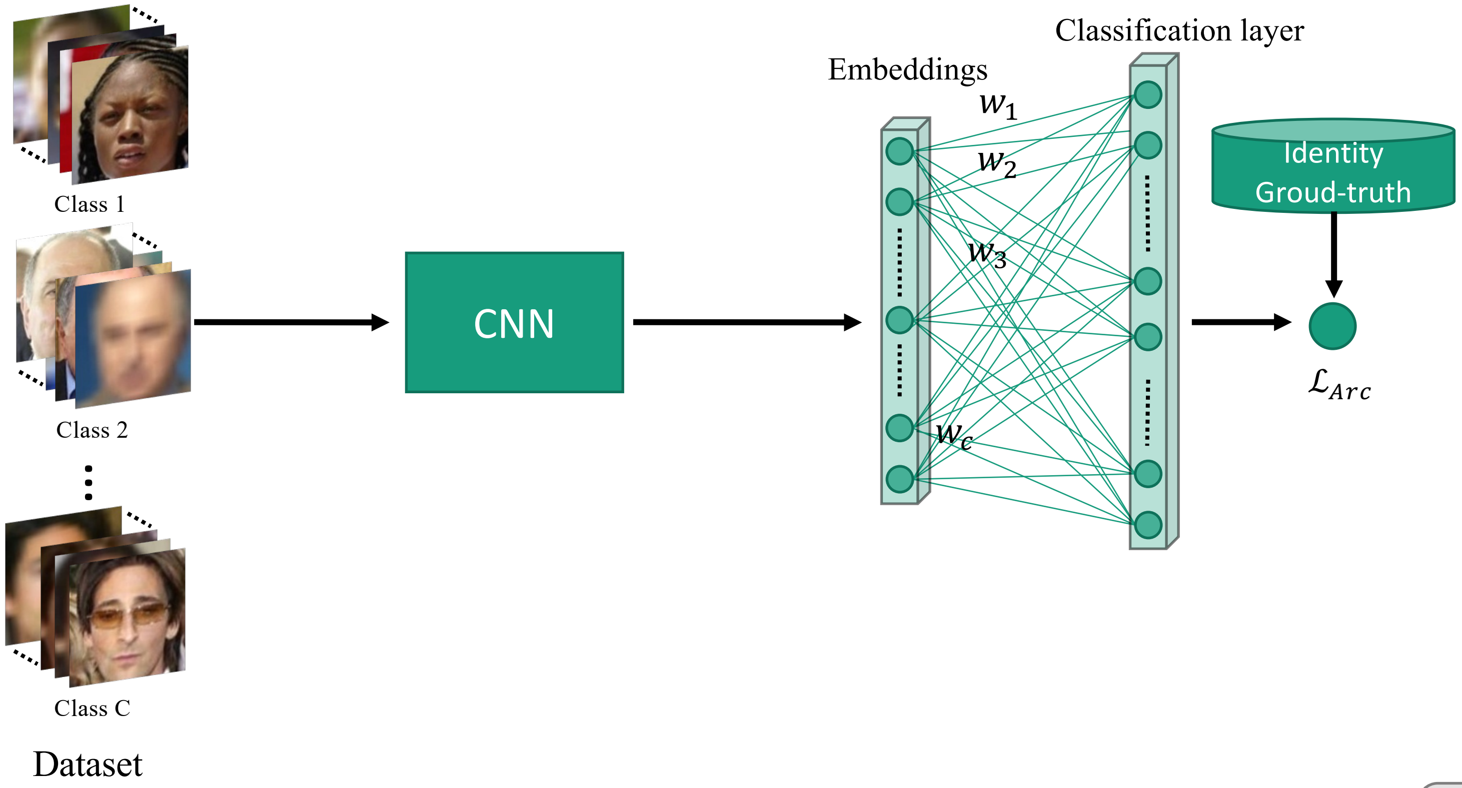
CR-FIQA: Face Image Quality Assessment by Learning Sample Relative Classifiability

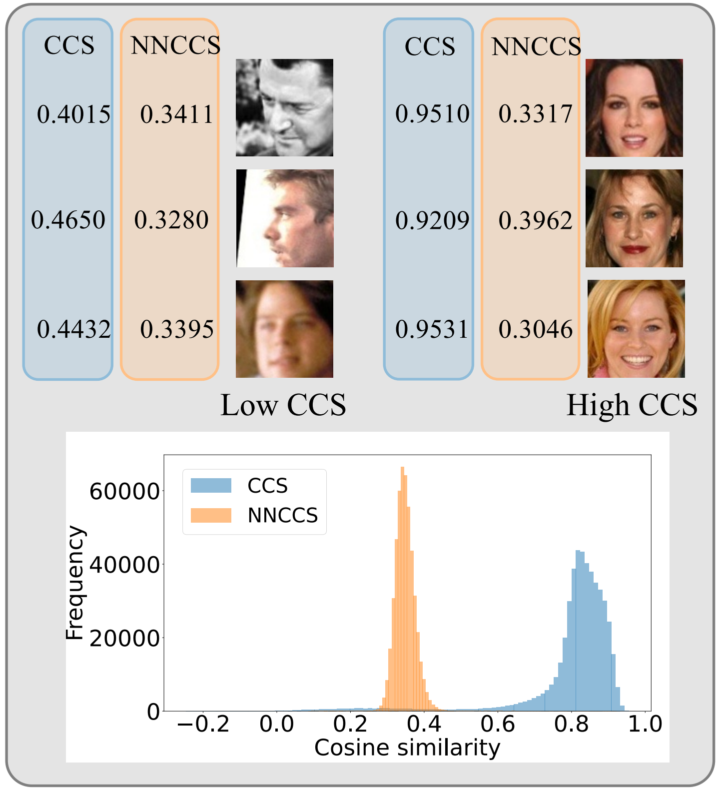
Fadi Boutros [1], Meiling Fang[1,2], Marcel Klemt[1], Biying Fu[1], Naser Damer[1,2]
[1]Fraunhofer Institute for Computer Graphics Research IGD, Darmstadt, Germany
[2]Department of Computer Science, TU Darmstadt, Darmstadt, Germany

Paper tag:TUE-PM-163 – Paper id: 7587

CR-FIQA

Face Image Quality Assessment by Learning Sample Relative Classifiability

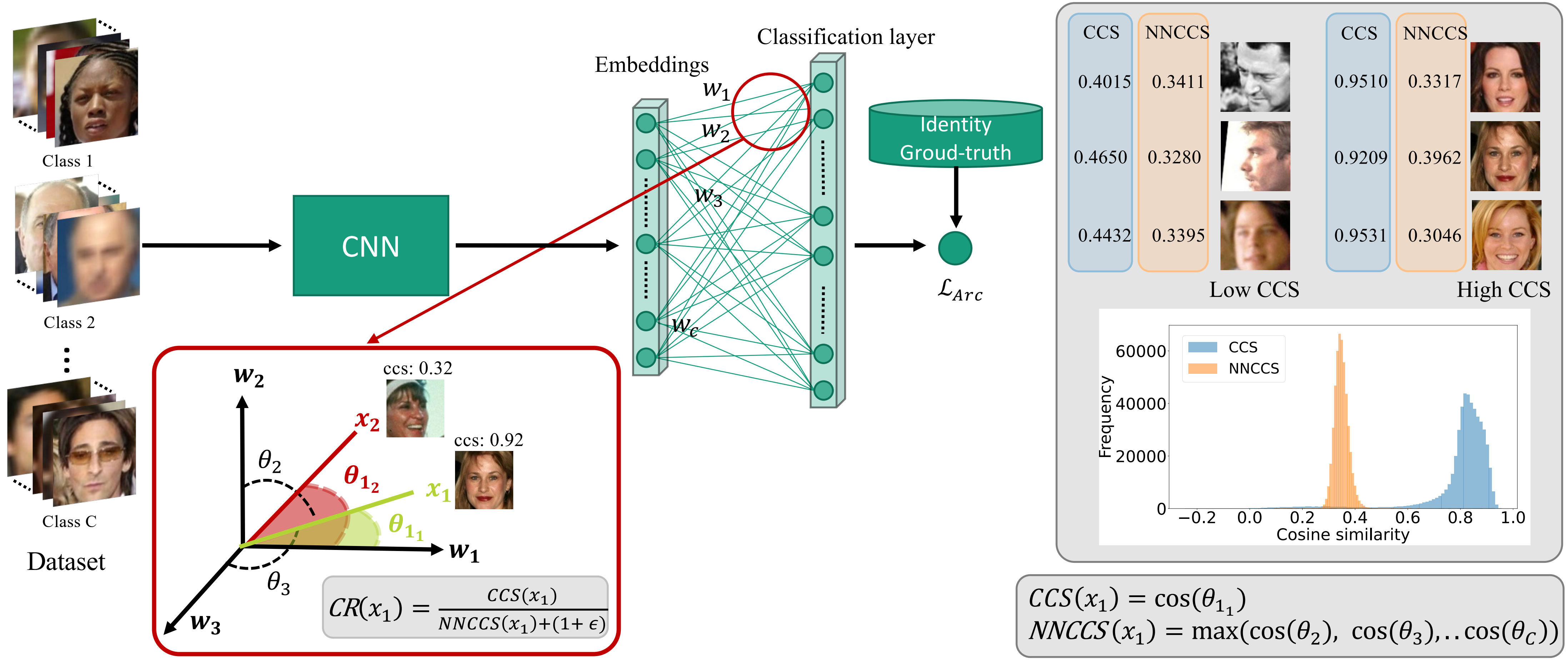




$$CCS(x_1) = \cos(\theta_{1_1})$$
$$NNCCS(x_1) = \max(\cos(\theta_2),\ \cos(\theta_3), .. \cos(\theta_C))$$

CR-FIQA

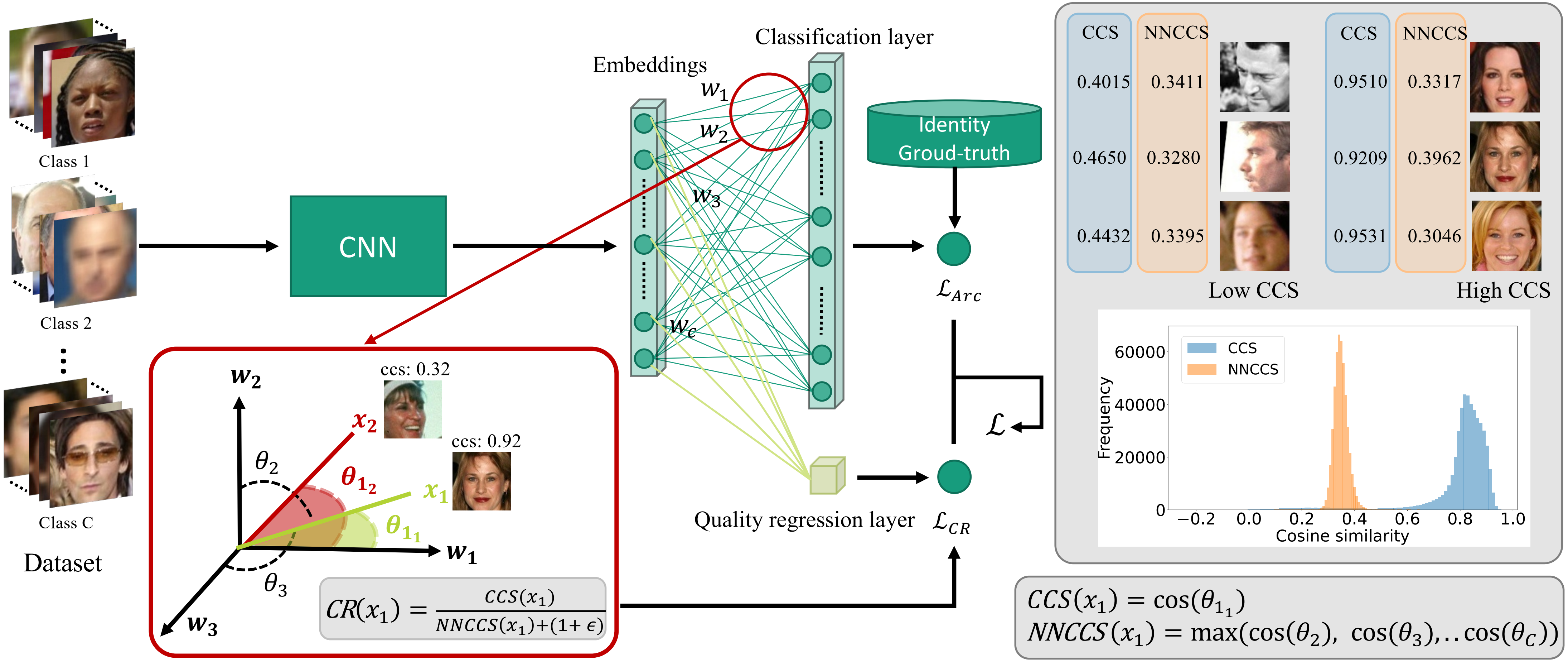
Face Image Quality Assessment by Learning Sample Relative Classifiability



$$CCS(x_1) = \cos(\theta_{1_1})$$
$$NNCCS(x_1) = \max(\cos(\theta_2),\ \cos(\theta_3), .. \cos(\theta_C))$$

CR-FIQA

Face Image Quality Assessment by Learning Sample Relative Classifiability



What is Face image quality (FIQ)?

- Face image utility indicates the utility (value) of an image to face recognition (FR) algorithms
- What face quality estimator should answer?

How usable (utility*) is the sample for automatic face recognition systems?



*Utility: the observed or predicted positive or negative contribution of an individual sample to the overall performance of a biometric system.

Why do we need FIQ?

- The performance of biometric recognition is driven by the quality of its samples

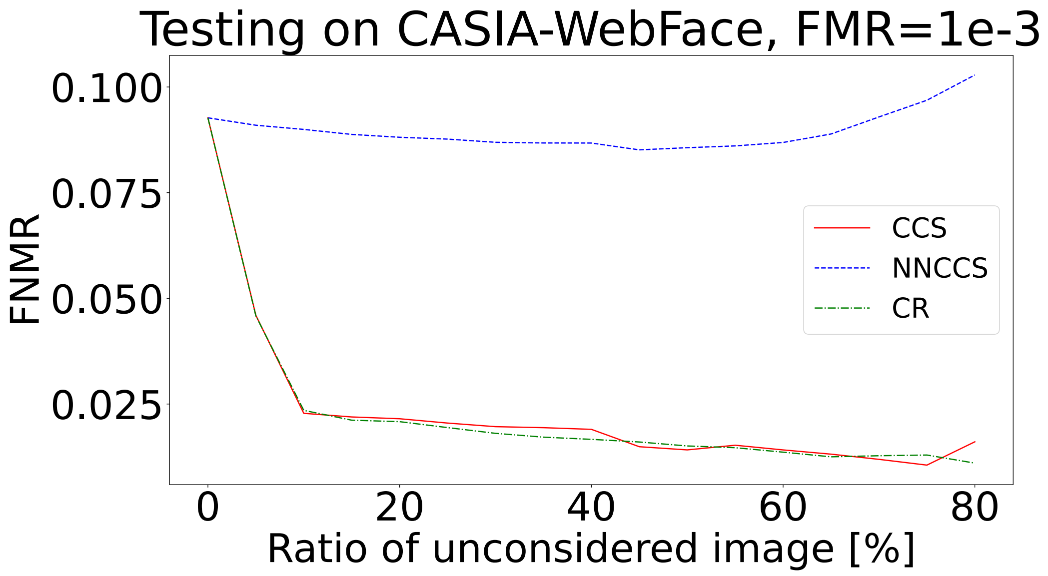
Biometric sample quality: estimate the suitability of a sample for recognition

- Some advantages of sample quality

  ✓ More robust enrolment
  ✓ More secure negative identification systems
  ✓ Enables quality-based fusion approaches

How can we evaluate face image quality estimator?

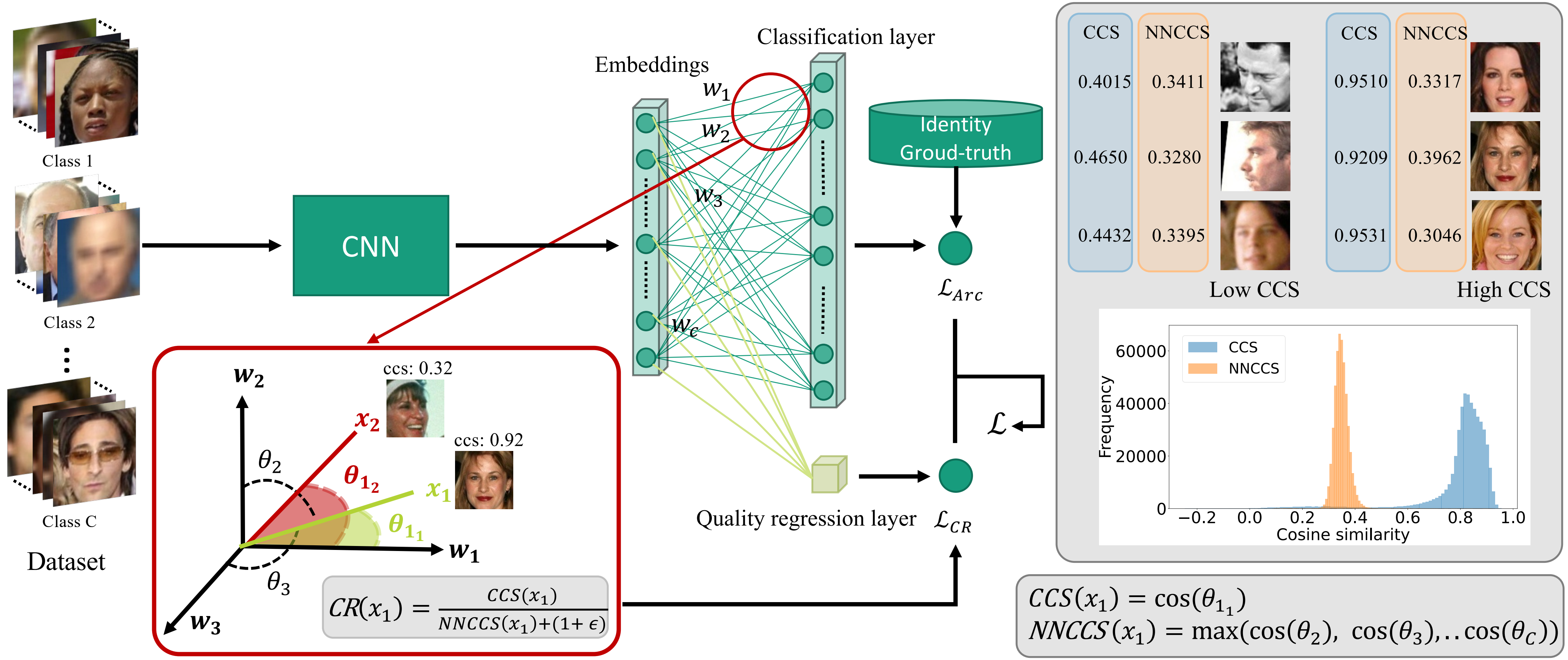
- If we removed the worse data (according to our quality estimation), do the face verification/identification performance on the rest of the data (higher quality) enhance?
- Error vs. Reject Curves (ERC): ERC plots the percentage of the neglected data (of the worst estimated quality) vs. a biometric verification/identification error rate.
- Area under the Curve (AUC) of the ERC: provide a quantitative aggregate measure of verification performance across all rejection ratios.



Paper tag:TUE-PM-163

CR-FIQA

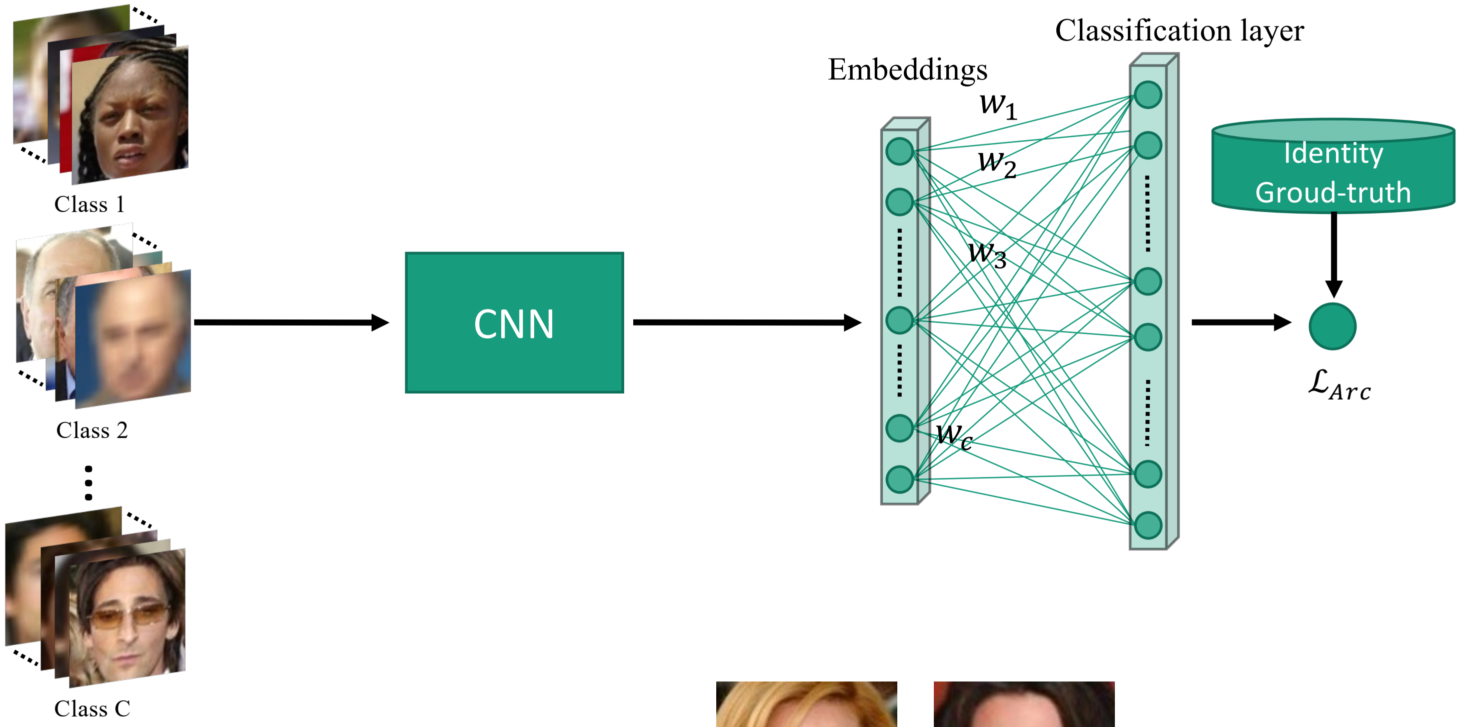
Face Image Quality Assessment by Learning Sample Relative Classifiability



$$CCS(x_1) = \cos(\theta_{1_1})$$
$$NNCCS(x_1) = \max(\cos(\theta_2),\ \cos(\theta_3), \ldots \cos(\theta_C))$$

CR-FIQA

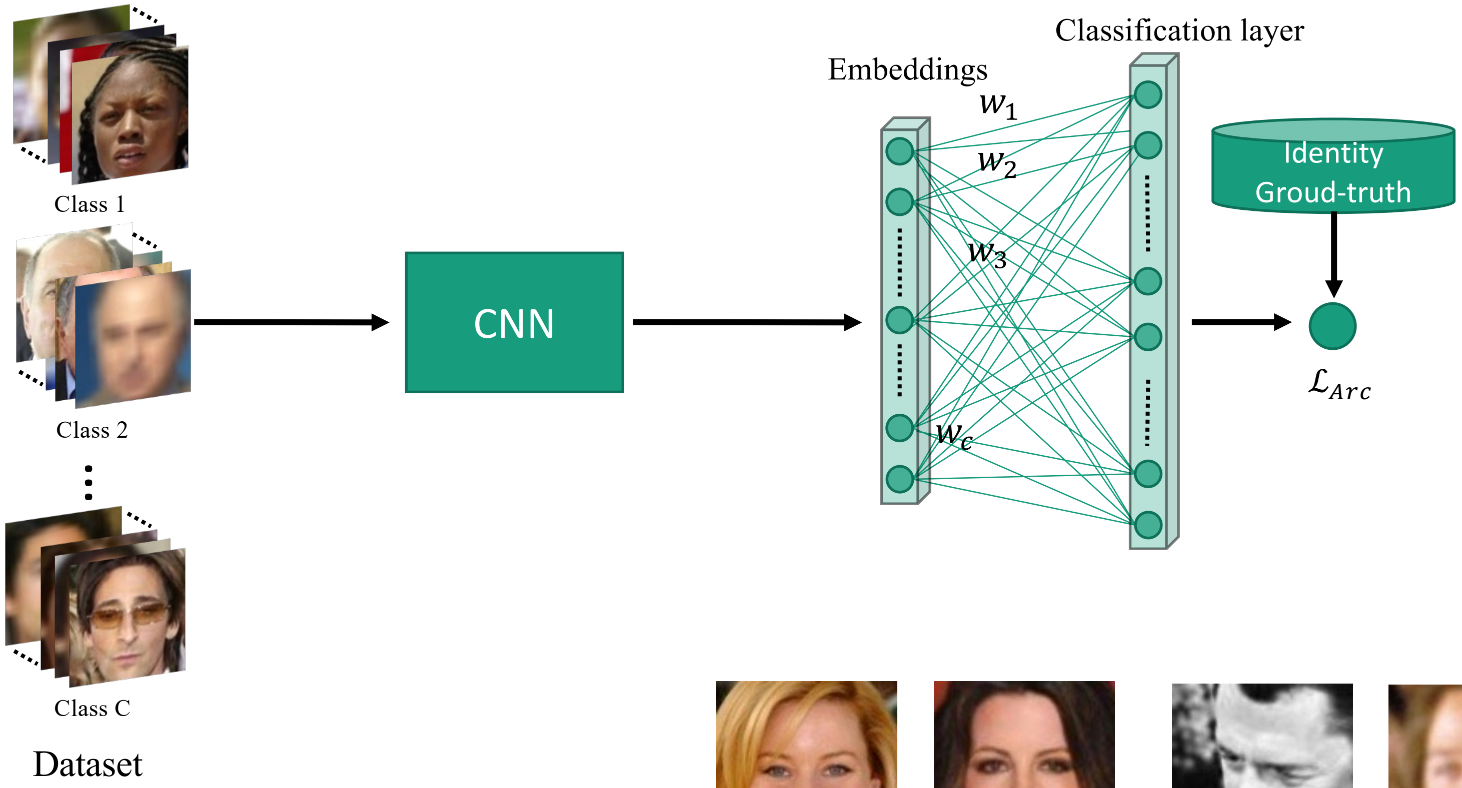
Face Image Quality Assessment by Learning Sample Relative Classifiability

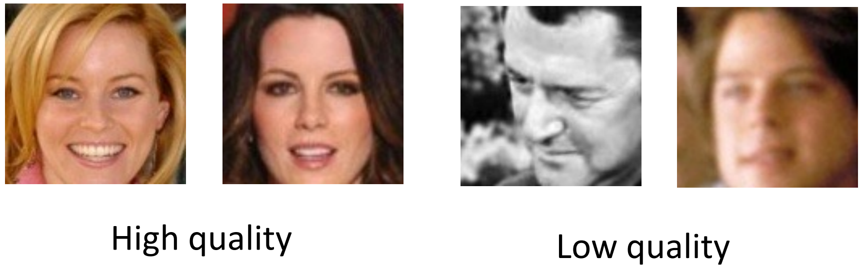


High quality

CR-FIQA

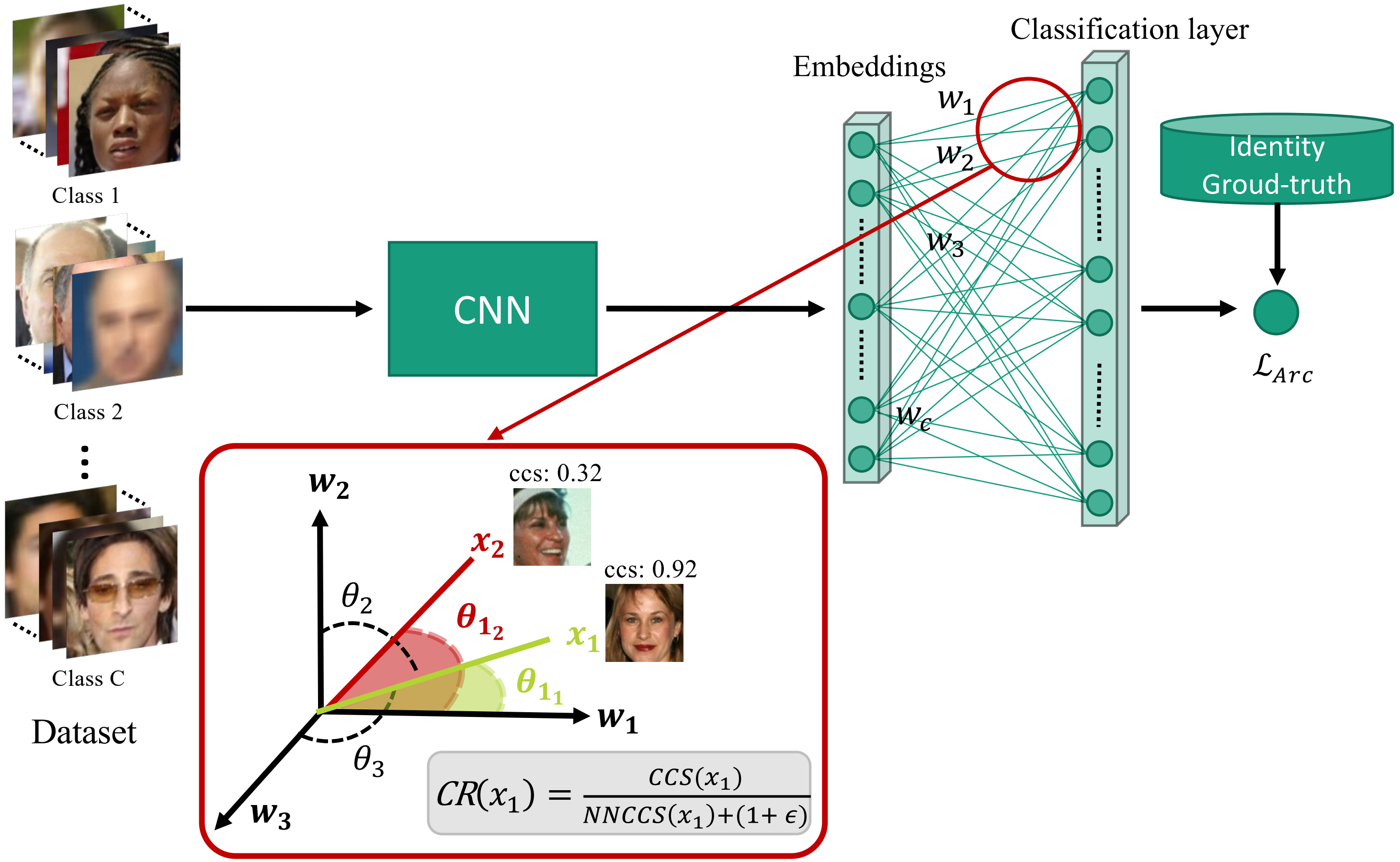
Face Image Quality Assessment by Learning Sample Relative Classifiability





CR-FIQA

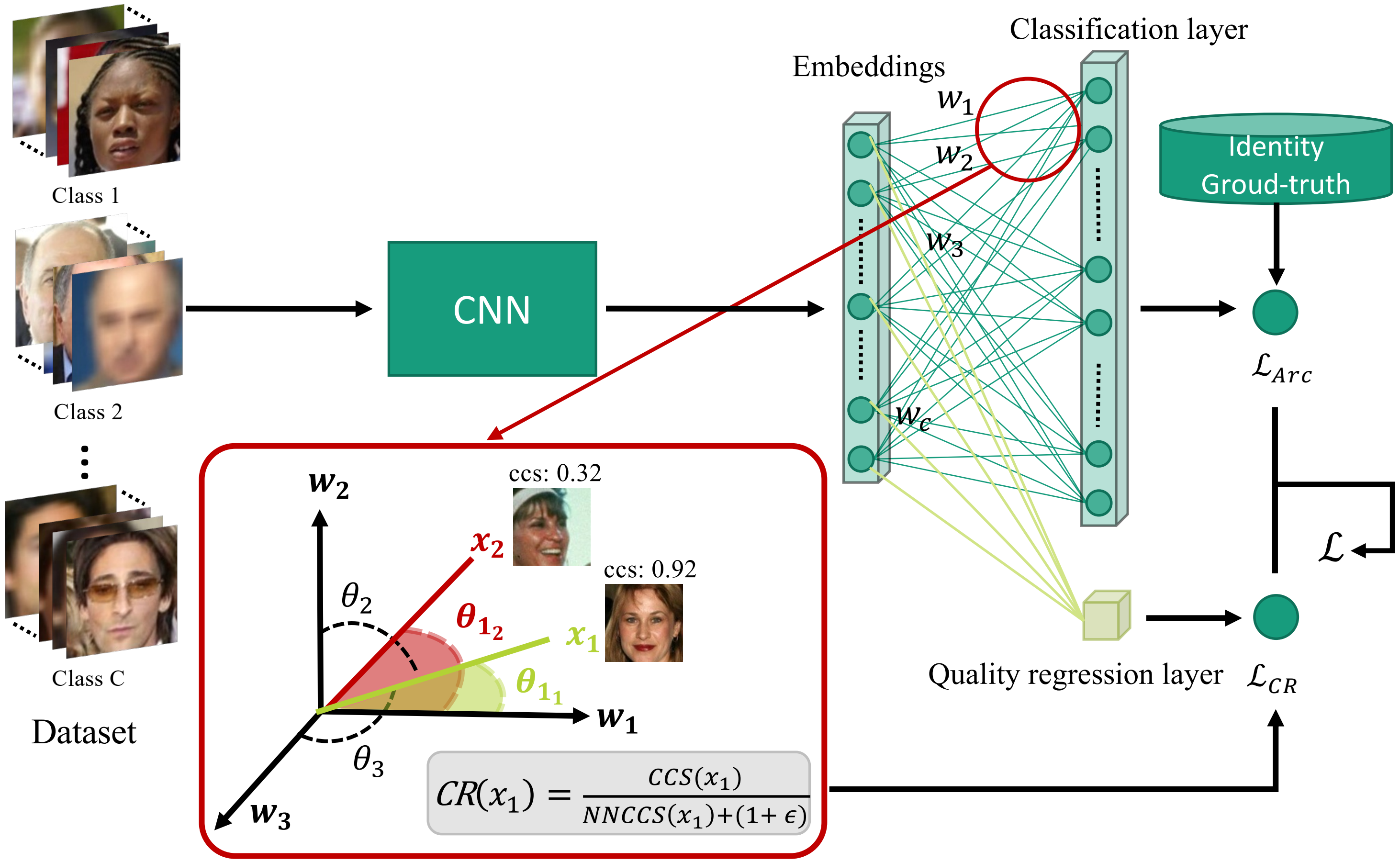
Face Image Quality Assessment by Learning Sample Relative Classifiability



if a given sample was hypothetically part of the FR model training (which it is not), how relatively close would it be to its class center?

CR-FIQA

Face Image Quality Assessment by Learning Sample Relative Classifiability



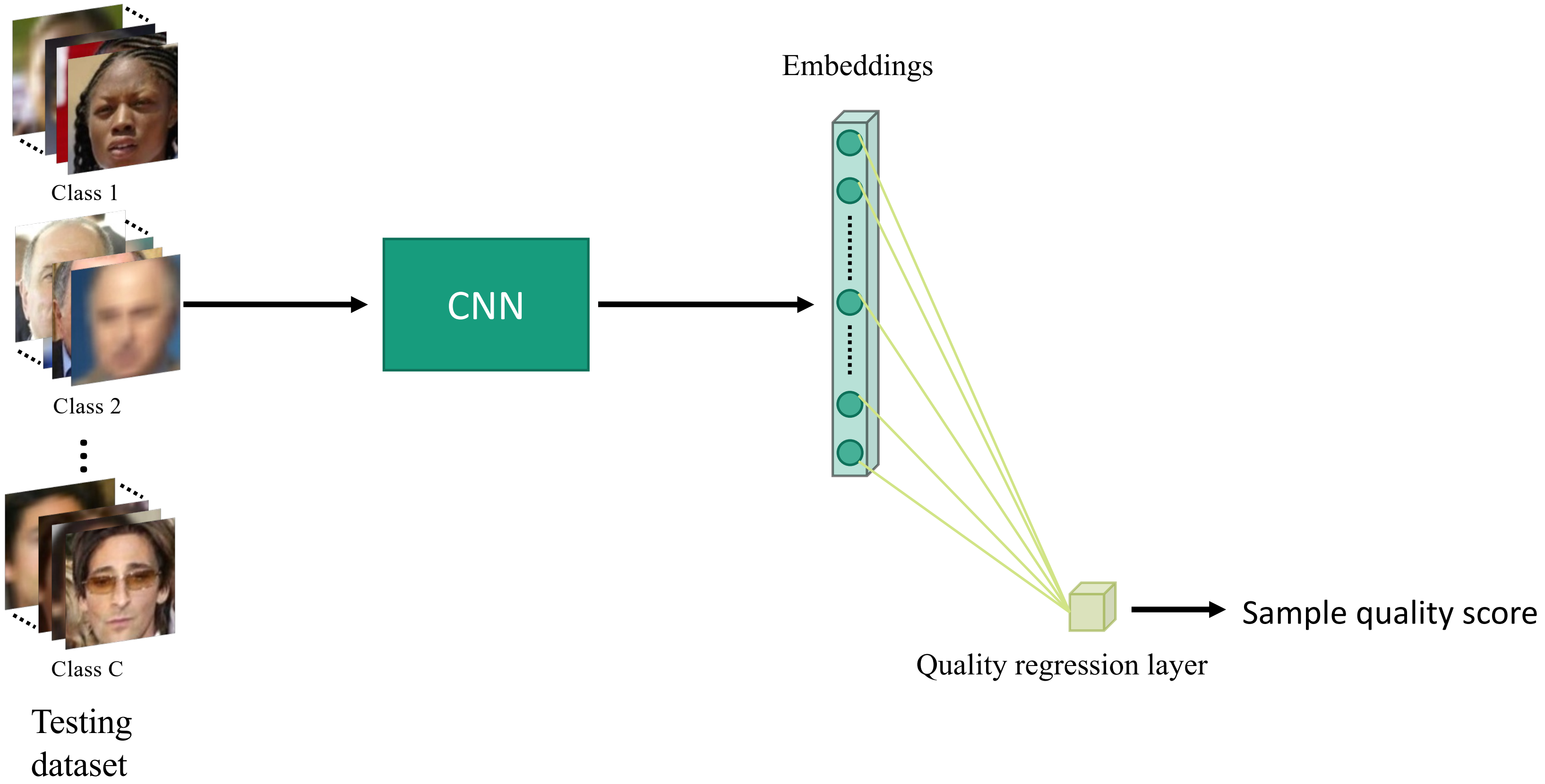
- **CCS** (Class Center Angular Similarity): a proximity between sample x and its class center
- **NNCCD** (Closest Nearest Negative Class Center Angular Similarity): a proximity between sample x and the nearest negative class center
- **Certainty Ratio (CR)**: a relative proximity of sample x to its class center and negative class center

$$CCS(x_1) = \cos(\theta_{1_1})$$
$$NNCCS(x_1) = \max(\cos(\theta_2),\ \cos(\theta_3), ..\cos(\theta_C))$$

Paper tag:TUE-PM-163

CR-FIQA

Face Image Quality Assessment by Learning Sample Relative Classifiability



CR-FIQA

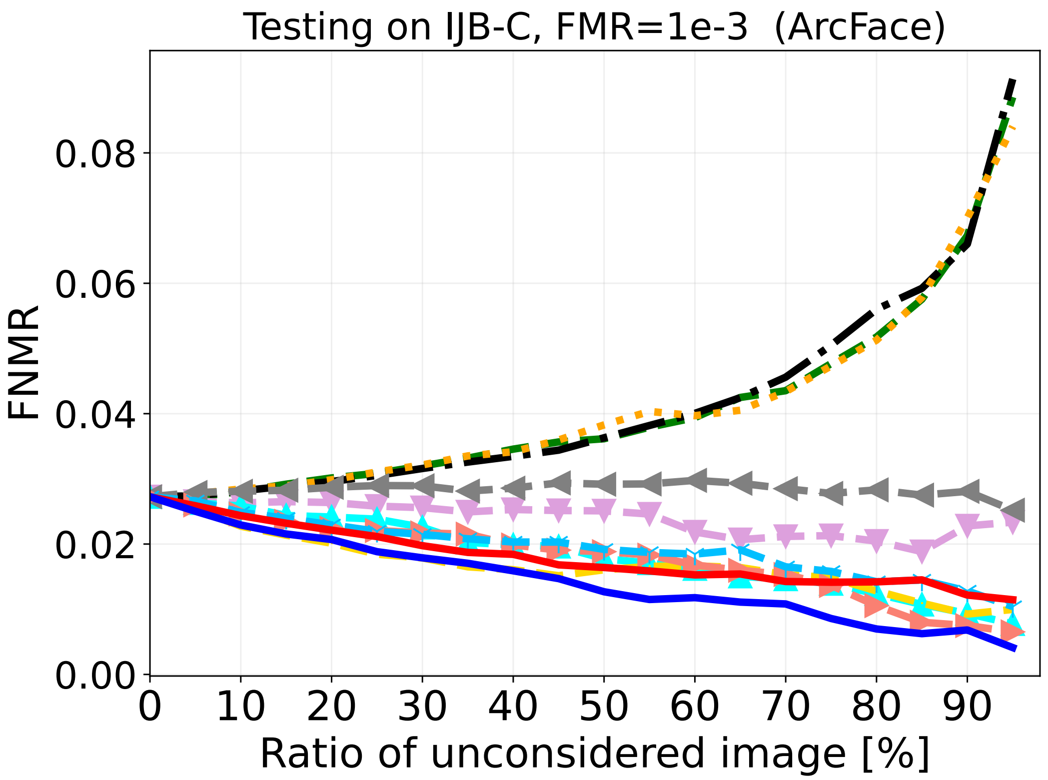
Evaluation

- CR-FIQA (S)
  - Training dataset: CASIA-WebFace
  - Base network architecture: ResNet-50
- CR-FIQA (L)
  - Training dataset: MS1MV2
  - Base network architecture: ResNet-100

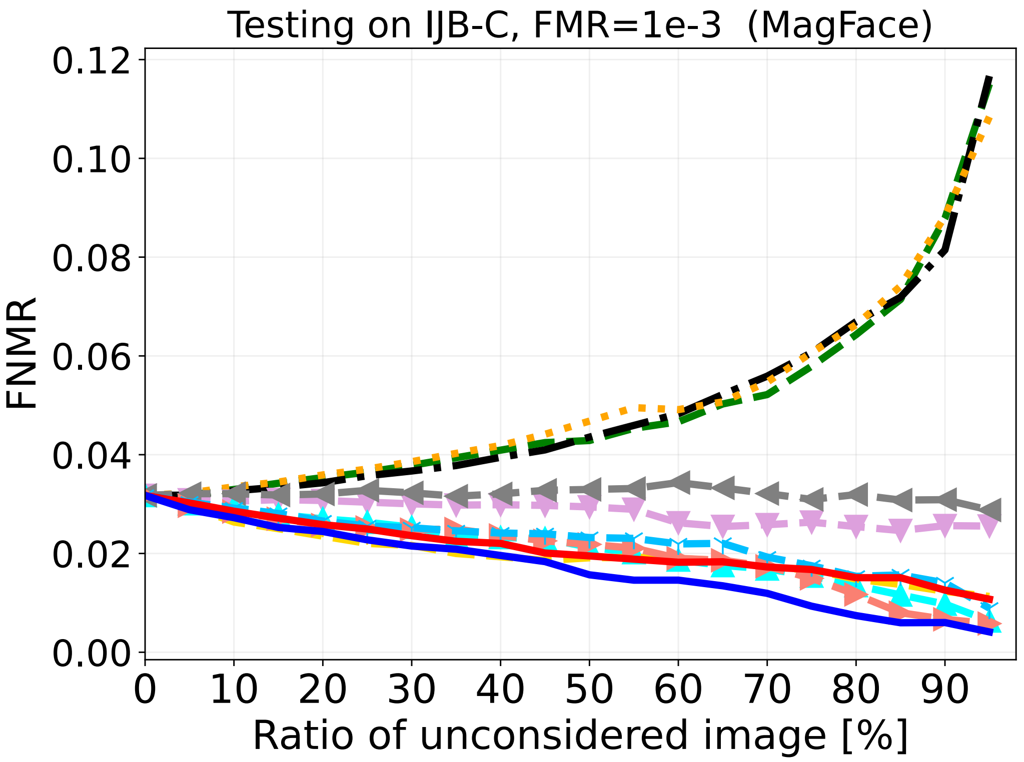
- Face recognition models
  - ElasticFace
  - ArcFace
  - MagFace
  - CurricularFace

- Evaluation benchmarks:
  - LFW
  - AgeDB-30
  - CFP-FP
  - CALFW
  - Adience
  - CPLFW
  - XQLFW
  - IJB-C

CR-FIQA: What did we achieve?



ArcFace



MagFace

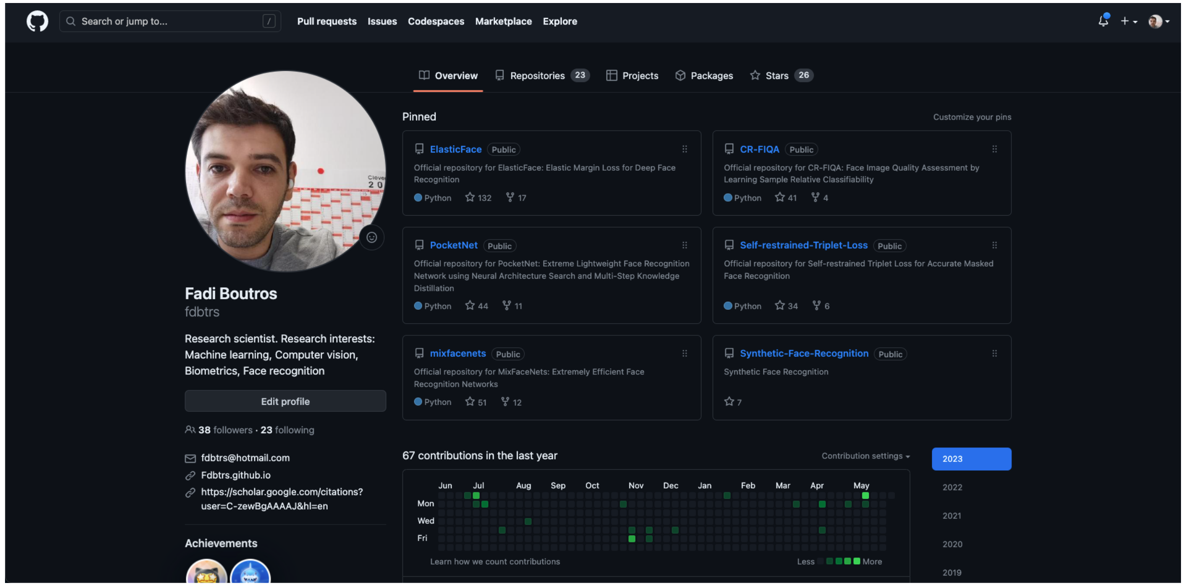
BRISQUE | RankIQA | DeepIQA | RankIQ | PFE | SER-FIQ | FaceQnet | MagFace | SDD-FIQA | CR-FIQA(S) (Our) | CR-FIQA(L) (Our)

CR-FIQA: What did we achieve?

Evaluation: quality scores as an embedding weighting term

	Quality Estimation	1:1 mixed Verification: TAR (%) at											
		ArcFace [5]			ElasticFace [3]			MagFace [28]			CurricularFace [17]		
		FAR=1e-6	FAR=1e-5	FAR=1e-4	FAR=1e-6	FAR=1e-5	FAR=1e-4	FAR=1e-6	FAR=1e-5	FAR=1e-4	FAR=1e-6	FAR=1e-5	FAR=1e-4
	-	*89.85*	94.47	96.28	89.15	94.54	96.49	85.67	93.08	96.65	90.46	94.89	96.58
IQA	BRISQUE [29]	86.65	93.62	95.98	85.68	93.51	95.65	81.11	90.64	94.82	88.16	93.98	96.29
	RankIQA [26]	86.37	93.61	95.83	86.71	93.46	96.00	80.78	90.75	94.86	88.16	94.11	96.22
	DeepIQA [2]	81.97	91.64	94.67	78.93	91.59	94.81	73.53	86.34	92.90	82.65	92.04	95.00
FIQA	RankIQ [4]	88.78	94.42	96.20	88.88	94.64	96.45	85.63	92.66	95.70	90.00	94.93	96.53
	PFE [35]	89.50	94.51	96.31	89.10	94.67	96.51	84.93	92.44	95.60	90.36	95.04	96.54
	SER-FIQ [36]	89.74	94.65	96.32	*90.05*	94.79	96.57	86.02	93.35	95.80	90.66	95.11	96.58
	FaceQNet [14, 15]	87.87	94.04	96.12	86.26	94.09	96.25	82.91	90.56	95.03	89.61	94.65	96.36
	MagFace [28]	89.49	94.41	96.22	89.37	94.69	96.46	85.75	92.71	95.54	90.34	95.02	96.50
	SDD-FIQA [32]	89.39	94.61	96.34	88.07	94.82	96.49	84.69	92.83	95.73	89.91	*95.12*	*96.63*
	CR-FIQA(S)(Our)	89.59	**94.78**	*96.35*	**90.30**	**94.97**	**96.63**	*86.45*	*93.48*	**95.95**	**90.82**	**95.13**	**96.64**
	CR-FIQA(L)(Our)	**90.16**	*94.75*	**96.36**	90.00	*94.92*	*96.58*	**87.12**	**93.67**	*95.90*	*90.79*	*95.12*	96.58

Code and pretrained models



https://github.com/fdbtrs/CR-FIQA

CR-FIQA: Face Image Quality Assessment by Learning Sample Relative Classifiability

Fadi Boutros [1], Meiling Fang[1,2], Marcel Klemt[1], Biying Fu[1], Naser Damer[1,2]
[1]Fraunhofer Institute for Computer Graphics Research IGD, Darmstadt, Germany
[2]Department of Computer Science, TU Darmstadt, Darmstadt, Germany

Paper tag:TUE-PM-163 – Paper id: 7587