Morphing Attack Detection Overview

Christoph Busch, Ulrich Scherhag, Christian Rathgeb, Kiran Raja, Raghu Ramachandra, Marta Gomez-Barrero, Daniel Fischer, Sergey Isadskiy

copy of slides available at:
https://www.christoph-busch.de/projects-mad.html

NIST-IFPC-2018, November 28, 2018
Overview

Agenda

- Introduction - Problem description
- Morphing Attack Detection - Scenarios and Methods
- Status: Face Morphing Attack Detection
- Future - what needs to be done?
- Conclusion
Problem Description
Face Morphing

The morphing attack was named and classified as vulnerability of a biometric system in Clause 8.3.8.1 of ISO/IEC FDIS 19792:

“… Examples of abnormal characteristics could include those with unusually large or small numbers of features. Such characteristics may not be representative of any human biometric characteristic but could be synthesised and copied to an artefact. Alternatively 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”
History - 2014

Integrated Project FIDELITY

- Fast and trustworthy Identity Delivery and check with ePassports leveraging Traveler privacy
- 4 years project (2012-2016)
  - European 7th Framework Programme
- Objectives:
  - To improve the ePassport issuing process
    - Security of birth certificates and other evidence of identity
    - Quality of biometric data in the chip
    - One individual one passport (duplicate enrolment check)
  - To demonstrate solutions that enable faster and more secure and efficient real-time authentication of individuals at border crossing
  - To protect privacy of the travel document holders with a privacy-by-design approach.

http://www.fidelity-project.eu/
Problem: Morphing Attacks

Verification against morphed facial images

- Probe sample of A
- Enrolment sample of A
- Enrolment morph M
- Enrolment sample of C
- Probe sample of C

Similarity:
- 0.87
- 0.65
- 0.59
- 0.94
- 0.03
Morphing Attack Detection (MAD)

Scenarios and Methods
Real world scenarios

- **No-reference morph detection**
  - One single facial image is analysed (e.g. in the passport application office)

Morphing Attack Detection Scenarios

Real world scenarios

• **No-reference morph detection**
  - One single facial image is analysed (e.g. in the passport application office)

• **Differential morph detection**
  - A pair of images is analysed - and one is a trusted Bona Fide image
  - Biometric verification (e.g. at the border)

Morphing Attack Detection (MAD) with texture analysis

- Image descriptors as hand-crafted features

**Face Pre-processing and Feature Extraction**

MAD with image descriptor

- Local Binary Pattern (LBP)
Face Pre-processing and Feature Extraction

MAD with image descriptor

- Binarized Statistical Image Features (BSIF)
Face Pre-processing and Feature Extraction

MAD with image descriptor

- Sharpness

\[ \delta x \]
\[ \delta y \]

Mean

Sharpness features
MAD with image descriptor

- Histogram of Gradients (HOG)
MAD with image descriptor

- Scale Invariant Feature Transform (SIFT)
- Speeded up Robust Features (SURF)
Face Pre-processing and Feature Extraction

MAD with image descriptor / forensic approach

- Photo Response Non-Uniformity (PRNU)

![Morph vs Bona Fide](image-url)
Morphing Attack Detection (MAD) with texture analysis

- Image descriptors as Deep features
Face Pre-processing and Feature Extraction

MAD with deep learning

- **Deep Features**
  - pre-trained Convolutional Neural Network (CNN)
  - OpenFace

No-Reference Morph Detection

MAD with deep learning

- **Feature level fusion** of Deep CNNs

MAD Evaluation Methodology
Face Morphing Attack evaluations are complex

- Evaluations must consider a dedicated methodology
  - see the following presentation by Marta Gomez-Barrero [SNR17]

Evaluations must consider many parameters

\[ \text{result} = f(\text{dataset-training, dataset-testing, morphing-attack, landmark-detector, feature-extractor, classifier, scenario (no-reference vs. differential), post-processing, printer, scanner}) \]

Quality of the passport image under investigation
- hopefully ICAO 9303 compliant
- ISO/IEC 39794-5 compliant
MAD Evaluation Methodology

Evaluations must consider many parameters

- For a **differential** MAD evaluation

\[ \text{result} = f(\text{dataset-training, dataset-testing, morphing-attack, landmark-detector, feature-extractor, classifier, scenario (no-reference vs. differential), post-processing, printer, scanner}) \]

- Quality of the **passport image** under investigation and quality of the **trusted probe image**

In our evaluation we use

- The FERET dataset for training
  

- The FRGCv2 dataset for testing
  

- Both data sets were filtered to reach ICAO compliance
Evaluations must consider many parameters

- Dataset preparation requires **pre-processing**

\[
\text{result} = f(\text{dataset-training, dataset-testing, morphing-attack, landmark-detector, feature-extractor, classifier, scenario (no-reference vs. differential), post-processing, printer, scanner})
\]

Facial images are **cropped and aligned** to a normalized **size**

Resulting images are

- **cropped** to 320x320 pixel
- aligned according to Dlib landmarks, such that eyes are at **identical coordinates**
MAD Evaluation Methodology

Evaluations must consider many parameters

- Morphing may require manual interaction (not desired)

\[ \text{result} = f(\text{dataset-training, dataset-testing, morphing-attack, landmark-detector, feature-extractor, classifier, scenario (no-reference vs. differential), post-processing, printer, scanner}) \]

Automated face morphing tools may introduce artifacts

In our evaluation we use

- Dlib / OpenCV
- FaceMorpher
Evaluations must consider many parameters

- From machine learning tools we select a classifier

\[ \text{result} = f(\text{dataset-training, dataset-testing, morphing-attack, landmark-detector, feature-extractor, classifier, scenario (no-reference vs. differential), post-processing, printer, scanner}) \]

Simplicity and generalisation capability are desired properties

In our evaluation we use

- Support Vector Machine (SVM)
  - with radial basis function as kernel
- AdaBoost
  - with 200 estimates and a decision stump
MAD Evaluation Methodology

Evaluations must consider many parameters
• Postprocessing might conceal morphing effects (e.g. smoothing)

\[ \text{result} = f(\text{dataset-training, dataset-testing, morphing-attack, landmark-detector, feature-extractor, classifier, scenario (no-reference vs. differential), post-processing, printer, scanner}) \]

...smoothing and other effects might be compensated by the attacker...

In our evaluation we show results for
• Sharpening
Results
Generalising evaluation - **differential** scenario

- **Differential morph detection**
  - A pair of images is analysed - and one is a trusted Bona Fide image
  - Biometric verification (e.g. at the border)
Generalising evaluation - **differential** scenario

- training on FERET, testing on FRGCv2
  - hand-crafted feature extractors perform well
  - no post-processing of morph images

<table>
<thead>
<tr>
<th>MAD-method</th>
<th>Classifier</th>
<th>Morphing Algorithm (Training)</th>
<th>Morphing Algorithm (Test)</th>
<th>D-EER</th>
</tr>
</thead>
<tbody>
<tr>
<td>LBP</td>
<td>SVM</td>
<td>Dlib und OpenCV</td>
<td>Dlib und OpenCV</td>
<td>0.0228</td>
</tr>
<tr>
<td>LBP (4x4 cells)</td>
<td>SVM</td>
<td>Dlib und OpenCV</td>
<td>Dlib und OpenCV</td>
<td>0.0997</td>
</tr>
<tr>
<td>LBP</td>
<td>AdaBoost</td>
<td>Dlib und OpenCV</td>
<td>Dlib und OpenCV</td>
<td>0.0645</td>
</tr>
<tr>
<td>LBP (4x4 cells)</td>
<td>AdaBoost</td>
<td>Dlib und OpenCV</td>
<td>Dlib und OpenCV</td>
<td>0.0471</td>
</tr>
<tr>
<td>BSIF</td>
<td>SVM</td>
<td>Dlib und OpenCV</td>
<td>Dlib und OpenCV</td>
<td>0.0775</td>
</tr>
<tr>
<td>BSIF (4x4 cells)</td>
<td>SVM</td>
<td>Dlib und OpenCV</td>
<td>Dlib und OpenCV</td>
<td>0.0656</td>
</tr>
<tr>
<td>BSIF</td>
<td>AdaBoost</td>
<td>Dlib und OpenCV</td>
<td>Dlib und OpenCV</td>
<td>0.0695</td>
</tr>
<tr>
<td>BSIF (4x4 cells)</td>
<td>AdaBoost</td>
<td>Dlib und OpenCV</td>
<td>Dlib und OpenCV</td>
<td>0.0742</td>
</tr>
<tr>
<td>OpenFace</td>
<td>SVM</td>
<td>Dlib und OpenCV</td>
<td>Dlib und OpenCV</td>
<td>0.1253</td>
</tr>
<tr>
<td>OpenFace</td>
<td>AdaBoost</td>
<td>Dlib und OpenCV</td>
<td>Dlib und OpenCV</td>
<td>0.1373</td>
</tr>
<tr>
<td>FaceNet</td>
<td>SVM</td>
<td>Dlib und OpenCV</td>
<td>Dlib und OpenCV</td>
<td>0.1943</td>
</tr>
<tr>
<td>FaceNet</td>
<td>AdaBoost</td>
<td>Dlib und OpenCV</td>
<td>Dlib und OpenCV</td>
<td>0.1745</td>
</tr>
<tr>
<td>LBP</td>
<td>SVM</td>
<td>FaceMorpher</td>
<td>FaceMorpher</td>
<td>0.0025</td>
</tr>
<tr>
<td>LBP (4x4 cells)</td>
<td>SVM</td>
<td>FaceMorpher</td>
<td>FaceMorpher</td>
<td>0.0023</td>
</tr>
<tr>
<td>LBP</td>
<td>AdaBoost</td>
<td>FaceMorpher</td>
<td>FaceMorpher</td>
<td>0.0453</td>
</tr>
<tr>
<td>LBP (4x4 cells)</td>
<td>AdaBoost</td>
<td>FaceMorpher</td>
<td>FaceMorpher</td>
<td>0.0000</td>
</tr>
<tr>
<td>BSIF</td>
<td>SVM</td>
<td>FaceMorpher</td>
<td>FaceMorpher</td>
<td>0.0253</td>
</tr>
<tr>
<td>BSIF (4x4 cells)</td>
<td>SVM</td>
<td>FaceMorpher</td>
<td>FaceMorpher</td>
<td>0.0085</td>
</tr>
<tr>
<td>BSIF</td>
<td>AdaBoost</td>
<td>FaceMorpher</td>
<td>FaceMorpher</td>
<td>0.0126</td>
</tr>
<tr>
<td>BSIF (4x4 cells)</td>
<td>AdaBoost</td>
<td>FaceMorpher</td>
<td>FaceMorpher</td>
<td>0.0695</td>
</tr>
<tr>
<td>OpenFace</td>
<td>SVM</td>
<td>FaceMorpher</td>
<td>FaceMorpher</td>
<td>0.1432</td>
</tr>
<tr>
<td>OpenFace</td>
<td>AdaBoost</td>
<td>FaceMorpher</td>
<td>FaceMorpher</td>
<td>0.1404</td>
</tr>
<tr>
<td>FaceNet</td>
<td>SVM</td>
<td>FaceMorpher</td>
<td>FaceMorpher</td>
<td>0.2054</td>
</tr>
<tr>
<td>FaceNet</td>
<td>AdaBoost</td>
<td>FaceMorpher</td>
<td>FaceMorpher</td>
<td>0.1745</td>
</tr>
</tbody>
</table>
MAD Evaluation

Generalising evaluation - differential scenario

- training on FERET, testing on FRGCv2
- now we focus on LBP only
  - and again no post-processing of morph images

<table>
<thead>
<tr>
<th>MAD-method</th>
<th>Classifier</th>
<th>Morphing Algorithm (Training)</th>
<th>Morphing Algorithm (Test)</th>
<th>D-EER</th>
</tr>
</thead>
<tbody>
<tr>
<td>LBP</td>
<td>SVM</td>
<td>Dlib und OpenCV</td>
<td>FaceMorpher</td>
<td>0.0153</td>
</tr>
<tr>
<td>LBP</td>
<td>AdaBoost</td>
<td>Dlib und OpenCV</td>
<td>FaceMorpher</td>
<td>0.0471</td>
</tr>
<tr>
<td>LBP</td>
<td>SVM</td>
<td>FaceMorpher</td>
<td>Dlib und OpenCV</td>
<td>0.0251</td>
</tr>
<tr>
<td>LBP</td>
<td>AdaBoost</td>
<td>FaceMorpher</td>
<td>Dlib und OpenCV</td>
<td>0.1369</td>
</tr>
</tbody>
</table>

We reach in the best case
- approx 1 % EER (between APCER and BPCER)
MAD Evaluation

Generalising evaluation - differential scenario

- training on FERET, testing on FRGCv2
- now we focus on LBP only
- post-processing of morph images with the sharpening operator

<table>
<thead>
<tr>
<th>MAD-method</th>
<th>Classifier</th>
<th>Morphing Algorithm (Training)</th>
<th>Morphing Algorithm (Test)</th>
<th>D-EER</th>
</tr>
</thead>
<tbody>
<tr>
<td>LBP</td>
<td>SVM</td>
<td>Dlib und OpenCV</td>
<td>FaceMorpher</td>
<td>0.0108</td>
</tr>
<tr>
<td>LBP</td>
<td>AdaBoost</td>
<td>Dlib und OpenCV</td>
<td>FaceMorpher</td>
<td>0.0414</td>
</tr>
<tr>
<td>LBP</td>
<td>SVM</td>
<td>FaceMorpher</td>
<td>Dlib und OpenCV</td>
<td>0.0417</td>
</tr>
<tr>
<td>LBP</td>
<td>AdaBoost</td>
<td>FaceMorpher</td>
<td>Dlib und OpenCV</td>
<td>0.1289</td>
</tr>
</tbody>
</table>

We still reach in the best case

- approx 1 % EER (between APCER and BPCER)
Future - What needs to be done?
MAD Evaluations on Digital Images

First investigations on morphing attack detection
- are on a **small** dataset
- Addressing only **digital** application process
  (applicable for New Zealand, Estonia, Ireland)

The upcoming evaluations
- NIST-FRVT-MORPH evaluation
- SOTAMD evaluation

will provide valuable insights
MAD Evaluations on Digital Images

Our submissions to NIST-FRVT-MORPH / SOTAMD:

- LBP-MAD proposed in [RRB16], [SRB18a] and [SRB18b]
- PRNU-MAD proposed in [DSRUB18a] and [DSRUB18b]


MAD Evaluations on Single Digital Images

Our submission to NIST-FRVT-MORPH:

- Classifiers for single image analysis
- No-reference morph detection
  - One single facial image is analysed (e.g. in the passport application office)
MAD Evaluations on Single Digital Images

Our submission to NIST-FRVT-MORPH:

• LBP-MAD classifier for single image analysis
  ‣ no-reference scenario

• feature vector
  ‣ 4 x 4 histograms, 256 values each
  ‣ Normalized histograms

• trained SVM on
  ‣ 1000 original images from FERET and FRGCv2
  ‣ 1000 morphs from FERET and FRGCv2
    - 2 morphing algorithms
    - 4 different post processing methods

• tested on
  ‣ 1000 original images from FERET and FRGC
  ‣ 1000 morphs from FERET and FRGC

We reach BPCER = 5.25% @ APCER = 5.80%
MAD Evaluations on Single Digital Images

Our submission to NIST-FRVT-MORPH:

• PRNU-MAD classifier for single image analysis
  ‣ no-reference scenario

• feature vector
  ‣ Noise residuals

• trained SVM on
  ‣ 1000 original images from FERET and FRGCv2
  ‣ 1000 morphs from FERET and FRGCv2
    - 2 morphing algorithms
    - 4 different post processing methods

• tested on
  ‣ 1000 original images from FERET and FRGC
  ‣ 1000 morphs from FERET and FRGC

We reach BPCER = 5.6%  @ APCER = 4.6%
What needs to be Done?

Evaluations must consider the printing process

- There are numerous parameters to explore for this

\[
\text{result} = f (\text{dataset-training, dataset-testing, morphing-attack, landmark-detector, feature-extractor, classifier, scenario (no-reference vs. differential), post-processing, printer, scanner})
\]

Printer / Scanner of choice

Resolution (spatial sampling rate)
What needs to be Done?

Multiple dimensions to explore:

- Large scale datasets evaluation in NIST FRVT MORPH
- **Generalisation** on public datasets
  ‣ FERET, FRGCv2, FEI, ARface
- Morphing mechanism
  ‣ Fantamorph, OpenCV, Splicing, GIMP, …
- **Number** of contributing **subjects** (broker model)
- The most effective **alpha-factor** (50:50 or 20:80)
- Random or **lookalike** morphs
  ‣ Same gender, same skin-color as selection criteria
- Digital samples versus digital-analog-digital transition
References

Publications available https://www.christoph-busch.de/projects-mad.html

This research is kindly sponsored by

- Federal Office for Information Security (BSI)
  - FACETRUST

- Center for Research in Security and Privacy:
  - LOEWE/BMBF CRISP [http://www.crisp-da.de/]

- Norwegian Research Council
  - IKTPLUSS SWAN [http://nislab.no/biometrics_lab/swan]

- Norwegian University of Science and Technology
Prof. Dr. Christoph Busch

Norwegian University of Science and Technology
Department of Information Security and Communication Technology
Teknologiveien 22
2802 Gjøvik, Norway
Email: christoph.busch@ntnu.no
Phone: +47-611-35-194