Face Image Quality Assessment (FIQA): Recent Advancements and Future Challenges

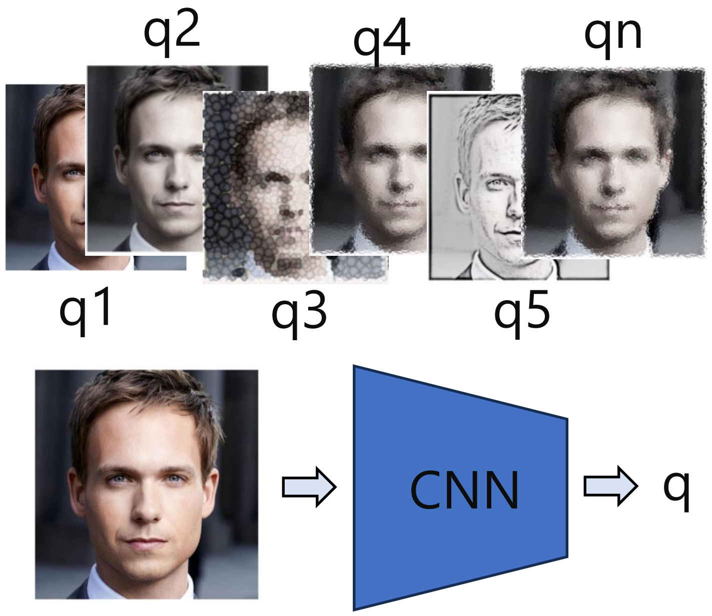
Prof. Vitomir Štruc, PhD

Faculty of Electrical Engineering
University of Ljubljana, Slovenia

vitomir.struc@fe.uni-lj.si

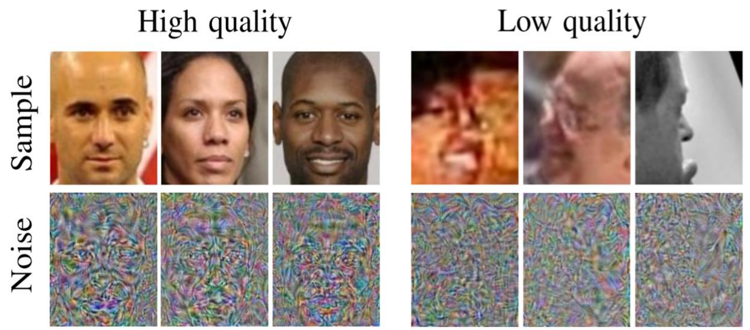
Talk Outline

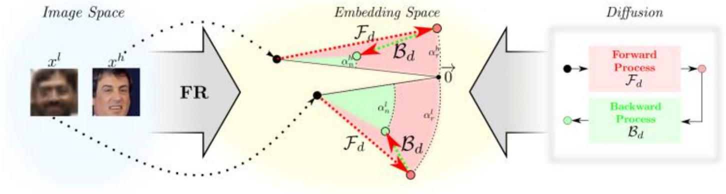
Part I: Face Image Quality Assessment



Background and Concepts

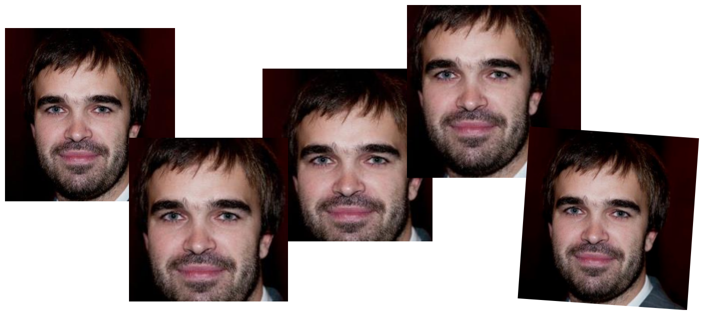
Part II: FaceQAN, DifFIQA and eDifFIQA

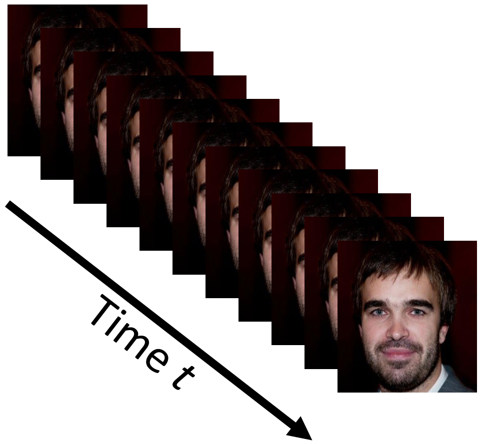




SOTA FIQA

Part III: Open Challenges





Part I: Background and Concpets

Publications available:
https://lmi.fe.uni-lj.si/en/vitomir-struc/publications/

GitHub: https://github.com/LSIbabnikz/

Face Image Quality Assessment

- Image Quality Assessment (IQA)
  - Image Quality:
    - The level of accuracy with which different imaging systems capture, process, store, compress, transmit and display the signals that form an image[1]
    - The weighted combination of all of the visually significant attributes of an image[1]
  - IQA methods can be:
    - Subjective or Objective
    - Related to Noise, Blur, Sharpness, Contrast, Color Accuracy, Distortions, etc.



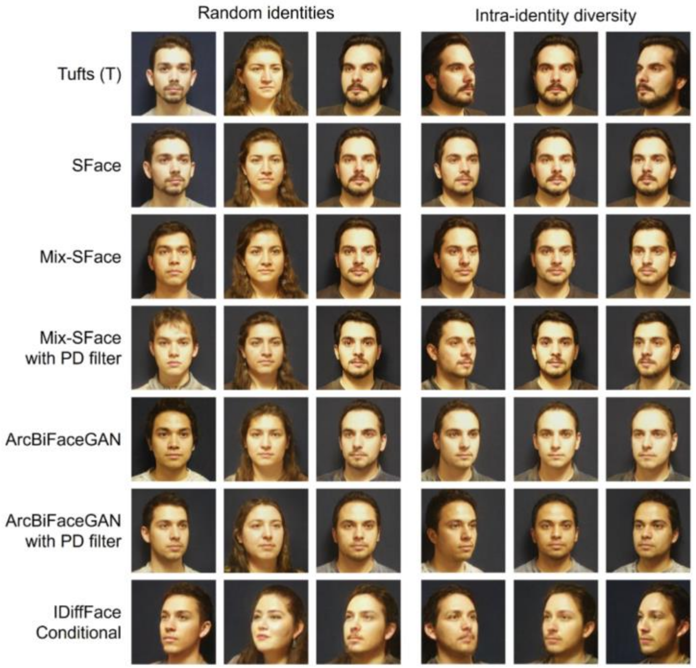


[1]Burningham, Norman; Pizlo, Zygmunt; Allebach, Jan P. (2002). "Image Quality Metrics". In Hornak, Joseph P. (ed.). *Encyclopedia of imaging science and technology*. New York: Wiley.

Face Image Quality Assessment

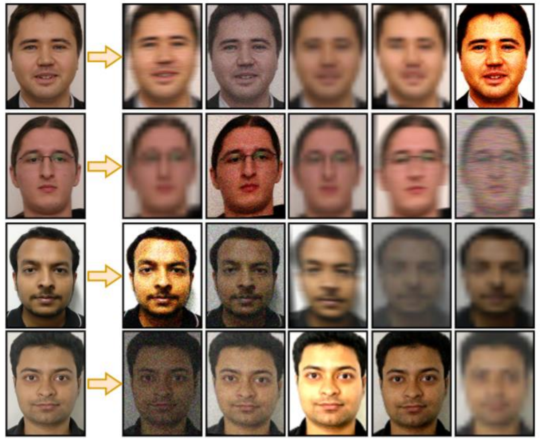
- Face Image Quality Assessment (FIQA)
  - IQA applied to face images
  - Important for multiple use-cases, e.g.:

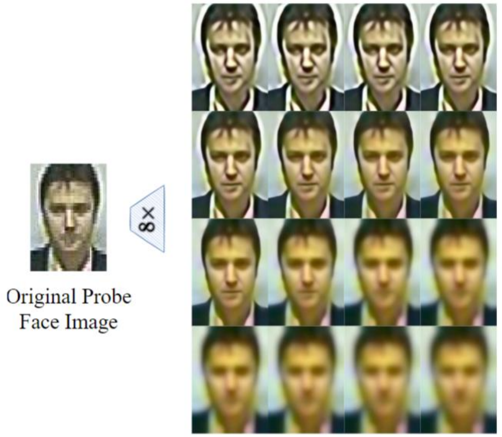
Synthetic Data Generation

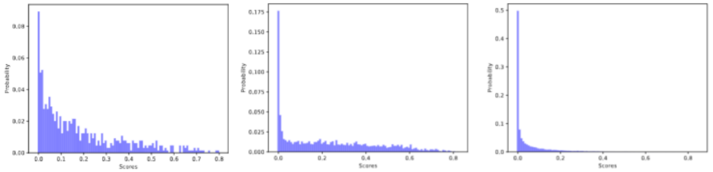


Sp.	Model	PD	CR-FIQA ↑
VIS	Tufts (H vs. T) [17]	–	1.743 ± 0.259
	StyleGAN2 [36]	–	2.141 ± 0.075
	DB-StyleGAN2 [18]	–	2.130 ± 0.009
	SFace [29]	–	1.779 ± 0.193
	Mix-SFace [29]	–	1.792 ± 0.164
		✓	1.824 ± 0.149
	ArcBiFaceGAN	–	1.841 ± 0.108
		✓	1.835 ± 0.140
	IDiff-Face (N) [31]	–	1.559 ± 0.332
	IDiff-Face [31]	–	1.843 ± 0.175
NIR	Tufts (H vs. T) [17]	–	1.575 ± 0.338
	StyleGAN2 [36]	–	2.047 ± 0.119
	DB-StyleGAN2 [18]	–	2.071 ± 0.102
	SFace [29]		1.553 ± 0.286
	Mix-SFace [29]		1.557 ± 0.239
		✓	1.642 ± 0.242
	ArcBiFaceGAN		1.666 ± 0.206
		✓	1.680 ± 0.199

Enhancement and Manipulation Schemes



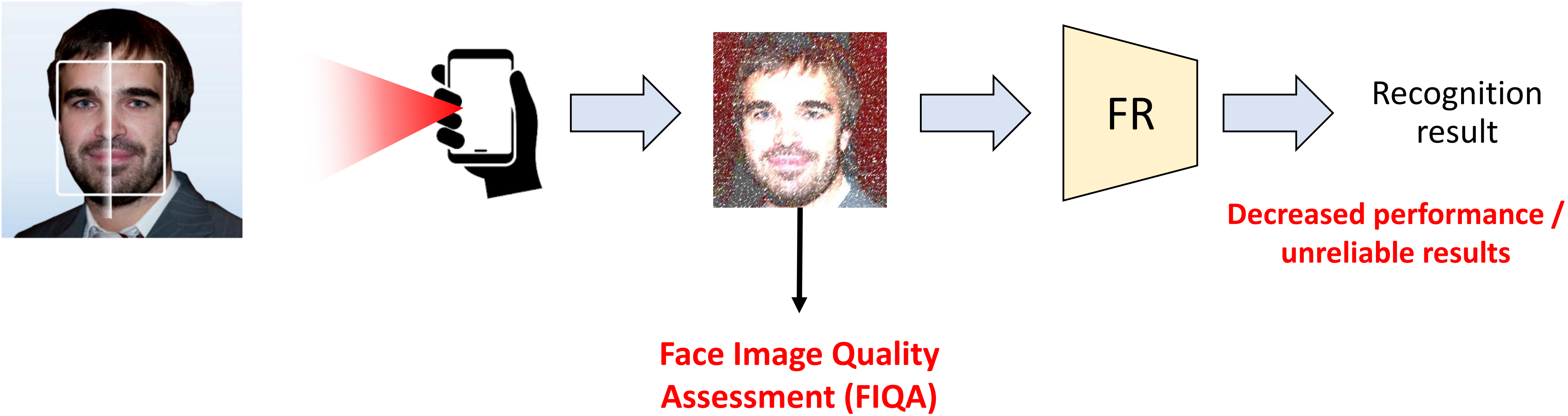




(a) SCFace probe images (g) Degraded DroneSURF galleries (h) Hallucinated DroneSURF probes

Face Image Quality Assessment

- Face Image Quality Assessment (FIQA)
  - Key use case in **Face Recognition**



  - Reject sample, use quality for recognition, etc.

Face Image Quality Assessment

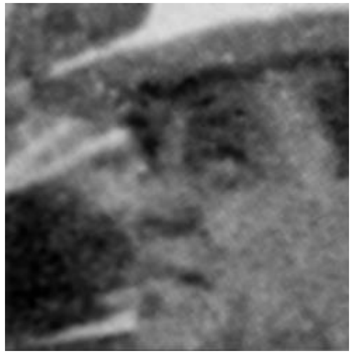
- Face Image Quality Assessment (FIQA)
  - ISO/IEC 29794-1 differentiates between three quality aspects [1]:
    - **Character** – attributes inherent to the biometric characteristic (faces) being acquired (e.g., skin texture, scars, etc.)
    - **Fidelity** – reflects the degree of similarity with the source face, e.g., blurred or noisy faces have low fidelity
    - **Utility** – the fitness of a sample for face recognition, e.g., confidence in the recognition result

High perceptual quality

High face image quality

High perceptual quality

Low face image quality



Low perceptual quality

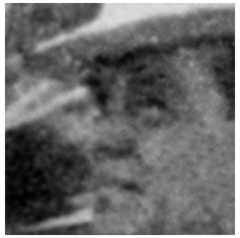
Low face image quality

Face Image Quality Assessment

FIQA

q

Unified Quality Score



$q_1 > q_2$

Higher scores mean better quality

Quality components:

- Face alignment
- Background uniformity
- Sharpness
- Mouth Closed
- Eyes open
- Head Pose Frontalness
- Oclussion level
- …

Face Image Quality Assessment

Existing solutions

Analytical:
Compute scores directly from the input sample.

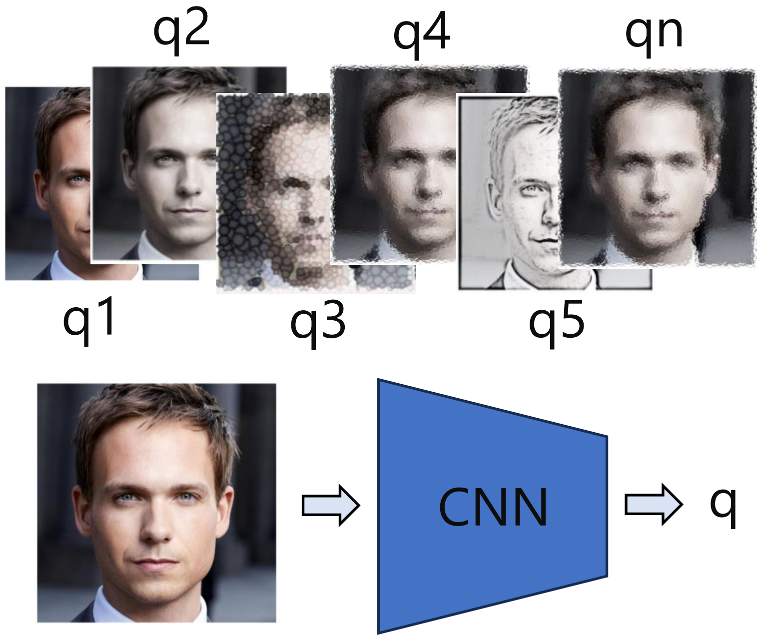
Sharpness

Level of occlusion

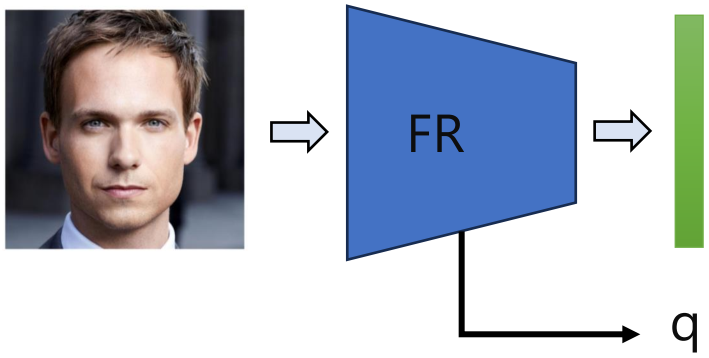
Frontalness

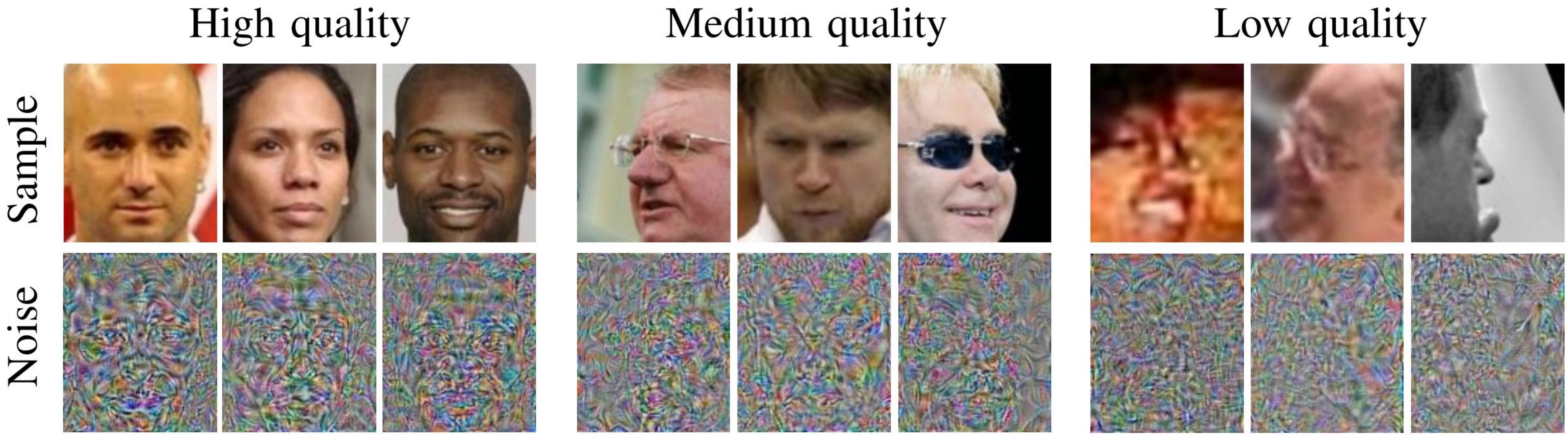
Unified QS

Regression based:
Train regression model on extracted pseudo-quality labels.



Model based:
Combine recognition and quality estimation tasks.





Part II-a: FaceQAN

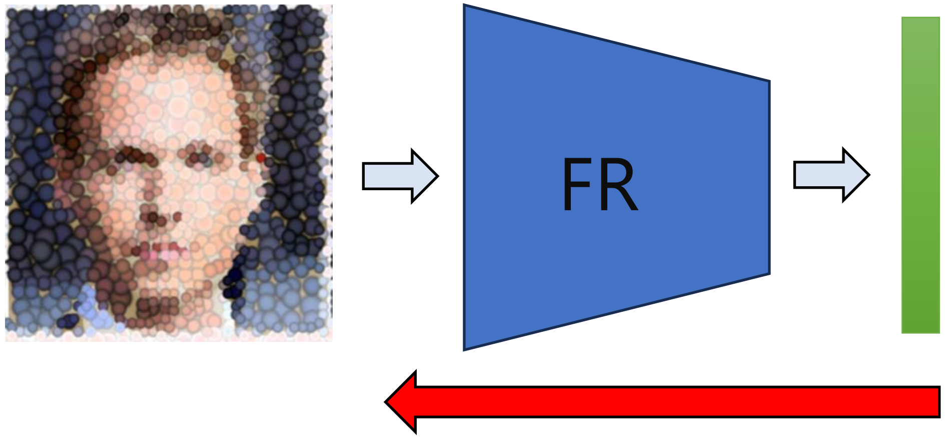
Ž. Babnik, P. Peer, V. Štruc, FaceQAN: Face Image Quality Assessment Through Adversarial Noise Exploration, ICPR 2022

Available on arXiv (Dec 2022):
https://arxiv.org/abs/2212.02127

GitHub: https://github.com/LSIbabnikz/FaceQAN

FaceQAN

- How difficult is it to perturb an embedding of a facial image?

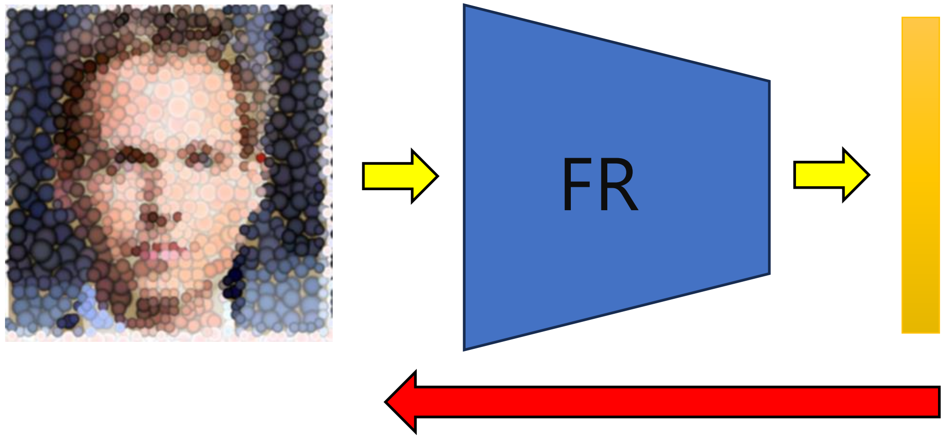


Perturbation through

Adversarial attack

FaceQAN

- How difficult is it to perturb an embedding of a facial image?

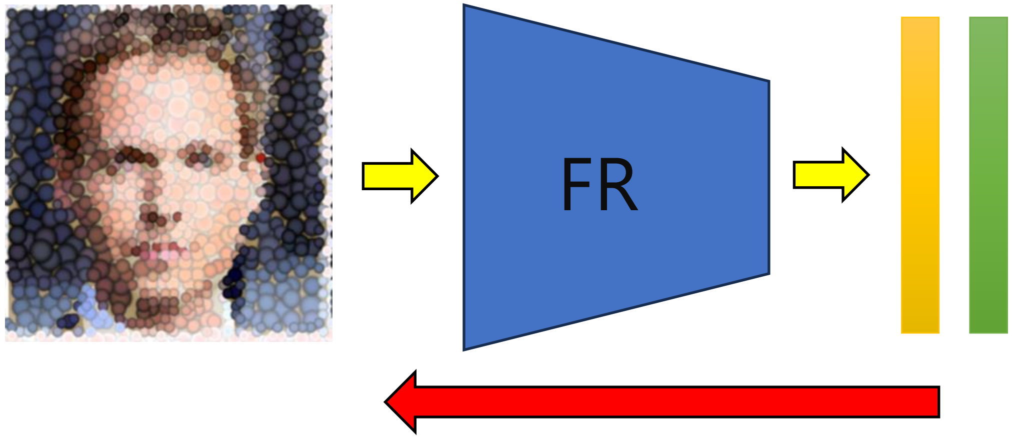


Perturbation through

Adversarial attack

FaceQAN

- How difficult is it to perturb an embedding of a facial image?



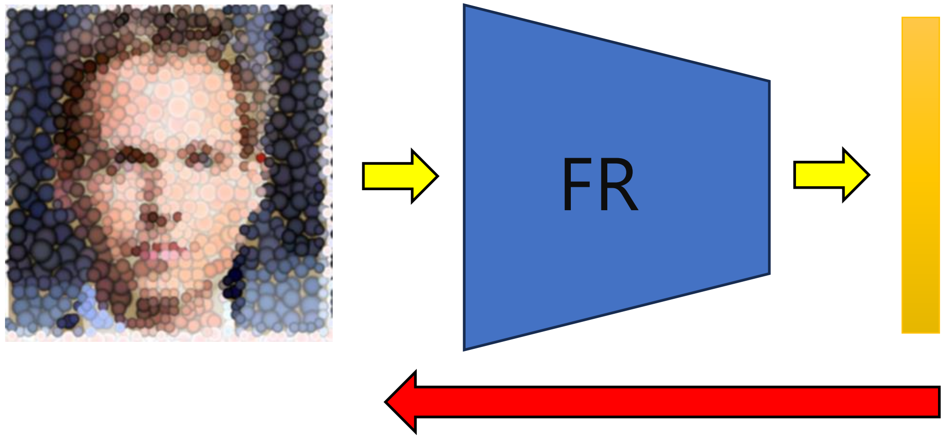
Perturbation through

Adversarial attack

- Stability of embedding => Face Image Quality (FIQ)

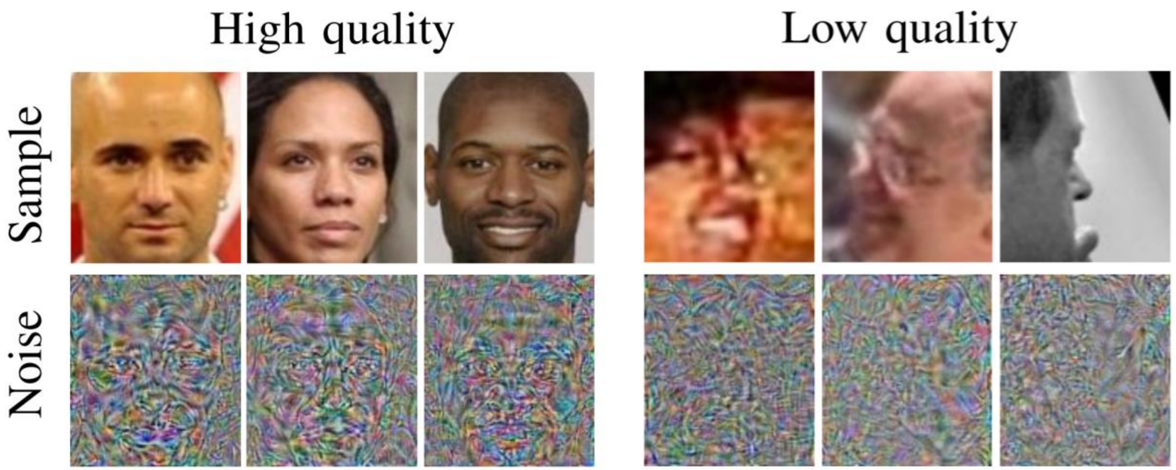
FaceQAN

- How difficult is it to perturb an embedding of a facial image?



Perturbation through

Adversarial attack

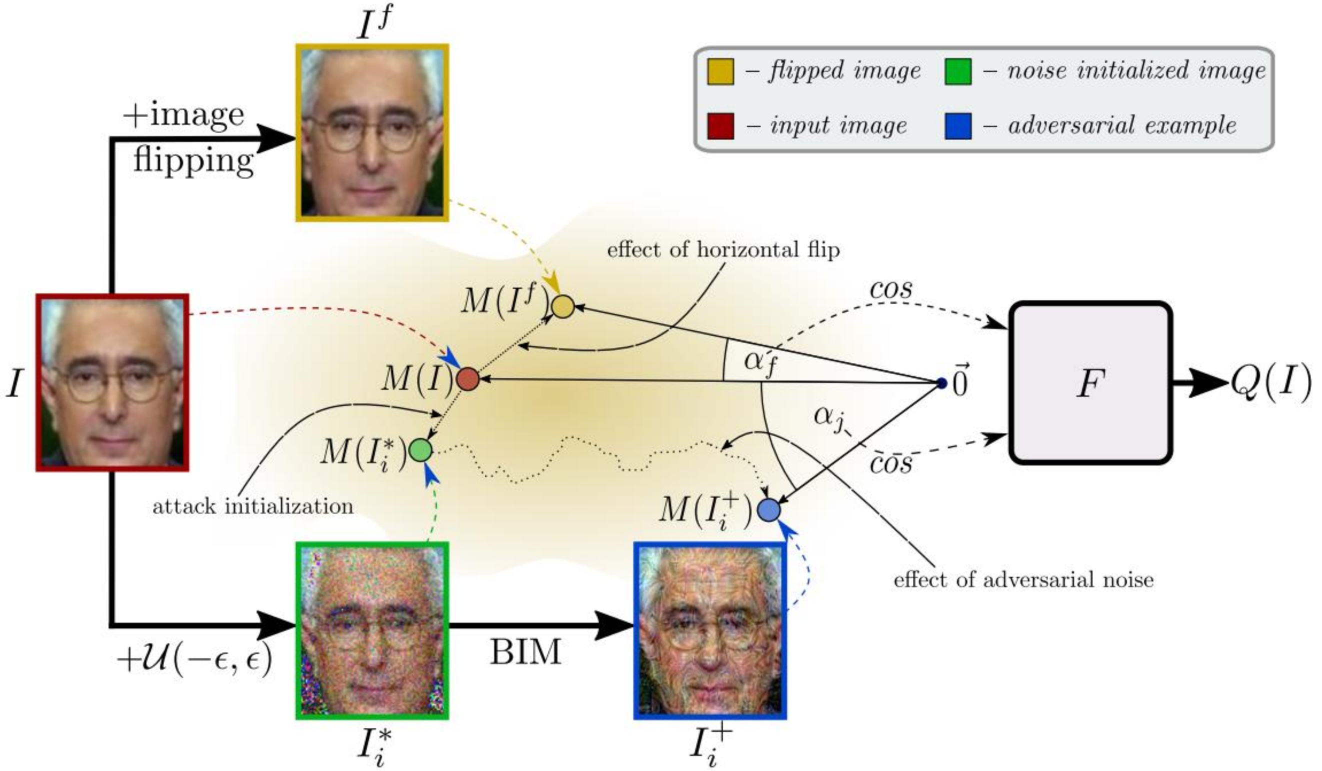


- Stability of embedding => Face Image Quality (FIQ)

FaceQAN

- **Key Components**

- Operates in embedding space
- Similarity-based adversarial attack (FGSM+BIM)
- Fixed number of iterations
- Pose modeled explicitly
- Repeated *k* times



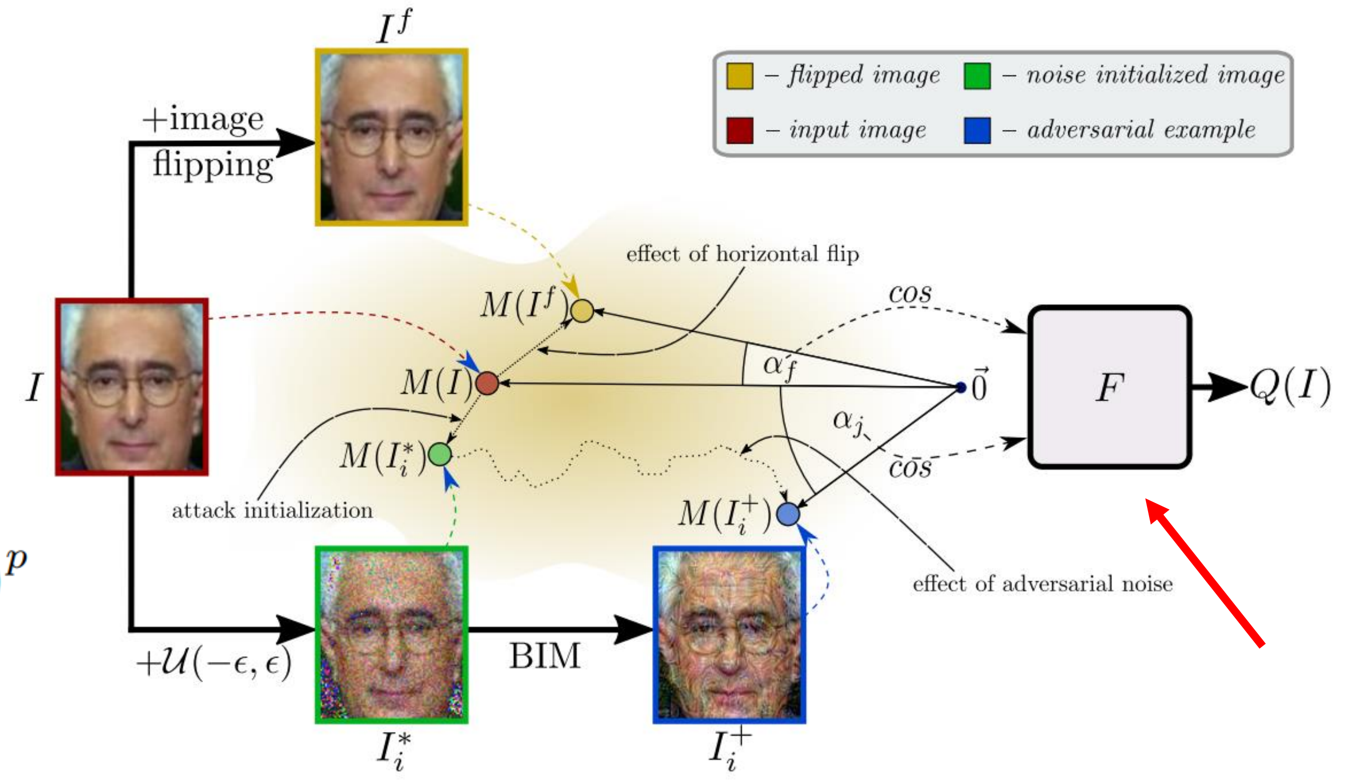
$$L(M(I^*), y) = 1 - \frac{M(I^*)^T \cdot y}{\|M(I^*)\| \|y\|}$$

FaceQAN

- **Key Components**
- Quality score calculation

$$\cos\alpha_i = S_i = \frac{y^T \cdot y_i^+}{\|y\|\|y_i^+\|}, \quad \cos\alpha_f = s_f = \frac{y^T \cdot y_f}{\|y\|\|y_f\|}$$

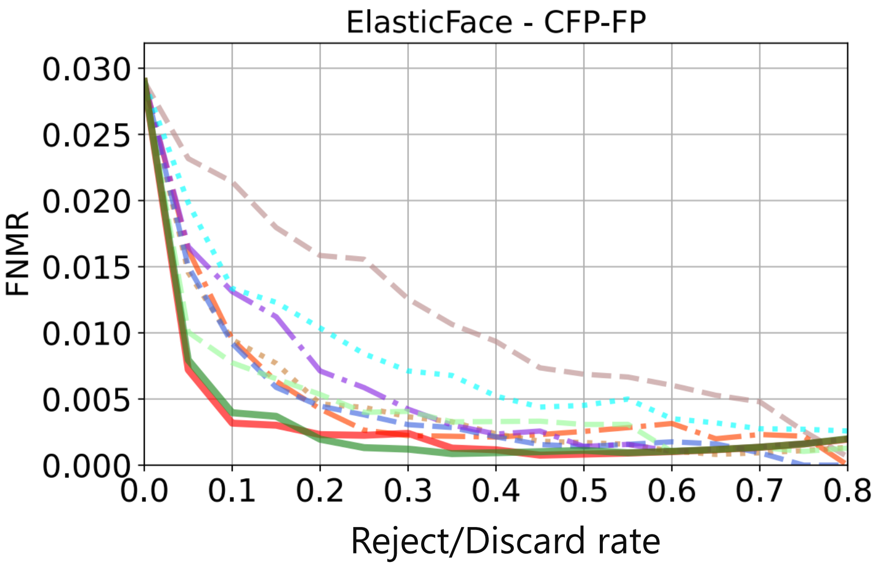
$$q_{adv} = \frac{\mu_S + 1}{2} \cdot \lfloor(1 - \sigma_S)\rfloor_{[0,1]}, \; Q = (\; q_{adv} \cdot s_f)^p$$



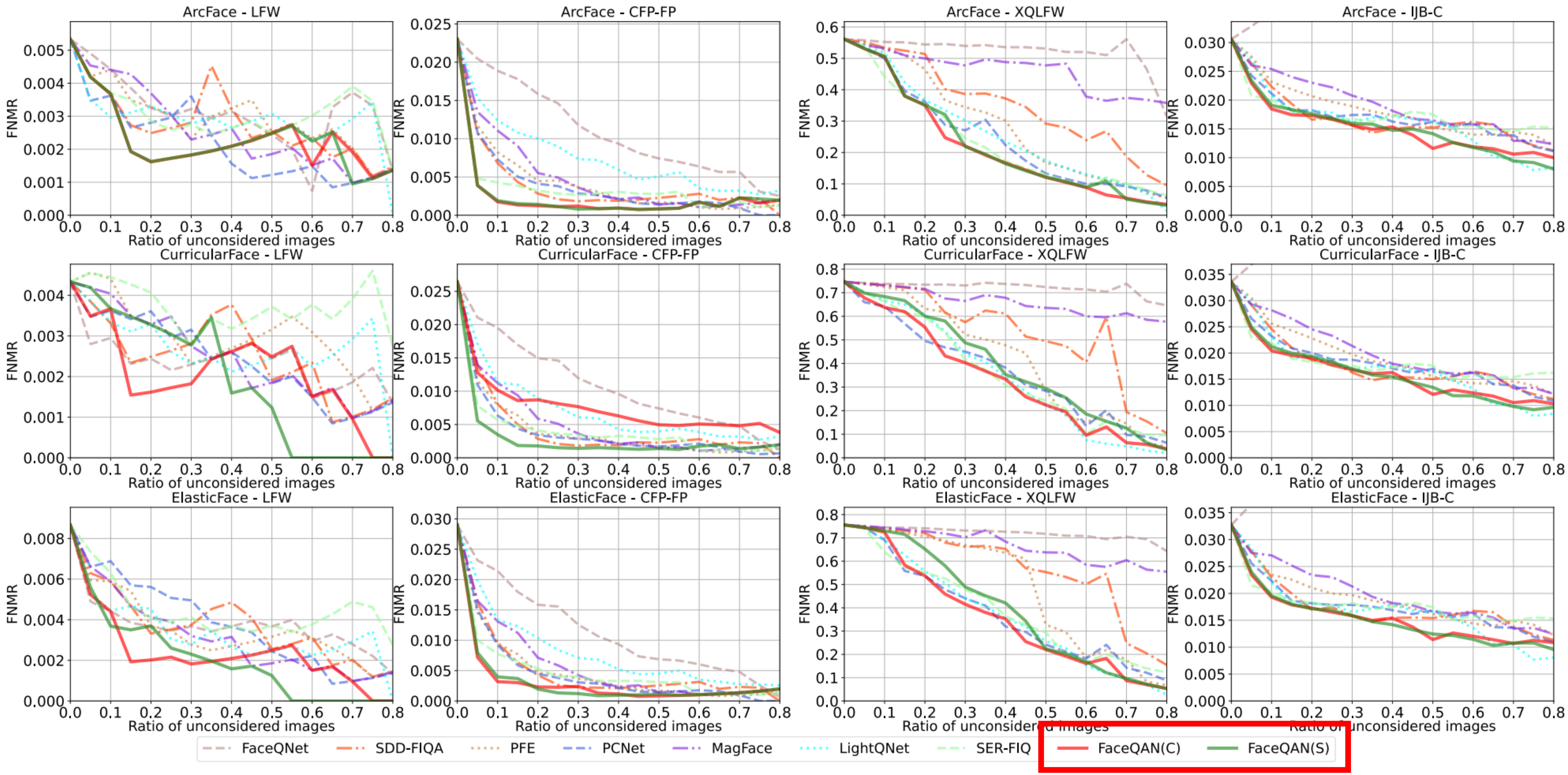
$$L(M(I^*), y) = 1 - \frac{M(I^*)^T \cdot y}{\|M(I^*)\|\|y\|}$$

FaceQAN - Experiments

- **Standard evaluation approach:**
  - **Error-Versus Discard Characteristic (EDC) curves**
    - *Show how the performance of FR models improves with increasing discard rates.*
  - **Area Under the Curve (AUC)**
    - *Lower value indicates better performance.*
  - **Benchmarks:**
    - *LFW, CFP-FP. XQLFW, IJB-C*
  - **FR models:**
    - *ArcFace, CurricularFace, ElasticFace – Same-Model, Cross-Model*
  - **FIQA methods:**
    - *FaceQnet, SDD-FIQA, PFE, PCNet, MagFace, LightQNet, SER-FIQ*



FaceQAN - Experiments



FaceQAN - Experiments

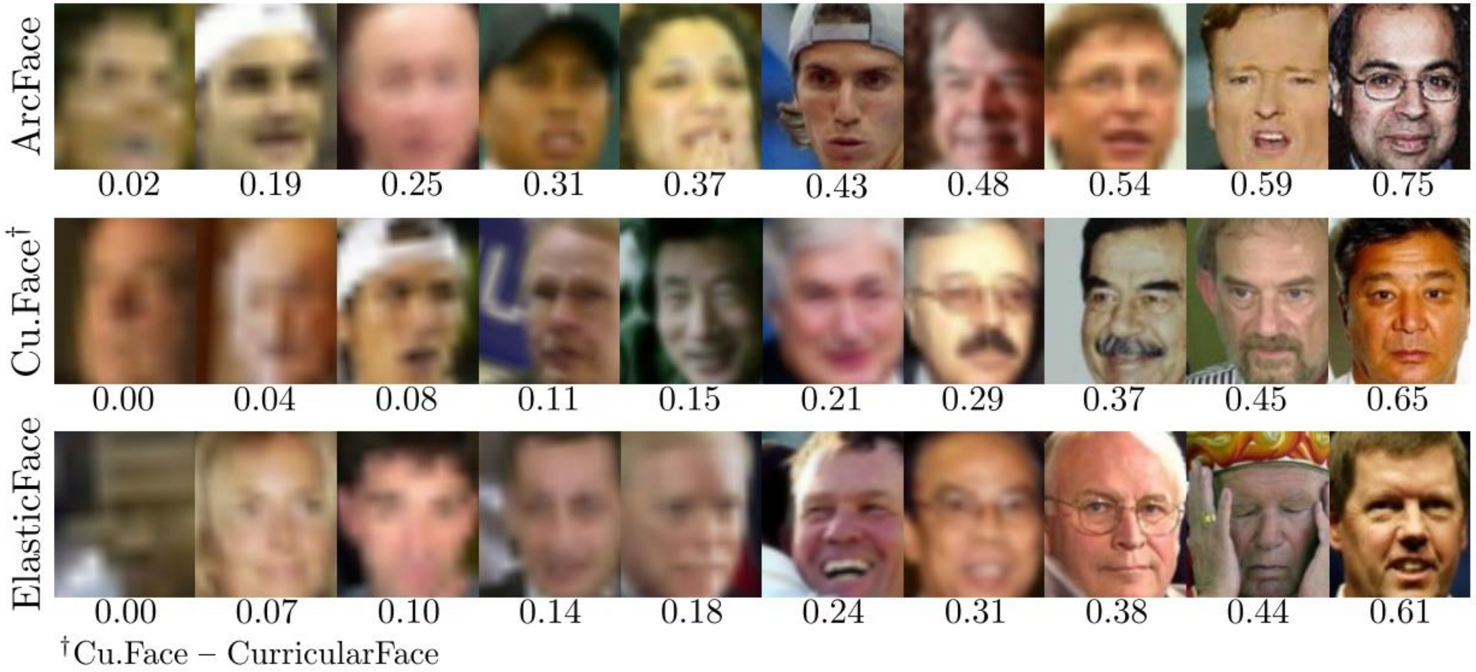


TABLE III
ANALYSIS OF THE TIME COMPLEXITY OVER CFP-FP USING COSFACE

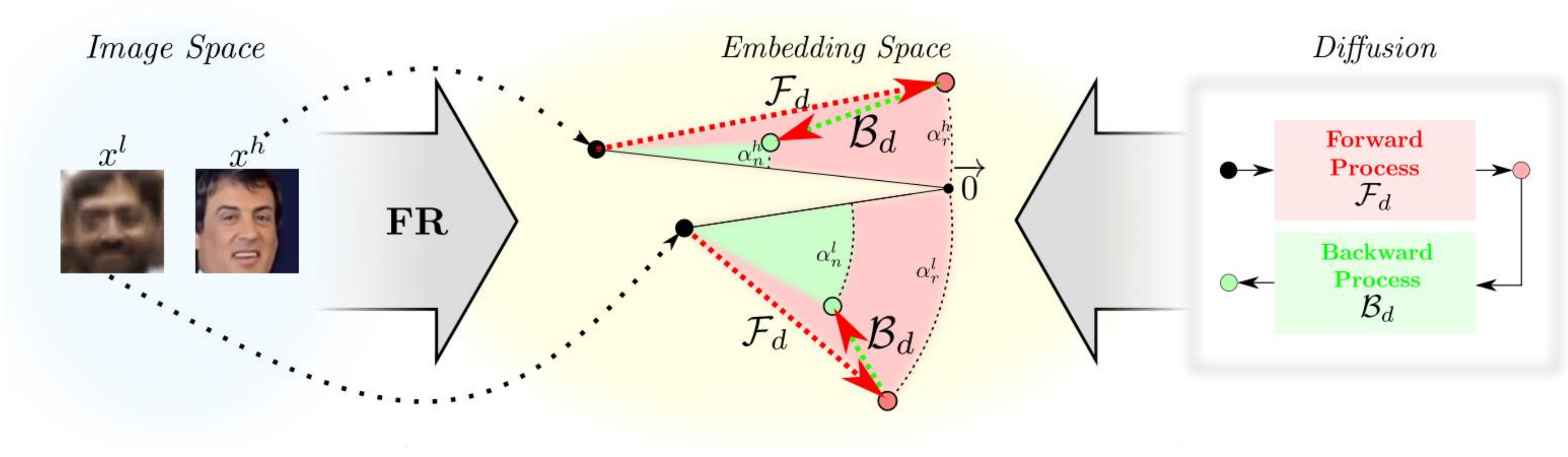
Complexity		**FaceQnet**	**SDD-FIQA**	**PFE**	**PCNet**	**MagFace**	**LightQnet**	**SER-FIQ**
t [s]	μ	0.0432	0.0006	0.0493	0.0175	0.0011	0.0535	0.1125
	σ	0.0026	0.0004	0.0275	0.0004	0.0004	0.0468	0.0418

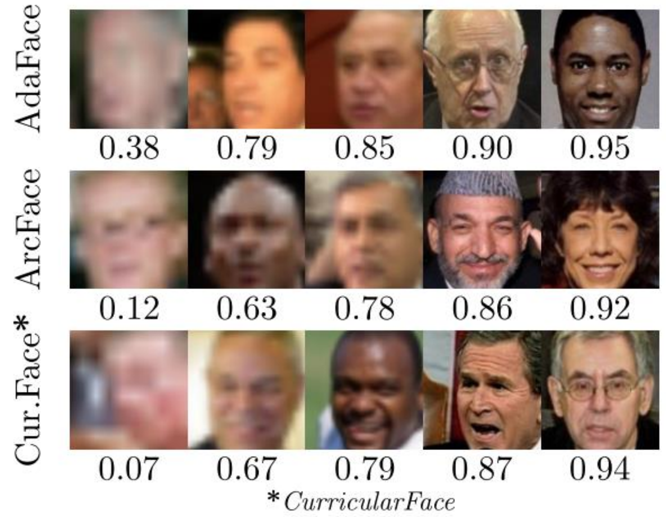
Complexity	**FaceQAN (ours)**				
	$k = 2$	$k = 5$	$^{\dagger}k = 10$	$k = 50$	$k = 100$
t [s] $(\mu \pm \sigma)$	0.21 ± 0.019	0.23 ± 0.018	0.30 ± 0.006	1.00 ± 0.027	1.87 ± 0.031

† Configuration used in experiments

FaceQAN – Take Aways

- Adversarial attacks are powerful for probing embedding stability
- FaceQAN achieves SOTA performance
  - Cross-model and Same-model setting
  - Across multiple benchmarks
  - Across multiple FR models
- Limitations:
  - Embedding stability = partial characterization of quality
  - Requires backprop over FR model - Very slow





Part II-b: DifFIQA

Ž. Babnik, P. Peer, V. Štruc, DifFIQA: Face Image Quality Assessment Using Denoising Diffusion Probabilistic Models, IJCB 2023

Available on arXiv (May 2023): https://arxiv.org/abs/2305.05768

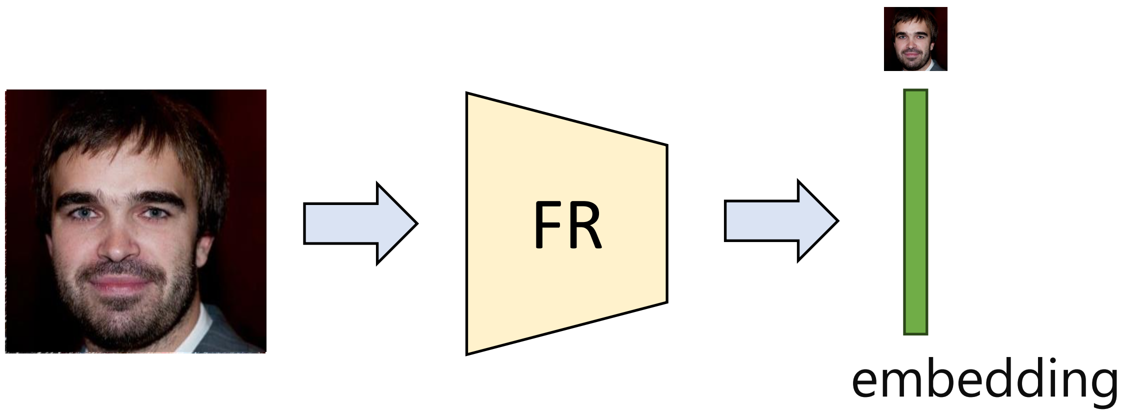
GitHub: https://github.com/LSIbabnikz/DifFIQA

DifFIQA: Quality Assessment Using Denoising Diffusion Probabilistic Models

- Face Image Quality Assessment based on two concepts:

Perturbation robustness

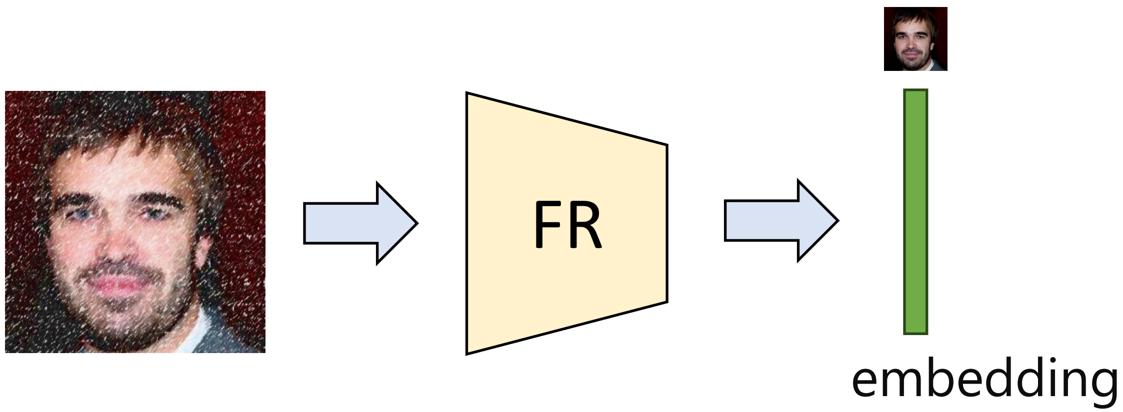
Reconstruction quality



DifFIQA: Quality Assessment Using Denoising Diffusion Probabilistic Models

- Face Image Quality Assessment based on two concepts:

Perturbation robustness



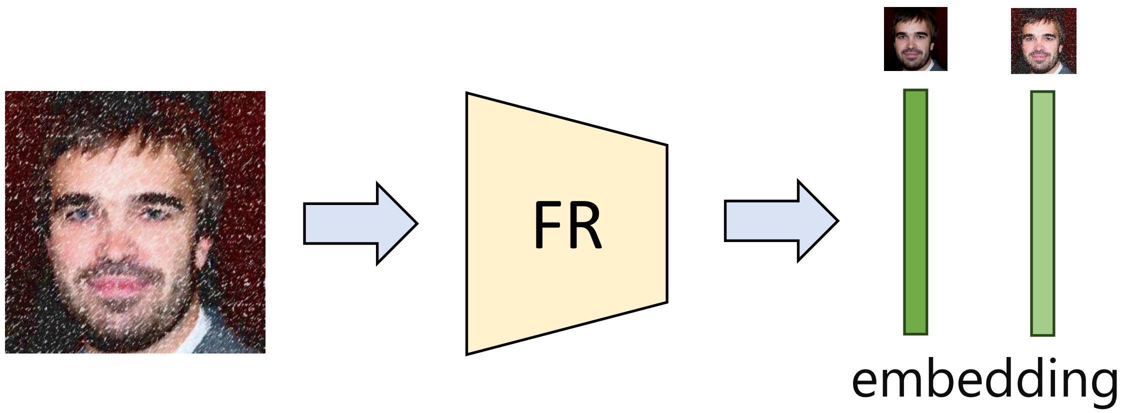
Reconstruction quality

DifFIQA: Quality Assessment Using Denoising Diffusion Probabilistic Models

- Face Image Quality Assessment based on two concepts:

Perturbation robustness

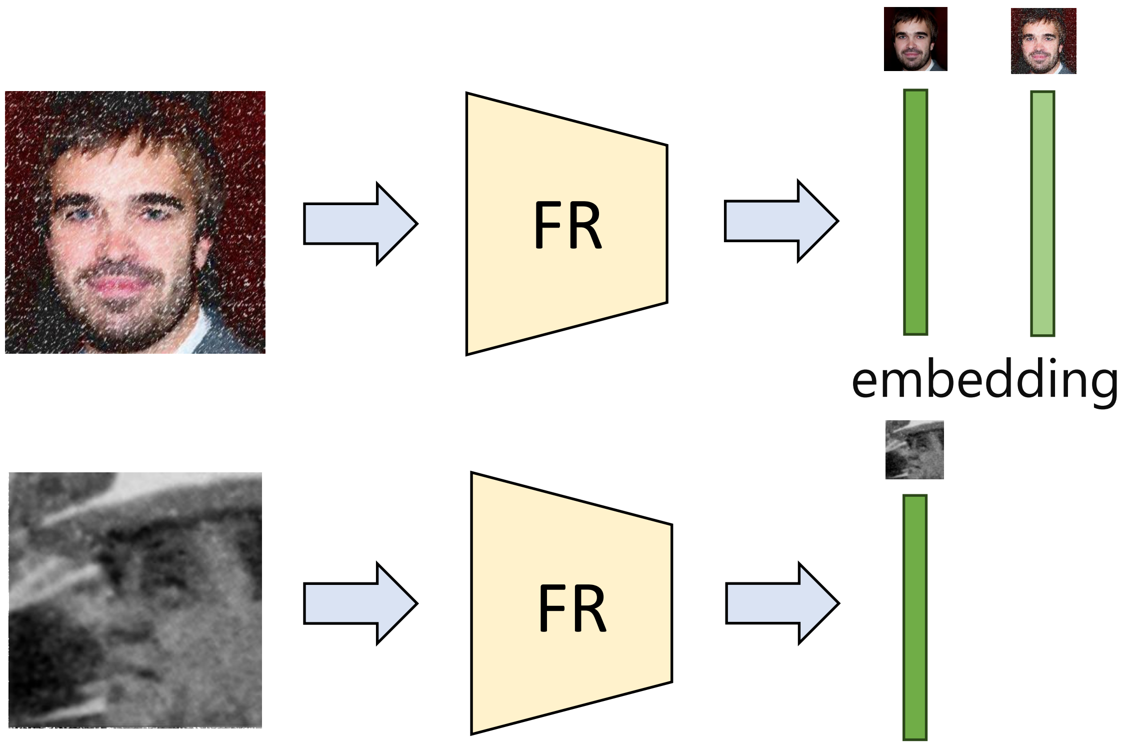
Reconstruction quality



DifFIQA: Quality Assessment Using Denoising Diffusion Probabilistic Models

- Face Image Quality Assessment based on two concepts:

Perturbation robustness



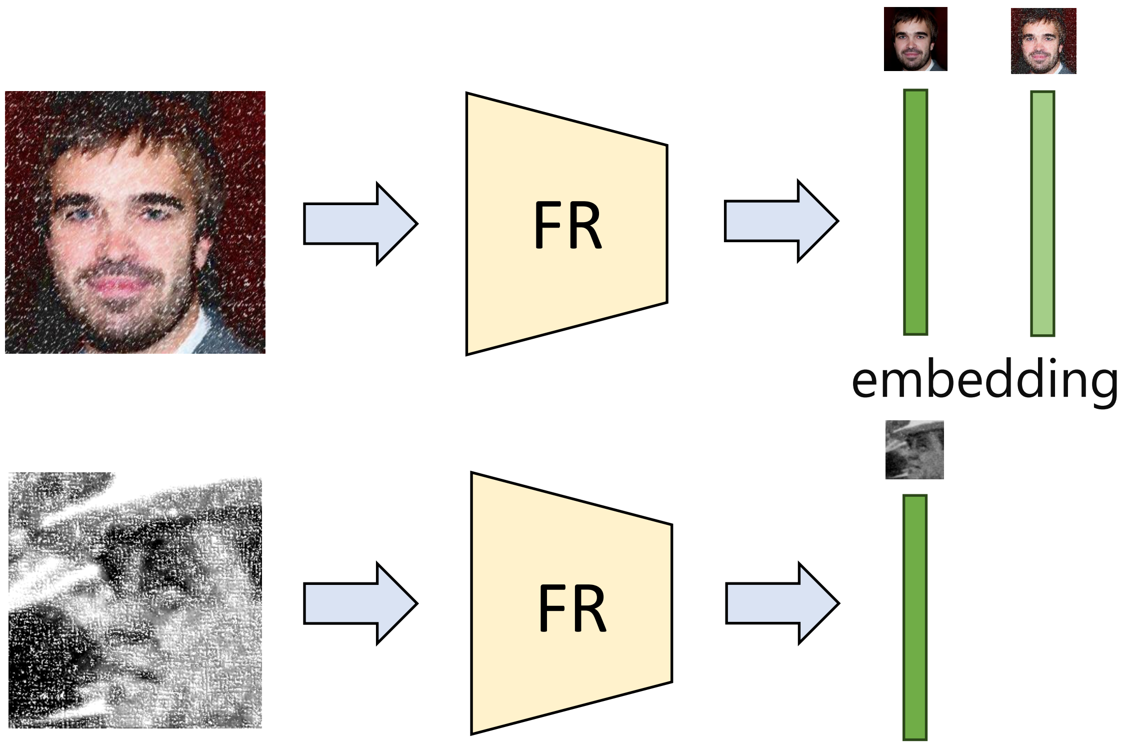
Reconstruction quality

DifFIQA: Quality Assessment Using Denoising Diffusion Probabilistic Models

- Face Image Quality Assessment based on two concepts:

Perturbation robustness

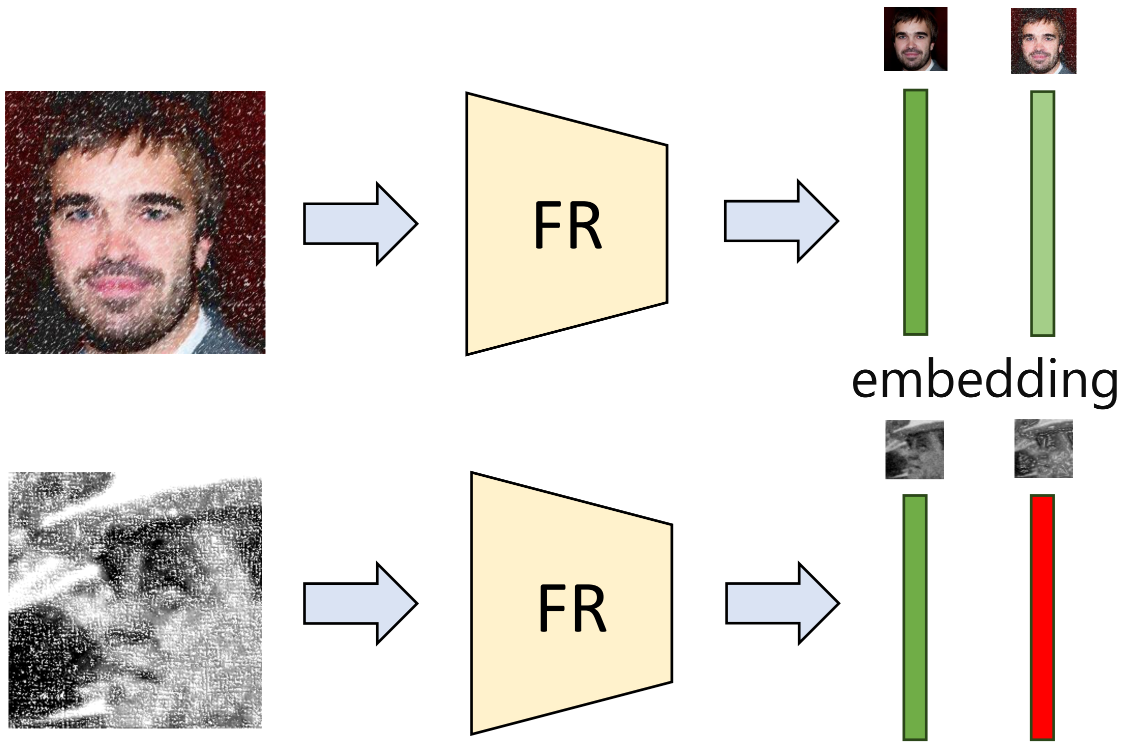
Reconstruction quality



DifFIQA: Quality Assessment Using Denoising Diffusion Probabilistic Models

- Face Image Quality Assessment based on two concepts:

Perturbation robustness

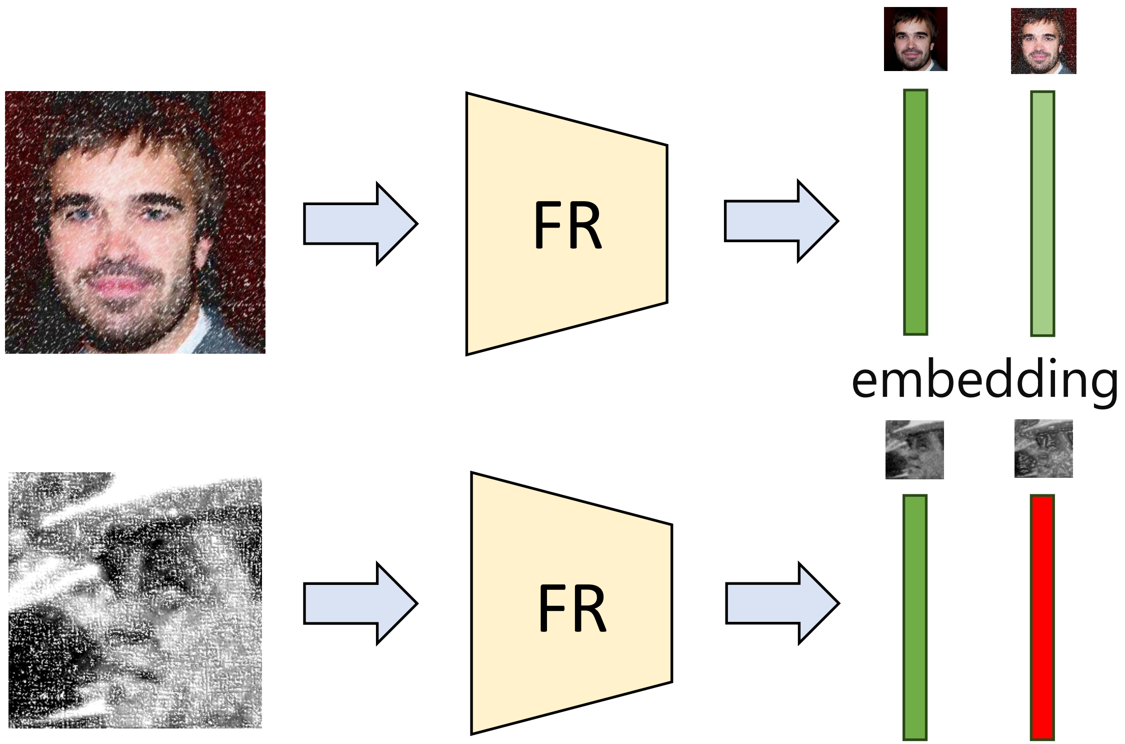


Reconstruction quality

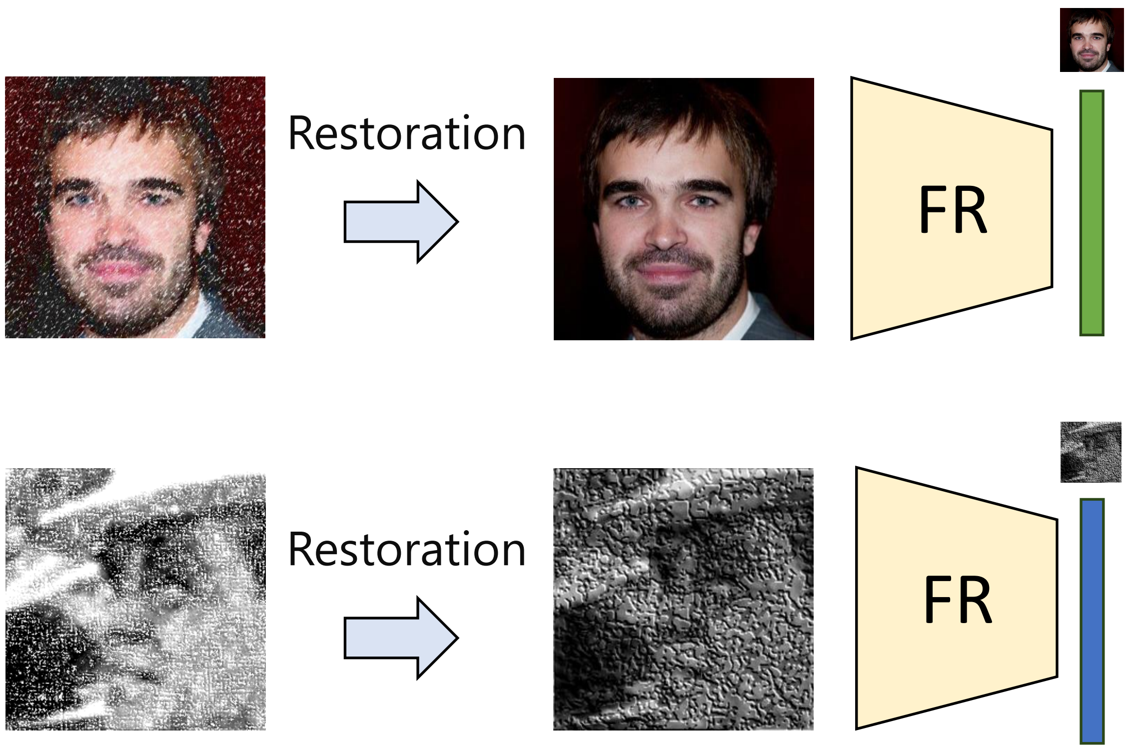
DifFIQA: Quality Assessment Using Denoising Diffusion Probabilistic Models

- Face Image Quality Assessment based on two concepts:

Perturbation robustness



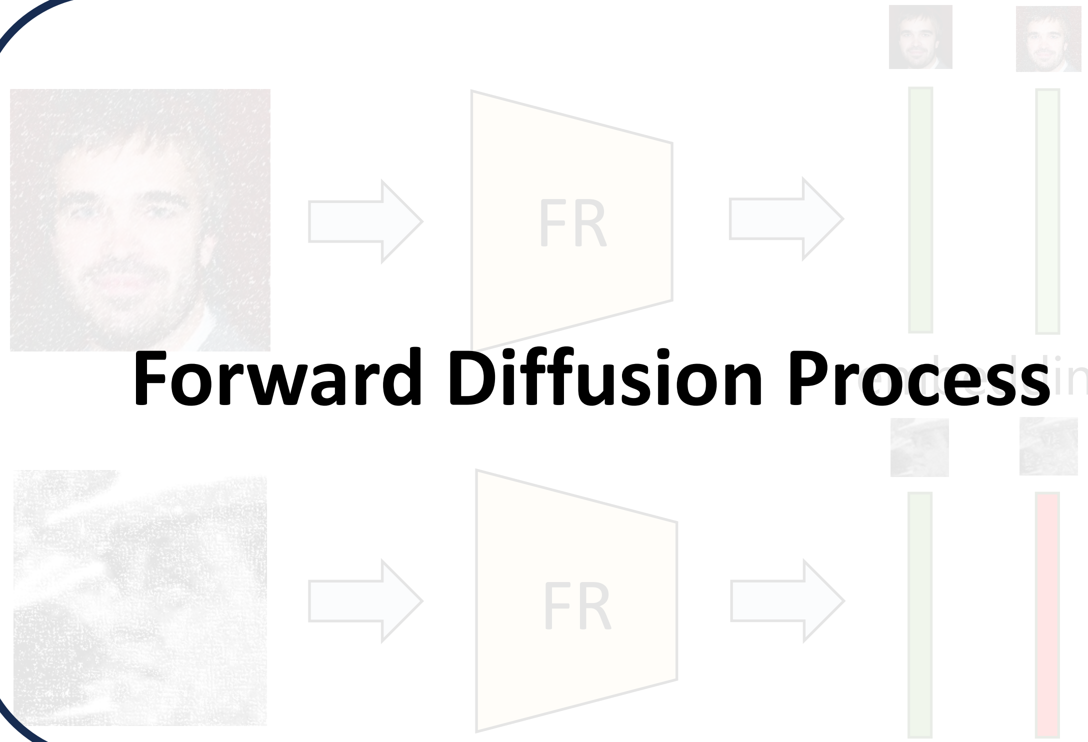
Reconstruction quality



DifFIQA: Quality Assessment Using Denoising Diffusion Probabilistic Models

- Face Image Quality Assessment based on two concepts:

Perturbation robustness



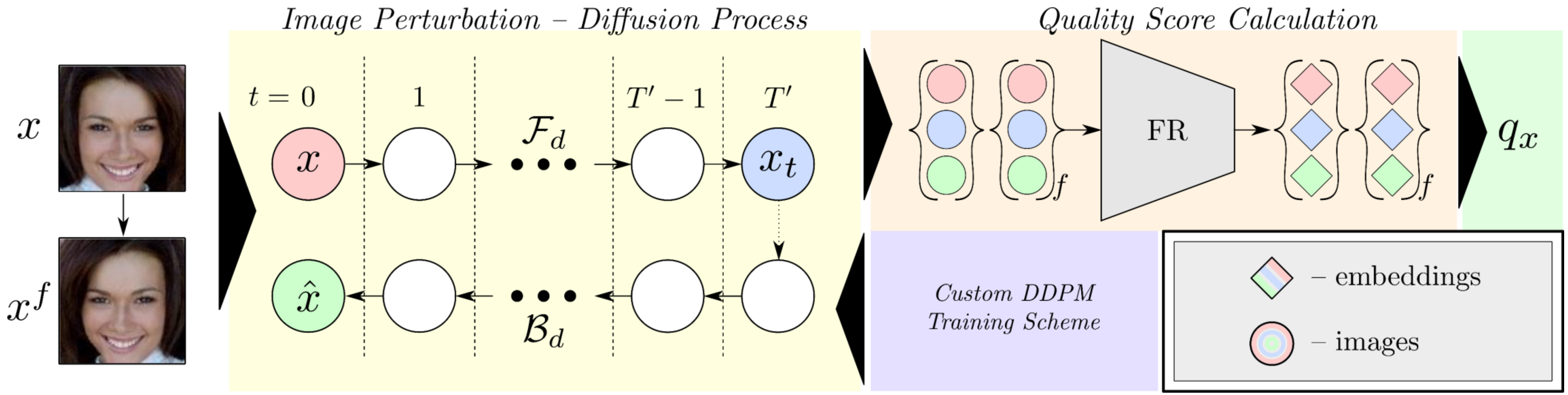
Forward Diffusion Process

Reconstruction quality

Backward Diffusion Process

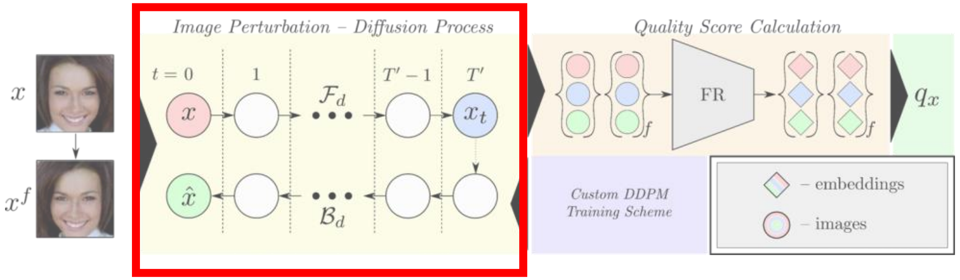
DifFIQA: Quality Assessment Using Denoising Diffusion Probabilistic Models

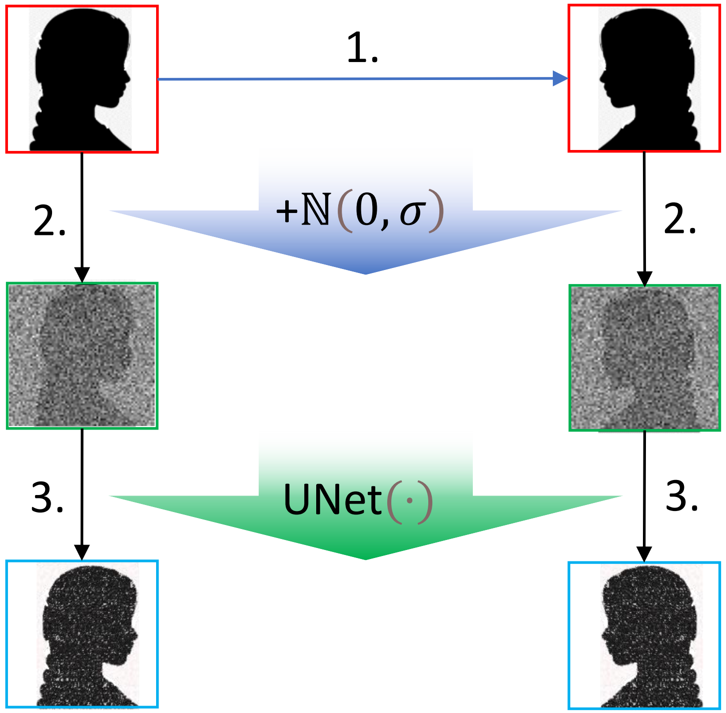
- DifFIQA (high-level) overview:
  - Step 1: Forward and backward diffusion (i.e., perturbations and reconstructions)
  - Step 2: Quality score calculation



DifFIQA: Quality Assessment Using Denoising Diffusion Probabilistic Models

- Step 1: Forward and backward diffusion





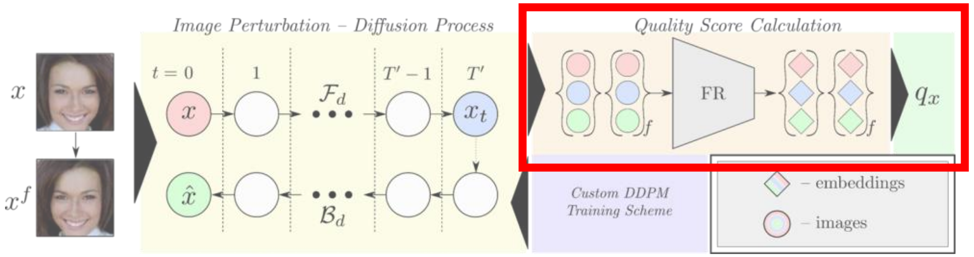
Flip image

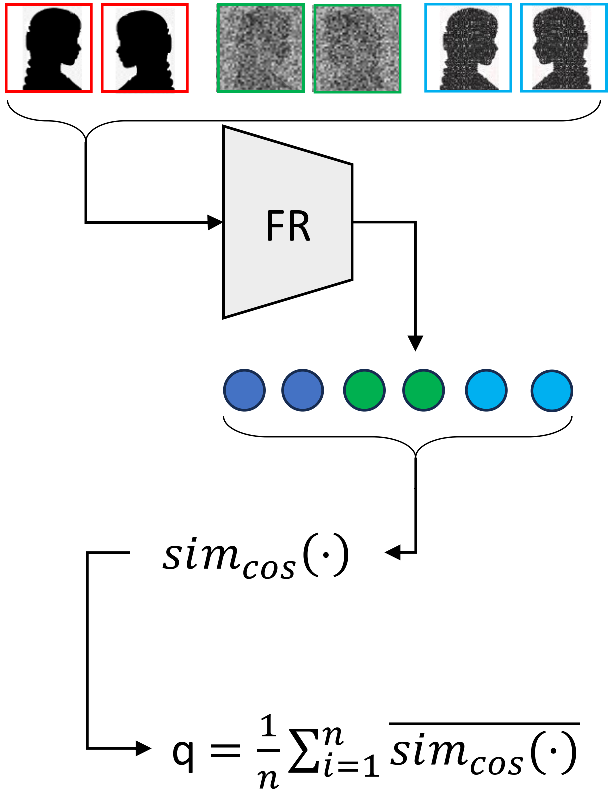
Forward diffusion – adds noise

Backward diffusion – generate reconstructions

DifFIQA: Quality Assessment Using Denoising Diffusion Probabilistic Models

- Step 2: Quality score calculation

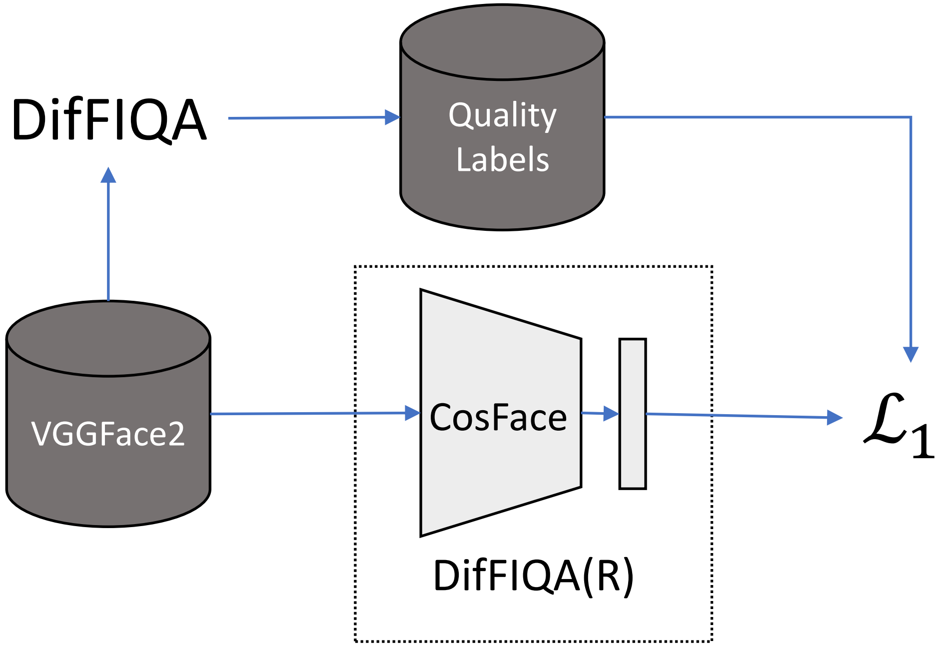




1. **Construct image embeddings:**
   - *using any target FR model,*
   - *for all three image pairs.*
2. **Calculate embedding similarity:**
   - *using cosine similarity,*
   - *comparing the input image embedding to all others.*
3. **Repeat process n-times and calculate average.**

DifFIQA: Quality Assessment Using Denoising Diffusion Probabilistic Models

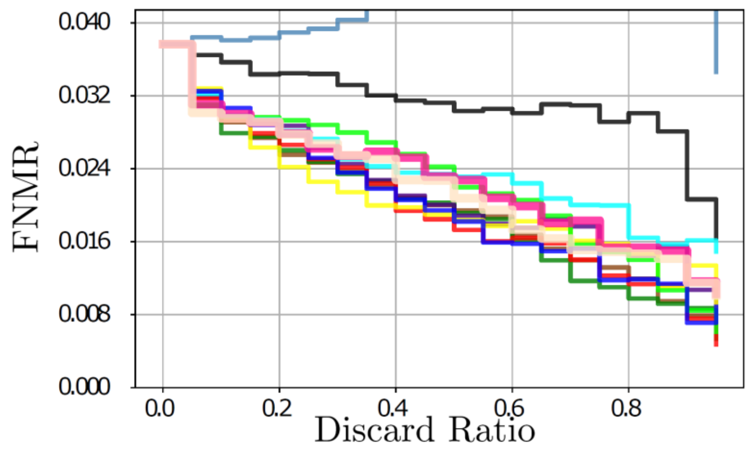
- Problem: Diffusion is slow!
- DifFIQA(R) - Knowledge distillation



Fine-tune a pretrained FR model using pseudo-quality labels extracted using DifFIQA.

Experimental Setup

- **Standard evaluation approach:**
  - **Error-Versus Discard Characteristic (EDC) curves**
    - *Show how the performance of FR models improves with increasing discard rates.*
  - **partial Area Under the Curve (pAUC)**
    - *Lower value indicates better performance.*
    - *Limit the discard ratio to (0.2) or (0.3).*
  - **Benchmarks:**
    - *Adience, CALFW, CFP-FP, CPLFW, IJB-C, LFW, XQLFW*
  - **FR models:**
    - *AdaFace, ArcFace, CurricularFace, CosFace*
  - **FIQA methods:**
    - *FaceQnet, SDD-FIQA, PFE, PCNet, MagFace, LightQNet, SER-FIQ, FaceQAN, CR-FIQA, FaceQgen*



Results (pAUC)

- **Average results over all benchmarks and FR models**

Results using discard ratio of (0.2).

FaceQnet [17]	SDD-FIQA [33]	PFE [39]	PCNet [44]	MagFace [31]	LightQNet [7]	SER-FIQ [40]	FaceQAN [3]	CR-FIQA [5]	FaceQgen [15]	DifFIQA	DifFIQA(R)
0.9458	0.8244	0.8197	0.8989	0.8253	0.8183	0.7985	**0.7519**	**0.7567**	0.8527	0.7591	**0.7518**

Results using discard ratio of (0.3).

FaceQnet [17]	SDD-FIQA [33]	PFE [39]	PCNet [44]	MagFace [31]	LightQNet [7]	SER-FIQ [40]	FaceQAN [3]	CR-FIQA [5]	FaceQgen [15]	DifFIQA	DifFIQA(R)
0.9315	0.7483	0.7497	0.8691	0.7635	0.7412	0.7292	0.6847	**0.6800**	0.7954	**0.6822**	**0.6768**

- Knowledge distilation improves the results of the base approach
- DifFIQA(R) outperforms SOTA methods.

Run-time

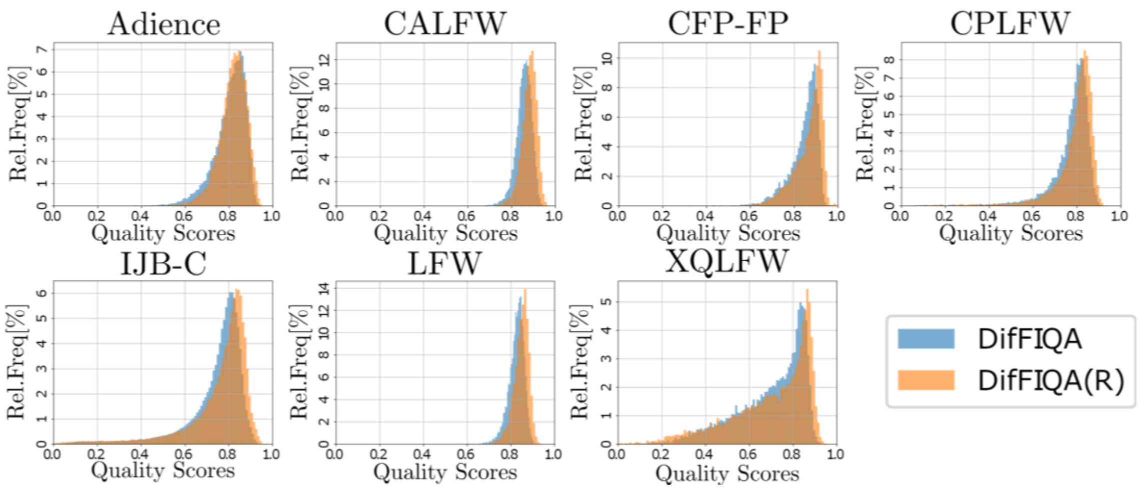
- **Hardware: Intel i9-10900KF CPU, 64 GB RAM in Nvidia 3090 GPU**

FIQA model	**Ours**	
	DifFIQA	**DifFIQA(R)**
Runtime ($\mu \pm \sigma$)	1074.62 ± 11.45	1.24 ± 0.36

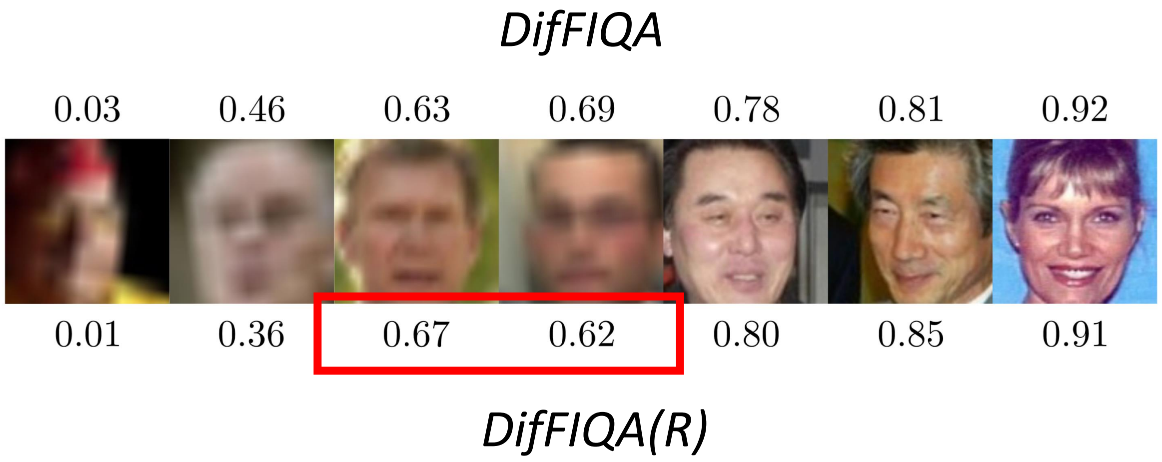
- Speed-up due to knowledge distillation is around **1000x**

DifFIQA and DifFIQA(R) Comparison

Comparison of quality score distributions

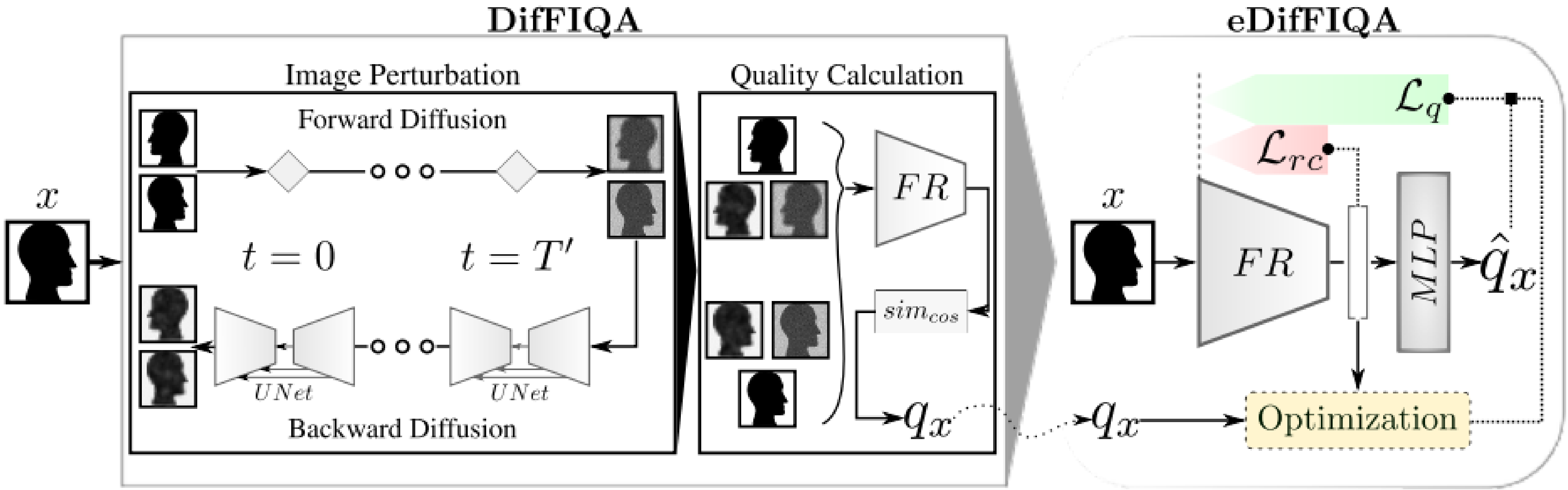


Visual comparison



DifFIQA – Take Aways

- DifFIQA: a new FIQA approach based on DDPMs
- DifFIQA(R): distilled version for fast prediction
- State-of-the-art performance across multiple datasets and face recognition models



Part II-c: eDifFIQA

Ž. Babnik, P. Peer, V. Štruc, eDifFIQA: Towards Efficient Face Image Quality Assessment based on Denoising Diffusion Probabilistic Models, TBIOM 2024

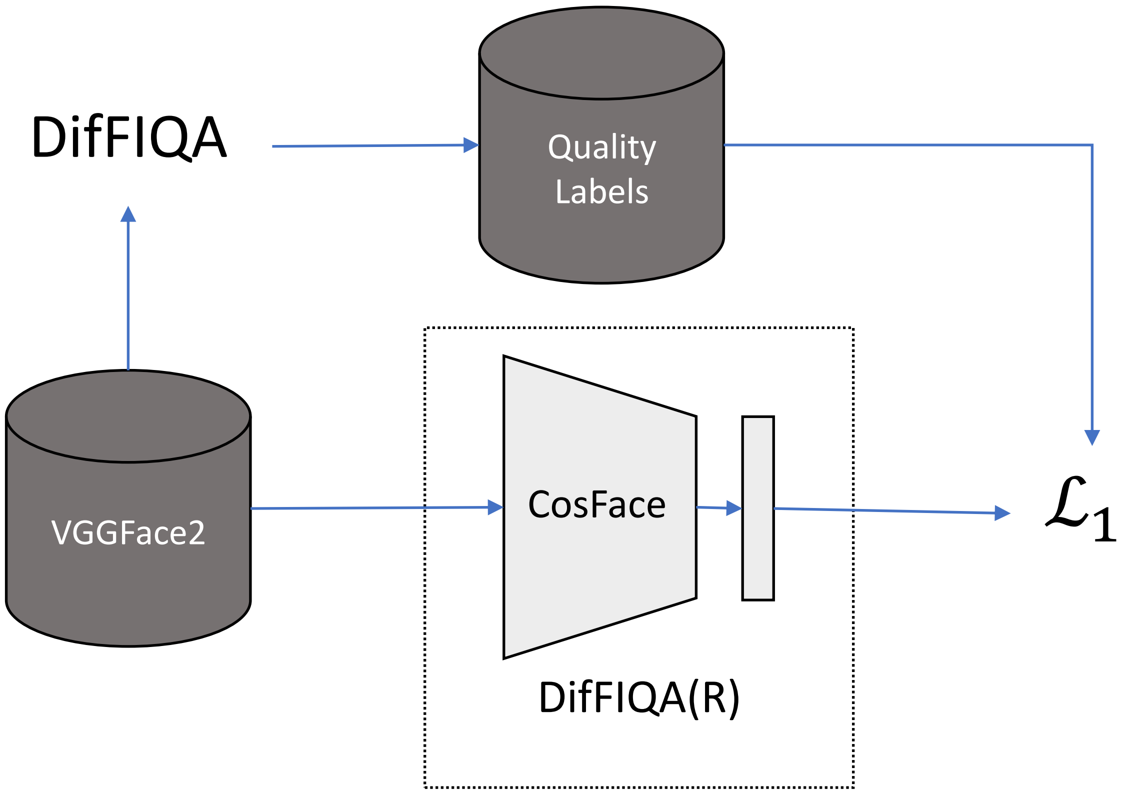
Available on IEEE Xplore in open access (May 2024):
https://ieeexplore.ieee.org/document/10468647/

GitHub: https://github.com/LSIbabnikz/eDifFIQA

eDifFIQA: Efficient Quality Assessment Using Denoising Diffusion Probabilistic Models

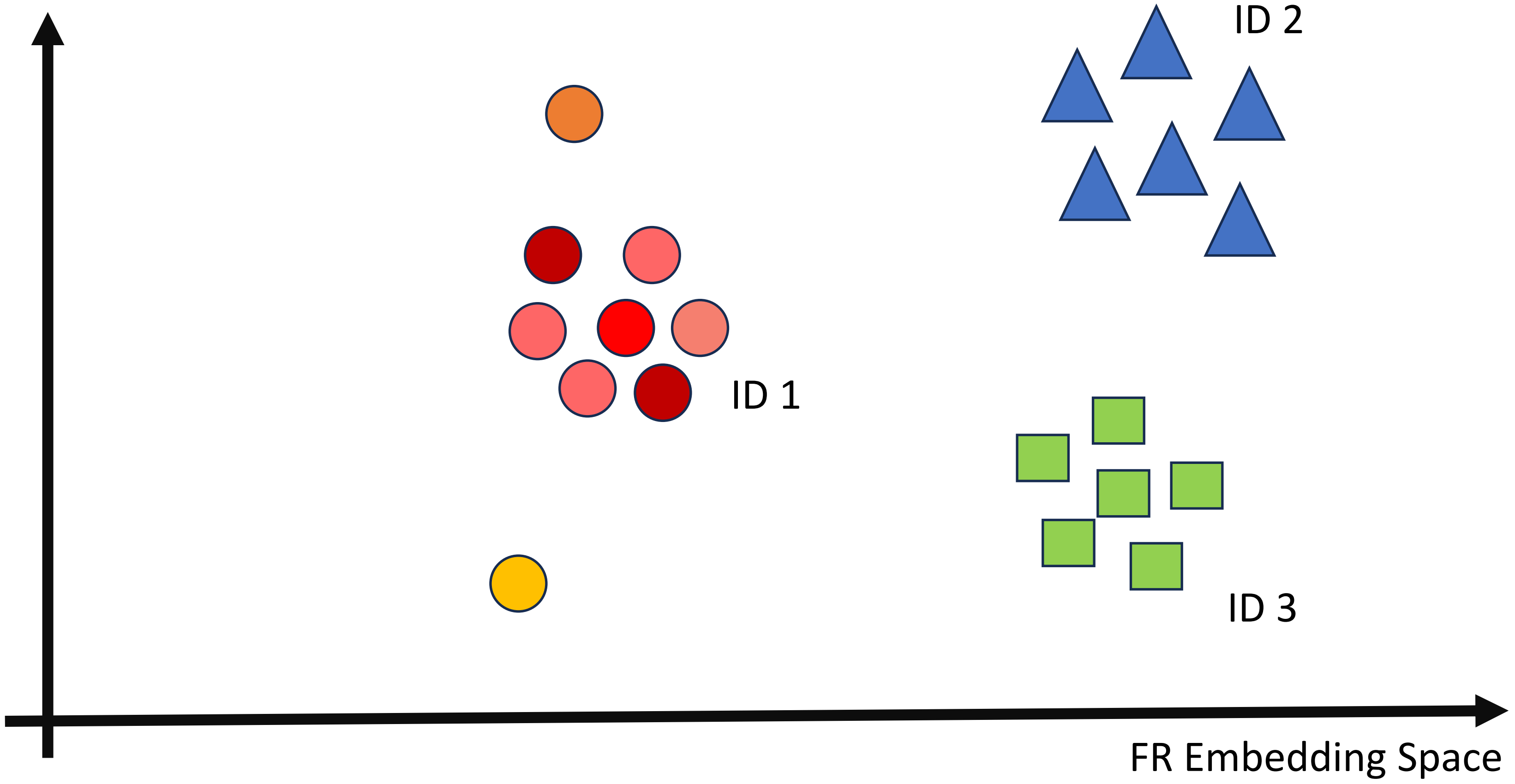
- eDifFIQA
  - Efficiency
  - Performance

- DifFIQA:
  - C1: Perturbation robustness
  - C2: Reconstruction quality



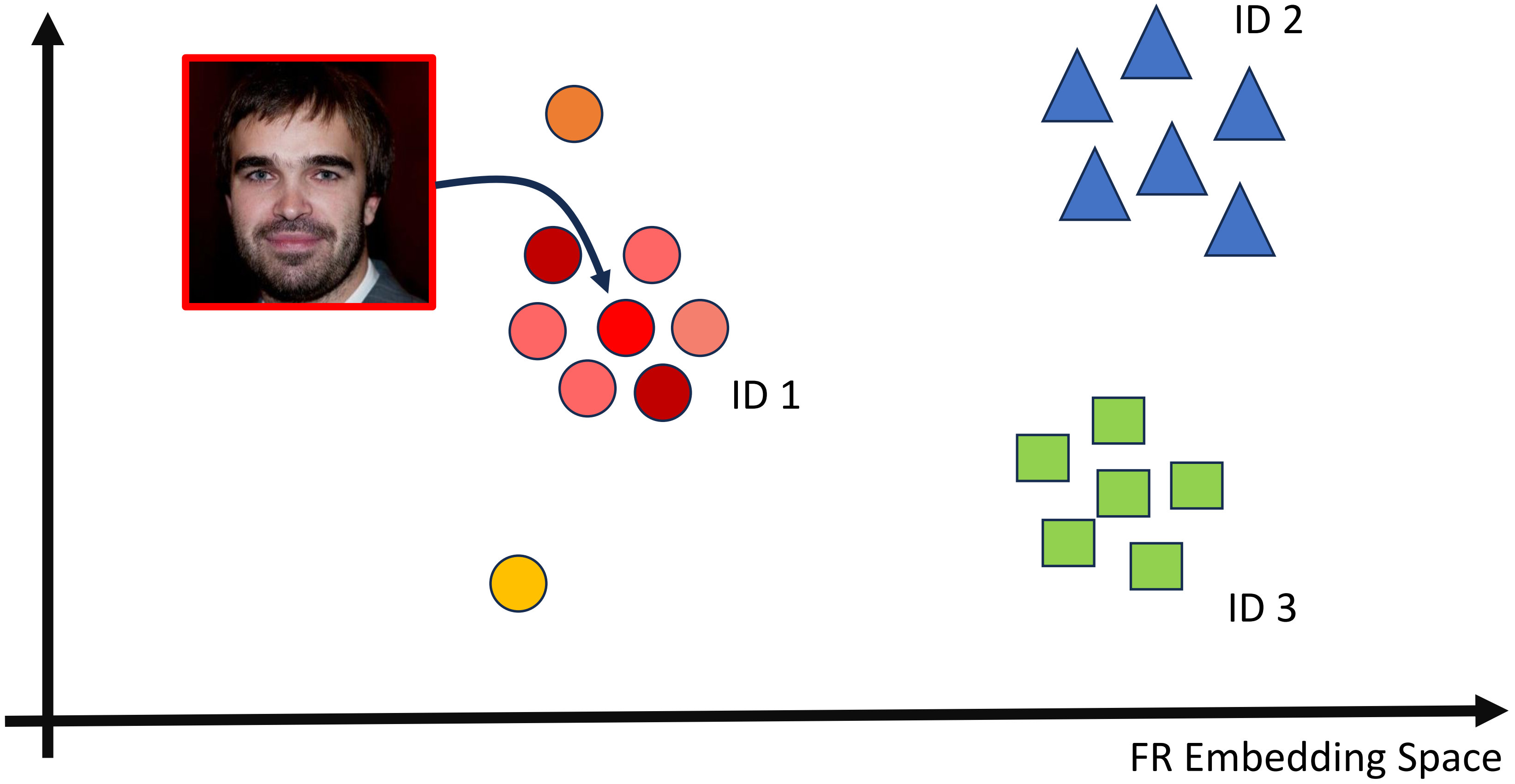
eDifFIQA: Efficient Quality Assessment Using Denoising Diffusion Probabilistic Models

- C3: Relative position in the embedding space



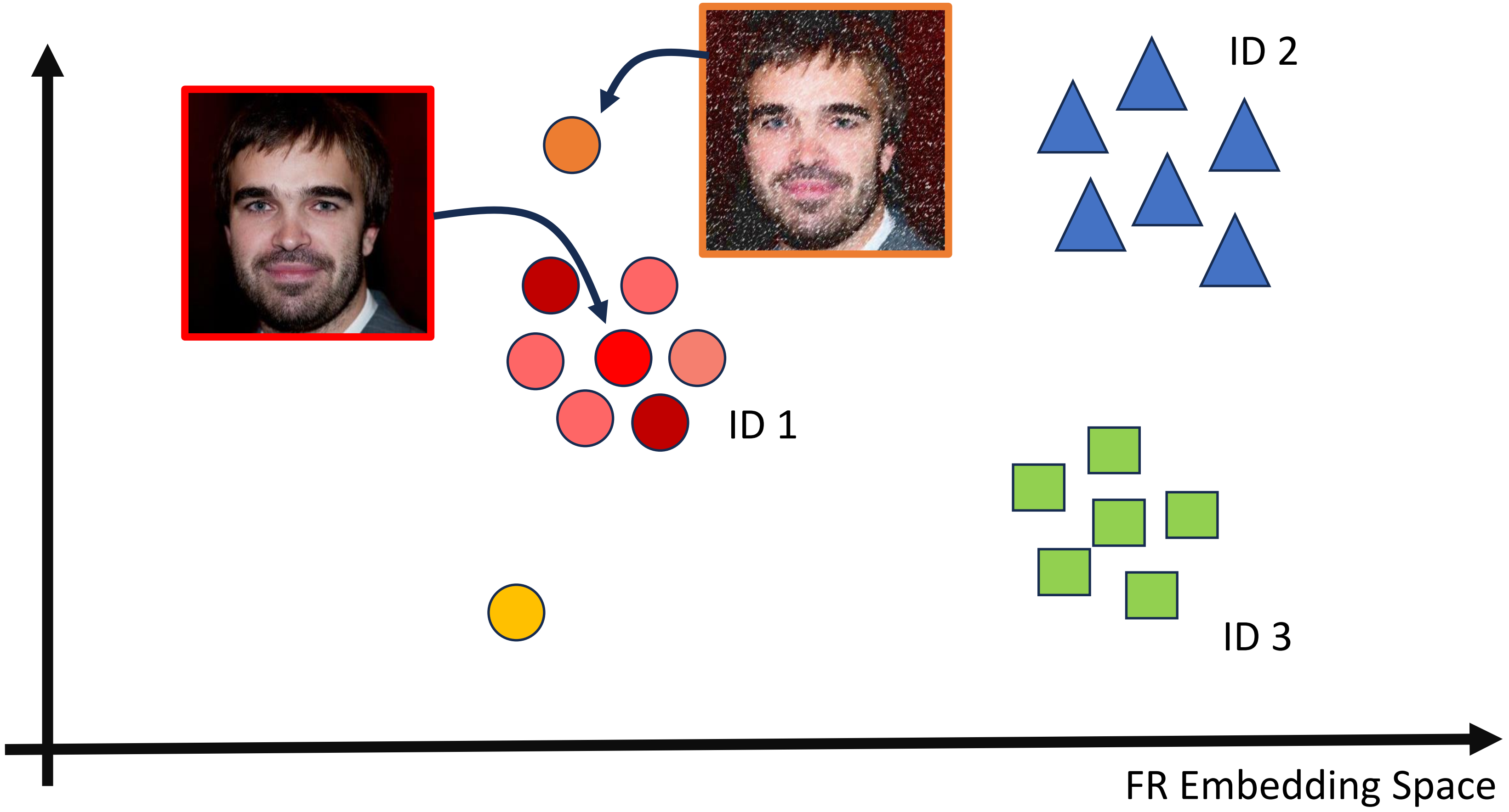
eDifFIQA: Efficient Quality Assessment Using Denoising Diffusion Probabilistic Models

- C3: Relative position in the embedding space



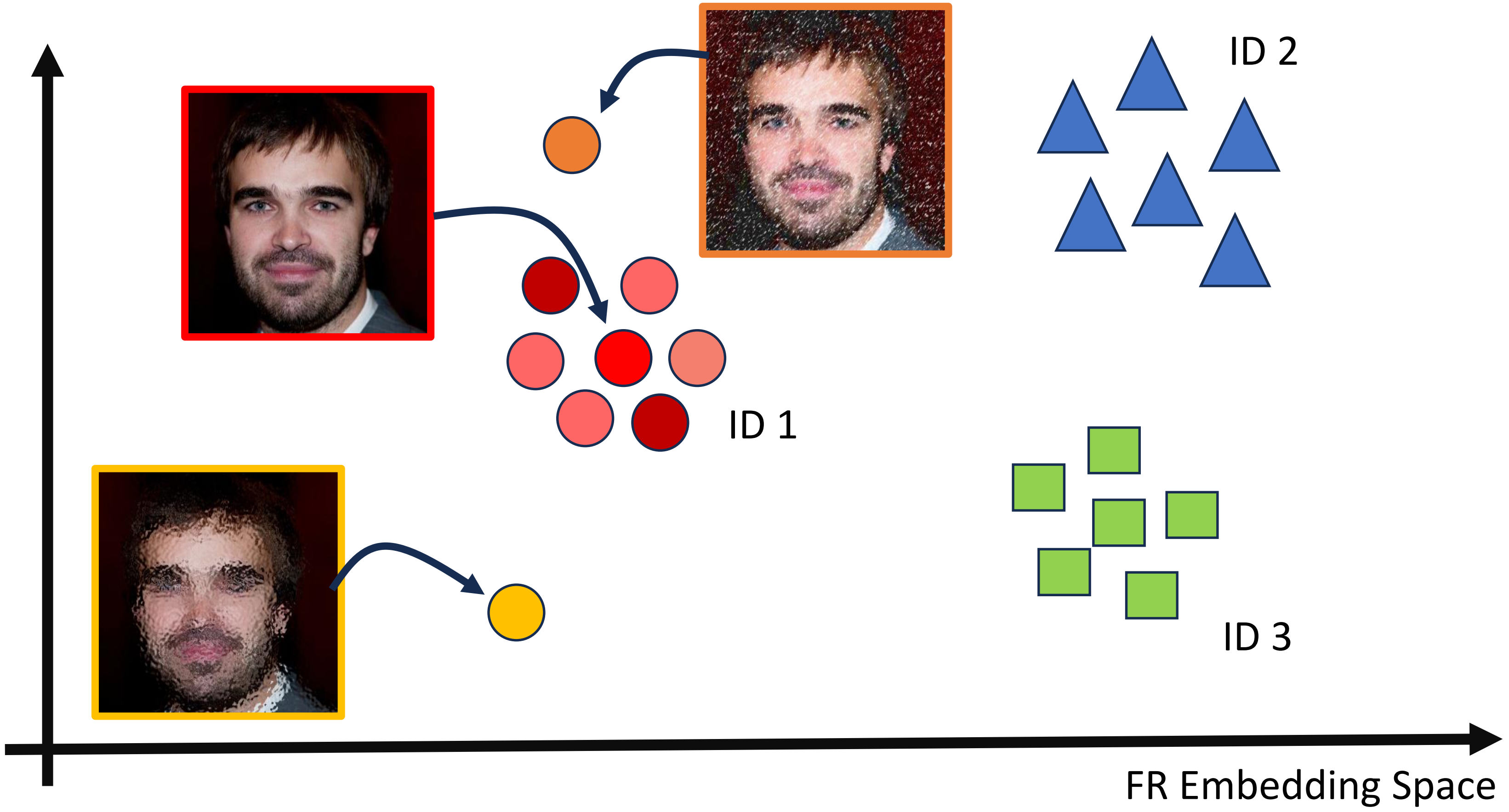
eDifFIQA: Efficient Quality Assessment Using Denoising Diffusion Probabilistic Models

- C3: Relative position in the embedding space



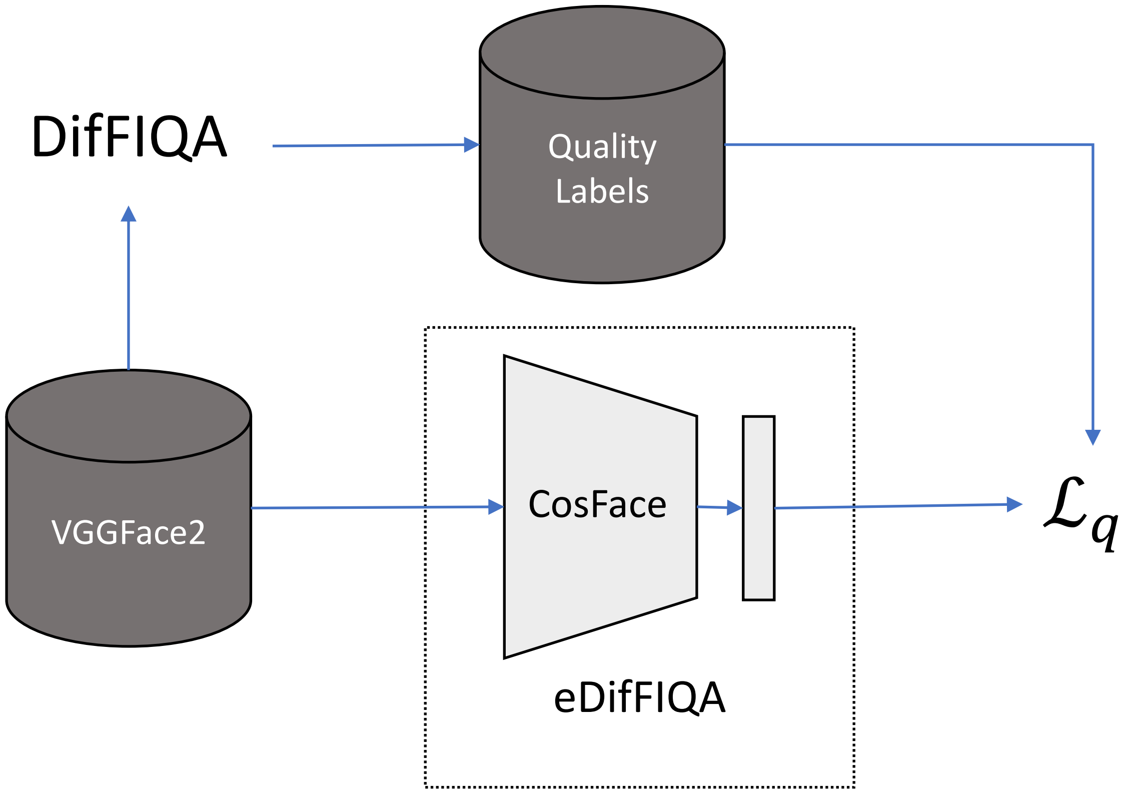
eDifFIQA: Efficient Quality Assessment Using Denoising Diffusion Probabilistic Models

- C3: Relative position in the embedding space



eDifFIQA: Efficient Quality Assessment Using Denoising Diffusion Probabilistic Models

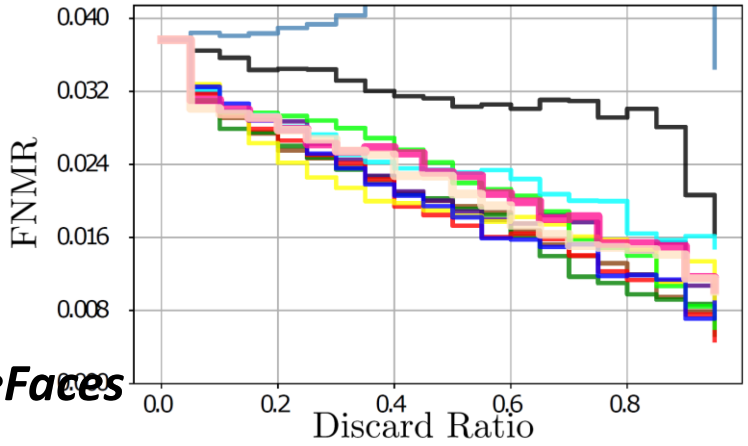
- eDifFIQA



$$q_x^o = \left(q_x - \epsilon \cdot \frac{\hat{e}_x^T \cdot \overline{e}_{C_i^{(p)}}}{\|\hat{e}_x\| \cdot \|\overline{e}_{C_i^{(p)}}\|}\right)$$

$$\mathcal{L}_q = \|\hat{q}_x - q_x^o\|$$

Experimental Setup



- **Standard evaluation approach:**
  - **EDC+pAUC**
  - **Benchmarks:**
    - *Adience, CALFW, CFP-FP, CPLFW, IJB-C, LFW, XQLFW,* ***YouTubeFaces***
  - **FR models:**
    - *AdaFace, ArcFace, CurricularFace, CosFace,* ***SwinFace***, ***TransFace***
  - **FIQA methods:**
    - *FaceQnet, SDD-FIQA, PFE, PCNet, MagFace, LightQNet, SER-FIQ, FaceQAN, CR-FIQA, FaceQgen, DifFIQA, DifFIQA(R)*
  - ***Backbones**:*
    - *eDifFIQA(L) – ResNet100, eDifFIQA (M) – ResNet50, eDifFIQA(S) – ResNet18*

Results (pAUC)

- *Results using discard ratio of (0.3).*

ArcFace - pAUC@FMR=10^{-3}(↓)

FIQA Model	Adience	CALFW	CFP-FP	CPLFW	IJB-C	LFW	XQLFW	$\overline{pAUC}$
FaceQnet [10]	0.943	0.955	0.702	0.878	1.224	0.884	0.899	0.926
SDD-FIQA [13]	0.783	0.901	0.497	0.734	0.720	0.808	0.774	0.745
PFE [18]	0.774	0.932	0.522	0.738	0.783	0.779	0.641	0.738
PCNet [11]	1.022	1.006	0.863	0.783	0.706	**0.623**	1.004	0.858
MagFace [19]	0.812	0.902	0.500	0.717	0.824	0.635	0.943	0.762
LightQNet [12]	0.789	0.913	0.582	0.752	0.721	0.745	**0.621**	0.732
SER-FIQ [14]	**0.767**	0.903	0.446	0.656	0.671	0.935	0.676†	0.722
FaceQAN [15]	0.824	0.941	**0.347**	0.677	0.673	0.624	0.667	0.679
CR-FIQA [21]	0.808	0.891	0.369	0.689	0.664	0.675	0.680	0.682
FaceQgen [32]	0.817	0.985	0.724	0.701	0.785	0.802	0.653	0.781
DifFIQA [5]	0.805	0.918	0.402	0.674	0.675	0.714	0.652	0.691
DifFIQA(R) [5]	0.801	0.898	0.408	**0.646***	**0.655***	0.708	0.653	0.681
eDifFIQA(S)	0.822	0.916	0.402	0.679	0.679	0.786	0.634*	0.703
eDifFIQA(M)	0.825	**0.846***	0.386	0.669	0.663	0.712	0.671	0.682
eDifFIQA(L)	0.793*	0.872	0.376*	0.667	0.663	0.705*	0.672	0.678

TransFace - pAUC@FMR=10^{-3}(↓)

FIQA Model	Adience	CALFW	CFP-FP	CPLFW	IJB-C	LFW	XQLFW	$\overline{pAUC}$
FaceQnet [10]	0.938	0.964	0.753	0.889	1.291	0.898	1.007	0.963
SDD-FIQA [13]	0.837	0.905	0.545	0.694	0.809	0.840	0.819	0.778
PFE [18]	0.834	0.930	0.546	0.711	0.879	0.810	0.798	0.787
PCNet [11]	0.998	1.003	0.837	1.010	0.804	**0.661**	0.654	0.853
MagFace [19]	0.867	0.898	0.512	0.682	0.930	0.673	0.935	0.786
LightQNet [12]	0.847	0.916	0.665	0.697	0.817	0.784	0.634	0.766
SER-FIQ [14]	**0.784**	0.901	0.430	0.647	0.773	0.935	0.585†	0.722
FaceQAN [15]	0.890	0.942	0.377	0.599	0.774	0.663	0.558	0.686
CR-FIQA [21]	0.832	0.887	0.379	0.614	0.765	0.714	0.580	0.681
FaceQgen [32]	0.858	0.985	0.761	0.699	0.877	0.834	0.739	0.822
DifFIQA [5]	0.863	0.917	0.405	0.597	0.782	0.736	0.529	0.690
DifFIQA(R) [5]	0.856	0.901	0.394	0.579	**0.744***	0.722*	**0.523***	0.674
eDifFIQA(S)	0.867	0.917	0.446	0.591	0.770	0.825	0.558	0.711
eDifFIQA(M)	0.861	**0.849***	0.384	0.577	0.759	0.750	0.553	0.676
eDifFIQA(L)	0.836*	0.872	**0.374***	**0.575***	0.755	0.735	0.528	0.668

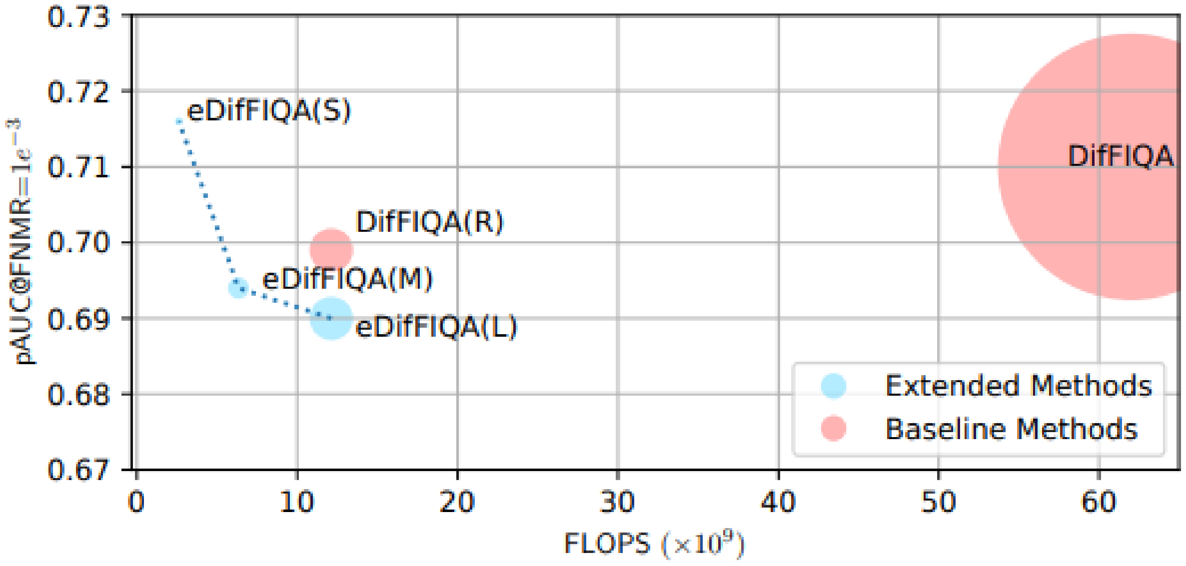
DifFIQA [5]	0.071	0.561	0.657	0.726	0.775	0.815	0.847	0.867	0.884	0.936
DifFIQA(R) [5]	0.007	0.448	0.452	0.659	0.684	0.760	0.837	0.845	0.845	0.942
eDifFIQA(S)	0.032	0.095	0.068	0.365	0.413	0.453	0.666	0.728	0.828	0.897
eDifFIQA(M)	0.000	0.176	0.056	0.469	0.539	0.575	0.734	0.775	0.826	0.870
eDifFIQA(L)	0.036	0.264	0.236	0.440	0.500	0.585	0.768	0.790	0.824	0.913

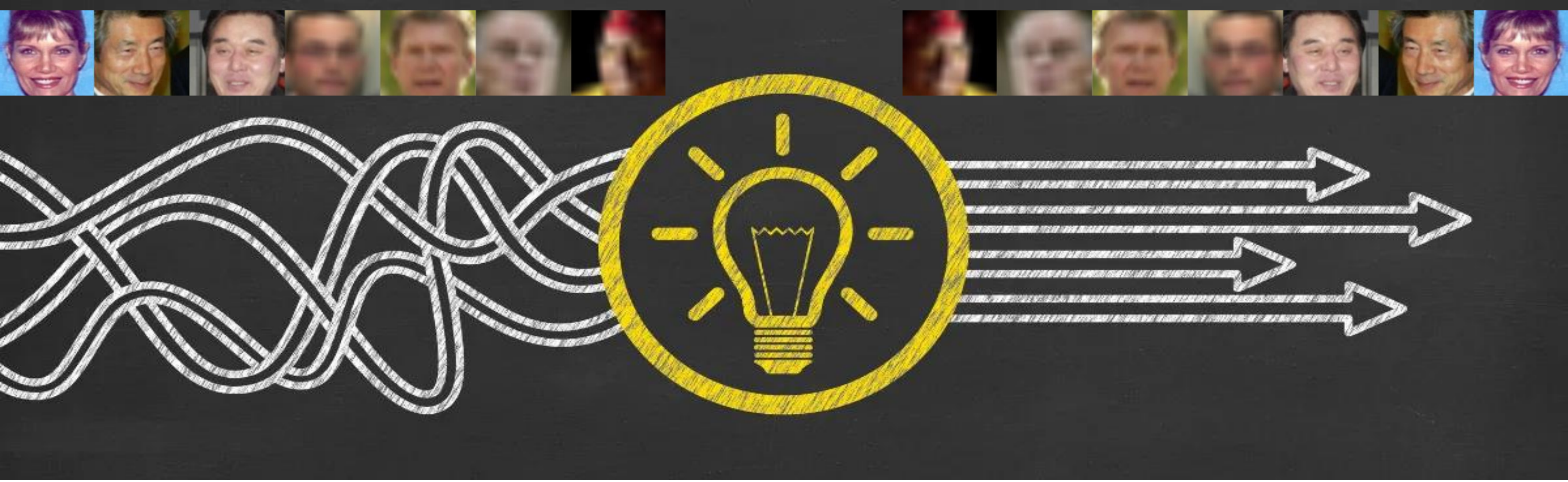
Run-time

- **Hardware: Intel i9-10900KF CPU, 64 GB RAM in Nvidia 3090 GPU**

FIQA model	Baseline methods [5]		eDifFIQA variants		
	DifFIQA	DifFIQA(R)	eDifFIQA(S)	eDifFIQA(M)	eDifFIQA(L)
Runtime ($\mu \pm \sigma$) [in ms]	1074.62 ± 11.45	1.24 ± 0.36	0.36 ± 0.77	0.63 ± 0.75	1.09 ± 0.77
# Parameters (10^6)	$142.87 + 65.15^{\dagger}$	65.15	24.55	44.11	65.68

† The sum refers to the number of parameters of the generative UNet model and the recognition model.





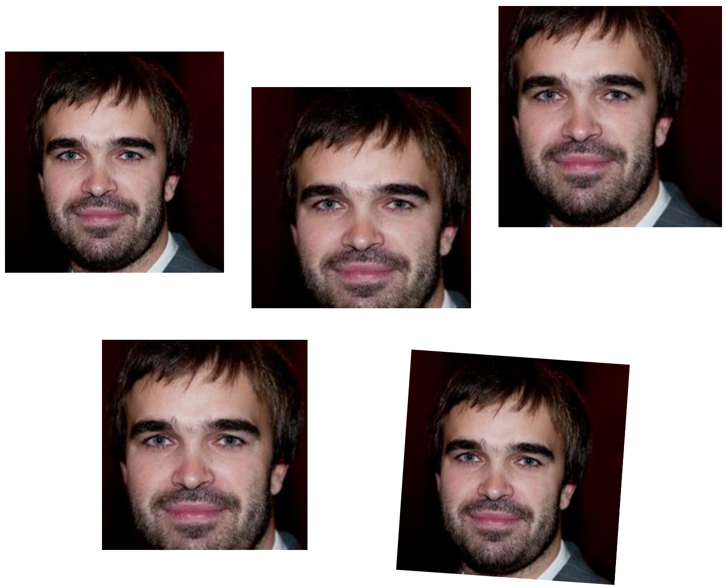
Part III: Open Challenges

Publications available:
https://lmi.fe.uni-lj.si/en/vitomir-struc/publications/

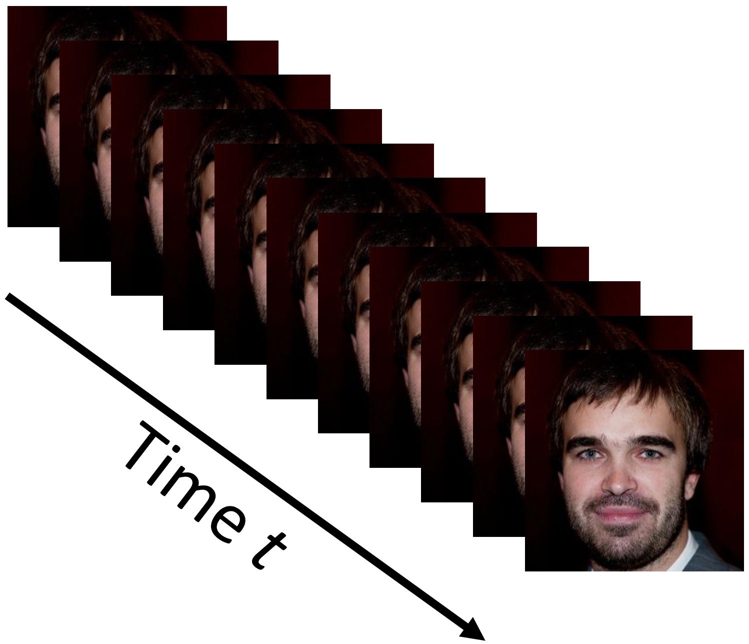
GitHub: https://github.com/LSIbabnikz/

FIQA – Open Challenges

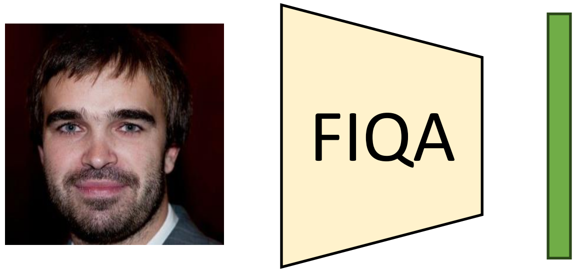
Detector variability:
Quality scores vary with resepect to detector variability



Video FIQA:
Large number of frames with similar quality



Beyond scalar FIQA:
Vector quality descriptions



Specialized FIQA:
For specific applications

Javna agencija za znanstvenoraziskovalno in inovacijsko dejavnost Republike Slovenije

Thank You!

ziga.babnik@fe.uni-lj.si
peter.peer@fri.uni-lj.si
vitomir.struc@fe.uni-lj.si

Publications

FIQA Code