Human Gender Classification Based on Hand Images Using Deep Learning

Rajesh Mukherjee, Asish Bera, Debotosh Bhattacharjee and Mita Nasipuri

December 7, 2021

Human Gender Classification based on Hand Images using Deep Learning

No Author Given

No Institute Given

Abstract. Soft biometrics such as the gender, age, etc. can offer relevant information for person identification. The hand-based modalities are widely studied for conventional biometric recognition for various applications. However, a little research attention is grown to tackle soft biometrics using hand images. In this paper, human gender classification is addressed using the frontal and dorsal hand images. For experimentation, we have created a new hand dataset at our University, denoted as U-HD [1], representing sufficient posture variations at an uncontrolled environment. We have collected the sample hand images of 57 persons to incorporate more user-flexibility in posing their hands that incur additional challenges to discriminate the gender of the person. Five state-of-the-art deep neural architectures are used as the backbones, and a simple deep model is used for the human gender discrimination. The method achieves the best 90.49% accuracy on the U-HD using the Inception-V3 model.

Keywords: Convolutional Neural Networks · Soft Biometrics · Hand Biometrics · Gender Recognition

1 Introduction

Soft biometric characteristics (e.g. gender, age, ethnicity, etc.) are very essential for several biometric, commercial, surveillance, and forensics applications [15], [16]. To discriminate an individual, these kinds of personal ancillary information (soft biometrics), mainly the human face, palmprint, hand shape, gait, and others are analysed over the years [28], [31]. The face and hand modalities are predominantly explored by the biometric community, mostly for the biometric recognition, and related other applications in the wider areas of artificial intelligence, machine leaning, and pattern recognition. Among several physiological modalities, hand-based traits such as palmprint, fingerprint, hand geometry, etc. are getting profound research interests [1], [7], [25], [29], [30]. Hand shape and palmprint are well-known traits, suitable for personal verification as a conventional (a.k.a. hard) as well as soft biometric traits [2], [5], [6], [7], [8], [9], [10], [11], [20]. Other than traditional biometric authentication, such as attendance maintenance, hand images can also be useful for anti-spoofing [12], liveness detection [14], CAPTCHA verification [11], gesture recognition [26], gender classification [25], and other pertinent applications. These diverse applications are also developed using other traits such as face, fingerprint, etc. (uni-modal) and/or their fusion

[1] University affiliation and author's information is anonymous for review process

(multi-biometrics) [17], [19], [21], [27], [32]. However, a little research interests are focused in the direction of soft biometrics relying on the frontal/dorsal hand images. Thus, our main objective is to determine the gender of a person from frontal and dorsal hand images which are collected at an unconstrained and contact-less environment with sufficient posture variations.

Traditional gender recognition systems follow a general pipeline with subsequent stages [33]. After pre-processing, diverse types of hand-crafted features are computed for classification [4]. Though, these methods offer a good solution, yet, suffer from certain limitations, such as manual supervision during feature extraction and selection. Many existing methods cannot produce high performance for a larger dataset. Thus, recent works are employing the convolutional neural networks (CNNs) to boost the performance for a larger dataset with less manual supervision. The advancements of CNNs have attracted a significant research attention which can outperform conventional methods. The effectiveness of CNNs is relies on salient feature extraction from multiple-stacked deep layers in an end-to-end pipeline which is used for a broad range of computer vision and pattern recognition problems [23], and biometrics [30].

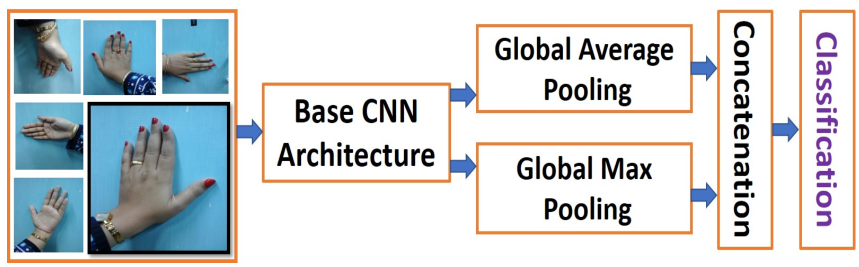


Fig. 1: Proposed method for gender recognition. The CNNs use input hand images, and compute the deep features which are pooled in two paths (global average pooling and global max pooling); and finally, combine them prior to classification. Sufficient pose-variations are considered during sample collection which is illustrated in the left-most grid for a randomly chosen female-user from U-HD.

Recently, the efficacy of CNNs has been explored using hand shape and palmprint [2], [5], [35], [36]. However, the application of CNNs for gender recognition is not fully explored. In this work, we propose a simple and effective baseline deep architecture for gender classification from hand images. A new hand dataset is created our University laboratory, denoted as U-HD. It contains a wider range of variations in hand pose which is novel than other existing dataset, such as 11k [2]. Sample images of the U-HD illustrating ample posture variations at different environmental conditions, are shown in Fig.2. The main contributions of this work are two-fold:

- A new hand database is created for gender classification, representing the frontal and dorsal hand images at unrestricted and sufficient pose variations.
- A fusion of pooled deep features, extracted from the state-of-the-art base CNN architectures is applied to assess the performance for gender classification.

The rest of the paper is organized as follows: related work is studied in Section 2, proposed method is described in Section 3, and experimental description is given in Section 4. Finally, the conclusion is drawn in Section 5.

2 Related Work

Earlier gender classification systems are developed using the region and boundary features of the hand based on Zernike moments, Fourier descriptors, and fusion strategy [4]. The geometric features of the palm images are also tested for gender classification in [33]. Geometrical measurements of the right hand are used for this purpose from 50 Iraqi adults in [3]. Recently, few approaches are devised using the CNNs to address this problem. One such relevant approach is 11k hands [2]. It provides a larger hand dataset for both hand biometric verification and gender classification using palmar and dorsal hand images. However, the sample images are captured at a controlled environment on a contrasting white background (see Fig.3). It follows a two stream CNN model, considering the color images and local binary pattern (LBP). It attains 94.2% and 97.3% accuracy for palmar and dorsal hand, respectively using the pre-trained AlexNet model. Global and Part-Aware Network (GPA-Net) also evaluates the performance on 11k hands for biometric verification [5]. A method for multiple attribute estimation using multi-CNNs from hand images of 11k hand dataset is presented in [24]. Gender and ethnicity classification using palmar hand images of the NTU-PI-v1 database is presented in [25]. Five state-of-the-art CNNs are fine-tuned and tested on the full hand, segmented, and palmprint (RoI) images, and 88.13% accuracy for gender classification is achieved using the DenseNet. In addition, a deep learning based method is proposed recently in this area using the fingerprints in [29].

In this direction, other modalities such as face, ear, palmprint, gait, etc. are also explored by the researchers [13], [22], [34], [18]. Inspired with this progress, we have studied a simple deep model for gender classification using a new growing hand dataset.

3 Methodology

The proposed method fuses two branches of pooled feature maps in the deep network, as illustrated in Fig. 1. Let, the original RGB hand image is $\mathtt{I} \in \mathtt{R}^{\mathtt{H}\times\mathtt{W}\times 3}$, where $\mathtt{H}$ and $\mathtt{W}$ represent the height and width of the input image with 3 color channels, respectively. We use a backbone convolutional neural network $\mathtt{N}(\sum_{\mathtt{i}} \omega_{\mathtt{i}}.\mathbf{x}_{\mathtt{i}} + \mathtt{b}_{\mathtt{i}})$ (weight ω_i, input x_i, and bias b_i at i^{th} layer) to extract the deep features $\mathtt{F} = \mathtt{N}(\mathtt{I}) \in \mathtt{R}^{h\times w\times c}$ from $\mathtt{I}$, where $\mathtt{h}$, $\mathtt{w}$, and $\mathtt{c}$ represent the height, width, and the number of channels of the convolutional feature maps respectively, obtained from the last layer of the base CNN. Next, we apply the global average pooling $\mathtt{p}_{\mathtt{avg}}$ (GAP) and global max pooling $\mathtt{p}_{\mathtt{max}}$ (GMP) to down-sample the spatial feature dimension by selecting the informative features via two different paths, $f_a = \mathtt{p}_{\mathtt{avg}}(\mathtt{F})$ and $f_m = \mathtt{p}_{\mathtt{max}}(\mathtt{F})$. It aggressively summarizes the features in the feature maps and maintains the translation invariance property. Both of the pooling are important to emphasize on the most important feature as well as the average value of the feature map. Therefore, mixing them can be an alternative choice to consider the benefits of both which is applied here. Now, these two feature vectors are combined to generate a high level convolutional feature map $\mathtt{f}_{\mathtt{c}} = \mathtt{f}_{\mathtt{a}} \oplus \mathtt{f}_{\mathtt{m}}$, where $\oplus$ is a linear concatenation function. Finally, the *softmax* predicts the probability ($\bar{y}$) to classify the gender from the input image, $\bar{\mathtt{y}}[0,1] = \mathtt{softmax}(\mathtt{f}_{\mathtt{c}})$. To solve binary classification problem, '0' is labeled for the 'female' and '1' is assigned as the 'male' class.

$$\mathtt{F} = \mathtt{N}(\mathtt{I}); \quad \text{where } \mathtt{F} \in \mathtt{R}^{h\times w\times c}; \ \mathtt{I} \in \mathtt{R}^{H\times W\times 3}; \quad \text{and } \mathtt{N}(\sum_{\mathtt{i}} \omega_{\mathtt{i}}.\mathbf{x}_{\mathtt{i}} + \mathtt{b}_{\mathtt{i}}) \quad \text{is a CNN} \tag{1}$$

$$f_a = p_{avg}(\mathtt{F}) \ \text{ and } \ f_m = p_{max}(\mathtt{F}); \ \ f_c = [f_a \oplus f_m]; \ \text{ and } \ \bar{y} = softmax(f_c) \tag{2}$$

3.1 Implementation

We have used five backbone CNNs and fine-tuned on the target dataset. Particularly, the VGG-16, VGG-19, Xception, ResNet-50, and Inception-v3 are used for feature extraction from the last layer of the respective model. Then, we apply the GAP and GMP via two different paths (see Fig.1). Finally, these two paths are concatenated and followed by the *softmax* layer to generate the output probabilities for recognition. The binary cross-entropy loss function is used for optimizing the training process.

$$L = \frac{1}{k}\sum_{i=1}^{k} -[y_i.log(p_i) + (1-y_i).log(1-p_i)] \tag{3}$$

where, p_i is the probability of class-female, and $(1-p_i)$ is the probability of male-class, k is the number of data samples. The stochastic gradient descent (SGD) optimizer is used with a learning rate of 0.01. The model is trained for 100 epochs with a mini batch-size of 16. The input image-size is $448 \times 448 \times 3$ and applied the standard data augmentation, horizontal-flip ± 0.2, and rotation ± 0.2. Pre-trained imagenet weights are used in the base $\mathtt{N}(.)$ for faster convergence in the learning task. The output feature dimension $(\mathtt{F})$ of Xception base CNN is $7 \times 7 \times 2048$, and it varies for other CNNs. After pooling (GAP and GMP), the feature space becomes $\mathtt{f_a} = \mathtt{f_m} = \mathtt{2048}$ and concatenation generates $\mathtt{f_c} = \mathtt{4096}$ features. A dropout regularization rate of 0.2 is also applied to avoid the overfitting problem. Lastly, the dense layer with *softmax* produces binary-class probabilities within [0,1] for the gender classification of the persons.

4 Experiments

We precisely describe the biometric sample collection from both sides of a hand. Next, some experimental analysis using five base CNNs are provided.

Table 1: Dataset description: frontal, dorsal, and combined both types hand images

Gender	Persons	Frontal		Dorsal		Combined Both	
		Training	Testing	Training	Testing	Training	Testing
Female	17	92	50	105	40	197	90
Male	40	220	112	254	82	474	194
Total	57	312	162	359	122	671	284

4.1 Proposed Hand Dataset (U-HD)

A new hand dataset is created at the departmental laboratory (Computer Science) of our University, named as U-HD. It contains the right- and left- hand frontal and dorsal images of 57 persons, including the UG and PG students, researchers, and professors, mostly are from Computer Science department. During image acquisitions, 15-17 hand images are collected from each person using a Samsung digital camera at 72 dpi with sufficient posture, lighting, and environmental variations. The image size is 1536×2048 pixels. The age of the persons ranges from 18 to 50 years, with Asian ethnicity. We have maintained the ethical issues and privacy of the users during the sample collection. There is no financial/commercial gain involved in this work. These are clearly specified to the participants based on which they have agreed upon to provide their biometric samples. Detailed specification about this dataset is given in Table 1.

The motivation of creating a new dataset is mainly the posture variations and user-flexibility. In this regard, sample images from 11k dataset are shown in Fig.3. It shows the clear environmental and image acquisition conditions, little posture variations, and high image quality (1200×1600, 96 dpi) which might not always be possible in read-world scenarios. In contrast, we consider more challenging image acquisition set-up and incorporate more variations. This dataset is still in growing phase. Sample images from our U-HD are shown in Fig.2.

Fig. 2: Sample hand images from U-HD. Left: intra-class posture variations (frontal and dorsal) of the same male-person. Right: inter-class pose variations (frontal and dorsal) of various female-person, showing different background and lighting conditions.

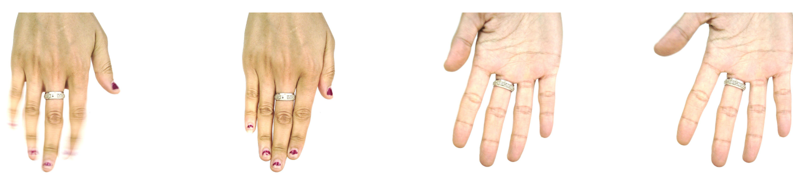


Fig. 3: Sample images from 11k Hand dataset [2]

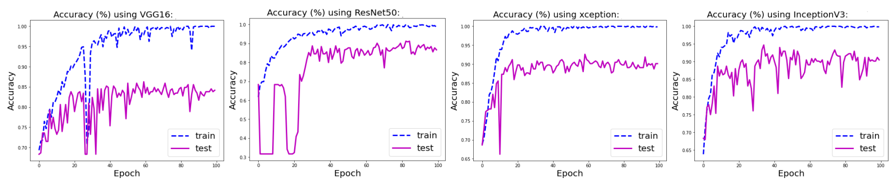
4.2 Gender Classification Results

The Receiver Operator Characteristic (ROC) curve is widely used metric for evaluating binary classification problems. Also, Area Under the Curve (AUC) computes the ability to distinguish between the classes which can be used as a summary of the ROC curve. A higher AUC indicates a better performance of the model to discriminate the female and male classes. Here, we have used the accuracy (%) and AUC (%) metrics.

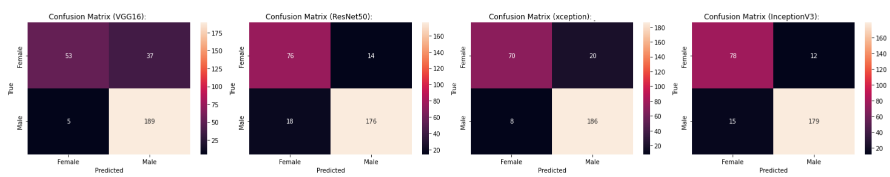
First, we evaluate the performance using the frontal and dorsal hand images separately. Then, we combine both types of samples together for assessment. The results are given in Table 2. The experiments are conducted in Google Colaboratory (Colab-Pro). As the data samples for each types are not sufficient for a deep framework, therefore, the models could not learn efficiently to render very good performance. As more samples are included in the dataset, the CNNs learn better and improve the performance. It is evident from the results in Table 2.

Table 2: Performance (%) of various CNNs for gender classification using the frontal, dorsal and combined both types hand images

CNN Model	Frontal		Dorsal		Combined Both	
	Accuracy	AUC	Accuracy	AUC	Accuracy	AUC
VGG-16	74.10	80.65	80.70	91.92	84.15	87.95
VGG-19	74.70	80.88	74.59	82.63	82.74	87.33
ResNet-50	82.53	87.58	90.16	94.42	88.73	92.20
Xception	**87.35**	**92.79**	92.62	93.70	90.14	93.44
Inception-V3	84.33	91.26	**92.62**	**94.59**	**90.49**	**94.25**

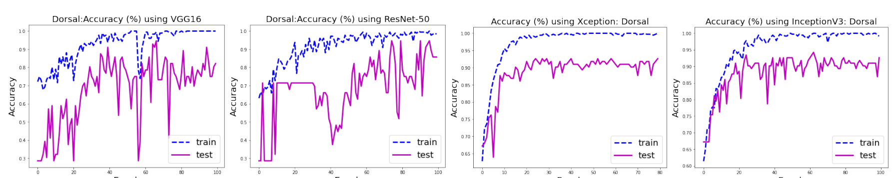


(a) VGG-16 (b) ResNet-50 (c) Xception (d) Inception-V3

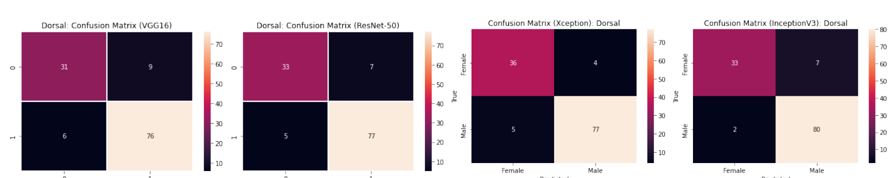


(e) VGG-16 (f) ResNet-50 (g) Xception (h) Inception-V3

Fig. 4: Gender classification performance on both hand-types using various CNNs



(a) VGG-16 (b) ResNet-50 (c) Xception (d) Inception-V3

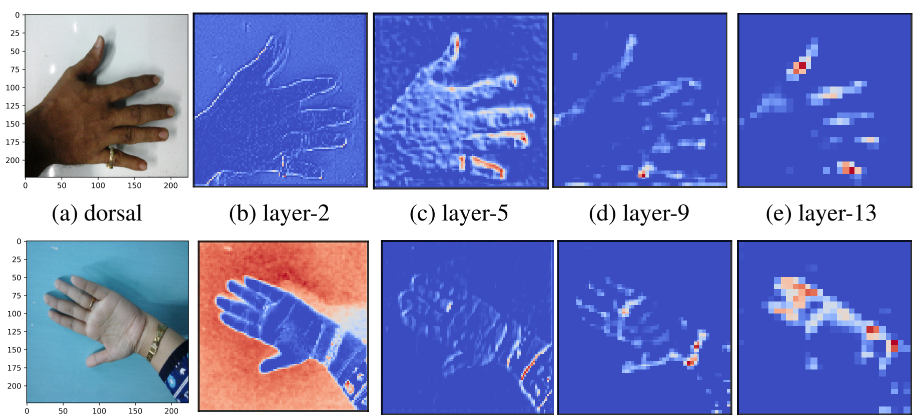


(e) VGG-16 (f) ResNet-50 (g) Xception (h) Inception-V3

Fig. 5: Performance on dorsal hands for gender classification. Best view: zoom-in

In addition, we have plotted the training and testing accuracy during the learning process for the combination of both types of samples in Fig.4. For a clear predictive evaluation, the confusion matrices of the various deep models are also shown. Confusion matrices are useful to conceptualize other standard metrics such as precision, recall, and specificity. In this test, the best performance (accuracy: 90.49% and AUC: 94.25%) is achieved using the Inception-V3. Also, the Xception has attained very close results. However, the VGG models can not perform well compared to the other CNNs. The VGG-16 and VGG-19 are also computationally expensive (in terms of model parameters of the network) than the other models such as Inception-V3. Therefore, the Inception and ResNet family can achieve better performance on this dataset.

Fig. 5 shows few exemplary results on dorsal hands, as shown in Fig.4. Though, promising results are obtained for this challenging task, yet, we can not compare the performance with existing methods directly due to dissimilar dataset. In near future, we plan to compare the performance of our dataset with others, such as 11k.



(a) dorsal (b) layer-2 (c) layer-5 (d) layer-9 (e) layer-13

(f) frontal (g) layer-2 (h) layer-5 (i) layer-9 (j) layer-13

Fig. 6: Feature map visualizations from various convolutional layers of VGG-16.

4.3 Feature Map Visualization

The feature maps from the various blocks of VGG-16 model are shown in Fig.6. Here, we have considered the last convolutional layers of the main blocks of VGG-16 for visualization, and the indices of those layers are: 2, 5, 9, and 13. It is clear that at the initial layers, the low-level feature maps can represent hand shape with finer details. For more deeper level, the CNN reflects more abstract high-level feature representation with less details of the hand shape. It is an inherent characteristic of any deep model.

5 Conclusion

In this paper, we have presented a new work on gender classification using the hand images which are collected at our laboratory. The performance is evaluated using five standard base CNNs on 57 female-male persons. The accuracy reflects that the dataset is more challenging than the existing one. Thus, there is an ample scope to improve the performance further. The dataset is under preparation and growing with a goal to collect more samples from more people. In future, we plan to develop a new deep model for both hard and soft biometrics and their fusion for wider applicability.

References

1. Abderrahmane, M.A., Guelzim, I., Abdelouahad, A.A.: Human age prediction based on hand image using multiclass classification. In: 2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI). pp. 1–5. IEEE (2020)
2. Afifi, M.: 11k hands: Gender recognition and biometric identification using a large dataset of hand images. Multimedia Tools and Applications **78**(15), 20835–20854 (2019)
3. Al-Kharaz, A.: Discriminating between genders using hand images of iraqi adults. In: 2021 IEEE 12th Control and System Graduate Research Colloquium (ICSGRC). pp. 29–33. IEEE (2021)
4. Amayeh, G., Bebis, G., Nicolescu, M.: Gender classification from hand shape. In: 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. pp. 1–7. IEEE (2008)
5. Baisa, N.L., Jiang, Z., Vyas, R., Williams, B., Rahmani, H., Angelov, P., Black, S.: Hand-based person identification using global and part-aware deep feature representation learning. arXiv preprint arXiv:2101.05260 (2021)
6. Bera, A., Bhattacharjee, D.: Human identification using selected features from finger geometric profiles. IEEE Transactions on Systems, Man, and Cybernetics: Systems **50**(3), 747–761 (2017)
7. Bera, A., Bhattacharjee, D., Nasipuri, M.: Hand biometrics in digital forensics. In: Computational Intelligence in Digital Forensics: Forensic Investigation and Applications, pp. 145–163. Springer (2014)
8. Bera, A., Bhattacharjee, D., Nasipuri, M.: Fusion-based hand geometry recognition using dempster–shafer theory. International Journal of Pattern Recognition and Artificial Intelligence **29**(05), 1556005 (2015)
9. Bera, A., Bhattacharjee, D., Nasipuri, M.: Finger contour profile based hand biometric recognition. Multimedia Tools and Applications **76**(20), 21451–21479 (2017)
10. Bera, A., Bhattacharjee, D., Nasipuri, M.: Finger biometric recognition with feature selection. In: The Biometric Computing, pp. 87–112. Chapman and Hall/CRC (2019)
11. Bera, A., Bhattacharjee, D., Shum, H.P.: Two-stage human verification using handcaptcha and anti-spoofed finger biometrics with feature selection. Expert Systems with Applications **171**, 114583 (2021)
12. Bera, A., Dey, R., Bhattacharjee, D., Nasipuri, M., Shum, H.P.: Spoofing detection on hand images using quality assessment. Multimedia Tools and Applications pp. 1–24 (2021)
13. Chai, T., Prasad, S., Wang, S.: Boosting palmprint identification with gender information using deepnet. Future Generation Computer Systems **99**, 41–53 (2019)
14. Czajka, A., Bulwan, P.: Biometric verification based on hand thermal images. In: 2013 International Conference on Biometrics (ICB). pp. 1–6. IEEE (2013)
15. Dantcheva, A., Elia, P., Ross, A.: What else does your biometric data reveal? a survey on soft biometrics. IEEE Transactions on Information Forensics and Security **11**(3), 441–467 (2015)
16. Dantcheva, A., Velardo, C., D'angelo, A., Dugelay, J.L.: Bag of soft biometrics for person identification. Multimedia Tools and Applications **51**(2), 739–777 (2011)
17. Edwards, T., Hossain, M.S.: Effectiveness of deep learning on serial fusion based biometric systems. IEEE Transactions on Artificial Intelligence (2021)
18. Greco, A., Saggese, A., Vento, M., Vigilante, V.: A convolutional neural network for gender recognition optimizing the accuracy/speed tradeoff. IEEE Access **8**, 130771–130781 (2020)
19. Guo, B.H., Nixon, M.S., Carter, J.N.: Fusion analysis of soft biometrics for recognition at a distance. In: 2018 IEEE 4th International Conference on Identity, Security, and Behavior Analysis (ISBA). pp. 1–8. IEEE (2018)

20. Izadpanahkakhk, M., Razavi, S.M., Taghipour-Gorjikolaie, M., Zahiri, S.H., Uncini, A.: Novel mobile palmprint databases for biometric authentication. International Journal of Grid and Utility Computing **10**(5), 465–474 (2019)
21. Jain, A.K., Dass, S.C., Nandakumar, K.: Soft biometric traits for personal recognition systems. In: International conference on biometric authentication. pp. 731–738. Springer (2004)
22. Jain, A., Kanhangad, V.: Gender classification in smartphones using gait information. Expert Systems with Applications **93**, 257–266 (2018)
23. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems **25**, 1097–1105 (2012)
24. Lin, Y.C., Suzuki, Y., Kawai, H., Ito, K., Chen, H.T., Aoki, T.: Attribute estimation using multi-cnns from hand images. In: 2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC). pp. 241–244. IEEE (2019)
25. Matkowski, W.M., Kong, A.W.K.: Gender and ethnicity classification based on palmprint and palmar hand images from uncontrolled environment. In: 2020 IEEE International Joint Conference on Biometrics (IJCB). pp. 1–7. IEEE (2020)
26. Nogales, R.E., Benalcázar, M.E.: Hand gesture recognition using machine learning and infrared information: a systematic literature review. International Journal of Machine Learning and Cybernetics pp. 1–28 (2021)
27. Nugrahaningsih, N., Porta, M.: Soft biometrics through hand gestures driven by visual stimuli. ICT Express **5**(2), 94–99 (2019)
28. Reid, D.A., Nixon, M.S., Stevenage, S.V.: Soft biometrics; human identification using comparative descriptions. IEEE Transactions on pattern analysis and machine intelligence **36**(6), 1216–1228 (2013)
29. Rim, B., Kim, J., Hong, M.: Gender classification from fingerprint-images using deep learning approach. In: Proceedings of the International Conference on Research in Adaptive and Convergent Systems. pp. 7–12 (2020)
30. Sundararajan, K., Woodard, D.L.: Deep learning for biometrics: A survey. ACM Computing Surveys (CSUR) **51**(3), 1–34 (2018)
31. Tome, P., Fierrez, J., Vera-Rodriguez, R., Nixon, M.S.: Soft biometrics and their application in person recognition at a distance. IEEE Transactions on information forensics and security **9**(3), 464–475 (2014)
32. Vasileiadis, M., Stavropoulos, G., Tzovaras, D.: Facial soft biometrics detection on low power devices. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. pp. 0–0 (2019)
33. Wu, M., Yuan, Y.: Gender classification based on geometry features of palm image. The Scientific World Journal **2014** (2014)
34. Yaman, D., Eyiokur, F.I., Sezgin, N., Ekenel, H.K.: Age and gender classification from ear images. In: 2018 International Workshop on Biometrics and Forensics (IWBF). pp. 1–7. IEEE (2018)
35. Yuan, Y., Tang, C., Xia, S., Chen, Z., Qi, T.: Handnet: Identification based on hand images using deep learning methods. In: Proceedings of the 2020 4th International Conference on Vision, Image and Signal Processing. pp. 1–6 (2020)
36. Zhu, J., Zhong, D., Luo, K.: Boosting unconstrained palmprint recognition with adversarial metric learning. IEEE Transactions on Biometrics, Behavior, and Identity Science **2**(4), 388–398 (2020)