On Learning Sensor-invariant Features For Fingerprint ROI Segmentation

Indu Joshi
IIT Delhi, India
indu.joshi@cse.iitd.ac.in

Ayush Utkarsh⋆
Independent Researcher, India
ayushutkarsh@gmail.com

Riya Kothari⋆
USC, USA
rskothar@usc.edu

Vinod K Kurmi
IIT Kanpur, India
vinodkk@iitk.ac.in

Antitza Dantcheva
Inria Sophia Antipolis, France
antitza.dantcheva@inria.fr

Sumantra Dutta Roy
IIT Delhi, India
sumantra@ee.iitd.ac.in

Prem Kumar Kalra
IIT Delhi, India
pkalra@cse.iitd.ac.in

Abstract—A fingerprint region of interest (roi) segmentation algorithm is designed to separate the foreground fingerprint from the background noise. All the learning based state-of-the-art fingerprint ROI segmentation algorithms proposed in the literature are benchmarked on scenarios when both training and testing databases consist of fingerprint images acquired from the same sensor. However, when testing is conducted on a different sensor, the segmentation performance obtained is often unsatisfactory. As a result, every time a new fingerprint sensor is used for testing, the fingerprint roi segmentation model needs to be re-trained with the fingerprint image acquired from the new sensor and its corresponding manually marked ROI. Manually marking fingerprint ROI is expensive because firstly, it is time consuming and more importantly, requires domain expertise. In order to save the human effort in generating annotations required by state-of-the-art, we propose a fingerprint roi segmentation model which aligns the features of fingerprint images derived from the unseen sensor such that they are similar to the once obtained from the fingerprints whose ground truth roi masks are available for training. Specifically, we propose a recurrent adversarial learning based feature alignment network that helps the fingerprint roi segmentation model to learn sensor-invariant features. Consequently, sensor-invariant features learnt by the proposed roi segmentation model help it to achieve improved segmentation performance on fingerprints acquired from the new sensor. Experiments on publicly available FVC databases demonstrate the efficacy of the proposed work.

Index Terms—Fingerprints, Biometrics, Deep Convolutional Neural Networks

I. INTRODUCTION

Biometric-based authentication systems are used for a gamut of applications such as law enforcement, border security, surveillance, etc. Among all the different biometric modalities which may require expensive sensing device or may not work reliably in uncontrolled settings, fingerprint is one of the most widely used modality. One of the key component which attributes to the robustness of a fingerprint matching system is the fingerprint region of interest (roi) segmentation module. A fingerprint roi segmentation module is dedicated towards segmenting foreground fingerprint region with clear ridge patterns from the background noise. Noise in a fingerprint can originate due to presence of oil, grease
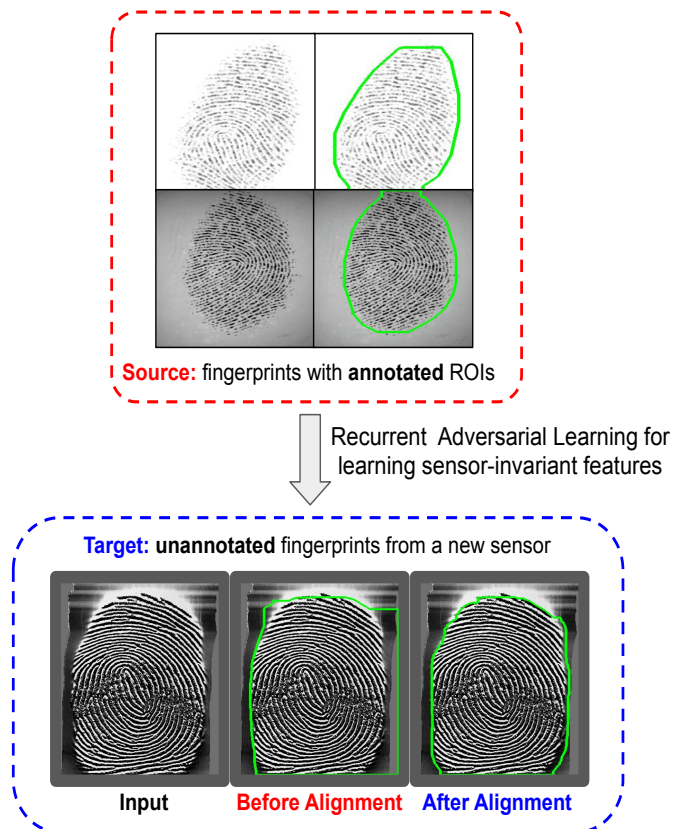

Fig. 1. Schematic diagram showcasing the benefit of proposed feature alignment method in improving segmentation performance on a sensor whose ground truth annotations are not available for training. For better understanding, roi marked fingerprints are presented instead of binary roi mask.

or dirt on surface of the fingerprint sensor used to acquire the fingerprints. A fingerprint roi segmentation module serves dual purpose. Firstly, it minimizes spurious minutiae (feature) detection which translates to improved matching performance. Secondly, it limits the matching to only foreground which reduces computational time for matching.

With the advent of Convolutional Neural Networks (CNNs) and its success in image processing applications, CNNs are

now the state-of-the-art architectures for fingerprint processing, including roi segmentation. However, a limitation of CNNs is that the model trained on one domain often does not generalize well on other domains. In this research, we study this limitation of CNNs in the context of fingerprint roi segmentation. Specifically, this work is based on the hypothesis that the fingerprint roi segmentation performance is likely to be unsatisfactory on an unseen fingerprint sensor. In such a case, state-of-the-art fingerprint roi segmentation models need to be retrained using pixel-level manually marked roi corresponding to the newly introduced sensor.

Manually marking roi every time a new fingerprint sensor is introduced is impractical. Pixel-level annotation is highly expensive due to the human effort and necessary domain expertise. In order to save the cost and time required for the annotations required by the state-of-the-art, we approach the roi segmentation on a different sensor as an unsupervised domain adaptation problem. We propose a *recurrent adversarial learning* framework to learn sensor-invariant features (see Figure 1) by aligning the features of training images (source) and fingerprint images corresponding to the target fingerprint sensor (without manually annotated roi).

The motivation for recurrent adversarial learning is derived from the human perception system as suggested in [44]. When humans observe a scene, their eye movements facilitate knowledge aggregation for scene understanding and also to refine perception [37]. Similarly, recurrent adversarial learning enables proposed RA-RUnet (Recurrently Aligned RUnet) to iteratively refine features such that sensor-invariant features are learnt.

II. RELATED WORK

A. Fingerprint ROI Segmentation

1) Roi segmentation using classical image processing: Filtering in spatial and Fourier domain are one the most widely approaches for roi segmentation [14], [43]. Morphological operations have also been explored for roi segmentation [7], [42]. Some approaches exploit information originating from fingerprint ridge orientations to segment roi [5], [41].

2) Roi segmentation using machine learning: Initial learning based approaches propose clustering of image pixels to discriminate foreground from background [8], [49]. Handcrafted texture and intensity features have also been explored [23]. Recent fingerprint roi segmentation techniques exploit convolutional neural networks [19], [38], [39].

B. Interoperability Across Fingerprint Sensors

Ross and Jain [32] conducted first ever study on cross-sensor matching performance and concluded poor interoperability. Initial approaches to improve interoperability include compensating for distortion [33] and resolution [13], [17]. Texture features have shown promising performance [3], [25]–[27]. Some approaches have exploited fusion based approaches [1], [2].

To summarize, these approaches primarily align minutiae to be obtain invariance from either difference in resolution or texture across varying sensors. Different from these, this research is concentrated on improving interoperability of fingerprint roi segmentation without requiring manual annotations for the newly introduced sensors.

C. Adversarial Learning

Several studies showcase the success of adversarial learning framework in a variety of applications such as image generation [12], [16], domain adaptation [9], image in-painting [28], [50], incremental learning and fairness leaning. All of these approaches optimize the network with an adversarial discriminator. In the adversarial learning literature, some methods are proposed to improve the discriminator for better learning. In [11], [29], multi discriminators are used to avoid the mode collapse problem faced during adversarial learning. A probabilistic discriminator is also explored for domain adaptation [20], [34]. Similarly, a Unet based discriminator [36] is used for image inpaining. In the fingerprints domain, adversarial learning is succesfully utilized for enhancement [18], [48].

D. Unsupervised Domain Adaptation

The source and target features are generally aligned by minimizing the domain discrepancy using maximum mean discrepancy (MMD) [45], correlation alignment (CORAL) [40], and adversarial metrics [10]. These discrepancy based methods are further extended to tackle the mode collapse problem [30] and reduce negative transfer in domain adaptation. Generative adversarial networks [12] are well explored to generated labeled images from the target domain [22], [35]. Other than these approaches, uncertainty calibration [21], [51] and distributional matching [24] are also applied to adapt the classifier for the target domain.

Recently, affinity transformation based model ASA-Net [52] is proposed to tackle the domain alignment problem in semantic segmentation. This model is based on the affinity relationship between adjacent pixels, termed as affinity space of source and target domain. A comparison between segmentation performance obtained by ASA-Net and proposed RA-RUnet is provided in Section VIII-C.

E. Research Contributions

To the best of our knowledge, this research is the first work in fingerprints domain to exploit adversarial learning for learning sensor-invariant features. Experiments demonstrate the efficacy of proposed work in improving generalization ability of fingerprint roi segmentation on new and unseen sensors without relying on the corresponding manually marked roi mask.

We study the most challenging scenario in which features from synthetic fingerprints are aligned with features of real fingerprints. Real fingerprints used in this study are acquired from a variety of fingerprint sensors employing disparate sensing technology i.e. optical, capacitive, thermal. Visualizations of
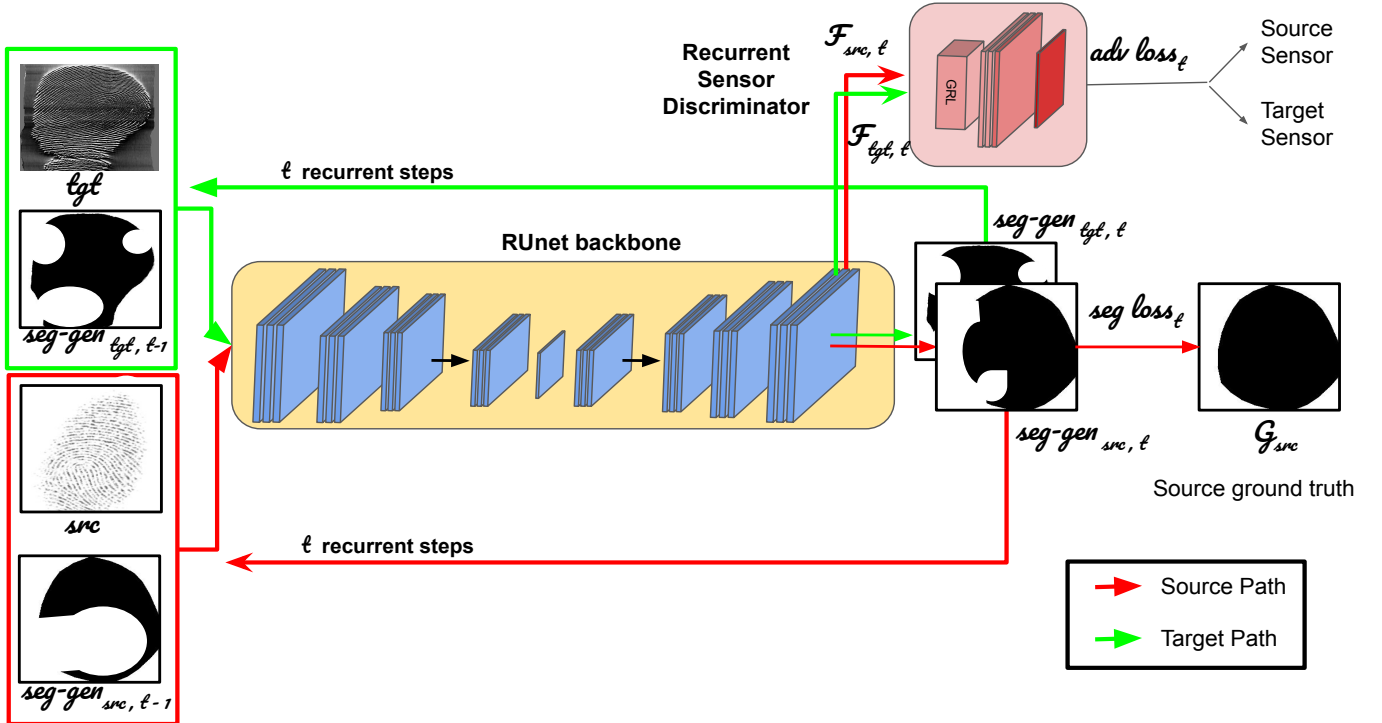
Fig. 2. **Flowchart of proposed RA-RUnet**. RA-RUnet has two sub-networks: RUnet backbone and Recurrent Sensor Discriminator. While RUnet is used for segmentation, the latter is dedicated to feature alignment. For an input fingerprint image,Runet iteratively generates a roi segmentation mask, while the recurrent sensor discriminator classifies the features as derived from source or target sensor. The adversarial loss penalizes the discriminator if it misclassifies, whereas it penalizes RUnet if the discriminator makes a correct classification. This framework helps the backbone to learn sensor-invariant features. Segmentation loss is defined for the images from the source sensor such that the segmentation mask is close to the groundtruth. Please note that annotations of only source sensor are used for training the proposed RA-RUnet.

neural activations obtained from Seg-Grad-Cam [46] are also presented to provide insights on the improved segmentation performance.

III. ALGORITHMS USED FOR BENCHMARKING

To find the most suitable backbone architecture for this problem, we benchmark the following three state-of-the-art semantic segmentation models on fingerprint roi segmentation under the proposed setting.

- **Unet:** Unet [31] is convolutional neural network that relies on an encoder-decoder network design, at its core. Unet is characterized by skip connections between encoding layer and decoding layers. These skip connections supply the required contextual information from neighbouring pixels and help the model to infer edge-level details. Owing to its compactness, Unet is one of the most effective network architecture for image processing applications involving small training datasets.
- **Ccnet:** Ccnet [15] is a convolution neural network with a criss-cross attention unit. Its attention unit is designed to combine contextual information from vertical and horizontal directions with a limited increase in the computational requirements. Moreover, it obtains global information from the image via a recurrent mechanism. As

a result, it learns contextual information from the image and obtains state-of-the-art segmentation performance.

- **Recurrent Unet (RUnet):** RUnet [47] is a variant of Unet. It is especially crafted for applications with small amount of training dataset. Consequently, it is highly suitable for fingerprint roi segmentation due to the small size of training dataset. To maintain compactness, RUnet introduces recurrent units into the baseline Unet architecture. These recurrent units iteratively refine both the internal state of the RUnet and the segmentation mask. This improves the segmentation performance of baseline Unet architecture. Recurrent structure promotes parameters sharing and therefore, helps to avoid overfitting by limiting the number of parameters.

IV. PROPOSED MODEL

As shown in Figure 2, proposed RA-RUnet consists of two networks: a segmentation network and a discriminator network. Given an input fingerprint image, the segmentation network recurrently generates a segmented roi map whereas the discriminator is dedicated to iteratively classify whether the input features correspond to fingerprints from the source or target sensor.

Benchmarking results presented in Section VIII-A reveal that Runet is the most effective baseline network for seg-

mentation. Therefore, proposed RA-RUnet employs RUnet as segmentation network and utilizes it recurrent architecture to improve feature alignment. Our contribution lies in introducing an adversarially trained recurrent sensor discriminator. Recurrent sensor discriminator ($\mathcal{D}$) iteratively classifies whether the recurrently generated feature maps of RUnet backbone network correspond to images from source or target sensor. This in effect, helps RUnet to learn sensor-invariant features.

To formalize, let us assume src and tgt denote input fingerprints acquired using source and target sensors respectively. For an iteration t, fingerprint image (x_{src} or x_{tgt}) is concatenated with the segmentation mask ($seg\_gen_{src,t-1}$ or $seg\_gen_{tgt,t-1}$ respectively) generated at iteration $t-1$ and forwarded through the network. The network is then optimized by a combination of weighted segmentation and adversarial loss. The relationship between $seg\_gen_{tgt,t-1}$ and $seg\_gen_{tgt,t}$ is the same as defined in the baseline RUnet [47] except a modification that now the optimization function also comprises of adversarial loss. Please note that only the ground truth roi mask of source, denoted by g_{src} are used to train the RA-RUnet. The segmentation and adversarial loss are defined as:

A. Segmentation Loss

Segmentation loss is defined as the cross entropy loss between the ground truth roi mask and the segmentation mask generated by RA-RUnet. This loss guides the model to generate segmentation mask close to the ground truth. As this study assumes that ground truth roi masks of target are not available for training, segmentation losss is only defined for the fingerprints originated from source sensor. Assume that G_f is the feature extractor network of backbone architecture. It is parameterised by θ_f. The source and target feature representation are obtained at t stage are:

$$f_{src,t} = G_f(seg\_gen_{src,t-1}, \theta_f) \tag{1}$$

$$f_{tgt,t} = G_f(seg\_gen_{tgt,t-1}, \theta_f) \tag{2}$$

The predicted mask for the source data is obtained by

$$seg\_gen_{src,t} = S_g(f_{src,t-1}, \theta_s) \tag{3}$$

where θ_s denotes the parameters of segmentation network. The segmentation loss is defined as:

$$seg\ loss_t = \frac{1}{n} \sum_{x=1}^{n} \mathcal{L}(seg\_gen_{src,x,t}, g_{src,x}) \tag{4}$$

where $g_{src,x}$ denotes the xth pixel of the ground truth segmentation mask. $gen_{src,x,t}$ denotes the xth pixel of the segmentation mask generated for input fingerprint from source sensor, at iteration t. $\mathcal{L}$ and n denote cross-entropy loss and the total number of pixels respectively.

B. Adversarial Loss

The adversarial loss is defined such that the backbone RUnet network is trained to minimize it while the discriminator ($\mathcal{D}$) is trained to maximize it. This is done by introducing a gradient reversal layer (GRL) [9] before the discriminator. The gradient reversal layer reverses the gradient, only during backpropagation. Due to GRL the backbone Runet is penalized if the features (f_{src} or f_{tgt}) are correctly classified by the discriminator. Due to the GRL based adversarial loss, RUnet learns the necessary sensor-invariant features for obtaining improved segmentation performance on the new target sensor. The discriminator network ($\mathcal{D}$) is trained by minimizing the sensor classification loss and is penalized if it misclassifies the domain of source and target fingerprint. Suppose θ_d are the parameters of discriminator network, the adversarial loss is defined as:

$$adv\ loss_t = \sum_{f \in (f_{src,t} \cup f_{tgt,t})} \mathcal{L}(\mathcal{D}(f, \theta_d), y) \tag{5}$$

where $y = 0$ if $f \in f_{src,t}$ and $y = 1$ if $f \in f_{tgt,t}$.

V. TRAINING AND TESTING

The objective function used to train the proposed RA-RUnet is defined as:

$$\mathcal{L}_{total} = \frac{1}{T} \sum_{t=1}^{T} \frac{1}{n} \sum_{x=1}^{n} [seg\ loss_t - \alpha\ adv\ loss_t] \tag{6}$$

where α denotes the weight of the adversarial loss and T denotes the number of iterations. Please note that the negative sign in the loss function implies the adversarial relationship between the features extractor and the discriminator. We incorporate the gradient reversal layer [10] to implement this in the proposed work. The loss is minimized using the standard backpropagation algorithm. While minimizing the total loss, the segmentation loss is minimized for the feature extractor, while adversarial loss is maximized. For the discriminator, the total loss is maximized; due to negative signs, the effective loss is minimized to train the discriminator. The over all model is optimized using the following objectives:

$$(\hat{\theta}_f, \hat{\theta}_s) = \arg \min_{\theta_f, \theta_s} \mathcal{L}_{total}(\theta_f, \theta_s, \theta_d) \tag{7}$$

$$(\hat{\theta}_d) = \arg \max_{\theta_d} \mathcal{L}_{total}(\theta_f, \theta_s, \theta_d) \tag{8}$$

In this research, we use $\alpha=1$ and $T=3$. RA-RUnet is trained for 5000 iterations. After training, discriminator is discarded and only a forward pass through the segmentation network (baseline RUnet) is required to obtain the segmentation mask. Analysis of hyper-parameter α is performed in Section VIII-E.

VI. DATABASES

All the experiments reported in this study are performed on the publicly available Fingerprint Verification Competition (FVC) databases. Source domain consists of FVC 2000DB4, FVC 2002DB4 and FVC 2004DB4 (subset B defined for

TABLE I
CHARACTERISTICS OF FVC DATABASES.

Database	Sensing Tech.	Domain	Size
2000 DB1	Optical	Target	300×300
2000 DB2	Capacitive	Target	256×364
2000 DB3	Optical	Target	448×478
2000 DB4	Synthetic	Source	240×320
2002 DB1	Optical	Target	388×374
2002 DB2	Optical	Target	296×560
2002 DB3	Capacitive	Target	300×300
2002 DB4	Synthetic	Source	288×384
2004 DB1	Optical	Target	640×480
2004 DB2	Optical	Target	328×364
2004 DB3	Thermal	Target	300×480
2004 DB4	Synthetic	Source	288×384

TABLE II
DICE SCORE OBTAINED BY VARIOUS STATE-OF-THE-ART SEGMENTATION
ALGORITHMS.

Database	Unet (↑)	Ccnet (↑)	RUnet (↑)
2000DB1	42.31	64.51	**71.24**
2000DB2	**76.62**	71.24	68.84
2000DB3	71.41	**91.58**	89.98
2002DB1	**97.79**	89.98	97.19
2002DB2	**92.77**	89.41	92.64
2002DB3	75.42	**91.80**	70.93
2004DB1	**98.86**	95.96	98.41
2004DB2	**92.50**	88.88	87.02
2004DB3	78.41	91.86	**94.36**

TABLE III
JACCARD SCORE OBTAINED BY VARIOUS STATE-OF-THE-ART
SEGMENTATION ALGORITHMS.

Database	Unet (↑)	Ccnet (↑)	RUnet (↑)
2000DB1	27.73	53.41	**57.39**
2000DB2	**65.23**	59.12	55.90
2000DB3	57.76	**85.47**	83.20
2002DB1	**95.70**	82.92	94.57
2002DB2	87.09	81.95	**87.29**
2002DB3	61.79	**85.36**	58.03
2004DB1	**97.76**	92.50	96.88
2004DB2	**86.32**	81.17	79.15
2004DB3	65.59	85.47	**89.67**

training). A total of 240 images are used as the source domain images. Target domain consists of FVC 2000DB1-DB3, FVC 2002DB1-DB3 and FVC 2004DB1-DB3 (subset A defined for testing) leading to a total of 7200 fingerprint images. Details about the FVC databases used in this study are presented in Table I. The manually marked roi masks are taken from [42][1].

VII. EVALUATION METRICS

To quantify the segmentation performance, this research is evaluated on two standard metrics used by the literature on segmentation: Dice [6] and Jaccard score [4].

$$Dice = \frac{2 \times TP}{(TP + FP) + (TP + FN)} \quad (9)$$

$$Jaccard = \frac{TP}{(TP + FP + FN)} \quad (10)$$

where TP, TN, FP and FN denote true positive, true negative, false positive and false negative with respect to groundtruth roi mask.

VIII. RESULTS AND DISCUSSIONS

A. Benchmarking Results

Table II and Table III report the dice and jaccard similarity scores respectively, obtained by the benchmarked algorithms. Although Unet outperforms other architectures on five databases out of nine, however, it performs significantly bad on some. RUnet on the other hand, has the best generalization ability. As the proposed research is intended to provide good generalization ability, we choose RUnet as the backbone architecture.

To give readers a perspective on how challenging the problem is, Table IV reports the dice and jaccard scores with and without using the annotations of roi masks target sensor images. RUnet (synthetic) represents the scenario when roi masks from only synthetic images (source domain) are used. RUnet (full) on the other hand, represents the case when training is conducted in a fully supervised fashion, using the fingerprint images and the corresponding annotated roi

[1]https://figshare.com/articles/dataset/Benchmark_for_Fingerprint_Segmentation_Performance_Evaluation/1294209

masks from the training set of target domain. As expected, the segmentation performance achieved by RUnet (synthetic) is significantly lower compared to RUnet (full). This compels us to propose aligning features of source and target domain to facilitate improved performance.

B. Improved Generalization Ability

Table V compares segmentation performance of the proposed RA-RUnet with baseline RUnet. Visualizations obtained from Seg-Grad-Cam [46], as shown in Figure 3, signify that higher activations around boundaries of foreground roi are obtained by RA-RUnet as compared to RUnet. As a result, RA-RUnet achieves better segmentation performance.

C. Comparison With State-of-the-art

Table VI compares the segmentation performance of proposed RA-RUnet with the recently proposed ASA-Net [52],

TABLE IV
PERFORMANCE OF RUNET WHEN TRAINED ON ONLY SYNTHETIC
FINGERPRINTS VERSUS FULL TRAINING SET.

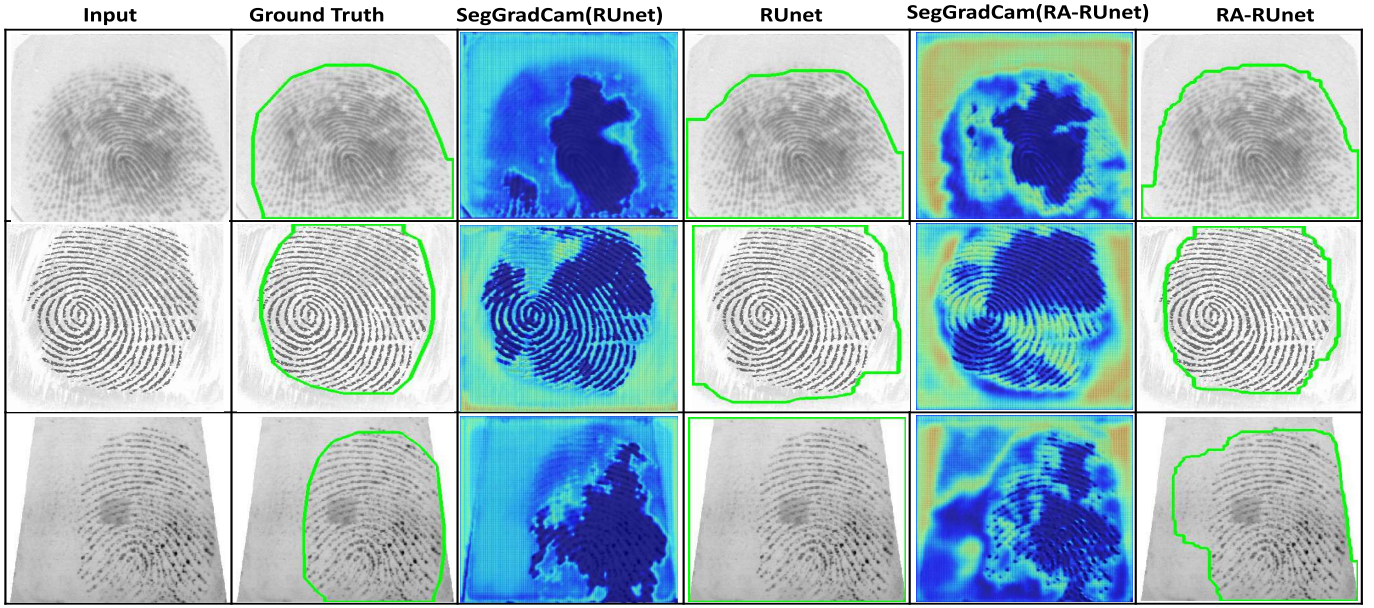| Database	Dice Score (↑)		Jaccard Similarity (↑)	
	RUnet (synthetic)	RUnet (full)	RUnet (synthetic)	RUnet (full)
2000DB1	71.24	**93.34**	57.39	**88.15**
2000DB2	68.84	**92.39**	55.90	**86.40**
2000DB3	89.98	**96.50**	83.20	**93.74**
2002DB1	97.19	**98.44**	94.57	**96.95**
2002DB2	92.64	**97.28**	87.29	**94.88**
2002DB3	70.93	**95.53**	58.03	**91.83**
2004DB1	98.41	**99.38**	96.88	**98.78**
2004DB2	87.02	**96.69**	79.15	**93.94**
2004DB3	94.36	**97.17**	89.67	**94.62**

Fig. 3. **Visualizations obtained using Seg-Grad-Cam** (best viewed in colour). Higher activations around boundaries are obtained by RA-RUnet compared to the baseline RUnet. This explains the improved generalization ability of RA-RUnet.

TABLE V
COMPARISON OF DICE SCORE AND JACCARD SIMILARITY OBTAINED BY
BASELINE RUNET AND PROPOSED RA-RUNET.

Database	Dice Score (↑)		Jaccard Similarity (↑)	
	RUnet	RA-RUnet	RUnet	RA-RUnet
2000DB1	71.24	**76.85**	57.39	**63.83**
2000DB2	68.84	**80.54**	55.90	**69.10**
2000DB3	89.98	**92.48**	83.20	**86.81**
2002DB1	**97.19**	95.77	**994.57**	91.96
2002DB2	**92.64**	90.52	**87.29**	83.83
2002DB3	70.93	**75.41**	58.03	**61.69**
2004DB1	**98.41**	97.75	**96.88**	95.62
2004DB2	87.02	**93.78**	79.15	**88.68**
2004DB3	94.36	**94.50**	89.67	**89.72**

TABLE VII
DICE SCORE OBTAINED BY VARIOUS STATE-OF-THE-ART SEGMENTATION
ALGORITHMS.

Database	RUnet (↑)	RA-RUnet ($T =1$) (↑)	RA-RUnet ($T =3$)(↑)
2000DB1	71.24	70.52	**76.85**
2000DB2	68.84	78.21	**80.54**
2000DB3	89.98	**93.21**	92.48
2002DB1	**97.19**	95.72	*95.77*
2002DB2	**92.64**	91.10	90.52
2002DB3	70.93	**88.35**	75.41
2004DB1	**98.41**	97.78	97.75
2004DB2	87.02	93.56	**93.78**
2004DB3	94.36	93.87	**94.50**

TABLE VI
COMPARISON OF DICE SCORE AND JACCARD SIMILARITY OBTAINED BY
ASA-NET [52] AND PROPOSED RA-RUNET.

Database	Dice Score (↑)		Jaccard Similarity (↑)	
	[52]	RA-RUnet	[52]	RA-RUnet
2000DB1	34.14	**76.85**	21.24	**63.83**
2000DB2	66.19	**80.54**	52.63	**69.10**
2000DB3	91.21	**92.48**	85.24	**86.81**
2002DB1	91.00	**95.77**	83.88	**91.96**
2002DB2	88.73	**90.52**	81.18	**83.83**
2002DB3	**87.48**	75.41	**78.18**	61.69
2004DB1	96.22	**97.75**	92.82	**95.62**
2004DB2	80.64	**93.78**	68.82	**88.68**
2004DB3	73.15	**94.50**	58.91	**89.72**

TABLE VIII
JACCARD SCORE OBTAINED BY VARIOUS STATE-OF-THE-ART
SEGMENTATION ALGORITHMS.

Database	RUnet (↑)	RA-RUnet ($T =1$) (↑)	RA-RUnet ($T =3$)(↑)
2000DB1	57.39	56.15	**63.83**
2000DB2	55.90	66.22	**69.10**
2000DB3	83.20	**87.88**	86.81
2002DB1	**94.57**	91.87	*91.96*
2002DB2	**87.29**	84.67	83.83
2002DB3	58.03	**79.84**	61.69
2004DB1	**96.88**	95.67	95.62
2004DB2	79.15	88.33	**88.68**
2004DB3	89.67	88.56	**89.72**

a state-of-the-art unsupervised domain adaptation method for semantic segmentation. Results reveal that the proposed RA-RUnet is better suited for fingerprints compared to ASA-Net as RA-RUnet significantly outperforms ASA-Net over eight out of nine databases.

D. Recurrent Discrimination Promotes Improved Segmentation

To showcase that recurrent discrimination yields improved segmentation, Table VII compares the segmentation performance when the discrimination is performed for just one iteration (T=1) compared to three (T=3). Please note, for both

TABLE IX
COMPARISON OF DICE SCORE AND JACCARD SIMILARITY OBTAINED BY PROPOSED RA-RUNET FOR DIFFERENT VALUES OF α.

Database	Dice Score (↑)					Jaccard Similarity (↑)				
	0.2	0.5	1	2	5	0.2	0.5	1	2	5
2000DB1	42.98	62.97	**76.85**	51.49	47.51	28.59	47.50	**63.83**	36.16	32.59
2000DB2	74.72	78.35	**80.54**	76.14	76.93	61.58	66.02	**69.10**	64.03	64.15
2000DB3	90.71	92.74	92.48	**92.92**	88.20	83.89	87.18	86.81	**87.42**	80.54
2002DB1	95.27	95.92	95.77	**95.95**	94.89	91.05	92.23	91.96	**92.28**	90.38
2002DB2	88.97	**90.66**	90.52	90.51	88.57	81.29	**84.03**	83.83	83.75	80.82
2002DB3	**79.49**	74.40	75.41	76.28	64.66	**66.91**	60.38	61.69	62.83	49.00
2004DB1	97.54	97.84	97.75	**97.90**	97.40	95.23	95.80	**95.92**	95.91	94.97
2004DB2	90.10	92.64	**93.78**	87.60	79.89	82.77	86.93	**88.68**	79.77	68.48
2004DB3	93.64	**94.94**	94.50	93.76	93.96	88.18	**90.48**	89.72	88.39	88.73

the cases, segmentation mask is iteratively refined thrice as done in baseline RUnet. Here T characterizes no of iterations used by the discriminator. For both the cases, as compared to the baseline RUnet, segmentation performance has significantly improved. However, model with T=3 outperforms the one T=1 on five databases. Thus, recurrent discrimination helps to improve performance without adding any more model parameters.

E. Effect of Hyper-parameter α

The loss function is a weighted combination of segmentation and adversarial loss characterized by hyper-parameter α. As reported in Table IX, we observe that initially, as α increases, contribution of adversarial loss increases. As a result, sensor-invariant features are learnt and the segmentation performance improves. However, beyond a point, increasing α leads to decrease in contribution of segmentation loss which negatively affects the segmentation performance. Therefore, a careful choice of α needs to be made in order to have a well balanced loss function and best segmentation performance.

IX. CONCLUSION AND FUTURE WORK

Presented research is the first work on improving fingerprint roi segmentation performance on a new sensor without requiring its corresponding manually marked roi. Towards this, benchmarking of three state-of-the-art semantic segmentation models is performed to find the best backbone network. Afterwards, recurrent adversarial learning based feature alignment is performed to ensure that fingerprints originating from source and target fingerprint sensors have similar features. Experiments and visualizations demonstrate that proposed feature alignment framework improves the activations and segmentation performance around boundaries. Additionally, comparison between single-level and recurrent adversarial learning based alignment is performed to give insights on the improved roi segmentation performance. In future, feature alignment can be explored for other modules of a fingerprint matching system.

ACKNOWLEDGMENT

REFERENCES

[1] F. Alonso-Fernandez, R. N. Veldhuis, A. M. Bazen, J. Fiérrez-Aguilar, and J. Ortega-Garcia. Sensor interoperability and fusion in fingerprint verification: A case study using minutiae-and ridge-based matchers. In *2006 9th International Conference on Control, Automation, Robotics and Vision*, pages 1–6, 2006.

[2] H. Alshehri, M. Hussain, H. A. Aboalsamh, and M. A. Al Zuair. Cross-sensor fingerprint matching method based on orientation, gradient, and gabor-hog descriptors with score level fusion. *IEEE Access*, 6:28951–28968, 2018.

[3] H. AlShehri, M. Hussain, H. A. AlSamh, and M. Zuair. Fingerprint verification system for cross-sensor matching based on lbp and sift descriptors and score level fusion. *International Journal of Computer Science and Software Engineering*, 7(3):47–51, 2018.

[4] S.-S. Choi, S.-H. Cha, and C. C. Tappert. A survey of binary similarity and distance measures. *Journal of Systemics, Cybernetics and Informatics*, 8(1):43–48, 2010.

[5] R. C. da Silva Vasconcelos and H. Pedrini. Fingerprint image segmentation based on oriented pattern analysis. In *VISIGRAPP (4: VISAPP)*, pages 405–412, 2019.

[6] L. R. Dice. Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302, 1945.

[7] M. F. Fahmy and M. Thabet. A fingerprint segmentation technique based on morphological processing. In *IEEE International Symposium on Signal Processing and Information Technology*, pages 000215–000220, 2013.

[8] P. M. Ferreira, A. F. Sequeira, and A. Rebelo. A fuzzy c-means algorithm for fingerprint segmentation. In *Iberian Conference on Pattern Recognition and Image Analysis*, pages 245–252, 2015.

[9] Y. Ganin and V. Lempitsky. Unsupervised domain adaptation by backpropagation. In *International conference on machine learning*, pages 1180–1189. PMLR, 2015.

[10] Y. Ganin and V. Lempitsky. Unsupervised domain adaptation by backpropagation. In *International Conference on Machine Learning*, pages 1180–1189, 2015.

[11] A. Ghosh, V. Kulharia, V. P. Namboodiri, P. H. Torr, and P. K. Dokania. Multi-agent diverse generative adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 8513–8521, 2018.

[12] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. In *Advances in neural information processing systems*, pages 2672–2680, 2014.

[13] Y. Han, J. Nam, N. Park, and H. Kim. Resolution and distortion compensation based on sensor evaluation for interoperable fingerprint recognition. In *The 2006 IEEE International Joint Conference on Neural Network Proceedings*, pages 692–698, 2006.

[14] C. Hu, J. Yin, E. Zhu, H. Chen, and Y. Li. A composite fingerprint segmentation based on log-gabor filter and orientation reliability. In *2010 IEEE International Conference on Image Processing*, pages 3097–3100, 2010.

[15] Z. Huang, X. Wang, L. Huang, C. Huang, Y. Wei, and W. Liu. Ccnet: Criss-cross attention for semantic segmentation. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 603–612, 2019.

[16] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE*

conference on computer vision and pattern recognition, pages 1125–1134, 2017.

[17] J. Jang, S. J. Elliott, and H. Kim. On improving interoperability of fingerprint recognition using resolution compensation based on sensor evaluation. In *International Conference on Biometrics*, pages 455–463, 2007.

[18] I. Joshi, A. Anand, M. Vatsa, R. Singh, S. D. Roy, and P. Kalra. Latent fingerprint enhancement using generative adversarial networks. In *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 895–903, 2019.

[19] I. Joshi, R. Kothari, A. Utkarsh, V. K. Kurmi, A. Dantcheva, S. D. Roy, and P. K. Kalra. Explainable fingerprint roi segmentation using monte carlo dropout. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV) Workshops*, pages 60–69, 2021.

[20] V. K. Kurmi, S. Kumar, and V. P. Namboodiri. Attending to discriminative certainty for domain adaptation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.

[21] V. K. Kurmi, S. Kumar, and V. P. Namboodiri. Attending to discriminative certainty for domain adaptation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 491–500, 2019.

[22] R. Li, W. Cao, S. Wu, and H.-S. Wong. Generating target image-label pairs for unsupervised domain adaptation. *IEEE Transactions on Image Processing*, 2020.

[23] E. Liu, H. Zhao, F. Guo, J. Liang, and J. Tian. Fingerprint segmentation based on an adaboost classifier. *Frontiers of Computer Science in China*, 5(2):148–157, 2011.

[24] Y.-W. Luo, C.-X. Ren, D. Dao-Qing, and H. Yan. Unsupervised domain adaptation via discriminative manifold propagation. IEEE, 2020.

[25] E. Marasco, Z. Chapman, and B. Cukic. Improving fingerprint interoperability by integrating wavelet entropy and binarized statistical image features. *Biosig*, 2016.

[26] E. Marasco, A. Feldman, and K. R. Romine. Enhancing optical cross-sensor fingerprint matching using local textural features. In *2018 IEEE Winter Applications of Computer Vision Workshops (WACVW)*, pages 37–43, 2018.

[27] E. Marasco, L. Lugini, B. Cukic, and T. Bourlai. Minimizing the impact of low interoperability between optical fingerprints sensors. In *2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, pages 1–8, 2013.

[28] D. Pathak, P. Krahenbuhl, J. Donahue, T. Darrell, and A. A. Efros. Context encoders: Feature learning by inpainting. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2536–2544, 2016.

[29] Z. Pei, Z. Cao, M. Long, and J. Wang. Multi-adversarial domain adaptation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, 2018.

[30] Z. Pei, Z. Cao, M. Long, and J. Wang. Multi-adversarial domain adaptation. In *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.

[31] O. Ronneberger, P. Fischer, and T. Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. pages 234 – 241, 2015.

[32] A. Ross and A. Jain. Biometric sensor interoperability: A case study in fingerprints. In *International Workshop on Biometric Authentication*, pages 134–145. Springer, 2004.

[33] A. Ross and R. Nadgir. A thin-plate spline calibration model for fingerprint sensor interoperability. *IEEE Transactions on Knowledge and Data Engineering*, 20(8):1097–1110, 2008.

[34] Y. Saatchi and A. G. Wilson. Bayesian gan. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, pages 3625–3634, 2017.

[35] S. Sankaranarayanan, Y. Balaji, C. D. Castillo, and R. Chellappa. Generate to adapt: Aligning domains using generative adversarial networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 8503–8512, 2018.

[36] E. Schonfeld, B. Schiele, and A. Khoreva. A u-net based discriminator for generative adversarial networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8207–8216, 2020.

[37] A. C. Schütz, D. I. Braun, and K. R. Gegenfurtner. Eye movements and perception: A selective review. *Journal of vision*, 11(5):9–9, 2011.

[38] P. B. S. Serafim, A. G. Medeiros, P. A. Rego, J. G. R. Maia, F. A. Trinta, M. E. Maia, J. A. F. Macêdo, and A. V. L. Neto. A method based on convolutional neural networks for fingerprint segmentation.

In *2019 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8, 2019.

[39] B. Stojanović, O. Marques, A. Nešković, and S. Puzović. Fingerprint roi segmentation based on deep learning. In *2016 24th Telecommunications Forum (TELFOR)*, pages 1–4, 2016.

[40] B. Sun and K. Saenko. Deep coral: Correlation alignment for deep domain adaptation. In *European Conference on Computer Vision*, pages 443–450. Springer, 2016.

[41] R. F. Teixeira and N. J. Leite. Unsupervised fingerprint segmentation based on multiscale directional information. In *Iberoamerican Congress on Pattern Recognition*, pages 38–46, 2011.

[42] D. H. Thai and C. Gottschlich. Global variational method for fingerprint segmentation by three-part decomposition. *IET Biometrics*, 5(2):120–130, 2016.

[43] D. H. Thai, S. Huckemann, and C. Gottschlich. Filter design and performance evaluation for fingerprint image segmentation. *PloS one*, 11(5):e0154160, 2016.

[44] Z. Tu and X. Bai. Auto-context and its application to high-level vision tasks and 3d brain image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 32(10):1744–1757, 2009.

[45] E. Tzeng, J. Hoffman, N. Zhang, K. Saenko, and T. Darrell. Deep domain confusion: Maximizing for domain invariance. *arXiv preprint arXiv:1412.3474*, 2014.

[46] K. Vinogradova, A. Dibrov, and G. Myers. Towards interpretable semantic segmentation via gradient-weighted class activation mapping (student abstract). In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, pages 13943–13944, 2020.

[47] W. Wang, K. Yu, J. Hugonot, P. Fua, and M. Salzmann. Recurrent u-net for resource-constrained segmentation. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2142–2151, 2019.

[48] Y. Xu, Y. Wang, J. Liang, and Y. Jiang. Augmentation data synthesis via gans: Boosting latent fingerprint reconstruction. In *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2932–2936, 2020.

[49] G. Yang, G.-T. Zhou, Y. Yin, and X. Yang. K-means based fingerprint segmentation with sensor interoperability. *EURASIP Journal on Advances in Signal Processing*, 2010(1):729378, 2010.

[50] R. A. Yeh, C. Chen, T. Yian Lim, A. G. Schwing, M. Hasegawa-Johnson, and M. N. Do. Semantic image inpainting with deep generative models. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 5485–5493, 2017.

[51] Z. Zheng and Y. Yang. Rectifying pseudo label learning via uncertainty estimation for domain adaptive semantic segmentation. *International Journal of Computer Vision*, pages 1–15, 2021.

[52] W. Zhou, Y. Wang, J. Chu, J. Yang, X. Bai, and Y. Xu. Affinity space adaptation for semantic segmentation across domains. *IEEE Transactions on Image Processing*, 2020.