FRONTEX
EUROPEAN BORDER AND COAST GUARD AGENCY

INTERNATIONAL CONFERENCE ON BIOMETRICS FOR BORDERS

International Conference on Biometrics for Borders

Morphing and Morphing Attack Detection Methods

CONFERENCE PROCEEDINGS

9-10 OCTOBER 2019
WARSAW

FRONTEX
EUROPEAN BORDER AND COAST GUARD AGENCY

INTERNATIONAL CONFERENCE ON BIOMETRICS FOR BORDERS

International Conference on Biometrics for Borders

Morphing and Morphing Attack Detection Methods

CONFERENCE PROCEEDINGS

Legal notice
The content of this publication does not necessarily reflect the official opinion of Frontex — the European Border and Coast Guard Agency — or any institution or body of the European Union. Frontex does not guarantee the accuracy of the data included in this report. Neither Frontex nor any person or company acting on behalf of Frontex is responsible for the use that may be made of the information contained in this report.

Frontex — European Border and Coast Guard Agency
Plac Europejski 6
00-844 Warsaw, Poland
T +48 22 205 95 00
F +48 22 205 95 01
frontex@frontex.europa.eu
www.frontex.europa.eu

Warsaw, February 2020
Research and Innovation Unit
Capacity Building Division

Luxembourg: Publications Office of the European Union, 2020

Graphic design by Softwin

Print version:
TT-04-20-167-EN-C

PDF:
TT-04-20-167-EN-N

FPI20.0019

Table of contents

List of acronyms

ABC	Automated Border Control
AI	Artificial Intelligence
APCER	Attack Presentation Classification Error Rate
API	Advance Passenger Information
BPCER	Bonafide Presentation Classification Error Rate
BSI	Germany Federal Office for Information Security
CGI	Computer-generated imagery
CNN	Convolutional neural network
EAB	European Association for Biometrics
EES	Entry Exit System
EU	European Union
eu-LISA	The European Union Agency for the Operational Management of Large-Scale IT Systems in the Area of Freedom, Security and Justice is an agency of the European Union
FRTV MORPH	Face Recognition Vendor Test (face morphing)
GAN	Generative adversarial network
GDPR	EU General Data Protection Regulation
IATA	International Air Transport Association
ICAO	International Civil Aviation Organization
ISO	International Organization for Standardization
MAD	Morphing Attack Detection
NGO	Non-Government Organisation
NIST	National Institute of Standards and Technology
PNR	Passenger Name Record
REST	REpresentational State Transfer
SIS	Schengen Information System
SML	Standard Meta Language
SVM	Support Vector Machine
UN	United Nations
VIS	Visa Information System

Foreword

Dear readers,

The first **International Conference on Biometrics for Borders** was held on 9-10 October 2019 in Warsaw, Poland. This first instalment of the ICBB proved to be a success with more than 200 distinguished international guests, representing authorities from EU and non-EU countries, numerous EU institutions, international organisations, research institutions and academia, the biometric community and industry, and affirming the Agency's commitment to organising a series of regular conferences dedicated to the topic of biometrics, its application in border control, and the opportunities and challenges it may pose to border management and border security.

Frontex plays a crucial role in the European Union's effort to safeguard the area of freedom, security and justice, and has become a cornerstone in guaranteeing an area of free movement without internal border checks. With the progressive expansion of the Agency's mandate, Frontex has more and more increased its activities to combat cross-border crime and help prevent terrorist attacks. Intelligence and information are crucial for formulating an appropriate response to real and potential threats at the EU's external borders. Frontex is continuously piloting new and innovative technologies to modernise the EU's border management, with the objective of striking an appropriate balance between increasing border checks and security screening, while facilitating smooth and fast border crossings of travellers visiting the European Union for business, tourism or study.

Biometric technologies are at the forefront of this effort, with the application of biometric technologies providing new opportunities to border management to facilitate legitimate travel while at the same time making borders more secure. We recognise that the large-scale integration of such technologies into the border control infrastructure does introduce new challenges to border security: Biometric systems can be attacked and subverted for the purpose of passing through border control undetected. This means that with the introduction of novel technologies supporting biometric recognition, countermeasures that can prevent, detect or defeat such attacks are needed. To sum up, for Frontex and the European border management community, the exploration and development of advanced biometric technologies and related capacities is crucial to border security.

Though the thematic spotlight of the 2019 conference was on morphing and its possible implications for border management, the overarching focus on biometrics remained strong. This is important, because the Agency

has the ambition to be a driving force in providing support and expertise to Member States and the European Institutions on the topic biometrics, and to lead on the range of possible applications and implications for borders. Current and future conference will continue to underline this ambition, aiming to establish a strong tradition of bringing national authorities, researchers, academics, industry and professionals from all over the world, to support and exchange on activities undertaken in this area, at EU level and beyond.

The success of the conference was ultimately due to the large number of international delegates who actively and enthusiastically participated in various panel discussions and thematic sessions. I would therefore like to express my gratitude to all the participants. We have learned from you, and hope that in return we have succeeded in offering you a valuable insight into the views and experiences of the border management community.

I wish you all pleasant reading, reflecting on some of the key messages we can take away from the conference. And of course, I hope we may welcome you in the future at our International Conference on Biometrics for Borders!

Fabrice Leggeri
Executive Director

Executive summary

The International Conference on Biometrics for Borders 2019 brought together experts and stakeholders from a variety of institutions from different disciplines, covering legislators, border management authorities, academia, NGOs and Industry Associations, and from different geographies including the EU (both Commission and Member States[1]), the United States and Australia. Despite their diversity they all shared an interest in the application of biometric technologies and solutions at borders.

The main purpose of the conference was to facilitate knowledge transfer and dialogue among the various stakeholders, as well as to summarise the key opportunities and challenges that biometric solutions at the border face both now and in the near future. As an example of a challenge the conference also looked in more detail into the threat to security potentially posed by morphing and discussed the latest updates on the status of potential morphing attack detection solutions.

A number of key themes came out of the conference:

Firstly, there was a general consensus that biometric solutions will continue to develop and will be an essential part of the border control process now and in the future, despite the variety of operational and implementation challenges they face, as they are essential to deal with the ever-increasing volumes of traveller facilitation whilst at the same time ensuring enhanced security.

Secondly, there is a continued need for improved coordination between the worlds of legislation, technology and the actual operational implementation of this technology at the border. It was emphasised that there needs to be an improvement in the testing of potential technological solutions, either in the field or in simulated field scenarios using specially prepared testing sites, which means that there has to be increased cooperation at an earlier stage between industry representatives and academic communities as well as with the end user community of border management and law enforcement bodies. In the EU context the European Border and Coast Guard Agency was perceived by many to be a natural potential leader of such projects.

Thirdly, there was an emphasis on the importance of training the end user of technology so that the solutions truly optimise the entire border crossing process and provide proper support to the human border guard. Training on biometrics should also be targeted at the decision maker level, as they also need to better understand the potential of biometric solutions, despite the inevitable evolving obstacles and vulnerabilities that come to the surface following the implementation of any new technology.

Fourthly, in terms of the threat of morphing attacks it was noted that currently none of the available Morphing Attack Detection (MAD) algorithms came close to offering an acceptable operational solution to this problem, and that there appeared to be a need for greater engagement from various stakeholders, including the commercial sector, to generate a volume of better algorithms that could provide potential operational solutions at the borders. Also noted was the need to develop and test other process- or human-based solutions to this problem, for example rolling out live enrolment of an applicant's image at passport document issuance or via the deployment of

1 In the context of these conference proceedings, Member States will refer to EU Member States and Schengen Associated Countries.

better human detection of morphed images at the borders. More work also needs to be done in Europe to standardise the measurement of quality criteria and the testing methodologies in the area of morph detection.

Fifthly, it was also noted how important it is to follow the GDPR[2] and national data protection regulations in the EU when testing new solutions and that access to properly managed data sets is of critical importance. It was stated that the likely solution will lie with the idea of bringing the algorithm to be tested to the data set rather than the other way around.

Sixthly, there was a feeling in particular from industry that there are big opportunities to be found in looking at the total travel process, not just the time spent by a traveller at a border crossing point. Potentially, easy efficiencies could be generated by the sharing of biometric data between key stakeholders and through the encouragement of processes where a previously enrolled traveller only has to have their data captured or checked once, not potentially multiple times. This in turn may help with the general societal acceptance of biometric solutions as it will be clearer what the convenience benefit is for the bona fide traveller of the continued roll-out of such solutions.

In conclusion it was stated that this kind of conference brings evident positive benefits for all the stakeholders in biometrics and that Frontex will look into organising more such events.

2 GDPR refers to the General Data Protection Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance), OJ L 119, 4.5.2016, p. 1–88.

DAY 1

Welcome address

Javier Quesada · Head of Unit, Research and Innovation, Frontex

On behalf of Frontex, Mr Javier Quesada opened the International Conference on Biometrics for Borders 2019. He welcomed all the delegates and reminded them of the programme and also that the conference topics would start by looking at the issues at a strategic level, in particular the role of biometrics at borders and the general role of Frontex, and would then proceed to dig deeper into the details concerning morphing and morphing attack detection methods, before returning again near the end of Day 2 to a more general level with discussions then focusing on what the future will bring for biometrics and border control. He stated that the format of the conference, with it being international in scope as well as multi-stakeholder, bringing together as it does in one place government, administration, technology vendors and academia, will he hope lead to a good exchange of ideas, contacts and networking opportunities.

He then proceeded to introduce the first of the two keynote speakers, Mr Fabrice Leggeri the Executive Director of Frontex, the European Border and Coast Guard Agency.

Keynote speech

Fabrice Leggeri · Executive Director, Frontex

Mr Leggeri started his speech by warmly welcoming all the delegates from different countries and also from various disciplines, including border administration, academia and the research community, as well as from technology companies. He went on to thank in particular the key stakeholders who have helped to organise the conference, in particular the European Commission and Finland as the Member State holding the current Presidency of the European Council as well as the key departments and people in Frontex who have been responsible for the conference organisation.

Moving to the main part of his speech he went on to state that in the modern world the application of technology to facilitate fast border crossings and protect citizens is essential, as the numbers of people involved and the nature of the evolving security threats simply makes the use of technological solutions inevitable.

He pointed out that this year alone it is expected that there will be more than 600 million legal crossings of the external border of the EU and it is estimated that this will rise to over 900 million by 2025, with one in three being made by citizens of third countries (countries from outside the EU or Schengen area). Given these numbers the key question

is therefore how, given these massive flows, to ensure a smooth and quick entry and exit at border crossings for the bona fide traveller while at the same time improving the detection of potential security threats.

In this world one of the key tools for the border guards will be technology – in particular biometric solutions. Mr Leggeri stated that these solutions if applied properly have the power to both speed up the facilitation of border crossing for the bona fide traveller while at the same time improving security detection and the efficiency of operations both at a national and at an EU level.

Biometrics is key for Frontex – and its mandate. Many activities are implemented but they all have one common goal: to protect and safeguard the area of free movement within the European Area of Freedom, Security and Justice without internal border checks. Mr Leggeri went on, however, to acknowledge that the large-scale application of biometric technologies at the borders also brings with it new challenges. These challenges are not only of a technological nature. They also relate to the impact that new technologies have on the procedures and processes followed at border crossing points. As such they present challenges for the border guards themselves, who are required to adapt and change their internal professional culture. It is also important to note that these technologies also impact on the issue of privacy and ethics in the context of the border guard's work.

For Frontex, the European Border and Coast Guard Agency, the technology must support the overall goal of border management, and it is because of this that conferences such as this one are so essential.

Frontex is a leading player in the expected new mandate in the field of integrated EU border control and will be playing a key role in the setting of standards to ensure that new policies can be translated into something which is of operational value.

As an example of how Frontex can play a positive role in facilitating the development, testing and piloting of new technological solutions, Mr Leggeri spoke about a pilot project which is being carried out at Lisbon Airport in cooperation with the Portuguese Immigration and Border Service to help test biometric technologies that capture biometrics on the move in a real operational environment. In this test travellers who are leaving the EU can be checked seamlessly using face recognition and touchless fingerprint scanning technology. This example demonstrates that Frontex facilitates the testing of new technology solutions, with targeted trainings in vulnerability assessments and the development of new best practice guidelines, so as to work towards creating a consistent and harmonised approach at all EU border crossing points as well as a similar passenger experience.

This, he stated, is just an example of how Frontex is committed to advanced technology solutions for borders and shows how this Agency offers expert support to both the European Commission as well as the Member States in this area. Frontex is and wants to be more and more a key player in this field for the foreseeable future. This conference is also a demonstration of this commitment, as it brings together different perspectives and areas of expertise.

To close his speech Mr Leggeri invited the conference participants to formulate a positive vision of the role of biometrics in leading to a future of faster, smoother facilitation and stronger security in a seamless travel paradigm. A future where the border guard feels supported and is mindful of his or her role in providing protection and service to citizens at their borders. Mr Leggeri called for the

conference to help secure the pathway to the proper usage of technology in order to deal with the vastly increased numbers as well as the new and evolving security threats, while always keeping in mind the mandate of Frontex to contribute to the securing of the EU area of security, freedom and justice with strong and properly functioning external borders. Mr Leggeri stated that he is greatly looking forward to hearing the perspectives from the conference delegates about these issues as well as about the role of Frontex in this area.

Finally, he invited the conference to listen to the next keynote speaker from the European Commission.

Keynote speech

Olivier Onidi · Deputy Director-General, Directorate-General for Migration and Home Affairs, European Commission

Mr Onidi started by saying how pleased he was to be invited by Frontex to attend this conference and give a keynote speech. He emphasised how strong a pillar Frontex is in the overall European security union and said that if there is no security at the external borders then there will be no security within the European Union – so these are matters of critical importance.

The Agency has also played a key role in enhancing the information flow between the Member States and the law enforcement community, ensuring the high standard of this information and of course working on solutions to the challenges which lie ahead. Mr Onidi then mentioned the work that has been done in the area of interoperability: the push to share more information but as well the desire to develop new ways of collecting new information. In this context he mentioned the work done to develop the new EES as well as on the upgrade of existing systems such as SIS and VIS. For these systems it is not only the collection of information, which is important, but also its quality, so the access to biometric data which enables a match to be made between a document and an individual person at a border crossing point. On this point in 2019 there has been a significant increase in the detection of false identities (+25%), which he sees as being both a result of improvements in the quality of checks as well as the fruit of the continuing development of new methods to falsify documents including the growth of morphing.

He went on to remind the audience that the role of biometrics was in a way unique as it requires a balance between security and fundamental rights to privacy and data protection. It is necessary to make better use of the information that is already collected and in this area he mentioned improvements to the Eurodac System, which now holds facial records in addition to fingerprints, as well as the planned EES which will require 2 biometric identifiers for the crossing of the external border, and improvements in the VIS system which will require much higher standards for the application photograph so as to increase the certainty later on as to the person's identity. Mr Onidi went on to state that he believed that recent enhancements in features to improve security did not mean that faster facilitation of visitors was not also possible at the same time. He believes that the pace of the discussion in the Member State governments and in the Commission needs to be picked up in order to give clear guidance to the technology providers, and in order to meet the deadline of having the EES operational within

2.5 years from now. He emphasised that the Commission would not wish to come back in 2-3 years to legislate to force harmonisation on solutions already deployed by Member States. Therefore, the focus is on having strong guidelines for harmonisation and interoperability, including at the level of the IT solutions being developed by eu-LISA, however in a framework that also ensures there is flexibility for each individual Member State. This harmonisation and interoperability are essential if there is to be a smooth transfer and analysis of data.

He went on to say that the Commission sees the role of the conference, as well as other initiatives in this area, to move in the direction of creating these guidelines, frameworks and standards which will then be a guide for the Member States.

He said he is looking for guidance from the conference in three key areas:

a) How to smoothly implement what has already been agreed upon
b) Input regarding enhanced security features, e.g. to counter morphing attacks
c) Training needs

He said that conferences such as this one are very important as they help law enforcement stay up to date with what technology can offer and therefore can help anticipate some of the law enforcement challenges which will come up in the future. He underlined that the work that Frontex and others do on new technology in the field of border management is essential for the EU to keep its capacity for strategic autonomy. He said that the EU does not want to rely on solutions and technology from the outside world to ensure its security. In these crucial areas they want to be able to develop and deploy the solutions "from within" when and if necessary.

To conclude Mr Onidi thanked Frontex and Director Fabrice Leggeri again for their work on organising the conference and also stated that Frontex is now very much playing the role in these discussions that the Commission wants, in particular in helping to identify new threats, analysing the forms of vulnerability in current and planned systems, and developing the best responses to them, as well as continuing to reach out to all stakeholders including civil society and industry.

OPENING PANEL DISCUSSION

Biometrics for Border Control and the role of Frontex

Moderator
Javier Quesada · Head of Unit, Research and Innovation, Frontex
Panellists
Aija Kalnaja · Director, Capacity Building Division, Frontex
Narjess Abdennebi · Chief Facilitation Section, International Civil Aviation Organisation
Jean Salomon · CEO, European Association for Biometrics
Hans de Moel · Director, Biometrics Institute
Laurent Beslay · Scientific Project Leader Law Enforcement Technologies and Citizen, Cyber and Digital Citizens' Security, Joint Research Centre, European Commission

The Moderator, Mr Quesada, introduced the panellists and the topic of the discussion. To start the session Mr Quesada asked all the panellists in turn to make a key statement which they plan to defend during the session. First off Aija Kalnaja stated that she represents the European Border and Coast Guard Agency on this panel, which she described as being the "European layer" of border management and

control and that she wished to present this point of view during the discussion.

Following her Narjess Abdennebi introduced ICAO as the UN specialist Agency which has as its goal the maintenance and growth of safe, secure and efficient civilian air transport globally, and that one of their strategic objectives is the promotion of passenger facilitation and security, which in their view are not seen as being necessarily in conflict with each other. In this she stated that she believed that their goals were very similar to the objectives of Frontex, as laid out by Mr Leggeri earlier in the morning.

Mr Salomon then summed up what he wanted to say in 6 words: "cooperation is key under controlled trust", by which he meant that given the situation was fluid with evolving aspects of technology and moving risks there was a need to find an overarching umbrella, or one constant, and he then suggested that perhaps this role could be played by Frontex, using its trustworthy operational transversal platform to provide sustained multi-level training.

Mr de Moel then stated as his introduction that the Biometric Institute which he is representing at today's conference has as its mission the promotion of the responsible and ethical use of biometrics and given this he would like (later, in the main part of his presentation) to draw a comparison with the responsible and ethical use of alcoholic beverages.

Mr Beslay then said that he would be focusing on quality, and that this should be perceived through the dimensions of assessment, management and improvement.

Aija Kalnaja

Ms Kalnaja started by saying that the reason why the European Border and Coast Guard Agency is under such a spotlight currently was connected of course to the migration pressures on Europe of 2015 which caught everyone by surprise. She said that the key goal now was to ensure that Europe was never again caught by surprise by such an event to the same extent. She then went on to note that we should remember that borders do not only represent security, they are meant to be crossed, as crossing borders means prosperity. We can see this most clearly in the European Union Schengen area where there are no borders, which is designed to enhance trade and interaction between these Member States. However, she emphasised that of course borders must also represent security and keeping citizens' homes safe. So, it is a balancing act between facilitation and security. She also reminded the audience that with regard to refugees there is also a moral dimension where we need to help those who are fleeing war.

She noted the globally unique structure of the European Union where the external borders are controlled by the Member States and the Frontex Agency together. She gave some statistics to illustrate the scale of the task. The external border of the European Union is 45 000 km long, of which 35 000 km is made up of the sea and there are a bit less than 1900 border crossing points. Also, in 2018 well over 500 million people crossed the EU external border. She stated that she mentions the statistics to put into context the discussion and as an introduction to the key question: how do we deal with these increasing flows?

Ms Kalnaja described Frontex as being the capability planners of the European border control and coast guard, but also that they represent the end user of border management. She claimed that as a person from a law enforcement background she realised that border control is not naturally innovative but rather reactive to challenges and rather looks for solutions which are on the market already and

on the table. Given this she sees the mandate for Frontex as being visionary, as it brings together border control with the end user and the capacity and ability to innovate. Technology has changed law enforcement and border control completely, but given this new technological world, what is the dream of the border law enforcement officer about this new future? Ms Kalnaja stated that this dream is that what can be done by technology at the border is done properly and that what needs to be done by human resources is done by human resources.

In the modern world, she argued, it is no longer enough to solve problems with more human resources, as the problems are too complex for this, so it is not about more human resources but about having the right people. She believes that this quality of human resource is more and more at a premium and this is why there must be a link with technology.

Ms Kalnaja then went on to talk about border management budgets. She emphasised that the amounts that are currently being discussed put substantial amounts into border control, with EUR 26 billion on the table between Member States and Frontex Agency in the new financial framework, with Frontex expected to get just over EUR 2 billion per year, up from the current approximate number of EUR 300 million. Ms Kalnaja described this as a "big step up" in resources. She said that the pooling of these resources will mean that issues of border management will become of more interest to research and industrial partners. She went on to state that this kind of forum, where the end user meets with research, academia and industry, is the only way forward as it allows for a proper exchange of perspectives which will hopefully lead to good solutions.

Ms Kalnaja then turned back to the role of Frontex, saying that the policy of border control is set at a higher level and that the Agency's job was to help implement it. In policy there are a lot of new initiatives as outlined by Mr Onidi from the Commission, including the EES, SIS and VIS upgrades and more. Ms Kalnaja repeated that she saw the role of Frontex in the dialogue with research, industry and academia as being about bringing the perspective of the end user to the table. It was emphasised that future scenarios will need to be worked up, which is not something that law enforcement are generally very good at, as they tend to focus only on the present or up to one year ahead. However, for technology development it will be necessary to think 10 years ahead, or even 20 as this is the time span often required to go from technology idea to end user execution on a mass scale. It is in this definition of the future scenario so as to define the "task" for technology development that Frontex will be able to deliver.

Although Frontex already has a record of working on technology innovation, the vision is to bring a more coordinated message from Europe as to what the future perspectives are, with Frontex being designated as a "senior user" of border technological solutions, hopefully providing a form of a "testing laboratory" for proposed new technological solutions, which will also ensure their ethical application.

The moderator then asked this speaker whether she believed that the human border guard will at some point be completely replaced by technology. Ms Kalnaja replied that she hoped not and pointed out that even with the development of artificial intelligence solutions, the algorithms are still written by humans. At this moment she thinks a bigger risk is that the technology will not be used properly at the borders due to a lack of training. She also pointed out that technology was better at dealing with standardised problems, but that problems that do not fit a standard approach will probably always have to be dealt with by a physical border guard.

Dr Narjess Abdennebi

This speaker first of all reminded the audience that the role of ICAO is to set standards for global civilian aviation, which take the form of 19 annexes, with annex 9 dealing with facilitation. This annex, although 70 years old, is being continually updated to consider technological innovations in the area of border clearance and control. To aid states they have developed the TRIP strategy, or "Traveller identification program", which requires regional and international cooperation for it to be implemented. She stated that the regional role and expertise of Frontex here is very helpful and that it could maybe be used to spread good border management solutions on other continents where they are needed, notably in Africa. The heart of the TRIP strategy are the guidelines and specifications for the machine-readable travel document – i.e. the passport. In order to ensure interoperability on a global scale all countries should follow the specifications when issuing their passports/travel documents to their citizens. Currently ICAO has a compliance program where they help states implement these specifications. The current specifications in terms of biometrics have only one identifier, which is the photograph, but it is possible that this issue will be re-opened at the expert level in ICAO.

Dr Abdennebi then went on to discuss the issue of passenger data exchange based on UN Resolution SC 2178 from 2014. ICAO has established a standard for API implementation, which can only be done once a state has legislated to regulate the use of this data. She then mentioned PNR data, which as of today is not yet a standard, however a recommendation in this area from the ICAO taskforce will be considered by the board in the near future, with key issues likely to be data protection and privacy.

At this point the moderator asked this speaker about her thoughts on the issue of what will be the approach in a world potentially with only digital documents –without a paper version. Dr Abdennebi answered by saying that e-Passports are not currently a standard because ICAO wants to ensure that all non-machine-readable passports are out of circulation first before potentially moving in this direction. She reminded the audience that although more than 130 countries have some form of e-Passport there is still work to be done to ensure that they are all fully compliant with the specifications. She mentioned that there is an ICAO expert panel that is working on a potential "paperless" traveller id solution, but that in fact their current recommendation is that the concept will be a hybrid one which combines a physical token with a digital one.

Jean Salomon

Mr Salomon started by returning to his theme of how we ensure cooperation and stability. He said he wanted to look at this in three ways, a) scalability b) synchronisation c) ubiquity of usage in the context of biometrics.

Firstly, in the context of scalability he asked whether we were able to operate at full scale using only one biometric measurement. He

stated that in fact if we were to extend a nationwide programme to allow a unique identification by using a single biometric among a registered population such as that of India, we would end up with thousands of potential matches to manually resolve. So, he stated the importance of understanding the limits of scalability, implying the joint use of multiple biometric identifiers. In this context he said we should also remember the additional data quality challenges related to the use of biometrics in roving "field" conditions, such as at night, at sea etc. The relatively benign and controllable environment of an international airport cannot be the only place where biometrics can be scaled up.

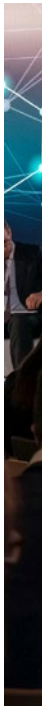
Secondly, Mr Salomon linked synchronisation to the behaviour of the various stakeholders in terms of operational legislation, looking for harmonisation between different legislative solutions. He first reminded the participants that although common standards were achieved quite rapidly in the document production of e-Passports across ICAO Member States, it took more than 10 years after that to ensure that the inspection systems used to read and control the newly developed e-Passports were converging into a harmonised, interoperable way during border crossing controls. This slowly developing harmonisation, through the generalisation of PKD services by ICAO, can induce serious security holes if not thoroughly implemented. Such gaps are worth addressing, even if not exclusively related to one of the topics of the conference - morphing. In terms of legislative synchronisation, he also used the example of the newly developed EU Entry-Exit System which will be able to cross-reference different database requests originating from different Member States, including for asylum seekers, while legislation needs to be supporting and synchronised in this area as well.

Thirdly, in terms of the ubiquity of biometrics usage, biometrics obviously will not only be used at borders, with all kinds of potential other applications including banking and healthcare. If we then look at this in the context of the "time to deter" and the famous "4 seconds to decide" at the border whether to admit someone or not, there is now the possibility of using information gathered before a person arrives at a border from other biometric usage points.

Mr Salomon also mentioned the importance of the enrolment process in biometrics, as per the well-known saying "garbage in, garbage out". He stated that if the biometric data collected at enrolment is of poorer quality, then the benefits of this data in future use will decrease rapidly. At least in theory it is possible to link biometrics at the border with big data information, e.g. PNR or credit card information, which is already the case for some ICAO Member States' air programs. However, on top of their intrinsic complexity, for such programs to be effective even to a limited extent would require harmonised legislation across jurisdictions with respect to data privacy, which makes it a tough problem to solve.

Mr Salomon concluded by outlining one other area where he believes that Frontex could play an important role. This is the issue of continuous training. How to keep border guards appropriately trained when there are such changing environments in terms of traffic volumes and extended risk factors is a major challenge, even as intensive staff recruitment is underway. In this context, he mentioned the example where iris enrolment is already perceived to be more "difficult" to do by the enrollers. He stated that the EAB was very much ready to help in this area in a potential "train the trainer of the trainer" role, because of the large time constraints needed to deploy Frontex forces.

The moderator thanked Mr Salomon and raised with him the issue of border authority acceptance of the possibility that some third



party would be responsible for collecting biometric data which they would then use. He stated that this is normally rejected as being "dangerous" and that it "can't be done" and that border authorities have to be solely responsible for this biometric data collection, which then by definition means that it can't be done "away from the border". Mr Salomon said this is the crux of the challenge for everyone. Education and training are only part of the solution, but what is needed is to create solid links between the different stakeholders. He reminded the audience that the airline industry is already good at this with their global exchange of data and went on to say that the airport ended up not being a "combat zone" fighting for data handling privileges. Their success in this area is a good example of what can be achieved if there is appropriate cooperation.

Mr Salomon went on to say that he sees Frontex as having the potential to be a leader in this area, by its experience in controlling and synchronising many of the stakeholders' aspects of operational risk assessment, a prerequisite for success in his initial summary that "cooperation is key under controlled trust".

Hans de Moel

Mr de Moel started by saying that initially he would like to focus on the issue that if you want to use biometrics responsibly and ethically, then you need proper data. He stated that he believes personally that ICAO had made a mistake in not mandating more than one biometric data group at the start on the e-Passport. Furthermore, Data Group 2 is defined as 'Encoded Face' (biometric template) and Data Group 5 as 'Displayed Portrait' (image). Currently Data Group 2 contains not a biometric template, but an image. So, the photograph in the traveller's document is sometimes of too low a quality for it be used properly.

Mr de Moel then returned to his opening analogy of the responsible use of alcoholic beverages. His point in this analogy is that even the gold standard of whisky – a bottle of premium single malt Scotch whisky – can be misused; for example, in Western films when we often see a bottle of whisky used to hit another character over the head. This is not the way a bottle of premium whisky should be used. Biometrics are a tool which can also be misused. Just like we would normally not mix a premium single malt Scotch whisky with Coca-Cola to make a cocktail, we normally would use a blended whisky for that, biometrics needs to work in the right environment and in the right way for it to be effective.

He then went on to present the evolution in the work of the Biometric Institute, comparing it to the evolution of the role of Frontex. Mr de Moel stated that one of the first tasks of the Biometric Institute was simply to connect people from end user, academia/research and industry. Later, as their organisation grew, they then started to share knowledge – so that people did not need to reinvent the wheel when it wasn't necessary. Knowledge should rather be shared so that everyone can benefit from it through cooperation. Now they are at the stage of thought leadership, so stating that this is what the new role of Frontex could be, preparing for new European legislation etc. He mentioned that the Institute has many different working groups and that he is the head of the borders one. In April they hold a conference on "Identities at the border" where they invite a mix of stakeholders to share what they are doing and learn from each other.

To conclude Mr de Moel said that Frontex should facilitate, guide but also lead in the area of biometrics at the border.

Laurent Beslay

To begin Mr Beslay introduced the institute he represents, describing it as the Commission's "in-house science service". They support in particular DG Home in policy development in the field of biometrics and border management and control, amongst of course many other supporting scientific services. He then returned to the theme of his initial statement: "Quality". Mr Beslay said that he believed personally that border management was primarily about knowledge management – so it is about ensuring that the right information is delivered to the right person at the right time, and the role of research is to support this by helping to provide the right tools.

In terms of quality, he stated that he would look at this first of all from the point of view of assessment. So how can we assess accurately and appropriately the quality level? To do this, tools are needed. Currently they are trying to contribute a quality metric algorithm, some other ones exist in this field, including from NIST, but also from the private sector, however Mr Beslay stated that there are still gaps in this area, e.g. on face. To develop these tools researchers need data sets with real representative field data, which he acknowledged is a challenge but one which can and must be overcome.

He then raised the issue of management. Operational activity needs to be fitted around the quality of the data, so the assessment process is crucial to know what the quality level actually is. So even though it is right to strive for higher quality, even imperfect quality can be useful if the operational procedures are calibrated to fit around this fact.

Thirdly Mr Beslay mentioned the concept of improvement. Research here has a role to play in improving the quality of even current processes, e.g. enrolment, with research conducted on hew ways of collecting biometric

modalities, for example using touchless technology. On this point he specifically welcomed the news about the tests being conducted at Lisbon airport. He also stated that improvements can be stimulated by competitions provided they use real data sets. In the context of improvements and the Joint Research Centre Mr Beslay also raised the recent work they have undertaken on the effects of ageing on fingerprint biometrics.

In conclusion he stated he wanted to raise 3 to 4 points. Firstly, he believes that privacy and data protection will in fact help them to improve quality, as it will enhance compliance and the quality of the data and hence how the systems work. Secondly Frontex should be a bridge between the researcher and the practitioner/end user. He stated that it was essential that they worked closer to each other, and he again mentioned in this context the issue of access to real data sets. Thirdly Mr Beslay also sees Frontex as a lighthouse –showing the direction and guiding in terms of priorities.

The Moderator thanked the panellists for their statements and turned to questions which had been submitted on-line from the audience.

Q1: How can we create technical solutions for Member States when Industry doesn't see a benefit to this?

Mr de Moel answered this question by suggesting the Horizon 2020 route, which is where Industry is sponsored by the European Commission to come up with solutions to selected problems. The problem needs to be defined and sent to the European Commission. The Moderator agreed with this but also stated that he believed that the user community had a certain expectation that Industry would also invest its own money in research, where the research programmes can be developed after discussions with Frontex about where the gaps are in the current processes.

Q2: AI is not only set by humans; what about machine learning and cognitive analysis?

Mr Salomon answered this question by stating that what is needed is a truly independent and unbiased assessment of the algorithms, for example control software which can detect any potential decision bias. Currently AI is not a panacea and that there must be control and trust, eventually reverting to humans.

Q3: How do we ensure that the introduction of new technology does not become a burden for border guards?

Ms Kalnaja answered this question by saying that if the border guards have set the needs and participated in the dialogue from the beginning of the process then this situation shouldn't arise. The technology should be seen as a tool to help them do their jobs. She also reiterated the importance of training, where she stated that it shouldn't only focus on today, which is the tendency in the law enforcement community. Instead of running behind the flow of technological development trying to catch up, rather the emphasis should be more on trying to identify and define the future up front and train for that. This is in fact not so difficult to organise, because research and development timelines are long, so this gives plenty of time for advance training. This will hopefully help to avoid the situation described in the question, however Ms Abdennebi emphasised that technology is not a choice – it is a necessity and has to be faced up to by the border guard community.

Q 4. Why do so few states use the electronic credentials of the e-Passport?

Ms Abdennebi replied that there are currently 70 nation-state participants; ICAO is actively encouraging further take-up and a master list will be published before the end of this year. She emphasised that states are investing in

this, but that it is perceived to be expensive so there is still work to do to encourage take-up.

Q5. Where is the decision made? Will machines take decisions, and what then is the future of the border guard?

Mr Beslay answered that in his view currently a decision only taken by a machine is not permitted legally. So for him, technology can offer great assistance tools, but they are there to help the real border guard to take a decision, and his institution's published recommendations for e.g. using the new upgraded SIS database follow this approach.

Q6. The technology is ready but legislation and legal systems lag behind. How can we change this?

Mr de Moel answered by saying that he partly agreed with the thesis inherent in this question. Technology is often ready but often legal systems don't allow it to be used. So what is needed is better coordination between legislation and technology. There is a big gap between politicians, lawyers and technology which contributes to this. As an example of this problem he shared the contradiction between the UN Security Council Resolution stating that information about a known terrorist suspect must be shared but that the sharing of biometric data is not allowed by the EU's data protection regulations. This is not such an issue outside Europe, but the general point remains that there are often gaps between legislation and the end user's needs.

At this point the Moderator closed the panel discussion and thanked the panellists for their input.

THEMATIC SESSION 1

The challenge of morphing for border control

Moderator

Dr Joseph Atick · Executive Chairman, ID4Africa

Panellists

Matteo Ferrara, Ph.D. · Department of Computer Science and Engineering, University of Bologna

Mei Lee Ngan · Computer Scientist, National Institute of Standards and Technology

Ronald Belser · Research and Development Advisor, National Office for Identity Data, Dutch Ministry of the Interior and Kingdom Relations

Prof. Dr Christoph Busch · Biometrics Laboratory, Norwegian University of Science & Technology and Hochschule Darmstadt, Germany

The Moderator introduced the session. He began by stating that he had been involved in the panels and decisions related to the choice by ICAO of face recognition as the main biometric for e-Passports between 1996 and 1998. At that time the three main challenges for face recognition technology were a) accuracy was not that high b) there was a lack of data sets of digital photographs for comparisons, and c)

there was huge concern about privacy and this technology was often seen through the paradigm of surveillance. At that time, however, the technology was not sufficiently developed for there to be real problem with surveillance.

The moderator went on to say that today obviously the world is very different due to technological development and that now he sees two key new challenges:

a) Face recognition has become very accurate and this in turn has led to potential applications of the technology which had not necessarily been foreseen, e.g. mass surveillance.
b) New vulnerabilities of face recognition have been uncovered even in legitimate applications

Mr Atick then introduced the main topic of this thematic session, namely the new vulnerabilities to the process of border control represented by the morphing of photographs in e-Passports – and he reminded the audience there are now more than 1 billion in the world today. The moderator then introduced the first presentation.

Dr Matteo Ferrara - The challenge of Morphing for border control

Dr Ferrara was one of the team of people who in 2014 first researched the vulnerability to morphing attacks of automated border gates. He started his presentation by giving a definition of morphing: it is not always related to facial changes, e.g. it is a special effects technique in the film industry. However, closer to home for border control he gave this working definition: "morphing is a special effect that transforms an image into another through seamless transition".

He then laid out a standardised process that a criminal who wishes to conceal their identity and travel by air could follow, when he knows that he cannot use a regular e-Passport. The technique does not involve doctoring or changing an already issued passport, but an attempt to use a morphed facial image during the proper e passport issuing process. This procedure requires the criminal to find a suitable accomplice who doesn't have a criminal record, whose face can be "morphed" with his own – in this way "hiding" some elements of the criminal's face as it is now mixed with some elements of the accomplice's facial image. In this way the e-Passport issuing office can issue a regular passport with a falsified photograph which can be used by either the accomplice or the criminal to pass an automated border check. It has been shown that it is possible using this technique to not only fool the issuing officer at the passport office but also the algorithms used at the automated gate at airport borders.

Dr Ferrara then went on to introduce the concept of the Morphing Factor or alpha which is the percentage of the criminal's face characteristics which are to be found in the morphed image. This factor needs to be addressed correctly so as fool both the passport issuing officer and the automatic border gates. They believe currently that the best-balanced morphing factor is between 20% and 30%.

Dr Ferrara emphasised that morphing in this context is not just a theoretical problem, but that it has already happened in the real world – and gave the example of a German activist's image morphed with the face of Federica Mogherini, then the EU High Representative for Foreign Affairs. This activist succeeded in using a morphed photograph to apply for a genuine German passport. Apart from this example it is not known whether and how many other successful morph attacks have been made on borders.

Dr Ferrara then presented various scenarios for how morphing detection may work. First of all, based on a single image, it should be possible for a morphing detection technology

to ascertain if this image has been morphed or not. Secondly if there was a second image available which was "trusted", i.e. it was taken "live" by a border officer, this could then be compared with the morphed image and the morphing could maybe then be detected using this comparison. He then presented more detail on the different technical solutions currently under consideration for morphing detection.

a) Micro- textual analysis/facial image colour
b) Topological learning – facial features/ shape analysis
c) Deep learning techniques

An additional option is to attempt a de-morphing process, so as to recover the original 2 images.

He concluded his presentation by saying that all of these approaches are potentially valuable – however none of them has reached an acceptable level of reliability due to three main reasons. Firstly, intra subject variations are often stronger and bigger than the changes introduced by morphing. Thus, over the standard passport lifespan of 10 years a subject's facial appearance can change significantly in terms of hair amount/style, facial hair, make-up, ageing process etc. These changes can be more pronounced than face morphing. Secondly the e-Passport issuance process often involves the applicant having a printed version of their photograph which the government officer then scans to get a digital version. This process in itself of printing and scanning erases most the small variations introduced by morphing, thus making detection on these images much harder. Finally, another problem is a lack of publicly available databases with morphed images, which have been made using a variety of techniques. This hampers research efforts as no single research group has a big enough "supply" of images to develop and test anti-morphing methods which work across different morphing techniques.

Ronald Belser – Morphing: The Invisible Risk

Mr Belser started by stating that morphing in the context of border control is a relatively new and undocumented phenomenon. He posed the question whether it is better to see morphing as a border control issue or is it actually a document issuance issue?

This speaker looks at the problem primarily from the perspective of his institute, which is involved in the issuance of travel documents, and he claims that it is indeed a very serious threat. They are working on finding any photo that has been morphed before document issuance, however, as has been stated previously there is currently no fool proof method of detection. Additionally, Mr Belser stated that there is currently no automated machine that can reliably detect morphs and that there is a lack of awareness of the problem amongst civil servants, with significant training needed. He then shared with the audience the various activities currently going on in the Netherlands to deal with the problem of morphing.

First of all, he mentioned an international workshop on morphing which was organised in 2017 and involved government, industry and international organisations and was an important step in information exchange and also in the attempts to get this topic onto the international agenda. After this workshop a consortium was formed, with the primary partners being the German Bundeskriminalamt, the Dutch National Office of Identity Data and Universities from Germany, Norway, Italy and the Netherlands who jointly made a successful bid for research funding from the EU and are now working on this SOTAMD project since February 2019. Mr Belser also mentioned a PhD thesis from Twente University on the topic of Morphing detection which should be completed by 2022.

He concluded his presentation by stating his belief that the problem of morphing will only be managed effectively if there is strong co-operation between government, industry, international bodies and research and academia.

Mei Lee Ngan - NIST FRVT MORPH: An Ongoing Morph Detection Evaluation

First of all, Ms Ngan gave a short introduction to the work of NIST in biometrics which has been going on for more than 50 years and in the area of face recognition for more than 20 years. To start with, she stated that the problem was that if the right tools are used with the right subjects it is possible to create disturbingly good morphed images that can look like both subjects. These morphs can fool not only humans but also automated face recognition systems. Since this issue was first highlighted by Dr Ferrara in 2014, research has been ongoing. Ms Ngan showed some results of NIST's ongoing tests of the automated border gates algorithms, including the threshold for the acceptance of facial images, including morphed ones. This shows that a significant number of morphed photos will be accepted by the face recognition algorithms in 2019, so in other words the problem still exists.

The speaker then went on to introduce NIST's FRVT Morph Detection Evaluation program, which evaluates various software solutions for detecting morphed facial images, with the first draft report of this work published recently for public review. Ms Ngan went into more detail concerning the research's findings on single image morph detection, and 2 image differential morph detection. In this research we have information both on those morphed images which are not detected (morph miss rate -APCER) as well as genuine images which are defined as being morphed (false detection rate -BPCER).

The tests also considered three levels of quality in terms of morphed images, firstly "less sophisticated morphs", then "better morphs" and finally "uncomfortably good" morphs. She went on to explain that in the context of border control one of the most important metrics is the false detection rate, as this has a direct bearing on the human resources required to clear up incorrect information from the automated gates.

So far, all the algorithms tested are from the Universities represented at the conference, however they are open to commercial companies, whom they encourage to take part by submitting their algorithms for testing. From the results presented we can generally conclude that morphing detection technology is still in its infancy as the detection rates are still at unacceptable levels for field deployment. The research continues, and Ms Ngan particularly encourages greater engagement from the commercial sector.

Professor Dr Christoph Busch - The Challenge of Morphing for Border Control

After thanking governmental and other supporters of the research programs in the anti-morphing area, Professor Busch went on to ask whether this vulnerability is a real problem in the context of border control. And if it is a real problem, what can be done about it and what will its impact be?

First of all, Professor Busch said that he believes it is a real problem and that this vulnerability is likely to already be of interest to people smugglers and facilitators of illegal immigration from outside the EU, so actions should certainly be taken to mitigate it. He went on to remind the audience that the potential problem of morphing in this space has been known since 2009 and of course was proven by the Bologna team in 2014. So it should not be a surprise.

Professor Busch then went on to give a short history of the work done to deal with this problem. He spoke about the Fidelity Project from 2014, which concluded with the main recommendation that all European countries should move to live enrolment for passport applications as soon as possible. This has now been initiated by only a very few countries.

He then turned his attention to what should be done now for the future. He broke these ideas down into a number of action points as follows:

1) Build consensus among stakeholders as to what should be done to secure the trusted link between an e-Passport and its holder is maintained or even strengthened, and that anti-morphing attack technological solutions need to be deployed. He then mentioned in this context the multi-stakeholder consortium of which he is a member which is working in this area: the iMARS Consortium, which brings together the institutions from which most of the speakers on this panel work, in academia, plus government agencies and the European Association of Biometrics. This consortium has already formulated an action

plan which should start as soon as possible, and the points from this action plan (MAD action plan) were then presented in the next slides.

2) Standardise the passport application process: Here it was recommended that the EU issue a mandate to move to a live enrolment passport application process in all countries and that this should be combined with technology that can detect photos that have been manipulated or morphed for whatever reason, as well as defences against presentation (silicon mask) attacks and with a device certification scheme.
3) Even after live enrolment for e-Passport applications is introduced there will still be passports with morphed photos in circulation, so image pair detection algorithms should be applied at the border to detect a morphed photo in a differential scenario (morphed photo compared with trusted photo), in addition to single image MAD solutions when there is only the morphed passport photo to assess.
4) More subjects in the databases – more ethnic variety, more morphing tools, a bigger range in the quality of the images. He proposed a joint effort to enlarge the database along these lines. Then testing using the on-line evaluation scheme of the University of Bologna (BOEP). The technical interfaces will be in line with the NIST Face Recognition Vendor test morph competition.
5) Standardise testing of MAD solutions. Find consensus in the testing community globally of standardised metrics and processes for assessing morphed facial images. In the end this can become an agreed ISO standard.

Professor Busch then concluded that in the end a standardised suite of predictive software will be needed which can be used at the border to analyse the quality of facial images and also give feedback as to what exactly should be done by the operator to make the image capture better in a biometrics on the move image capture scenario, in terms of pose, lighting, illumination and other variables.

Finally, he mentioned the importance of training staff to foster best practices at the border, where he hopes that Frontex and the iMARS consortium can be involved. He also mentioned the need for the training of communication specialists in this area, as this issue is likely to get more media coverage in the near future.

The Moderator then asked the audience for direct questions from the floor.

The first question, from a representative of the French border guard authority, was regarding the current lack of minimum standards as to the results achieved by existing face recognition devices. He stated that each producer currently set their own standard. This obviously hinders interoperability, even within the EU. Professor Busch answered by saying that this is indeed a problem and that the same principle should be applied as when the current standard for fingerprint biometrics was set. The process for setting a uniform standard for face recognition started with a proposal from NIST on the table, which has been initially accepted, and that he thinks that in 2 to 3 years it should be ready to be implemented as an ISO. On this point Mr Belser stated that in the case of the enrolment process there are certain guidelines in the area of the issuance of ID and travel documents at a national Member State level contained in relevant European Union guidelines (Article 6) but that these are only guidelines and are not currently mandatory.

The next question from the audience was whether, given that we are assuming that better quality facial images will improve both face recognition and morph attack detection,

has it been considered that this could also create a drawback in the form of bigger/better images leading to longer "read" times at the border, and thus longer passenger processing times which will thus effect operations? The questioner asked whether it was likely that in setting standards a balance would be struck regarding image quality and the size of the digital file. In answer to this Professor Busch stated that indeed the capacity of the chip in the e-Passport was an issue in 2005 when the first guidelines for the facial image and fingerprint were set in terms of picture resolution, however he also stated that it could be possible in the future to have two images stored, a basic one on current guidelines and a high-quality one which would only be "read" at a second line inspection if the base one had tripped a morph detector.

After a follow-up question from the moderator, Professor Busch confirmed that to date no research has been done on the relationship between facial image quality and the morphing detection rate, but that as part of the MAD plan he wanted this to be dealt with in the future. Ms Ngan from NIST confirmed that they plan to run their tests starting with high-quality, high-resolution images and then gradually decrease the quality threshold, and this is something they are looking at. At this point she also raised the issue of face recognition technology and twins. Currently the technology is unable to differentiate between twins. In the past it was believed that a skin texture algorithm had been developed that could make this differentiation but only on very high-resolution images. Ms Ngan challenged the face recognition community to consider whether they needed to solve the issue of better distinguishing between twins and even siblings in order to better solve the issue of morphs.

The Moderator then further questioned the issue of skin texture, asking Dr Ferrara whether he was testing skin texture algorithms vs shape or deep learning. Dr Ferrera answered that skin texture analysis is generally better than shape algorithms, and he reiterated that the quality of the image is very important in this area as well. The printing and scanning process in itself normally renders any skin texture differentiation analysis impossible as this level of detail has been lost. Dr Ferrera went on to state that hardware processing speeds are increasing so rapidly that in a relatively short period of time it should be possible to read and analyse high resolution facial images in real time – so this issue will evolve.

A further question from the floor asked whether morphed images had ever been tested against images of the individuals when they were older – e.g. 10 years later. Ms Ngan stated that at a following panel they will be discussing a similar project where they searched through databases for matches to morphed photographs. Dr Ferrera stated that he was not aware of any research that is directly in this area, however he emphasised that changes due to ageing are normally much bigger than the changes that are attempted on morphed images and that this is the critical issue.

Then a questioner asked whether there was any cross referencing done in terms of research between these biometric based studies and other studies being conducted on "recognition" of neural structures or molecules, so there may be other scientific areas which could yield useful input. Professor Busch mentioned in his answer some work that has been done on multi-sensor recognition which could yield some valuable learnings for the biometrics area.

As a final question a delegate asked about whether image "beautification" programmes could become a problem in the future in the morphing area in that a photo's "improvement" for genuine reasons could be interpreted as a morphing issue. The moderator stated that if a passport application was made with a "beautified" image then this should maybe be rejected at this moment of enrolment as it is too far away from reality. Professor Busch said any manipulation of the image was a problem – whatever its cause. As with other issues of this kind there are only two solutions: either live enrolment or an algorithm that can detect such manipulations.

The moderator then thanked the panel and closed the session.

THEMATIC SESSION 2

National approaches to prevent and detect morphing

Moderator

Dinusha Frings · Research Manager, National Office for Identity Data, Dutch Ministry of the Interior and Kingdom Relations

Panellists

Dr Gert Jan de Nijs · Senior Project Manager, Dutch Vehicle Authority

Kari Kanto · Senior Advisor, National Police Board of Finland

Dr Uwe Seidel · Senior Scientific Director, Head of Section KT 5 – IT Forensics and Documents, German Federal Criminal Police Office

Dr Rebecca Heyer · Department of Defence Science and Technology, Australian Government

As an introduction the Moderator stated that this session was designed to present interesting and innovative ideas and tools related to anti-morphing activity in the context of identity documents which have been initiated by various Member States and other countries. She then introduced the first speaker.

Dr Gert Jan de Nijs - Deduplication driver license database: a research study

The presenter introduced his talk by saying that his project relates to the issuance of driving licences in the Netherlands (which is treated as a National Identity Document) but that although not strictly related to borders it does deal with biometrics and face recognition issues. In the Netherlands driving licences are issued centrally and they operate a data-base including photographs of about 12 million subjects. In fact, Dr de Nijs stated that since 2017 they are using face matching as part of the process for driver licence renewal, checking if the historical photo corresponds to the one presented in the new application.

They are also looking for duplicate photos in their database, and the software has been finding them. They believe in the majority of cases the duplicates are genuine and can be explained as either being twins or "look-alikes". When the research was initiated, they also thought that human error may play a part and they are looking for what are the main reasons in proportion to the other potential reasons. They also added morphing detection after their invitation to the conference.

Dr de Nijs then set out how they approached the duplicate image search, cross referenced to date of birth, which in the end took approximately 10 weeks to conduct. For the purposes of their analysis they defined a duplicate when it was found to be over a certain confidence threshold. As expected, the main reason for duplicates would appear to be twins, however their research also showed that when the required confidence ratio increased above a certain point the twin ratio started to drop. As of now they do not have an explanation for this. They found that if the confidence ratio is set at 0.9, then 0,11% of the database may be twins but this number would then need to be filtered. After this twin filtering they ended up with about 1800 people out of a 12 million-person database. The analysis of this is currently ongoing but he stated that they have already found errors resulting from previous applications, twins and "look-alikes".

As an offshoot of this research they also conducted a fraud simulation test using morphed pictures. Their hypothetical scenario was that a Person A (with photo) was attempting to get a driver's licence using the identity of Person B. To conduct this test, they took volunteers and then took a new photo of them, which they then morphed with another photo based on a 50/50 alpha. They then conducted a hypothetical driving licence renewal application using the morphed photo for about 100 people.

Dr de Nijs presented the summary findings of this research where in two of the five samples there was information relevant for further analysis.

The Moderator then introduced the next speaker, Mr Kari Kanto.

Kari Kanto – Morph detection activities at the Finnish Police

Mr Kanto introduced his presentation by saying that they have been doing tests in this area for a few years now. They had access to 500,000 photos from their registry, and a variety of morphing algorithms, some homemade and others bought commercially.

For the single image scenario, they tested deep convolutional neural networks, however Mr Kanto believes the results indicate that there are no universal signs of morphing common to all morphing algorithms, and that any single-image morph detector is therefore bound to be more or less algorithm-specific. The algorithm-specific artefacts are also easy to remove.

For the image pair scenario, on the other hand, they found that the so-called de-morphing method is very vulnerable to "benign" differences between the images, such as age and picture quality, which greatly raised the number of false alerts, thus making the method operationally unfeasible.

They also tested how vulnerable newer face recognition algorithms are to morphing attacks. Here the results are a bit more optimistic meaning that continued development in this area may lead to a reduction in the "space" where morphing attacks can currently potentially thrive. The latest tests make use of Generative Adversarial Networks (GAN).

Mr Kanto then introduced a platform they are currently working on which is a centralised Image Analysis Server. This is a system which can be accessed by other systems to check image quality and look for signs of manipulation, including morphing. The analyses are performed by modules that are easy to add or replace, with potentially several different modules doing similar tests with different algorithms. Following on from this Mr Kanto then proposed that it might be a good idea to create a common interface specification for morph detection tools – a de-facto standard. This would speed up the testing, as many organisations could participate in it and if a good solution is found it can be spread more quickly. As part of this there could also be shared repositories/trusted hubs for downloading. He suggested that this could be organised by Frontex or Europol and stated that the dissemination of tools should be restricted to trusted parties in order to avoid "broadcasting" the exact capability of law enforcement in this area to potential criminals. Mr Kanto also believes that Frontex and Europol could be made responsible for collating and managing the data sets that are needed to effectively test the new algorithms.

The Moderator then introduced the next speaker.

Dr Uwe Seidel - National approaches to prevent and detect morphing

Dr Seidel started his presentation by reminding the audience about when this issue officially hit the news in Germany and Europe last

year, with the morphing of a photo of Federica Mogherini with a German activist which led to the morphed image being used on successful German passport application. Although of course the issue of morphing attacks per se has been known about for longer.

In their response to this stunt the German Ministry of the Interior stated that they are considering the issue of live enrolment for passport applications, however the speaker pointed out to the audience that as this is a devolved state responsibility in Germany (so beyond the direct remit of the Federal government) it would be challenging to change the methodology as there are 5300 offices for ID and passport applications right now, with multiple software vendors and many regional data centres. However, there is some good news in that there is a common software core to all the deployed solutions which deals with biometric enrolment.

The speaker then went into more detail as to how the current enrolment process works in Germany. For the vast majority of the enrolment offices (97.3%) the passport image is taken at a photographer and later on scanned in the municipality offices – only a very small minority of offices currently offer live enrolment. He then went on to explain that for the Federal Government the anti-morphing strategy requires a two-fold approach. Firstly, security needs to be improved for national passport enrolment, which primarily means investments in live enrolment kiosks (or certified photographers) and also developing the appropriate legislation. Secondly, for the control of third country national passports, investments must be made in anti-morphing technological solutions. He then set a task for the audience asking them for their opinion on what would be an acceptable false alarm rate at the border. He said this information would be used to help them with a research project they are currently working on.

Dr Seidel then went on to summarise the legislation being worked on in this area in Germany and at the EU level. At the EU level there is currently a new regulation for identity cards, but which could also have implications for passports where it says that Member States "could consider" live enrolment. He personally sees this as being too weak a recommendation but is aware that this was due to Member State input. In Germany work on legislation to make live enrolment for fingerprints and facial images mandatory for both ID cards and passports is currently being worked on and

this will be treated as a Federal initiative including from a financing perspective.

As a final part of the presentation, the speaker then gave a summary of current research projects in the European Union in this space. This includes the current EU-supported consortium working in this area (previously referred to partners) – the EU -ISF-SOTAMD project and the H2020 funding round which specifically calls for projects in the area of morphing detection. In addition to this of course there are also projects happening at the Member State level, for example in Germany the BSI, which tests various technologies and defences against morphing and presentation attacks.

The Moderator then introduced the final speaker on this panel.

Dr Rebecca Heyer - Morph detection by humans: an Australian perspective

Dr Heyer started her presentation by saying that her focus today was on the human factor and the role that humans can and do play in morph detection. She said that in the past her department had been asked by the Government to assess software packages that claimed to detect morphs but that their experience at that time (2014) was that none of them were effective and that to a large extent they relied on humans to interpret their alerts. They also found that not only was the technology not consistent but that the humans were also not consistent in this area.

Their interest then moved more specifically to the human factor. Firstly, they looked at what they call the issue of a "Morph persona", this being a situation where an identity document has been issued to someone using a morphed image and that this document then comes up for renewal. Their conclusion was that neither face recognition software nor human experts were likely to spot these morph personas. They decided to try to build further on the knowledge already there about the human capabilities in morph detection, in particular what variables impact upon human detection, what are the tell-tale features that might lead to detection and finally how this information could be used operationally.

The speaker then presented this research project in more detail, showing first of all how they tested the method of presentation of the facial image, i.e. mocked up on an ID document vs on its own on a computer screen, then how they tested the face quality in terms of an original image vs a print/scanned one, then whether there was any difference in the results between novices and likely experts. As experts they chose people who were likely to be possible experts in this area from other law enforcement Agencies.

In detail they ran 300 detection trials: 100 face images on ID cards, 100 on a 4x6cm screen, and 100 on an 8x10cm screen. The trials ran with a total of 108 participants (66 novices and 42 experts) testing the variables as presented above. The participants were simply shown an image and asked if it was real or a fake. They were given no information about morphing previous to the test. If they said it was a fake they were then asked why. The morphed images in the test had not been extensively worked on so they were not "no effort" morphs.

On the issue of method of presentation, the general conclusion was that morph detection is hard across all formats. At least 90% of CGIs (totally computer generated) were detected, however 10% of these still passed if the image was on an ID document. For morphs the 2-person morph was definitely the hardest to detect and on an ID card was called out at the same rate as real faces. In terms of face quality of digital vs prints/scans it was seen as being consistently harder to detect a morph once

an image has been scanned and printed, and that this process also greatly increased the number of false detections. In terms of experts and novices, there were no significant differences in their morph detection rates except in the case of false detections, where experts were significantly less likely than the novices to call a real face a fake. Another difference was that the experts tended to be more precise and technical in their reasons for calling a fake, whereas the novices tended more to use their gut feeling.

An overall theme emerged in the results: the fakes were described as being "creepy", "lifeless" or "weird". Dr Heyer presented a hypothesis on this point which they call the "Uncanny valley". This presents robotic limbs, humanoids and humans on one axis with the other being likeable versus creepy. This shows that as we go from robotic arm to cartoon human figure there is a steady increase in likeability but that there then occurs a sharp drop in likeability (uncanny valley) when we get to humanoid robots with skin who are perceived to be very creepy and then again the likeability grows rapidly from this low base up to a regular human face at the top. Various morphs are to be found on this line as it increases rapidly but the best version of the fake face would seem to be those which are GAN generated, which do often seem to have a lifelike quality about them.

Dr Heyer then moved to the issue of how detection can be improved at an operational level, with the focus on training. She stated that training does seem to make a difference and that it should focus on specific features on a facial image, on telling the difference between "artefacts" related to morphing and those related to printing/scanning and also that "gut feel" is also important, at least in terms of asking for a second opinion, or referral to a technological solution. She also stated that other research (Robertson) has shown that general awareness of the issue does work – particularly in the case of low performers.

She concluded that technological solutions remain very important but that humans will always be needed, in particular in the context of remediation. So if technology detects an issue, who is then to verify this and deal with the issue on the ground? She believes that this should be considered now in these discussions.

The moderator then moved on to the question and answer session. The questions for this session were given on-line.

Question 1: Is there any work being done to track and record the actual cases of morphing at the borders or is the evidence so far anecdotal?

In answer to this Dr Seidel stated that he himself knew of between five and ten real cases at borders but that he believes this kind of figure understates the scale of the real problem. In the cases which he knows about the information was shared with relevant other countries on a bilateral basis. However, there is no "central database" of such cases currently.

Question 2: There was a clarification question regarding the Australian research presented by Dr Heyer.

She stated that the respondents in the research had no specific time limit when they were asked to look at various facial images, however the responses were still spontaneous.

Question 3: Is there any research on beautification manipulations vs morphing or other techniques?

Mr Kanto responded that their test data did include algorithmically "beautified" photos but that the effect was not conclusive. Dr Seidel stated that they may take this into account in future tests on the database.

Question 4: Concerning the issues of access of data sets: Given the problems in this area related to GDPR, it may be better to bring potential solutions to be tested to the recognised designated data sets – not the other way around.

Dr Seidel agreed that the idea of "bringing the algorithm to the data" was likely to be the best route in the EU.

Question 5: Do any of the face recognition software vendors have morphing detection as part of their package?

Mr Kanto stated in response that he had spoken about the topic with one vendor but they had not been interested. Dr Seidel stated that the commercial sector is not enthusiastic about this, probably because they are concerned that their current products will not do well in these tests. Dr Heyer stated that she believed vendors had not so far had success in this area and at the moment didn't perceive it as being a commercial priority.

Question 6: How can we ensure the credibility of the digital photo on a pan-European basis?

Mr Kanto replied that this is a political issue. Dr Seidel stated that this kind of issue in terms of general standards was part of the Fidelity program, where in his view the solutions from a technical point of view are ready and on the table, but it is now a question of legislation. The Moderator reminded the panel and audience that in Article 6 there are some guidelines in this area and that they hope that regulations will be developed further to deal with this issue.

To conclude the session, the moderator asked Dr Seidel what the feedback had been to him regarding an acceptable false rejection rate for morphing detection. He thanked the audience for all the responses and said there was a big range of answers from 0,1% to 10%, but that most people seemed to agree it should definitely be less than 2%. He said he may give more detail on this during the next day's session.

As a final question the moderator then asked Mr Kanto, given that he had been fairly pessimistic as to the chances of improving the detection of morphing from its current rather low level, how he saw the future in this area. Mr Kanto stated that he does in fact believe that there is a ceiling to what can be done when regular photos are used. In addition, he believes that expertly made morphs will remain difficult to detect. We need to hope however that these represent a small minority of the potential morphs.

The moderator then closed this session.

DAY 2

THEMATIC SESSION 3

Ongoing research in the area of morphing and morphing attack detection methods

Moderator

Mei Lee Ngan · Computer Scientist, National Institute of Standards and Technology

Panellists

Dr Sébastien Marcel · Senior Researcher, Head of Biometrics Security and Privacy group, Idiap Research Institute

Dr Andrey Makrushin · Postdoctoral Researcher, Advanced Multimedia Security Lab (AMSL), Otto von Guericke University of Magdeburg

Dr David J. Robertson · Lecturer in Psychology, School of Psychological Sciences and Health, University of Strathclyde

Ulrich Scherhag, Ph.D. Student · da/sec Biometrics and Internet-Security Research Group, Centre for Research in Security and Privacy (CRISP), and Darmstadt University of Applied Sciences

The Moderator introduced the session and the panellists. She then asked the first speaker to give his presentation.

Dr Sebastien Marcel - Vulnerability of face recognition to deep morphing: morphing with deep fakes

Dr Marcel introduced the Institute for which he works and the topic of his talk, which is Deep Morphing, or Deep Fakes. The concept of Deep Fakes has appeared more recently, mostly in social media where internet users create fake faces of politicians or celebrities.

Deep Fakes are made from videos not images, and are created by a computer – so they are much faster to create than the typical static image morph. They can also mimic movement and facial expressions. Deep fakes are created by taking a source video with a face and then selecting a target face in video format. Using GAN neural networks, the computer then teaches itself how to convert the target face into the source face. Then in the future when there is a new video including the target face, it is easy for the computer to swap this face for the source one in a well-edited way.

As examples of the kind of deep fakes he is talking about from social media he showed two examples with actors, one where Jennifer Lawrence's face had been replaced by Steve Buscemi's and another where Amy Adams' face had turned into that of Nicholas Cage. He stated that these examples looked like they were fun but were not really a threat to a person's identity as they are done for entertainment reasons. However, as a research project they decided to try and see whether face recognition algorithms could be fooled by morphs created using this technology.

Dr Marcel then proceeded to summarise how the research project was set up. They selected 16 pairs of people from a video database, swapped the faces within each pair using the GAN approach, and came out with two versions, a low quality and a high quality

one. He then presented some videos following the format of first the face donor face, then the original target face and finally the generated face.

After this he then went on to describe the Vulnerability Assessment tests which they carried out on these deep morphs -both low and high quality. They tested against 2 open source GAN face recognition algorithms, VGS and FaceNet, with the most important result being that for the first of these algorithms they incorrectly identified a deep morph as being genuine 85% of the time and for the second one 95%. He stated that interestingly as has been seen to be the case with presentation attacks the algorithm that is the most accurate in a normal environment is also the one which is the most vulnerable.

This speaker then moved on to the issue of whether it is possible to detect these deep morphs. Given that this is work in a new area, there is no real database to compare to and no reference for this work, they decided to create a very simple baseline which will be useful for future research. The research showed a relatively high detection rate for low-quality morphs but a problematic one for higher-quality morphs.

In conclusion, the speaker outlined that deep fakes already exist as deep morphs and they already can fool standard face recognition algorithms, and that generally the higher the quality of the deep fake the harder it is to detect. For the future he stated that more work needs to be done to understand better the technology of deep fakes in the context of setting the balance between two faces, as in this technology you are taking a face and changing it into another face, it is not (like in standard morphing) a blend of faces with proportions being retained from the 2 faces, it being rather a one to one swap of faces. They are therefore planning further work to see if it is possible to retain information from face No. 1 in the newly formed face using this technology. They also need to work on better detection methods for these deep morphs. Finally, he said they are developing an open project so that others should be able to check their algorithms on their data sets. On this point he completed his presentation.

Dr Andrey Makrushin - Distributed and GDPR/IPR compliant benchmarking of facial morphing attack detection services

Dr Makrushin introduced his presentation on benchmarking of morph attack detection technologies by emphasising that they must be GDPR and IPN compliant in the EU. He stated that he was presenting the joint work of five German government, industry and research institutions – which was part of a Federal government funded project "Ananas" on "Benchmarking of Morph Attack Detection Services".

The basis for this research project, Dr Makrushin stated, was the current status that human assessment was prone to error and unable to detect high quality morphs, and that automated face recognition systems tended to have very high false acceptance rates. Given this, new algorithms were needed in this field.

He went on to explain that there are at least five research groups working in this area but that comparison of their results was difficult given the image privacy restrictions of GDPR and also researchers may be unwilling to submit their algorithms for analysis due to IPR, as in fact it may be in their commercial interest not to hand over their algorithms.

Dr Makrushin then went into more detail to explain the restrictions of GDPR and IPR. With GDPR the standard way around these data restrictions is to bring the algorithm to the authorised database – not the other way around. For IPR issues source code is often not shared and the benchmarking tests must be carried out, keeping full respect to the rights of the code's owners.

According to this speaker there are currently two categories of benchmarks a) Public and independent but with the data unknown to the MAD developers, which means that it is hard to work with the "verdict" of such benchmarks (e.g. University of Bologna, NIST) b) individual benchmarks which are aimed at understanding and assessing a particular MAD development and allow for more tinkering with the test data to enable further development.

He contended that the public benchmarks are sometimes difficult to use due to the rigidity of their requirements and the standardised format required from the data sent to them. As an answer to these issues he then presented his consortium's approach, called "Ananas benchmarking infrastructure".

Dr Makrushin then presented the features of this proposed new benchmarking infrastructure, the main points in his view being that it is secure, GDPR compliant and more "user friendly" and transparent than other potential options in terms of the benchmarking analysis conducted. Different aspects of this benchmarking system are hosted by the various entities who make up the research consortium.

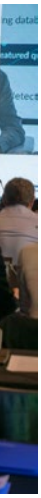
The speaker then presented some results of the test run of this benchmark from the end of 2018. In this they tested six MAD algorithms against 680 genuine images and 12 000 specially developed morphs. The key evaluation metrics were false positive rate and false negative rate.

To finish, Dr Makrushin gave his key conclusions about this benchmarking solution and the test that was run. This solution uses REST technology, allowing it to be fully GDPR/IPR compliant. MAD solutions can be used individually and not only as part of the framework, and the benchmarking results are reproducible and transparent. The test run only showed a part of the framework's capabilities. Dr Makrushin concluded by inviting anyone interested in getting more information about this benchmarking solution to contact him or others in the consortium.

Dr David J. Robertson - Morphed Passport Photo Detection by Human Observers

Dr Robertson started by stating that in his field, the variation possible within one face is of great interest, which he illustrated by photographs of actor Harrison Ford. Despite all the variations caused by age, hair style and pose, because the face is very familiar viewers ignore the variations and immediately find the shared attributes. So in this area, humans are experts.

Dr Robertson went on however to state that this "expertise" with familiar faces unfortunately doesn't transfer to new or unfamiliar faces. In fact, humans have very limited insight into how people's unfamiliar faces vary from each other, which can lead to easy false acceptance of faces being the same when in fact they are from two different people.



This of course causes "real world" problems, for example at passport control, and also in the criminal justice system. Research has shown that the standard error range for humans who are shown two unfamiliar faces which are "look-alikes" (not the same person) is between 10 and 30%. Dr Robertson then proceeded to present the top-line results of various studies that they have conducted in this area.

In Study 1 the research participants were shown in each set three passports, first the original genuine, second another genuine one of a person who looks rather like person 1(to check opportunity fraud), and thirdly a passport with a morphed image of the two people with different "levels" of morph percentages – however a 50/50 morph was taken as being the "gold standard" morphing proportion for the test. It is important to note that in this test the participants were not aware that they may be shown a morphed or manipulated photo. In terms of results around 8-10% of people wrongly claimed that person No. 2 was the same as person No. 1 (opportunistic fraud). However, 68% of people incorrectly stated that the 50/50 morphed image was the same person as person No. 1, clearly showing the potential scale of the problem with human assessment of morphed images.

In Study 2 the participants were given some knowledge about morphing and about certain tell-tale signs to look out for (particularly in low quality morphs) before the test. As in Study 1, they were then asked if there was a) a match, b) a mismatch or c) a morphed image. The results showed that again around 8% of respondents incorrectly stated that person No. 2 was the same as person No. 1 –without change from Study 1, however now the detection of the morph was much higher, with only 21% of respondents now claiming that the morphed image was the same as the image of person No. 1. In effect, in this test a morphed image was only 13% more realistic than a straight look-alike.

Dr Robertson stated that although obviously a 13% false acceptance rate is far too high for a security environment, the tests do show that even simple instructions/training can make a substantial difference in the effectiveness of human assessment of morphs.

In Study 2, which Dr Robertson then went on to present, the key issue addressed was whether the human ability to detect morphs could be improved or not. In this study, two groups, a) trained group and b) guidance group, were assessed for their morph detecting ability, with the trained group receiving a greater amount of basic short training in morph detection than the guidance group. The results showed that across the different levels of morphing percentages, the "trained" group were consistently better at morph detection than the control group, which suggests that investment in training on this issue could be worthwhile. Dr Robertson however pointed out that even for the trained group the false acceptance rates are still too high for a security environment. As another caveat he stated that when the performance of the research participants was looked at on an individual level, it was shown that the biggest improvements came for the people who were starting from the lowest levels of competence, and that for people who were better at the beginning the improvements were not necessarily statistically valid.

Dr Robertson then went on to look at the issues of whether "super recognisers" could become "super morph detectors". He stated that the general consensus in his academic discipline is that experience and training do not seem to improve face recognition. The general view is that face recognition ability in humans is an individual skill – much like singing, where you have small groups at the bottom and the top who are abnormally bad or gifted, with the bulk of people coming somewhere in the middle with average abilities. So, in face recognition there are "super recognisers" who

are unusually naturally good at these kinds of tasks. Thus, the question for this analysis was: if you are a "super recogniser" with outstanding face recognition abilities will you also be an outstanding morph detector? The analysis conducted by Dr Robertson using the base data from Studies 1 and 2 (see above) concludes that there is a moderate correlation and that more research should be done here to develop and potentially confirm this potential hypothesis further.

In conclusion Dr Robertson stated that it was difficult for humans to detect morphed images but that some awareness building/training seems to help make it more effective. There may be a link between human face recognition skills and morph detection dependent on further research, and current data supports the training of the staff dealing with these issues. However, as an intro into a potential discussion he added the caveats that his work has been done on low- to medium-quality morphs, so that it is possible that this affected the results; in addition, ethnicity may be a factor but they only tested Caucasian photographs. Dr Robertson would like to encourage users and practitioners to cooperate with academia, for example in sharing real cases. Finally, he stated that he was still in favour of a balanced approach to MAD using a mix of technology and humans, however with the highest quality morphs it is possible that these will always remain beyond the abilities of humans to detect.

Ulrich Scherhag - Face morphing detection: Issues and challenges

This speaker started by talking about the MAD schemes which are currently operating. First of all, there is no reference detection (Single Image MAD) where the assessor has no other image to compare the morphed image against. In this approach, after a phase of feature extraction and face pre-processing, the image is then analysed by an algorithm searching for morphs and we get an answer of either bona fide or morph. These algorithms can be useful at certain moments when there is only one image, for example at document issuance, but they tend to focus on certain artefacts, so with high quality morphs they may

not be effective. Another problem is that they can be influenced relatively easily by the process of printing and scanning. In the second MAD scheme, there is a reference-based detection scheme or differential MAD in which the facial image can be compared against a trusted probe image, taken for example at the e-gate. Then the differences between the 2 images are analysed by the algorithm. This approach is less susceptible to print/scanning issues as there is a comparison being made.

After presenting these two basic MAD schemes, Mr Scherhag then went on to categorise the MAD approaches under each scheme presented above. Under Single Image MAD there are categories of texture descriptors, forensic image analysis and deep-learning approaches, and under Differential MAD there is the comparison of two facial images and a reversal of the morphing process (de-morphing).

The speaker then talked about the general concept of MAD based on deep face representations using neural networks and how this works for morph detections based on deep face. This system also compares 2 images (potential morph and trusted live capture) and leads to the computer decision that the image is either bona fide or a morph. The speaker said methods based on this approach to neural networks use CNNs which are not trained on morphed images, so the problem of over-fitting is avoided and in addition they are trained to extract the relevant information from a facial image and are more robust against issues to do with variations in pose, lighting etc. between images.

Mr Scherhag then outlined the database requirements to evaluate such algorithms. A realistic database at the border should contain 2 images, firstly the passport image and secondly the probe image taken at some point at the gate, which is of lower quality. For their research they created a version of such a scenario, and the speaker presented some sample images both genuine and morphed using various algorithms. He stated that various experiments are still ongoing, however they already have some initial observations, namely that the performance levels are promising and they seem to be quite robust against print/scan issues, and compression, also that high-quality morphs that leave fewer artefacts are clearly harder to detect; the quality of the images that the algorithms are trained on doesn't seem to matter that much to the results; and finally the algorithms work best the better the quality of the probe image, so in fact the more realistic the probe image(with greater variation in illumination, pose etc.) the harder the task is for the algorithm. Finally, 2 further conclusions are that the SVMs seem to be the best machine learning algorithms of those used, and there is some evidence that in these specific MAD tasks the open source algorithms perform better than the commercial one, which is unusual in general in face recognition.

To conclude, Mr Scherhag raised a number of issues or challenges that he feels should be up for discussion, in particular the need for shared evaluation metrics and protocols, the need to be able to make general conclusions from specific MAD approaches and the issue of testing and access to real databases, and also the issue of transparency of results for others. This issue is currently unsatisfactory as there is often not enough information given about the algorithms, even in academic papers, for the results to be verified by others.

The Moderator then posed selected questions taken from the floor via the Slido application.

Question 1: Will trained humans be able to verify a morph alert provided by an algorithm?

In answer to this Dr Robertson said that it would appear that the best approach for now would be to try and pair the best algorithms with the best humans. He went on to say that in the situation outlined in the question it could be the case that the morph is so refined that it is beyond the capacity of the human perceptual system to add anything more in terms of verification etc, so it depends on the quality of the morph that the algorithm detected. However, he reiterated that it probably would be best to combine the best algorithm with best performing human.

Question 2: What training should be included for border guards regarding morph detection?

Dr Robertson replied that he believes that the existing test to find the best people in face recognition, the Cambridge Test, universally recognised as the gold standard in this area), should be somehow adapted to create a second gold standard test to find the super morph detectors. In terms of specific types of techniques to be used in training for morph detectors, he stated that we should bear in mind that the training so far that they have tested was rather basic and applied to rather low-quality morphs. More work needs to be done to find the cues that humans may be able to see on higher quality morphs. If this is done it could then be possibly implemented in training.

Question 3: Is trying to detect morphs a hopeless endeavour by either machine or human – given the pessimism shown by one panel member yesterday?

Dr Makrushin replied that he didn't see it like this and that in his view algorithms based on de-morphing have a real chance of being able to detect morphs at an applicable level of accuracy. Dr Robertson added that people should read the paper he referenced by Dr Kramer in the UK, which is published in open source so is publicly available – which was also rather pessimistic in its conclusions regarding the highest quality morphs. However, he emphasised that there needs to be greater cooperation between science and the end user practitioners, as currently we don't know how prevalent high-end morphs are, but that even without that he doesn't believe that it's a hopeless case, even looking at it from the human perspective.

Question 4: What does the research community need to create breakthroughs in morph detection?

Mr Scherhag answered that in his view frameworks for comparison of results were necessary, as well as access to databases with real (or realistic) images so that solid algorithms could be developed.

Question 5: What are the biggest barriers that research faces in this area?

Dr Makrushin answered by saying that the biggest barrier was always funding! If there were more funding, then better people would be able to create better algorithms.

Question 6: Are there any studies done on Super Recognisers to determine whether an image is morphed?

Dr Robertson answered this by saying that there is some research in the pipeline on this done in conjunction with the University of Greenwich. The results are just out but he hasn't seen the detailed data yet. It's too early to say if a connection can be proven between Super Recognisers and expertise in morph detection.

Question 7: Given privacy issues is there a problem with sharing images of public photos available on the internet?

Dr Robertson informed or reminded the audience that for the database that they work with, the participants gave consent to allow their images to be used in certain defined scientific research. He asked what the procedure was for the getting the database images for the researchers working on the algorithms. Dr Makrushin agreed that there were many public databases with images, but that the problem was they were often not ICAO compliant – so they were not the kind of images that are used for passports in terms of pose and quality etc.

The moderator then thanked the panellists and closed the session.

THEMATIC SESSION 4

The application of biometric technologies at our borders: An industry perspective

Moderator
Darek Saunders · Head of Sector *(acting)*, Border Security Research Observatory, Research and Innovation Unit, Frontex
Panellists
Lukasz Kubik · Secunet Security Networks AG
Jürgen Mathwich · T3K and Austrian Institute of Technology (AIT)
Brage Strand · Mobai
Jan-Willem ter Hennepe · Idemia

The moderator introduced the session and asked for the first speaker to deliver his presentation.

Lukasz Kubik - Biometrics and the EES

Mr Kubik started his presentation by saying that in his view one of the biggest challenges for biometrics at borders will be the ongoing work and implementation of the EU EES (entry exit system). First of all, he reminded the audience about what the EES is, namely a database which will register every entry and exit to the EU Schengen Area by third-country nationals. The system is designed to help monitor travel flows, manage visa overstayers and help with immigration decisions as well as help to fight cross-border crime.

Scheduled to start in 2022, the system will require the registration of all third-country nationals (biometrics: face and 4 fingerprints) including children (biometrics: faces only). This will require synchronisation with not only border authorities but also immigration, visa issuance and law enforcement at a national level. In terms of what information will be stored, each third-country national will have an individual file including personal data, document information and the biometrics. The file will be created at the border at the moment of first entry, including the live capture of biometrics, and then subsequently every visit will be recorded with an electronic stamp. The data will normally be kept for three years. Mr Kubik stated that the proposed live capture of biometrics at the border is in itself is a huge challenge given that the facial image must be similar in quality to the existing passport photo.

He then moved on to the challenges, which with a program of this scale are numerous. Firstly, the database itself will be vast, with twin tasks of identification and verification. He went through the specifications for facial images for which quality thresholds have been set by the European Commission, with the photos to be frontal image type compliant with the ISO standards. As an example of the challenge in this the speaker showed an example of a photo taken with an adjustable height function vs a fixed height. The system and ISO guidelines will give no or only limited possibilities to post process the photo image, so it is important that everything is right the first time as far as is possible. As for fingerprints the process is also a sizeable one with 4 fingerprint captures required, and pre-set quality norms will also have to be met.

In terms of biometrics capture at the border and the creation of the traveller file, it

will obviously take longer than is currently the case to cross the border and this in itself will create challenges for all the stakeholders in the travel chain. To illustrate this point the speaker presented the findings of a pilot project conducted in Germany in which compared to the current border crossing time the planned capture of four fingerprints and facial biometrics will increase the time spent at the border by three times per traveller. A very significant amount. As he said, attempts will have to be made to reduce this.

He then moved to the third part of his presentation, titled "Why this matters". In this section he first of all presented the issue of the critical importance of the quality of the facial image on biometric recognition performance as defined by NIST at the end of 2018. The general point is that to achieve fast processing times it is essential to have good quality biometrics in the system, as otherwise it will often require human verification with all the time aspects that this involves. So to achieve the low error rates that are essential the biometrics captured must be of high quality. This in turn requires certain quality ensuring steps to be followed while capturing the biometrics. This will of course be easier to achieve at an airport than at a land border either by car or train. For self-service kiosks there also must be very good and clear guidance to the travellers.

Jurgen Mathwich - Cutting-edge biometric technologies facilitating travel and securing borders

As an introduction, Mr Mathwich started by looking at the issue of security versus convenience at the border by asking: are these in conflict with one another or not? In terms of increasing security, this can lead to an increased number of checks, the time and issues related to establishing a unique identification, document fraud is on the rise and sophisticated criminals are more tech aware than ever. On the other hand, in terms of the bona fide travellers' expectations in terms of convenience at the border the expectation is free movement, frictionless security checks, short or preferably no queues, increased security and privacy protection. He presented this conundrum in the form of a triangle consisting of three constraints: Security, Privacy and Convenience – with his thesis being that you can't have all three maximised at once. If the demand is for more security, then privacy or convenience will be compromised. But if you want more convenience then likely security and privacy may be lowered. So, the speaker asked whether technology can help reduce this innate conflict, not to eradicate it but to reduce the power of these constraints.

Mr Mathwich then went on to set out how he sees the requirements for today's cutting-edge biometric solutions. He presented this under three headings: a) easy to use and fast, b) secure and integrated and c) mobile and smart. In terms of easy to use it is important that technology can be integrated into the existing workflow processes, for b) secure and integrated it needs to be interoperable, and for c) the mobile revolution should be taken advantage of.

The speaker then laid out in more detail what kind of technological solution should be considered, that it should use on the move biometrics for smartphone-based identification, it should have improved authentication and it should be effective against presentation attacks.

Mr Mathwich then presented as an example of the kind of technology he is referring to a product from his company: a mobile fingerprint capturing device which is basically software which then utilises the hardware already available in a smartphone – a high-quality camera and operating system. This device can quickly and easily capture fingerprints on the move, with passengers feasibly being able to scan their own fingerprints earlier. He then showed a video which demonstrated this same technology but in a dedicated robust hardware device, suitable for a border guard working in imperfect conditions, e.g. outside and in the cold.

To conclude, Mr Mathwich presented a vision of how biometrics may work at the borders of the future, using technologies which are still in the research phase such as iris at a distance, palm vein reading and passport reading at a distance. He believes that an appropriate suite of these technologies can help to reduce the constraints of security, privacy and convenience all pulling in opposite directions which he presented in the first slide.

Brage Strand - World leading attack detection on biometric systems

First of all, this speaker introduced the company, a spin off from the Norwegian Biometrics Laboratory at the Norwegian University of Science and Technology. Their specialisation in biometrics is Biometric Authentication

or stand-alone presentation attack detection using multi-modal biometrics of face, iris and ocular. They are also focused on the problem of face morphing detection and have a patent pending in this area. They are also participating in the benchmarking tests currently being carried out, such as NIST's FRVT.

Mr Strand then outlined what he sees as 3 Must-Win battles related to morphing. The first one is "Strong attack detection in the travel workflow". As part of this he emphasised the importance of training and the human input of border guards, which he contends will have a strong role in this area. It is important he stated that they themselves also become supporters and ambassadors for biometrics. The solution is also of course technological, in terms of morph attack detection algorithms and dealing with potential presentation attacks. Mr Strand stated that his company wanted to approach attack detection as a service, where various potential technological tools can provide support to existing processes, whether in document issuance or in ID authentication.

The second "must-win battle" is "Rethink the process of issuing passports". Mr Strand pointed out that only some countries have moved to live enrolment and that many are yet to take this step, so the risk is still high. He mentioned as an example the UK passport application portal, where the applicant can upload their own photo which is very convenient and user friendly and of course in his home country of Norway where they have now live enrolment. The key issue, he maintained, is how to maintain trust in the system. Trust is built on the system getting adequate information, being able to avoid fraud and attacks and remains cost efficient. Mr Strand went on to say that the gold standard was likely to be live enrolment but that there may be ways to improve the "user-friendly" on-line enrolment approaches by creating a kind of hybrid where live enrolment for first passports is supplemented by an on-line version with built in attack detection technology for passport renewals, lost passports etc. Whatever the solution, it is essential that trust in the documents and the system is maintained and built on.

Mr Strand then went on to present his third "must-win battle", which is "world class usability and collaboration". With the help of training, the border guards will be offering a total service to the traveller including technology solutions, so the total user experience is very important. He also believes that convenience is a very important aspect of biometrics, and this is often how other users of biometrics such as banks talk about it. Biometrics at the border can also perhaps be seen in this context as an enabler which will also connect to other stakeholders such as airports and airlines who are responsible for the total traveller experience.

The speaker then asked if there were any people in the audience who were end users of biometrics who were not in law enforcement or border guards. There was someone from the driving licence authority. Mr Strand then said that his point was that biometrics is not only an issue at borders but for a number of industries including other public administration and financial services and that collaboration between these sectors on the issue of morphing will also be very important, of course as well as collaboration on other issues which have been already raised at the conference such as datasets, benchmarking and between academia and end user communities. Mr Strand maintained that the more collaboration there is and the more "safe spaces" for exchange of ideas and testing, the shorter the lead times will be for the introduction of new technologies.

In conclusion, he challenged the audience to think about what they can do to deal with all three "must win battles" in their areas when they leave the conference.

Jan-Willem ter Hennepe - Application of biometric technologies at our borders: An industry perspective

The speaker introduced his presentation by saying that he will be taking the user perspective on these issues rather than a scientific one: the perspective of the traveller, the border guard or the airport. Recently he has been working on the upcoming EES programme, where he believes that the user perspective is also the most important.

Mr ter Hennepe then outlined some basic information about the Idemia company and its scale of operations. It is a global corporation with its headquarters in France and is one of the largest companies working in this area. In terms of biometrics some of their biggest projects are the Biometric Matching System of the Visa Information System in the EU, the AFIS systems for law enforcement agencies the world over and the gigantic ''AADHAAR' ABIS program in India.

He then proceeded to look more specifically at the issues of the traveller and border management by laying out his view on the key future challenges and opportunities. In terms of challenges, he mentioned increasing pressure on government budgets, the 11 billion air passengers expected by 2030, the 68 million lost or stolen travel documents every year and the continued threat of terrorist attacks. In terms of future opportunities there is the increased use of biometrics, further innovations in artificial intelligence and the better use of existing information about travellers, e.g. when the traveller leaves a plane and crosses a border on arrival, he/she is the same verified person they were when they crossed the border and boarded the plane at departure.

The speaker contended that given the scale of the issues to be addressed we may need to conclude that the way systems and processes work today may not be sufficient to deal with it. The way borders are currently managed may no longer be sustainable in the future: traveller flows will simply be too large and the time checking takes will be too long. There are many inefficiencies in the current systems, for example continuous repeat actions of checking. In the future in his view there is likely to be a move to a more person-centric and risk-based approach, and not so much document-centric as is currently the case. Information will be gathered during every interaction with the authorities and they will be less dependent on checks which happen specifically at the border crossing.

For the next part of his presentation, Mr ter Hennepe brought everyone back to earth with the reality of the EES system. He said that this system, the details of which had been presented earlier, is focused very much on the security aspect rather than the facilitation of the traveller or the perspective of the airport. The proposed EES will put a strain on time and space availability in airports. However, there are areas where biometrics can help from a practical perspective. These are with the queues, processing time, traveller interaction and self-service systems. The speaker went on to outline what are thus the key operational cornerstones of the EES system: biometric performance, using quality algorithms, the need to understand traveller behaviour and traveller flows (including where to go and what to do), taking exceptions into account because there is another plane of 300 people coming right up behind. Tailor-made solutions integrated with national Border Control systems are key as there is no general solution to all issues/locations. EES is also about identity management, requiring the need for constant GDPR compliance.

The presenter then showed a diagram of the regulated 4 lanes at the external border that will manage the upcoming EES system,

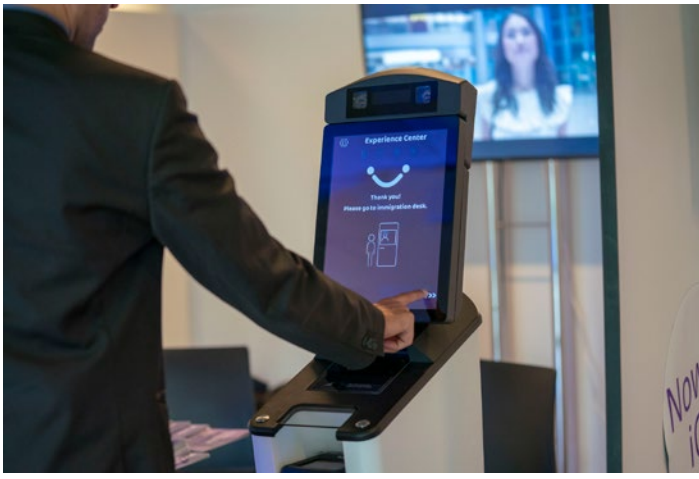
consisting of Standard ''EU'', third country nationals, all passports and finally Member State Registered Traveller Programmes, and then discussed the different biometrics required or suggested for each. An EES National Facilitation Program (RTP) can use whatever biometrics they wish. It could well be multimodal, making it inherently more secure. On this point the speaker reminded the audience that for example irises cannot be morphed.

Returning to the issue of biometrics, Mr ter Hennepe stated that we need to be aware that biometrics are not a commodity. The quality of biometrics – enabling fast and secure capture and matching - is critical. Biometric technology has evolved greatly compared to what is even now in operation, most commonly in airports.

The speaker then presented a number of current product solutions with their associated benefits for some of the issues raised. Firstly, MFace for face recognition, an example of how biometrics have evolved. MFace does not require additional light or height adjustment, supporting fast processing and a traveller-centric approach. This product is in use in Singapore Changi Airport, successfully processing millions of passengers. Then "OneLook" for dual biometrics in one device, capturing and matching face and iris in one ''move''. Lastly MWave for contactless fingerprint capture.

He then talked about a biometric solution for cruise ships as part of a public-private partnership between US CBP, RCCL and Idemia, operational in three USA seaports. Using Idemia MFace, live facial images are verified against the government picture database in CBP's TVS for entry and exit from the cruise ships.

To conclude his presentation, Mr ter Hennepe stated that the key was user-centric biometrics to support the border processes where the biometric tools need to have superior



algorithms to work faster. They need to be ergonomic and intuitive in their use. In this way you can achieve secure borders which can be passed quickly by bona fide travellers.

The Moderator thanked the panellists and then asked the audience for any questions.

Question 1: Please identify the key barriers to Industry in the area of development of biometrics

One panellist mentioned access to data sets that comply with GDPR as an important barrier that needs to be overcome; another said collaboration between all the various stakeholders was a critical issue as this gives the information needed to see what is the best solution; and finally, a panellist said that a key point was the sharing of data between stakeholders.

Question 2: An audience member asked the Idemia representative to elaborate a bit further on the Registered Traveller program and Known Digital Traveller Initiative which he mentioned in his presentation.

Mr ter Hennepe replied that as regards the Digital Traveller, this was a pilot program that involved cooperation between certain

governments, airlines and Industry suppliers to look at the future of how travel might look. The Registered Traveller Program can be a very important tool to facilitate regular travellers, who if they register with biometrics may be able to proceed through the gate without a document check.

Question 3: Concerned about the differences in approach to height adjustment for facial image reading between Secunet and Idemia solutions.

Mr ter Hennepe stated that their solution has in-built features which mean that height adjustment is not needed, e.g. they have two cameras. The Secunet representative focused his reply on what are the ISO requirements for the facial image and in particular what is required if post processing is needed.

Question 4: The Moderator asked the Secunet representative what to do about his point that their research showed the processing time for third-country nationals may increase by three times from the current baseline.

He answered that the use of self-service systems needs to be promoted, with user guidance to shift focus to the automated systems. If travellers are using the self-service solution, then although each of them individually may spend longer at the kiosk than they would at a manual border post, because of the scaling possibilities a greater number of passengers will be processed faster than used to be the case.

Finally, there was a comment from the floor from a representative of the IOM (Institute of Migration) that it would be good if both private corporations and government agencies gave a gender breakdown when talking about the number of their employees.

The Moderator thanked her for her comment and thanked the panel for their presentations and input.

CLOSING PANEL DISCUSSION

The way ahead for Borders and Biometrics

Moderator
Ted Dunstone · CEO of Biometix and Head of the Biometrics Institute "Biometric Security and Integrity Expert Group" (BSIEG)
Panellists
Marc Sulon · Head of Unit, Information Systems for Borders, Migration and Security, Directorate-General for Migration and Home Affairs, European Commission
Guido Brockmann · Head of Sector, Product Management, eu-LISA
Arun Vemury · Director · Biometric and Identity Technology Center · U.S. Department of Homeland Security
Nayra Perez · Head of Office · Data Protection · Frontex
Rasa Karbauskaite · Head of Sector *(acting)*, Standards and Capacity Development, Research and Innovation Unit, Frontex

The Moderator introduced the topic of the panel discussion and the panel. To start things off he said he would give his view as to what the key themes are for biometrics at the border over the next ten years. Then each panellist would be able to make their presentations.

Mr Dunstone started by saying that this is an exciting and active time for biometrics at the borders, with a lot of initiatives at various stages of implementation. There are obviously the initiatives which are coming from the EU, e.g. the EES, but a key issue will be how these kind of actions interact with other global initiatives such as the IATA - 1 ID project, the World Economic Forum Known Traveller and the Digital Identity project, as well as a variety of other industry initiatives. So how these initiatives will come together to scope the future is still an open question.

He then went on to talk about the morphing issue, highlighting that even though, as many people have stated at the conference, the issue is certainly challenging, he maintained that such vulnerabilities and challenges are nothing new and should be seen through the general lens of risk management at the border - in fact as part of a wider issue of border management and biometrics at the border. On this issue he also commended the training materials developed by Frontex, which are available both on-line and in person. He then moved on to one other topic that has been raised frequently at the conference, which is the potential to share data while respecting privacy constraints across countries and between countries and organisations in order to foster collaboration, with a key issue being standards. At the level of data there are standards but at the higher level of systems there is still much work to do on the issue of standardisation. To close his perspective, he said he believed it was very important to focus on end user generated research, so that major challenges such as queues at the border can be avoided and that not only security is enhanced but that the end user experience is also a positive one.

He then introduced the first speaker.

Marc Sulon

Mr Sulon started by saying that he would start with what there is today in terms of standards and legislation before moving to the future. He stated that even today there are technologies, regulations and standards to be implemented but that what is missing is that they are implemented all together. As

an example, he said that most of the focus at the conference so far has been on borders at airports, however he maintained that these locations are the easiest because everything is controlled. He went on to state that member states who have land borders tend to look at the issues differently, as capturing biometrics from people in cars or buses was obviously a different issue. Here he said that it was necessary to appreciate that these countries' perspectives were that the need to capture biometrics to increase security needed to be set against what implementations were actually workable at a land border without greatly increasing waiting times for entry.

Workable means that it should be possible to enrol people and facilitate them at land borders as well as in airports, where you can have self-service kiosks and e-gates to deal with the traveller flows. Mr Sulon also pointed out that at the land border if large queues are created e.g. to enrol third country nationals the biggest number of people who will be standing in the queues will actually be EU citizens, as the reality at the land border is that there is one queue to be facilitated. The speaker stated that as two-thirds of the people crossing the external borders of the EU are actually EU citizens, and that their facilitation at land borders and at seaports and train stations should be addressed first and then separately the issue of third-country nationals should be dealt with. In the case of the enrolment of biometrics at land borders of third-country nationals, he stated that it is desirable that this be done while the travellers are still sitting in their cars, without having to get out.

He then went on to discuss the issue of how to facilitate previously registered travellers, whether EU citizens or third-country nationals, who have been pre-vetted so that their identities are known, their car is known etc. In this case he said it should be possible to facilitate these people so that they do not need to get out of their cars and also that their car does not even need to stop; the process should be at a similar speed to the tolling areas for cars on French motorways. He stated that this vision was already possible with existing technology and that such a solution is probably more challenging that what will happen at airports but that it will solve a bigger problem as the longest queues are currently at land borders.

Mr Sulon raised the issue of the anticipation of border control. He stated that self-capture and enrolment of biometric data via an application on the traveller's own mobile device is possible under the current regulations but that the issue is how to guarantee that this data is correct vs what is in the passport and that it hasn't been modified in some way and can be sent and processed without any possibility of changing it. So this will be a border crossing, but with a pre-enrolment of biometric data which will allow for a seamless border crossing without stopping if it's not necessary.

In general, facilitation should be enhanced and sped up for people who are known already, in terms of identity and an automatic verification travel document automatically (Schengen Master List implementation); in the first instance EU citizens but later also third-country nationals with visas. The visa system also needs to be facilitated and integrated into these proposed changes and automated so that those people who are known are also facilitated faster.

Mr Sulon then went on to say that it was important that biometrics be implemented consistently and coherently. There is currently a discussion regarding the way facial images will be enrolled. It has turned out that a facial image is in fact a composite of several images, and that the question is how to ensure that a biometric facial image captured in Member State a) will be able to be verified by a Member State b) which may be using different software etc. In fact, there are standards lacking here and it is surprising that this issue has come out so late after the invention of facial image identification. Mr Sulon said the compatibility of the systems at the actual borders with the central EES system was a big challenge today, not in 10 years. He stated that a solution to this was necessary from industry and that it is essential to have technological solutions to comprehensively implement what is already planned and is in fact currently in the implementation phase, e.g. EES, and not issues related only to the future. His key message was that in fact the future is now.

The moderator then introduced the next speaker.

Rasa Karbauskaite

Ms Karbauskaite started by saying that she was speaking as a representative of the border management community. One of the issues that has been raised is business processes and process optimisation, and this is indeed a major challenge when it comes to the implementation of different kinds of technology at the different kinds of border. In particular there are large differences between border crossing points in terms of the constraints for dealing with traveller flows, with differing space availability etc. a major issue. She reminded the audience that there are 451 land border crossing points to the EU Schengen zone, so it is important that the processes are optimised and that the stakeholders are supported. She went on to state that best practice process optimisation needs to be shared both internally in the EU but also potentially externally with partners as well, and that

research and testing was also of critical importance, and Frontex has considerable expertise in the area of the operational impact of technological solutions.

Ms Karbauskaite continued by referring to the Biometrics on the Move pilot at Lisbon airport and that this will enable a lot of learning in terms of ergonomics and processes as well as technology, and that pilots or tests are always best conducted in real environments, not in simulations. In terms of performance, trust in the biometrics is very important for the border management community, and here the issue is: how can this trust be built when we don't yet know how the technology will actually perform. She asked: what are the tools to test the technology? Not only on single cases, but on multiple usage. Conversations with end users have shown that they often don't have the capacity to test these solutions, so they are relying in fact on the vendors. If these tests are not done regularly and correctly then vulnerabilities will continue to arise. In her view industry should be more proactive in providing solutions to potential vulnerabilities.

In terms of future trends there is the issue of seamless travel, which has pros and cons, and also paperless travel, with a key issue being to what extent the digital identity can be trusted. Also, it needs to be considered what was said by the ICAO representative on the first day of the conference that it seems more likely that there will be a hybrid approach in this area, not completely paperless after all.

Ms Karbauskaite then went on to consider the issue of changing the "border mindset'" concerning the value that biometric technology can bring. It needs to be seen by political leaders as an opportunity, not only a risk. The knowledge gap in terms of technology needs to be filled by training, not only of border guards but also of leaders. Defining these issues only in the context of vulnerabilities, be it presentation attack or morphing attack detection, is too narrow a definition of the training problem; it is more an issue of the overall role of biometrics, what it can provide and how the process will change.

To conclude she raised the issue of standardisation, a theme which has arisen several times at the conference. She stated that standardisation is not only about technology, but that in the context of border control management it is crucial that standardisation is end user driven.

Guido Brockmann

Mr Brockmann started by reminding the audience of what eu-LISA is responsible for, in particular the large-scale IT systems that stand behind current EU border management: the Schengen information system and the VIS. They are also operationally responsible for the development of the EES in terms of implementation and then interoperability with national systems.

He noted that the challenges are huge, with 1900 external border crossing points which will all need to be equipped with compatible material. To illustrate the scale of the task he mentioned that the US-Canada border has 100 border crossing points, and the

US-Mexico border, 50. In addition there is the issue of the variety of border crossing points in the EU, ranging from large land border crossing points between Poland and Ukraine, small ones between Croatia and Bosnia and Herzegovina and then somewhere like the Frankfurt airport. Additionally, the data needs to be collected in this variety of points and then also to be available at every border crossing point – with similar processing times. In addition to all this there is the issue of sea borders in general and of course more specifically the challenges which will be faced at French seaports with ferry connections to the UK after Brexit, when UK citizens will become third-country nationals. Of course, the main challenge he stated is with the first-time traveller (after introduction of EES) who will have to give their biometric data at the first EU external border that they encounter, in a process which according to the regulation should take no longer than 23 seconds for bona fide travellers. Afterwards, for repeat entries and exits, the process will become much simpler as his/her biometrics will then be in the system; the process should be seamless and take no longer than 2 seconds.

Mr Brockmann said the issue is also, as has been raised by others, how to test a system of this magnitude. This requires eu-LISA to reach out not only to Member States but also to academia for input as to how to test not only the performance but also the accuracy of this new proposed system. It is very important to ensure that the algorithms are based on an ethnically diverse range of facial images, as only then can they deal properly with the reality of the border crossing point. Another issue he raised is the lack of international standards for biometrics, which also need to be developed, preferably globally. Access to databases is also an issue for testing, as the system must be scalable and accurate from Day 1.

To conclude, Mr Brockmann said that the overall objective was how to retain an open but

secure EU, with the user experience of the bona fide traveller preferably enhanced from the current benchmark, but that simultaneously the potential threats can be found and forwarded to a potential second line inspection.

Arun Vemury

This speaker started by stating that there has been a huge improvement in the robustness of biometric technologies and reminded the audience that these technologies are being increasingly used because they work very well. He also believes that quite a lot of what has been presented at the conference so far does not accurately reflect how good biometric technologies actually are. He went on to say that he does not see facilitation as being the enemy of enhanced security and that he believes that both can be achieved simultaneously.

Mr Vemury went on to state that good cooperation with industry depends on giving them a clear set of requirements, as an example he mentioned a recent test conducted by the U.S. Department of Homeland Services in 2019 called the Biometric Technology Rally test, organised for different vendors who were given certain constraints and then had to come up with the best solution using biometric technology, including matching, to meet the set objectives. However, within the framework set they could innovate their approach. The test also involved signage and information for the sample traveller.

Concerning face recognition systems, he went on to say that it is important to remember that the sensor matters, and the quality of the camera is also very important – this has an impact on the effectiveness of the capture across all demographics. It is important to realise, he said, that biometrics are not a silver bullet, they are simply a means to reduce risk. Mr Vemury underlined that is very important to understand the algorithms. Face Recognition Vendor Tests (FRVT) conducted by NIST show that there are significant differences in the quality performance of different algorithms, and that not all matching algorithms are top tier. On the issue of interoperability, he emphasised how important this is and that you cannot assume that data from one system will work just as well in a different system.

Human factor errors are much more of an issue at the border than system ones, Mr Vemury maintained, however at the end of the day it doesn't actually matter if an error is system or human, the result is the same in terms of longer processing time and consequently a longer queue. On this point he gave the example of travellers not knowing where to look to have their photograph taken, as the camera in front of them "doesn't look like a camera". Solving this issue takes time but is it a system failure, a human failure or a failure in proper explanatory signage? He reminded the audience of the versatility of the border guard who can do many things versus the system which can only do a few, highly specialised things. The technology should be a complement to what the border guards can do, in effect if the system can take away some of the heavy load of ID document checking then the border guard is freed up to do more value-added activities such as risk assessment. The technology is therefore a force multiplier to deal with scaling populations.

In terms of testing and standards, more work needs to be done. Mr Vemury also emphasised that it is far more challenging to organise real world tests than laboratory testing, but it is also important to set up the test requirements correctly and be very clear what you want measured, which may not be the same as what the technology's own log measures. As an example, he stated that cameras sometimes simply cannot read that they have a person standing in front of them, but the camera's own log does not record this as an

error, as it simply doesn't have any record of this person being there. However, from a system performance point of view such a situation is obviously an error and the frequency of this happening should be tested.

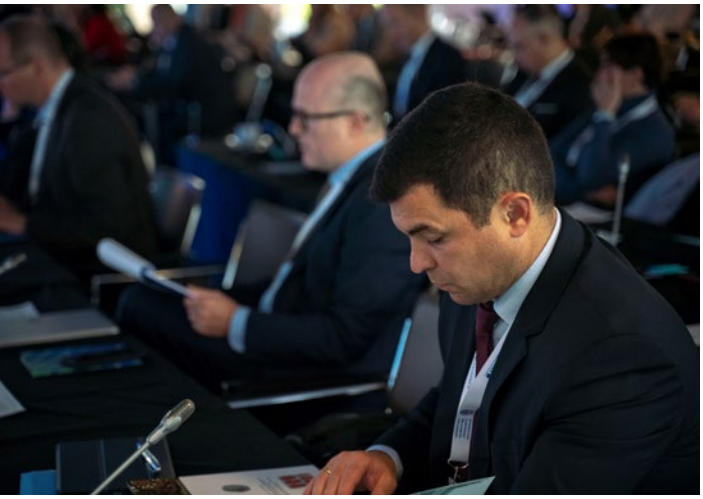
To conclude, the speaker went back to the learnings from the Biometric Technology Rally test they organised in 2019 . Mr Vemury stated that it is important not only to test the individual performance of the camera systems and the matching algorithms but also their interoperability - how they could potentially work together. This is a very important issue, as in fact of the 80 or so possible combinations that took part in the test only 12 combinations ended up meeting the requirements of under 5 seconds and over 95% accuracy for the matching rate.

Nayra Perez

At the start of her presentation Ms Perez said that for the future of biometrics in the context of data protection we should assume that the GDPR will still be in operation in 10 and 20 years, as this was certainly the intention when it was introduced – to be a long-term solution. She said that most of the talk about data protection and the GDPR at the conference had presented data protection in a negative light, but that at the same time most speakers had also said that it was important. She stated that she felt the previous speakers at the conference for the most part didn't seem to know how the legislation worked. Firstly, she mentioned the raised issue of data quality, but in this area in fact the legislation also says that data quality is very important and should be accurate, reliable and kept up to date, especially for minors and seniors.

Secondly, the issue of data sharing: this is allowed, under certain legal conditions. In terms of data transfer this is allowed to third countries if they meet certain standards in terms of data protection. If the country itself doesn't have these data protection regulations, then it can be done between the 2 parties under special data protection agreements. She mentioned that it was important to use lawyers or advisors who have expertise in this area. On testing, Ms Perez also said that there are specific articles in the legislation which deal with the issue of the usage of personal data and scientific research but that it sounded from a number of the speakers that they seem not to be aware of this. Biometric data is in a special category of data in the regulations as its misuse could have particularly bad consequences. To conclude on this point, she said that the legislation actually allows a lot of things to be done with data but that they must be processed following the correct procedures.

She went on to outline one of the key concepts behind the GDPR, namely "purpose". So, what is the intended reason or purpose why the data needs to be processed? The purpose of using biometrics for testing will be different from the purpose of enhancing security, i.e. the purpose of the capture at the border itself. Both of these are dealt with in different parts of the legislation. As a corollary to purpose is also the concept of proportionality. Is the data collection being proposed proportionate to the proposed usefulness of the biometric data capture? At the foundation of these principles is common sense, she emphasised. If it is planned to create large data sets with biometric data to help with testing then it will be necessary to consider what are the risks, and as part of that what is the worst-case scenario if someone hacks into this data for the individuals whose data it is. It is also necessary to recognise that risks can also come from within in terms of inaccurate data or incorrectly processed data by border guards or law enforcement. To conclude, Ms Perez stated that data protection in her view is not a barrier to security – it is



a part of security, and law enforcement officers need to be appropriately prepared and trained that they can achieve security while respecting data privacy regulations –more specifically, in defining the purpose properly and proportionately of what it is they want to do with a particular data set.

The Moderator thanked the panellists for the interesting thoughts that they had bought to the table. He then asked the floor if there were any questions, in particular on the issue of data protection with regard to biometrics. The Moderator himself asked if from a data protection perspective there was any difference between a facial image, which might be seen as being more generally "visible", and for example an individual's fingerprints which have connections to forensics and police databases and which might therefore be perceived as being more sensitive. In response the data specialist from Frontex said that the law does not distinguish on this point between different biometrics, but that the key was getting good quality advice and in defining the process of how the data will be dealt with and what are the risks associated with this, including worst-case scenarios.

A conference delegate then laid out their view on the problems which law enforcement faces with regard to dealing with crime and the current privacy regulations regarding biometrics, saying that they potentially hinder the possibility to capture and use video camera evidence. He asked that it be borne in mind that cameras used in this way proportionately are essential and useful tools in fighting crime. In response the data protection specialist said that as a comment she agreed with this point – it is always a question of balance. Each issue should be addressed from both sides and perspectives, which in her view are not incompatible.

Mr Sulon replied to a question about the current readiness of the EU and Member States in terms of EES implementation, by saying that the Member States are working very hard on many aspects of this issue: implementing EES, reforming the visa application system and implementing interoperability, and changes to the SIS. From a technical perspective he believes the implementation is proceeding well with no big issues, however the IT systems are only part of this as in fact the EES will impact the border processes – which are in fact a combination of human and automated solutions that need to work together. So, the processes at the border crossing points need to be adapted to take advantage of what the technology will be able to deliver. Mr Brockmann added that eu-LISA is reaching out to the Member states and has biometric working groups with each one in order to ensure that the new system works from Day 1. In answer to another question submitted on-line he said that there were no plans to move the systems to the cloud and that they will keep it all on site. In reference to a question regarding whether there were centrally set benchmarks for the operation of ABC gates, Rasa Karbauskaite replied that as of yet this has not happened but that she believes that it needs to. Additionally, she stated that the EU

probably needs to consider setting up a testing centre, something equivalent to the United States' Homeland Security technology testing centre, to properly test new potential technological solutions.

The Moderator then thanked the panel and closed the session.

Closing remarks

Javier Quesada · Head of Unit, Research and Innovation, Frontex

Javier Quesada thanked in particular the last panel for their summary of the opportunities and challenges ahead for biometrics and borders. Mr Quesada stated that when organising the conference, they considered the issue that biometrics is applicable in many different places and industries than just border control and also of course that it is a global technological phenomenon, not just related to the EU. This was of course the reason why the conference is international. They are aware that the topic is much bigger one that just the focus of the conference, but they are not sorry that the focus of the conference was on biometrics at the border, as the issues here and now in this area are truly pressing.

Mr Quesada went on to say that he is happy that the business case for biometrics at the border seems to now be universally accepted, with a lot of talk of both facilitation and security at the conference as well as the underlying necessity of concern for data protection and privacy.

Mr Quesada then went on to follow up on remarks made by Aija Kalnaja on Day 1 of the conference about the concern that the individual border guard feels properly supported by the new technology, through training and also thanks to new legal regulations. He believes there is a lot of work to do in this area and that there is probably a gap between the rhetoric and the reality from the border guard's perspective. He also noted that there were not that many actual operational border guards present at the conference. He strongly emphasised that although the conference did not feature the viewpoint of the individual border guard as much as perhaps had been hoped, Frontex was absolutely committed to representing the interests of the border guard community.

Mr Quesada said that the discussion concerning opportunities and challenges had a much stronger focus on the latter than the former. Whatever the reasons for this, e.g. the complexity of the current situations, or too much implementation work, it is nevertheless a concern. The conference saw a lot of reference to the fact that many algorithms are not very effective and that there are problems with data sets, and that the traveller flow experience may indeed worsen in the near future. In addition to this the acquisition of technology by Member States from industry vendors is perceived not as an opportunity but as an interoperability challenge at the EU level, with a lack of shared trust in technological solutions. However, he stated that everyone does seem to agree that the border guard should be given the power to make the best decisions that a human being can make with the technology implemented to empower him or her in this task, as a significant proportion of the mundane checking tasks will then be automated. It is still not clear, however, how this vision will be implemented in many cases.

Mr Quesada then informed the delegates of some early results of the previously mentioned pilot test for biometrics on the move which has been running at Lisbon Airport. Since the start of the test on the September 8th there have been 92 biometric enrolments using on the move technology for face and fingerprints and of this 72 were able to proceed seamlessly without problems. However, this means that 20 had problems which required the intervention of an officer. So in all it can be said that

these top line results clearly illustrate the nature of the challenges ahead.

On the subject of morphing he stated that the conference had fulfilled its objective of simply sharing more widely an appreciation of the issue and also summarising where things stand in this area at the moment. He felt that this had been achieved very well.

He then went on to say that he had in particular remembered a question from a delegate who questioned "whether it has to be done this way at all", or whether the processes can be completely re-thought and done in a different way. He is aware that this won't be a popular way of thinking for the Commission and legislators but that it is a serious point that should perhaps be addressed at future conferences.

To conclude, Mr Quesada said that he believed that the conference clearly achieved one of its primary objectives, which was to bring together biometric experts from different perspectives with various representatives of the end user community. This networking opportunity had been successfully created. The 20 or so industry exhibition stands were also in his view a clear success and had been of interest to delegates. He then said that of course everyone will have their own opinion about the conference, so he encouraged the delegates who hadn't already done so to go onto the application and fill in the conference evaluation form. He said that Frontex wanted to organise another conference next year, so their advice via the evaluation form will be very important in starting to prepare for that.

In closing he thanked the Frontex and external staff responsible for the details of the conference organisation and wished all the panellists and delegates a safe journey home.

Programme

DAY 1 – 9 October 2019

08.00 – 09.00 On-site registration

09.00 – 10.15 Welcome address

Javier Quesada · Head of Unit, Research and Innovation, Frontex

Keynote Speeches

Views on biometrics and its application in border control

Keynotes

Fabrice Leggeri · Executive Director, Frontex

Olivier Onidi · Deputy Director-General, Directorate-General for Migration and Home Affairs, European Commission

10.15 – 10.45 Coffee Break and exhibition viewing

10.45 – 12.15 Panel Discussion

Biometrics for border control and the role of Frontex

In this panel discussion the panellists will discuss the role of Frontex in providing support and expertise to its end user community on the topic of biometrics. Focusing on biometrics, its application in border control, and the challenges it may pose to border security, it will explore the way in which Frontex may interact with the biometric community to mainstream added value for end-users.

Moderator

Javier Quesada · Head of Unit, Research and Innovation, Frontex

Panellists

Aija Kalnaja · Director, Capacity Building Division, Frontex

Narjess Abdennebi · Chief Facilitation Section, International Civil Aviation Organisation

Jean Salomon · CEO, European Association for Biometrics

Hans de Moel · Director, Biometrics Institute

Laurent Beslay · Scientific Project Leader Law Enforcement Technologies and Citizen, Cyber and Digital Citizens' Security, Joint Research Centre, European Commission

12.15 – 13.45 Lunch and exhibition viewing

13.45 - 15.00 Thematic Session 1

The challenge of Morphing for border control

Morphing attacks in the context of border control is a relatively new and undocumented phenomenon. In this thematic session the panellist will help define what morphing is, discuss morphing and its implications for

border management, and propose possible actions aimed at mitigating the threat of morphing attacks.

Moderator

Dr Joseph Atick · Executive Chairman, ID4Africa

Panellists

Matteo Ferrara, Ph.D. · Department of Computer Science and Engineering, University of Bologna

Mei Lee Ngan · Computer Scientist, National Institute of Standards and Technology

Ronald Belser · Research and Development Advisor, National Office for Identity Data, Dutch Ministry of the Interior and Kingdom Relations

Prof. Dr Christoph Busch · Biometrics Laboratory, Norwegian University of Science and Technology and Hochschule Darmstadt, Germany

15.00 – 15.30 Coffee Break and exhibition viewing

15.30 – 16.45 Thematic Session 2

National approaches to prevent and detect morphing

National authorities from around the world are actively engaged in research activities and the development of policy tools aimed at addressing and overcoming the threat of morphing attacks. In this thematic session, representatives from different national authorities will present and discuss novel approaches to morphing.

Moderator

Dinusha Frings · Research Manager, National Office for Identity Data, Dutch Ministry of the Interior and Kingdom Relations

Panellists

Dr Gert Jan de Nijs · Senior Project Manager, Dutch Vehicle Authority

Kari Kanto · Senior Advisor, National Police Board of Finland

Dr Uwe Seidel · Senior Scientific Director, Head of Section KT 5 – IT Forensics and Documents, German Federal Criminal Police Office

Dr Rebecca Heyer · Department of Defence Science and Technology, Australian Government

15.45 – 17.00 Wrap-up Day 1

19.30 – 24.00 Conference Dinner

DAY 2 – 10 October 2019

09.00 – 09.15 Opening

09.15 – 10.30 Thematic Session 3

Ongoing research in the area of morphing and morphing attack detection methods

In this thematic session researchers from renowned academic institutions will share and discuss their latest research and developments in the area of morphing as well as its potential impact for border control.

Moderator

Mei Lee Ngan · Computer Scientist, National Institute of Standards and Technology

Panellists

Dr Sébastien Marcel · Senior Researcher, Head of Biometrics Security and Privacy group, Idiap Research Institute

Dr Andrey Makrushin · Postdoctoral Researcher, Advanced Multimedia Security Lab (AMSL), Otto von Guericke University of Magdeburg

Dr David J. Robertson · Lecturer in Psychology, School of Psychological Sciences and Health, University of Strathclyde

Ulrich Scherhag, Ph.D. Student · da/sec Biometrics and Internet-Security Research Group, Centre for Research in Security and Privacy (CRISP), and Darmstadt University of Applied Sciences

10.30 – 11.00 Coffee Break and exhibition viewing

11.00 – 12.30 Thematic Session 4

The application of biometric technologies at our borders: An industry perspective

In this final thematic session border management authorities and other participants will be given first-hand insight into how the industry prepares for the increasing use of biometrics at borders, and how industry is developing cutting-edge biometric technologies aimed at both facilitating travel and securing borders, while complying with new regulations and policies.

Moderator

Darek Saunders · Head of Sector *(acting)*, Border Security Research Observatory, Research and Innovation Unit, Frontex

Panellists

Lukasz Kubik · Secunet Security Networks AG

Jürgen Mathwich · T3K and Austrian Institute of Technology (AIT)

Brage Strand · Mobai

Jan-Willem ter Hennepe · Idemia

12.30 – 14.00 Lunch and exhibition viewing

14.00 – 15.30 Panel Discussion

The way ahead for Borders and Biometrics

Following the thematic sessions dedicated to one of many challenges the use of biometrics introduces – morphing - this closing panel discussion aims return the focus to the broader theme of biometrics in the context of border control. Joined by key representatives from the broader biometrics and border management community, this panel will explore some of the main themes, challenges and opportunities presented by the widespread adoption of biometrics for identity verification at the border, such as the planned implementation of future information systems, the importance of standards set at international level, rights and privacy implications, and the role and ambitions of Frontex and the border management community with respect to biometrics.

Moderator

Ted Dunstone · CEO of Biometix and Head of the Biometrics Institute "Biometric Security and Integrity Expert Group" (BSIEG)

Panellists

Marc Sulon · Head of Unit, Information Systems for Borders, Migration and Security, Directorate-General for Migration and Home Affairs, European Commission

Guido Brockmann · Head of Sector, Product Management, eu-LISA

Arun Vemury · Director · Biometric and Identity Technology Center · U.S. Department of Homeland Security

Nayra Perez · Head of Office · Data Protection · Frontex

Rasa Karbauskaite · Head of Sector *(acting)*, Standards and Capacity Development, Research and Innovation Unit, Frontex

15.30 – 16.00 Closing Remarks

Research abstracts

Vulnerability of Face Recognition to Deep Morphing

Pavel Korshunov and Sebastien Marcel´
Idiap Research Institute, Martigny, Switzerland
{pavel.korshunov,sebastien.marcel}@idiap.ch

Abstract

It is increasingly easy to automatically swap faces in images and video or morph two faces into one using generative adversarial networks (GANs). The high quality of the resulted deep-morph raises the question of how vulnerable the current face recognition systems are to such fake images and videos. It also calls for automated ways to detect these GAN-generated faces. In this paper, we present the publicly available dataset of the Deepfake videos with faces morphed with a GANbased algorithm. To generate these videos, we used open source software based on GANs, and we emphasize that training and blending parameters can significantly impact the quality of the resulted videos. We show that the state of the art face recognition systems based on VGG and Facenet neural networks are vulnerable to the deep morph videos, with 85.62% and 95.00% false acceptance rates, respectively, which means methods for detecting these videos are necessary. We consider several baseline approaches for detecting deep morphs and find that the method based on visual quality metrics (often used in presentation attack detection domain) leads to the best performance with 8.97% equal error rate. Our experiments demonstrate that GAN-generated deep morph videos are challenging for both face recognition systems and existing detection methods, and the further development of deep morphing technologies will make it even more so.

1. Introduction

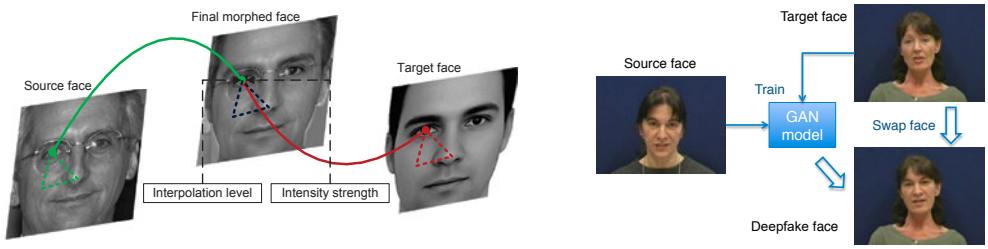
Recent advances in automated video and audio editing tools, generative adversarial networks (GANs), and social media allow the creation and the fast dissemination of high quality tampered video content. Such content already led to appearance of deliberate misinformation, coined 'fake news', which is impacting political landscapes of several countries [2]. A recent surge of videos (started as obscene) called Deepfakes[1], in which a neural network is used to train a model to replace faces with a likeness of someone else, are of a great public concern[2]. Accessible open source software and apps for such face swapping lead to large amounts of synthetically generated Deepfake videos appearing in social media and news, posing a significant technical challenge for detection and filtering of such content.

Although the original purpose of GAN-based Deepfake is to swap faces of two people in an image or a video, the resulted synthetic face is essentially a morph, i.e., a *deep morph*, of two original faces. The main difference from more traditional morphing techniques is that deep-morph can seamlessly mimic facial expression of the target person and, therefore, can also be successfully used to generate convincing fake videos of people talking and moving about. However, to understand how threatening such videos can be in the context of biometric security,

1 Open source: https://github.com/deepfakes/faceswap
2 BBC (Feb 3, 2018): http://www.bbc.com/news/technology-42912529 [3]https://www.snapchat.com/

we need to find out whether these deep-morphed videos pose a challenge to face recognition systems and whether they can be easily detected.

Traditional face morphing (Figure 1a illustrates the morphing process) has been shown to be challenging for face recognition systems [3, 16] and several detection methods has been proposed since [10, 18, 9]. For the GAN-based deepmorphing, until recently, most of the research was focusing on advancing the GAN-based face swapping [6, 8, 12, 14]. However, responding to the public demand to detect these synthetic faces, researchers started to work on databases and detection methods, including image and video data [15] generated with a previous generation of face swapping approach Face2Face [19] or videos collected using Snapchat[3] application [1]. Several methods for detection of Deepfakes have also been proposed [7, 21, 5].



(a) Morphing faces (b) Generating Deepfake faces

Figure 1: Comparing morphing and GAN-based face swapping techniques.

In this paper, we focus on evaluating the vulnerability of face recognition systems to Deepfake videos where real faces are replaced by GAN-generated images trained on the faces of two people. The resulted synthetic face is essentially a deep morph of two people. The database was created using the open source software with cyclic GAN model[4] (see Figure 1b for illustration), which is developed from the original autoencoder-based Deepfake algorithm[1]. We manually selected 16 similar looking pairs of people from publicly available VidTIMIT database[5]. For each of 32 subjects, we trained two different models (see Figure 2 for examples), referred to in the paper as the low quality (LQ) model, with 64×64 input/output size, and the high quality (HQ) model, with 128×128 size. Since there are 10 videos per person in VidTIMIT database, we generated 320 videos corresponding to each version, resulting in total 620 videos with faces swapped. For the audio, we kept the original audio track of each video, i.e., no manipulation was done to the audio channel.

We assess the vulnerability of face recognition to deep morph videos using two state of the art systems: based on VGG [13] and Facenet[6] [17] neural networks. For detection of the deep morphs, we applied several baseline methods from presentation attack detection domain, by treating deep morph videos as digital presentation attacks [1], including simple principal component analysis (PCA) and linear discriminant analysis (LDA) approaches, and the approach based on image quality metrics (IQM) and support vector machine (SVM) [4, 20].

3 https://www.snapchat.com/
4 https://github.com/shaoanlu/faceswap-GAN
5 http://conradsanderson.id.au/vidtimit/
6 https://github.com/davidsandberg/facenet

To allow researchers to verify, reproduce, and extend our work, we provide the database coined DeepfakeTIMIT of Deepfake videos[7], face recognition and deep morph detection systems with corresponding scores as an open source Python package [8].

(g) Original 1 (h) Original 2 (i) LQ swap 1 → 2 (j) HQ swap 1 → 2 (k) LQ swap 2 → 1 (l) HQ swap 2 → 1

Figure 2: Screenshot of the original videos from VidTIMIT database and low (LQ) and high quality (HQ) deep morphs.

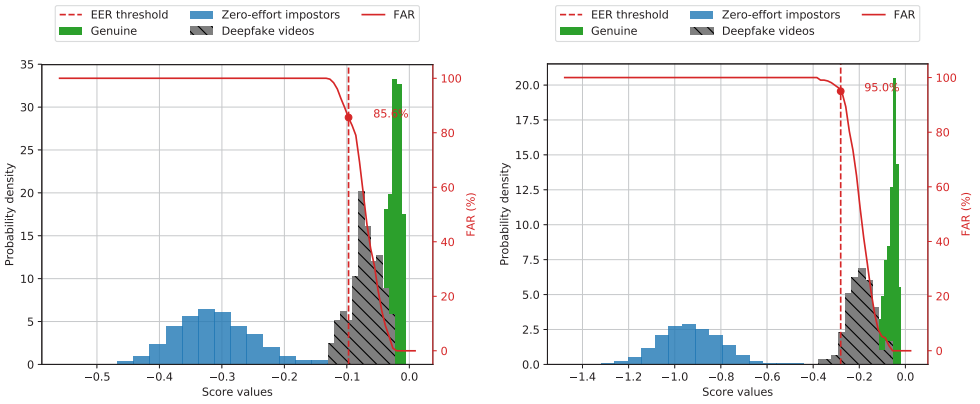
2. Database of deep morph videos

As the original data, we took video from VidTIMIT database[5]. The database contains 10 videos for each of 43 subjects, which were shot in controlled environment with people facing camera and reciting predetermined short phrases. From these 43 subject, we manually selected 16 pairs in such a way that subjects in the same pair have similar prominent visual features, e.g., mustaches or hair styles. Using GAN-based algorithm based on the available code[4], for each pair of subjects, we generated videos where their faces are replaced by a GAN-generated deep morphs (see the example screenshots in Figure 2). For each pair of subjects, we have trained two different GAN models and generated two versions of the deep morphs:

1. The low quality (LQ) model has input and output image (facial regions only) of size 64×64. About 200 frames from the videos of each subject were used for training and the frames were extracted at 4 fps from the original videos. The training was done for 10°000 iterations and took about 4 hours per model on Tesla P40 GPU.
2. The high quality (HQ) model has input/output image size of 128×128. About 400 frames extracted at 8 fps from videos were used for training, which was done for 20°000 iterations (about 12 hours on Tesla P40 GPU).

Also, different blending techniques were used when generating deep morph videos using different models. With LQ model, for each frame from an input video, generator of the GAN model was applied on the face region to generate the fake counterpart. Then a facial mask was detected using a CNN-based face segmentation algorithm proposed in [12]. Using this mask, the generated fake face was blended with the face in the target video. For HQ model, the blending was done based on facial landmarks (detected with publicly available MTCNN model [22]) alignment between generated fake face and the original face in the target video. Finally, histogram normalization was applied to the blended result to adjust for the lighting conditions, which makes the result more realistic (see Figure 2).

7 https://www.idiap.ch/dataset/deepfaketimit
8 Source code: https://gitlab.idiap.ch/bob/bob.report.deepfakes



(a) VGG-based face recognition

(b) FaceNet-based face recognition

Figure 3: Histograms show the vulnerability of VGG and Facenet based face recognition to high quality deep morphs.

2.1. Evaluation protocol

When evaluating vulnerability of face recognition, for the *licit* scenario without the deep morph videos, we used the original VidTIMIT5 videos for the 32 subjects for which we have generated corresponding deep morph videos. In this scenario, we used 2 videos of the subject for enrollment and the other 8 videos as probes, for which we computed the verification scores.

From the scores, for each possible threshold ϑ, we computed commonly used metrics for evaluation of classification systems: false acceptance rate (FAR) and false reject rate (FRR). Threshold at which these FAR and FRR are equal leads to an equal error rate (EER), which is commonly used as a single value metric of the system performance.

To evaluate vulnerability of face recognition, in *tampered* scenario, we use deep morph videos (10 for each of 32 subjects) as probes and compute the corresponding scores using the enrollment model from the *licit* scenario. To understand if face recognition perceives deep morphs to be similar to the genuine original videos, we report the FAR metric computed using EER threshold ϑ from *licit* scenario. If FAR value for deep morph videos is significantly higher than the one computed in *licit* scenario, it means the face recognition system cannot distinguish synthetic videos from originals and is therefore vulnerable to deep morphs.

Database	Detection system	EER (%)	FRR@FAR10% (%)
LQ deep morph	Pixels+PCA+LDA	39.48	78.10
	IQM+PCA+LDA	20.52	66.67
	IQM+SVM	3.33	0.95
HQ deep morph	IQM+SVM	8.97	9.05

Table 1: Baseline detection systems for low (LQ) and high quality (HQ) deep morph videos. EER and FRR when FAR equal to 10% are computed on Test set.

When evaluating deep morph detection, we consider it as a binary classification problem and evaluate the ability of detection approaches to distinguish original videos from deep morph videos. All videos in the dataset, including genuine and fake parts, were split into training (*Train*) and evaluation (*Test*) subsets. To avoid bias during training and testing, we arranged that the same subject would not appear in both sets. We did not introduce a development set, which is typically used to tune hyper parameters such as threshold, because the dataset is not large enough. Therefore, for deep morph detection system, we report the EER and the FRR (using the threshold when *FAR* =10%) values on the *Test* set.

3. Vulnerability of face recognition

We used publicly available pre-trained VGG and Facenet architectures for face recognition. We used the *fc7* and *bottleneck* layers of these networks, respectively, as features and used cosine distance as a classifier. For a given test face, the confidence score of whether it belongs to a pre-enrolled model of a person is the cosine distance between the average feature vector, i.e., model, and the features vector of a test face. Both of these systems are state of the art recognition systems with VGG of 98.95% [13] and Facenet of 99.63% [17] accuracies on labeled faces in the wild (LFW) dataset.

We conducted the vulnerability analysis of VGG and Facenet-based face recognition systems on low quality (LQ) and high quality (HQ) face swaps in VidTIMIT[5] database. In a *licit* scenario when only original videos are present, both systems performed very well, with EER value of 0.03% for VGG and 0.00% for Facenet-based system. Using the EER threshold from *licit* scenario, we computed FAR value for the scenario when deep morph videos are used as probes. In this case, for VGG the FAR is 88.75% on LQ deep morphs and 85.62% on HQ deep morphs, and for Facenet the FAR is 94.38% and 95.00% on LQ and HQ deep morphs respectively. To illustrate this vulnerability, we plot the score histograms for high quality deep morph videos in Figure 3. The histograms show a considerable overlap between deep morph and genuine scores with clear separation from the zero-effort impostor scores (the probes from *licit* scenario).

From the results, it is clear that both VGG and Facenet based systems cannot effectively distinguish GAN-generated synthetic faces from the original ones. The fact that more advanced Facenet system is more vulnerable is also consistent with the findings about presentation attacks [11].

4. Detection of deep morph videos

We considered several baseline deep morph detection systems:

- *Pixels+PCA+LDA: use raw faces as features with PCA-LDA classifier, with 99% retained variance resulting in 446 dimensions of transform matrix.*
- *IQM+PCA+LDA: IQM features with PCA-LDA classifier with 95% retained variance resulting in 2 dimensions of transform matrix.*
- *IQM+SVM: IQM features with SVM classifier, each video has an averaged score from 20 frames.*

The systems based on image quality measures (IQM) are borrowed from the domain of presentation (including replay attacks) attack detection, where such systems have shown good performance [4, 20]. As IQM feature vector, we used 129 measures of image quality, which include such measures like signal to noise ratio, specularity, bluriness, etc., by combining the features from [4] and [20].

The results for all detection systems are presented in Table 1. The results demonstrate that the IQM+SVM system has a reasonably high accuracy of detecting deep morph videos, although videos generated with HQ model pose a more serious challenge. It means that a more advanced techniques for face swapping will be even more challenging to detect.

5. Conclusion

In this paper, we demonstrated that state of the art VGG and Facenet-based face recognition algorithms are vulnerable to the deep morphed videos from DeepfaTIMIT database and fail to distinguish such videos from the original ones with up to 95.00% equal error rate. We also evaluated several baseline detection algorithms and found that the techniques based on image quality measures with SVM classifier can detect HQ deep morph videos with 8.97% equal error rate.

However, the continued advancements in development of GAN-generated faces will result in more challenging videos, which will be harder to detect by the existing algorithms. Therefore, new databases and new more generic detection methods need to be developed in the future.

References

[1] A. Agarwal, R. Singh, M. Vatsa, and A. Noore. Swapped! digital face presentation attack detection via weighted local magnitude pattern. In *IEEE International Joint Conference on Biometrics (IJCB)*, pages 659–665, Oct 2017.

[2] H. Allcott and M. Gentzkow. Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2):211–236, 2017.

[3] M. Ferrara, A. Franco, and D. Maltoni. The magic passport. In *IEEE International Joint Conference on Biometrics (BTAS)*, pages 1–7, Sep. 2014.

[4] J. Galbally and S. Marcel. Face anti-spoofing based on general image quality assessment. In *International Conference on Pattern Recognition*, pages 1173–1178, Aug 2014.

[5] D. Guera and E. J. Delp. Deepfake video detection using recurrent neural networks. In¨ IEE *International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, pages 1–6, Nov 2018.

[6] P. Isola, J. Zhu, T. Zhou, and A. A. Efros. Image-to-image translation with conditional adversarial networks. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5967–5976, July 2017.

[7] P. Korshunov and S. Marcel. Vulnerability assessment and detection of Deepfake videos. In *International Conference on Biometrics (ICB 2019)*, Crete, Greece, June 2019.

[8] I. Korshunova, W. Shi, J. Dambre, and L. Theis. Fast face-swap using convolutional neural networks. In *IEEE International Conference on Computer Vision (ICCV)*, pages 3697–3705, Oct 2017.

[9] R. S. S. Kramer, M. O. Mireku, T. R. Flack, and K. L. Ritchie. Face morphing attacks: Investigating detection with humans and computers. *Cognitive Research: Principles and Implications*, 4(1):28, Jul 2019.

[10] A. Makrushin, T. Neubert, and J. Dittmann. Automatic generation and detection of visually faultless facial morphs. In *Proceedings of International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP)*, pages 39–50. INSTICC, SciTePress, 2017.

[11] A. Mohammadi, S. Bhattacharjee, and S. Marcel. Deeply vulnerable: a study of the robustness of face recognition to presentation attacks. *IET Biometrics*, 7(1):15–26, 2018.

[12] Y. Nirkin, I. Masi, A. T. Tuan, T. Hassner, and G. Medioni. On face segmentation, face swapping, and face perception. In *IEEE International Conference on Automatic Face Gesture Recognition (FG)*, pages 98–105, May 2018.
[13] O. M. Parkhi, A. Vedaldi, and A. Zisserman. Deep face recognition. In *BMVC*, 2015.
[14] H. X. Pham, Y. Wang, and V. Pavlovic. Generative adversarial talking head: Bringing portraits to life with a weakly supervised neural network. *arXiv.org*, 2018.
[15] A. Rossler, D. Cozzolino, L. Verdoliva, C. Riess, J. Thies, and M. Nießner. Faceforensics: A large-scale video dataset for forgery¨ detection in human faces. *arXiv.org*, 2018.
[16] U. Scherhag, C. Rathgeb, J. Merkle, R. Breithaupt, and C. Busch. Face recognition systems under morphing attacks: A survey. *IEEE Access*, 7:23012–23026, Feb. 2019.
[17] F. Schroff, D. Kalenichenko, and J. Philbin. FaceNet: A unified embedding for face recognition and clustering. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 815–823, June 2015.
[18] C. Seibold, W. Samek, A. Hilsmann, and P. Eisert. Detection of face morphing attacks by deep learning. In C. Kraetzer, Y.-Q. Shi, J. Dittmann, and H. J. Kim, editors, *Digital Forensics and Watermarking*, pages 107–120, Cham, 2017. Springer International Publishing.
[19] J. Thies, M. Zollhfer, M. Stamminger, C. Theobalt, and M. Niener. Face2Face: Real-time face capture and reenactment of RGB videos. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2387–2395, June 2016.
[20] D. Wen, H. Han, and A. K. Jain. Face spoof detection with image distortion analysis. *IEEE Transactions on Information Forensics and Security*, 10(4):746–761, April 2015.
[21] X. Yang, Y. Li, and S. Lyu. Exposing deep fakes using inconsistent head poses. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 8261–8265, May 2019.
[22] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao. Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters*, 23(10):1499–1503, Oct 2016.

Face Morphing Detection: Issues and Challenges

J. Merkle[1], C. Rathgeb[1,2], U. Scherhag[2], C. Busch[2], R. Breithaupt[3]
[1]*secunet Security Networks AG, Essen, Germany*
johannes.merkle@secunet.com, christian.rathgeb@secunet.com
[2]*Hochschule Darmstadt, Darmstadt, Germany*
ulrich.scherhag@h-da.de, christoph.busch@h-da.de
[3]*Federal Office for Information Security (BSI), Bonn, Germany,*
ralph.breithaupt@bsi.bund.de

Abstract: Recently, facial recognition systems have been found vulnerable to morphing attacks. In these attacks, the facial images of two (or more) individuals are combined (morphed) and the resulting morphed facial image is then presented during registration as a biometric reference. If the morphed image is accepted, it is likely that all individuals that contributed to the morphed facial image can be successfully authenticated against it. Morphing attacks thus pose a serious threat to facial recognition systems, in particular in border control scenarios, where the reference image is often provided in printed form by the applicant. This paper provides a rough overview of the current state-of-the-art methods for detecting morphed facial images, and discusses issues and challenges in the development and evaluation of morphing attack detection methods.

Keywords: face recognition; face morphing attacks; morphing attack detection; vulnerability analysis; issues and challenges

INTRODUCTION

Image morphing techniques can be used to combine information from two (or more) images into one image. Morphing techniques can also be used to create a morphed facial image from the biometric face images of two individuals, of which the biometric information is similar to that of both individuals. An example of a morphed facial image (hereinafter referred to as "morph") is shown in Figure 1.

Figure 1: Example of a morphed facial image. The morph was created with FantaMorph. On the left and right the contributing subjects are depicted and in the middle the resulting morph (image source: Hochschule Darmstadt, BSI).

In many countries, the facial image used for an electronic travel document is provided by the applicant either in analogue (i.e. print on paper) or digital form. Therefore, an *attacker* (e.g., a wanted criminal or a foreigner not eligible for entry to the Schengen area) could morph his face image with the face image of a similar looking *accomplice*, and the accomplice could apply for a passport or another electronic travel document with that image. It should be noted that morphed facial images look realistic and may be similar enough to both individuals to deceive human examiners [1][2]. This was showcased in Germany by members of the political activist group Peng! Kollektiv, who succeeded without any problem in applying for a passport with a morphed face image[1]. Both, the attacker and the accomplice can then be successfully verified against the morphed image so that the attacker can also use the electronic travel document issued to the accomplice to pass through an automatic border control (or even human inspections at border crossings). If more than two images are morphed, this usually reduces the attacker's chances of success if his characteristics are weaker in the resulting morph. The risk of the described *morphing attack* (MA) [3] is increased by the fact that realistic looking morphed facial images can be generated by unskilled persons. This can be done with the help of an easy-to-use morphing software for facial images, e.g., FantaMorph[2], which is either freely available or can be purchased at a reasonable price.

VULNERABILITY ANALYSIS

When analyzing the vulnerability of face recognition systems to MAs, it is obvious to augment the metrics for evaluation of presentation attacks, in which an attacker, for example, holds a photograph of another subject in front of the camera. The *Impostor Attack Presentation Match Rate* (IAPMR) [4] introduced in ISO/IEC 30107-3 represents a standardized metric for evaluating the impact of a presentation attack. The IAPMR is defined as follows: the proportion of impostor attack presentations species in which the target reference is matched in a full-system evaluation.

However, the disadvantage of the IAPMR metric for the evaluation of MAs is that it is calculated from individual attacks and therefore only reflects the probability of success of one of the subjects involved in the attack. In fact, however, two different scenarios can be relevant:

1. Only the attacker wants to be successfully authenticated by the face recognition system. In this scenario it is assumed that an accomplice was able to successfully apply for a passport, i.e. a human inspection of the morphed image was already overcome when the application was submitted. In such a scenario an asymmetric morphing of images, so that attacker and accomplice(s) contribute with different weights (a.k.a. alpha factors) to the morphed image, can be useful. An asymmetrical morphing can also be realized by procedures which morph the faces only in the inner area and the outer area (with forehead, hair, ears, neck) is taken only from one of the two initial images. It is usually assumed in the literature that the face of the accomplice contributes more to the morph than that of the attacker and that the outside area of the accomplice is used, because the risk of the picture being rejected during the application process is then lower.

1 Peng! Kollektiv, MaskID: https://pen.gg/de/campaign/maskid/
2 FantaMorph, Abrasoft: http://www.fantamorph.com/

However, since serious consequences (e.g., criminal prosecution[3]) are hardly to be expected in the case of a rejection in the application process, the reverse case, in which the accomplice is represented to a lesser extent in the morph, would also be conceivable. If only the attacker is to be successfully verified, the IAPMR can be used as a metric to evaluate the overall system's vulnerability. Care should be taken to ensure that the morphs used in the evaluation can at least overcome human inspections when presented by the accomplice, so that they are accepted when the application is made.

2. All individuals contributing to the morph want to be successfully authenticated against the morphed facial image. In such a scenario a symmetrical morphing of images is more realistic, i.e. attackers and accomplices contribute equally to the morphed image[4]. This scenario cannot be evaluated using the IAMPR and motivated the introduction of new evaluation metrics [5]. The comparison of a morphed facial image with a face image of a contributing subject is called a paired morph comparison. A MA is successful, if all involved subjects have been successfully verified. Hence, the minimum (for similarity values) or the maximum (for distance values) of all paired morph comparisons is of particular interest. Motivated by ISO/IEC 30107-3 [4], the *Mated-Morph-Presentation-Match-Rate* (MMPMR) is proposed in [5] to evaluate the effect of a MAs on the overall system.

MORPHING ATTACK DETECTION

In order to detect MAs, so-called *morphing attack detection* (MAD) techniques must be developed, which allow reliable differentiation between morphs and bona fide (i.e., genuine) facial images. If a potentially morphed facial image is detected in the course of an automatic border control, it can be inspected in a second step, e.g., by a border official, or the identity of the suspect can be checked using the fingerprints stored on the electronic passport. A particular challenge is the detection of analog morphs, i.e. after they have been printed and scanned, since many artefacts that indicate morphing can be lost due to the print-scan transformation. This is particularly relevant for passports from countries such as Germany, where an application with facial images in analogue form is still the rule.

DETECTION SCENARIOS

MAD procedures can be divided into two classes, see Figure 2, according to the scenario under consideration:

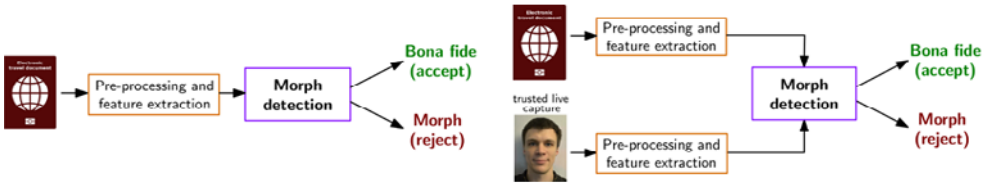


Figure 2: Morphing attack detection scenarios. Left: Single image Morphing Attack Detection, Right: Differential Morphing Attack Detection (image source: Hochschule Darmstadt).

3 It is also questionable whether an application for a passport with a morphed picture is a criminal offence, since even a morphed picture technically represents a photograph of the passport holder, which is clearly what is required, for instance by the German Passport Act.

4 However, the outside area can be taken over by only one subject to avoid possible morphing artefacts in that area.

- *Single image MAD*: These approaches examine a single face image, for example, when checking the authenticity of a passport without reference or directly when applying for a passport, and check whether it has been morphed. For this purpose, the image is examined for potential traces of a morphing process. This class of MAD procedures is also known as no-reference MAD or forensic MAD.

- *Differential MAD*: These procedures compare the potentially morphed reference image with a trusted probe image, e.g., a live image from an eGate for automatic border control. This class of MAD procedures is also referred to as image pair-based MAD.

Basically, the two approaches differ in that single image MAD approaches aim to detect certain artefacts induced by the morphing process (e.g., "ghost artefacts" in which structures of the original images overlap), while differential MAD methods analyze the features of the potentially morphed facial image and the live image of a face, e.g. by estimating difference vector between both feature vectors. It can be assumed that carefully created morphs contain only a few recognizable artefacts (if any), which after a print-scan process (i.e., when providing an analog facial image) are probably very difficult to detect. Single image MAD procedures can depend heavily on the training data used and can only detect the artefacts learned during training. This can greatly limit the generalizability of these methods. For these reasons, differential MAD procedures are generally to be seen as more promising.

In recent years, numerous approaches for the automated detection of MAs have been presented. A detailed overview is given in [3]. The majority of works is based on the single image scenario. Despite promising results reported in many studies, the reliable detection of morphed facial images is still an open research task. In particular, the generalizability and robustness of the published approaches could not yet be proven. The results are hardly comparable and comprehensible. The vast majority of publications use internal databases of the respective research groups for training and testing. In addition, different evaluation metrics are used in the publications, and some even state error rates of zero without specifying the number of samples. Since most implemented MAD procedures are not made publicly accessible, no comparative independent evaluation of the detection performance is possible (without cooperation with the respective authors).

Furthermore, most publications only use images from a single database and morphs generated with a single algorithm for training and testing, so that the generalization capability of the methods cannot be assessed across different databases and morphing methods. In publications on differential MAD, the comparison images used often show a low variance with respect to poses, facial expressions and illumination and are usually produced shortly after the reference image - in real scenarios such as border control, a much higher variance is to be expected. In addition, most studies neglect the probable application of image post-processing techniques by an attacker, such as subsequent image sharpening, and the print-scan transformation.

SINGLE IMAGE MAD APPROACHES

The single-image MAD approaches can be categorized into three classes: Texture descriptors, e.g., in [6], forensic image analysis, e.g., in [7], and methods based on deep neural networks, e.g., in [8]. These differ in the artefacts they can potentially detect. A brief overview is given in Table 1.

Table 1: Categories of singe image MAD approaches.

Category	Analyzed artefacts
Texture descriptors	Smoothened skin texture, ghost artefacts/ half-shade effects (e.g., on pupils, nostrils), distorted edges, offset image areas
Forensic image analysis	Sensor pattern noise, compression artefacts, inconsistent illumination or color values
Deep-learning approaches	All possible artefacts learned from a training dataset

DIFFERENTIAL MAD APPROACHES

Differential MAD can be categorized into approaches that perform a biometric comparison directly with the two facial images, e.g., in [9], and algorithms that attempt to reverse the (potential) morphing process, e.g., in [10]. In the former category, features from both face images, the potentially morphed facial image and the probe image, are extracted and then compared. The comparison of the two feature vectors and the classification as bona fide comparison or MA is usually done using machine learning techniques. By specifically training these procedures for the recognition of MAs, they can - in contrast to facial recognition algorithms - learn to recognize specific patterns within the differences between the two feature vectors for these attacks. This has already been demonstrated for features derived from general purpose texture descriptors. While training a deep neural network from scratch in order to learn discriminative features for MAD requires a high amount of training data, pre-trained deep networks can be employed.

The second type of differential MAD procedure aims at reversing the morphing process in the reference image ("de-morphing") by using a probe image. If the reference image was morphed from two images and the probe image shows a person contributing to the morph (the attacker), the face of the accomplice would ideally be reconstructed, which would be rejected in a subsequent comparison with the probe image using biometric face recognition; if, on the other hand, a bona fide reference image is available, the same subject should still be recognizable after the reversal of a presumed morph process with the probe image, and thus the subsequent comparison of the facial recognition process should be successful.

MAD BASED ON DEEP FACE REPRESENTATIONS

For both single image MAD and differential MAD, a straightforward approach is to train a classifier on deep features computed by existing convolutional neural networks (CNNs) for

biometric face recognition. The advantage of this approach is that it benefits from the strength of CNNs to extract relevant features from image data but does not require the large amount of data typically necessary to train a CNN. While the features extracted by face recognition networks have not been trained to detect morph attacks, at least in the differential scenario, they might still be very useful for MAD: As the morphed face image does not only contain biometric features of the attacker but also those of the accomplice, its deep face features should, at least in certain aspects, considerably deviate from those detected in the probe image. The vulnerability of face recognition networks to morph attacks does not necessarily imply that the features extracted by those are not eligible for MAD but can also be explained by an inaptly chosen classification method (which is typically based on simple geometric distances). Thus, one can hope that a new classifier trained for MAD on deep face features may be able to recognize the characteristic differences in the features between morphs and probe images.

In [11], deep face representations, i.e., VGG-Face16 and VGG-Face2, have been employed to train machine learning-based classifiers for single-image MAD. Promising detection rates have been reported in the presence of printing/scanning and heterogeneous image sources.

In a preliminary study of the authors, conducted in the course of the FACETRUST project, deep face features of both commercial and open source face recognition systems were employed to develop differential MAD. Deep face representations extracted from reference and probe images were combined, e.g., by element-wise subtraction or concatenation, and the resulting vectors were then used to trained machine learning-based classifiers for differential MAD.

The following conclusions regarding performance/generalizability are reached:

- *Detection performance*: the detection performances achieved are promising and highly robust with respect to image post-processing, i.e., image compression, image resizing and even print-scan transformation. This is a clear advantage over MAD based on texture descriptors, which is typically quite sensitive to post-processing, particularly in more challenging scenarios. Moreover, in some cases it turned out to be favorable to perform training on digital images, which have not been printed and scanned, to obtain improved detection rates even for scanned images.

- *Heterogeneous morphing algorithms*: morphs generated by morphing algorithms which produce obvious artefacts, e.g., clearly visible ghost artefacts, were generally detected with higher accuracy. Furthermore, the recognition performance slightly degrades if training and evaluation sets contain morphs generated by different morphing algorithms.

- *Heterogeneous databases*: if training and testing is conducted on heterogeneous face image databases which contain face images with different conditions, e.g., variations in pose and lightning, detection performance is negatively affected. On databases obtained from subsets of the publicly available FERET and the FRGCv2 face database, experiments revealed higher detection accuracy on the FERET subset in which probe images only contain slight variations in expression and pose as opposed to the FRGCv2 subset, which additionally comprises probe images with variations in lightning and focus. It can be concluded that

strong variations in lightning and focus of probe images represent especially challenging conditions for differential MAD.

- *Machine learning-based classifiers*: among the tested machine learning-based classifiers, i.e., AdaBoost, Gradient Boosting, Random Forest and Support Vector Machine (SVM), SVM-based classifiers generally revealed most competitive detection performance across the vast majority of conducted experiments.

- *Commercial vs. open-source*: while commercial face recognition algorithms frequently outperform corresponding open-source implementations, this is not necessarily the case for MAD. Precisely, for the task of MAD, deep face representations obtained from open-source algorithms, e.g. FaceNet or ArcFace, might be better suited, compared to deep features extracted by commercial face recognition systems.

ISSUES AND CHALLENGES

In research on MAD, there are various open questions and challenges:

Evaluation metrics: Even though initial efforts have already been made to introduce them, standardized metrics for evaluating the performance of MAD procedures are not yet available; these should be defined uniformly (ideally as an international standard) and applied in publications on MAD procedures in order to enable a meaningful comparison of the proposed approaches.

Evaluation protocols: To obtain reproducible and statistical significant results performance evaluations of proposed MAD approaches should be transparent and based on sufficient data. Used face databases must be split into subject-disjoint sets for training and evaluation. Reporting the used number of sample and conducted amount of comparisons is essential in order to interpret obtained results in a meaningful way.

Generalizability of MAD approaches: The majority of the MAD methods published so far - in particular the single image MAD methods - aim at the detection of artefacts that can easily be avoided, e.g., clearly visible ghost artefacts, double compression artefacts and changed image noise patterns. Hence, reported detection rates tend to be over-optimistic. In contrast, research should focus on the development of MAD methods that detect artefacts that are difficult to avoid. In addition, MAD approaches are, like any classification task, susceptible to overfitting to training data. Therefore, when evaluating MAD approaches, images of which source and properties differ from those of the training data, i.e., images from other databases and morphs created with other techniques, should be employed.

For border control scenarios, MAD techniques need to be robust against print-scan transformations, resizing and strong compression of reference images. Similarly, in the case of differential MAD, considerable variance of illumination, background, pose, appearance (hair, beard, glasses, etc.) and aging (up to 10 years for passports) can be expected in probe images. In order to be applicable to these scenarios, MAD approaches should be trained and evaluated on images exhibiting these characteristics.

Unfortunately, post-processing steps applied to reference images like printing/scanning and strong image compression have been found to cause drastic drops in the detection performance at least for single image MAD, since artefacts caused by morphing vanish in the post-processed reference. In order to reduce this issue in the long term, responsible authorities should raise the requirements for image quality, resolution and size of face images to be stored in electronic travel documents. Eventually, the susceptibility of the passport issuance processes can be eliminated by using live enrolment stations.

Figure 3: From left to right: original reference; reference printed, scanned (300 dpi), resized (360x465 pixels) and compressed (JPEG 2000, 15KB); probe with slight rotation; probe with changing expression and variation in illumination.

Databases: Currently, the publicly available facial image databases do not represent the characteristics and variance of real-world scenarios. To the authors' knowledge, there is no public database containing a large number of printed and scanned facial images. Furthermore, there is no database comprising face images which fulfill the conditions of reference and probe images needed to simulate a realistic boarder control scenario, i.e., containing both images conforming to the ICAO specifications for passport photographs and images resembling all variations (in particular aging) to be expected for live images in a border control. Figure 3 depicts face images taken from the FRGCv2 database which reflect at least some of the variance expected in a real border control scenario. In addition, there is just one database with morph images of good quality that has been made available[5], and the creation of morphs of high quality is still laborious with publicly available tools.

In order to overcome this issue, border control agencies could collect large databases with images that resemble the characteristics of images typically met in border control scenarios. These images should comprise bona fide reference images taken in accordance with ICAO requirements [12] as well as high-quality morphs of these (created with various methods). To all reference images realistic post-processing steps (e.g., printing and scanning, resizing to approx. 400x500 pixels and JPEG-2000 compression to 15KB) should be applied. The database should also contain corresponding probe images with realistic distribution of illumination, pose, appearance and aging. It should also be taken into account that in morph attacks, the variance between reference and probe is likely to be smaller than for bona fide authentication attempts. Ideally, such database would be made available to researchers for the development

5 https://www.linkedin.com/pulse/new-face-morphing-dataset-vulnerability-research-ted-dunstone

and evaluation of MAD methods. If operational data cannot be made available due to data protection legislation, images could be captured with volunteers under realistic conditions, e.g., using automatic border control gates.

The detection performance of differential MAD approaches can be influenced by the quality of the captured probe image. It is well-known that high recognition performance can only be achieved if the quality of the captured facial data is sufficient. As stressed in a recent study [13] by the Joint Research Centre (JRC) of the European Union, algorithms must be incorporated to ensure a robust determination of the face image quality.

Transparency: In scientific publications, the MAD procedures are usually presented in a way that they cannot easily be re-implemented by third parties without considerable effort while resulting re-implementations hardly achieve comparable recognition performance. Implementations of MAD procedures should therefore be made publicly available in order to guarantee the reproducibility of results that were achieved on public data. It is expected that the planned benchmark program of the National Institute of Standards and Technology (NIST) [14] will enable a quantitative comparison of published approaches in the near future. Border control agencies could support this program by providing realistic image data or information on the characteristics and variance of the images to be expected in border control scenarios.

SUMMARY

Morph attacks pose a high security risk to modern facial recognition systems in particular for border control. To counteract this, reliable methods for morph attack detection must be developed. Various research groups from the fields of image processing and biometrics have recently published scientific papers on this topic, and several publicly funded research projects are currently dealing with this problem. However, research in this field is still in its infancy and does typically not address the variance of the image data available in border control scenarios. The development of MAD approaches that are effective and robust in real-world scenarios will require a considerable amount of future research as well as close collaborations with border guard agencies.

ACKNOLEDGEMENTS

This work was partially supported by FACETRUST project of the Federal Office for Information Security (BSI).

REFERENCES

[1] M. Ferrara, A. Franco and D. Maltoni, „On the Effects of Image Alterations on Face Recognition Accuracy", in Face Recognition Across the Imaging Spectrum, Springer International Publishing, 2016.

[2] J. D. Robertson, A. G. Mungall, D. Watson, A. K. Wade, J. S. Nightingale and S. Butler, „Detecting morphed passport photos: a training and individual differences approach", Cognitive Research: Principles and Implications, 2018.

[3] U. Scherhag, C. Rathgeb, J. Merkle, R. Breithaupt, and C. Busch. Face recognition systems under morphing attacks: A survey. IEEE Access, 7:23012–23026, 2019.

[4] ISO/IEC JTC1 SC37 Biometrics, ISO/IEC IS 30107-3:2017, IT – Biometric presentation attack detection – Part 3: Testing and Reporting, 2017.
[5] U. Scherhag, A. Nautsch, C. Rathgeb, M. Gomez-Barrero, R. N. J. Veldhuis, L. Spreeuwers, M. Schils, D. Maltoni, P. Grother, S. Marcel, R. Breithaupt and R. Ramachandra, „Biometric Systems under Morphing Attacks: Assessment of Morphing Techniques and Vulnerability Reporting", in Proceedings of the 2017 International Conference of the Biometrics Special Interest Group (BIOSIG), 2017.
[6] R. Ramachandra, K. B. Raja and C. Busch, „Detecting morphed face images", in Proceedings of the 8th International Conference on Biometrics Theory, Applications and Systems (BTAS), 2016.
[7] C. Kraetzer, A. Makrushin, T. Neubert, M. Hildebrandt and J. Dittmann, „Modeling Attacks on Photo-ID Documents and Applying Media Forensics for the Detection of Facial Morphing", in Proceedings of the 5th ACM Workshop on Information Hiding and Multimedia Security - IHMMSec '17, 2017.
[8] C. Seibold, W. Samek, A. Hilsmann and P. Eisert, „Detection of Face Morphing Attacks by Deep Learning", in Digital Forensics and Watermarking, 2017.
[9] U. Scherhag, C. Rathgeb and C. Busch, „Towards detection of morphed face images in electronic travel documents", in Proceedings of the 13th IAPR Workshop on Document Analysis Systems (DAS), 2018.
[10] M. Ferrara, A. Franco und D. Maltoni, „Face Demorphing", IEEE Transactions on Information Forensics and Security, 2018.
[11] M. Ferrara, A. Franco, D. Maltoni: „Face morphing detection in the presence of printing/scanning and heterogeneous image sources", arXiv:1901.08811, 2019.
[12] International Civil Aviation Organization, ICAO Doc 9303, Ma-chine Readable Travel Documents - Part 9: Deployment of Biometric Identification and Electronic Storage of Data in MRTDs (7th edition), 2015.
[13] J. Galbally, P. Ferrarra, R. Haraksim, A. Psyllos and L. Beslay, "Study on Face Identification Technology for its Implementation in the Schengen Information System", Publications Office of the European Union, 2019.
[14] M. Ngan, P. Grother and K. Hanaoka, „Performance of Auto-mated Facial Morph Detection and Morph Resistant Face Recognition Algorithms", National Institute of Standards and Technology (NIST), 2018.

Distributed and GDPR/IPR Compliant Benchmarking of Facial Morphing Attack Detection Services

Andrey Makrushin[1], Christian Kraetzer[1], Gregor Mittag[2], Hermann Birkholz2, Uwe Rabeler[3], Andreas Wolf[3], Clemens Seibold[4], Anna Hilsmann[4], Peter Eisert[4], Lukasz Wandzik[5], Raul Vicente Garcia[5], Jana Dittmann[1]

[1] *Otto-von-Guericke University of Magdeburg*
andrey.makrushin@ovgu.de, {kraetzer, jana.dittmann}@cs.iti.uni-magdeburg.de
[2] *DERMALOG Identification Systems GmbH*
{gregor.mittag, hermann.birkholz}@dermalog.com
[3] *Bundesdruckerei GmbH*
{uwe.rabeler, andreas.wolf}@bdr.de
[4] *Fraunhofer HHI*
{clemens.seibold, anna.hilsmann, peter.eisert}@hhi.fraunhofer.de
[5] *Fraunhofer IPK*
{lukasz.wandzik, raul.vicente}@ipk.fraunhofer.de

Abstract: Having started in June 2016 and lasting till May 2020, the research project ANANAS, funded by the German Federal Ministry of Education and Research (BMBF), was the first inter-institutional research initiative established for designing Morphing Attack Detection (MAD) methods. The project is the prompt response to the paper "The Magic Passport" published by Ferrara et al. in 2014, demonstrating the threat that the face morphing attack (FMA) poses to the identity verification procedure based on facial photographs, including the usage of electronic Machine Readable Travel Documents (MRTD). Apart from the considerable scientific contribution reflected in a large number of international publications, an important result of the project is the distributed framework of MAD services, face image databases as well as the benchmarking service allowing for performing General Data Protection Regulation (GDPR) and Intellectual Property Rights (IPR) compliant evaluation. This paper introduces the MAD benchmarking framework by reporting its infrastructure, communication protocol and the results of an exemplary evaluation run. The design of the framework brought together the expertise of industrial companies with the innovative power of research institutions. The framework enables a statistically significant performance evaluation of MAD services. The individual MAD services as well as their combination may help countering the threat posed by FMA. The framework offers a research tool which could be used not only by project members but also by external parties.

Keywords: Face Morphing Attack, Morphing Attack Detection, Project ANANAS, GDPR and IPR Compliant Benchmarking

1. INTRODUCTION

Biometric verification plays an increasingly important role in our daily life. Automated identity (ID) document checking technology increasingly supports border control officers and partly allows for automation of processes. Back in 2004, the International Civil Aviation Organization (ICAO) selected face as the primary biometric trait used with an MRTD. Along with all advantages of face verification there are serious security concerns caused by the vulnerability in the submission process of facial photographs [2]. If a morphed photograph appears in a document, both border guards and automated face recognition (AFR) systems are very likely to

accept any of the constituent individuals [3][11], abolishing the unique links between individuals and their ID documents. Although protection from FMA is a young research field, several research groups have already designed and prototypically implemented a bunch of morphing attack detection (MAD) approaches [12]. The still missing part of the research is the fair and GDPR/IPR compliant benchmarking of MAD approaches.

The trustworthiness of a benchmarking process highly depends on the input data for experiments. A benchmark maintainer should prepare two image sets: genuine face images and morphed face images. By collecting genuine face images, one should bear in mind that facial photographs, in particular those of a high quality, are regarded as personal biometric data which is protected by GDPR. Sharing of such data with third parties is prohibited by the European Union law warranting the image donors' right to request image removal at any moment. As a consequence, facial photographs must be stored on a protected media disabling the option of copying the data to any uncontrolled media. Putting the data into the public domain is prohibited. An elegant solution is not to grant access to the database, but to ask the developers of MAD approaches to submit their algorithms for evaluation. It is important that the benchmarking is conducted by an independent body ensuring that the algorithms are not misused, e.g., disassembled or offered to third parties without consent of the owner. While preparing morphed face images, one should bear in mind that the morphing approach used by a perpetrator may differ from that used for generating input data for an experiment.

While the number of publications on MAD approaches is growing fast (see recent overviews [12][15]), the efforts on benchmarking of such methods are slow on the uptake. All former efforts on benchmarking of MAD approaches can be assigned to one of two categories: public benchmarks with the data unknown to MAD developers and individual benchmarks with self-collected or public data aiming at understanding the characteristics of MAD approaches and improving their performance.

Currently there exist two public benchmarks for MAD approaches: the FVC-onGoing Face Morphing Challenge maintained by the UNIBO [4] and the FRVT MORPH maintained by the NIST [14]. Both challenges provide i/o interfaces and encourage the potential participants to submit MAD solutions that are compliant with the given runtime environment specifications. The submissions are supposed to be executed on local servers of the organizers with undisclosed genuine and morphed face images and the benchmarking results are supposed to be publicly reported. At the time this paper was drafted (June 2019), the Web sites of both challenges reported no participation results. The aforementioned public benchmarks share the same drawbacks. MAD algorithms have to be re-implemented to comply with the very restrictive run-time environment and additionally the numbers of submissions and test runs per participant are limited. Moreover, the composition of the test dataset cannot be changed, which restricts the understanding of how specific image characteristics influence the error rates of a MAD algorithm. On the one hand, all these constraints make sense for a public benchmark, because otherwise the system would be prone to sensitivity attacks conducted by participants aiming at dominating the challenge by creating detectors that perfectly fit to the test dataset instead of generally preventing FMA. On the other hand, to better understand the shortages of their MAD solutions, the researchers are forced to come up with alternative (non-public)

benchmarking. Early scientific publications on specific aspects of benchmarking [5][6][13], were not implemented into a fully operational public benchmark.

Our proposed benchmarking framework is a network of RESTful Web services including those automatically generating high quality morphed images, MAD services, private face image databases, and a meta-database of image IDs. The benchmarking Web service, which is a core part of our framework, requests the image IDs that meet certain criteria from the meta-database. Then the images are derived from one of the image databases and sent to the MAD services. The responses of these MAD services are stored in the benchmarking log. A user communicates with the benchmarking Web service only by defining the criteria for image selection and by choosing the MAD services to test. All components of the framework and the communication between them remain hidden.

Experiments demonstrated the effectiveness of the benchmarking engine: We compared the performances of selected existing MAD services in an exemplary run on a particular dataset. The main benefits of such benchmarking are the secure transfer of private biometric images, the possibility to develop MAD solutions on different platforms using any programming language, the non-disclosure of the implementation details of MAD solutions, and the possibility to run an individually configured benchmark for any registered user. This design (together with a service-level agreement signed by the maintainers on the stateless nature of their services) ensures GDPR and IPR compliance. Since the MAD services are completely independent and maintained by different institutions, new services can be easily added to the benchmarking framework. The same applies to image databases requiring only the registration of new images in the meta-database. The presented framework was designed and implemented as a part of the research project ANANAS enabling a fair comparison of MAD services provided by the project partners and a better understanding of advantages and shortcomings of certain MAD approaches.

2. METHODS

While designing the infrastructure for the benchmarking framework the focus was on security, sustainability and extendibility. The first core part of the framework is the infrastructure based on Representational State Transfer (REST) technology. The framework is in its nature a network of RESTful Web services which exchange data in a JSON format. The JSON objects are sent via HTTP POST requests. Web services are independent of each other so that they can be easily replaced or extended by the other ones. There are no restrictions on the hardware, operating system, or programming language for MAD algorithm providers. The code of Web services is undisclosed and the inter-service as well as client-service communication is done via the data exchange protocols.

The second core part is the secure storing and transfer of images. Datasets of face images are stored persistently on the servers of the image providers. The images are transferred via encrypted channels and the image consumers (e.g., MAD Web services) keep these images exclusively in a protected volatile memory. The images for benchmarking can be selected using a database containing image meta data such as image ID, image type (genuine or morphed), image characteristics, and characteristics of a data subject. The annotations to morphed images include references to all images used for their generation.

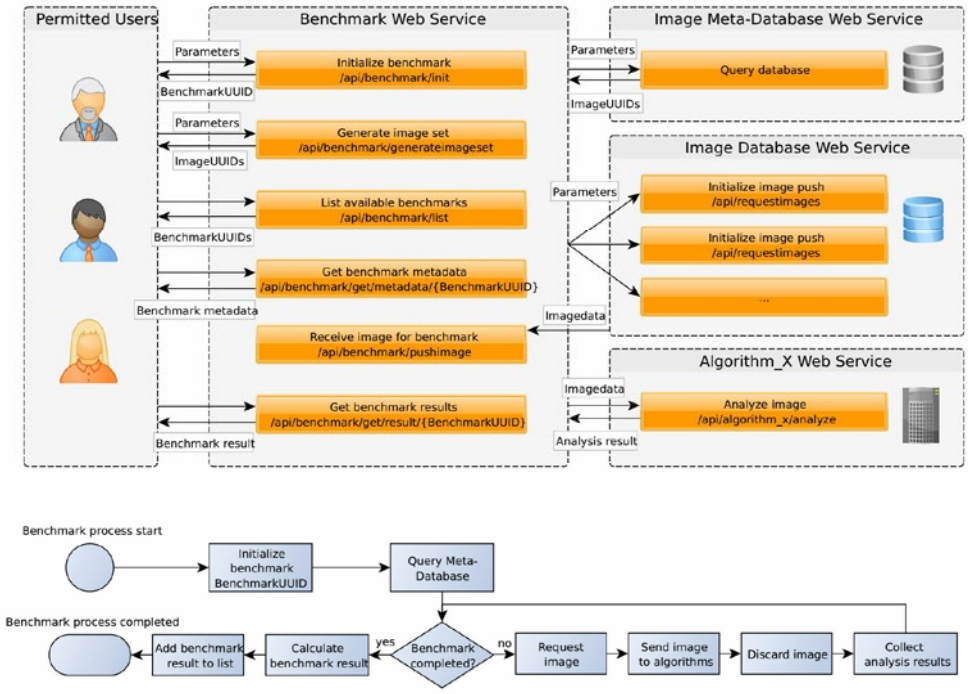


Figure 1: Architecture of the benchmarking engine and benchmarking workflow

The benchmarking framework architecture and the benchmarking workflow are illustrated in Figure 1. The *Benchmark Web Service* serves as an entry point for benchmarking requests. *Permitted Users* can initialize a benchmark by specifying a list of images or image meta data, a list of MAD algorithms, and a list of image manipulation options. From the given image meta data, a request is constructed and sent to the *Image MetaDatabase Web Service*. This meta-database stores meta data on all images hosted by the project members and links these images to unique identifiers. The result of a database request is merged with the explicitly provided image list. Each element of an image list contains the image ID as well as the ID of the image owner. Using this information several requests are constructed and sent to *Image Database Web Services* hosted by the corresponding vendors. The *Image Database Web Services* return the requested images to the *Benchmark Web Service*, which passes them to the image manipulation pipeline configured by the image manipulation options mentioned above. Afterwards, the images are sent to the selected MAD Web services (called *Algorithm_X Web Services*). The results of image evaluations are returned to the *Benchmark Web Service* and stored persistently.

The results of a benchmarking run can be requested via a unique BenchmarkUUID. These results can be reproduced by running a new benchmark with the same parameterization. In the case of several benchmarks running in parallel and operating with same images, the analysis requests are sent only once to the corresponding MAD web services and the evaluation results are used for all benchmarks. The MAD services support both "blind" detection based on a passport photograph only and detection in the presence of a "live" photograph. Thanks to the integrated image manipulation engine, the influence of anti-forensic approaches to

the detection performance can be evaluated. The interface based on Web services enables easy integration with the Automated Border Control (ABC) reference environment of Bundesdruckerei GmbH, Germany. The list of Web service providers is given in Table 1. Note that Table 1 also includes the morphing Web services which are strictly speaking not a part of the benchmarking framework. However, these Web services are used to fill the image databases with morphed face images.

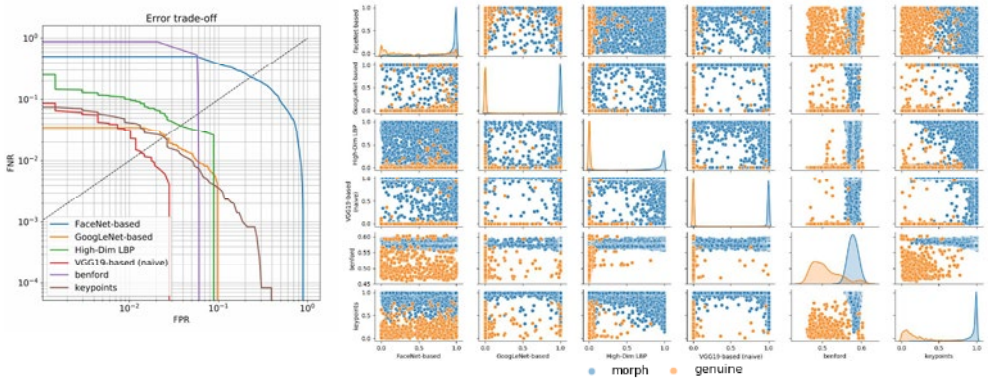
By May 2019, the total number of face images registered in the meta-database exceeded five million. Currently, ten MAD algorithms are available: four hosted by OVGU, four by HHI, and two by IPK.

Table 1: Hosts of Web services

Benchmark Web service	- DERMALOG Identification Systems GmbH (Dermalog)
Meta-Database Web service	- Fraunhofer Heinrich-Hertz Institute (HHI)
Image Database Web services	- DERMALOG Identification Systems GmbH (Dermalog) - Fraunhofer Heinrich-Hertz Institute (HHI) - Fraunhofer Institute for Production Systems and Design Technology (IPK) - Otto-von-Guericke University of Magdeburg (OVGU)
Algorithm_X Web services	- Fraunhofer Heinrich-Hertz Institute (HHI) - Fraunhofer Institute for Production Systems and Design Technology (IPK) - Otto-von-Guericke University of Magdeburg (OVGU)
Morphing Web services	- Fraunhofer Heinrich-Hertz Institute (HHI) - Fraunhofer Institute for Production Systems and Design Technology (IPK) - Otto-von-Guericke University of Magdeburg (OVGU)

3. FINDINGS AND ARGUMENT

Here, we report the results of an exemplary benchmarking run lasting from November 1st to December 3rd, 2018. Based on 680 genuine images (605 male/75 female) selected from the PUT face database [8] we generated 12000 morphed face images with two algorithms, 6000 each. Both morphing algorithms are deployed as Web services, the first one by OVGU [13] and the second one by HHI [16]. There are 5321/679 male/female morphs generated by OVGU Web service and 5730/270 male/female morphs by HHI Web service. The ethnicities of the data subjects are Caucasian, Latin, and Middle Eastern. For the purpose of quality assessment, all morphed face images were compared with constituent face images using the Dermalog Face Recognition software [1] as a commercial off-theshelf product. The requirement for the inclusion of a morphed face image into the dataset is that both comparison scores exceed 80% similarity.



a) b)

Figure 2: (a) DET curves and (b) discriminatory power of the MAD Web services

We benchmarked six MAD algorithms: *FaceNet-based* [18], *GoogLeNet-based* [16], *High-Dim LBP* [18], *VGG19-based (naive)* [17], *benford* [10], and *keypoints* [9]. For performance evaluation, we used standard metrics for binary classification problems: FPR vs. FNR. Note that since we endeavour to detect morphs, morphed images are considered as positive samples and genuine images as negative. FPR is the fraction of genuine images that are falsely classified as morphed images (also called *false alarms*) and FNR is the fraction of morphed images that are falsely classified as genuine images (also called *misses*). Within the context of presentation attack detection the FPR is denoted as BPCER and the FNR as APCER [7]. The Detection Error Trade-off (DET) graph in Figure 2(a) demonstrates the performances of the MAD algorithms at different operating points. The discriminatory power of the evaluated algorithms is visualized in Figure 2(b). The diagrams on the main diagonal show the matching score distributions of morphed (blue) and genuine (orange) samples while non-diagonal diagrams reveal the potential for fusion by demonstrating the pair-wise correlation between MAD algorithms. A point on a non-diagonal diagram represents an image by a pair of matching scores resulting from two different MAD algorithms (one on the X-axis and another on the Y-axis). A matching-score fusion of MAD algorithms is expected to improve the recognition performance if the blue and orange points can be clearly separated by a diagonal line. Horizontal or vertical separation lines indicate the domination of one MAD algorithm over another and a limited potential for fusion. Note that the MAD services operate at fixed thresholds, i.e. whether the algorithm classifies an image as morphed or genuine critically depends on the chosen decision boundary.

Table 2: Detection performance of the MAD Web services, the best performances are highlighted

	FPR	FNR	FNR @ 0.01% FPR	FNR @ 0.1% FPR	FNR @ 1% FPR	FNR @ 10% FPR	EER
FaceNet-based	43.53%	12.54%	59.23%	59.09%	57.69%	37.65%	23.21%
GoogLeNet-based	3.97%	1.52%	**4.06%**	**4.03%**	3.66%	0.30%	2.53%
High-Dim LBP	0.15%	18.02%	25.96%	25.53%	10.03%	1.45%	3.77%
VGG19-based (naive)	1.03%	2.75%	10.29%	9.24%	**3.13%**	**0.00%**	**1.36%**
benford	53.09%	0.00%	99.94%	99.38%	93.82%	0.00%	5.97%
keypoints	4.71%	1.14%	7.93%	7.66%	4.19%	0.36%	2.43%

Table 2 shows the factual FPR and FNR values with the predefined thresholds used by MAD services, theoretical FNR values at particular levels of FPR, and the Equal Error Rates (EER). Regarding the EER and the detection performance at FPR higher than 1%, the best algorithm is *VGG19-based (naive)* followed by *GoogLeNet-based* and *keypoints*. At low levels of FPR (0.1% and lower), the *VGG19-based (naive)* has FNR of over 9%, while the *GoogLeNet-based* of around 4% and *keypoints* of under 8%. The error rates of *High-Dim LBP, benford* and *FaceNetbased* are strongly imbalanced requiring more careful selection of decision thresholds. Observing the EER values, we see that with properly selected decision thresholds the MAD services demonstrate solid detection performance. However, these error rates are too high clearly indicating that the algorithms are still not mature for practical application.

4. CONCLUSIONS

Thanks to its design based on the REST technology, the presented benchmarking framework is a powerful and flexible tool for GDPR/IPR compliant evaluation of MAD approaches. Our proposed interfaces for MAD services support both "blind" detection based on a passport photograph only and detection in the presence of a "live" photograph. The integrated image manipulation tools enable for evaluation of the influence of anti-forensics. The benchmarking results are reproducible and transparent for the benchmark users.

The demonstrated benchmarking run does not cover all capabilities of the framework, but gives an idea how the benchmark can be configured and how the results can be visualized.

Due to the flexibility of interfaces, the MAD services can be used not only as a part of the framework, but also individually. Currently, the MAD Web services are integrated into the ABC reference environment of Bundesdruckerei GmbH, Germany.

The proposed benchmarking framework can be used as an alternative to FVC-onGoing Face Morphing Challenge and NIST FRVT MORPH having an advantage of providing more flexibility by developing of MAD approaches and granting more transparency in test image datasets. Parties interested in benchmarking MAD approaches are invited to contact the authors in order to register as users of the framework so that they can browse through existing benchmarks as well as configure and run own ones.

5. ACKNOWLEDGEMENT

This work has been funded in part by the German Federal Ministry of Education and Research (BMBF) through the research programme ANANAS under the contract no. FKZ: 16KIS0508 (BDR), 16KIS0509K (OVGU), 16KIS0510 (DERMALOG), 16KIS0511 (HHI) and 16KIS0512 (IPK).

6. REFERENCES

1. Dermalog Face Recognition, https://www.dermalog.com/products/software/face-recognition, online, 06.06.2019
2. M. Ferrara, A. Franco, and D. Maltoni, "The Magic Passport," Proc. IEEE Int. Joint Conf. on Biometrics, pp. 1–7, 2014

3. M. Ferrara, A. Franco, and D. Maltoni, "On the Effects of Image Alterations on Face Recognition Accuracy," in Face Recognition Across the Electromagnetic Spectrum, T. Bourlai (Ed.), pp. 195–222, Springer, 2016
4. FVC onGoing: Face Morphing Challenge, https://biolab.csr.unibo.it/FVCOnGoing/UI/Form/BenchmarkAreas/Benchmark AreaFMC.aspx, online, 06.06.2019
5. M. Gomez-Barrero, C. Rathgeb, U. Scherhag, and C. Busch, "Predicting the vulnerability of biometric systems to attacks based on morphed biometric information," IET Biometrics 7(4): 333-341, 2018
6. M. Hildebrandt, T. Neubert, A. Makrushin, and J. Dittmann, "Benchmarking Face Morphing Forgery Detection: Application of StirTrace for Impact Simulation of Different Processing Steps," Proc. 5th Int. Conf. on Biometrics and Forensics (IWBF), 2017
7. ISO/IEC JTC1 SC37 Biometrics, ISO/IEC IS 30107-3:2017 Information technology - Biometric presentation attack detection - Part 3: Testing and reporting
8. A. Kasiński, A. Florek, and A. Schmidt, "The PUT face database," Image Processing and Communications 13:59-64, 2008
9. C. Kraetzer, A. Makrushin, T. Neubert, M. Hildebrandt, and J. Dittmann, "Modeling Attacks on Photo-ID Documents and Applying Media Forensics for the Detection of Facial Morphing," Proc. 5th ACM Workshop on Information Hiding and Multimedia Security (IH&MMSec), pp. 21–32, 2017
10. A. Makrushin, C. Kraetzer, T. Neubert, and J. Dittmann, "Generalized Benford's Law for Blind Detection of Morphed Face Images," Proc. 6th ACM Workshop on Information Hiding and Multimedia Security (IH&MMSec), pp. 49-54, 2018
11. A. Makrushin, T. Neubert, and J. Dittmann. "Automatic generation and detection of visually faultless facial morphs," Proc.
12th Int. Joint Conf. on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 6: VISAPP, pp. 39– 50, 2017
12. A. Makrushin and A. Wolf, "An Overview of Recent Advances in Assessing and Mitigating the Face Morphing Attack," Proc. 26th European Signal Processing Conference (EUSIPCO), pp. 1017-1021, 2018
13. T. Neubert, A. Makrushin, M. Hildebrandt, C. Kraetzer, and J. Dittmann, "Extended StirTrace Benchmarking of Biometric and Forensic Qualities of Morphed Face Images," IET Biometrics 7(4):325–332, 2018
14. NIST FRVT MORPH, https://www.nist.gov/programs-projects/face-recognition-vendor-test-frvt-morph, online, 06.06.2019
15. U. Scherhag, C. Rathgeb, J. Merkle, R. Breithaupt, and C. Busch, "Face Recognition Systems Under Morphing Attacks: A Survey," IEEE Access 7:23012-23026, 2019
16. C. Seibold, W. Samek, A. Hilsmann, and P. Eisert, "Detection of Face Morphing Attacks by Deep Learning," Proc. 16th Int. Workshop on Digital Forensics and Watermarking (IWDW), pp. 107-120, 2017
17. C. Seibold, W. Samek, A. Hilsmann, and P. Eisert, "Accurate and Robust Neural Networks for Security Related Applications Exampled by Face Morphing Attacks," CoRR abs/1806.04265, 2018
18. L. Wandzik, G. Kaeding, and R. Vicente-Garcia, "Morphing Detection Using a General- Purpose Face Recognition System," Proc. 26th European Signal Processing Conference (EUSIPCO), pp. 1012-1016, 2018

Morped Passport Photo Detection by Human Observers

David J. Robertson[1]
[1] *School of Psychological Sciences and Health, University of Strathclyde, Glasgow, UK.*
E-Mail: david.j.robertson@strath.ac.uk
Phone: +44 (0)141 548 4461
Lab: https://www.strathacpl.com/

Abstract

Introduction: The use of fraudulent passports for identity verification represents a significant threat to national security. Modern passports contain counterfeit prevention measures (e.g., printed patterns visible only under specific artificial illumination) which make any attempts to alter or duplicate the document itself unlikely to go unnoticed. As a result, fraudsters are now known to be focusing on obtaining FOG (fraudulently obtained but genuine) passports. FOG passports are real documents which are wrongly issued to fraudulent applicants, and they arise when a confederate, who holds a genuine passport, submits a renewal application with the photo of a similar looking client. If the mismatch between the renewal image and the image held on file goes undetected, a FOG passport is issued which can be used illegally by the client individual. In a recent advancement in this approach, criminals are seeking to increase their success rate by submitting a morphed passport photo, an image of the confederate and the client which has been digitally blended together and which retains a likeness of both individuals. Border security agencies have only recently detected the use of passport morphs, and research is required to ensure that the relevant agencies and practitioners stay one step ahead of these criminal attacks. Here we use applied psychological science to quantify morph detection rates, to assess the effectiveness of a morph detection training task, and to evaluate the use of individuals who show a high aptitude on a test of unfamiliar face matching as a potential countermeasure.

Keywords: Passport Morphs, Identity Fraud, Face Recognition, Biometrics, Border Security

Methods: Across two published studies and a total of four experiments, we assess passport morph detection in human observers, both in a 'spot the morph' task and in a passport matching context (i.e. match, mismatch or morph decision).

Findings and Argument: Across these experiments we show that morph detection rates are often at or near chance level across tasks, that facial identification aptitude as measured by the Glasgow Face Matching Test (GFMT) and the Models Face Matching Test (MFMT) only partially predict morph detection performance, suggesting that super-recognisers may be only a moderately effective counter-measure, and finally we show that rudimentary morph awareness and detection training can significantly increase detection rates.

Conclusion and Recommendations: From our studies, we conclude that morph detection is a challenging task and one which is highly prone to error. However, effective counter-measures, at least for human observers, include morph awareness information, training, and the selection of individuals with a natural aptitude for facial identification (i.e. super-recognisers).

References

Robertson, D. J., Kramer, R. S. S., & Burton, A. M. (2017). Fraudulent ID using face morphs: experiments on human and automatic recognition. PLOS One, 12(3), [e0173319]. https://doi.org/10.1371/journal.pone.0173319

Robertson, D. J., Mungall, A., Watson, D. G., Wade, K. A., Nightengale, S. J., & Butler, S. (2018). Detecting morphed passport photos: a training and individual differences approach. Cognitive Research: Principles and Implications, 3(27), 1-11. https://doi.org/10.1186/s41235-018-0113-8

Face Morphing Attacks: What needs do be done

C. Busch, S. Caillebotte, U. Seidel, F. Knopjes, D. Maltoni, M. Ferrara, R. Veldhuis, L. Spreeuwers, K. Raja, R. Raghavendra, M. Gomez-Barrero, C. Rathgeb

Abstract: The intention of this paper is to summarize, what countermeasures are needed to mitigate the threat of electronic passports with morphed images. This paper indicates, what Frontex and the concerned national governments can and should do as short term and as long term countermeasures. The suggestions presented in this contribution are based on the discussions of the SOTAMD and iMARS project consortia which proposed a harmonized European approach to tackle the morphing attack threat with joint forces from industry, academia and governmental agencies.

Keywords: face recognition; face morphing attacks; vulnerability analysis; border control

Introduction

The problem of morphing attacks has been addressed in the biometric research community only recently, despite it was already identified back in 2004 in the presentation by Matthew Lewis and Philip Statham at the Biometrics Consortium Conference (BCC). Five years later in 2009 the morphing attack was classified as *vulnerability* of a biometric system in ISO/IEC FDIS 19792 stating: *"... a synthesised characteristic could be injected electrically during a replay attack or planted in the reference database. - feature sets comprising amalgamations of biometric features from 2 or more individuals, e.g. morphed facial images"*. However it took until 2014, before the feasibility of face morphing attacks was first demonstrated in the FIDELITY project by Matteo Ferrara, Annalisa Franco and Davide Maltoni and published in their IJCB paper "*The magic passport*" [Fer14]. Only then researchers started to investigate countermeasures for the problem [Ram16][Sch19]. In 2017 the iMARS consortium[1] was formed with the joint research resources from industry, academia and governmental agencies and is seeking now support from the EU H2020 research program. In 2019 the Dutch National Office for Identity Data and the German Bundeskriminalamt were tasked by European Commission DG Home, to investigate the State-Of-the-Art of Morphing Detection (SOTAMD)[2] by collecting an initial morphing test dataset and by evaluating currently available academic morphing attack detection solutions.

In order to maintain the control on migration of third country nationals, refugees and asylum seekers with the established procedures, Europe should immediately start an action to secure the trusted link between a MRTD and the document holder and to develop and deploy technical mechanisms that can detect a morph passport at borders. This paper describes the necessary steps that should be taken to protect European borders against the threat of morphing attacks. The following chapters describe how Frontex and government agencies of European Member States can support this process.

---

1 The iMARS consortium consists of Idemia, NTNU, University Bologna, University Twente, Hochschule Darmstadt, University Leuven, Dutch National Office for Identity Data, German Bundeskriminalamt, Vision-Box, Cognitec, IBS, EAB and various end users (border control agencies)

2 SOTAMD partners are Dutch National Office for Identity Data, German Bundeskriminalamt, University of Bologna, University of Twente, NTNU and Hochschule Darmstadt

Needs to re-establish a Trusted Link

Unfortunately, in many ICAO Members States and most European Member States, the facial image used for an electronic travel document is provided by the applicant in printed or digital form and not taken by means of 'live-enrolment' in an controlled environment such as a municipality office. Moreover, some countries even operate smartphone-based enrolments, such as the application process for the passport card in Ireland. This fundamental weakness must be stopped **immediately** and a European regulation should enforce that all Member States switch to live enrolment, as it is already operational e.g. in Norway and Sweden. Only then, with full control of the biometric capture process by a civil servant in the passport application office, trust in the link of passport holder to reference data can be assured. The iMARS consortium has proposed to define a secure ID Document application process, which is robust against enrolment fraud such that it will be made more difficult to apply for an ID document with a photograph that has been morphed or manipulated otherwise (e.g. data subjects that want to look younger or more beautiful) in order to influence the biometric recognition process, or by presenting a fraudulent document (in the case of first-time issuance or renewal). Citizens living abroad require a specific use case: the only feasible process for an EU citizen, living abroad, far away from an embassy or a consulate, where she/he can apply for an ID document, could be an application (i.e. passport renewal) from home, which would require specific precautions to prevent enrolment fraud. On the other side of the spectrum, even the seemingly secure live enrolment at a passport office requires precautions to detect a case that someone tries to enrol with a well-crafted facemask (i.e., conducting a presentation attack with a morphed face image on the mask).

The iMARS consortium proposed to define:

- Technical specifications for serving those use cases. Such specifications can be used in a new European regulation on passport application.
- The specifications could also include solutions that secure a wide range of application processes against enrolment fraud (e.g., live-enrolment with kiosk).
- Requirements for Presentation Attack Detection (PAD), to avoid for instance that a silicon mask is used against a face capture device in a live enrolment process.

Moreover, the regulation should define that for facial reference images, which are stored in the ICAO 9303 Logical Data Structure (LDS), the capture device certification scheme will be recorded in the data interchange format, as defined in the new extensible interchange format ISO/IEC 39794 [ISO39794]. This way, the future receiver of the facial reference image can have assurance that the image was captured with live enrolment and thus can be considered trustworthy.

As the passport application process in non-European states cannot be regulated, Europe should through its stakeholders initiate the discussion process, to suggest in the upcoming revision of ICAO 9303 [ICAO9303] a secondary mandatory biometric identifier (iris or fingerprint reference images). Note that ICAO 9303 already allows in Data Group 3 the storage of finger images

and in Data Group 4 the storage of iris images. In fact, the USA have recently decided to include an iris image into future passports. A future European border control point could then in case of suspicion regarding a potentially morphed face image rely on fingerprint or iris recognition for traveler identification beyond doubt. For iris, this comes with the advantage that face and iris probe images could be captured with one single capture device.

Further activities of European stakeholders to initiate a new EU regulation are needed.

Need to detect automatically Morph Passports at Borders

Given the validity period of electronic passports, after the future EU-wide transition to live enrolment European border management must anticipate that European passports are presented at least for the next 10 years[3] potentially containing morphed image; as well as passports of third-country nationals beyond the reach of the EU. One of the main goals of the iMARS consortium is to propose efficient solutions for border control points to detect ID documents containing manipulated/morphed images. The Morphing Attack Detection (MAD) solution is expected to enable efficient and reliable automatic data authenticity checks and elevate the process and security of biometric technology to a level that allows operational deployment. For deployed and potentially fraudulent passports, the MAD solutions suggested by iMARS will analyse those potentially manipulated documents. iMARS will provide solutions for the border control process on the one hand based on a differential analysis, where the images stored on the ID document are compared with a trusted live image of the ID document holder, while the capture process is run under supervision.

The iMARS consortium suggests for the processes at European border the development of explicit and implicit image pair detection algorithms (differential MAD – see Figure 1): iMARS explicit image pair based detection algorithms will use image pairs (morphed and bona fide) to setup various models for differential MAD. Regarding Implicit Image Pair detection algorithms, the use of the Deep Neural Network (DNN) approach, along with other methods, will be used to reach the expected progress.

Figure 1: Morphing attack detection scenarios. Left: Single image Morphing Attack Detection, Right: Differential Morphing Attack Detection

The challenge is that MAD systems can, to date, not generalize across databases (e.g. of different sample quality) and can either reliably detect morphed images stemming from a print and scan process yet. Thus, the two pressing objectives are to:

3 As Europe cannot impact passport application procedures in non-European countries, one should expect morphed passport to be presented at European borders way beyond that date.

- Improve the detection accuracy and its MAD's capability to generalize across databases.

- Measure detection accuracy as a function of the quality (e.g. are the 90-pixel inter-eye-distance sufficient).

The robustness of differential MAD will depend on both the different resolution of the enrolment images (i.e., in the passport), and on the quality of face image data and real-life noise (i.e., illumination, pose and shadow) that are commonly encountered in ABC systems. Specific challenges foreseen are to:

- Achieve high morphed detection rate without increasing the false rejection rate, even in the presence of image variations (i.e., pose, lighting, accessories, etc.)

- Adjust MAD for high-quality and low-quality low-resolution morphs. Get a better idea of the characteristics of these morphs in order to improve the detection performance.

A general challenge lies in the geometric transformation of the trusted live face image to the passport image geometry with a sufficient accuracy, i.e., accurate registration of the images is required. Further, the morph composition will have hyperparameters such as the percentage of the two images, contributing to the morph, in the morphing process, which needs to be estimated in the differential detection process.

Further research as suggested by the iMARS consortium is needed.

Need to detect automatically Morph Passports in Forensic Investigations

In order to support forensic investigations, the iMARS consortium suggests the development of explicit single image detection algorithms based on advanced feature extraction methods, which is especially relevant when no trusted image reference of the passport owner is available (see Figure 1). Regarding iMARS implicit single image detection, DNNs will be trained with large-scale face morph databases. As a consequence, robustness of the deep learning-based morph detector based on DNNs will be improved. Advanced machine learning techniques, e.g. transfer-learning, will be analysed, which will deal with different levels of quality, as the iMARS databases will cope with the problem of variability of face sample quality. This class of MAD solutions is also known as no-reference MAD or forensic MAD.

In a forensic investigation, the examination is based on the relatively low-resolution digital image stored in the passport, which has been processed by the authority or passport producer. While a morphed image may be visually indistinguishable to humans, the signal artefacts may be detected by MAD solutions. However, in carefully designed morphed images, the signal artefacts can be attenuated or completely suppressed. Further as the print and scan process tends to hide morphing artefacts, digital forensic tools are confronted with a challenge to detect alterations. As MAD methods mostly rely on trainable classifiers such as a Support Vector Machines (SVMs), their detection capability is linked to the quality of the training data. In addition, the artefacts have to pass the underlying feature extractor, i.e. even if there are artefacts in the image, those may not be reflected in the extracted features leading to misclassification of the MAD classifier. In explicit methods, the feature extraction is manually designed and the

classifier needs to find a proper decision boundary. Thus, the number of parameters to be estimated is low compared to the number of parameters to be learned in end-to-end approaches.

Advanced feature extraction methods and image forensic techniques will be employed to improve the MAD performance. Further techniques will be applied to enhance robustness and generalizability, e.g. fusion of multiple MAD subsystems.

iMARS will also categorize and consider admissible image processing steps in contrast to typical manipulative processing steps, and will analyse and model their traces in image data. This knowledge will strengthen the best detection methods and will identify rules to facilitate the distinction between correct bona fide and manipulated images.

Consequently, for implicit morphing attack detection, the number of training data used for end-to-end learning via CNNs has to be larger than for explicit methods. In addition, to avoid learning spurious correlations, sufficient variation in the training data is mandatory. Finally, it should be noted that, given the independence in the underlying concept, the fusion with other approaches is a promising direction.

Further research as suggested by the iMARS consortium is needed.

Need to compose Test Data and establish an Online Evaluation Platform

Testing of MAD solution can't be done without appropriate data. To tackle this issue, the SOTAMD consortium has composed a database of 150 individuals. Multiple passport enrolment images have been captured over the Summer 2019 with a typical eMRTD issuance process, including print and scan from the facial images, and from at least two automated border gates facial samples have been acquired (e.g., from the German BEC testing gate in Bonn-Siegburg). This dataset can be considered as a high quality data set. The data was split into a subset that is used for the morphing process and a disjoint subset that serves as bona fide image in a differential morph detection trial. Morphed face images were generated by each academic SOTAMD partner[4] with three different selected morphing algorithms. To mimic the application process as close as possible, both bona fide and morphed images are printed using professional photo printing devices and then scanned afterwards.

The iMARS consortium suggests to augment this initial data and to contribute an additional dataset (around 10,000 digital morphed images) with multiple enrolment and border gate probe images per subject, which are captured in a variety of illumination conditions. This new data will stem from challenging operational conditions at 5 selected borders (e.g. Cyprus, France, Greece, Israel, Portugal). This data will constitute the iMARS mixed-quality dataset. Moreover, the iMARS consortium suggested to contribute high quality morphs. The main challenges obtaining good morphs of images lie in: i) the mapping of the corresponding image positions or elements, and ii) the proper fusion of the image texture information. While automated morphing strategies lead to visually appealing results in certain scenarios, the results can degrade considerably under conditions such as pose variation. Although the ISO/IEC 19794-5 standard

4 Academic SOTAMD partners are University of Bologna, University of Twente, Norwegian University of Science and Technology and Hochschule Darmstadt [5] https://biolab.csr.unibo.it/FVCOnGoing

[ISO19794] defines face poses close to zero degrees, in practice this does not hold true in all cases. The main challenges for successful morphing strategies thus are:

- Minimizing image artefacts generated by morphing to limit as much as possible human intervention.
- Establishing the best morphing factor (also known as α factor) to maximize the probability to fool the human officer during enrolment and the automatic recognition system at the border.
- Developing automatic methods for aligning critical areas such as nostrils and irises.
- Aligning the eyes between the two images to reduce visual artefacts in the face.
- Avoiding algorithms that widely introduce visual artefacts such as shadows etc.

The University of Bologna is currently extending the existing FVConGoing platform[5] with new benchmarking services for differential morph detection and thereby developing the Bologna-Online-Evaluation-Platform (BOEP) platform. The new SOTAMD dataset will be stored in a highly protected environment, not exposed directly to the internet. It is suggested that the iMARS mixed-quality dataset is added, as soon as it becomes available. Further, it is suggested that BOEP will be extended in order to benchmark also non-reference MAD mechanisms. The datasets will be accessible through BOEP and provide open access benchmark tests. Thus, with BOEP, Frontex and the national border control agencies will be able to evaluate if the MAD State-of-the Art meets the operational requirements. The technical interfaces are by design equivalent to the benchmark portal of the NIST Face Recognition Vendor Test (FRVT) MORPH Competition [NISTFRVT]. However the functionality of BOEP will exceed FRVT-MORPH. The BOEP will provide a dedicated benchmark environment that can allow different tests on different selection pre-conditions (e.g., lookalikes, random selected pairs, or skin-color similarities, etc.). The quality and method of the morphing technique influences the ability of detection. In order to cover a broad spectrum of attacks the generated database has to cover a broad range of morphing strategies.

Hosting the data on the BOEP will enable researchers and operators to submit algorithms for online evaluation, without the need that confidential data has to travel to an evaluation lab. This will hence allow:

- Testing on lookalikes, same demographic subgroup, versus random selected pairs
- Testing with variation of morph algorithms, alpha values, and resolution.
- Testing with various facial image quality

The data that will be available on the BOEP will constitute a scenario test. In order to evaluate the impact of operational deployment, border control agencies shall be motivated by Frontex to contribute any real case data for differential MAD cases or forensic MAD cases to the University of Bologna, such that an additional benchmark with real case morph images can be

offered. Specifically, for the differential MAD testing case, the collected data on the BOEP has no time gap between the passport image and the trusted image from the ABC gate. Thus, image pairs from real cases, where there might be a time gap of up to ten years, is of great interest.

Further morphing attack data as suggested by the iMARS consortium is needed.

Need to standardize Testing of MAD Solutions

When analyzing the vulnerability of face recognition systems to morphing attacks, the need to augment the metrics for evaluation of presentation attacks is obvious. The *Impostor Attack Presentation Match Rate* (IAPMR) [ISO30107] introduced in ISO/IEC 30107-3, represents a standardized metric for evaluating the impact of a presentation attack. Contrary to PAD evaluations, for a morphing attack all individuals contributing to the morph want to be successfully authenticated against the morphed facial image. This scenario cannot be evaluated using the IAMPR and thus motivated the introduction of new evaluation metrics [Sch17]. A Morphing Attack (MA) is only successful if all involved subjects have been successfully verified. Motivated by ISO/IEC 30107-3 [4], the *Mated-Morph-Presentation-Match-Rate* (MMPMR) is proposed in [Sch17] to evaluate the effect of a MAs on the overall system. This metric that has been established in the academic literature and should be further developed in an international ISO/IEC standard. The iMARS consortium suggested to anchor the MAD evaluation methodology in the by ISO/IEC 30107 multipart standard.

Standardization of measures to define the threats and the efficacy of countermeasures in a quantifiable and objective way has taken first steps. The challenge is to:

- Find consensus in the MAD research community and formulate a narrow set of relevant metrics.
- Standardise metrics to evaluate the performance of MAD methods and the vulnerability of biometric recognition systems to morphing attacks.

The iMARS consortium will initiate the ISO/IEC standardisation process of metrics to evaluate the performance of MAD methods and measures for the vulnerability of biometric recognition systems to morphing attacks. Identified methods will be validated with live images acquired on operational eGates.

An international standard for MAD testing as suggested by the iMARS consortium is needed. Border control agencies of EU Member State shall be motivated by Frontex to participate in this standardisation process.

Need to develop Face Image Quality Metrics

Assessment of face image quality is vital to capture samples that are sufficiently good in term of illumination, sharpness, or pose, such that the probe sample can verify an individual's identity accurately and reliably. The framework for biometric sample quality is well described in ISO/IEC 29794-1:2016 [ISO29794]. The essential definition is that a quality measure shall represent the quality of the source (e.g. the skin for a fingerprint recognition system) but also the fidelity of

the sensor (i.e., is the image signal representing the source?). An expression of a quality score must be in the range of 0 to 100 (poor to best quality). Moreover, a quality score must be capable of predicting recognition performance (i.e. a sample with a low-quality score will likely not reach a high similarity score in a later recognition). Such correlation of quality scores and low false–non-match rate can be observed with the Error-versus-reject-curve (ERC).

The iMARS consortium suggests the development of an automatic face image quality assessment software, which can predict recognition accuracy and provide actionable feedback to the data capture subject and/or to the operational personnel. Such software will serve the needs of passport enrolment agencies but also European agencies (e.g. FRONTEX and EU-LISA) that must control quality of data in their databases. The resulting software prototype to automate image quality assessment will form the basis for a technical contribution to an international standard ISO/IEC 29794-5 (as revision of the previously existing technical report). This work will be the equivalent to the successful NFIQ2.0 metric for fingerprint images.

Once predictive face quality software is available, MAD evaluation can be adapted to the three relevant scenarios (i.e., ID Document issuance, border control, and forensic investigation) that should be used to evaluate the MAD solutions. The iMARS consortium suggested to adjust to the use cases corresponding to the different dimensions of quality: Image quality will differentiate use cases involving, on the one hand, always high quality, versus, on the other hand, a high-quality image during enrolment and a low-quality image at the border gate (due to poor illumination or pose variations caused by distracting factors at the gate). Once a predictive metric is available, the impact of face image quality on biometric recognition and MAD performance can be evaluated, e.g. the correlation of quality of acquired face images for an image pair based morphing attack detector can be measured. Moreover, the impact of face image quality on biometric recognition performance as well as morphing attack detection will be benchmarked.

Need to train operating Border Officers and Communication Personnel

The iMARS consortium is committed to deliver sustainable solutions for Border control operators. In interaction with Frontex and governmental agencies of European member states, the iMARS consortium suggests to address the usability/ergonomics requirements defined by the operators. Furthermore, the iMARS consortium suggests to develop best practices and a training curriculum for improving the officers' skills on manipulated/morphed image and document fraud detection while also respecting fundamental rights: training will lead to better mutual understanding and knowledge transfer. The iMARS consortium will design a training curriculum to reinforce end-users' skills on MAD solutions and to transmit professional expertise gained during the project. This curriculum will allow increasing their detection of manipulated images or document fraud, will enhance their skills, but also show them that the tools will not replace, but complement, their expertise. The iMARS consortium will team up with Frontex and their training procedures.

Training of operators' communication personnel is also considered to mitigate public excitement and explain attack resolving solutions against morphing attacks, once the threat is reported in the media.

Impact and Conclusion

If the needs elaborated in this contribution are implemented, then a strong impact can be expected for the security of European borders. The SOTAMD and iMARS consortium will facilitate reproducible research by employing a unified platform to allow standardised online evaluation. Application oriented specific benchmarks will be defined and developed. This will lead to a situation where error rates of differential MAD approach are reduced considerably – thus the chance that a criminal can fool an ABC system while keeping the amount of false morphing warnings will be quite limited.

The efficiency of the iMARS MAD solutions will entice practitioners to use them: the reduction of false alarm rate (Bona Fide Presentation Classification Error Rate - BPCER) of morphing attack detectors will allow the technology to be deployed at border with neglectable interference of the passenger flow and the outcomes can be successfully integrated into the existing ePassport life-cycle in a reasonable amount of time. When MAD solutions are deployed as first-line control solution, they will not be a substitute for a human expertise, which will always require that results need to be confirmed by an empowered agent in the second line in case of alarm or doubt.

The standardisation activities performed during the project will also contribute to the reproducibility of the tests performed and are of global benefit for all ICAO members. The standardisation project on face image quality ISO/IEC 29794-5 will be initiated and supported.

Within the iMARS project, new strategies to prevent cross-border crime will be proposed and implemented. Findings of the project will be consolidated in form of guidelines, which could be used by civil servants to take transparent and reliable decisions. For example, in case of doubt, biometric verification at the border shall be done with a second biometric identifier, such as fingerprint or iris reference images.

References

[Fer14] M. Ferrara, A. Franco, and D. Maltoni, The Magic Passport , in proceedings International Joint Conference on Biometrics (IJCB), Clearwater, Florida, USA, pp.1-7, October 2014..

[Sch19] U. Scherhag, C. Rathgeb, J. Merkle, R. Breithaupt, and C. Busch. Face recognition systems under morphing attacks: A survey. IEEE Access, 7:23012–23026, 2019.

[ISO19794] ISO/IEC JTC1 SC37 Biometrics, ISO/IEC 19794-5:2005, Biometric data interchange format - Part 5: Face image data, 2005.

[ISO29794] ISO/IEC JTC1 SC37 Biometrics, ISO/IEC 29794-1:2016, Biometric sample quality – Part 1: Framework, 2016.

[ISO39794] ISO/IEC JTC1 SC37 Biometrics, ISO/IEC 39794-1:2019, Extensible biometric data interchange format – Part 1: Framework, 2019.

[ISO30107] ISO/IEC JTC1 SC37 Biometrics, ISO/IEC 30107-3:2017, Biometric presentation attack detection – Part 3: Testing and Reporting, 2017.

[ICAO9303] International Civil Aviation Organization, ICAO Doc 9303, Ma-chine Readable Travel Documents - Part 9: Deployment of Bio-metric Identification and Electronic Storage of Data in MRTDs (7th edition), 2015.

[NISTFRVT] U.S. NIST Face Recognition Vendor Test – Morph, https://www.nist.gov/programs-projects/frvt-morph

[Sch17] U. Scherhag, A. Nautsch, C. Rathgeb, M. Gomez-Barrero, R. N. J. Veldhuis, L. Spreeuwers, M. Schils, D. Maltoni, P. Grother, S. Marcel, R. Breithaupt und R. Ramachandra, „Biometric Systems under Morphing Attacks: Assessment of Morphing Techniques and Vulnerability Reporting," in Proceedings of the 2017 International Conference of the Biometrics Special Interest Group (BIOSIG), 2017.

[Ram16] R. Ramachandra, K. B. Raja und C. Busch, „Detecting morphed face images," in Proceedings of the 8th International Conference on Biometrics Theory, Applications and Systems (BTAS), 2016.

Industry exhibitors

www.thalesgroup.com

www.dermalog.com

www.idemia.com

www.sita.aero

www.secunet.com

www.griaule.com

www.integratedbiometrics.com

www.ingroupe.com

www.biorugged.com

www.ait.ac.at

www.t3k-forensics.com

www.bpiservices.eu

www.readid.com

www.jenetric.com

www.laxtongroup.com

www.secomitalia.com

www.mobai.bio

www.globalsat.com

www.thermofisher.com

www.iecisa.com

Annex

Message from the Finnish Presidency of the Council of European Union[1]

Olli-Poika Parviainen ·State Secretary, Ministry of Interior, Finland

Dear Colleagues,

Identity verification is at the heart of authorities' activities. We must be able to identify individuals correctly, and they must not have an opportunity to circumvent control by using different identities.

Identification is increasingly based on technological solutions rather than paper and ink. This puts pressure on the technology used by authorities and in travel documents. Many parties have a malicious intent to break the integrity of the technology.

For example, facilitators of illegal immigration fabricate manipulated data to exploit hacked technology for criminal activity, such as human smuggling. Criminal networks or terrorists gain advantages from concealing identities and travel histories. A false identity is a way to avoid alerts. State actors can exploit the vulnerability of the system, for example for espionage, sabotage or other hybrid interference. Instead of a single travel document, they may target whole systems, services, servers or databases. Increasingly even at European scale.

Societal development and technological advances are making us increasingly dependent on technology. There are virtually no backup systems. In the future, for example, if European information and backup systems fail, it will be practically impossible to return to manual activities on a large scale and on a long-term basis.

However, technology is not primarily a threat, but an opportunity. It facilitates more reliable and streamlined services that are easier to manage. It results in smoother transport and improved legal certainty for authorities.

Designing and using systems securely improves the efficiency of authorities. Artificial intelligence and machine learning, for example, are important new tools that will become mainstream in government activities and citizen services in only a couple of years.

Ensuring data protection and data security is the cornerstone of the legitimacy of the system. We are pleased to see that a great deal of attention is paid to this all across the EU.

We must also ensure that legal operators are always one step ahead of criminals and other hostile parties. This conference is a good example: gathering in Warsaw to share experiences, solve challenges together, disseminate best practices and build bridges between Member States, administrative branches, EU institutions and agencies, research and industry.

Money, staff and time are always in limited supply. Smooth cross-border traffic keeps society's wheels turning. We need to create efficient and cost-effective methods, for example by harnessing automation to ensure that society remains safe and secure.

1 Due to unforeseen circumstances the originally planned keynote speech on behalf of the Finnish Presidency was cancelled. The scheduled speaker, Mr. Olli-Poika Parviainen, therefore shared with the participants of the ICBB2019 this message on behalf of the Presidency.

National and EU security are challenged by security threats related to technological systems, which the authorities must be able to prevent. Hybrid and cyber threats in particular have become a key concern when it comes to system vulnerability. These threats must be taken into account in the design of machinery and equipment and in the coding of software.

On the other hand, the development of technology and artificial intelligence has created increasing opportunities for improving safety and security and developing tools for authorities. In particular, AI, machine learning and robotics applications can be used to replace routine activities and free up resources for other tasks.

The objective of the Finland's Presidency is highlighting the further development of effective external border control. One of the elements is the completion of the work on the regulation on the European Border and Coast Guard Agency, which gives Frontex a stronger mandate.

As for the information system regulations, the Finnish Presidency is also ready to finalise negotiations on open regulations with the European Parliament in order to avoid delays to the implementation of the systems.

Finland's Presidency coincides with the start of the implementation phase. Finland will promote the implementation of measures already agreed in various Council configurations. The EU has agreed on new systems (EES, ETIAS, ECRIS-TCN), on improving old systems (VIS, SIS, Eurodac) and on the interoperability of the systems.

The importance of biometric technologies is increasing. The deployment phase will be challenging, because both the technology and the processes are in constant flux.

During Finland's Presidency, we would like to draw particular attention to the implementation of the Entry/Exit System (EES), as its implementation and enforcement at all European border crossing points will be a huge undertaking.

Data security, data protection and privacy protection must be considered. Biometric identifiers are particularly sensitive information. Processing such information, therefore, cannot be based solely on the perspective or the needs of security authorities.

The future looks challenging as the globalisation will further increase the movement of people, threats will become increasingly cross-border and decreasing technology prices will increase the opportunities of criminals to act. Still, the resources of authorities will not increase.

A comprehensive approach is needed. Preventing the use of morphing and other image manipulations requires effective issuing and inspection processes, easy-to-use and secure technology and skilled personnel. Development work cannot be carried out in silos, and, without conferences such as this one, counter-measures will become fragmented. We have many challenges and open questions in the implementation of the new EU systems.

The foundation of society's core functions will erode without reliable methods of identity verification. Let us hope that this event will spur a wide-ranging exchange of ideas and bring a new understanding of the challenges we face.

I wish you all a successful conference.

On behalf of the Finnish Presidency,

9 October 2019

Olli-Poika Parviainen
State Secretary
Ministry of Interior

FRONTEX
EUROPEAN BORDER AND COAST GUARD AGENCY

Plac Europejski 6
00-844 Warsaw, Poland

T +48 22 205 95 00
F +48 22 205 95 01

frontex@frontex.europa.eu
www.frontex.europa.eu

Print version:
TT-04-20-167-EN-C

PDF:
TT-04-20-167-EN-N

FPI20.0019

Publications Office
of the European Union