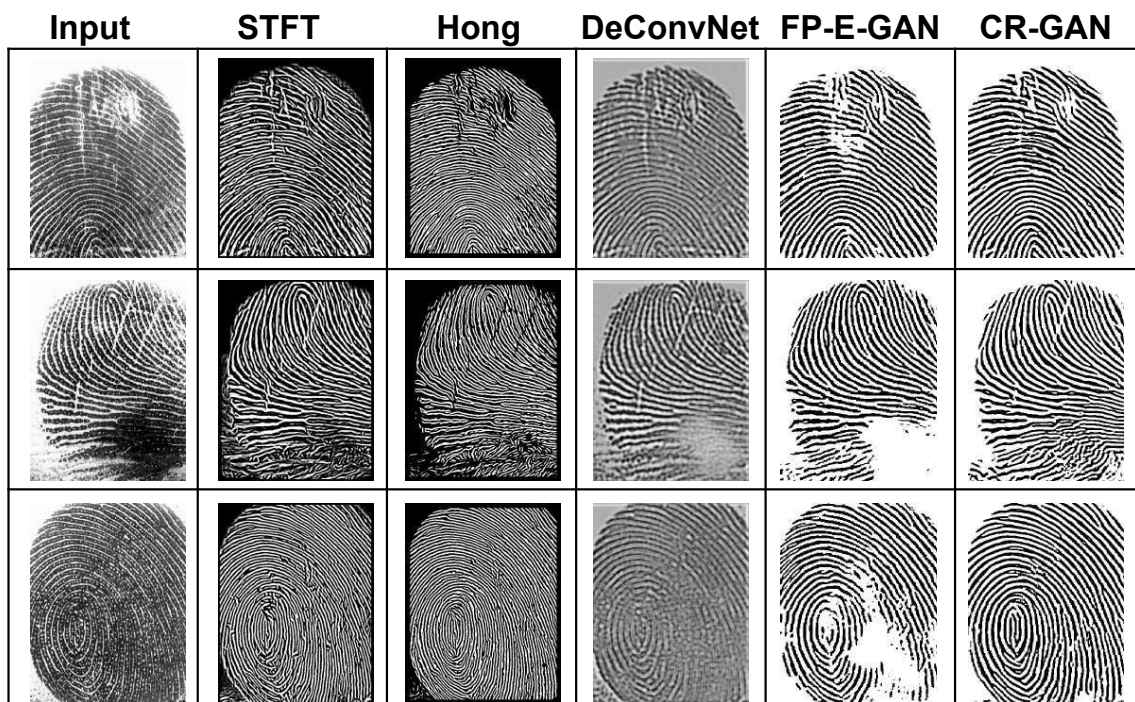
Fig. 21: Examples illustrating difficult instances and contrasting cutting-edge algorithms for fingerprint enhancement.

Database	Jaccard Similarity (↑)		Dice Score (↑)	
	CR-RUnet	SE-RUnet	CR-RUnet	SE-RUnet
2000DB1	**86.97**	49.66	**92.71**	34.82
2000DB2	**84.87**	61.69	**91.55**	49.08
2000DB3	**93.04**	86.81	**96.16**	78.47
2000DB4	**88.68**	86.65	**93.94**	78.44
2002DB1	**96.65**	94.66	**98.29**	90.41
2002DB2	**93.93**	69.22	**96.73**	57.73
2002DB3	**92.75**	60.07	**96.11**	43.91
2002DB4	**88.67**	66.23	**93.93**	51.38
2004DB1	**98.64**	96.72	**99.31**	93.94
2004DB2	**95.46**	84.85	**97.65**	75.94
2004DB3	**94.90**	84.93	**97.35**	76.95
2004DB4	**95.17**	61.58	**97.48**	46.79

Table 17 Fingerprint ROI segmentation performance degrades after incorporating both CRU and SE block into RUnet architecture. However, CRU performs significantly better than SE block.

of datasets. We hypothesize that better ROI segmentation performance of RUnet in contrast to CRUnet is observed due to potential overfitting by CRUnet as a result of more model parameters. However, as the increase in model parameters is not so significant, the performance doesn't drop significantly.

To validate the above-mentioned hypothesis, we perform an additional experiment. We compare CRU with the SE block. The fingerprint ROI segmentation model designed after introducing the SE block into RUnet is named *SE-RUnet*. Table 17 compares the fingerprint ROI segmentation performance after the introduction of CRU (CR-RUnet) and SE block (SE-RUnet). The fingerprint ROI segmentation performance of SE-RUnet drops significantly compared to CR-RUnet. These results signify that although the fingerprint ROI segmentation performance does not improve significantly after introducing channel level attention, however, CRU significantly outperforms the cutting-edge channel attention method SE block. These findings support the assertion that the suggested CRU is a better-suited channel-level attention mechanism for fingerprint preprocessing compared to the state-of-the-art.

8 Conclusion

This research presents a channel refinement unit (CRU), a channel-level attention mechanism for deep fingerprint preprocessing models. Extensive experimentation confirms that CRU generalizes on different choices of deep architectures for fingerprint preprocessing. Visualization of correlation matrix indicates that CRU reduces correlation among features, explaining the improved generalization ability. However, as CRU introduces additional parameters into the fingerprint preprocessing models, CRU is better suited for applications with enough training data. In the future, several other applications in the domain, such as presentation attack detection, as well as deep models for fingerprint feature extraction and matching, can benefit from the proposed CRU.

9 References

1 Stojanović, B., Marques, O., Nešković, A. and Puzović, S.: 'Fingerprint ROI Segmentation Based on Deep Learning'. Proc. Telecommunications Forum, 2016. pp. 1 – 4
2 Li, J., Feng, J. and Kuo, C.C.J.: 'Deep Convolutional Neural Network for Latent Fingerprint Enhancement', *Signal Processing: Image Communication*, 2018, **60**, pp. 52 – 63
3 Joshi, I., Utkarsh, A., Singh, P., Dantcheva, A., Dutta.Roy, S. and Kalra, P.K.: 'On Restoration of Degraded Fingerprints', *Multimedia Tools and Applications*, 2022, pp. 1–29
4 Hong, L., Wan, Y. and Jain, A.: 'Fingerprint Image Enhancement: Algorithm and Performance Evaluation', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1998, **20**, (8), pp. 777 – 789
5 Gottschlich, C. and Schönlieb, C.B.: 'Oriented Diffusion Filtering for Enhancing Low-Quality Fingerprint Images', *IET Biometrics*, 2012, **1**, (2), pp. 105 – 113
6 Turroni, F., Cappelli, R. and Maltoni, D.: 'Fingerprint Enhancement using Contextual Iterative Filtering'. Proc. International Conference on Biometrics (ICB), 2012. pp. 152 – 157
7 Ramos, R.C., de Lima.Borges, E.V.C., Andrezza, I.L.P., Primo, J.J.B., Batista, L.V. and Gomes, H.M.: 'Analysis and Improvements of Fingerprint Enhancement from Gabor Iterative Filtering'. SIBGRAPI Conference on Graphics, Patterns and Images, 2018. pp. 266 – 273
8 Wang, W., Li, J., Huang, F. and Feng, H.: 'Design and Implementation of Log-Gabor Filter in Fingerprint Image Enhancement', *Pattern Recognition Letters*, 2008, **29**, (3), pp. 301 – 308
9 Gottschlich, C.: 'Curved-Region-Based Ridge Frequency Estimation and Curved Gabor Filters for Fingerprint Image Enhancement', *IEEE Transactions on Image Processing*, 2011, **21**, (4), pp. 2220 – 2227
10 Chikkerur, S., Cartwright, A.N. and Govindaraju, V.: 'Fingerprint Enhancement using STFT Analysis', *Pattern Recognition*, 2007, **40**, (1), pp. 198 – 211
11 Ghafoor, M., Taj, I.A., Ahmad, W. and Jafri, N.M.: 'Efficient 2-Fold Contextual Filtering Approach for Fingerprint Enhancement', *IET Image Processing*, 2014, **8**, (7), pp. 417 – 425
12 Yoon, S., Feng, J. and Jain, A.K.: 'On Latent Fingerprint Enhancement'. Biometric Technology for Human Identification VII. vol. 7667, 2010. pp. 766707 – 766716
13 Hsieh, C.T., Lai, E. and Wang, Y.C.: 'An Effective Algorithm for Fingerprint Image Enhancement Based on Wavelet Transform', *Pattern Recognition*, 2003, **36**, (2), pp. 303 – 312
14 Jirachaweng, S. and Areekul, V.: 'Fingerprint Enhancement Based on Discrete Cosine Transform'. Proc. International Conference on Biometrics (ICB), 2007. pp. 96 – 105
15 Feng, J., Zhou, J. and Jain, A.K.: 'Orientation Field Estimation for Latent Fingerprint Enhancement', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, **35**, (4), pp. 925 – 940
16 Yang, X., Feng, J. and Zhou, J.: 'Localized Dictionaries Based Orientation Field Estimation for Latent Fingerprints', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014, **36**, (5), pp. 955 – 969
17 Chen, C., Feng, J. and Zhou, J.: 'Multi-Scale Dictionaries Based Fingerprint Orientation Field Estimation'. Proc. International Conference on Biometrics (ICB), 2016. pp. 1 – 8
18 Liu, S., Liu, M. and Yang, Z.: 'Sparse Coding Based Orientation Estimation for Latent Fingerprints', *Pattern Recognition*, 2017, **67**, pp. 164 – 176
19 Chaidee, W., Horapong, K. and Areekul, V.: 'Filter Design Based on Spectral Dictionary for Latent Fingerprint Pre-Enhancement'. Proc. International Conference on Biometrics (ICB), 2018. pp. 23 – 30
20 Cao, K. and Jain, A.K.: 'Latent Orientation Field Estimation via Convolutional Neural Network'. Proc. International Conference on Biometrics (ICB), 2015. pp. 349 – 356
21 Qu, Z., Liu, J., Liu, Y., Guan, Q., Yang, C. and Zhang, Y.: 'Orienet: A Regression System for Latent Fingerprint Orientation Field Extraction'. Proc. International Conference on Artificial Neural Networks, 2018. pp. 436 – 446
22 Sahasrabudhe, M. and Namboodiri, A.M.: 'Fingerprint Enhancement using Unsupervised Hierarchical Feature Learning'. Proc. IAPR- and ACM-sponsored Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP), 2014. pp. 1 – 8
23 Schuch, P., Schulz, S. and Busch, C.: 'De-Convolutional Auto-encoder for Enhancement of Fingerprint Samples'. Proc. International Conference on Image Processing Theory, Tools and Applications (IPTA), 2016. pp. 1 – 7
24 Rama, R.K. and Namboodiri, A.M.: 'Fingerprint Enhancement using Hierarchical Markov Random Fields'. Proc. IEEE International Joint Conference on Biometrics (IJCB), 2011. pp. 1 – 8
25 Svoboda, J., Monti, F. and Bronstein, M.M.: 'Generative Convolutional Networks for Latent Fingerprint Reconstruction'. Proc. IEEE International Joint Conference on Biometrics (IJCB), 2017. pp. 429 – 436
26 Qian, P., Li, A. and Liu, M.: 'Latent Fingerprint Enhancement Based on Dense-UNet'. Proc. International Conference on Biometrics (ICB), 2019. pp. 1 – 6
27 Wong, W.J. and Lai, S.H.: 'Multi-Task CNN for Restoring Corrupted Fingerprint Images', *Pattern Recognition*, 2020, **101**, pp. 107203 – 107213
28 Joshi, I., Anand, A., Vatsa, M., Singh, R., Dutta.Roy, S. and Kalra, P.: 'Latent Fingerprint Enhancement using Generative Adversarial Networks'. IEEE Winter Conference on Applications of Computer Vision (WACV), 2019. pp. 895 – 903
29 Joshi, I., Anand, A., Dutta.Roy, S. and Kalra, P.K. 'On Training Generative Adversarial Network for Enhancement of Latent Fingerprints'. In: AI and Deep Learning in Biometric Security. (, 2021. pp. 51 – 79
30 Joshi, I., Kothari, R., Utkarsh, A., Kurmi, V.K., Dantcheva, A., Dutta.Roy, S., et al.: 'Explainable Fingerprint ROI Segmentation using Monte Carlo Dropout'. IEEE Winter Conference on Applications of Computer Vision Workshops (WACVW), 2021. pp. 60 – 69
31 Joshi, I., Utkarsh, A., Kothari, R., Kurmi, V.K., Dantcheva, A., Dutta.Roy, S., et al.: 'Data Uncertainty Guided Noise-Aware Preprocessing of Fingerprints'. International Joint Conference on Neural Networks (IJCNN), 2021. pp. 1 – 8
32 Joshi, I., Dhamija, T., Kumar, R., Dantcheva, A., Roy, S.D. and Kalra, P.K.: 'Cross-Domain Consistent Fingerprint Denoising', *IEEE Sensors Letters*, 2022,
33 Joshi, I., Prakash, T., Jaiswal, B.S., Kumar, R., Dantcheva, A., Dutta.Roy, S., et al.: 'Context-aware Restoration of Noisy Fingerprints', *IEEE Sensors Letters*, 2022,
34 Schuch, P., Schulz, S. and Busch, C.: 'Survey on the Impact of Fingerprint Image Enhancement', *IET Biometrics*, 2017, pp. 102 – 115
35 Hu, C., Yin, J., Zhu, E., Chen, H. and Li, Y.: 'A Composite Fingerprint Segmentation Based on Log-Gabor Filter and Orientation Reliability'. Proc. IEEE International Conference on Image Processing (ICIP), 2010. pp. 3097 – 3100
36 Thai, D.H., Huckemann, S. and Gottschlich, C.: 'Filter Design and Performance Evaluation for Fingerprint Image Segmentation', *PloS One*, 2016, **11**, (5), pp. 1 – 31
37 Thai, D.H. and Gottschlich, C.: 'Global Variational Method for Fingerprint Segmentation by Three-Part Decomposition', *IET Biometrics*, 2016, **5**, (2),
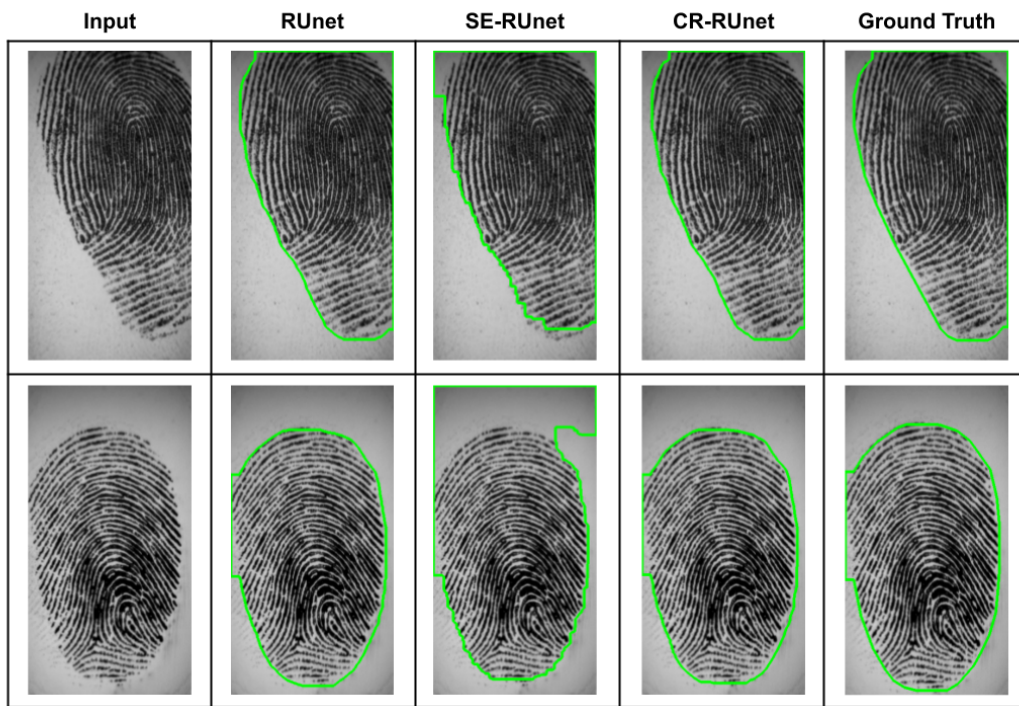
Fig. 22: Fingerprint ROI segmentation performance after incorporating both CRU and SE block into RUnet architecture.

pp. 120 – 130

38 Fahmy, M.F. and Thabet, M.: 'A Fingerprint Segmentation Technique Based on Morphological Processing'. Proc. International Symposium on Signal Processing and Information Technology, 2013. pp. 215 – 220

39 Teixeira, R.F. and Leite, N.J.: 'Unsupervised Fingerprint Segmentation Based on Multiscale Directional Information'. Proc. Iberoamerican Congress on Pattern Recognition, 2011. pp. 38 – 46

40 da Silva.Vasconcelos, R.C. and Pedrini, H.: 'Fingerprint Image Segmentation Based on Oriented Pattern Analysis.'. vol. 11, 2019. pp. 405 – 412

41 Wu, C., Tulyakov, S. and Govindaraju, V.: 'Robust Point-Based Feature Fingerprint Segmentation Algorithm'. Proc. International Conference on Biometrics (ICB), 2007. pp. 1095 – 1103

42 Yang, G., Zhou, G.T., Yin, Y. and Yang, X.: 'K-Means Based Fingerprint Segmentation with Sensor Interoperability', *EURASIP Journal on Advances in Signal Processing*, 2010, **2010**, (1), pp. 1 – 12

43 Ferreira, P.M., Sequeira, A.F. and Rebelo, A.: 'A Fuzzy C-Means Algorithm for Fingerprint Segmentation'. Proc. Iberian Conference on Pattern Recognition and Image Analysis, 2015. pp. 245 – 252

44 Lei, W. and Lin, Y.: 'A Novel Dynamic Fingerprint Segmentation Method Based on Fuzzy C-means and Genetic Algorithm', *IEEE Access*, 2020, **8**, pp. 132694 – 132702

45 Ferreira, P.M., Sequeira, A.F., Cardoso, J.S. and Rebelo, A.: 'Robust Clustering-Based Segmentation Methods for Fingerprint Recognition'. Proc. International Conference of the Biometrics Special Interest Group (BIOSIG), 2018. pp. 1 – 5

46 Liu, E., Zhao, H., Guo, F., Liang, J. and Tian, J.: 'Fingerprint Segmentation Based on An Adaboost Classifier', *Frontiers of Computer Science in China*, 2011, **5**, (2), pp. 148 – 157

47 Zhu, Y., Yin, X., Jia, X. and Hu, J.: 'Latent Fingerprint Segmentation Based on Convolutional Neural Networks'. IEEE Workshop on Information Forensics and Security (WIFS), 2017. pp. 1 – 6

48 Ezeobiejesi, J. and Bhanu, B. 'Latent Fingerprint Image Segmentation using Deep Neural Network'. In: Deep Learning for Biometrics. (Springer, 2017. pp. 83 – 107

49 Serafim, P.B.S., Medeiros, A.G., Rego, P.A., Maia, J.G.R., Trinta, F.A., Maia, M.E., et al.: 'A Method Based on Convolutional Neural Networks for Fingerprint Segmentation'. International Joint Conference on Neural Networks (IJCNN), 2019. pp. 1 – 8

50 Sankaran, A., Jain, A., Vashisth, T., Vatsa, M. and Singh, R.: 'Adaptive Latent Fingerprint Segmentation using Feature Selection and Random Decision Forest Classification', *Information Fusion*, 2017, **34**, pp. 1 – 15

51 Khan, A.I. and Wani, M.A.: 'Patch-Based Segmentation of Latent Fingerprint Images using Convolutional Neural Network', *Applied Artificial Intelligence*, 2019, **33**, (1), pp. 87 – 100

52 Wang, W., Yu, K., Hugonot, J., Fua, P. and Salzmann, M.: 'Recurrent U-Net for Resource-Constrained Segmentation'. Proc. IEEE International Conference on Computer Vision (ICCV), 2019. pp. 2142 – 2151

53 Itti, L., Koch, C. and Niebur, E.: 'A model of saliency-based visual attention for rapid scene analysis', *IEEE Transactions on pattern analysis and machine intelligence*, 1998, **20**, (11), pp. 1254–1259

54 Corbetta, M. and Shulman, G.L.: 'Control of goal-directed and stimulus-driven attention in the brain', *Nature reviews neuroscience*, 2002, **3**, (3), pp. 201–215

55 Hu, J., Shen, L. and Sun, G.: 'Squeeze-and-excitation networks'. Proceedings of the IEEE conference on computer vision and pattern recognition, 2018. pp. 7132–7141

56 Gao, Z., Xie, J., Wang, Q. and Li, P.: 'Global second-order pooling convolutional networks'. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019. pp. 3024–3033

57 Yang, Z., Zhu, L., Wu, Y. and Yang, Y.: 'Gated channel transformation for visual recognition'. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2020. pp. 11794–11803

58 Qilong.Wang, P.Z.P.L.W.Z. Banggu Wu and Hu, Q.: 'Eca-net: Efficient channel attention for deep convolutional neural networks'. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020.

59 Singh, P., Verma, V.K., Mazumder, P., Carin, L. and Rai, P.: 'Calibrating cnns for lifelong learning', *Advances in Neural Information Processing Systems*, 2020, **33**, pp. 15579–15590

60 Singh, P., MAZUMDER, P. and Namboodiri, V.: 'Accuracy booster: Performance boosting using feature map re-calibration'. Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2020.

61 Verma, V.K., Singh, P., Namboodri, V. and Rai, P.: 'A "network pruning network" approach to deep model compression'. Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2020.

62 Mnih, V., Heess, N., Graves, A., et al.: 'Recurrent models of visual attention', *Advances in neural information processing systems*, 2014, **27**

63 Ba, J., Mnih, V. and Kavukcuoglu, K.: 'Multiple object recognition with visual attention', *arXiv preprint arXiv:14127755*, 2014,

64 Xu, K., Ba, J., Kiros, R., Cho, K., Courville, A., Salakhudinov, R., et al.: 'Show, attend and tell: Neural image caption generation with visual attention'. International conference on machine learning, 2015. pp. 2048–2057

65 Hu, J., Shen, L., Albanie, S., Sun, G. and Vedaldi, A.: 'Gather-excite: Exploiting feature context in convolutional neural networks', *Advances in neural information processing systems*, 2018, **31**

66 Singh, P., Mazumder, P., Rai, P. and Namboodiri, V.P.: 'Rectification-based knowledge retention for continual learning'. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021. pp. 15282–15291

67 Singh, P., Mazumder, P. and Namboodiri, V.P.: 'Context extraction module for deep convolutional neural networks', *Pattern Recognition*, 2022, **122**, pp. 108284

68 Li, J., Wang, J., Tian, Q., Gao, W. and Zhang, S.: 'Global-local temporal representations for video person re-identification'. Proceedings of the IEEE/CVF international conference on computer vision, 2019. pp. 3958–3967

69 Liu, Z., Wang, L., Wu, W., Qian, C. and Lu, T.: 'Tam: Temporal adaptive module for video recognition'. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021. pp. 13708–13718

70 Srivastava, R.K., Greff, K. and Schmidhuber, J.: 'Training very deep networks', *Advances in neural information processing systems*, 2015, **28**

71 Li, X., Wang, W., Hu, X. and Yang, J.: 'Selective kernel networks'. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2019. pp. 510–519

72 Yang, B., Bender, G., Le, Q.V. and Ngiam, J.: 'Condconv: Conditionally parameterized convolutions for efficient inference', *Advances in Neural Information Processing Systems*, 2019, **32**

73 Chen, Y., Dai, X., Liu, M., Chen, D., Yuan, L. and Liu, Z.: 'Dynamic convolution: Attention over convolution kernels'. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020. pp. 11030–11039

74 Wang, F., Jiang, M., Qian, C., Yang, S., Li, C., Zhang, H., et al.: 'Residual attention network for image classification'. Proceedings of the IEEE conference on computer vision and pattern recognition, 2017. pp. 3156–3164

75 Woo, S., Park, J., Lee, J.Y. and Kweon, I.S.: 'Cbam: Convolutional block attention module'. Proceedings of the European conference on computer vision (ECCV), 2018. pp. 3–19

76 Park, J., Woo, S., Lee, J.Y. and Kweon, I.S.: 'Bam: Bottleneck attention module', *arXiv preprint arXiv:180706514*, 2018,

77 Roy, A.G., Navab, N. and Wachinger, C.: 'Recalibrating fully convolutional networks with spatial and channel "squeeze and excitation" blocks', *IEEE transactions on medical imaging*, 2018, **38**, (2), pp. 540–549

78 Misra, D., Nalamada, T., Arasanipalai, A.U. and Hou, Q.: 'Rotate to attend: Convolutional triplet attention module'. Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2021. pp. 3139–3148

79 Song, S., Lan, C., Xing, J., Zeng, W. and Liu, J.: 'An end-to-end spatio-temporal attention model for human action recognition from skeleton data'. Proceedings of the AAAI conference on artificial intelligence. vol. 31, 2017.

80 Du, W., Wang, Y. and Qiao, Y.: 'Recurrent spatial-temporal attention network for action recognition in videos', *IEEE Transactions on Image Processing*, 2017, **27**, (3), pp. 1347–1360

81 Fu, Y., Wang, X., Wei, Y. and Huang, T.: 'Sta: Spatial-temporal attention for large-scale video-based person re-identification'. Proceedings of the AAAI conference on artificial intelligence. vol. 33, 2019. pp. 8287–8294

82 Yang, J., Zheng, W.S., Yang, Q., Chen, Y.C. and Tian, Q.: 'Spatial-temporal graph convolutional network for video-based person re-identification'. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2020. pp. 3289–3299

83 Chen, L., Zhang, H., Xiao, J., Nie, L., Shao, J., Liu, W., et al.: 'Sca-cnn: Spatial and channel-wise attention in convolutional networks for image captioning'. Proceedings of the IEEE conference on computer vision and pattern recognition, 2017. pp. 5659–5667

84 Singh, P., Mazumder, P., Karim, M.A. and Namboodiri, V.P.: 'Calibrating feature maps for deep cnns', *Neurocomputing*, 2021, **438**, pp. 235–247

85 Roy, R., Joshi, I., Das, A. and Dantcheva, A. '3D CNN Architectures and Attention Mechanisms for Deepfake Detection'. In: Handbook of Digital Face Manipulation and Detection. (, 2022. pp. 213–234

86 Qin, Z., Zhang, P., Wu, F. and Li, X.: 'Fcanet: Frequency channel attention networks'. Proceedings of the IEEE/CVF international conference on computer vision, 2021. pp. 783–792

87 Zhang, H., Dana, K., Shi, J., Zhang, Z., Wang, X., Tyagi, A., et al.: 'Context encoding for semantic segmentation'. Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, 2018. pp. 7151–7160

88 Sutton, R.S., McAllester, D., Singh, S. and Mansour, Y.: 'Policy gradient methods for reinforcement learning with function approximation', *Advances in neural information processing systems*, 1999, **12**

89 Hochreiter, S. and Schmidhuber, J.: 'Long short-term memory', *Neural computation*, 1997, **9**, (8), pp. 1735–1780

90 Spillmann, L., Dresp.Langley, B. and Tseng, C.H.: 'Beyond the classical receptive field: The effect of contextual stimuli', *Journal of Vision*, 2015, **15**, (9), pp. 7–7

91 Joshi, I., Grimmer, M., Rathgeb, C., Busch, C., Bremond, F. and Dantcheva, A.: 'Synthetic Data in Human Analysis: A Survey', *arXiv preprint arXiv:220809191*, 2022,

92 Puri, C., Narang, K., Tiwari, A., Vatsa, M. and Singh, R.: 'On Analysis of Rural and Urban Indian Fingerprint Images'. Proc. International Conference on Ethics and Policy of Biometrics, 2010. pp. 55 – 61

93 Sankaran, A., Vatsa, M. and Singh, R.: 'Multisensor Optical and Latent Fingerprint Database', *IEEE Access*, 2015, **3**, pp. 653 – 665

94 Tiwari, K. and Gupta, P.: 'Fingerprint Quality of Rural Population and Impact of Multiple Scanners on Recognition'. Proc. Chinese Conference on Biometric Recognition, 2014. pp. 199 – 207

95 Maltoni, D., Maio, D., Jain, A.K. and Prabhakar, S.: 'Handbook of Fingerprint Recognition'. (, 2009)

96 Wang, Z., Bovik, A.C., Sheikh, H.R. and Simoncelli, E.P.: 'Image Quality Assessment: From Error Visibility to Structural Similarity', *IEEE Transactions on Image Processing*, 2004, **13**

97 Choi, S.S., Cha, S.H. and Tappert, C.C.: 'A Survey of Binary Similarity and Distance Measures', *Journal of Systemics, Cybernetics and Informatics*, 2010, **8**, (1), pp. 43 – 48

98 Ndajah, P., Kikuchi, H., Yukawa, M., Watanabe, H. and Muramatsu, S.: 'An Investigation on the Quality of Denoised Images', *International Journal of Circuit, Systems, and Signal Processing*, 2011, **5**, (4), pp. 423 – 434

99   Dice, L.R.: 'Measures of the Amount of Ecologic Association Between Species', *Ecology*, 1945, **26**, (3), pp. 297 – 302

100 NIST. 'NBIS- NIST Biometric Image Software'. (, . http://biometrics.idealtest.org/

101 Ronneberger, O., Fischer, P. and Brox, T.: 'U-Net: Convolutional Networks for Biomedical Image Segmentation'. Proc. International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2015. pp. 234 – 241

102 Oktay, O., Schlemper, J., Folgoc, L.L., Lee, M., Heinrich, M., Misawa, K., et al.: 'Attention U-Net: Learning Where to Look for the Pancreas', *arXiv preprint arXiv:180403999*, 2018,

103 Karabulut, D., Tertychnyi, P., Arslan, H.S., Ozcinar, C., Nasrollahi, K., Valls, J., et al.: 'Cycle-Consistent Generative Adversarial Neural Networks based Low Quality Fingerprint Enhancement', *Multimedia Tools and Applications*, 2020, **79**, (25), pp. 18569 – 18589

104 Joshi, I., Utkarsh, A., Kothari, R., Kurmi, V.K., Dantcheva, A., Dutta.Roy, S., et al.: 'Sensor-Invariant Fingerprint ROI Segmentation using Recurrent Adversarial Learning'. International Joint Conference on Neural Networks (IJCNN), 2021. pp. 1 – 8

105 Joshi, I. 'Advanced Deep Learning Techniques for Fingerprint Preprocessing'. IIT Delhi, 2021

106 Joshi, I., Utkarsh, A., Kothari, R., Kurmi, V.K., , Dantcheva, A., et al. 'On Estimating Uncertainty of Fingerprint Enhancement Models'. In: Digital Image Enhancement and Reconstruction. (, 2022.