Face Quality @ Idemia

Criteria Analysis

Stéphane Gentric, Ph. D.
Associate Professor @TelecomParis
Chief AI Scientist
Fellow Expert

IDEMIA

Warning: Copyright Notice

Context

ISO/IEC 19794-5:(2005, 2011) and 39794-5:2019 aims to define face image data.

The ISO criteria are a list of relevant recommendations but mainly without quantitative assessment.

However, even today, this is the main document used to drive face image qualities for any purpose : from enrolment in a biometric system, to printing of faces on ID-Documents.

For example, European regulation states that the EES (Entry Exit System) facial quality algorithm shall be comprehensible in terms of the ISO/IEC 19794-5:2011 criteria.

Many Customers ask to assess quality of face images in order to ensure the biometric performance of their system.

The correlation between ISO criteria and biometric performance needs to be investigated.

FFA – Full Frontal Assessment

FFA is one global value and one explanation is case of failure.

Actionable criteria from ISO/IEC 19794-5:2011 are used

Internal thresholds are set to limit the total False Rejection Rate

FFA v1 was developed in 2020

FFA v2 is under development and aims at improving previous version with additional features and evaluations :

- **Performance**

  We update algorithms and add new criteria

  We label images and validate behaviour on extreme cases

  We perform large validation on operational datasets and wild images

- **Link with biometric performance**

  At a very selective threshold, in 1:N @FPIR=0,1% with N=100 Millions

- **Selectivity on existing ISO documents**

  All Countries have defined rules/process to produce ISO.

Operational Datasets

Few customers allow us to compute statistics on their images

Main results presented here, come from a dataset of 3.5 millions pairs on face images coming from different persons, coming from various countries around the globe

Images are sequestered on customer data centre. Visual inspection is not possible.

For each pairs, we have :

One image coming from an ID-Document, supposed to follow ISO-2005 or 2011 recommendations

One image acquired live, using various acquisition systems

Lots of qualities on both images

A biometric score with a recent Idemia Face Recognition algorithm

Results presented here also incorporate tests on :

internal datasets with ground truth estimated by human

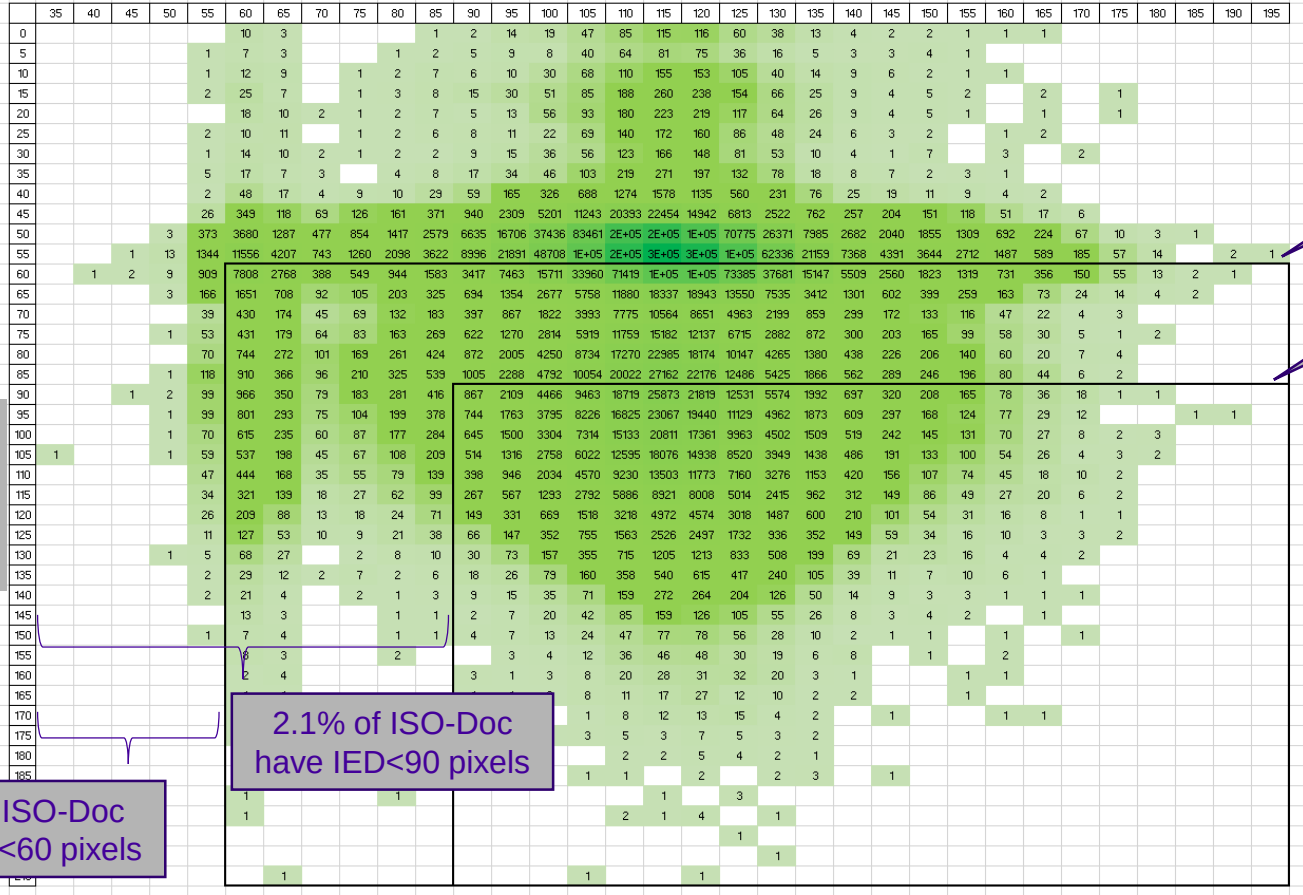
large datasets from images taken without any constraints (named “Wild” Images)

Dataset repartition with IED (Inter Eyes Distance)

3.5 Millions pairs of face images

ISO-Doc IED

Live IED

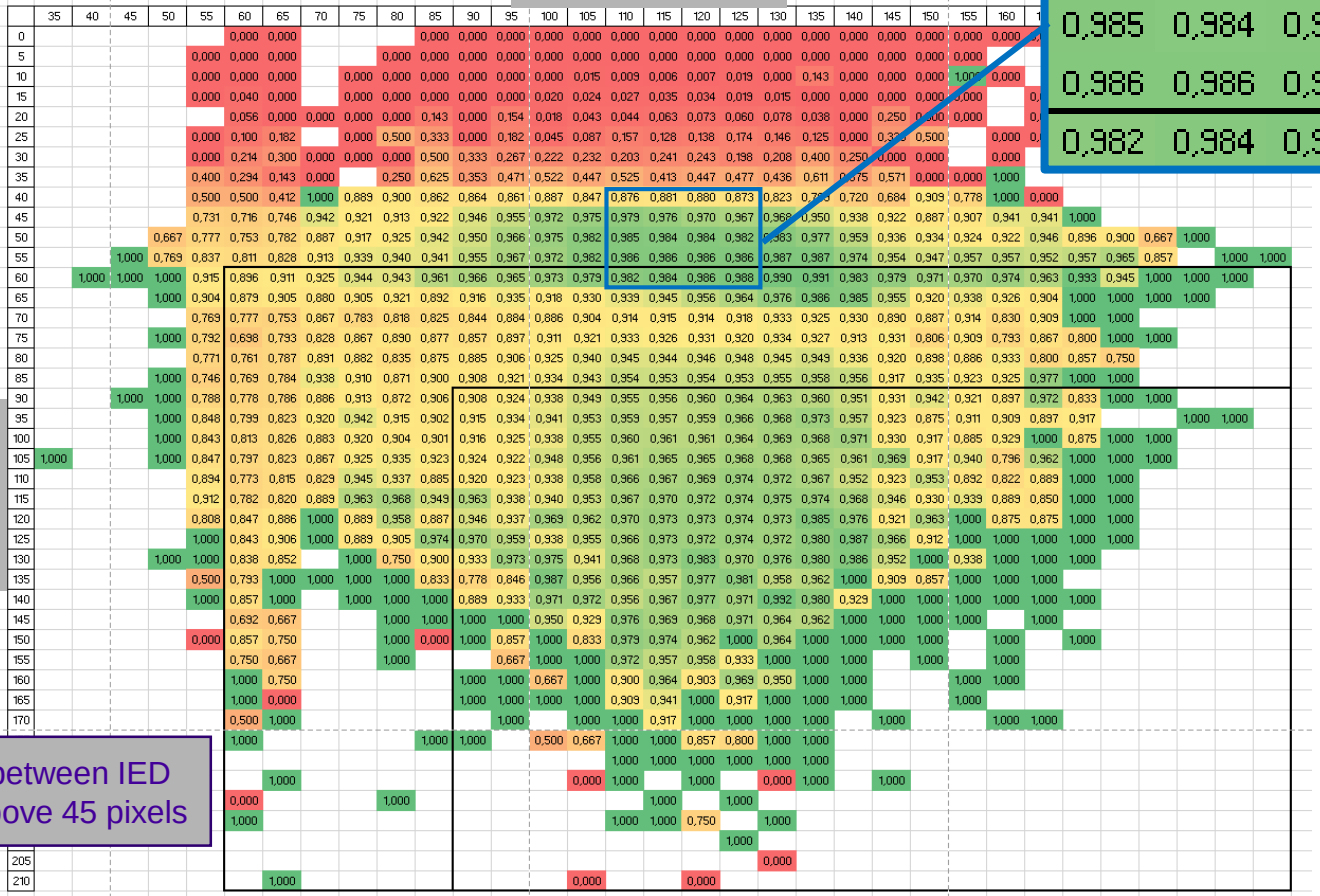


Correlation between IED and Accuracy

Global accuracy :
TPIR = 97,4%
@FPIR = 0.1%
for N = 100 Millions

ISO-Doc IED

Live IED



0,876	0,881	0,880	0,873
0,979	0,976	0,970	0,967
0,985	0,984	0,984	0,982
0,986	0,986	0,986	0,986
0,982	0,984	0,986	0,988

No strong link between IED and accuracy above 45 pixels

Face Position

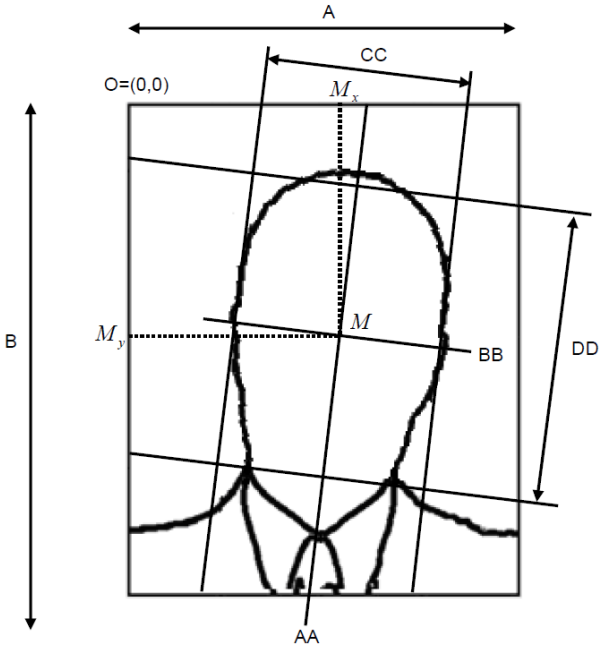
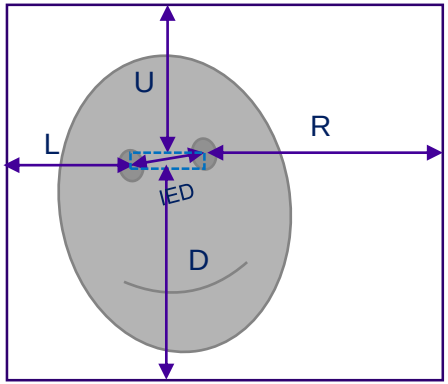


Figure 14 — Geometric characteristics of the Full Frontal Face image

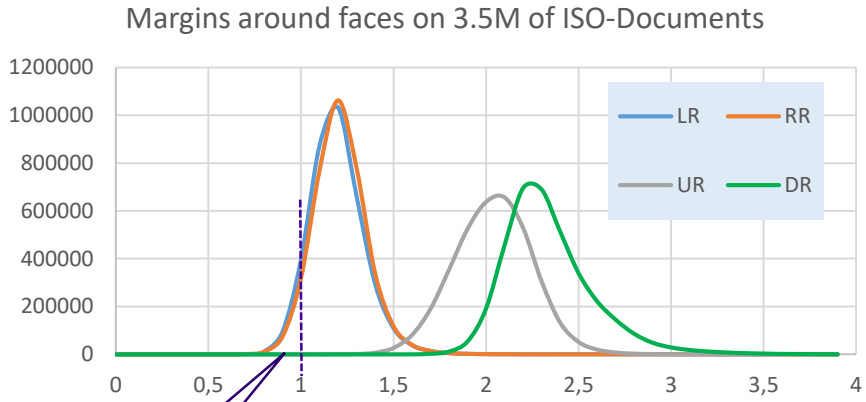
Image from ISO/IEC 19794-5:2011



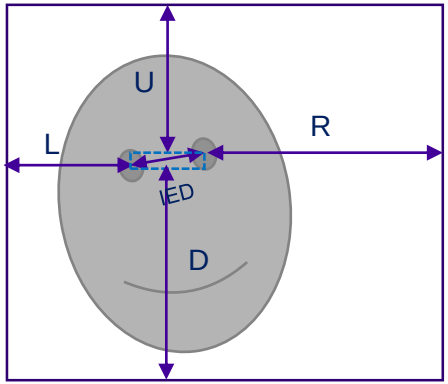
LR = L / IED
RR = R / IED
UR = U / IED
DR = D / IED

- Exact head size are hard to determine when occluded by hairs. Using head size leads to unstable evaluation of face position.
- FFA v1 & FFA v2 use eyes positions.
- Margins around eyes are computed relatively to IED

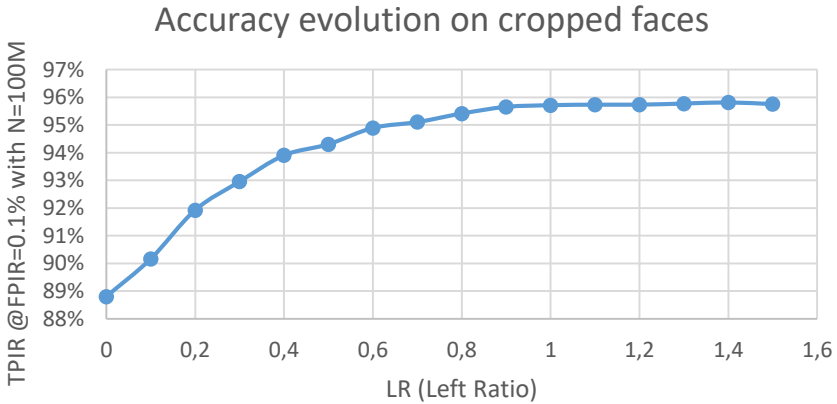
Face Position



3% of have
RR or LR < 1.0



LR = L / IED
RR = R / IED
UR = U / IED
DR = D / IED



	min value FFAv1	min value FFAv2	max value FFAv2
LR	1.0	0.8	2.0
RR	1.0	0.8	2.0
UR	1.0	1.3	2.7
DR	1.5	1.7	3.2

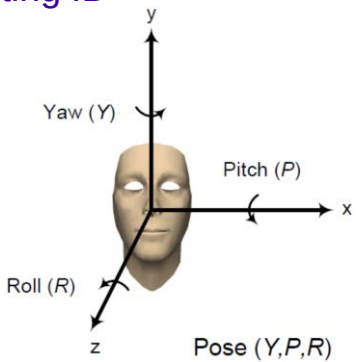
Face Pose

- Requirement for pose are +/- 8° for Roll and +/- 5° for Yaw and Pitch
- From eyes position, on frontal faces, roll is easy to compute and very precise.
- Estimation of Yaw & Pitch by human on a single image is very imprecise. Without ground truth the direct evaluation of Yaw & Pitch estimators is not possible.

- Yaw & Pitch have been learned from 3DModel fitting. Operation threshold have been set to limit reject on existing ID-Documents.

	threshold	ISO-Doc	Live
Yaw	10,3°	0,098%	6,230%
Pitch	15,0°	0,132%	2.679%
Roll	8°	0,167%	3,673%

- Subject tends to be less frontal on live acquisition systems.



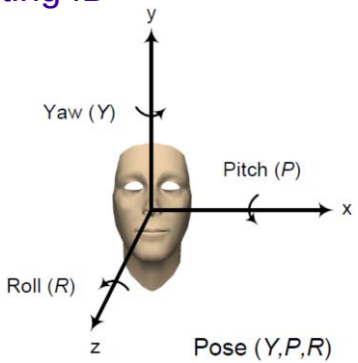
Face Pose

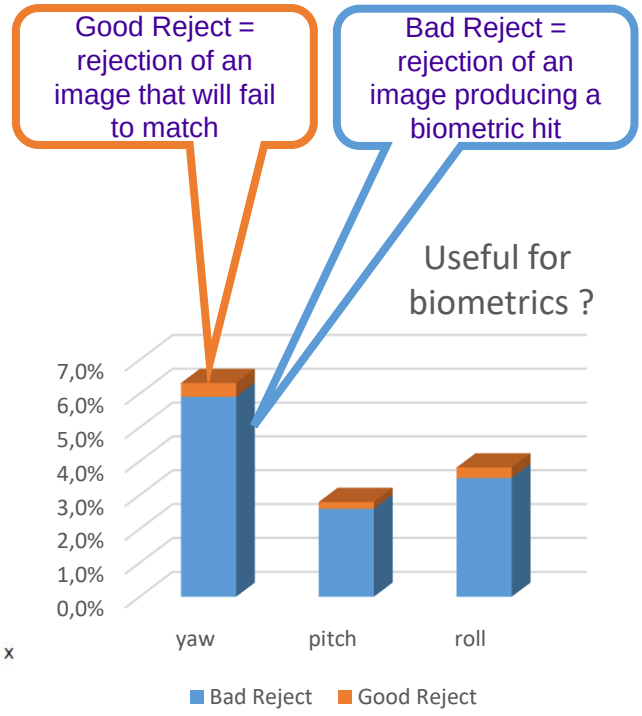
- Requirement for pose are +/- 8° for Roll and +/- 5° for Yaw and Pitch
- From eyes position, on frontal faces, roll is easy to compute and very precise.
- Estimation of Yaw & Pitch by human on a single image is very imprecise. Without ground truth the direct evaluation of Yaw & Pitch estimators is not possible.

- Yaw & Pitch have been learned from 3DModel fitting. Operation threshold have been set to limit reject on existing ID-Documents.

	threshold	ISO-Doc	Live
Yaw	10,3°	0,098%	6,230%
Pitch	15,0°	0,132%	2.679%
Roll	8°	0,167%	3,673%

- Subject tends to be less frontal on live acquisition systems.

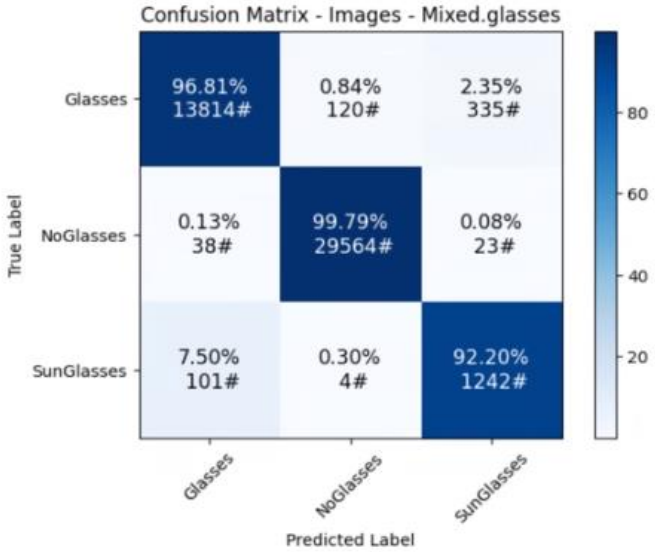




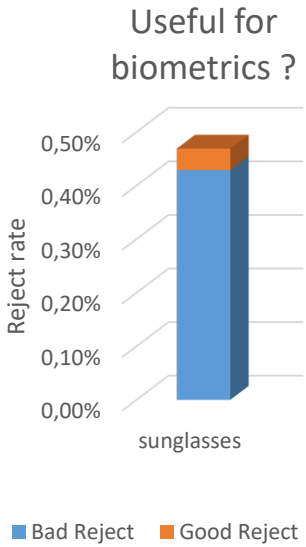
Occlusion - SunGlasses

- Generic Occlusion detection is not mature enough to be used at a low false positive rate.
- We focus on two specific occlusions: Sunglasses and Sanitary Masks.

- Differences between opaque sunglasses and some semi-transparent glasses may be small, however sunglasses detection works well without false detection on ISO-Documents.



	ISO-Doc	Live	Wild
SunGlasses	0,014%	0,440%	1.8%



Occlusion – “Sanitary” face Mask

- Even if this is not a dedicated criteria in ISO document, today it seems mandatory to check if there is a mask or not.
- Detection of sanitary mask was a hot topic in 2020. It is improved in FFA v2.

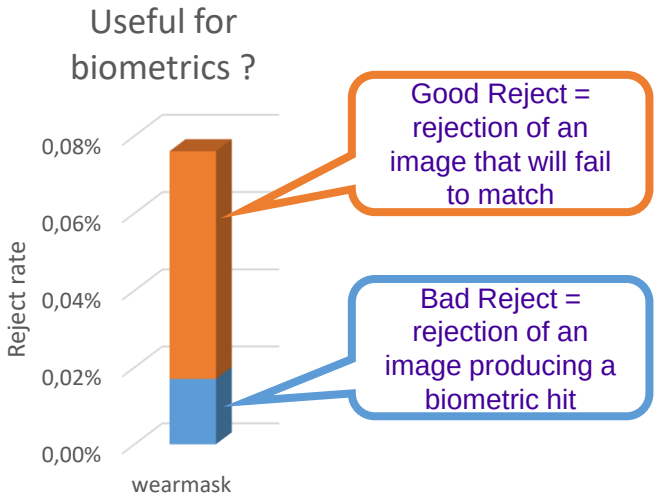
- Almost no false detection on existing Iso-Documents

	ISO-Doc	Live	Wild
No_Mask	0,002%	0,074%	0,159%

- Performance have been evaluated on 2 larges annotated datasets

Datasets	FP	FN
Synthetic Mask	0,011%	0,004%
Annotated	0,280%	0,280%

- Mask do impact biometric performance. Lots of mask detection leads to biometric rejections



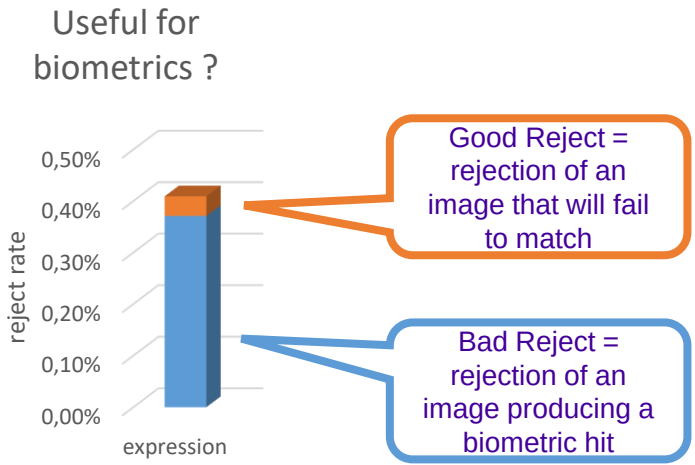
Mouth Open and Expression

- Regarding mouth open, expression and teeth visibility, ISO versions are more and more precise from 2005 to 2011, then 2019.
- We predict one criteria mixing these constraints
- People tend to smile more often on their document than during live authentication

	ISO-Doc	Live	Wild
Mouth Open	0,292%	0,138%	7.27%

- Expression do not affect a lot Biometric performance

	Score	Opening	smile	teeth visible
bad quality	0.0-0.2	Wide open	Yes	Yes
	0.2-0.4	Open	Yes	Yes
	0.4-0.6	Open	No	Yes/No
	0.6-0.8	Almost closed	Yes	Partly
	0.8-0.9	Almost closed	No	No
	0.8-0.9	Closed	Yes	No
good quality	0.9-1.0	Closed	No	No

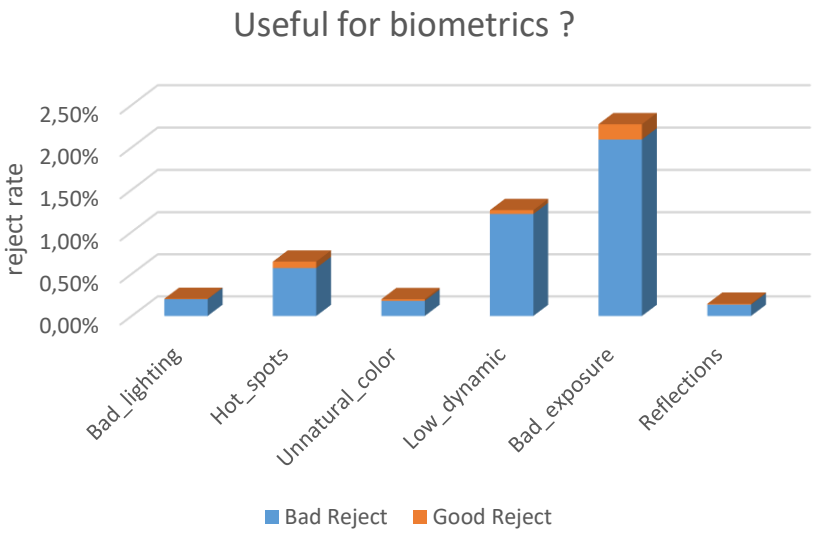


Illumination

- Lots of criteria regarding illumination are described into ISO documents, without actionable definitions. Today Biometric algorithms are very robust to most illumination environments.
- Defining a ground truth is also a complex task.
- FFA v2 handles 6 additional criteria independently with conservative decision thresholds.

	ISO-Doc	Live	Wild
Bad_lighting	0,025%	0,185%	1,16%
Hot_spots	0,020%	0,636%	0,45%
Unnatural_color	0,009%	0,182%	0,45%
Low_dynamic	0,074%	1,185%	0,11%
Bad_exposure	0,073%	2,126%	4,28%
Glasses_reflections	0,007%	0,144%	0,03%

- Some Live Acquisition may have been done in low light environments.



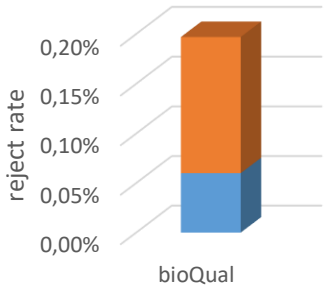
Biometric Quality

- CNN learned to mimic biometric scores
- Biometric Quality is the best predictor for biometric performance.
- Used in FFA to remove low quality images not explained by ISO criteria
- Return a normalized value in 0-100

- Almost no false reject on ID-Documents

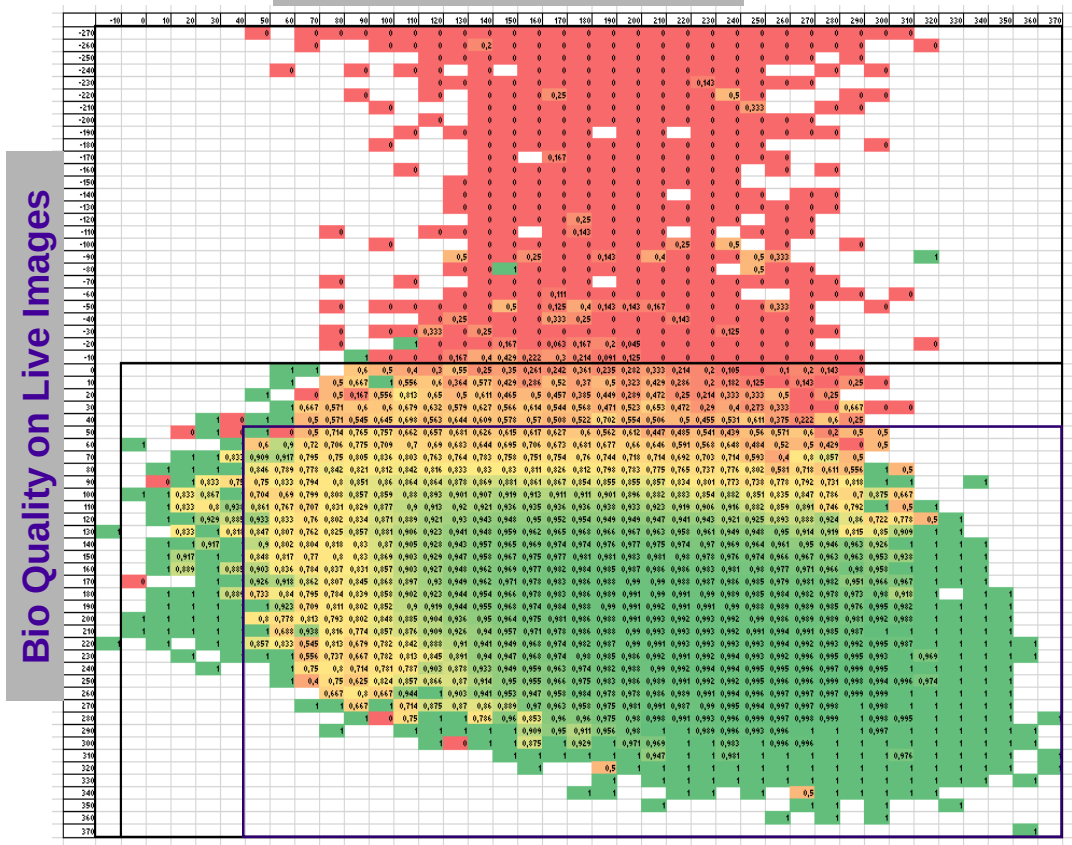
	ISO-Doc	Live
Bio Quality	0,011%	0,185%

Useful for biometrics ?



Bio Quality on ISO-Doc

Bio Quality on Live Images



FFA pipeline

FFA returns one global value and one explanation is case of failure.

FFA is a pipeline. It checks criteria in a fixed order.

Some algorithms are less accurate when an other criteria fails : Open Mouth with Mask, Pose on cropped faces…

On this test on 3.5 millions, we have 3.3% of reject.

With FFA, 96.7% of existing ID-Documents are considered compliant.

			ratio	number
1	IED too small		2,032%	71857
2	position	Left Ratio	0,024%	861
3		Right Ratio	0,032%	1123
4		Up Ratio	0,288%	10181
5		Down Ratio	0,115%	4053
6	pose	Yaw	0,093%	3272
7		Pitch	0,127%	4475
8		Roll	0,157%	5537
9	occlusion	wear mask	0,002%	59
10		sunglasses	0,012%	431
11	expression/mouth		0,214%	7562
12	illumination	Bad lighting	0,020%	717
13		Hot spots	0,015%	539
14		Unnatural color	0,008%	271
15		Low dynamic	0,073%	2565
16		Bad exposure	0,051%	1794
17		Glasses Reflections	0,005%	173
18	Biometric Quality		0,011%	383
	total		3,27%	
	ISO compliant with FFA:		96,7%	

Conclusion

- The FFA evaluates ISO criteria having an impact on biometric performance
  - FFA ensures the relevance of the explanation in case of rejection
  - FFA is set to limit rejections on images able to produce high biometric scores
  - The rejection rate on existing ISO documents is very limited
- The FFA anticipates future European regulations on Artificial Intelligence by improving explainability
- The FFA will help customers to reduce acquisition throughput by avoiding unnecessary multiple acquisitions, while ensuring the biometric performance of the central system.

Perspectives on ISO/IEC 29794-5:202x

Today, at Idemia we come across two main uses where customers refers to ISO/IEC 19794-5:2011

- **Biometric purpose**
  - Is the quality good enough for comparison ?

  (as central system provider or as image acquisition device provider)
- **Document Printing purpose**
  - Is the quality good enough to print this image on a document ?

  Printed documents are used by humans and algorithms

When it comes to implementation, qualities are better defined by a purpose that by a semantic definition.
As biometric performance increases, the relevance of some criteria evolves. Strict compliance with these criteria may limit the deployment of some applications (free flow boarding, frictionless acquisition…).

The scope of future normative document related to face has to be clearly defined.

Questions ?

Stephane.gentric@idemia.com

IDEMIA

Warning: Copyright Notice