Mobile Face Recognition Systems: Exploring Presentation Attack Vulnerability and Usability

Heinz Hofbauer • Luca Debias • Andreas Uhl
University of Salzburg, Department of Computer Sciences
Jakob Haringer Str. 2, 5020 Salzburg, Austria

UNIVERSITÄT SALZBURG
WAVELAB
Multimedia Signal Processing and Security Lab

Abstract

We have evaluated face recognition software to be used with hand held devices (smartphones).
We contrast the robustness against presentation attacks with the systems usability during regular use, and highlight where currently state of commercial of the shelf systems (COTS) stand in that regard.
We will look at the results specifically under the tradeoff between acceptance, linked with usability, and security, which usually negatively impacts usability.

Main Results

- Biometric verification seems to work well.
- Liveness detection is far too strict.
  - Glasses break liveness detection
  - Sunlight (or other bright light) breaks liveness detection.
- The tradeoff between usability and security is currently very bad.
- More complicated/expensive attacks, like masks, did worse than replay attacks.

Introduction

Basic information

- We were tasked by a company with evaluating the usability and security of face recognition systems.
- System is used to unlock a smartphone via selfie-based face recognition.
- Liveness detection on client side.
- Matching is done server side.

Constraints, shortcomings

- Testing time limited, due to licensing issues.
- Limited number of users/attempts.
- Should be seen as a showcase for obvious problems which happen outside a "lab setup".

What do we want to know?

- How secure are these systems?
  That is, how hard is it for an adversary to unlock the phone.
- How usable are these systems?
  That is, how easy is it for a legitimate user to unlock the phone.

What are we doing? We will look at the security of the two systems under test PassiveSys and ActiveSys, with the goal of unlocking the device with minimal fuss on the part of the user.

Liveness detection modes of the different tools

- PassiveSys
  - *video* seems to take a video
  - *lessvid* less stringent version of *video*
  - *image* simply takes a picture
- ActiveSys
  - *blink*: user has to keep still and blink on cue
  - *arrow*: requires turning the head to steer an arrow along a line to a target, when the arrow and target align the user has to blink

Usability and Baseline

Results are split between liveness detection test (LD) and verification results (Match). The presence of glasses in the probe (Pr.) and gallery (Gal.) images is given as well.

Baseline for PassiveSys for modes video, lessvid, image.

video

Pr.	Gal.	LD	Match
yes	yes	0/20	0/20
no	no	13/20	13/20
no	yes	10/20	10/20
yes	no	0/20	0/20

lessvid

Pr.	Gal.	LD	Match
yes	yes	4/20	4/20
no	no	18/20	18/20
no	yes	12/20	12/20
yes	no	9/20	9/20

image

Pr.	Gal.	LD	Match
yes	yes	6/20	6/20
no	no	20/20	20/20
no	yes	18/20	18/20
yes	no	19/20	19/20

Baseline for ActiveSys for modes blink and arrow.

blink

Pr.	Gal.	LD	Match
yes	yes	20/20	20/20
no	no	16/20	16/20
no	yes	12/20	12/20
yes	no	12/20	12/20

arrow

Pr.	Gal.	LD	Match
yes	yes	18/20	18/20
no	no	20/20	20/20
no	yes	20/20	20/20
yes	no	20/20	20/20

- The presence of glasses increases the error rate of the liveness detection.
- Matching always worked when liveness detection was passed.
- *video* overall rejected almost 72% of all attempts
- *arrow* seems to reject less than *blink*, even though the task is more complicated.
- Failure of modes which took several seconds, e.g. *arrow*, became frustrating very fast.

Usability and Baseline Outdoors

- Suspicion: Failure for images with glasses was due to reflection on the glasses.
- We performed another test in natural sunlight, during a bright day. Facing was either *towards* the sun or *away* from it.

PassiveSys split for modes. video and lessvid

Facing	LD	Match
towards	0/20	0/20
away	0/20	0/20

image

Facing	LD	Match
towards	10/20	10/20
away	20/20	20/20

ActiveSys split for modes. blink

Facing	LD	Match
towards	14/20	14/20
away	9/20	9/20

arrow

Facing	LD	Match
towards	10/20	10/20
away	19/20	19/20

- Clear impact of lighting conditions on the liveness detection.
- Verification always worked.

Repeat of the experiment in a controlled environment with:

- Different strengths light (1,3,6).
- Diffused or Spot light.
- Frontal, side or back illuminataion.

per 10 attempts

			LD under Intensities					
			spot			diffuse		
System	Mode	Dir.	1.0	3.0	6.0	1.0	3.0	6.0
PassiveSys	*video*	front	0	0	0	0	0	0
PassiveSys	*lessvid*	front	4	2	3	10	5	5
PassiveSys	*image*	front	8	8	9	10	9	10
ActiveSys	*blink*	front	8	8	6	2	5	3
ActiveSys	*arrow*	front	9	9	10	10	7	8
PassiveSys	*video*	side	0	0	0	0	0	0
PassiveSys	*lessvid*	side	3	4	2	7	5	0
PassiveSys	*image*	side	9	8	9	10	7	8
ActiveSys	*blink*	side	3	4	4	4	4	6
ActiveSys	*arrow*	side	3	5	1	10	5	3
PassiveSys	*video*	back	0	0	0	0	0	0
PassiveSys	*lessvid*	back	5	3	1	4	3	1
PassiveSys	*image*	back	5	9	6	2	0	1
ActiveSys	*blink*	back	6	7	5	7	5	3
ActiveSys	*arrow*	back	9	10	10	10	9	10

- When liveness was detected the user was also correctly verified.
- Spot light worse than diffuse light.
- Frontal light illuminates the subject and has the least influence.
- Side light produces an uneven illumination and worse results.
- Back light should mess up the exposure and leave the face in shadow, it is still better than sidelight.

Presentation Attack: Replay Attacks

- Record an image or video and present that to the device instead of the genuine face.
- *video* mode will no longer be used (not practical given the light test).
- 20 tries per attack were performed

		Successes with Replay		
System	Mode	print	screen	video
PassiveSys	*lessvid*	0	1	0
PassiveSys	*image*	12	17	20
ActiveSys	*blink*	—	—	0
ActiveSys	*arrow*	—	—	5

- When liveness was detected the user was also correctly verified.
- Higher quality/effort reproductions have a higher success rate.
- The *arrow* mode is easier to pass than the *blink* mode, even though more 'user' interaction is required.

Ease of Attack

- Only for PassiveSys/*image* combination with the screen attack.
- How degraded an image can still unlock the device?
- The number of successful attacks out of 10 attempts is given.

Degrad.	Strength		
Type	low	med.	high
Noise	10	10	10
Blur	10	10	6
Resolution	10	10	6/0*

*liveness was detected 6 times, but verification was passed 0 times

Gaussian Blur

low med. strong

Noise

low med. strong

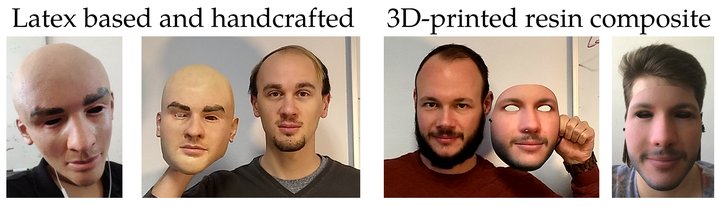
Resolution

low med. strong

Presentation Attack: Masks

We used two attack types based on wearable masks:

Latex based and handcrafted — 3D-printed resin composite



			Latex Mask		Resin Mask	
per 20 attemps						
System	Mode		LD	M	LD	M
PassiveSys	*lessvid*		5	0	10	0
PassiveSys	*image*		10	0	20	0
ActiveSys	*blink*		0	0	10	0
ActiveSys	*arrow*		0	0	16	0

- Better interaction allows bypass of liveness detection.
- Mask reproduction quality fails for matching.
- Higher quality effert *does not* produce better results.
  The latex mask took three times as long to acquire and was four times as expensive as the 3D-printed mask.

Usability versus Security

Presentation attack levels based on time, expertise and equipment.

Threat	Level A	Level B	Level C
Time	short	>3 days	>10 days
Expertise	anyone	practice needed	extensive skill required
Equipment	readily available	requires planning	specialized
Biometric source	readily available	difficult to obtain	difficult to obtain
Example	paper print of image	paper mask or video	3D face reconstruction

Compare threat level, presentation attack success rate and usability per mode and system.

System	Mode	Attack	Threat	Suc. Rate	Usability
PassiveSys	*lessvid*	Image	Level A	5%	
PassiveSys	*lessvid*	Video	Level B	0%	44.2%
PassiveSys	*lessvid*	Mask	Level C	0%	
PassiveSys	*image*	Image	Level A	85%	
PassiveSys	*image*	Video	Level B	100%	77.5%
PassiveSys	*image*	Mask	Level C	0%	
ActiveSys	*blink*	Video	Level B	0%	69.2%
ActiveSys	*blink*	Mask	Level C	0%	
ActiveSys	*arrow*	Video	Level B	25%	89.2%
ActiveSys	*arrow*	Mask	Level C	0%	

- Level C is more costly and doesn't improve over Level B and A.
- PassiveSys system is unusable: Either the usability of a mode is low (*lessvid*) or the success rate of attack is high (*image*).
- ActiveSys is better. Tradeoff between usability and security.