## Biometric Systems Lesson 11bis: Beware of failures



## Maria De Marsico demarsico@di.uniroma1.it





### Strong biometrics?



- Spoofing is a common issue to address for biometric security
  - Single biometric traits both present limitations and can be attacked
- Using biometric equipment might be difficult for average users
- It is possible to mention a number of 'ugly' stories ...



## Apple Touch ID



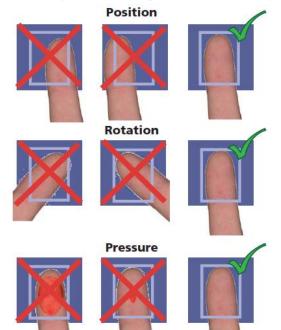
- Apple's iPhone 5s and later, and the iPad Air 2 and later, and the MacBook Pro 2016, all boast a Touch ID fingerprint sensor, which allows to ditch one's passcode in favour of one's fingerprint.
- Earlier models come with first-gen Touch ID; the iPhone 6s and later, and the 2016 MacBook Pro, come with the faster and generally more reliable second-gen Touch ID.
- However ... as with most technology, sometimes things can go wrong ...
- The page at https://www.macworld.co.uk/how-to/iphone/touch-id-fixes-what-do-if-iphone-ipad-fingerprint-scanner-isnt-working-3489429/ takes through the troubleshooting steps ... if Touch ID fingerprint sensor isn't working (Jan 2017)

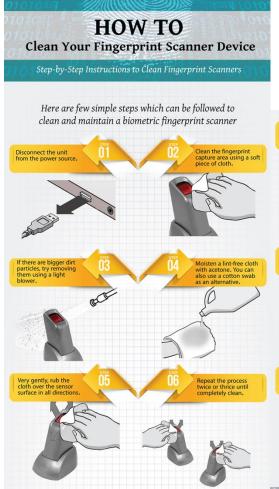


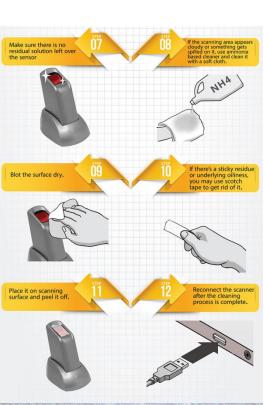
## Suggestions

#### 1010101010101 1010101010 1010101010101

- Among the suggestions:
- Clean the scanner
- Position your finger correctly







Bayometric, 1743 Park Avenue, San Jose, CA 95126

Phone: +1 (408) 940-3955

Email: sales@bayometric.com

-----



## Only a problem for Apple?



- Error: Fingerprint scanner dirty OnePlus Forums
  - https://forums.oneplus.net/threads/error-fingerprint-scanner-dirty.576316/
     (July 2017)
- Has anybody else had issues with "Dirty fingerprint scanner"?
   (Verizon LG G5)
  - https://forum.xda-developers.com/verizon-lg-g5/help/issues-dirty-fingerprint-scanner-t3351795



#### How difficult to break?





- First, the fingerprint of the enrolled user is photographed with 2400 dpi resolution.
- The resulting image is then cleaned up, inverted and laser printed with 1200 dpi onto transparent sheet with a thick toner setting.
- Finally, pink latex milk or white woodglue is smeared into the pattern created by the toner onto the transparent sheet.
- After it cures, the thin latex sheet is lifted from the sheet, breathed on to make it a tiny bit moist and then placed onto the sensor to unlock the phone.
- Note: This process has been used with minor refinements and variations against the vast majority of fingerprint sensors on the market.





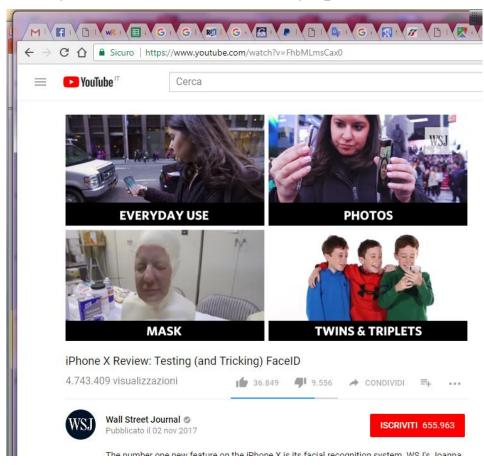
#### **March 2017**

• **WIRED:** almost immediately after Apple announced Face ID, WIRED began scheming to spoof Apple's facial recognition system. They claim to have ultimately spent

thousands of dollars on every material we could imagine to replicate a genuine user's face, down to every dimple and eyebrow hair. All they claim to have learned from their rather expensive experiment is that Face ID is, at the very least, far from trivial to spoof.

#### **November 2, 2017**

• Video on YouTube by Wall Street Journal: it seems unbreakable







#### **November 2, 2017**

• Video on Repubblica RepTech: test with twin sisters









Pubblicato il 09 nov 2017





#### **November 14, 2017**

• WIRED again! Attaullah Malik and Sana Sherwani made that discovery earlier this month, when their fifth-grade son, Ammar Malik, walked into the bedroom of their Staten Island home to admire their new pair of iPhone Xs just after they'd set up Face ID. "There's no way you're getting access to this phone," the older Malik remembers his wife telling her son, in a half-joking show of strictness.

Malik offered to let Ammar look at his phone instead, but the boy picked up his mother's, not knowing which was which. And a split second after he looked at it, the phone unlocked... The parents were shocked.







#### **November 28, 2017**

• Second successful attempt by Bkav ... This time they reproduced in front of witnesses the overall enrollment/testing process



In the new test the mask is printed in 3D like the previous one, made with a mixture of crushed stone and polymers and with two-dimensional images of the eyes (crushed stone seems the secret!)





#### **July 2003**

• CBS news (https://www.cbsnews.com/news/fingerprints-infallible-evidence/): "The FBI has long maintained that fingerprint identification is an exact science that can be used to match prints with 100 percent certainty. But recently, the bureau was forced to admit that three top analysts all made the same mistake when they swore that fingerprint evidence linked Oregon lawyer Brandon Mayfield to the terrorist bombings in Madrid. Mayfield spent two weeks in jail before the FBI admitted its error and offered an apology. Mayfield is one of a small but growing number of Americans who've been jailed on the basis a fingerprint match that turned out to be false. And as Correspondent Lesley Stahl first reported last year, it also happened to Rick Jackson."



#### **July 2003**

**CBS news (continued)**: "On TV, fingerprints are matched automatically, by computer. But that never happens in real-life forensic work ... snip ... a computer, like the one at the Bureau, sorts through a database of 44 million sets of prints, but that merely narrows the search. And then the humans, the fingerprint examiners, make the actual match the way they always have by eye. In fact, it's arduous work involving human judgment. An examiner can spend hours, even days, analyzing a fragment of a fingerprint. Usually, police lift only fragments at a crime scene, and they can be contaminated or distorted. The first thing an examiner compares is what's called ridge flow. ... snip ... not only has there never been a study of the reliability of crimescene fingerprint matching, there are no agreed-upon standards for what constitutes a match. "There's complete disagreement amongst fingerprint examiners themselves as to what they need to see in order to declare a match," ... snip .... In Italy, for example, examiners say they have to see 16 or 17 points of similarity. In Brazil, it's 30; in Sweden, it's seven points; and in Australia, it's 12. And most examiners in the United States, including those at the FBI, don't even use a point system."





#### October 2005

#### ABC news

(http://abcnews.go.com/Technology/DyeHard/story?id=1202813): "A single fingerprint found at the scene of a crime is such powerful evidence that it's almost an automatic conviction. Fingerprints never lie: Juries have been told that for more than a century. But a criminologist at the University of California, Irvine, has documented 22 cases, most involving violent crimes, in which fingerprint evidence turned out to be dead wrong, usually discovered after defendants had served time for crimes they did not commit. The fingerprints didn't lie. But the experts who matched them with a suspect were wrong, and subsequently had to admit it, according to court records analyzed by Simon Cole, assistant professor of criminology, law and society at Irvine. Cole's exhaustive research argues that the "zero error rate" claimed by fingerprint experts needs serious retooling. His findings are published in the current issue of the Journal of Criminal Law & Criminology." Perfect Method, Imperfect Analysts





#### **March 2007**

#### • The Guardian

(https://www.theguardian.com/science/2007/mar/23/crime.penal): " The reliability of fingerprint evidence has been called into question by a study that tested whether forensic experts make consistent judgments on print matches. Despite the perceived infallibility of fingerprint evidence, the study found that experts do not always make the same judgment on whether a print matches a mark at a crime scene when presented with the same evidence twice. ... snip ... The study by Itiel Dror, a psychologist at Southampton University, suggests otherwise. "I wanted to see if it is as objective and scientific as it claims to be," he said. "I wanted to see if the same expert would make the same decision on the same fingerprint if it is presented in a different context." He presented six fingerprint experts from various countries including the UK, the US and Australia with eight marks from crime scenes (called latent prints) and eight inked marks from suspects. "





#### **March 2007**

**The Guardian (continued)**: "The experts, who had 35 years experience between them, had all given judgments on the pairs of prints in previous court cases - four as matches and four as exclusions. But Professor Dror engineered the experiment so that none of them knew they were participating in a study, something that he says makes the study much more powerful. "If people know they are studying them they behave differently, especially if you are studying errors," he said. Of the 48 tests, the experts changed their decision in six cases and only two of the experts were consistent with their previous decision in all of their eight cases. They were more likely to change their decision if given contextual information, such as "the suspect has confessed", that conflicted with their previous judgment. "The same expert on the same fingerprint can make totally conflicting decisions, depending on the context," said Dr Dror, who presented his results at the British Psychological Society's annual meeting in York.





#### February 2017

**Dietrich College of Humanities and Social Sciences – Carnegie Mellon** University (https://www.cmu.edu/dietrich/news/newsstories/2017/february/fingerprint-science.html): "Fingerprints can be useful for eliminating suspects or linking crimes together, but as an identification tool, they are not as infallible as TV shows and judges and jurors are made to believe," said Kadane, the Leonard J. Savage University Professor of Statistics, Emeritus, in the Dietrich College of Humanities and Social Sciences.... Snip ... a fingerprint analyst may observe common characteristics between the mark left at a crime scene and a fingerprint on file. However, there is no current scientific basis to estimate the number of people who share these characteristics. In particular, there is no science to support the conclusion that only one person, the person whose fingerprint is on file, could have left the mark." **Perfect Method?** 





#### September 2017

**American Association for the Advancement of Science** (https://www.aaas.org/news/fingerprint-source-identity-lacksscientific-basis-legal-certainty): A new AAAS working group report on the quality of latent fingerprint analysis states: "We have concluded that latent print examiners should avoid claiming that they can associate a latent print with a single source and should particularly avoid claiming or implying that they can do so infallibly, with 100% accuracy ... snip ... Forensic examiners have long proclaimed high levels of certainty that latent prints, based on their analysis, originated from an "identified" person, statements that multiple reports have called "scientifically indefensible." Studies by the National Research Council in 2009, a National Institute of Standards and Technology's working group on latent fingerprint analysis in 2012, and, most recently, the President's Council of Advisors on Science and Technology in 2016, reached similar conclusions. Such assertions have led to false arrests and convictions."





#### September 2017

**American Association for the Advancement of Science (continued): "In** reality, there is not, at present, an adequate scientific basis for either claim," the AAAS report says. "There is no basis for estimating the number of individuals who might be the source of a particular latent print. Hence, a latent print examiner has no more basis for concluding that the pool of possible sources is probably limited to a single person than for concluding it is certainly limited to a single person. ... snip ... The "Forensic Science Assessments: A Quality and Gap Analysis of Latent Fingerprint Analysis" report makes clear that while latent fingerprint examiners can successfully rule out most of the population from being the source of a latent fingerprint based on observed features, insufficient data exist to determine how unique fingerprint features really are, thus making it scientifically baseless to claim that an analysis has enabled examiners to narrow the pool of sources to a single person."



#### Contombor 2017

#### September 2017

- **American Association for the Advancement of Science (continued):** "Another way to get around cognitive bias, the report states, is to improve the ability of automated fingerprint systems. Already, automated systems play an important role in helping law enforcement officials quickly cull through thousands of fingerprints to identify those with features "most similar" to latent fingerprints under review, the report notes. The systems, however, are unable to match a fingerprint lifted from a crime scene to one gathered earlier by authorities from a known source, nor can they determine whether a comparison of two prints – one from a crime scene, the other from police records – is valid, the report says. Still, the report points to the promise of automated systems, saying they could become effective in determining when a crime scene print matches a known print and in weighing the legal strength of a fingerprint analysis indicating that a pair of prints originated from the same person. "It is possible that automated fingerprint identification systems could evolve over time," the report says."
- https://mcmprodaaas.s3.amazonaws.com/s3fspublic/reports/Latent%20Fingerprint%20Report%20FINAL%209\_14.pdf?i9xGS\_EyMHnIPLG6INI UyZb66L5cLdlb



### Reality vs. fiction



- It's a scene that is played over and over again on TV. A grainy image appears on screen, a person says "enhance," a computer beeps, the image sharpens, and a red square flashes around a face.
- Some variation on "Got 'em!" is inevitably what's said next as a pair of detectives grab their jackets on the way out the door to nab the newly identified suspect...
- But that's not at all what happened in the days after April 15, 2013, when terrorists struck at the Boston Marathon.



## Reality vs. fiction



- Thanks to a survivor, the investigators were able to isolate images in which the first suspect, who we now know to be Tamerlan Tsarnaev, appeared.
- They followed him frame by frame by hand until footage revealed a co-conspirator.
- Images of the two subjects were in FBI database thanks to their driving licences but ...
- ...the brothers were identified the old fashioned way: the FBI held a press conference, publicized photographs of the suspects, and asked for the public's assistance.
- Automated face recognition simply wasn't up to the task.



## Why is it still difficult?

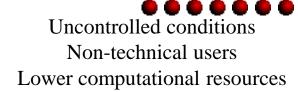
- The still insurmountable hurdle in the quest for highly accurate automated face recognition, at least for now, is *how* our **minds** resolve that information into a recognizable face
- Facial recognition systems are of limited use when fed grainy, low-resolution images, captured at a distance, from a cellphone camera or surveillance video, as with the images captured before, during and after the violence unleashed at the 2013 Boston Marathon.



## Iris: usability vs accuracy

Increasing popularity of mobile biometrics

















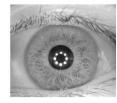


## Iris: usability vs accuracy – MICHE dataset

CASIA DATASET (LATEST VERSION)







CASIA\_INTERVAL



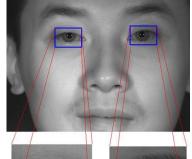


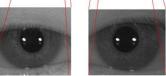


CASIA\_LAMP









CASIA\_IRISTHOUSA



Why not suited for unattended mobile testing?
A few mobile devices at present are equipped with (cheap and easy usable) NIR sensors



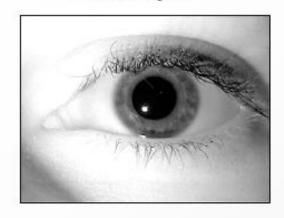
### Iris: usability vs accuracy – MICHE dataset

ICE (Iris Challenge Evaluation) COMPETITIONS by NIST





Left Eye



Why not suited for unattended mobile testing?
A few mobile devices at present are equipped with (cheap and easy usable) NIR sensors



### Iris: usability vs accuracy – MICHE dataset





First crucial difference: images in visible light Second crucial difference: uncontrolled conditions, normal equipment

Why not much suited for unattended mobile testing? Very high image resolution



#### MICHE dataset



- Images captured
  - in visible light
  - by «normal» user-level mobile devices



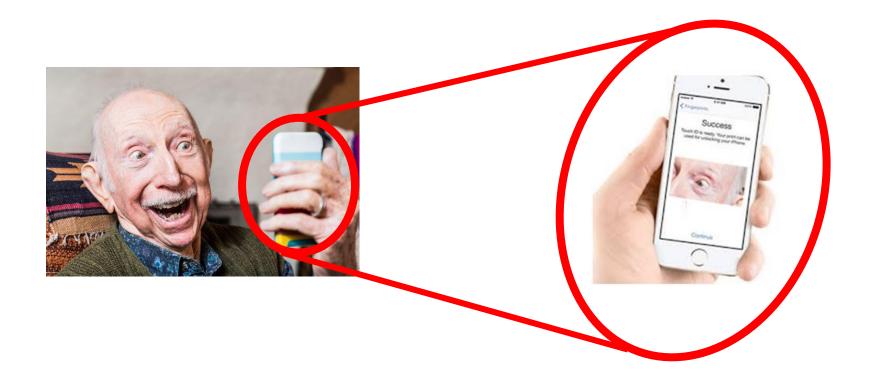




#### MICHE dataset



• Images captured by the user in uncontrolled/unattended conditions

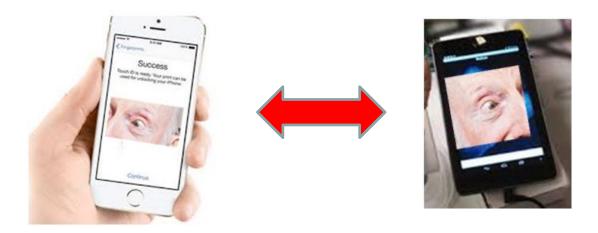




#### MICHE dataset

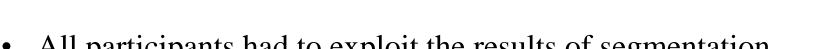


• Images captured by different devices for cross-device matching





## MICHE II competition setup – common segmentation

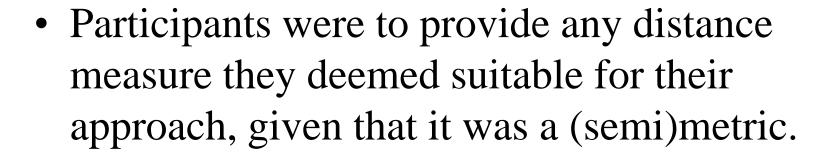


- All participants had to exploit the results of segmentation provided by the best algorithm in MICHE I competition\*
- For all participants, tests were repeated at BIPlab according to the common protocol

<sup>\*</sup>Haindl, M., Krupi cka, M., 2015. Unsupervised detection of non-iris occlusions. Pattern Recognition Letters 57, 60–65.



## MICHE II competition setup – distance measure



$$D: I_a \times I_b \to [0,1] \subset \mathbb{R}$$

1. 
$$D(I_a, I_a) = 0$$

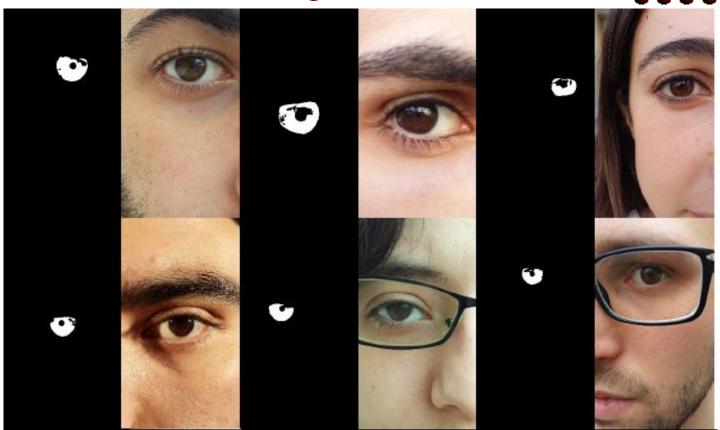
2. 
$$D(I_a, I_b) = 0 \rightarrow I_a = I_b$$

3. 
$$D(I_a, I_b) = D(I_b, I_a)$$



## MICHE II competition setup –

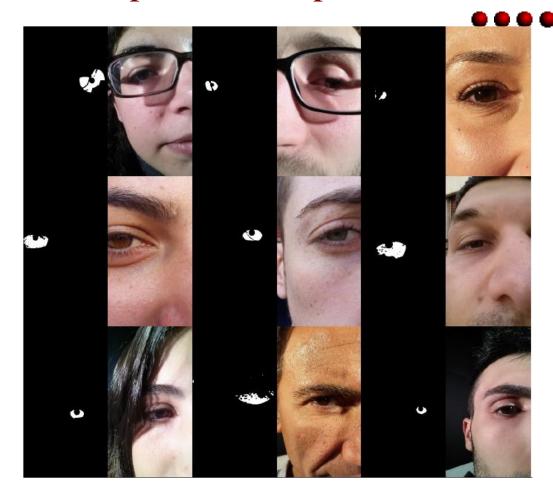
«good» cases



- Good segmentation
- Sufficient (and connected) iris region for matching



## MICHE II competition setup – «bad» cases



- Poor segmentation (possible fragments)
- Possibly insufficient iris region for matching







- Identification (1:N) = Recognition Rate (RR)
- Verification (1:1) = Area Under (ROC) Curve (AUC)

# MICHE II competition – participants



-----

- tiger miche
  - matching by a combination of a popular iris coding approaches and a periocular biometric processing based on the Multi-Block Transitional Local Binary Patterns (LBP); results are fused at score-level to improve the system accuracy.
- karanahujax
  - SIFT (Scale-Invariant Feature Transform) descriptors
  - Model1: Deep Learning in unsupervised mode
  - Model2: Deep Learning in supervised mode
- Raja
  - multi-patch deep features using deep sparse filters, in order to obtain robust features for iris recognition; the
  - patch level representation is obtained by dividing the iris image into a number of patches, that are mapped onto a collaborative subspace to perform classification via maximized likelihood; features are also extracted from the whole iris
- irisom
  - Two proposals with Self Organizing Map (SOM) of size  $5\times5$  and  $10\times10$
- FICO\_matcher
  - V1: distance measure considers color, texture and cluster descriptors
  - V2: distance measure only considers color and cluster descriptors
- otsedom
  - it exploits both Machine Learning paradigms, and Computer Vision techniques; well known descriptors are computed, such as LBP, LPQ, and WLD. They are used individually in order to construct a classifier, and then subsets of them are combined to outperform the obtained accuracy; the final algorithm combines the best five descriptors to obtain a robust dissimilarity measure of two given iris images.
- ccpsiarb (different combinations of computer vision transformations and machine learning-based classifiers)
  - ccpsiarb 17: Edge transformation + IB1 classifier
  - ccpsiarb 2: Equalize transformation + IB1 classifier
  - ccpsiarb\_42: Gaussian transformation + IB1 classifier



## MICHE II competition – results for same-device matching

A la a vith ve	GS4 vs GS4				
Algorithm	RR	AUC	<b>Global Score</b>	Time(seconds)	
tiger_miche	1,00	1,00	1,00	1,72	
karanahujax_Model1	0,83	0,97	0,90	5,72	
karanahujax_Model2	0,83	0,95	0,89	4,62	
irisom_10_10	0,77	0,88	0,82	3,45	
otsedom	0,67	0,94	0,80	42,37	
FICO_matcher_V1	0,67	0,89	0,78	1,00	
Irisom_5_5	0,63	0,88	0,75	3,00	
ccpsiarb_42	0,63	0,81	0,72	289,27	
ccpsiarb_17	0,63	0,81	0,72	61,43	
ccpsiarb_2	0,63	0,81	0,72	265,38	
FICO matcher V2	0,50	0,79	0,65	0,57	

karanahujax versions achieve comparable results, as well as ccpsiarb

irisom is less stable

the most significant difference in performance is achieved by FICO\_matcher versions

time was not considered for ranking, but we observe that best methoda also achieve acceptable processing times (single match)



## MICHE II competition –

results for same-device matching.

Alacovithus	IP5 vs IP5				
Algorithm	RR	AUC	Global Score	Time(seconds)	
tiger_miche	1,00	1,00	1,00	1,71	
karanahujax_Model1	1,00	1,00	1,00	5,63	
karanahujax_Model2	0,93	0,98	0,96	4,69	
FICO_matcher_V1	0,87	0,98	0,92	1,01	
Irisom_5_5	0,87	0,93	0,90	3,05	
irisom_10_10	0,83	0,92	0,88	3,51	
otsedom	0,63	0,92	0,78	41,10	
ccpsiarb_17	0,70	0,85	0,77	54,30	
FICO_matcher_V2	0,57	0,93	0,75	0,57	
ccpsiarb_2	0,63	0,86	0,75	252,84	
ccpsiarb_42	0,63	0,81	0,72	283,39	

karanahujax versions are still consistent, as well as different versions of ccpsiarb

irisom is stable this time (lower resolution = less noise?)

again, the most significant difference in performance is achieved by FICO\_matcher versions

as for time, the same observation regarding the best methods hold, but it is worth observing



## MICHE II competition –

## results for cross-device matching (hardest)

Algorithm	All vs ALL				
Algorithm	RR	AUC	Global Score	Time(seconds)	
tiger_miche	1,00	0,99	0,99	1,71	
karanahujax_Model2	0,92	0,86	0,89	4,65	
karanahujax_Model1	0,88	0,76	0,82	5,68	
irisom_10_10	0,80	0,78	0,79	3,48	
otsedom	0,63	0,93	0,78	41,74	
Irisom_5_5	0,75	0,79	0,77	3,03	
FICO_matcher_V1	0,73	0,80	0,77	1,00	
ccpsiarb_17	0,68	0,83	0,75	57,64	
ccpsiarb_2	0,65	0,82	0,74	259,27	
ccpsiarb_42	0,65	0,81	0,73	286,20	
FICO_matcher_V2	0,48	0,73	0,61	0,57	

karanahujax versions are still consistent, as well as different versions of ccpsiarb

irisom is quite stable

the most significant difference in performance is achieved by FICO\_matcher versions

as for time, the same observation regarding the best methods hold, but it is worth observing the speed of FICO matcher V2



## MICHE II competition – final results



		ALLvsA	GS4vsGS	Ip5vsIP	Final
Rank	Algorithm	LL	4	5	Score
1	tiger_miche	0,99	1,00	1,00	1,00
3	karanahujax_Model2	0,89	0,89	0,96	0,91
4	irisom_10_10	0,79	0,82	0,88	0,83
5	FICO_matcher_V1	0,77	0,78	0,92	0,82
6	otsedom	0,78	0,80	0,78	0,79
7	ccpsiarb_17	0,75	0,72	0,77	0,75

we choose karanahujax\_Mode2 instead of karanahujax\_Model (they got the same final score) since it has a more stable behaviour across operational settings, and also achieves better results in cross-device matching



#### MICHE II - Final observations

- In general, methods presented in different versions are quite stable across test coditions, except for FICO\_matcher
- The best metods also generally require less computational time per single match
- The better the ranking achieved, the more stable the method with respect to the test setting
- All methods provide consistently lower performances in ALLvsALL condition.
- On the average, the images over which the best results are achieved in same-device settings come from IP5 (mean value= 0,89), and the achieved results further present a lower standard deviation (0,086). Lower resolution = less noise?

	ALLvsALL	GS4vsGS4	Ip5vsIP5
Average Global Score achieved by proposed algorithms	0,85	0,87	0,89
Variance of Global Score achieved by proposed algorithms	0,090596	0,0943836	0,0865148



#### MICHE II - Conclusions



- The performance achieved by the methods participating in MICHE-I were not comparable to other investigation performed so far on iris recognition.
- However, the best methods participating in MICHE-II, focusing on feature extraction and recognition only, achieve extremely promising results.
- The most interesting aspect is that images were acquired in uncontrolled conditions and in visible light, that are widely recognized as extremely adverse conditions.
- On the other hand, most mobile devices are equipped with high-resolution RGB cameras, and not with Near Infrared (NIR) sensors that would allow better accuracy.
- Results suggest that it is worth further exploring possible improvements of the available techniques.



#### Some references



- Cole, S. A. (2004). More than zero: Accounting for error in latent fingerprint identification. *J. Crim. l. & Criminology*, 95, 985.
- Dror, I. E., & Mnookin, J. L. (2010). The use of technology in human expert domains: challenges and risks arising from the use of automated fingerprint identification systems in forensic science. *Law, Probability & Risk*, 9(1), 47-67.
- Ahmed, N. U., Cvetkovic, S., Siddiqi, E. H., Nikiforov, A., & Nikiforov, I. (2017). Combining iris and periocular biometric for matching visible spectrum eye images. *Pattern Recognition Letters*, *91*, 11-16.
- Ahuja, K., Islam, R., Barbhuiya, F. A., & Dey, K. (2017). Convolutional neural networks for ocular smartphone-based biometrics. *Pattern Recognition Letters*, *91*, 17-26.
- Raja, K. B., Raghavendra, R., Venkatesh, S., & Busch, C. (2017). Multi-patch deep sparse histograms for iris recognition in visible spectrum using collaborative subspace for robust verification. *Pattern Recognition Letters*, *91*, 27-36.
- Abate, A. F., Barra, S., Gallo, L., & Narducci, F. (2017). Kurtosis and skewness at pixel level as input for SOM networks to iris recognition on mobile devices. *Pattern Recognition Letters*, *91*, 37-43.
- Galdi, C., & Dugelay, J. L. (2017). FIRE: Fast Iris REcognition on mobile phones by combining colour and texture features. *Pattern Recognition Letters*, *91*, 44-51.
- Aginako, N., Castrillón-Santana, M., Lorenzo-Navarro, J., Martínez-Otzeta, J. M., & Sierra, B. (2017). Periocular and iris local descriptors for identity verification in mobile applications. *Pattern Recognition Letters*, *91*, 52-59.
- Aginako, N., Echegaray, G., Martínez-Otzeta, J. M., Rodríguez, I., Lazkano, E., & Sierra, B. (2017). Iris matching by means of Machine Learning paradigms: A new approach to dissimilarity computation. *Pattern Recognition Letters*, *91*, 60-64.