Unified Quality Score based on MagFace

08.11.2023

Johannes Merkle

secunet Security Networks AG

https://de.wikipedia.org/wiki/StyleGAN

Motivation

- Face quality is a pre-requisite for successful deployment of face recognition
- Error rates are strongly dependent on face image quality
- Quality assurance of face images is paramount
- Good facial image quality assessment (FIQA) algorithms required
- OFIQ will be an open source implementation of FIQA algorithms
  - Funded by BSI

Face Image Quality

- For face recognition, the facial images should meet various requirements
  - Capture related: illumination, exposure, sharpness, background, …
  - Subject-related: Frontal pose, neutral expression, no occlusions, …
- Some are known to contribute to low error rates, others are imposed by regulations, e.g. for passport images

Face Image Quality

- Beside assessing the individual requirements, OFIQ will output a Unified Quality Score
- Acc. to ISO/IEC 29794-5, it shall be trained to predict the **utility** of the face image
  - Degree to which it supports biometric recognition performance (ISO/IEC 29794-1)
- Thus, the Unified Quality Score relates to the error rates of face recognition

QS = 95	QS = 85	QS = 62	QS = 42
a	b	c	d

Figure 7 – Four face images with example image quality values.

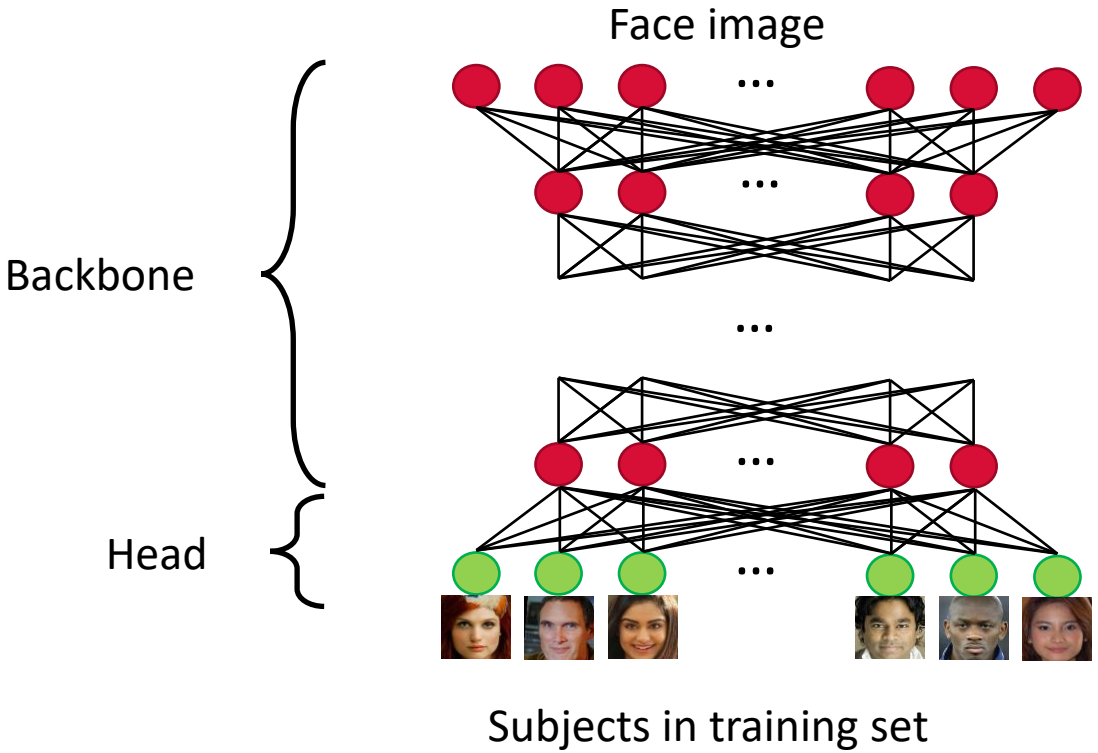
From: FATE Quality Concept Document

MagFace

- MagFace: A Universal Representation for Face Recognition and Quality Assessment
  - Meng, Zhao, Huang & Zhou, CVPR 2021
- Defines a new loss function for the training of face recognition CNNs
- Leads the CNN to learn quality-aware feature representations
- Extends the ArcFace loss
  - Enforce correlation between quality and magnitude (length) of feature representation

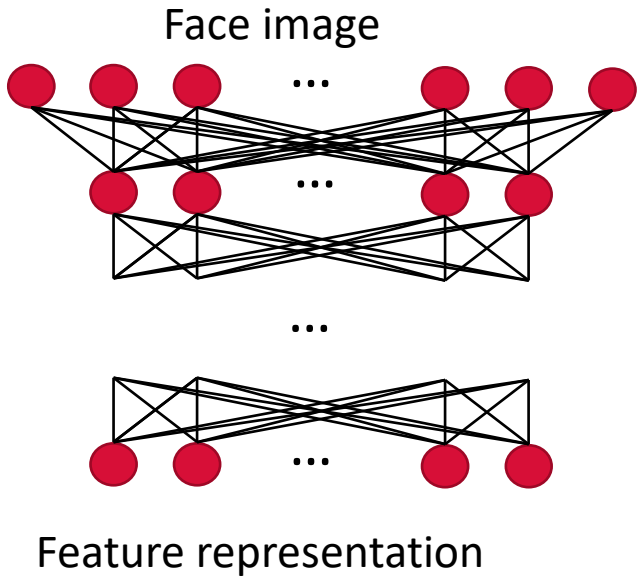
MagFace

- Most modern face recognition CNNs are trained on classification
  - Class = subject in training set
  - Final layer(s) (“head”) perform classification



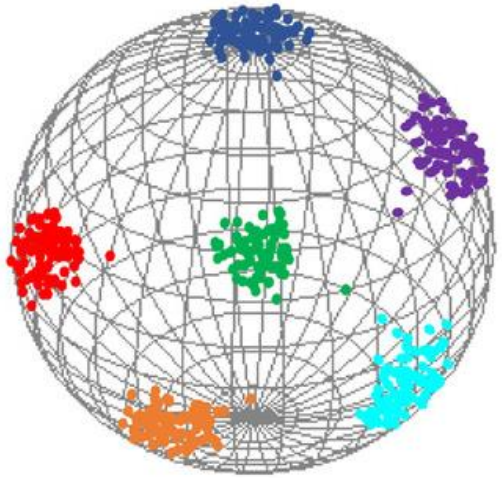
MagFace

- Most modern face recognition CNNs are trained on classification
  - Class = subject in training set
  - Final layer(s) (“head”) perform classification
  - After training, the head is removed
    → output: internal feature representations
  - Loss function ensures that distance of feature representations measure similarity of face images
  - Distance measure depends on loss function
    → e.g. Euclidean or cosine distance



MagFace

- ArcFace loss:
  - Similarity measured by geodetic distance on hypersphere
  - Feature representations are normalized to unit length
  - Additional margin is enforced between classes (subjects) by penalizing too short distance to other class centres
  - Length (magnitude) of feature representation is irrelevant



Distances of representations on hypersphere (colors = classes)

Zhang, Gong, Chen: Face recognition based on adaptive margin and diversity regularization constraints. IET Image Processing 15 (5). 2021

MagFace

- MagFace loss:
  - Enforced class margin depends on length of feature representation
  - Penalty for too short distance to other classes decreases with representation length
    i.e. longer feature representation → larger margin required
  - Additional regularization term to maximize representation length
- Leads the CNN to learn quality-aware feature representations
  - Typically, poor quality images result in higher variance of representations
    → Smaller distance (margin) to other classes
    → Shorter feature representation
- Results in a combined CNN for face recognition and unified quality score

MagFace

- Authors have published different MagFace models on Github
  - iResNet100
  - iResNet50
  - iResNet50FP16 (trained with float16)
  - iResNet18
- Models have been evaluated with respect to the quality score

Evaluation Methodology

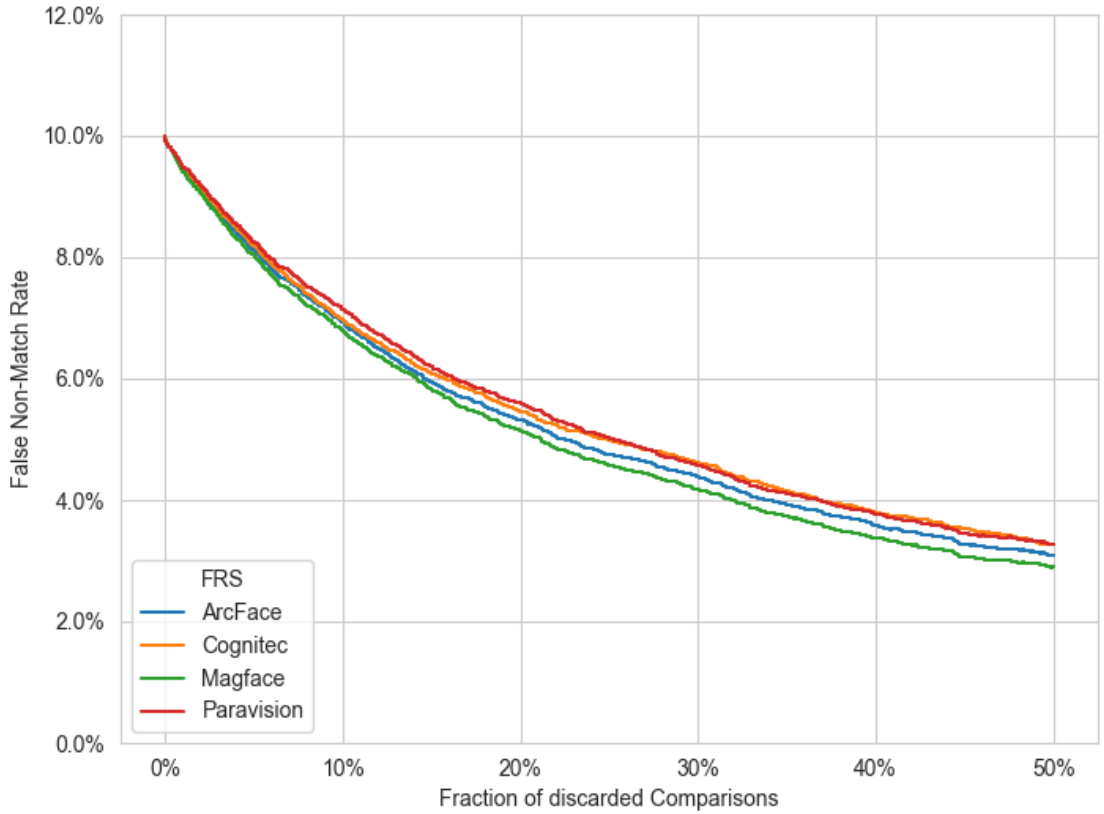
- Unified quality score doesn’t assess any specific property of the image
  - No evaluation with ground-truth data possible
- It predicts the image’s influence on biometric performance
- Typically, quality has a higher impact on FNMR than on FMR
- Thus, we evaluate the correlation of the score with the FNMR
- Precisely, we use Error-versus-Discard Characteristics (EDC)
  - Analyses the decline of the FNMR when using the FIQA algorithm for quality assurance

Evaluation Methodology

- Face recognition algorithms
  - Commercial products: Paravision (1.0.6) and Cognitec (9.3.2.0)
  - Open source CNNs: ArcFace (ResNet100) and MagFace (ResNet100)
- Test set: 50.000 images from VGGFace2
  - Large of variance of quality and variety of quality issues

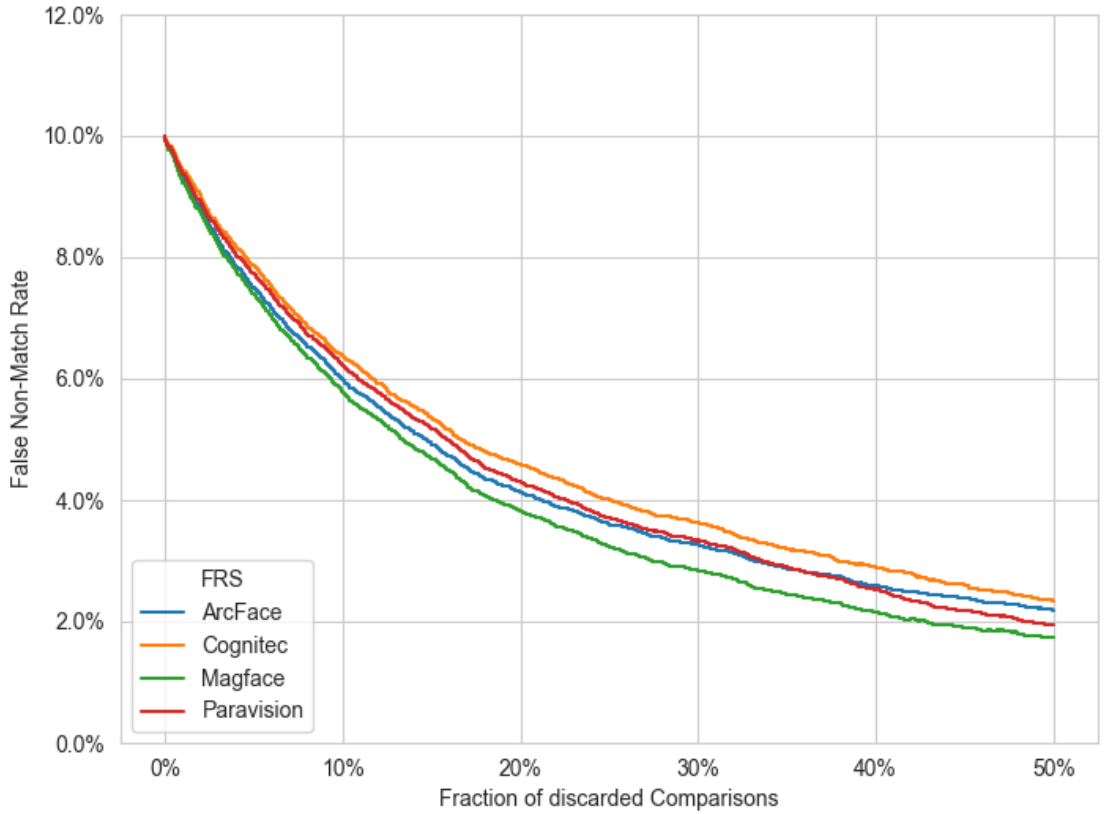
Evaluation Results

- iResNet18



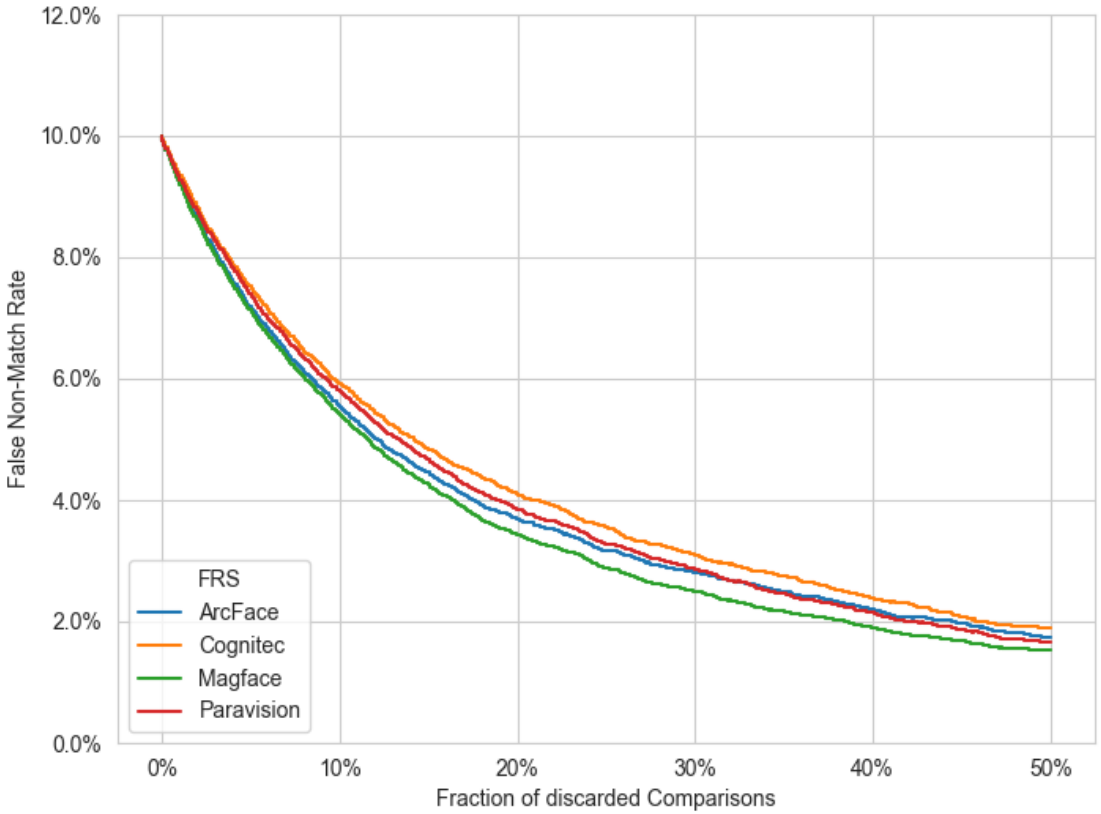
Evaluation Results

- iResNet50



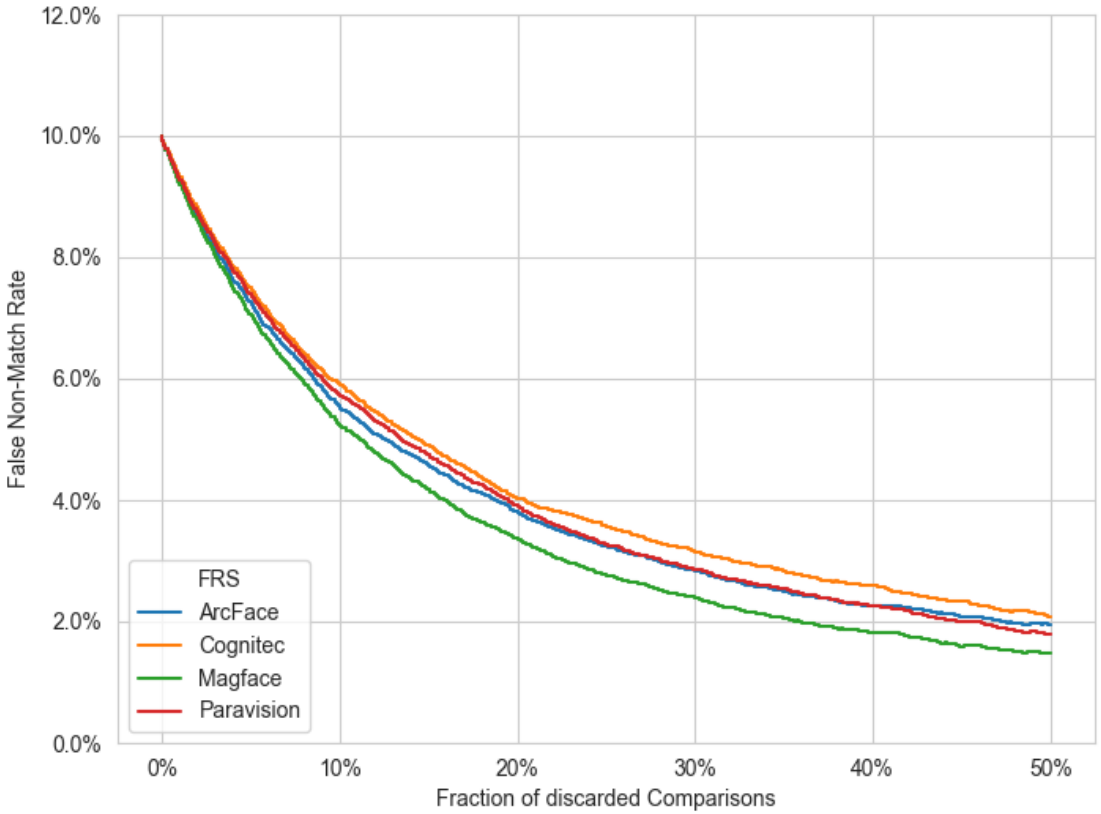
Evaluation Results

- iResNet50FP16



Evaluation Results

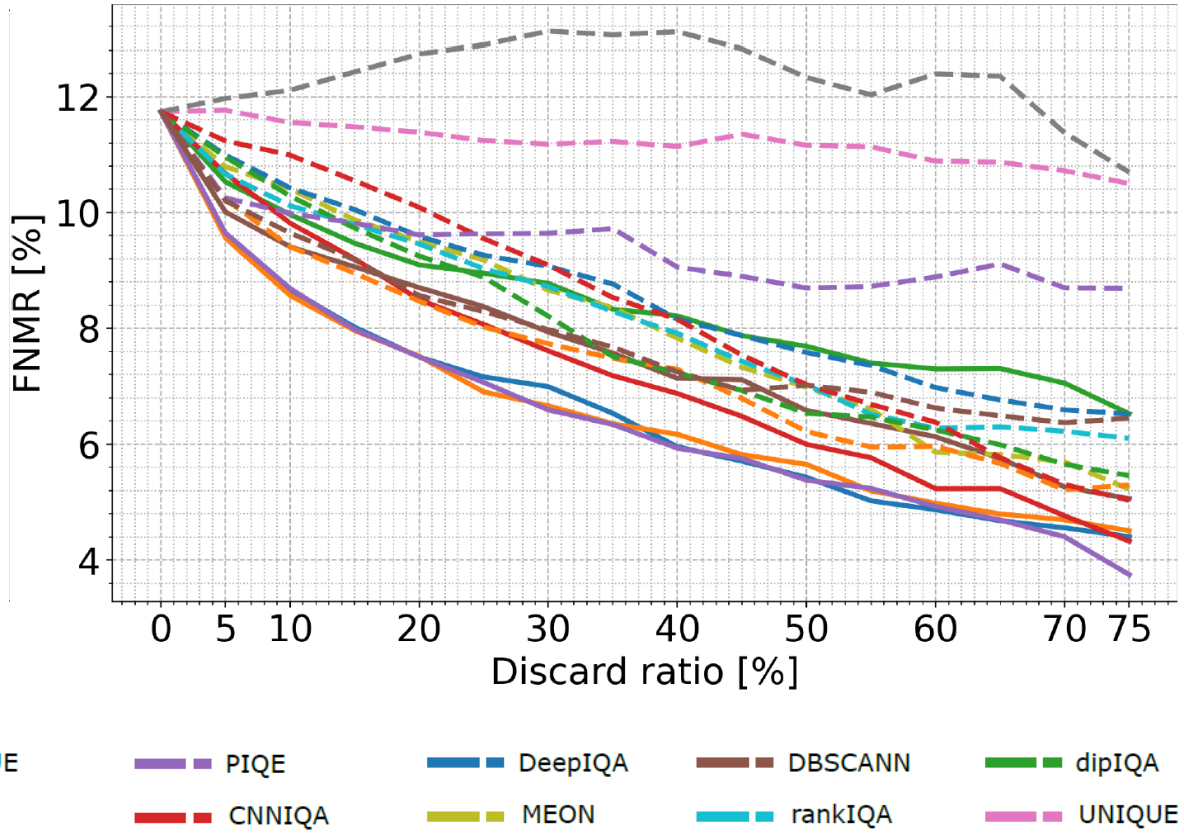
- iResNet100



Evaluation Results

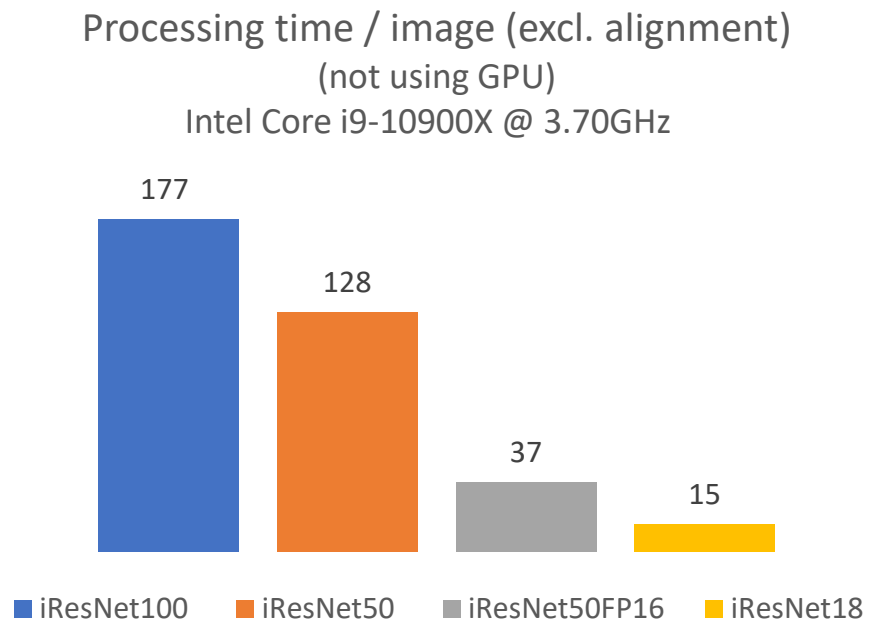
- Comparison of MagFace (iResNet100) with other FIQA
  - On VGGFace2
  - ArcFace as face recognition algorithm
  - Top performance, on par with PFE and SDD-FIQA

Fu, Chen, Henniger & Damer: A Deep Insight into Measuring Face Image Utility with General and Face-specific Image Quality Metrics. IEEE/CVF Winter Conference on Applications of Computer Vision. 2022



Evaluation Results

- iResNet50 floating point 16 very efficient



Summary and Outlook

- MagFace approach shows good predictive performance as FIQA algorithm
- Best trade-off with processing time achieved for model iResNet50 FP16

- Future improvements:
  - Fusion with criteria-specific algorithms of OFIQ, e.g. head pose, illumination, eyes open
    - Requires training on large data set
    - Outputs of algorithms (incl MagFace) are features
    - Labels computed from comparison scores

secunet