IWBF 2020

PROGRAMME

	1st DAY: APRIL 29TH
8:30-9:00	**Welcome & Opening Session**
9:00-10:00	Keynote Talk **Professor Peter Eisert** *Explainable AI for Face Morphing Detection*
10:00-10:45	Oral Session I **Presentation Attack Detection**
	Coffee Break
11:00-12:30	Oral Session II **Biometric Recognition**
	Lunch
13:30-14:40	**Industry Session** *BioID, Yoonik, CardioID Technologies*
14:40-15:35	**Posters Spotlight Session**
15:35-15:45	**Demos Spotlight Session**
	Coffee Break
16:00-17:30	**Posters & Demos Session**

	2nd DAY: APRIL 30TH
8:30-9:00	**Doctoral Consortium Session**
9:00-10:00	Keynote Talk **Professor Zeno Geradts** *Forensic Aspects and the Analysis of Deep Fake Videos*
10:00-11:10	Oral Session III **Image Forensics**
	Coffee Break
11:25-12:30	Oral Session IV **Emerging Biometrics**
	Lunch
13:30-14:10	Oral Session V **Text Forensics**
14:10-15:20	Oral Session VI **Security**
	Coffee Break
15:40-16:00	**Best Paper Award & Honorable Mentions**
16:00-16:30	**Closing Session**

(ALL TIMES ARE CEST - CENTRAL EUROPEAN SUMMER TIME)

KEYNOTE TALKS

KEYNOTE TALK

EXPLAINABLE AI FOR FACE MORPHING ATTACK DETECTION

Deep learning has received an enormous interest for many data analysis tasks. In various applications, including biometrics and forensics, such methods outperform classical approaches in terms of accuracy and classification performance. However, deep learning usually provides only black box decisions, which are critical in most security and safety related applications. Here, it is desirable to know why a neural network has come to a particular decision in order to verify a decision or to modify the system in case of mis-classifications.

In this talk, state of the art methods, like layerwise relevance propagation (LRP), will be presented that enable the explanation and visualization of neural network decisions. This will be illustrated for the particular case of CNN based face morphing attack detection. It is shown that not only plausibility of decisions can be determined but also generality and attack detection performance can be improved, making a system more robust to unknown future threats.

Professor Peter Eisert
IAPR Invited Keynote Speaker

Peter Eisert is Professor for Visual Computing at the Humboldt University Berlin and heading the Vision & Imaging Technologies Department of the Fraunhofer Institute for Telecommunications - Heinrich Hertz Institute Berlin, Germany. He is also a Professor Extraordinaire at the University of Western Cape, South Africa. He received the Dipl.-Ing. degree in Electrical Engineering "with highest honors" from the Technical University of Karlsruhe, Germany, in 1995 and the Dr.-Ing. degree "with highest honors" from the University of Erlangen-Nuremberg, Germany, in 2000.

In 2001, he worked as a postdoctoral fellow at the Stanford University, USA, on 3D image analysis as well as facial animation and computer graphics. In 2002, he joined Fraunhofer HHI, where he is coordinating and initiating numerous national and international 3rd party funded research projects with a total budget of more than 15.6 Million Euros.

He has published more than 150 conference and journal papers and is Associate Editor of the International Journal of Image and Video Processing as well as in the Editorial Board of the Journal of Visual Communication and Image Representation. His research interests include 3D image/video analysis and synthesis, face and body processing, image-based rendering, computer vision, computer graphics in application areas like multimedia, security and medicine.

KEYNOTE TALK

FORENSIC ASPECTS AND THE ANALYSIS OF DEEPFAKE VIDEOS

Since the advent of AI and neural networks (especially the GAN-networks), it is easier to make a realistic deepfake video. In the early days, digital video manipulation was easy to detect, since many artefacts were visible. However, with the software that is nowadays available, also on the consumer market, it is easier to produce those so-called deepfake videos. In this presentation, we will focus on the making of deepfakes, including the technique of face morphs.

The forensic detection of deepfake is based on the detection of artefacts and discontinuities in the video, which can be accomplished manually, with a neural network or a combination of both. The Photo Non Uniformity Response (PRNU) technique is an effective way of detecting if a deepfake has been encountered. Nevertheless, the detection of deepfakes is never fully solved. Like with all the anti-spoofing methods, once the detection technique is known, it is possible to develop ways to prevent detection, for example, by spoofing the PRNU patterns.

Professor Zeno Geradts

Zeno Geradts is a senior forensic scientist at the Netherlands Forensic Institute of the Ministry of Security and Justice at the Forensic Digital Biometrics Traces departement. He is expert witness in the area of forensic (video) image processing and biometrics such as manipulation detection on deepfakes.

Within the team Forensic Big Data Analysis, Geradts works in research on artificial intelligence and images and video. He works within the European Project ASGARD on Forensic big data analysis. He is President of the American Academy of Forensic Science 2019-2020 and chairman of the ENFSI Forensic IT Working group.

Since September 1st 2014, Zeno Geradts has been a full professor on Forensic Data Science by special appointment at the University of Amsterdam for 1 day a week.

INDUSTRY SESSION

INDUSTRY SESSION

With liveness detection emerging to become a mandatory fraud prevention measure for online services, the German company BioID strongly focuses on R&D in this field. Liveness detection for facial recognition serves as an effective deterrent against online frauds ensuring session/transaction security. In terms of GDPR, it is an effective user consent mechanism (non-repudiation). CEO Ho Chang will be sharing over 20 years of technological development and presenting use cases for biometric user presence verification.

YooniK was founded in 2019 by a team of founders who decided to put their combined experience and track record in (1) Machine Learning and Biometrics, (2) Privacy by Design, (3) Travel and Hospitality Products and (4) Business Growth, to deliver the next big global tech disruption: Biometrics in Things™ (BiT). YooniK started from a future vision of immersive and seamless guest experiences. This requires the environment to recognize you and adapt to you. BiT™ technology endows connected objects with intelligence to recognize people and behave differently to provide guests with a tailored unique experience. Whatever object, whatever device, whatever process, BiT™ will make any experience YooniK. At the core, YooniK is connecting consumers to producers. With our platform, through the use of facial recognition, any object can become an immediate selling point of any producer to any consumer. Wherever you are, you may find an intelligent object and use your face and your YooniK ID to acquire any service or access.

CardioID is a technological company founded in 2014, developing innovative applications around cardiac signals. Our heterogeneous team, accumulating over 10 years of R&D experience, combines skills from the biomedical, machine learning, software, electronics, hardware manufacturing, and business development fields. The presentation will focus on how CardioID technologies has incorporated ECG biometrics on the product portfolio, in particular we will show the automotive industry use case, and the insertion of the Advanced Driver Assistance Systems (ADAS) ecosystem.

ORAL SESSIONS

ORAL SESSION I: **PRESENTATION ATTACK DETECTION**

#70

Can GAN Generated Morphs Threaten Face Recognition Systems Equally as Landmark Based Morphs? - Vulnerability and Detection

Sushma Venkatesh, Haoyu Zhang, Raghavendra Ramachandra, Kiran Raja, Naser Damer, Christoph Busch

The primary objective of face morphing is to combine face images of different data subjects (e.g. an malicious actor and an accomplice) to generate a face image that can be equally verified for both contributing data subjects. In this paper, we propose a new framework for generating face morphs using a newer Generative Adversarial Network (GAN) - StyleGAN. In contrast to earlier works, we generate realistic morphs of both high-quality and high resolution of 1024×1024 pixels. With the newly created morphing dataset of 2500 morphed face images, we pose a critical question in this work. (i) Can GAN generated morphs threaten Face Recognition Systems (FRS) equally as Landmark based morphs? *Seeking an answer, we benchmark the vulnerability of a Commercial-Off-The-Shelf FRS (COTS) and a deep learning-based FRS (ArcFace). This work also benchmarks the detection approaches for both GAN generated morphs against the landmark based morphs using established Morphing Attack Detection (MAD) schemes.*

#55

Interpretable Biometrics: Should We Rethink How Presentation Attack Detection is Evaluated?

Ana F. Sequeira, Wilson Silva, João Ribeiro Pinto, Tiago Gonçalves, Jaime S. Cardoso

Presentation attack detection (PAD) methods are commonly evaluated using metrics based on the predicted labels. This is a limitation, especially for more elusive methods based on deep learning which can freely learn the most suitable features. Though often being more accurate, these models operate as complex black boxes which makes the inner processes that sustain their predictions still baffling. Interpretability tools are now being used to delve deeper into the operation of machine learning methods, especially artificial networks, to better understand how they reach their decisions. In this paper, we make a case for the integration of interpretability tools in the evaluation of PAD. A simple model for face PAD, based on convolutional neural networks, was implemented and evaluated using both traditional metrics (APCER, BPCER and EER) and interpretability tools (Grad-CAM), using data from the ROSE Youtu video collection. The results show that interpretability tools can capture more completely the intricate behavior of the implemented model, and enable the identification of certain properties that should be verified by a PAD method that is robust, coherent, meaningful, and can adequately generalize to unseen data and attacks. One can conclude that, with further efforts devoted towards higher objectivity in interpretability, this can be the key to obtain deeper and more thorough PAD performance evaluation setups.

ORAL SESSION II: **BIOMETRIC RECOGNITION**

#19

ActGAN: Flexible and Efficient One-shot Face Reenactment

Ivan Kosarevych, Marian Petruk, Markian Kostiv, Orest Kupyn, Mykola Maksymenko, Volodymyr Budzan

This paper introduces ActGAN – a novel end-to-end generative adversarial network (GAN) for one-shot face reenactment. Given two images, the goal is to transfer the facial expression of the source actor onto a target person in a photo-realistic fashion. While existing methods require target identity to be predefined, we address this problem by introducing a "many-to-many" approach, which allows arbitrary persons both for source and target without additional retraining. To this end, we employ the Feature Pyramid Network (FPN) as a core generator building block – the first application of FPN in face reenactment, producing finer results. We also introduce a solution to preserve a person's identity between synthesized and target person by adopting the state-of-the-art approach in deep face recognition domain. The architecture readily supports reenactment in different scenarios: "many-to-many", "one-to-one", "one-to-another" in terms of expression accuracy, identity preservation, and overall image quality. We demonstrate that ActGAN achieves competitive performance against recent works concerning visual quality.

#71

ComSupResNet: A Compact Super-Resolution Network for Low-Resolution Face Images

Aashish Rai, Vishal Chudasama, Kishor Upla, Kiran Raja, Raghavendra Ramachandra, Christoph Busch

Typically in face recognition based applications, a certain degree of resolution is required for effective feature extraction and comparison. Many practical systems such as surveillance applications violate the requirement by capturing Low-Resolution (LR) face images due to wider angle of imaging or longer stand-off to the camera. Such gap in the requirement versus the practice has led numerous works to investigate approaches to super-resolve the face images that span from classical dictionary based methods to recent deep learning based approaches. In this work, we propose a compact and computationally efficient Convolutional Neural Network (CNN) to increase the spatial resolution of a LR face image to obtain a High-Resolution (HR) face image with an upscaling factor of up to ×8 which we refer as ComSupResNet. Contrary to other earlier works, the proposed architecture in the compact network comprises a progressive residual propagating asymmetrical architecture with three modules: low-frequency and high-frequency feature extraction modules and reconstruction module. In addition to designing a new architecture, we also exercise care to reduce the number of parameters to approx: 1 M as compared to similar earlier work which has more than 30 M parameters. As a second aspect, we present the generalization aspect for the learned network in a cross-database setting by training the network on CelebA dataset while evaluating it both on CelebA and LFW datasets. Through empirical evaluations, we demonstrate the gain in high fidelity reconstruction in terms of structural similarity and Peak-Signal-to-Noise Ratio (PSNR) despite the compactness of the model.

#65

Periocular Biometrics in Head-Mounted Displays: A Sample Selection Approach for Better Recognition

Fadi Boutros, Naser Damer, Kiran Raja, Raghavendra Ramachandra, Florian Kirchbuchner, Arjan Kuijper

Virtual and augmented reality technologies are increasingly used in a wide range of applications. Such technologies employ a Head Mounted Display (HMD) that typically includes an eye-facing camera and is used for eye tracking. As some of these applications require accessing or transmitting highly sensitive private information, a trusted verification of the operator's identity is needed. We investigate the use of HMD-setup to perform verification of operators using periocular regions captured from inbuilt camera. However, the uncontrolled nature of the periocular capture within the HMD results in images with a high variation in relative eye location and eye opening due to varied interactions. Therefore, we propose a new normalization scheme to align the ocular images and then, a new reference sample selection protocol to achieve higher verification accuracy. The applicability of our proposed scheme is exemplified using two handcrafted feature extraction methods and two deep learning strategies. We conclude by stating the feasibility of such a verification approach despite the uncontrolled nature of the captured ocular images, especially when proper alignment and sample selection strategy is employed.

#40

Towards adept hand-crafted features for ocular biometrics

Ritesh Vyas

This article presents a hand-crafted feature descriptor for ocular recognition, which as opposed to the deep-learning based approaches, is free from any kind of learning. The proposed approach is able to mitigate the limitations of iris recognition, such as poor iris segmentation, partial or covered iris. The proposed approach leverages the unique texture present in the periocular region, which can provide complementary details along with the iris modality, or can act as a potential stand alone trait. The proposed descriptor is evaluated on three benchmark databases, namely VISOB, CrossEyed and MICHE. Two of these databases (VISOB and MICHE) provide eye images captured through the smartphones, whereas the third database provides standard eye images registered in visible as well as near-infrared wavelengths. Hence, the evaluation reported in this article becomes a comprehensive one. The experimental results exhibit that the proposed approach proves to be suitable in challenging evaluation frameworks.

ORAL SESSION III: **IMAGE FORENSICS**

#45

Robust video source recognition in presence of motion stabilization

Pasquale Ferrara, Laurent Beslay

Video source attribution is getting a growing interest from researchers, law enforcement officers and forensic analysts. The capability of linking a video recording with its source device enables to search out who has generated a video recording. Such a feature finds immediate application in fighting against technology enabled crimes such as digital piracy and child abuse online. Currently, the most powerful techniques rely on the unique noise traces left by each camera sensor within any visual content, widely known as Photo Response Non-Uniformity. However, in the case of videos, the increasing adoption of digital motion stabilization interferes with the extraction of reliable noise patterns. In such a context, this paper describes a novel methodology for creating a robust reference video PRNU from still images for source camera recognition. Moreover, we provide a novel optimized strategy to compare two different PRNUs extracted from videos in presence of motion stabilization. The conducted experimental evaluation highlights the strength of the proposed methods.

#51

Development of a score-to-likelihood ratio model for facial recognition using authentic criminalistic data

Anna Leida Mölder, Isabelle Enlund Åstrom, Elisabet Leitet

Automated face matching systems have emerged as a useful tool for identification purposes in criminal investigations. In a forensic context it is desirable to evaluate the findings from such comparisons as probabilities in terms of a likelihood ratio. When comparing two biometric samples, many facial recognition systems produce a score value as the output. The score describes the relative similarity between the two facial images. To obtain the likelihood ratio, it is necessary to construct a statistical model for score-to-likelihood ratio conversion. The model is highly dependent on the available training data and ideally it should reflect the relevant population as closely as possible. In order to construct a general model applicable on a national level, we use data from a national mugshot database as training data. In a full crossmatch drawing from 51563 records, we develop and evaluate five different models in a Bayesian statistical framework using a total of 9000 facial comparisons with equal distribution between same source and different source scores.

#52

Use of Scene Geometry Priors for Data Association in Egocentric Views

Huiqin Chen, Emanuel Aldea, Sylvie Le Hégarat-Mascle, Vincent Despiegel

The joint use of dynamic, egocentric view cameras and of traditional overview surveillance cameras in high-risk contexts has become a promising avenue for advancing public safety and security applications, as it provides more accurate localization and finer analysis of individual interactions. However, the strong scene scale changes, occlusions and appearance variations make the egocentric data association more difficult than the standard across-views data association. To address this issue, we propose to use two independent geometric

priors and integrate them with the classic appearance cues into the objection function of the data association algorithm. Our results show that the proposed method achieves significant improvement in terms of the association accuracy. We highlight the attractive use of geometric priors in across-views data association and its potential for supporting pedestrian tracking in this context.

ORAL SESSION IV: **EMERGING BIOMETRICS**

#46

Which Ear Regions Contribute to Identification and to Gender Classification?

Di Meng, Sasan Mahmoodi, Mark S. Nixon

Previous studies in biometrics have shown how gender can be determined from images of ears for recognition, but without specificity. In this paper, we use model-based analysis and deep learning methods for gender classification from ear images. We use these methods to determine the differences between female and male ears. We confirm the identification performance and then the gender discrimination before analyzing which ear parts contribute most to performance. To this end, we compare the heatmaps of different genders with identification heatmaps. It appears from the heatmaps that ears encode females and males differently and we show how this can lead to successful gender discrimination and to increase insight into the process of identification of people by their ears. This could lead to gender identification in surveillance imagery, even when the face is concealed and provides a potential focus for future gender research.

#61

Bilateral Symmetry in Central Retinal Blood Vessels

Sangeeta Biswas, Johan Rohdin, Martin Drahansky

Symmetry can be defined as uniformity, equivalence or exact similarity of two parts divided along an axis. While our left and right eyes clearly have a high degree of external bilateral symmetry, it is less obvious to what degree they have internal bilateral symmetry. This is especially true for central retinal blood vessels (CRBVs) which are responsible for supplying blood to retinas and also can be used as a strong biometric. In this paper, we study whether CRBVs of the left and right retinas possess strong enough bilateral symmetry so that we reliably tell whether a pair of CRBVs of the left and right retinas belongs to a single person. We evaluate and analyse the performance of both human and neural network based bilateral CRBVs verification. By experimenting on a large publicly available data set, we confirm that CRBVs have bilateral symmetry to some extent.

#21

Contactless Finger Knuckle Authentication under Severe Pose Deformations

Ajay Kumar

Contactless biometrics identification using finger knuckle images has shown significant potential for the e-business and forensic applications. One of the key challenges in accurately matching the real-world contactless finger knuckle images is related to the knuckle pattern deformations that are involuntarily generated due to finger pose changes. Earlier work in this area therefore acquired fixed pose finger images for the authentication and therefore the performance achieved from such images cannot reflect the expected performance under the deployment scenarios. This paper adopts a new approach to accurately match such finger knuckle images and presents the first attempt to authenticate finger-knuckle patterns under severe pose changes. This approach attempts to correct pose related deformations by identifying the correspondence between a fixed number of chosen points between two matched images. The match score is computed using

local feature descriptors, at each of these correspondence points, and consolidated to generate average match score. The experimental results are presented in this paper, both using two-session and single-session index finger knuckle images from 221 different subjects, using a publicly available database. These results are outperforming and indicate the merit of spatial-domain approach to match deformed finger knuckle images using a fixed number of correspondence points.

ORAL SESSION V: **TEXT FORENSICS**

#42

Ensemble Method for Sexual Predators Identification in Online Chats

Muhammad Ali Fauzi, Patrick Bours

Cyber grooming is a compelling problem worldwide nowadays and many reports strongly suggest that it becomes very urgent to tackle this problem to protect the children from sexual exploitation. In this study, we propose an effective method for sexual predator identification in online chats based on two-stage classification. The purpose of the first stage is to distinguish predatory conversations from the normal ones while the second stage aims to tell apart between the predator user and the victim within a single predatory conversation. Finally, some unique predators are derived from the second stage result. We investigate several machine learning classifiers including Naive Bayes, Support Vector Machine, Neural Network, Logistic Regression, Random Forest, K-Nearest Neighbors, and Decision Tree with Bag of Words features using several different term weighting methods for this task. We also proposed two ensemble techniques to improve the classification task. The experiment results on PAN12 dataset show that our best method using soft voting based ensemble for first stage and Naïve Bayes based method for the second stage obtained an F0:5-score of 0.9348, which would place as number one in the PAN12 competition ranking.

#23

Forensic Authorship Analysis of Microblogging Texts Using *N*-Grams and Stylometric Features

Nicole Mariah Sharon Belvisi, Naveed Muhammad, Fernando Alonso-Fernandez

In recent years, messages and text posted on the Internet are used in criminal investigations. Unfortunately, the authorship of many of them remains unknown. In some channels, the problem of establishing authorship may be even harder, since the length of digital texts is limited to a certain number of characters. In this work, we aim at identifying authors of tweet messages, which are limited to 280 characters. We evaluate popular features employed traditionally in authorship attribution which capture properties of the writing style at different levels. We use for our experiments a self-captured database of 40 users, with 120 to 200 tweets per user. Results using this small set are promising, with the different features providing a classification accuracy between 92% and 98.5%. These results are competitive in comparison to existing studies which employ short texts such as tweets or SMS.

ORAL SESSION VI: **SECURITY**

#35

Secure Triplet Loss for End-to-End Deep Biometrics

João Ribeiro Pinto, Jaime S. Cardoso, Miguel V. Correia

Although deep learning is being widely adopted for every topic in pattern recognition, its use for secure and cancelable biometrics is currently reserved for feature extraction and biometric data preprocessing, limiting achievable performance. In this paper, we propose a novel formulation of the triplet loss methodology, designated as secure triplet loss, that enables biometric template cancelability with end-to-end convolutional neural networks, using easily changeable keys. Trained and evaluated for electrocardiogram-based biometrics, the network revealed easy to optimize using the modified triplet loss and achieved superior performance when compared with the state-of-the-art (10.63% equal error rate with data from 918 subjects of the UofTDB database). Additionally, it ensured biometric template security and effective template cancelability. Although further efforts are needed to avoid template linkability, the proposed secure triplet loss shows promise in template cancelability and non-invertibility for biometric recognition while taking advantage of the full power of convolutional neural networks.

#44

Efficient Fingervein Sample Image Encryption

Sanjay Shekhawat, Heinz Hofbauer, Bernhard Prommegger, Andreas Uhl

Efficient sample encryption techniques are investigated for fingervein data. We propose an approach where it suffices to encrypt 0.5% of the sample JPEG2000 bitstream and thereby completely disable biometric recognition. Evaluations with 5 different recognition schemes on two different datasets reveal that results are stable across all techniques considered as long as the start of the bitstream is encrypted.

#05

Security Assessment of Partially Encrypted Visual Data: Using Iris Recognition as Generic Measure

Martin Rieger, Jutta Hämmerle-Uhl, Andreas Uhl

Security assessment of partially encrypted visual data is known to be difficult. We show that a set of known image quality measures turn out not to be good predictors for encryption strength, especially if a different extent of recognisability is present in the data. Iris recognition applied to encrypted sample data is proposed to assess the protection strength of the employed encryption scheme. When choosing the settings as identified in this work, iris recognition performance turns out to be a viable predictor for security of encrypted data, in a more consistent manner compared to image quality measures.

POSTER SESSION

POSTER SESSION

#03

Template Protection on Multiple Facial Biometrics in the Signal Domain under Visible and Near-Infrared Light

Simon Kirchgasser, Luca Debiasi, Rudolf Schraml, Heinz Hofbauer, Andreas Uhl, Jonathan Boyle, James Ferryman

Template protection techniques like cancellable biometrics have been introduced in order to overcome privacy issues in biometric applications. We conduct an ISO/IEC Standard 24745 compliant assessment of block re-mapping and warping focusing on recognition performance issues as well as security and unlinkability aspects. Both of these template protection schemes are applied on a multi-biometrics dataset in the signal (image) domain. The dataset includes 2D face, iris and periocular images which have been acquired not only using visual light (VIS) but also near-infrared light (NIR). With respect to the used data, this is the first study that applies and evaluates cancellable template protection methods in the signal domain on VIS/NIR 2D face, iris and periocular biometrics.

#06

Analysing a Vein Liveness Detection Scheme

Thomas Herzog, Andreas Uhl

We examine a previously published liveness detection method for guarding against presentation attacks on vein recognition systems which employs motion magnification on video frames, and develop three new attacks that circumvent the proposed protective scheme. The first pair of attacks are direct attacks or presentation attacks, and involve presenting a fake sample with rhythmic motions to the biometric system. The third attack is an indirect attack that feeds the biometric system a synthetic video signal designed to circumvent the liveness detection scheme. Results show that the analysed liveness detection system must not be used as a standalone technique. We conclude by recommending improvements to the analysed scheme to harden against attacks of the kind we presented and to avoid having to combine it with other presentation attack detection techniques.

#10

Advanced Multi-Perspective Enrolment in Finger Vein Recognition

Bernhard Promegger, Andreas Uhl

Finger vein recognition deals with the recognition of subjects based on their venous pattern within the fingers. It has been shown that its recognition accuracy heavily depends on a good alignment of the acquired samples. There are several approaches that try to reduce the impact of finger misplacement. However, none of these approaches is able to prevent all possible types of finger misplacements. As finger vein scanners are evolving towards contact-less acquisition, alignment problems, especially due to longitudinal finger rotation, are becoming even more important. Along with rotation detection and correction, capturing the vein pattern from multiple perspectives, as e.g. in multiple-perspective enrolment (MPE), is a way to tackle the problem of longitudinal finger rotation. Involving multiple cameras increases cost and complexity of the capturing devices,

and therefore their number should be kept to a minimum. Perspective multiplication for MPE (PM-MPE) successfully reduces the number of cameras needed during enrolment while keeping the recognition rates at a high level. So far, (PM-)MPE has only been applied using Maximum curvature features (MC). This work analyses further approaches to improve the their recognition rates and investigates the applicability of (PM-)MPE to recognition schemes using features other than MC.

#11

Spatial-Temporal Omni-Scale Feature Learning for Person Re-Identification

Aida Pločo, Andrea Macarulla Rodriguez, Zeno Geradts

State-of-the-art person re-identification (ReID) models use Convolutional Neural Networks (CNN) for feature extraction and comparison. Often these models fail to recognize all the intra- and inter-class variations that emerge in person ReID, making it harder to discriminate between data subjects. In this paper we seek to reduce these problems and improve performance by combining two state-of-the-art models. We use the Omni-Scale Network (OSNet) as our CNN to test the Market1501 and DukeMTMC-ReID datasets for person ReID. To fully utilize the potential of these datasets, we apply the spatial-temporal constraint which extracts the camera ID and timestamp from each image to form a distribution. We combine these two methods to create a hybrid model titled Spatial-Temporal OmniScale Network (st-OSNet). Our model attains a Rank-1 (R1) accuracy of 98.2% and mean average precision (mAP) of 92.7% for the Market1501 dataset. For the DukeMTMC-reID dataset our model achieves 94.3% R1 and 86.1% mAP, hereby surpassing the results of OSNet by a large margin for both datasets (94.3%, 86.4%, 88.4%, 76.1%, respectively).

#22

Detecting Deepfakes with Metric Learning

Akash Kumar, Arnav Bhavsar, Rajesh Verma

With the arrival of several face-swapping applications such as FaceApp, SnapChat, MixBooth, FaceBlender *and many more, the authenticity of digital media content is hanging on a very loose thread. On social media platforms, videos are widely circulated often at a high compression factor. In this work, we analyze several deep learning approaches in the context of deepfakes classification in high compression scenarios and demonstrate that a proposed approach based on metric learning can be very effective in performing such a classification. Using less number of frames per video to assess its realism, the metric learning approach using a triplet network architecture proves to be fruitful. It learns to enhance the feature space distance between the cluster of real and fake videos embedding vectors. We validated our approaches on two datasets to analyze the behavior in different environments. We achieved a state-of-the-art AUC score of 99.2% on the Celeb-DF dataset and accuracy of 90.71% on a highly compressed Neural Texture dataset. Our approach is especially helpful on social media platforms where data compression is inevitable.*

#28

Comparison-Level Mitigation of Ethnic Bias in Face Recognition

Philipp Terhörst, Mai Ly Tran, Naser Damer, Florian Kirchbuchner, Arjan Kuijper

Current face recognition systems achieve high performance on several benchmark tests. Despite this progress, recent works showed that these systems are strongly biased against demographic sub-groups. Previous works

introduced approaches that aim at learning less biased representations. However, applying these approaches in real applications requires a complete replacement of the templates in the database. This replacement procedure further requires that a face image of each enrolled individual is stored as well. In this work, we propose the first bias-mitigating solution that works on the comparison-level of a biometric system. We propose a fairness-driven neural network classifier for the comparison of two biometric templates to replace the systems similarity function. This fair classifier is trained with a novel penalization term in the loss function to introduce the criteria of group and individual fairness to the decision process. This penalization term forces the score distributions of different ethnicities to be similar, leading to a reduction of the intra-ethnic performance differences. Experiments were conducted on two publicly available datasets and evaluated the performance of four different ethnicities. The results showed that for both fairness criteria, our proposed approach is able to significantly reduce the ethnic bias, while it preserves a high recognition ability. Our model, build on individual fairness, achieves bias reduction rate between 15.35% and 52.67%. In contrast to previous work, our solution is easy to integrate into existing systems by simply replacing the systems similarity functions with our fair template comparison approach.

#32

Vulnerability Assessment and Detection of Makeup Presentation Attacks

Christian Rathgeb, Pawel Drozdowski, Daniel Fischer, Christoph Busch

The accuracy of face recognition systems can be negatively affected by facial cosmetics which have the ability to substantially alter the facial appearance. Recently, it was shown that makeup can also be abused to launch so-called makeup presentation attacks. In such attacks, an attacker might apply heavy makeup to achieve the facial appearance of a target subject for the purpose of impersonation. In this work, we assess the vulnerability of a widely used open-source face recognition system, i.e. ArcFace, to makeup presentation attacks using the publicly available Makeup Induced Face Spoofing (MIFS) and FRGCv2 databases. It is shown that the success rate of makeup presentation attacks in the MIFS database has negligible impact on the security of the face recognition system. Further, we employ image warping to simulate improved makeup presentation attacks which reveal a significantly higher success rate. Moreover, we propose a makeup attack detection scheme which compares face depth data with face depth reconstructions obtained from RGB images of potential makeup presentation attacks. Significant variations between the two sources of information indicate facial shape alterations induced by strong use of makeup, i.e. potential makeup presentation attacks. Conceptual experiments on the MIFS database confirm the soundness of the presented approach.

#39

A Method for Curation of Web-Scraped Face Image Datasets

Kai Zhang, Vítor Albiero, Kevin W. Bowyer

Web-scraped, in-the-wild datasets have become the norm in face recognition research. The numbers of subjects and images acquired in web-scraped datasets are usually very large, with number of images on the millions scale. A variety of issues occur when collecting a dataset in-the-wild, including images with the wrong identity label, duplicate images, duplicate subjects and variation in quality. With the number of images being in the millions, a manual cleaning procedure is not feasible. But fully automated methods used to date result in a less-than-ideal level of clean dataset. We propose a semi-automated method, where the goal is to have a clean dataset for testing face recognition methods, with similar quality across men and women, to support comparison of accuracy across gender. Our approach removes near-duplicate images, merges duplicate subjects, corrects mislabeled images, and removes images outside a defined range of pose and quality. We conduct the curation on the Asian Face Dataset (AFD) and VGGFace2 test dataset. The experiments show that a state-of-the-art

method achieves a much higher accuracy on the datasets after they are curated. Finally, we release our cleaned versions of both datasets to the research community.

#53

Deep Learning Based Stress Prediction From Offline Signatures

Hakan Yekta Yatbaz, Meryem Erbilek

Soft-Biometric measurements are now increasingly adopted as a robust means of determining individual's nonunique characteristics with the emerging models that are widely used in the deep learning domain. This approach is clearly valuable in a variety of scenarios, specially those relating to forensics. In this study, we specifically focus on stress emotion, and propose automatic stress prediction technique from offline signature biometrics using well-known deep learning architectures such as AlexNet, ResNet and DenseNet. Due to the limited number of research that study emotion prediction from offline handwritten signatures with deep learning methods, best to our knowledge this is the first experimental study that presents empirical achievable prediction accuracy around 77%.

#58

Human Emotion Distribution Learning from Face Images using CNN and LBC Features

Abeer Almowallad, Victor Sanchez

Human emotion recognition from facial expressions depicted in images is an active area of research particularly for medical, security and human-computer interaction applications. Since there is no pure emotion, measuring the intensity of several possible emotions depicted in a facial expression image is a challenging task. Previous studies have dealt with this challenge by using label-distribution learning (LDL) and focusing on optimizing a conditional probability function that attempts to reduce the relative entropy of the predicted distribution with respect to the target distribution, which leads to a lack of generality of the model. In this work, we propose a deep learning framework for LDL that uses convolutional neural network (CNN) features to increase the generalization of the trained model. Our framework, which we call EDL-LBCNN, enhances the features extracted by CNNs by incorporating a local binary convolutional (LBC) layer to acquire texture information from the face images. We evaluate our EDL-LBCNN framework on the s-JAFFE dataset. Our experimental results show that the EDLLBCNN framework can effectively deal with LDL for human emotion recognition and attain a stronger performance than that of state-of-the-art methods.

DOCTORAL CONSORTIUM SESSION

DOCTORAL CONSORTIUM SESSION

Biometrics as forensic evidence: some reflections from the Italian Criminal proceeding's point of view

Ernestina Sacchetto

Biometric science is not free of errors, especially considering its typical statistical – probabilistic nature. The systematic use of biometric technologies in the criminal proceeding could lead to some issues in terms of the reliability of the results from their application and the compatibility between the discipline in question, constitutional principles and typical procedural guarantees. However, because of the speed of the recent technology development, the criminal proceeding seems not to be able to do without the contribution offered by biometric science and the reason is that both the science and the trial, even with different approaches, have the different objective of the reconstruction of the causal connection. Biometric data have to be analyzed and evaluated in terms of scientific accuracy and it is necessary to understand their value case by case. The most crucial risks for forensic science in its recent digital dimension need to have legislative clarification aimed at standardising the discipline, also in terms of interpretability of technology, distinguishing one biometric data from another, implementing different applications since the specificity of each data element requires different legal solutions. Otherwise, the danger is that the procedural guarantees could be weakened (such as the principle of contradictory, the principle of reasonable duration of trial and the right of defence) in addition to some issues related to individual guarantees such as the right to privacy.

DEMO SESSION

DEMO SESSION

#D1

Multimodal biometrics in vehicles: a simulator study

Pedro Costa, André Lourenço, Carlos Carreiras, David Velez, Pedro Mendes Jorge, Arnaldo Abrantes

This demo presents a simulator that combines different state-of-the-art technologies for in-vehicle environment biometric recognition. The platform integrates Intel RealSense RGB-Depth camera for facial analysis and recognition, together with heart biometrics from CardioID Technologies CardioWheel. Additionally, it allows to compare the driver physiological signals (face and heart monitoring) with behavioral data extracted from the act of driving. The demo enables the interaction of known users, registration of new drivers and identification or authentication in both scenarios. Biometrics are used for initial identification (to access and start the vehicle, as for further personalization of vehicle settings) and for continuous authentication, for security purposes.

#D2

Face Completion Demonstrator

Philipp Terhörst, Jonas Henry Grebe, Naser Damer, Fadi Boutros, Florian Kirchbuchner, Arjan Kuijper

Recent face recognition systems works well on unconcealed faces. However, recognizing faces containing occlusions is considerably more challenging. To expand the ability of facial recognition systems to deal with partially concealed pictures our generative deep learning model with image-to-image translation offers the possibility of inpainting the concealed part of the face. It allows us to remodel any obfuscated part of the face and thus, facial recognition systems to work.
