ACADEMIC GRADUATION MONITORING REPORT

2022

European Association for Biometrics (EAB)
version: 2023-09-16
Email: secretariat@eab.org

RESEARCH MONITOR CONTENT

PREFACE

During the 2019 EAB general assembly it was proposed to compose an annual academic graduation monitoring report, which should provide information about academic theses that are completed in EAB member institutions.

Such report should contain lists of entries of Bachelor-, Master- or PhD-theses and a short summary of each thesis.

EAB is proud to provide now an overview of the research going on in Europe. If you are member of EAB and you can contribute information about your graduated students. In order to facilitate the data collection, a webform, accessible to EAB members, has been added to the EAB website, in which author and contact information can be provided as well as a title, and abstract and an optional link to the report. The webform can be found here: https://eab.org/information/academic_report.html

This report was composed by the European Association for Biometrics (EAB) for its members. If you are not EAB member yet – please join and share the non-profit spirit of EAB. We are grateful for your continuous support of the EAB initiatives through your membership.

MONITOR
PHD-THESES

JOÃO RIBEIRO PINTO - SEAMLESS MULTIMODAL BIOMETRICS FOR CONTINUOUS PERSONALISED WELLBEING MONITORING

Full Title: Seamless Multimodal Biometrics for Continuous Personalised Wellbeing Monitoring
Institution: Universidade do Porto, Portugal
Supervisor: Jaime S. Cardoso; Miguel Velhote Correia
URL: https://jtrpinto.github.io/files/pdf/thesis_jtrpinto_2022_vt1.pdf
Link description: Thesis Full-Text
Contact email: jtrpinto@gmail.com

Abstract:

Artificially intelligent perception is increasingly present in the lives of every one of us. Vehicles are no exception, as advanced driver assistance systems (ADAS) help us comply with speed limits, keep within the lanes, and avoid accidents. In the near future, pattern recognition will have an even stronger role in vehicles, as self-driving cars will require automated ways to understand what is happening around (and within) them and act accordingly. Within pattern recognition, biometrics offer promising applications in vehicles, from keyless access control to the automatic personalisation of driving and environmental conditions based on the recognised driver. Similarly, wellbeing monitoring technologies have long attracted attention to the possibility of recognising activity, emotions, sleepiness, or stress from drivers and passengers. However, these two topics are starkly opposed, since wellbeing recognition relies on intrasubject variability while biometrics thrives on intersubject variability. Despite their differences, biometric recognition and wellbeing monitoring could (and should) coexist. Continuous identity recognition from seamlessly acquired data could be used to personalise wellbeing monitoring models and attain improved performance. These personalised models could be the key to more robust ways of monitoring drivers' drowsiness and attention and avoiding accidents. In a broader sense, they could be applied to all vehicle occupants, paving the way towards the accurate recognition of activity, emotions, comfort, and even violence episodes in shared autonomous vehicles. This doctoral work focused on advancing in-vehicle sensing through the research of novel computer vision and pattern recognition methodologies for both biometrics and wellbeing monitoring. The main focus has been on electrocardiogram (ECG) biometrics, a trait well-known for its potential for seamless driver monitoring. Major efforts were devoted to achieving improved performance in identification and identity verification in off-the-person scenarios, well-known for increased noise and variability. Here, end-to-end deep learning ECG biometric solutions were proposed and important topics were addressed such as cross-database and long-term performance, waveform relevance through explainability, and interlead conversion. Face biometrics, a natural complement to the ECG in seamless unconstrained scenarios, was also studied in this work. The open challenges of masked face recognition and interpretability in biometrics were tackled in an effort to evolve towards algorithms that are more transparent, trustworthy, and robust to significant occlusions. Within the topic of wellbeing monitoring, improved solutions to multimodal emotion recognition in groups of people and activity/violence recognition in in-vehicle scenarios were proposed. At last, we also proposed a novel way to learn template security within end-to-end models, dismissing additional separate encryption processes, and a self-supervised learning approach tailored to sequential data, in order to ensure data security and optimal performance. Following the results of this work, one can conclude that truly personalised wellbeing is yet to be achieved. However, this work has built a strong framework to support future work towards the goal of integrating biometric recognition and wellbeing monitoring in a multimodal, seamless, continuous, and realistic way. Overall, this doctoral work led to numerous contributions to biometrics and wellbeing monitoring in general, resulting directly in twenty-four scientific publications in major biometrics and pattern recognition venues. Its quality and impact have been recognised by the scientific community with over three hundred citations and multiple awards, including the EAB Max Snijder Award 2022.

LAZARO JANIER GONZALEZ-SOLER - GENERALISABLE PRESENTATION ATTACK DETECTION FOR MULTIPLE TYPES OF BIOMETRIC CHARACTERISTICS

Full Title: Generalisable Presentation Attack Detection for Multiple Types of Biometric Characteristics
Institution: Hochschule Darmstadt
Supervisor: Prof. Dr. Christoph Busch, Prof. Dr. Marta Gomez-Barrero, Prof. Dr. Andreas Heinemann
Contact email: lazaro-janier.gonzalez-soler@h-da.de

Abstract:

Biometric systems have experienced a large development in recent years since they are accurate, secure, and in many cases, more user convenient than traditional credential-based access control systems. In spite of their benefits, biometric systems are still vulnerable to attack presentations (APs), which can be easily launched by a fraudulent subject without having a wide expert knowledge. This way, he/she can gain access to several applications, such as bank accounts and smartphone unlocking, where biometric systems are frequently deployed. In order to mitigate such threats and increase the security of biometric systems, the development of reliable Presentation Attack Detection (PAD) algorithms is of utmost importance to the research community. In the context of PAD, we explore in this Thesis different strategies and methods in order to improve the generalisation capability of PAD schemes. To that end, we propose the definition of a semantic common feature space which successfully discriminates bona fide presentations (BPs) from APs. In essence, this process is seeking for those significant features extracted from known-attacks samples that are observed in unknown-attacks. In addition, we explore several handcrafted techniques in order to build a reliable description of features per biometric characteristics studied. The experimental evaluation shows that a common feature space can be computed through the fusion between generative models and discriminative approaches. Remarkable detection performances for high-security thresholds lead to the construction of a convenient (i.e., low BP rejection rates) and secure (i.e., low AP acceptance rates) PAD subsystem.

SUSHMA VENKATESH - MORPHING ATTACK DETECTION

Full Title: Robust Algorithms For Face Morphing Attack Detection: Database, Vulnerability and Detection
Institution: NTNU
Supervisor: Christoph Busch and Kiran Raja
Link description: https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/3022013
Contact email: christoph.busch@ntnu.no

Abstract:

Biometrics has emerged as a promising technology for automated recognition of individuals. The stored biometric characteristics are used to recognise an individual and hence biometrics technology plays a major role in security-related applications and stands in the front-line for authentication of data subjects. Biometrics has a wide range of applications in law enforcement, surveillance, banking, border control, medical records, time and attendance tracking etc. As the inherent biometrics characteristics don’t undergo any changes, biometric technology has shown the best performance for authentication of data subjects. Though biometrics technology is promising for person authentication, attackers may employ various techniques like presentation attacks and adversarial attacks to impersonate an enrolled individual with interest to obtain unauthorised access to the system. In addition to these attacks, in relatively recent times, biometric are attacked in the facial image enrolment stage, especially in the ID related applications, by performing face morphing. Facial morphing is initially performed for entertainment purposes, but gradually it has been used to attack face recognition systems. The face morphing process combines two different facial identities to generate a single facial image with the facial representations of both identities. An attacker may use the morphed facial image to enrol it in the ID documents (like driving license, passport). Since the morphed facial image shows a high resemblance to both facial identities, the ID document can be claimed by both identities. This indicates the severity of facial morphing and the necessity of morphing attack detection mechanisms to avoid the security lapse. Hence the primary objective of this thesis is the development of face morphing attack detection techniques using hand-crafted and deep learning approaches. During this doctoral work, morphing attack detection approaches are developed for both digital and print-scan datasets. To empirically evaluate the performance of the newly generated morphing attack detection approaches, various face morphing databases are generated using landmark and deep learning-based GAN techniques. Furthermore, the vulnerability of face recognition systems to face morphing attacks with ageing co-variate is evaluated. To this extent, this doctoral thesis contributes with the novel morphing attack detection approaches.

MARCO SANTOPIETRO - AN EXPLORATION OF DYNAMIC BIOMETRIC PERFORMANCE USING DEVICE INTERACTION AND WEARABLE TECHNOLOGIES

Full Title: An exploration of dynamic biometric performance using device interaction and wearable technologies
Institution: University of Kent
Supervisor: Richard Guest
URL: https://kar.kent.ac.uk/98627/1/79santopietro2022phdfinal.pdf
Contact email: r.m.guest@kent.ac.uk

Abstract:

With the growth of mobile technologies and internet transactions, privacy issues and identity check became a hot topic in the past decades. Mobile biometrics provided a new level of security in addition to passwords and PIN, with a multitude of modalities to authenticate subjects. In this thesis we explore the verification performance of behavioural biometric modalities, as previous studies in literature proved them to be effective in identifying individual behaviours and guarantee robust continuous authentication. The scope of this project is to assess the performance and stability of authentication models for mobile and wearable devices, with ceremony based tasks and a framework that includes behavioural and electrocardiogram biometrics. The results from our experiments suggest that a fast verification, applicable to real life scenarios (e.g. login or transaction request), with a single sample request and the considered modalities (Swipe gestures, PIN dynamics and electrocardiogram recording) can be performed with a stable performance. In addition, our novel fusion method implemented greatly reduced the authentication error. We are confident that our findings presented in this thesis will contribute to the enhancement of identity verification on mobile and wearable technologies.

MATTHEW BOAKES - A PERFORMANCE ASSESSMENT FRAMEWORK FOR MOBILE BIOMETRICS

Full Title: A Performance Assessment Framework for Mobile Biometrics
Institution: University of Kent
Supervisor: Richard Guest; Farzin Deravi
URL: https://kar.kent.ac.uk/97792
Link description: Kent Academic Repository
Contact email: Matthew.Boakes@Gmail.com

Abstract:

This project aims to develop and explore a robust framework for assessing biometric systems on mobile platforms, where data is often collected in non-constrained, potentially challenging environments. The framework enables the performance assessment given a particular platform, biometric modality, usage environment, user base and required security level. The ubiquity of mobile devices such as smartphones and tablets has increased access to Internet-based services across various scenarios and environments. Citizens use mobile platforms for an ever-expanding set of services and interactions, often transferring personal information, and conducting financial transactions. Accurate identity authentication for physical access to the device and service is, therefore, critical to ensure the security of the individual, information, and transaction. Biometrics provides an established alternative to conventional authentication methods. Mobile devices offer considerable opportunities to utilise biometric data from an enhanced range of sensors alongside temporal information on the use of the device itself. For example, cameras and dedicated fingerprint devices can capture front-line physiological biometric samples (already used for device log-on applications and payment authorisation schemes such as Apple Pay) alongside voice capture using conventional microphones. Understanding the performance of these biometric modalities is critical to assessing suitability for deployment. Providing a robust performance and security assessment given a set of deployment variables is critical to ensure appropriate security and accuracy. Conventional biometrics testing is typically performed in controlled, constrained environments that fail to encapsulate mobile systems' daily (and developing) use. This thesis aims to develop an understanding of biometric performance on mobile devices. The impact of different mobile platforms, and the range of environmental conditions in use, on biometrics' accuracy, usability, security, and utility is poorly understood. This project will also examine the application and performance of mobile biometrics when in motion.

FADI BOUTROS - EFFICIENT AND HIGH PERFORMING BIOMETRICS: TOWARDS ENABLING RECOGNITION IN EMBEDDED DOMAINS

Full Title: Efficient and High Performing Biometrics: Towards Enabling Recognition in Embedded Domains
Institution: Fraunhofer Institute for Computer Graphics Research - IGD
Supervisor: Arjan Kuijper
URL: https://tuprints.ulb.tu-darmstadt.de/21571/
Contact email: fadi.boutros@igd.fraunhofer.de

Abstract:

The growing need for reliable and accurate recognition solutions along with the recent innovations in deep learning methodologies has reshaped the research landscape of biometric recognition. Developing efficient biometric solutions is essential to minimize the required computational costs, especially when deployed on embedded and low-end devices. This drives the main contributions of this work, aiming at enabling wide application range of biometric technologies. Towards enabling wider implementation of face recognition in use cases that are extremely limited by computational complexity constraints, this thesis presents a set of efficient models for accurate face verification, namely MixFaceNets. With a focus on automated network architecture design, this thesis is the first to utilize neural architecture search to successfully develop a family of lightweight face-specific architectures, namely PocketNets. Additionally, this thesis proposes a novel training paradigm based on knowledge distillation (KD), the multi-step KD, to enhance the verification performance of compact models. Towards enhancing face recognition accuracy, this thesis presents a novel margin-penalty softmax loss, ElasticFace, that relaxes the restriction of having a single fixed penalty margin. Occluded faces by facial masks during the recent COVID-19 pandemic presents an emerging challenge for face recognition. This thesis presents a solution that mitigates the effects of wearing a mask and improves masked face recognition performance. This solution operates on top of existing face recognition models and thus avoids the high cost of retraining existing face recognition models or deploying a separate solution for masked face recognition. Aiming at introducing biometric recognition to novel embedded domains, this thesis is the first to propose leveraging the existing hardware of head-mounted displays for identity verification of the users of virtual and augmented reality applications. This is additionally supported by proposing a compact ocular segmentation solution as a part of an iris and periocular recognition pipeline. Furthermore, an identity-preserving synthetic ocular image generation approach is designed to mitigate potential privacy concerns related to the accessibility to real biometric data and facilitate the further development of biometric recognition in new domains.

MATTHEW BOAKES - A PERFORMANCE ASSESSMENT FRAMEWORK FOR MOBILE BIOMETRICS

Full Title: A Performance Assessment Framework for Mobile Biometrics
Institution: University of Kent
Supervisor: Richard Guest
URL: https://kar.kent.ac.uk/97792/1/72boakes2022phdfinal_.pdf
Contact email: r.m.guest@kent.ac.uk

Abstract:
This project aims to develop and explore a robust framework for assessing biometric systems on mobile platforms, where data is often collected in non-constrained, potentially challenging environments. The framework enables the performance assessment given a particular platform, biometric modality, usage environment, user-base and required security level. The ubiquity of mobile devices such as smartphones and tablets has increased access to Internet-based services across various scenarios and environments. Citizens use mobile platforms for an ever-expanding set of services and interactions, often transferring personal information, and conducting financial transactions. Accurate identity authentication for physical access to the device and service is therefore critical to ensure the security of the individual, information, and transaction. Biometrics provide an established alternative to conventional authentication methods. Mobile devices offer considerable opportunities to utilise biometric data from an enhanced range of sensors, alongside temporal information on the use of the device itself. For example, cameras and dedicated fingerprint devices can capture front-line physiological biometric samples (already used for device log-on applications and payment authorisation schemes such as Apple Pay) alongside voice capture using conventional microphones. Understanding the performance of these biometric modalities is critical to assessing suitability for deployment. Providing a robust performance and security assessment given a set of deployment variables is critical to ensure appropriate security and accuracy. Conventional biometrics testing is typically performed in controlled, constrained environments that fail to encapsulate mobile systems' daily (and developing) use. This thesis aims to develop an understanding of biometric performance on mobile devices. The impact of different mobile platforms, and the range of environmental conditions in use, on the accuracy, usability, security, and utility of biometrics is poorly understood. This project will also examine the application and performance of mobile biometrics when in motion.

MONITOR
MASTER-THESES

ROMAN KESSLER - FACE MORPHING

Full Title: Analysis of face embeddings to facilitate image pre-selection for face morphing
Institution: Hochschule Darmstadt and NTNU
Supervisor: Christoph Busch, Juan Tapia and Kiran Raja
Contact email: christoph.busch@h-da.de

Abstract:

Face Morphing Attacks pose a novel threat to the security of identification documents. The fusion of the face images of two or more – similarly looking – individuals during the application process for a new travel document (i.e., passport) or identity card enables both individuals to travel with the same document. In order to develop algorithms to detect morphing attacks, large data sets of morphed face images are needed, for which in turn many similarly looking individuals need to be paired. The study at hand uses face embeddings of openly accessible face recognition models to describe similarity between individuals. It aims at finding appropriate face recognition models, metrics to quantify similarity, morphing algorithms to fuse facial images of paired individuals, and soft biometric characteristics to analyze the attack potential of face morphs. Results demonstrate, that image pre-selection based on Cosine or Euclidean distances between face embeddings highly improves the attack potential of morphs. Especially ArcFace and MagFace provide valuable face embeddings to quantify similarity for pre-selection. Both open source, as well as Commercial Off-The-Shelf Face Recognition Systems get fooled by morphed faces. Landmark-based, closed source morphing algorithms pose high risk for any of the tested Face Recognition Systems. On the other hand, MagFace embeddings further emerge as valuable means to detect morphed face images. Soft biometrics characteristics however were only partially relevant to predict morph success, if morphing has been conducted within similar age, gender, and race groups. The results emphasize that face embeddings are valuable instruments on both sides of the morphing attack, image pre-selection for face morphing and detection of morphed faces.

XINYU TIAN - RECONSTRUCTION-BASED ANOMALY DETECTION WITH MACHINE LEARNING

Full Title: Reconstruction-based Anomaly Detection with Machine Learning for High Throughput Scanning Electron Microscope Defect Inspection
Institution: Universiteit Twente - Zilverling Service Desk
Supervisor: Luuk Spreeuwers
URL: https://essay.utwente.nl/92906/
Link description: Reconstruction-based Anomaly Detection with Machine Learning for High Throughput Scanning Electron Microscope Defect Inspection
Contact email: l.j.spreeuwers@utwente.nl

Abstract:

With rapid advancement of high throughput wafer inspection systems, there are growing demands on more efficient and accurate defect inspection algorithms. This work focuses on investigating the feasibility of applying learning based algorithms to the image processing pipeline of the next generation wafer inspection tools. An end-to-end reconstruction-based anomaly detection methodology is proposed with two types of reconstruction models based on Principal Component Analysis and Convolutional Autoencoder. Both methods have demonstrated competitive performance compared to the current defect detection algorithm on multiple datasets with different patterns and various defect types.

WASSIM KABBANI - REAL FACES IN THE LATENT SPACE

Full Title: Real Face Image Embedding into StyleGAN2 Latent Space
Institution: DTU and NTNU
Supervisor: Christoph Busch and Marcel Grimmer
Contact email: christoph.busch@ntnu.no

Abstract:

Biometric systems are vital components in our current security infrastructure. Face recognition systems, in particular, have been receiving a lot of interest recently due to their convenience and unobtrusive nature. However, due to various challenges in the operational environments of these systems, the conditions under which samples are obtained during the verification process are often different from those present when obtaining reference face templates. To increase the variability in the stored reference templates, the individual is typically asked to provide multiple samples during the registration process. Since this is not always feasible or sufficient, an automated and controllable method of introducing variability in the dataset of an existing biometric system or in a training dataset that can be used to train algorithms used in the biometric recognition process is needed. With the purpose of exploring such methods, state-of-the-art approaches for translating real face images into the latent space of Generative Adversarial Network models are investigated. These networks are capable of generating highly realistic synthetic face images that are non-distinguishable from real face images, and embedding real-face images into their latent space will allow the manipulation of different attributes of these faces. Therefore, introducing a tool to increase the variability of the reference images in an automated and controlled way is needed.

DAVIDE GHIANI - FEATURE ENGINEERING FOR THE DEVELOPMENT OF AN EXPLAINABLE ANOMALY DETECTION SYSTEM IN CROWD ANALYSIS

Full Title: Feature engineering for the development of an explainable anomaly detection system in crowd analysis
Institution: University of Cagliari
Supervisor: Gian Luca Marcialis
Contact email: marcialis@unica.it

Abstract:

Nowadays, we face problems due to the progressive overpopulation of cities where events like demonstrations, festivals, parades, or other sorts of people gathering can raise several security issues. For safety, it is necessary to involve security forces to monitor the crowds through CCTV systems. However, this task is not easy and requires high attention levels for extended periods. To facilitate it, many smart cities are adopting intelligent surveillance systems, which require the application of advanced techniques of crowd analysis. It is easy to imagine a significant advantage in terms of efficiency if it could suggest an operator monitor a specific scene where the system has noticed a threat. Our work fits into this open issue by studying crowd dynamics to signal the occurrence of potentially dangerous events for the safety of individuals, placing it in the anomaly detection field. We start from the work in [15], which aims to detect anomalies using a threshold on the central bin of descriptive histograms, which are related to the groups of people counted in a window comprehensive of a limited number of frames over the video. This mechanism is based on the description of the rapidity of groups' aggregation and disintegration present in a given scenario, which can effectively describe the level of panic according to our working hypothesis. This method assumes that, under normal conditions, movements of individuals between groups or between the filmed context and the area outside, and mainly their speed, follow a general trend from which any possible panic scenario deviates. However, the limit of this system is that the threshold used to determine whether the description of a video window is suspicious should be adapted from video to video. No threshold can generalize the best behavior. Indeed, several factors can affect the description of the videos and differentiate the cases, from noise and disturbances to the big difference in the dynamics of some scenarios. This implies a phase of choosing the best threshold upstream of the detection process, otherwise, the results will not be satisfactory. For this reason, in the present work, we decided to start with a statistical analysis of the system's output described in [15]. By collecting the histograms in sets divided by clustering algorithm and camera angle, we studied the occurrence of values assumed by the central bin, obtaining more clarity on the results exposed in [15]. We then set ourselves the problem of how to go beyond this issue. It was decided to introduce a machine learning module and train it using the feature for extended or at least enhancing more elements of it. This is because going beyond the mechanism proposed in [15], we hypothesized that the information content contributed by the extended feature could be greater than the descriptive potential of the central bin alone. However, we did not use the feature provided by the trit-based temporal descriptor as it was. To provide the models to be trained with the best feature, we analyzed the activation frequency of each bin on a set-by set basis and then selected only the most representative bins as feature components. We noticed that specific bins were recurrently among the most active ones regardless of the set. We paid more attention to these and tried to interpret their meaning at the description level of the dynamics. With this analysis, we achieved a reduction in feature size, greater clarity in the representation of the feature space, and significantly better results. This machine learning procedure is carried on by training an explainable classifier, a support vector machine, and a neural network. The proposed system takes as input of the different classifiers the same feature, a description of a surveillance camera video obtained with the descriptor discussed in [15], and returns a binary classification that aims to distinguish between normal and abnormal situations in the scene. Given the need for a machine learning system, we also dealt with structuring the data, especially their labeling. Unfortunately, there is no state-of-the-art guideline on performing the labeling procedure since different choices lead the system to satisfy different tasks. We labeled the dataset functionally for obtaining training on a balanced number of samples between normal and abnormal examples. Furthermore, the labeling method highlights the anomalous event's consequence, beyond the strict task of detecting the anomaly trigger.

JOÃO ISIDORO - MORPHING ATTACKS TO FACE RECOGNITION SYSTEMS

Full Title: Morphing Attacks to Face Recognition Systems
Institution: Instituto Superior Tecnico - Universidade de Lisboa, Portugal
Supervisor: Paulo Lobato Correia
URL: https://fenix.tecnico.ulisboa.pt/cursos/meec21/dissertacao/846778572214598
Contact email: paulo.lobato.correia@tecnico.ulisboa.pt

Abstract:

With the widespread usage of face recognition systems they can be targets of various types of attacks, such as presentation attacks, where the reproduction of an image or video is presented to a camera, or morphing attacks, where facial images from more than one person are merged together so that the resulting image resembles the intervening individuals faces. This Thesis addresses the problem of detecting face morphing attacks, notably when two facial images are morphed together and the result is used to issue an official identification document. The Thesis proposes a de-morphing solution, using a neural network, attempting to reconstruct the face of the person that is not present at an Automatic Border Control (ABC) system and whose image was used to create the morphed image included in the identification document being verified. The development of the solution highlighted the need of having a representative set of images in the database used for training the system, and discusses solutions to improve the quality of the training data. The de-morphing system has shown an interesting performance, allowing to detect the occurrence of face morphing attacks, and the recovered face can be used to identify the accomplice (or victim of identity theft) that contributed to the morphed face included in the ID document.

PIA CAVASINNI DI BENEDETTO - PGAN TO REDUCE FACE RECOGNITION BIAS

Full Title: Using PGAN to create synthetic face images to reduce bias in biometric systems
Institution: Sapienza University of Rome
Supervisor: Maria De Marsico
Contact email: demarsico@di.uniroma1.it

Abstract:

Balancing classes in datasets of images of unbalanced people's faces is an important problem to be solved both ethically and biometrically, as underrepresented classes could be cultural minorities in the world. Recent studies in deep generative networks have given the possibility to generate synthetic images starting from visual attributes. This study proposes a progressive GAN (generative adversarial network) model, called A2F_P, which gives the possibility to create synthetic images from specific attributes, such as age, gender, and ethnicity, with the aim of balancing an unbalanced dataset of facial images. A convolutional neural network for age, gender and ethnicity detection is the mean to do experiments to understand the benefit of balancing. Two CNN-based models will be analyzed and compared: Model_Unbal, trained on unbalanced data, and Model_Bal trained on balanced data. Thanks to this comparison, it is possible to notice the improvement of the metrics for the balanced model.

VID KRIŽNAR - DETECTION OF GENERATIVE ADVERSARIAL NETWORK ARTEFACTS AS AN AID FOR DETECTING DEEPFAKES

Full Title: Detection of generative adversarial network artefacts as an aid for detecting deepfakes
Institution: Faculty of Computer and Information Science, University of Ljubljana
Supervisor: Borut Batagelj, Peter Peer, Vitomir Štruc
Contact email: borut.batagelj@fri.uni-lj.si

Abstract:
We present a novel approach to deepfake detection. Deepfake is a type of media, usually a picture or video, in which a part of the picture, most frequently face or body, has been digitally modified. Deepfakes are often used with ill intentions, such as spreading misinformation or opinion formulation. Modification of digital media usually leaves traces, a so-called digital artefacts. Artefacts can be defined as irregularities in digital media, which are unwanted consequences of modification. We present five methods for detecting deepfakes by detecting artefacts of generative adversarial networks. We evaluate the presented methods on seven different deepfake databases, which are further divided into those that are primarily generated by a generative adversarial network and those that are not. We show that the presented methods achieve promising results on the prepared databases.

ALVILS STŪRE - 3D FACE IMAGE GENERATION

Full Title: 3D Face Image Generation using GANs
Institution: DTU and NTNU
Supervisor: Christoph Busch and Marcel Grimmer
Contact email: christoph.busch@ntnu.no

Abstract:

The availability of biometric data has been a crucial factor in the success of any state of the art recognition system. Often, datasets are collected primarily for biometric recognition and not utilized for further processing, such as training and evaluation of biometric systems. Gathering and labelling qualitative and discriminative data is a time consuming and often prohibitively expensive process, and with the introduction of the European General Data Protection Law (GDPR), especially in the field of biometrics, acquisition, sharing, and usage of the data are highly restricted. Human faces are attractive biometric modalities due to their uniqueness and nonintrusiveness. Nowadays, many real world applications replace traditional authentication mechanisms with face recognition systems in order to increase overall security and trustworthiness. Most of the research on face recognition is performed using 2D images, but in order to improve overall face recognition performance in a wide variety of the inthe wild conditions, additional data needs to be acquired, which in most cases can be an expensive and time consuming process. The recent development of generative adversarial networks (GANs) has proven to be a potential alternative to acquiring a large amount of biometric data for a fraction of the time and cost compared to a manual collection. Synthetic data usage is widely used for autonomous vehicles, healthcare, biometric systems, and other industries, as it provides a way to generate new data from the same sample space, including face data, that can comply with GDPR restrictions. In the thesis, we propose a pipeline for generating controllable 2D head samples from 3D Morphable Models (3DMMs) with different head poses, head shapes and expressions, textures, camera, and lighting conditions. Additionally, the domain gap between the generated synthetic data and real data is evaluated for some cases of how the face recognition system performs on different poses and lighting conditions within the image in comparison with real data in order to see if synthetic data can be applied in finding these edge cases where face recognition system may give inaccurate or less accurate results in face verification.

DARIAN TOMAŠEVIĆ - GENERATING OCULAR IMAGES

Full Title: Generating ocular images with deep generative models
Institution: University of Ljubljana, Slovenia
Supervisor: Peter Peer and Vitomir Štruc
URL: https://repozitorij.uni-lj.si/Dokument.php?id=160862&lang=eng
Link description: The PDF is available for the thesis
Contact email: vitomir.struc@fe.uni-lj.si

Abstract:

Most modern segmentation techniques for ocular images are based on deep learning methods and are thus critically dependent on large-scale annotated datasets. Unfortunately, suitable datasets are labour-intensive to gather and often raise privacy concerns. To address these issues, we present a novel framework, called BiOcularGAN, capable of generating large-scale synthetic datasets of photorealistic ocular images, in both the visible and the near-infrared light spectrum, along with corresponding segmentation masks. The framework is centered around an innovative Dual-Branch StyleGAN2 model, which facilitates the generation of high-quality aligned bimodal images. By exploiting latent features of the model, the framework is also able to produce extremely accurate segmentation masks of the synthetic images, based only on a handful of manually labeled examples, therefore minimizing human involvement. We evaluate the BiOcularGAN framework through extensive experiments across five diverse ocular datasets and analyze how bimodal data generation affects the quality of produced images and masks. In addition, we showcase that the generated data can be employed to train highly successful deep segmentation models, which can generalize well to other real-world datasets.

X. NAME UPON REQUEST - AR FACE FILTERS AND THE GDPR: LOST IN DEFINITIONS

Full Title: AR Face Filters and the GDPR: Lost in definitions
Institution: Leiden Law School, Advanced Studies Program in Law and Digital Technology degree
Supervisor: E. J. Kindt
Contact email: els.kindt@kuleuven.be

Abstract:

Main findings This thesis demonstrates that AR face filter systems’ functionality is in many ways similar to an FRS, yet is not precisely the same. Based on the author’s knowledge, no attempts have been made by any scholars to differentiate between the two by going through the relevant patents, describing the steps in detail, and comparing the two. Such distinction between the two systems can assist the regulators in understanding how to treat this novel technology by taking inspiration from its very similar counterpart -FRSs. Furthermore, this project scrutinizes the quality of the GDPR’s Article 9 wording through the lens of technology neutrality by assessing the relevant privacy and security risks that could arise due to AR face filters’ processing. A significant finding based on such risk analysis is that AR face filter systems could profile individuals with a precision that cannot be found in biometric systems that use only one biometric characteristic as their input due to incorporating multiple modalities of biometrics. The analysis of Article 9(1) offered in this thesis project differs from the previous scholarly works in this field as it goes beyond the identification-verification-categorization understanding of the functions of biometric systems and introduces modification as a new function that biometric technologies could perform.

MARK LUKEK - FACE HALLUCINATION WITH DUAL AUTOENCODERS

Full Title: Face hallucination with dual autoencoders and coupled latent spaces
Institution: University of Ljubljana, Slovenia
Supervisor: Vitomir Štruc
URL: https://repozitorij.uni-lj.si/Dokument.php?id=155731&lang=slv
Link description: The PDF is available for the thesis - in Slovene
Contact email: vitomir.struc@fe.uni-lj.si

Abstract:

Recently, convolutional models based on neural networks have achieved great success in super-resolution using a single input image, called Single-Image-SuperResolution or SISR. Such models are very flexible and efficient in non-linear mapping of low resolution images to high-resolution ones. In this work, we present a novel super-resolution procedure based on two autoencoders and coupled latent spaces. The first autoencoder is capable of reconstructing low-resolution images, while the second one is capable of reconstructing high-resolution images. The latent spaces of the two autoencoders are connected by a linking network, which allows for the conversion between the low- and high- resolution latent spaces. Using the low-resolution encoder, the linking network and the high-resolution decoder, it is possible to efficiently upscale an arbitrary low-resolution input image. The results of the above method area tested on four datasets, CASIA-WebFace, LFW, QMUL-TinyFace and QMUL-SurvFace. Part of the CASIA-WebFace database was used to train all models, the rest for testing. The QMUL-TinyFace and QMUL-SurvFace databases are used to verify the system performance on real images, where we do not have low- and high-resolution pairs. Finally, the results of the super-resolution model are further compared with existing approaches such as bicubic interpolation, SRCNN and SRGAN. When frontal face images are used as input, our approach outperforms bicubic interpolation and the SRCNN model. The faces are more pronounced and smoother, but contain less high-resolution details than faces produced by SRGAN.

EVANGELOS GOUGOULIS DIMITRIADIS - SIXTHSENSE : OUTDOOR COLLISION AVOIDANCE ASSISTANT FOR THE BLIND AND VISUALLY IMPAIRED.

Full Title: Sixthsense : outdoor collision avoidance assistant for the blind and visually impaired.
Institution: Universiteit Twente - Zilverling Service Desk
Supervisor: Luuk Spreeuwers
URL: https://essay.utwente.nl/93532/
Link description: Sixthsense : outdoor collision avoidance assistant for the blind and visually impaired.
Contact email: l.j.spreeuwers@utwente.nl

Abstract:

The present work aims to provide a novel outdoor collision assistant for the blind and visually impaired (BVI) population. Regardless of the rapid technological advancements of recent years, current BVI navigation and collision avoidance aids are still in a primitive stage. Our main goal is to create a system that can warn BVI users about nearby obstacles in a reliable and intuitive manner. To address this challenge, we investigate the integration of cutting-edge object detection and depth data in a small and mobile form factor. For this reason, a compact but powerful passive stereo depth camera (OAK-D) is used. In an attempt to limit the overwhelming quantity of information that previous navigation aid systems pass to the user, we explore the prioritization of collision alerts. More specifically, obstacle warnings are prioritized based on their estimated collision time with the BVI. Communication with the user is achieved via a smart vest, which allows him to continue using other widely-adopted navigation tools like the white cane. The smart vest comprises of a 3x3 grid of vibration actuators spread in its back area, indicating this way the location of approaching obstacles. Additionally, various obstacle proximity levels can be conveyed through the vibration intensity of the haptic feedback. For convenience, the camera hardware is mounted in the front area of the smart vest. The results of testing the developed system in a variety of situations suggest its ability to provide reliable outdoor collision assistance to the blind and visually impaired.

HAMZA ALI - ENHANCING GENERALIZABILITY OF FACE PRESENTATION ATTACK DETECTION

Full Title: Intra-identity PatchSwap: On the Generalizability of Face Presentation Attack Detection
Institution: Fraunhofer Institute for Computer Graphics Research (FraunhoferIGD)
Supervisor: Meiling Fang
Contact email: hamza.ale@gmail.com

Abstract:

With the widespread deployment of face recognition systems, face presentation attack detection (PAD) plays an essential role in mitigating their vulnerabilities. Face PAD is employed before the identification system to detect if the presented face is an attack. However, most of the existing face PAD methods tend to overfit on the training data and fail to generalize well on unknown attacks in a real-world scenario. The main reason for such poor generalizability is that existing face PAD datasets are limited in quantity and diversity. Moreover, recent PAD works leverage pixel-wise supervision strategy and show great progress in face PAD performance. Nevertheless, obtaining accurate pixel- wise labels is a challenging task. To alleviate these issues, we propose the plug-n-play PatchSwap approach in this thesis. The proposed PatchSwap method maximizes limited data utilization and generates more challenging bonafide/attack samples and partial attacks by swapping patches between training data by a well-designed strategy. Meanwhile, their pixel-wise labels are correspondingly updated. As a result, the augmented training samples contain more complex attack patterns, benefiting robust feature learning. Furthermore, we demonstrated the proposed PatchSwap method combined with three prevailing backbones: ResNet, DenseNet, and MixFaceNet. The extensive experiments were performed on four benchmark datasets under both intra-dataset and cross-dataset scenarios. We also conducted several detailed ablation studies to explore the effect of patch types, selected candidate identity, and the probabilities controlling the swapping process. The experimental results show that the proposed PatchSwap approach achieved significant performance improvement. For example, the ACER value on the most challenging Protocol-4 of Oulu-NPU decreased from 20.51% achieved by DenseNet baseline to 3.41% by DenseNet-PatchSwap.

YU LINGHU - OPEN-SET FACE RECOGNITION WITH ENTROPIC OPEN-SET LOSS

Full Title: Open-Set Face Recognition with Entropic Open-Set Loss
Institution: University of Zurich
Supervisor: Manuel Günther
URL: https://www.merlin.uzh.ch/publication/show/22304
Contact email: siebenkopf@googlemail.com

Abstract:

The goal for the open-set face recognition is to identify the unseen subjects and do not assign them to any known subject with high confidence. There are two types of subjects involved in the task: the ones that we are interested in and have labels, i.e. known subjects; the ones that we do not care about and have no labels (we use −1 in the experiment instead), i.e. unknown subjects. We build a complete face recognition pipeline through Bob. ArcFace R100 network, as a feature extractor, has a good performance on the IJB-C dataset. Our goal is to add an extra network after ArcFace to enhance its power on open-set face recognition tasks. We attempt three cases: first, the unknown subjects have never appeared in the training; second, the unknown subjects appear in both training and testing; third, the unknown subjects only appear in the testing, and they are replaced by the adversarial samples generated from the knowns in the training. The training unknowns have no overlap with the testing unknowns in case three. Plain softmax loss and entropic open-set loss are applied to the first two cases, respectively, and objectosphere loss is used for the second and third cases. We prove that those models create a high True Positive Identification Rate especially when the False Positive Identification Rate is small. Replacing the unknown subjects in case two to the adversarial samples as in case three is successful without performance degradation. One flaw is that the magnitude separation property of the entropic open-set loss and objectosphere loss is not apparent. When working with the adversarial samples, the situation is worse.

BJØRN IVAR NIELSEN - CONTINUOUS AUTHENTICATION ON AN SSH CONNECTION

Full Title: Continuous Authentication on an SSH Connection
Institution: NTNU
Supervisor: Patrick Bours
URL: https://ntnuopen.ntnu.no/ntnu-xmlui/discover
Contact email: patrick.bours@ntnu.no

Abstract:

With a shift to more remote-based work, that is only accelerated by the COVID19 pandemic, new ways of ensuring the identity of users of IT systems are important. The classical approach of username and password for authentication is vulnerable to stolen credentials. By stealing credentials of a user, an adversary can act on the system as the rightful owner of the credentials. Continuous authentication can be implemented to increase the chance of discovering an intruder, and secures the system by revoking access to the account suspected of not being controlled by the owner. The goal of this thesis has been to investigate whether analysis of keystroke dynamics on keystroke data captured at the server side of an SSH session can be used to identify the rightful owner of an account. A publicly available data set has been used as the source of data for our testing, and eBPF has been used to extract the decrypted SSH traffic on the server. The functionality of a commonly used distance measure has been compared on both sides of an SSH session. Different network stress has been applied to investigate the impact of network-introduced interference. A stability of the functionality of keystroke dynamics has been observed at the server side in normal network behavior, whereas a network under stress shows signs of heavily impacting the functionality of keystroke dynamics on the captured data of the SSH channel. With these observations we can state that it is possible to conduct continuous authentication on data capered on the server side of an SSH channel, but in unstable network condition the process will experience degradation.

NOAH CHAVANNES - MULTI-TARGET ADVERSARIAL ATTACKS WITH LOTS

Full Title: Multi-Target Adversarial Attacks with LOTS
Institution: University of Zurich
Supervisor: Manuel Günther
URL: https://www.merlin.uzh.ch/publication/show/22468
Contact email: siebenkopf@googlemail.com

Abstract:

Face recognition systems are on the rise and are being widely used throughout the industry. With the advance of face recognition systems, more and more adversarial attacks are emerging. Layerwise Origin-Target Synthesis is one such attack in which the image of a source person is iteratively modified so that a face recognition system identifies it as another person. We extend this approach by allowing one input image to mimic multiple targets simultaneously. We further improve the loss function of the approach by including additional components that measure the structural similarity between the original image and the adversarial image. We evaluate our new method quantitatively with experiments and conduct an empirical analysis with 73 participants to investigate the relationship between human perception and similarity metrics. Our results show that we can successfully perform multi-target attacks and keep perturbations minimal. We also show how different source-target constellations affect the quality of adversarial images. Lastly, we demonstrate that the similarity metrics used to measure the size of perturbations are not perfect predictors of human perception.

MURIEL VAN DER SPEK - VASCULAR BIOMETRIC IMAGING PROCEDURE.

Full Title: Understanding and Modelling the Vascular Biometric Imaging Procedure.
Institution: Universiteit Twente
Supervisor: Luuk Spreeuwers
URL: https://essay.utwente.nl/91672/
Link description: Understanding and Modelling the Vascular Biometric Imaging Procedure.
Contact email: l.j.spreeuwers@utwente.nl

Abstract:

In finger vascular biometric images, captured with near-infrared (NIR) light, several structures are visible, from which the exact origin is unknown. These include the vagueness, width and intensity of the vessel projection, the brightness of the two joints and the intensity of the dark area between the joints. These features vary per subject. To understand the origin of these elements, the imaging procedure is mimicked using a simplified mathematical model of the finger. This model creates images similar to real finger vascular images incorporating basic anatomy and corresponding optical properties. Using the model, the origin of several features is examined. The vessels appear vague, because the projection is actually a shadow caused by the strong scattering of the bone. The intensity of the finger (besides the vessels) is directly dependent on both tissue consistency (amount of absorption/scattering) and finger anatomy (path length of the photons). This research gives an insight on the vascular imaging procedure and this knowledge can be used in future research on vascular biometric identification, by incorporating additional features from the images.

SARA TRAMONTE - MULTIMODAL EMOTION RECOGNITION WITH A 3-INPUT CNN

Full Title: Multimodal emotion recognition with a 3-input CNN
Institution: Sapienza University of Rome
Supervisor: Maria De Marsico
Contact email: demarsico@di.uniroma1.it

Abstract:

In this thesis work, the aim is to perform emotion recognition using a multimodal approach that exploits convolutional neural networks with more than one input. Multimodal approaches allow different modalities to cooperate in order to achieve generally better performances because different features are extracted from different pieces of information. In this work, from videos, the facial frames, the optical flow computed from consecutive facial frames and the MEL spectrograms are extracted from videos and combined together in different ways to understand which modality combination works better. For this purpose, Python’s libraries Keras and Tensorflow are used, to create a model that is able to extract relevant spatio-temporal features from the video frames, and relevant features from the MEL spectrograms computed from the video audios, which then are used to train the model. Several experiments are run on the models by first taking one modality at a time into consideration and after that good accuracy results are found on each modality, the models are concatenated to create a final model that allows multiple inputs. The performances of the proposed models are shown through accuracy results and confusion matrices, demonstrating that the 3-input CNN is more accurate than every CNN with a single kind of input. For the experiments the datasets used are BAUM-1 and RAVDESS, which both collect two distinguished sets of videos based on the different intensity of the expression, that is acted or spontaneous, providing the representations of the following emotional states that will be taken into consideration: angry, disgust, fearful, happy and sad.

MELISSA TIJINK - FUSING FORENSIC FEATURES AND A FACE RECOGNITIONSYSTEM ON LOOKALIKE FACES.

Full Title: Fusing Forensic Features and a Face RecognitionSystem on Lookalike Faces.
Institution: Universiteit Twente - Zilverling Service Desk
Supervisor: Luuk Spreeuwers
URL: https://essay.utwente.nl/94250/
Link description: Fusing Forensic Features and a Face RecognitionSystem on Lookalike Faces.
Contact email: l.j.spreeuwers@utwente.nl

Abstract:

Face Recognition Systems are popular and widelyused, however their performance on challenging cases can stillbe improved. One of the challenges are lookalikes, which aresubjects who look similar, but have a different identity. In thisresearch the dependence of the score computation of an existingFace Recognition System on face regions will be analyzed. Byoccluding parts of the face and visualizing the change in score asheatmaps, the cases of mated, (random) non-mated and lookalikepairs can be compared. The heatmaps show that the regions ofthe eye(brows) and nose are important for mated and lookalikepairs. The next step is to investigate whether the performanceof the Face Recognition System can be improved by fusing itsoutput with forensic features. The idea is to force the wholesystem to pay attention to details in important facial regions.A methodology is presented to automatically retrieve forensicfeatures from an image. Several fusing strategies are comparedwith a focus on the case of lookalikes. Results show that, althoughoverall performance is not significantly improved, comparableresults with a face recognition system are reached and severalfused systems show potential on individual lookalike cases.

SEBASTIAN BUNDA - LIMITED RESOURCE OPTIMIZATION FOR FACE RECOGNITION NEURAL NETWORKS: SUB-BYTE QUANTIZATION OF MOBILEFACENET USING QKERAS.

Full Title: Limited Resource Optimization for Face Recognition Neural Networks: Sub-byte quantization of MobileFaceNet using QKeras.
Institution: Universiteit Twente - Zilverling Service Desk
Supervisor: Luuk Spreeuwers
URL: https://essay.utwente.nl/90930/
Link description: Limited Resource Optimization for Face Recognition Neural Networks: Sub-byte quantization of MobileFaceNet using QKeras.
Contact email: l.j.spreeuwers@utwente.nl

Abstract:

Face recognition is one of the most populair biometric identification systems and as such is widely used. With the growing need for digital personal data security, it is crucial to seek solutions to work on personal devices. To stimulate these developments, the computational and memory footprint of these face recognition systems should be reduced to fit on edge devices. Based on the populair MobileNetV2, MobileFaceNet is a very efficient face recognition neural network with 99.15% accuracy on the LFW dataset with a model size of only 4MB using a 32-bit representation. This work presents a method to reduce the bit length of MobileFaceNet in the form of QMobileFaceNet using sub-byte quantization. This is achieved by first identifying the most strategic use of the QKeras library enabling sub-byte dynamic fixed-point quantization. This work shows that 8-bit and 4-bit versions of QMobileFaceNet can be obtained with 98.68% and 98.63% accuracy on the LFW dataset which reduces footprint to 25% and 12.5% of the original weight respectively. Both show an accuracy loss similar to the performance described by other quantization methods applied on MobileNetV2. Using mixed-precision, an accuracy of 98.17% can be achieved whilst requiring only 10% of the original weight footprint.

ROMANO FERLA - EXPLORING THE GANFORMER FOR FACE GENERATION : INVESTIGATING THE SEGMENTATION AND SMILE AUGMENTATION POTENTIAL

Full Title: Exploring the GANformer for Face Generation : investigating the segmentation and smile augmentation potential
Institution: Universiteit Twente - Zilverling Service Desk
Supervisor: Luuk Spreeuwers
URL: https://essay.utwente.nl/90496/
Link description: Exploring the GANformer for Face Generation : investigating the segmentation and smile augmentation potential
Contact email: l.j.spreeuwers@utwente.nl

Abstract:

Advancing the research in face applications is limited by proprietary databases and increasing data protection regulations, synthetically generated databases may provide a solution. In this work the GANformer, a hybrid generative image model, is explored for this application. While only trained for unconditioned face generation like many other models, this works shows the potential of two use cases. First, the unique implementation of the attention is examined for the application of segmentation. Results indicate segmenting behaviour is present, though post-processing is needed before its implementation in synthetic databases. Second, real labeled faces are reconstructed in latent space to find latent directions describing disentangled attributes. This concept is brought in practice by augmenting neutral to smiling faces, but could be applied on other expressions and attributes as well. In both the segmentation and the smile augmentation the results indicate that the GANformer is able to be used for multiple applications in synthetic database generation. This work can be use as basis as it opens up two directions for further research.

XINYI ZHANG - FACIAL VIDEO RECOGNITION VIA 3D CONVOLUTIONAL NETWORKS

Full Title: Facial Video Recognition via 3D Convolutional Networks
Institution: University of Zurich
Supervisor: Manuel Günther
URL: https://www.merlin.uzh.ch/publication/show/22314
Contact email: siebenkopf@googlemail.com

Abstract:

Face recognition has been popular in the video recently. As the development of deep learning, various CNNs models are implemented into face recognition such as ResNet, MobileNet, Mo- bileFaceNet. During this experiment, we verified that the light CNN model – Stacked2D, and 3D MobileFaceNet can extract features from several frames at the same time on the video dataset (YoutubeFaces). First, the baseline model – the original 2D MobileFaceNet combined ArcFace loss function model is trained from the face recognition task. Then, this model is implemented as the feature extractor in bob framework, which can construct a face recognition pipeline easily. Using the same process, the Stacked2D and 3D MobileFaceNet models with Arcface are trained using YTF dataset. In the end, we run the video recognition pipeline in bob framework and com- pare the results using different models. In this experiment, we verify that it is feasible to use 2D, Stacked2D, and 3D MobileFaceNet models in video face recognition, and the model with larger frames input can perform better because it can capture more spatial and temporal information from video data.

GIANPAOLO PERELLI - A DEEPFAKE DETECTOR IN THE WILD

Full Title: A Deepfake Detector In The Wild
Institution: University of Cagliari
Supervisor: Gian Luca Marcialis & Giovanni Puglisi
Contact email: marcialis@unica.it

Abstract:
Deepfake is a rapidly growing phenomenon that worries the whole globe because, unfortunately, there are always those who use technological progress incorrectly. The news reports more and more cases of "deep fake abuse" in which artificial intelligence is used improperly, for example, to modify pornographic images or videos by inserting the faces of unwitting women, to make false statements that can be used for political purposes or for defaming public figures. Thanks to various tools in circulation, generating deep fake has become simple and easily accessible, increasing the need to reach an even better level of deep fake recognition. Many state-of-the-art methods achieve excellent results on certain datasets. Still, often, when the test images transform in terms of size or jpeg compression, the accuracy of the results quickly degrades. It is very important to have a robust classifier for these transformations as they represent the typical operations that images undergo before circulating on the world wide web. This Thesis project presents the study of a method capable of extracting particularly discriminating information in the frequency range through the use of the discrete cosine transform. The limits of the method placed under certain conditions and the way to overcome them are faced.

GIANLUCA MASELLI - MULTIMODAL PERSONALITY RECOGNITION

Full Title: Bi-modal Automatic Personality Recognition using Deep Learning
Institution: Sapienza University of Rome
Supervisor: Maria De Marsico
Contact email: demarsico@di.uniroma1.it

Abstract:

Especially in the field of biometric systems, distinguishable human traits are widely studied. Among them, we have both physical and behavioral traits making a subject unique with respect to others. Well-known psychical traits are fingerprints, face, iris, hand geometry and DNA. By contrast, behavioral traits are the ones that depend on the person’s behavior pattern. The class of identifiers under the so-called "behaviormetrics" are related to gait, keystroke dynamics, signature and behavioural profiling. Particularly interesting is the concept of behavioral pattern which is related to repetitive actions, tasks, or behaviors taken by a subject without thinking about them. This regard both toxic and non-toxic behavioral patterns. In fact, also toxic behavioral patterns are done automatically and are likely to continue due to the associated reward. The problem with behavioral patterns is related to the difficulty in changing them since they are sometimes strictly correlated to a subject’s personality. For example, anxious people tend to smoke cigarettes many times a day. It is then important to recognize toxic behaviors beforehand in order to find a way to correct them. More than anything, the relationship between personality and health behaviors has been studied in order to identify those personality qualities which could lead to risky behaviors in young people. Most of the research relied on the Five-factor model of personality to link the broad personality traits to risky behaviors. It is then intuitive how a specific personality trait could affect the behavior of a subject with respect to the others in early and subsequent stages of life. This can include several aspects such as changes in lifestyle, jobs, relationships, and so on. In the present work, we exploit the concept of personality traits in order to implement a deep learning system aiming to automatically detect and infer personality in video clips where subjects speak toward the camera. We truly believe that extending this research could help humans in their daily lives with several real-world applications.

NOUR ELDIN ALAA BADR - REPRESENTATIVE FACE PAD

Full Title: Momentum Contrast for Representative Face Presentation Attack Detection
Institution: Fraunhofer Institute for Computer Graphics Research (FraunhoferIGD)
Supervisor: Meiling Fang
Contact email: noureldinalaa93@gmail.com

Abstract:

With the widespread usage of using face recognition systems, they became vulnerable to presentation attacks encountered by attackers. To tackle this issue, face presentation attack detection (PAD) methods are implemented. However, these methods have several shortcomings including the generalizability of unknown attacks. This thesis targets two main problems that face PAD methods. The first problem that this work target to solve is databases annotation problems. Annotating databases with labels is time-consuming, to solve this problem, a representative learning model (MoCo framework in this thesis) is used as it focuses on unsupervised learn- ing databases. The second problem that this work target is the insufficient PAD data. Most PAD databases are manually collected especially presentation attack samples, thus they are labor-intensive and small-scale. This thesis target this problem by training the model on a face recognition database such as CASIA-Web database which is a very large-scale public facial recognition database, not a PAD database, which is collected randomly in the wild where images are diverse from illumination, sensors, identity. This work proves that using face recognition databases to learn face representation, can be adapted to be used in detecting presentation attacks and the model can benefit from using extra existing face recognition data besides the model becomes more familiar with diverse setups and illuminations within face images. Finally, the classification model suggested by the state-of-art MoCo, is extended by applying pseudo labeling to it, which improved the general results.

JESPER BANG - CONTACTLESS FINGERPRINT PRESENTATION ATTACK DETECTION

Full Title: Ensuring Security in Biometrics - Presentation Attack Detection for Contactless Fingerprint Recognition
Institution: DTU & h_da / ATHENE
Supervisor: Jascha Kolberg & Jannis Priesnitz
Contact email: jascha.kolberg@h-da.de

Abstract:

In the current modern world, biometrics has become an integrated part of people’s daily lives. These systems provide fast and secure identification and verification for, e.g., mobile devices, removing the need for slower knowledge-based authentication. As the usage of biometric systems for recognition increase, so does the attacks against them. Therefore, Presentation Attack Detection methods are essential to help fight spoofing attempts and ensure system security. This work utilizes multiple versions of Convolutional Neural Networks and an Autoencoder to identify attack presentations and seeks to improve these with hyperparameter tuning and added processing steps. Such steps involve obtaining relevant regions of interest from the images by masking and gaining information for classification using image Power Spectrum Energy (PSE). This work also proposes different fusions of presentation attack detection methods. These included combinations of some of the best performing versions of CNNs, the AE and the PSE. At a fixed Bona Fide Presentation Classification Error Rate (BPCER) of 0.2%, the fusions performed from 0% Attack Presentation Classification Error Rate (APCER) with Detection Equal Error Rate (D-EER) of 0% up to 10% APCER with 4% D-EER. The default performance of the autoencoder was 95.83% APCER on the contactless fingerprints. This work improved performance of the autoencoder significantly, and through model fusions even greater improvements were achieved. Most of the fusions were also suitable for a mobile scenario where an operator would be hand-holding the capture device.

AHMAD FOROUGHI - AVATAR, CARICATURE AND SKETCH FACE RECOGNITION

Full Title: Avatar, Caricature and Sketch Face Recognition
Institution: Hochschule Darmstadt
Supervisor: Christian Rathgeb & Christoph Busch
URL: https://dasec.h-da.de/hda-cdfdb/
Contact email: christian.rathgeb@h-da.de

Abstract:

The accuracy of face recognition skyrocketed in past years and its robustness towards various covariates has been shown, such as variations in pose or age. In recent years, a considerable amount of research efforts has been devoted to cross-domain face recognition aiming at comparing facial images obtained from domains that are different in capture technologies, e.g. visible spectrum versus infrared, or signal representations, e.g. photographs versus sketches. Yet, various relevant domains have hardly been explored and a lack of public databases hampers the development of new algorithms.In this work, we introduce the HDA Cross-Domain (HDA-CD) face image database comprising 1,400 face images from three different domains including avatars, caricatures, and sketches. Said face images were manually generated using popular mobile apps. In a benchmark, we evaluate commercial and open-source state-of-the-art facial analysis methods on the HDA-CD database including face detection and recognition. For the latter task, generated facial images are compared against their original counter-parts. The HDA-CD database is made publicly available at: https://dasec.h-da.de/hda-cdfdb/

ALJAŽ ĐUKIĆ - MULTITASK IRIS SEGMENTATION

Full Title: Development of a multitask model for iris segmentation
Institution: University of Ljubljana, Slovenia
Supervisor: Vitomir Štruc
URL: https://repozitorij.uni-lj.si/Dokument.php?id=155727&lang=slv
Link description: The PDF is available for the thesis - in Slovene
Contact email: vitomir.struc@fe.uni-lj.si

Abstract:

The human iris is considered an extremely safe and reliable physiological modality and is thus often used in biometric recognition systems. A crucial pre-processing step for reliable and accurate iris recognition lies in iris segmentation, a process that determines which part of the captured image belongs to the iris. Iris segmentation has in recent years shifted from traditional algorithms to deep learning approaches, which have many advantages. In our work, we follow the trend of using deep learning for solving the task of iris segmentation as we try to further improve the achieved accuracy of iris segmentation using multi-task learning. For this purpose, we develop and evaluate different single-task and multi-task learning models, whose architecture is based on the classic U-Net network, which we additionally modify. We also assess the effect of using different auxiliary tasks and loss weights on the iris segmentation accuracy. Besides the auxiliary task of image inpainting, we also evaluate the performance of models built using the auxiliary tasks of image denoising and colourization of grey images. The chosen models are trained and evaluated on the MOBIUS and SBVPI datasets, where the auxiliary task of image inpainting achieves the best performance among the tested multi-task learning auxiliary tasks on both datasets. The iris segmentation performance of the single-task learning model is improved by using the multi-task learning model with image inpainting chosen as the auxiliary task only when evaluated on the SBVPI dataset, which we contribute to the differences between the datasets. We also demonstrate that choosing bigger auxiliary tasks' loss weights adversely impacts the performance of iris segmentation because of their increased influence on the training of the models.

MD RAAHIM AL AMIN - CONDITIONAL DATA-DRIVEN GENERATION OF FINGERPRINT IMAGES TAKING ACCOUNT OF PRIVACY-RELATED ATTRIBUTES

Full Title: Conditional data-driven generation of fingerprint images taking account of privacy-related attributes
Institution: Otto-von-Guericke University Magdeburg
Supervisor: Prof. Dr. Jana Dittmann, Dr. Andrey Makrushin
Contact email: andrey.makrushin@ovgu.de

Abstract:
Considering the key traits of biometric data such as uniqueness and permanence, human fingerprints have been widely used for identification and authorization processes. Fingerprints are subjected to the explicit consent of the bearer, therefore databases of real fingerprints are barely publicly available. Moreover, security of the biometric data has recently been a great concern as these data are constantly being used by its bearer for authorization and authentication. In addition to that, restrictions are posed by the cross-border regulations regarding the handling of biometric data. For these reasons, any research that needs fingerprint data might be legally obstructed. Artificial fingerprint generation could be a feasible solution to this particular problem as the synthesized fingerprints are not linked to any individual. Previously, researchers have generated synthetic fingerprints with statistical model-based approaches. At present, much research has been done on data-driven synthetic fingerprint generation to create realistic fingerprints. Focusing on dermatoglyphics which is referred to as the study of fingerprint patterns that reflects several privacy-related attributes of human e.g., gender, age, etc., the general goal of this work is to be able to generate fingerprint images with the aforementioned attributes with a data-driven approach. A Generative Adversarial Network(GAN) is used to generate fingerprints containing these attributes and validated by NFIQ2, NEUROtechnology VeriFinger, and Fréchet Inception Distance in terms of quality and realistic appearance, identity preservation, and presence of attributes respectively. Among several attributes, this work focuses on three of them. They are gender, age, and skin diseases. Gender is divided into two categories: male and female. Age is divided into two groups: young, and old without specifying a certain age range. And lastly, two categories of skin diseases: Fingertip Eczema and Warts. The GAN architecture is trained with completely synthetic fingerprint images created from a synthetic fingerprint generator.

WOUTER COUWENBERGH - DETECTION OF AREAS OF INTEREST IN BENDING STRAIN DATA FOR PIPELINES THROUGH 1D OBJECT DETECTION

Full Title: Detection of areas of interest in bending strain data for pipelines through 1D object detection
Institution: Universiteit Twente - Zilverling Service Desk
Supervisor: Luuk Spreeuwers
URL: https://essay.utwente.nl/93957/
Link description: Detection of areas of interest in bending strain data for pipelines through 1D object detection
Contact email: l.j.spreeuwers@utwente.nl

Abstract:

Bending strain is a metric used to evaluate the build-up of stress within pipelines and subsequently their risk of breaking. Its analysis is a time-consuming and tedious process mostly involving manual evaluation which this paper aims to address. Bending strain data is 1D, similar to time series. However, in this domain detection and localization of areas of interest is still a relatively new field. U-Net has shown to work well on similar 1D data such as EEG and ECG data, but object detection algorithms have yet to be used. This paper will therefore explore the feasibility of applying a YOLO v4 based model to detect areas of interest within 1D bending strain data. The model's performance will be evaluated using a 1D U-Net as a baseline, as it has already been established to work well on 1D data. Though, to allow for direct comparison with YOLO, the segmentation map of U-Net will be converted into a set of bounding boxes. The results show that U-Net outperforms YOLO in terms of detecting bends (with an average precision of 0.71 and 0.51 respectively), but that YOLO outperforms U-Net in detecting strain areas (with an average precision of 0.084 and 0.065 respectively). Moreover, while both models achieve comparable results (suggesting that YOLO performs on par with U-Net on 1D data), they are still found lacking in performance especially when detecting strain areas. In the best-case scenario, using an IoU threshold of 0.5, both models were able to attain an average precision of about 0.15, which is not sufficient to be used. Using the same threshold, however, both models were able to achieve an average precision of about 0.8 for bends which is a lot more promising. It was later found that some possible inconsistencies within the data and the labelling of said data might be the cause for this performance disparity. Future work using this data should therefore first aim to standardize the data and remove any inconsistencies. Thereafter, the focus of any future work should be on improving the detection performance of strain areas within bending strain data.

ANNA FRIDTUN AAREKOL - GRAPH THEORETICAL APPROACH TO ONLINE PREDATOR DETECTION

Full Title: Graph Theoretical Approach to Online Predator Detection
Institution: NTNU
Supervisor: Patrick Bours
URL: https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/3014533
Contact email: patrick.bours@ntnu.no

Abstract:

Many children spend much time online, watching videos, playing games, or talking with friends or strangers on social media. Many different online platforms are created targeted at children. The Internet has enabled kids to meet new friends and stay in touch with each other without physically meeting. Although these platforms may contribute significantly to children's social life, they may also pose threats to the children. The online platforms give easy access to conversations with children, even for people with bad intentions. On these platforms, predators can come in contact with children with a low risk of getting disclosed. This master thesis aims to find a method for recognizing predators online using a graph-theoretical approach. There are research projects that have already studied online predator detection. Most of the research in this area uses textual analysis for the task, many with promising results. The methods involve recognizing specific words or phrases that a predator would use that are unusual for children. There are multiple challenges with this approach. First, when making a predator detection system that analyses text, it can only function if used in the language it was developed for. It will be impossible to create such a system independent of the language. Secondly, the text messages on a chat platform are often informal and contain many slang words. This makes it challenging for machines to interpret what the messages mean. To avoid the challenges posed by the textual analysis, we use a graph-theoretical approach to detect predators online. Using a real-world data set collected from a social network for children, graph representations of the network will be used to detect predators. The users will be represented as nodes, and the messages between the users as edges. The main goal of the thesis is to study if it is possible to recognize a predator by studying the properties of the nodes in the graph. We have, throughout the study, designed and implemented a set of features that has been used in various clustering algorithms. From the results of the clustering algorithms, we have discovered multiple users that we considered likely to be predators. To assess some specific users in more detail, we studied anonymized text messages from relevant users and concluded whether the users were predators or not. We concluded that a graph theoretical approach can be used for online predator detection. However, in the future, both unsupervised and supervised learning in static and dynamic graphs should be studied further for predator detection to find more precise methods to find users with abnormal behavior.

MARKO GROFFEN - EXPLORING IDENTITY MATCHING FOR LOW QUALITY IMAGES WITH THE HELP OF A PIPELINE FOR SYNTHETIC FACE GENERATION

Full Title: Exploring identity matching for low quality images with the help of a pipeline for synthetic face generation
Institution: Universiteit Twente - Zilverling Service Desk
Supervisor: Luuk Spreeuwers
URL: https://essay.utwente.nl/94003/
Link description: Exploring identity matching for low quality images with the help of a pipeline for synthetic face generation
Contact email: l.j.spreeuwers@utwente.nl

Abstract:

In this paper we propose a pipeline for synthetic low resolution face generation as an alternative to image downsampling and real-world low resolution datasets for use in forensic cases. The goal of the pipeline is recreating physically accurate low resolution face images in a 3D space. We were able to incorporate a state-of-the-art conditioned machine learning algorithm to generate a realistic synthetic high resolution gallery dataset, by combining StyleGAN2 and attribute based latent space exploration. Using single image 3D reconstruction and a physically based renderer, an identity preserving pipeline was introduced that allows for one-to-one gallery to low resolution probe dataset generation, while enabling flexible pose, lighting, resolution and compression adjustment. Using volumetric path tracing, subsurface light scattering within human skin was emulated. To get further insight into our pipeline output, facial recognition experiments were conducted using state-of-the-art commercial and open-source facial recognition software and super-resolution upscaling using a convolutional neural network. Comparison to the real-world face image dataset SCFace was also conducted to test for potential applicability of our pipeline in forensic cases. Lack of accurate optical aberration and sensor characteristics resulted in a significantly different facial recognition performance on our synthetic dataset, making current application of our pipeline in forensic scenarios unfit. Thorough description of design choices and background make our research however an interesting stepping-stone for future research.

AMINA MOKHTAR - OBJECT DETECTION FOR TOP VIEW IMAGES IN DAIRY FARMING

Full Title: Object Detection For Top View Images In Dairy Farming
Institution: Universiteit Twente
Supervisor: Luuk Spreeuwers
URL: https://essay.utwente.nl/92045/
Link description: Object Detection For Top View Images In Dairy Farming
Contact email: l.j.spreeuwers@utwente.nl

Abstract:

Detecting cows, people, and robots on farms helps assist farmers in their daily tasks. This work explores different methods for object detection in top view images, explicitly investigating the Transformer-based networks. Besides that, a semi-supervised method used for data collection will be presented. This technique helps overcome the dataset's easiness and overpower any data imbalance problems. The results show that transformer-based models perform better than non-transformer-based models even though they predict horizontal bounding boxes. The transformer-based models achieved an average precision of 0.985 and 28 FPS, while the non-transformer-based model achieved an average precision of 0.956 with 15.4 FPS. On top of that, the developed data collection method improved the detection accuracy by 5.9%. Furthermore, we trained Deformable DETR to estimate the rotated bounding boxes to solve the challenge of oriented object detection. As a result, the Rotated Deformable DETR model achieved an average precision score of 0.984.

MONITOR
BACHELOR-THESES

MARTA ROBLEDO - EXPLORING KEYSTROKE DYNAMICS SYSTEMS FOR MOBILE DEVICES

Full Title: Exploring keystroke dynamics systems for mobile devices
Institution: Universidad Autonoma de Madrid
Supervisor: Ruben Vera-Rodriguez
Contact email: ruben.vera@uam.es

Abstract:

Due to the expansion of technologies used to perform increasingly sensitive tasks (such as any action that requires the user's identity to be always verified), it is important to pay attention to research on improving the security of mobile devices by designing continuous subject authentication systems. This is a particularly complex problem because it must be addressed in such a way that it interferes as little as possible with the user's experience. Biometric recognition mechanisms are an emerging topic, based on the fact that each person shows physically different features (e.g. fingerprint or iris) and, in addition, manifest behaviors that can distinguish them from others (e.g. the way they walk or they keystroke dynamics). The most recent lines of research point to the use of generic yet flexible and optimizable algorithms, such as methods based on Deep Learning. Mobile devices provide a large amount of data thanks to the information they can collect from the environment through sensors. This information enables the analysis of device usage dynamics. This data is of vital importance for Deep Learning based systems, since it represents the set of information on which the algorithm should rely on in order to learn. In this Bachelor Thesis, the previous research on user verification systems based on keystroke dynamics are first reviewed and then new options with better performance are explored, whose cost is not high and that work in a transparent way to the user. This work starts from a model developed by the BiDA-Lab research group (called TypeNet), which is based on Long Short-Term Memory neural networks, and then modifications are made in order to study the influence of certain factors, such as the type of keyboard used, the duration of the typing sequences and the network input features. The extracted results are quantified in terms of network performance and analyzed considering those factors that refer to the quality of the user experience. It is concluded that there is a need for keyboards capable of extracting key press and key release data accurately. It is also important to have sufficiently long typing sequences in order to carry out a thorough analysis. In addition, the importance of making an appropriate choice of network input features is noted, since they largely determine the performance of the network.

PABLO CANCELA - EXPLORING ON-LINE SIGNATURE VERIFICATION SYSTEMS FOR MOBILE DEVICES

Full Title: Exploring on-line signature verification systems for mobile devices
Institution: Universidad Autonoma de Madrid
Supervisor: Ruben Vera-Rodriguez
Contact email: ruben.vera@uam.es

Abstract:

In this Thesis, it is presented for the first time ever a wide finger-based online signature database captured with general purpose mobile devices: BehavePassDB. Studies involving well known verification systems have been carried out not only to test the acquired data quality, but also intending to adapt these systems to our acquisition scenario: day-to-day signatures. First of all, all BehavePassDB data was meticulously revised. Pre-processing, adaptation, correction and segmentation of the signatures was completed before any experimental use of the database. Immediately after completing this work, the student and some other members of BiDA Laboratory actively participated in making various forgeries of each genuine signature, being an essential part in every database created with verification systems development purposes. Afterwards, the forgeries went through a similar process to the genuine ones (revision and preprocessing), and they were finally added to the database, leading to the beginning of the main part of this thesis: the experiments. At first, experiments were focused on exploring online signature verification systems which have demonstrated to obtain really satisfactory results in past studies. Not only in order to anticipate its performance with the new data, but also to establish a reference for other systems to outperform. Once this first approach was carried out and the student was able to get deep in the understanding of how these systems work, a new goal was set: the adaptation of the system to our database. The main system used in this thesis is deep learning based. Therefore, the adaptation consisted in exploring the parameters and re-training the network with our new data. The purpose is likewise manage to get a robust spoofing system able to unequivocally identify every user, but this time adapted to daily scenarios with general purpose mobile devices. This is a challenging task due to the lower quality of this data compared to other databases captured with specialized acquisition devices. Finally, once the system was evaluated and all the results were obtained, the student drew some conclusions and gave a list of interesting ideas for future researching regarding this thesis.

MARINA IULIANA AUR - WEARABLE GAIT RECOGNITION

Full Title: Gait recognition via wearable-captured signals
Institution: Sapienza University of Rome
Supervisor: Maria De marsico
Contact email: demarsico@di.uniroma1.it

Abstract:

The work is a follow-up of previous projects to improve the performance of gait recognition using features extracted from wearable accelerometer signals.

ISIDOR CLAUDIU DAMIAN - SCLERA RECOGNITION

Full Title: Biometrie Oculari: Segmentazione di Sclera e suoi Vasi Sanguigni (Ocular biometrics: segmentation of the sclera and its blood vessels)
Institution: Sapienza University of Rome
Supervisor: Maria De Marsico
Contact email: demarsico@di.uniroma1.it

Abstract:

The work aims at processing eye images to extract the sclera region and then further segment it to extract the pattern of blood vessels. The main steps of the devised method can be grouped in sclera-related and vessel-related. The sclera-related steps entail a preprocessing phase including the color space conversion from RGB to HSV and Gaussian blur, then a k-mean clustering to extract the different regions and a post-processing via morphological operators to refine the result. The vessel-related steps include CLAHE preprocessing, adaptive thresholding, and a final postprocessing to delete artifacts.

LUKA MARKIĆEVIĆ - EAR SUPERRESOLUTION

Full Title: Super-resolution with the application on ear images
Institution: University of Ljubljana, Slovenia
Supervisor: Peter Peer and Vitomir Štruc and Žiga Emeršič
URL: https://repozitorij.uni-lj.si/Dokument.php?id=161145&lang=slv
Link description: The PDF is available for the thesis
Contact email: vitomir.struc@fe.uni-lj.si

Abstract:

Super-resolution (SR) represent a class of image enhancing methods that boosts image resolution. This is useful in various areas, such as visually enhancing photographs or improving person recognition performance. This undergraduate thesis focuses on Single Image Super Resolution of ears. One of the earliest ways to address the issue of super-resolution was interpolation, but achieved limited success. The latest improvements in SR have been made feasible by deep neural networks, which significantly improved performance. We evaluated the performance of the Enhanced Deep Residual Network (EDSR) and Shifted Windows Transformer Network (SwinIR) for image super-resolution of ears. Using the AWE dataset, which consists of 16,665 images of ears of various sizes, shapes, and orientations, we trained four models: two on EDSR and two on SwinIR networks, each with scaling factor of two and four. Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM) performance measures were used to evaluate the two different model designs. SwinIR achieves a superior PSNR and SSIM, however, the visual results seem to be highly similar.

MATTHIAS MYLAEUS - LOW-RESOLUTION FACE RECOGNITION USING RANK LISTS

Full Title: Low-Resolution Face Recognition Using Rank Lists
Institution: University of Zurich
Supervisor: Manuel Günther
URL: https://www.merlin.uzh.ch/publication/show/22681
Contact email: siebenkopf@googlemail.com

Abstract:

The field of automatic face recognition has experienced a significant boost in recent years since the use of artificial neural networks was introduced. Face recognition today poses a critical element for everyday life, supporting various tasks from security, surveillance, and access control all the way to unlocking smartphones. In many different situations, such as changing illumination and faces covered with scarfs or glasses, recognition networks have come to shine and achieve the most accurate results. However, when it comes to recognizing faces from a far distance, they start struggling and leave room for improvement. This thesis discusses the usage of a reference database. Images are compared to it, resulting in a signature. Then, rather than comparing a probe image directly to the gallery, their signatures are compared. The idea of rank lists is set side-by- side with the standardization of those signatures to evaluate whether more accurate results can be achieved. However, for all experiments, results show that the usage of a reference database does not outperform the direct comparison. Further research using more extensive databases and various network models is needed to ensure which approach is more accurate.

ŽIGA ROT - SCLERA VASCULATURE SEGMENTATION

Full Title: Sclera vasculature segmentation from visual data using deep learning
Institution: University of Ljubljana, Slovenia
Supervisor: Vitomir Štruc
URL: https://repozitorij.uni-lj.si/Dokument.php?id=160146&lang=slv
Link description: The PDF is available for the thesis - in Slovene
Contact email: vitomir.struc@fe.uni-lj.si

Abstract:

This thesis presents the implementation of a model that can extract scleral vasculature from image data – a biometric feature which can be used in iris-based biometric recognition systems to enhance robustness and accuracy. The model consists of two stages. The first stage is used for extraction of the region of interest – the sclera, from which then the next stage segments the vascular structure. We justify our design with two experiments. In the first one, we show the impact of prior extraction of the region of interest on the final output. In the second one, we present the difference in segmentation quality between binary and multi-class versions of sclera segmentation.

SEBASTIAN SCHACHNER - COMPRESSED FACE IMAGE QUALITY ASSESSMENT

Full Title: Auswirkung von Kompression auf die Bildqualität für Gesichtserkennung
Institution: Hochschule Darmstadt
Supervisor: Torsten Schlett
URL: https://dasec.h-da.de/sebastian-schachner-successfully-defended-his-bachelor-thesis
Contact email: torsten.schlett@h-da.de

Abstract:

This work compares different compression schemes with respect to face recognition. Face recognition is an essential part of today’s society, as digital identification of people is becoming more common, for example with digital ID cards. For this purpose, a reference image is stored on the device, which is compressed to be stored in a space-saving manner. For this different types of compression are available. In order to achieve the smallest possible storage sizes, lossy compression types are used. For the compression the schemes Joint Photographic Experts Group (JPEG), JPEG 2000 and JPEG XL are chosen, as well as the downscaling of images in the Portable Network Graphics (PNG) scheme. In addition, an autoencoder approach is used that uses an 128 dimensional embedding and compressed residual images. The residual images are created from the difference of a blurred image, generated by decoding the embedding, and the original image. Two experiments were conducted for the analysis. The first experiment uses images cropped to relevant face informations in regards to face recognition. The second experiment uses portrait images. Both experiments are analyzed for image quality and comparability using different target sizes. For the Face Image Quality Assessment (FIQA), five models based on Deep Convolutional Neural Network (DCNN)s and two models estimating the sharpness of the images are used. For image comparisons the ArcFace model is used. Mated image pairs, non mated image pairs and original images mated with compressed images are compared. In addition, we analyse the deviations of the actual sizes of the compressed images to the target sizes and to what extent the quality assessment match the mated comparison values. In general it could be shown that JPEG XL could achieve the best quality preservation and the best comparison values. Furthermore, it could be shown that JPEG XL achieved the target values best. However, it was shown that quality assessment models do not always correlate with the comparison values. Therefore, it could be concluded that the influence of compression for face recognition depends on the target size and the compression type. Thus, for strong compression, the smallest influence could be observed for JPEG XL and for scaled PNG images the largest influence on face recognition was measured. However, at low compression the approaches were largely similar. This work can be used as a basis for further research. On the one hand, this work can be used to further improve the estimation of image quality models, and on the other hand, it can be used as a basis to develop better compression methods, especially for compression regarding face recognition.

WUFEI YANG - FAIRNESS IN PAD

Full Title: Bias Exploration and Mitigation in Face Presentation Attack Detection systems
Institution: Fraunhofer Institute for Computer Graphics Research (FraunhoferIGD)
Supervisor: Meiling Fang
Contact email: ahfieywf@gmail.com

Abstract:
With the widespread application of face recognition systems, face presentation attack detection (PAD) plays a critical role to protect the security and credibility of the system. A growing number of researchers have investigated the biases of face recognition systems, and their results demonstrated the existence of demographic and attribute biases in recognition systems. However, the fairness of face PAD system has not attracted much attention. A key problem is the insufficient demographic and attribute annotations of face PAD data. Hence, this thesis first combines six face PAD databases: CASIA-FASD, REPLAY- ATTACK, MSU-MFSD, HKBU-MARs V1+, OULU-NPU, WFFD , which consisting of print, replay, 3D mask, wax face attacks. Furthermore, identities from this combined database are manually annotated with one demographic label and six facial attribute labels. Second, this thesis explores and analyzes demographic bias and additionally facial attribute bias in face PAD methods by using this combined database. To enable the bias study, one hand-crafted feature based model LBP-MLP, and three deep learning based models: ResNet50, DeepPixBis, and LMFD-PAD, are adopted. In addition to report PAD performance, a modified fairness discrepancy rate (FDR) is introduced to further determine the system fairness. The experimental results point out that deep learning based PAD models trained only on female or male group are unfairer than models trained on the fused data (including female and male). In addition, models trained on fused data and only on occlusion group indicate higher fairness than models training only on non-occlusion data. To further mitigate system bias, a modified version of PatchSwap, named cross-identity PatchSwap in this thesis, is introduced to enable patch substitution between identities with different gender and facial attributes. Despite the significantly improved PAD performance achieved by cross-identity PatchSwap, the FDR results also suggest that this approach is able to improve the system fairness for different gender groups when models are trained on fused data and on male data.