Editorial
Introduction to the Special Issue on Biometrics at a Distance in the Deep Learning Era

I. Introduction

BIOMETRICS at a distance is a relevant topic in different areas, primarily video surveillance for forensics on one side and subject recognition for smart ambient in a completely different context. Both classes of applications are relevant nowadays, of course, for different motivations and present common problems. Signal processing, especially related to images, plays a core role in addressing such issues. In this respect, distance is a classical adverse condition that still calls for effective and efficient solutions since it increases the effect of lack of resolution, different points of view, and other hindering factors. The recent achievements of deep learning (DL) architectures provide promising inspiration to tackle related issues. Accordingly, this special issue (SI) of the IEEE Journal of Selected Topics in Signal Processing (J-STSP) is dedicated to using DL for addressing biometrics at a distance.

A total of 36 papers were submitted to this SI, where 14 papers were accepted after a rigorous review process. Three papers in the final collection aim at improving the performance of people re-identification systems, including the use of visible and infrared images, proposing a new retrieval and verification loss function, and suggesting the fusion of visual appearance and soft biometrics. Five works focus on faces, addressing diverse issues such as face image completion or dealing with low-resolution faces, even those affected by turbulences. Three works deal with other human body parts, such as the eye, the palm, and the finger knuckle. Two papers use human motion as a biometric feature. Finally, a new dataset of pedestrians is introduced for soft biometrics at a distance. A more detailed introduction to the SI topics follows.

II. Summary of the SI

The section below highlights the main contributions of the papers included in this SI.

A. People Re-Identification

The main goal of people re-identification in images is to develop computer algorithms that can accurately match images of the same person taken from different camera views or at different times, even when the images may be of different qualities, angles, or lighting conditions. This problem is often studied in the context of surveillance and security applications, where tracking individuals across multiple cameras and locations is necessary. The goal is to develop algorithms that can reliably and efficiently match images of individuals, even in challenging real-world conditions, to improve security and safety.

Liu et al. [A1] tackle the challenge of visible-infrared person re-identification (VI-ReID) by introducing Aligned Grayscale Modality (AGM), a unified dark-line spectrum to reformulate dual-mode learning as a gray-gray single-mode learning problem. AGM generates grayscale modality from visible images and uses a style transfer model to convert infrared images to grayscale, significantly reducing modality discrepancy. A multi-granularity feature extraction network is then introduced to conduct feature-level alignment. Local features, specifically head-shoulder features, are exploited to complement global information and form a stronger feature descriptor for person Re-ID. Experimental results on SYSU-MM01 and RegDB datasets demonstrate that the proposed method outperforms state-of-the-art methods in cross-modality retrieval performance.

Yuan et al. [A2] tackle the interesting problem of automating the design of loss functions with specific evaluation metrics for re-identification (re-ID). This task can be formulated as a retrieval or verification task. However, the possible inconsistencies between hand-designed loss functions and the used evaluation metrics can often degrade the model performance when using retrieval- or verification-based loss functions. The authors propose the Parameterized Retrieval & Verification (RV) Loss to tackle this problem and decrease the required workload. It achieves the automatic searching of loss functions and jointly optimizes RV tasks while introducing parameterized functions to replace non-differentiable operations in RV evaluation metrics. A single formula can represent different evaluation metric approximations by a family of parameterized functions. Though outperforming existing loss functions on several re-ID methods, the authors also sketch the still open problems of their proposed approach. Parameterized RV Loss is quite demanding regarding time and hardware resources for searching for better parameters, and the model performance may degrade without enough search time. Therefore, the optimization of the search strategy is an interesting open challenge.

Finally, Khan et al. [A3] tackles person re-identification by effectively integrating two features: appearance-based features used in most of the existing person re-identification methods;

Digital Object Identifier 10.1109/JSTSP.2023.3269211

and soft biometrics (or attributes). The authors introduce a multi-scale pyramid attention (MSPA) model to leverage their complementary capabilities by considering size and aspect ratio changes depending on the distance from the camera.

B. Face Processing and Recognition

Face recognition technology is becoming essential in various aspects of modern life, from securing personal devices to identifying individuals in public places. However, one of the biggest challenges of face recognition is its accuracy at a distance, where facial features are less distinguishable. Despite this challenge, advancements in DL are making significant progress in improving the technology's ability to accurately recognize faces even at a distance, which has great potential to improve security and efficiency across multiple industries. Several papers on this topic are included in this SI.

Mei et al. [A4] address a topic of great significance that has not yet received adequate attention. Specifically, in the context of face recognition over extremely long distances, atmospheric turbulence causes significant degradation of images. The authors' objective is to remove the effect of turbulence and restore face images while preserving their identities. To achieve this, they have utilized Generative Adversarial Networks (GANs) for image restoration.

Face biometrics are widely used in smart cities but often exhibit performance disparities across gender and ethnic groups. Previous research has mainly focused on high-resolution images, neglecting low-resolution images captured by surveillance cameras at a distance. In [A5], Atzori et al. propose a framework for exploring demographic disparities in low-resolution face recognition. They use a deep generative approach to convert high-resolution face images to low-resolution ones and evaluate the performance of state-of-the-art face recognition models on artificially degraded images from five datasets. The results demonstrate significant disparities across gender and ethnic groups, highlighting the need for timely interventions.

Dosi et al. [A6] explore the problem of identifying the occluded facial features and recognizing a person by non-occluded biometric features. To this aim, they propose and test a novel Segmentation based Disguise Guided Dropout Network (Seg-DGDNet). The proposed approach is divided into disguise detection to discard occluded regions from the recognition process and the following disguised face recognition, resulting in more accurate performance. In practice, the proposed segmentation-based disguise-guided dropout algorithm uses an intelligent dropout technique for discarding occluded regions known as disguise-guided dropout. The performance of the proposed model is analyzed on various datasets under different test scenarios, represented by different disguise types and image resolutions. The experimental results demonstrate the superior performance of the Seg-DGDNet model when compared against the state-of-the-art models for low-resolution face recognition and disguised face recognition. The approach relies on some knowledge of the occlusions present in the images. Therefore an open problem is to detect other facial occlusions, such as fake beards/moustaches and other accessories, or to make the method occlusion type-independent.

Face image quality (FIQ) is key to successful facial recognition on surveillance visual material. However, existing automated methods for assessing FIQ are computationally expensive and limited in their ability to provide specific information about low-quality factors. To address these issues, Macarulla et al. [A7] propose the multitask explainable quality networks (XQNets) that can provide quality values and associated facial and environmental attributes, and automatically learn how each attribute contributes to the quality value. It also introduces a dataset-agnostic quality pairing protocol to ensure fair evaluations. The experimental results on benchmark datasets show that XQNets generalize well across different datasets and outperform state-of-the-art methods.

Finally, Su et al. [A8] focuse on the task of completing facial images based on parsing feature maps. The authors note that deep learning methods often do not consider the semantic structure information of face images. This can result in incomplete or unreasonable results. The authors propose a facial image completion network based on face parsing features to address this problem. The network has two sub-networks: the face parsing map prediction network and the facial image completion network. The face parsing map prediction network predicts the parsing features of the missing regions. The facial image completion network uses these features to guide the facial image completion network to fill the missing areas. The proposed method is validated on two public datasets: CelebA-HQ and Flickr-Faces-High-Quality (FFHQ).

C. Other Body Parts

Apart from the full body or face, other specific body parts can be used for identification. In this section, we present contributions that focus on eyes, finger knuckles, and palmprints.

Taha et al. [A9] present a new learning model that uses only eye movement profiles to authenticate drivers continuously. This is an improvement over existing models that rely on limited or inconvenient modalities. The authors use LSTM and dense networks to learn temporal characteristics from the eye movement profiles. They focus on low-rate devices to increase affordability, which presents challenges due to limited quality measurements. They build their datasets using two low frame-rate devices, Autocruis and GazePoint, and achieve state-of-the-art authentication performance with as low as 30 seconds of frame length.

In [A10], Zhou et al. the distance between the capture device and the biometric sample is lower than the usual one underlying re-identification and, in general, biometrics at a distance. However, the presented approach presents both theoretical and practical interesting aspects. The paper proposes a complete deep neural network-based contactless finger knuckle identification framework. The first problem tackled by the authors is knuckle detection and segmentation. The novelty of

the approach is achieving accurate results with complex backgrounds, which are realistically observed in real-world applications. In addition, the authors design and introduce a new loss function to learn knuckle features from images captured at a relative distance. Experimental results presented in this paper use five different public databases, using challenging protocols and, most of all, cross-database performance evaluation. Also, for this work, despite the outperforming results, open problems trace the line for future work. The first problem common to several deep learning-based approaches is the limited availability of public data for suitable training. The involuntary deformations represent the second problem in the knuckle patterns during the access. This calls for more effective algorithms to handle knuckle deformations for real-world applications. A final mentioned limitation is using a two-stage approach, where one network is trained for knuckle segmentation. In contrast, another network extracts the knuckle features for matching. It would be worth incorporating the features extracted for knuckle detection directly into the recognition process.

Finally, Liang et al. [A11] discusses the challenges of localizing the region of interest (ROI) for touchless palmprint recognition (PPR) in real-world scenarios with complex backgrounds and hand poses. Existing methods based on hand skeletons and landmarks have been insufficiently accurate due to annotation errors. To address this issue, the authors propose a palm keypoint localization neural network (PKLNet) that combines information on the hand region, palm boundary, and finger valley edges to achieve accurate and robust keypoint localization. The network includes a two-stage neural network that effectively adopts the transformer framework to perform palm region segmentation and ROI keypoint coordinate regression. The authors also developed a weakly-supervised training strategy based on conventional palmprint ROI localization methods via palm image synthesis and data augmentation. The results of extensive experiments demonstrate that the proposed PKLNet is robust to palm rotation, translation, and inference from complex backgrounds, ensuring the usability of touchless PPR in real-world application scenarios.

D. Motion-Based Identification

Human motion as a biometric feature at a distance is a growing area of interest in biometrics. It can identify unique gait or body movement patterns, making it useful in security and surveillance. Unlike other biometric features, it can be detected from a greater distance, making it valuable for monitoring large crowds or identifying individuals in challenging environments. It is less affected by changes in appearance, making it reliable. Accurate methods for analyzing human motion can enhance security and public safety.

In this Special Issue, Dehshibi et al. [A12] discuss the potential impact of chronic pain on biometric data analysis and propose a new method to address this issue. To achieve this, the authors use data from inertial measurement unit (IMU) and surface electromyography (sEMG) sensors to classify pain levels and pain-related behavior in the EmoPain database. They propose a sparsely-connected recurrent neural networks (s-RNNs) ensemble with the gated recurrent unit (GRU) that incorporates multiple autoencoders using a shared training framework. Additionally, they fuse hand-crafted features derived from information-theoretic approaches with represented features in the shared hidden state to compensate for variations in the temporal dimension. The experiments demonstrate that the proposed method outperforms the state-of-the-art approaches in classifying pain levels and pain-related behavior.

Regarding using gait motion as a biometric feature, current gait recognition methods consider appearance, posture, and temporal information separately. To address this, Hsu et al. [A13] propose a new framework called GaitTAKE that uses a learned temporal attention mechanism to fuse global and local appearance features and temporal aggregated human pose features. Their experiments show that GaitTAKE achieves superior performance in gait recognition, with competitive performance on the CASIA-B, OU-MVLP, GREW, and Gait3D datasets.

E. New Datasets

Compiling datasets for training and testing biometric models at a distance is important to develop accurate and reliable biometric systems operating in real-world scenarios. Several datasets have been created and publicly released for research in recent years. However, the existing large-scale datasets for soft biometrics were mainly annotated for global soft biometrics like gender and age only. In contrast, in [A14], Hassan et al. constructs a dataset of soft biometrics with rich annotations: categorical annotations for 46 soft biometrics; and comparative annotations for 26 soft biometrics. Also, the dataset contains person images at multiple distances to appropriately evaluate the effect of distances on soft biometrics.

III. Concluding Remarks

This Special Issue covers advances in biometric research, especially in relation to the factor of distance and the use of Deep Learning. The high-quality papers of the Special Issue, "Biometrics at a Distance in the Deep Learning Era," address challenging topics in this area, such as people re-identification, face recognition, motion-based biometrics, and body part-related methods (eyes and hands). We hope this compilation of works helps to advance rapidly in the field of biometrics at a distance.

Manuel J. Marín-Jiménez,
University of Granada
Granada 18071, Spain

Shiqi Yu,
Department of Computer Science and Engineering
Southern University of Science and Technology
Shenzhen 518055, China

YASUSHI MAKIHARA,
Institute for Advanced Co-Creation Studies
Osaka University
Suita 565-0871, Japan

VISHAL M. PATEL,
Department of Electrical and Computer Engineering
Johns Hopkins University
Baltimore, MD 21218, USA

MANEET SINGH,
Indraprastha Institute of Information Technology
Delhi
Delhi 110020, india

MARIA DE MARSICO,
Department of Computer Science
Sapienza University
00185 Roma, Italy

ACKNOWLEDGMENT

The authors would like to thank all authors who submitted their papers to this issue. Our guest editors would also like to express our sincere gratitude to all reviewers.

APPENDIX
RELATED ARTICLES

[A1] H. Liu, D. Xia, and W. Jiang, "Towards homogeneous modality learning and multi-granularity information exploration for visible-infrared person re-identification," *IEEE J. Sel. Topics Signal Process.*, vol. 17, no. 3, May 2023, doi: 10.1109/JSTSP.2022.3233716.

[A2] X. Yuan, X. Xu, Z. Wang, K. Zhang, W. Liu, and R. Hu, "Searching parameterized retrieval & verification loss for re-identification," *IEEE J. Sel. Topics Signal Process.*, vol. 17, no. 3, May 2023, doi: 10.1109/JSTSP.2023.3250989.

[A3] S. U. Khan et al., "Visual appearance and soft biometrics fusion for person re-identification using deep learning," *IEEE J. Sel. Topics Signal Process.*, vol. 17, no. 3, May 2023, doi: 10.1109/JSTSP.2023.3260627.

[A4] K. Mei and V. M. Patel, "LTT-GAN: Looking through turbulence by inverting GANs," *IEEE J. Sel. Topics Signal Process.*, vol. 17, no. 3, May 2023, doi: 10.1109/JSTSP.2023.3238552.

[A5] A. Atzori, G. Fenu, and M. Marras, "Demographic bias in low-resolution deep face recognition in the Wild," *IEEE J. Sel. Topics Signal Process.*, vol. 17, no. 3, May 2023, doi: 10.1109/JSTSP.2023.3249485.

[A6] M. Dosi et al., "SEG-DGDNet: Segmentation based disguise guided dropout network for low resolution face recognition," *IEEE J. Sel. Topics Signal Process.*, 2023.

[A7] A. Macarulla, L. Unzueta, Z. Geradts, M. Worring, and U. Elordi, "Multi-task explainable quality networks for large-scale forensic facial recognition," *IEEE J. Sel. Topics Signal Process.*, vol. 17, no. 3, May 2023, doi: 10.1109/JSTSP.2023.3267263.

[A8] S. Su, X. Shao, L. He, H. Lin, Y. Zuo, and Z. Qiang, "Face image completion method based on parsing features maps," *IEEE J. Sel. Topics Signal Process.*, vol. 17, no. 3, May 2023, doi: 10.1109/JSTSP.2023.3262357.

[A9] B. Taha, S. N. A. Seha, D. Y. Hwang, and D. Hatzinakos, "EyeDrive: A deep learning model for continuous driver authentication," *IEEE J. Sel. Topics Signal Process.*, vol. 17, no. 3, May 2023, doi: 10.1109/JSTSP.2023.3235302.

[A10] Z. Zhou and A. Kumar, "Completely contactless and online finger knuckle identification for real world applications," *IEEE J. Sel. Topics Signal Process.*, vol. 17, no. 3, May 2023, doi: 10.1109/JSTSP.2023.3254148.

[A11] X. Liang, D. Fan, J. Yang, W. Jia, G. Lu, and D. Zhang, "PKLNet: Keypoint localization neural network for touchless palmprint recognition based on edge-aware regression," *IEEE J. Sel. Topics Signal Process.*, vol. 17, no. 3, May 2023, doi: 10.1109/JSTSP.2023.3241540.

[A12] M. Dehshibi et al., "Pain level and pain-related behaviour classification using GRU-based sparsely-connected RNNs," *IEEE J. Sel. Topics Signal Process.*, vol. 17, no. 3, May 2023, doi: 10.1109/JSTSP.2023.3262358.

[A13] H. Hsu, Y. Wang, C.-Y. Yang, J.-N. Hwang, H. L. U. Thuc, and K.-J. Kim, "Learning temporal attention based keypoint-guided embedding for gait recognition," *IEEE J. Sel. Top. Signal Process.*, to be published, doi: 10.1109/ICIP46576.2022.9897409.

[A14] B. Hassan, M. Fiaz, H. H. R. Sherazi, and U. J. Butt, "Annotated pedestrians: A dataset for soft biometrics estimation for varying distances," *IEEE J. Sel. Topics Signal Process.*, vol. 17, no. 3, May 2023, doi: 10.1109/JSTSP.2023.323449.

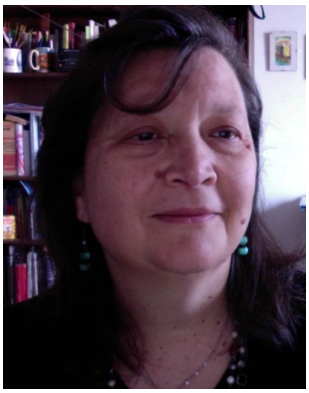
Manuel J. Marín-Jiménez (Member, IEEE) received the B.Sc., M.Sc., and Ph.D. degrees from the University of Granada, Granada, Spain. He was a visiting Student with the Computer Vision Center, Barcelona, Spain, Vislab-ISR/IST of Lisboa, Portugal, and Visual Geometry Group, Oxford, U.K., and visiting Researcher with INRIA-Grenoble (Perception and THOTH teams) and Human Sensing Lab (CMU). He is the co-author of more than 70 technical papers at international venues and a Reviewer of top computer vision and pattern recognition journals. He was also an Associate Professor (tenure position) with the University of Cordoba, Cordoba, Spain. In addition, he is currently a Member of the Governing Board of AERFAI (Spanish Association of *Pattern Recognition and Image Analysis*). His research interests include object detection, human-centric video understanding (including biometrics), visual SLAM, and machine learning.

Shiqi Yu received the B.E. degree in computer science and engineering from Chu Kochen Honors College, Zhejiang University, Hangzhou, China, in 2002, and the Ph.D. degree in pattern recognition and intelligent systems from the Institute of Automation, Chinese Academy of Sciences, Beijing, China, in 2007. He is currently an Associate Professor with the Department of Computer Science and Engineering, Southern University of Science and Technology, Shenzhen, China. He was an Assistant Professor and Associate Professor with the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, from 2007 to 2010, and an Associate Professor with Shenzhen University, Shenzhen, from 2010 to 2019. His research interests include computer vision, image processing, pattern recognition, and artificial intelligence. He was a Member of Program Committees of CCBR 2017, IJCB 2021 and PRCV 2022. Since 2018, he has been the Vice Director of IAPR/IEEE Winter School on Biometrics.

Yasushi Makihara received the B.S., M.S., and Ph.D. degrees in engineering from Osaka University, Suita, Japan, in 2001, 2002, and 2005, respectively. He is currently a Professor of the Institute for Advanced Co-Creation Studies, Osaka University. His research interests include computer vision, pattern recognition, and image processing, and including gait recognition. He has obtained several honors and awards, including the 2nd International Workshop on Biometrics and Forensics (IWBF 2014), IAPR Best Paper Award, 9th IAPR International Conference on Biometrics (ICB 2016), Honorable Mention Paper Award. He was an Associate Editor-in-Chief for *IEICE Transactions on Information and Systems*, Associate Editor for *IPSJ Transactions on Computer Vision and Applications* (CVA), the Program Co-Chair of the 4th Asian Conference on Pattern Recognition (ACPR 2017), area chairs of ICCV 2019, 2021, 2023, CVPR 2020, 2023, and ECCV 2020, 2022.

Vishal M. Patel received the Ph.D. degree in electrical engineering from the University of Maryland, College Park, MD, USA, in 2010. He is currently an Associate Professor with the Department of Electrical and Computer Engineering (ECE), Johns Hopkins University, Baltimore, MD, USA. Before joining Hopkins, he was an A. Walter Tyson Assistant Professor with the Department of ECE, Rutgers University, Newark, NJ, USA, and a Research Faculty Member with the University of Maryland Institute for Advanced Computer Studies (UMIACS). He was the recipient of a number of awards, including the 2021 Pierre-Simon Laplace Early Career Technical Achievement Award, 2021 NSF CAREER Award, 2021 IAPR Young Biometrics Investigator Award (YBIA), 2016 ONR Young Investigator Award, 2016 Jimmy Lin Award for Invention, A. Walter Tyson Assistant Professorship Award, Best Paper Awards at IEEE AVSS 2017 & 2019, IEEE BTAS 2015, IAPR ICB 2018, IEEE ICIP 2021, and two Best Student Paper Awards at IAPR ICPR 2018. He was also the recipient of the 2022 IEEE Transactions on Circuits and Systems for Video Technology Best Paper Award. He is currently an Associate Editor for the IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, *Pattern Recognition* Journal.

Maneet Singh received the Bachelor of Technology degree in computer science from the Indraprastha Institute of Information Technology Delhi, New Delhi, India, in 2015, and the Ph.D. degree from the Indraprastha Institute of Information Technology Delhi. She has authored more than 20 publications, including journals and peer-reviewed conferences. Her research interests include machine learning and deep learning with applications in face recognition. Her research was the recipient of the Best Poster Award at the IEEE International Joint Conference on Biometrics 2017. She was also the recipient of the Google Anita Borg Memorial Scholarship in 2015 and Best Teaching Assistant Award for two consecutive years in 2014 and 2015. She was a Reviewer for several journals, such as *Pattern Recognition*, *Information Fusion*, *Computer Vision and Image Understanding*, and IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY. She is also the Women in Engineering Coordinator of the IEEE Biometrics Council.



Maria de Marsico is currently a Full Professor with the Department of Computer Science, Sapienza University of Rome Rome, Italy. She is also a Member of the Biometrics and Image Processing Laboratory (BIPLab), University of Salerno, Fisciano SA, Italy. She has authored or coauthored more than 200 papers in top-ranked international journals and international conferences, and is a Member of several Program Committees in International Conferences in Biometrics, *Pattern Recognition*, and *Human-Computer Interaction*. Her research interests include image processing, human-computer interaction, biometric systems, and multimodal interaction, and recently on the cooperation of the last two fields in supporting biometric interaction in smart ambient. She is the Editor in Chief of *Pattern Recognition Letters*, an Associate Editor for *Pattern Recognition*, and Associate Editor for IEEE TRANSACTION ON BIOMETRICS, BEHAVIOR, AND IDENTITY SCIENCE. Since 2013, she has been the Program Chair of the International Conference on Pattern Recognition Applications and Methods, an IAPR-endorsed Conference. She is a Member of ACM, European Association for Biometrics (EAB), and Institute for Systems and Technologies of Information, Control and Communication (INSTICC).