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

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
Email: fadi.boutros@igd.fraunhofer.de

1. Supplementary Material

This supplementary material complements the main submission by providing:

1. Complementary ERC curves with AUC values for all the FR models and benchmarks to complement and support the ablation study section (Section5) of the main manuscript.
2. Samples images from the 8 benchmarks with quality scores achieved by our CR-FIQA and SOTA methods.
3. Quality score distribution of the evaluation benchmarks achieved by our CR-FIQA and SOTA methods.
4. ERC (FNMR at FMR1e-4 vs reject) curves that provide a complement to the AUC reported in Table 1 of the main manuscript.
5. ERC (FNMR at FMR1e-3 vs reject) curves using MagFace and Curricular FR models that provide a complement to the AUC reported in Table 1 and Figure 4 of the main manuscript.
6. More details on the databases and benchmarks.
7. Discussion of the potential social impacts.
8. Details on further existing assets used in the work.
9. A discussion on the technical limitations of the presented work.

1.1. Complementary Result for Ablation Study

Figures 1, 2, 3 and 4 present a comparison between ERCs (FNMR at FMR1e-3) of CR-FIQA(S), CCS-FIQA(S), CR-FIQA(S) (On top) and CCS-FIQA(S) (On top) on the evaluation benchmarks. Figures 5, 6, 7 and 8 present a comparison between ERCs (FNMR at FMR1e-4) of CR-FIQA(S), CCS-FIQA(S), CR-FIQA(S) (On top), and CCS-FIQA(S) (On top) on the evaluation benchmarks. These ERC curves are complementary to the ablation study presented in main manuscript (Section 5). In the main submission, these ERC curves are presented for ArcFace [5] FR model on Adience [7], AgeDb-30 [18], CALFW [27] and CFP-FP [20] in Figure 3 (main submission) and discussed on ablation study section (Section 5 of main submission). However, in this supplementary material we opt to provide the evaluation mentioned in Lines 583-597 on all considered FR models and evaluation benchmarks to stress the conclusion of our ablation study (Section 5 of main submission). This again points out the benefits of CR of CCS (thus the NNCCS term in equation 4 of the main submission), as well as the simultaneously training rather than on the top learning.

1.2. Histogram of CCS and NNCCS

Figure 9 shows an insight into the CCS and NNCCS values distribution of the training datasets (CASIA-WebFace [25] and MS1MV2 [5]). Figure 9a shows an enhanced visualisation of the same plot shown in Figure 1 (main submission) based on the R50(CASIA) model and discussed in lines 338-342 of the main submission. Figure 9b shows CCS and NNCCS values distribution of MS1MV2 dataset obtained from ResNet-100 (R100(MS1M-V2)) model to provide an additional illustration of the CCS and NNCCS value distribution on another training setup (model and dataset). On both models one can notice that the CCS and NNCCS values vary between samples.

1.3. Quality score distribution

Figure 11 presents the quality score distribution of the evaluation benchmarks achieved by our CR-FIQA and the SOTA methods, all normalized to have a range between 0 and 1. One can notice in the distributions, that for the XQLFW dataset where the data contains extreme low and extreme low quality samples by design, this two groups of quality is only visible in our CR-FIQA, PFE, MagFace, SDD-FIQA, as well as the methods that were used to label the qualities when constructing the XQLFW, i.e. SER-FIQA and BRISQUE.

1.4. Sample images with quality scores

Figure 10 shows sample images of the evaluation benchmarks with quality score values obtained from our CR-FIQA the SOTA methods. These images in Figure 10 illustrate samples of different benchmarks with quality score values. It is important to mention that, although the quality scores are normalized between 0 and 1, the higher quality score values across FIQA methods do not mean that the method points out a relative higher quality estimation than the other methods. For example, SER-FIQ method resulted always in relatively high quality score value when it is compared to other SOTA methods. However, as show in Figure 11, the quality score value range of SER-FIQ is higher when compared to other SOTA methods.

1.5. FIQA performance as ERC (FNMR at FMR1e-4 vs reject) curves

Figures 12, 13, 14 and 15 present ERC (FNMR at FMR1e-4 vs reject) curves for all the evaluation settings. These ERC curves illustrates the curves producing the AUC (FNMR at FMR1e-4) presented in Table 1 of the main submission. Such ERC curves are shown in Figure 4 in the main submission and Figure 18 in supplementary material on FNMR at FMR1e-3 and discussed in details in Section 6. However, we present also in this supplementary material the ERC curves on another FNMR, FNMR at FMR1e-4. These ERC curves also correspond to the AUC values presented in Table 1 of the main submission.

1.6. FIQA performance as ERC (FNMR at FMR1e-3 vs reject) curves using MagFace and CurricularFace FR models

Figure 18 presents ERC (FNMR at FMR1e-3 vs reject) curves for all the evaluation benchmarks using MagFace and CurricularFace FR models. These ERC curves also correspond to the AUC values presented in Table 1 of the main submission and discussed in details in Section 6.

1.7. CR-FIQA as feature extraction

The evaluation of CR-FIQA(L) backbone as feature extraction, which is not the goal of this work, on mainstream FR benchmarks is presented in Table 1. The considered benchmarks are LFW [10], AgeDB-30 [18], CFP-FP [20], CALFW [27], CPLFW [26] and IJB-C [14]. We followed the evaluation metrics defined in the utilized benchmarks as follows: LFW (accuracy), CALFW (accuracy), CPLFW (accuracy), CFP-FP (accuracy), AgeDB-30 (accuracy) and IJB-C (TAR at FAR1e-4). Although, the presented solution in this paper does not aim, and is not presented as, a solution to extract face embeddings, but rather a FIQA solution, the reported evaluation results (Table 1) are very comparable to the recent SOTA models trained under a similar training setting and only using the face recognition loss.

1.8. Datasets

This section presents the description and license information of the used datasets in our work.

Adience [7]: Adience was used to estimate the age and gender from face images acquired in challenging and in the wild conditions. Adience dataset contains 26,580 images across 2,284 identities, where the images were captured as close to the real-world condition as possible, under all variations in appearance, pose, illuminations, and image quality. Adience license is limited to research purposes only. Detailed information on database creation and licensing can be found in [7] and `https://talhassner.github.io/home/projects/Adience/Adience-main.html`.

AgeDB-30 [18]: AgeDB is an in-the-wild dataset for age-invariant face verification evaluation, containing 16,488 images of 568 identities. Every image is annotated with respect to the identity, age, and gender attribute. In our case, we report the performance for AgeDB-30 (30 years age gap) as it is the most reported and challenging subset of AgeDB. More details on the collection process can be found in [18] and the details on the license are presented in `https://ibug.doc.ic.ac.uk/resources/agedb/`.

LFW [10]: Labeled Faces in the Wild (LFW) is an unconstrained face verification dataset. The LFW contains 13,233 images of 5749 identities collected from the web. The LFW is licensed under CC-BY-4.0, and more information on database creation can be found in [10] and `http://vis-www.cs.umass.edu/lfw/`.

CFP-FP [20]: Celebrities in Frontal-Profile in the Wild (CFP-FP) [20] dataset addresses the comparison between frontal and profile faces. CFP-FP dataset contains 7,000 images across 500 identities, where 10 frontal and 4 profile image per identity. More information can be found in [20] and `http://www.cfpw.io/`.

CALFW [27]: The Cross-age LFW (CALFW) dataset [27] is based on LFW with a focus on comparison pairs with the age gap, however not as large as AgeDB-30. Age gap distribution of the CALFW is provided in [27]. It contains 3000 genuine comparisons, and the negative pairs are selected of the same gender and race to reduce the effect of attributes. The detailed information on database creation can be found in [27] and `http://whdeng.cn/CALFW/`.

CPLFW [26]: The Cross-Pose LFW (CPLFW) dataset [26] is based on LFW with a focus on comparison pairs with pose differences. CPLFW contains 3000 genuine comparisons, while the negative pairs are selected of the same gender and race. More information can be found in [26] and http://whdeng.cn/CPLFW/.

XQLFW [12]: The Cross-Quality LFW (XQLFW) is derived from the LFW dataset. The XQLFW maximizes the quality difference, which contains only more realistic synthetically degraded images when necessary and is used to investigate the influence of image quality. XQLFW is licensed under the MIT License, and the detailed information can be found in [12] and https://martlgap.github.io/xqlfw/.

IJB-C [14]: The IARPA Janus Benchmark–C (IJB-C) [14] is a video-based face recognition dataset provided by the Nation Institute for Standards and Technology (NIST). It is an extension of the IJB-B [24] dataset with a total of 31,334 still images and 117,542 frames of 11,779 videos across 3531 identities. IJB-C is made available under different Creative Commons license variants. Detailed information on database creation can be found in [14] and https://www.nist.gov/programs-projects/face-challenges.

CASIA-WebFace [25]: CASIA-Webface consists of 494,141 face images from 10,757 different identities. A prepossessed (aligned and cropped) version of CASIA-WebFace is available in InsightFace (https://insightface.ai/) repository under Dataset-Zoo https://github.com/deepinsight/insightface/tree/master/recognition/_datasets_. The code and the databases of InsightFace is under MIT licence (https://github.com/deepinsight/insightface/blob/master/LICENSE).

MS1MV2 [5, 8]: The MS1MV2 is a refined version [5] of the MS-Celeb-1M [8] containing 5.8M images of 85K identities. A prepossessed (aligned and cropped) version of MS1MV2 is available in InsightFace (https://insightface.ai/ repository under Dataset-Zoo https://github.com/deepinsight/insightface/tree/master/recognition/_datasets_. The code and the database of InsightFace is under MIT licence (https://github.com/deepinsight/insightface/blob/master/LICENSE).

1.9. Use of existing assets

The results of the SOTA FIQA methods are produced based on the official code provided by each of these works. Table 2 presents the used SOTA methods along with link to their code repositories and licences.

The utilized FR models to report the verification performance at different quality rejection rates are ArcFace [5], ElasticFace (ElasticFace-Arc) [3], MagFace [16], and CurricularFace [11]. The link to the official code repository and license for each of the employed FR models are provided in the following:

- ArcFace [5] is provided under MIT license https://github.com/deepinsight/insightface/blob/master/LICENSE and the official pretrained model and code is published under the link https://github.com/deepinsight/insightface.
- MagFace [16] is provided under Apache License 2.0 https://github.com/IrvingMeng/MagFace/blob/main/LICENSE and the official pretrained model and code is published under the link https://github.com/IrvingMeng/MagFace.
- CurricularFace [11] is provided underMIT license https://github.com/HuangYG123/CurricularFace/blob/master/LICENSE and the official pretrained model and code is published under the link https://github.com/HuangYG123/CurricularFace/.
- ElasticFace [3] is provided under Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) license https://github.com/fdbtrs/ElasticFace/blob/main/README.md and the official pretrained model and code is published under the link lhttps://github.com/fdbtrs/ElasticFace.

1.10. Release of implementation and pre-trained models

The implementation and pre-trained models are released publicly under the Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) license. The code is available under https://github.com/fdbtrs/CR-FIQA

1.11. Potential societal impacts

We stress that our efforts in the advancement of FIQA and thus, face recognition, are aimed at enhancing the security, convenience, and life quality of the members of society, e.g. enabling convenient access to financial and health services [6] and enhancing the security of border checks within clear legal frameworks and users consent [1, 23]. We acknowledge, however reject, the possible malicious or illegal use of this and other machine learning-based technologies. Such a use of face recognition can involve the processing of face images for barometric recognition purposes out of legal framework and without the consent of the individual to create user/group profiles or the not consent use of face recognition in functionalities beyond the identity recognition itself [15].

Model	LFW Acc (%)	AgeDB-30 Acc (%)	CFP-FP Acc (%)	CALFW Acc (%)	CPLFW Acc (%)	IJB-C TAR at FAR1e-4
ArcFace [5]	99.82	98.15	98.27	95.45	92.08	96.28
ElasticFace [3]	99.80	98.35	98.67	96.17	93.27	96.49
MagFace [16]	99.83	98.17	98.46	96.15	92.87	96.65
CurricularFace [11]	99.80	98.32	98.37	96.20	93.13	96.58
CR-FIQA (L) (Ours)	99.80	98.17	98.49	96.15	92.90	96.23

Table 1. The verification performances of CR-FIQA (L) as feature extraction models on mainstream bookmarks and compared to the recent SOTA face recognition models.

Method	Code link	License
SER-FIQA [22]	https://github.com/pterhoer/FaceImageQuality	Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) license https://github.com/pterhoer/FaceImageQuality/blob/master/README.md
FaceQnet [9]	https://github.com/uam-biometrics/FaceQnet	no specific license provided by the authors
MagFace [16]	https://github.com/IrvingMeng/MagFace	Apache License 2.0 https://github.com/IrvingMeng/MagFace/blob/main/LICENSE
SDD-FIQA [19]	https://github.com/Tencent/TFace/tree/quality	Extension of Apache License Version 2.0 https://github.com/Tencent/TFace/blob/master/License.txt
rankIQ [4]	https://jschenthu.weebly.com/projects.html	This toolbox is made available for research purpose only as stated in README.md of code webpage
BRISQUE [17]	http://live.ece.utexas.edu/research/quality/BRISQUE_release.zip	Free usage is stated in the readme file contained in the project
PFE [21]	https://github.com/seasonSH/Probabilistic-Face-Embeddings	MIT License https://github.com/dmaniry/deepIQA/blob/master/LICENSE
rankIQA [13]	https://github.com/xialeiliu/RankIQA	MIT License https://github.com/xialeiliu/RankIQA/blob/master/LICENSE
DeepIQA [2]	https://github.com/dmaniry/deepIQA	MIT License https://github.com/dmaniry/deepIQA/blob/master/LICENSE

Table 2. The official released code links and licenses of the FIQA methods reported in this work. The results of the FIQA methods in the main submission are produced and reported based on their official released code and strictly following their licenses.

1.12. Limitation of the proposed approach

Unlike methods where the FIQA does not require to train a quality regression [16, 21, 22] our CR-FIQA requires a training a regression. However, this only required to be done once and the resulting model can be used to estimate quality for multiple efficiently FR models as demonstrated by the result.

References

[1] 740 ILCS/14. Biometric Information Privacy Act (BIPA). Public act 095-994, Illinois General Assembly, 2008. 3

[2] Sebastian Bosse, Dominique Maniry, Klaus-Robert Müller, Thomas Wiegand, and Wojciech Samek. Deep neural networks for no-reference and full-reference image quality assessment. *IEEE Trans. Image Process.*, 27(1):206–219, 2018. 4

[3] Fadi Boutros, Naser Damer, Florian Kirchbuchner, and Arjan Kuijper. Elasticface: Elastic margin loss for deep face recognition. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, CVPR Workshops 2022, New Orleans, LA, USA, June 19-20, 2022*, pages 1577–1586. IEEE, 2022. 3, 4

[4] Jiansheng Chen, Yu Deng, Gaocheng Bai, and Guangda Su. Face image quality assessment based on learning to rank. *IEEE Signal Process. Lett.*, 22(1):90–94, 2015. 4

[5] Jiankang Deng, Jia Guo, Niannan Xue, and Stefanos Zafeiriou. Arcface: Additive angular margin loss for deep face recognition. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2019, Long Beach, CA, USA, June 16-20, 2019*, pages 4690–4699. Computer Vision Foundation / IEEE, 2019. 1, 3, 4

[6] e-Aadhaar - Unique Identification Authority of India. https://eaadhaar.uidai.gov.in/, 2015. 3

[7] Eran Eidinger, Roee Enbar, and Tal Hassner. Age and gender estimation of unfiltered faces. *IEEE Trans. Inf. Forensics Secur.*, 9(12):2170–2179, 2014. 1, 2

[8] Yandong Guo, Lei Zhang, Yuxiao Hu, Xiaodong He, and Jianfeng Gao. Ms-celeb-1m: A dataset and benchmark for large-scale face recognition. In Bastian Leibe, Jiri Matas, Nicu Sebe, and Max Welling, editors, *Computer Vision - ECCV 2016 - 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part III*, volume 9907 of *Lecture Notes in Computer Science*, pages 87–102. Springer, 2016. 3

[9] Javier Hernandez-Ortega, Javier Galbally, Julian Fiérrez, and Laurent Beslay. Biometric quality: Review and application to face recognition with faceqnet. *CoRR*, abs/2006.03298, 2020. 4

[10] Gary B. Huang, Manu Ramesh, Tamara Berg, and Erik Learned-Miller. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Technical Report 07-49, University of Massachusetts, Amherst, October 2007. 2

[11] Yuge Huang, Yuhan Wang, Ying Tai, Xiaoming Liu, Pengcheng Shen, Shaoxin Li, Jilin Li, and Feiyue Huang. Curricularface: Adaptive curriculum learning loss for deep

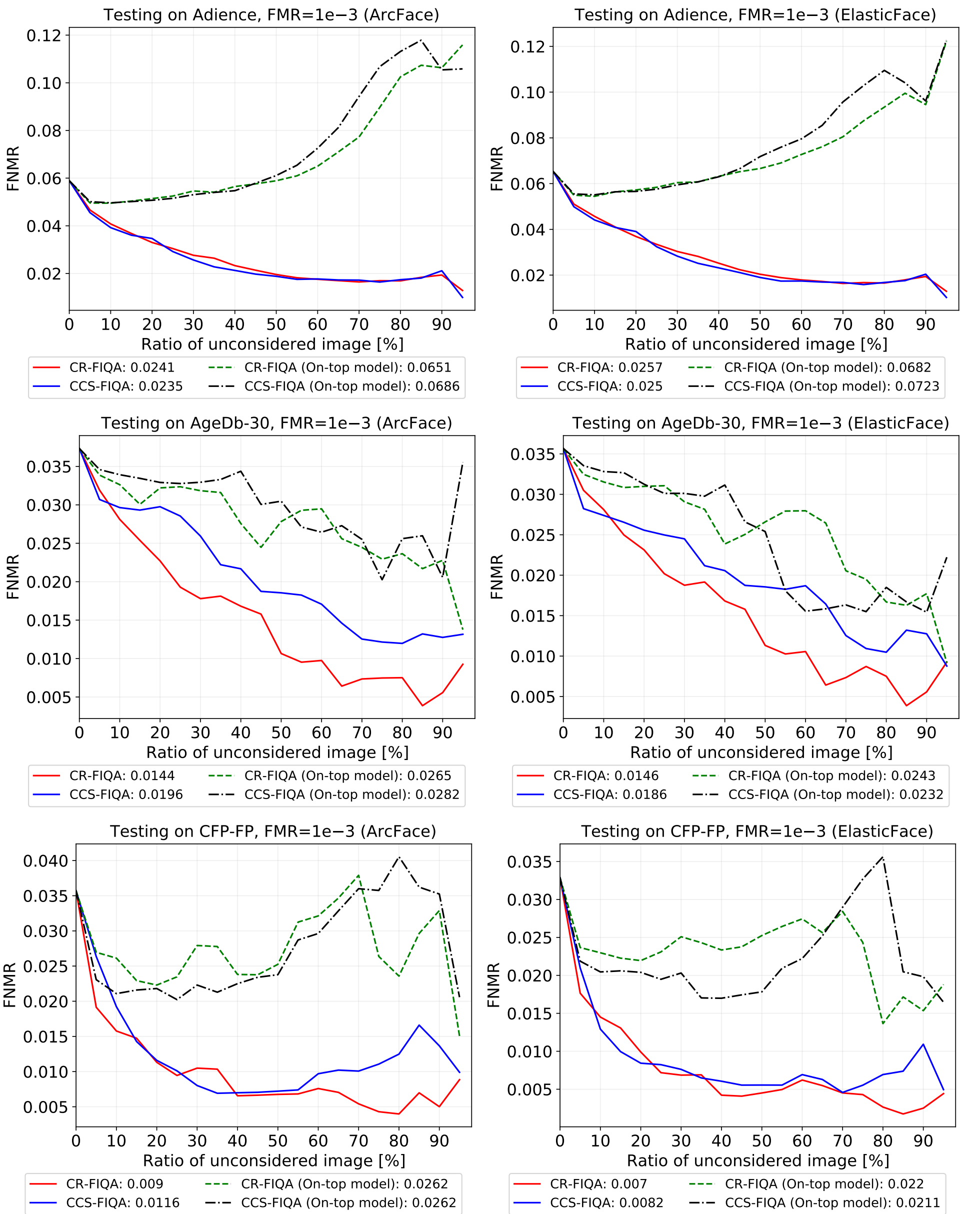


Figure 1. ERC comparison between CR-FIQA(S), CCS-FIQA(S), CR-FIQA(S) (On top) and CCS-FIQA(S) (On top). The plots show the effect of rejecting samples of lowest quality, on the verification error (FNMR at FMR1e-3) using ArcFace and ElasticFace models on Adience, AgeDb-30 and CFP-FP benchmarks . CR-FIQA(S) and CCS-FIQA(S) outperformed the on-top solutions, and CR-FIQA(S) performs generally better than CCS-FIQA(S) (curve decays faster with more rejected samples). AUC values are mentioned under the plots.

face recognition. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2020, Seattle, WA, USA, June 13-19, 2020*, pages 5900–5909. Computer Vision Foundation / IEEE, 2020. 3, 4

[12] Martin Knoche, Stefan Hörmann, and Gerhard Rigoll. Cross-quality LFW: A database for analyzing cross- resolution image face recognition in unconstrained environments. In *16th IEEE International Conference on Automatic Face and Gesture Recognition, FG 2021, Jodhpur, India, December 15-18, 2021*, pages 1–5. IEEE, 2021. 3

[13] Xialei Liu, Joost van de Weijer, and Andrew D. Bagdanov. Rankiqa: Learning from rankings for no-reference image quality assessment. In *IEEE International Conference on Computer Vision, ICCV 2017, Venice, Italy, October 22-29, 2017*, pages 1040–1049. IEEE Computer Society, 2017. 4

[14] Brianna Maze, Jocelyn C. Adams, James A. Duncan, Nathan D. Kalka, Tim Miller, Charles Otto, Anil K. Jain, W. Tyler Niggel, Janet Anderson, Jordan Cheney, and Patrick

Grother. IARPA janus benchmark - C: face dataset and protocol. In *2018 International Conference on Biometrics, ICB 2018, Gold Coast, Australia, February 20-23, 2018*, pages 158–165. IEEE, 2018. 2, 3

[15] Blaz Meden, Peter Rot, Philipp Terhörst, Naser Damer, Arjan Kuijper, Walter J. Scheirer, Arun Ross, Peter Peer, and Vitomir Struc. Privacy-enhancing face biometrics: A comprehensive survey. *IEEE Trans. Inf. Forensics Secur.*, 16:4147–4183, 2021. 3

[16] Qiang Meng, Shichao Zhao, Zhida Huang, and Feng Zhou. Magface: A universal representation for face recognition and quality assessment. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2021, virtual, June 19-25, 2021*, pages 14225–14234. Computer Vision Foundation / IEEE, 2021. 3, 4

[17] Anish Mittal, Anush Krishna Moorthy, and Alan Conrad Bovik. No-reference image quality assessment in the spatial domain. *IEEE Trans. Image Process.*, 21(12):4695–4708, 2012. 4

[18] Stylianos Moschoglou, Athanasios Papaioannou, Christos Sagonas, Jiankang Deng, Irene Kotsia, and Stefanos Zafeiriou. Agedb: The first manually collected, in-the-wild age database. In *2017 IEEE CVPRW, CVPR Workshops 2017, Honolulu, HI, USA, July 21-26, 2017*, pages 1997–2005. IEEE Computer Society, 2017. 1, 2

[19] Fu-Zhao Ou, Xingyu Chen, Ruixin Zhang, Yuge Huang, Shaoxin Li, Jilin Li, Yong Li, Liujuan Cao, and Yuan-Gen Wang. SDD-FIQA: unsupervised face image quality assessment with similarity distribution distance. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2021, virtual, June 19-25, 2021*, pages 7670–7679. Computer Vision Foundation / IEEE, 2021. 4

[20] Soumyadip Sengupta, Jun-Cheng Chen, Carlos Domingo Castillo, Vishal M. Patel, Rama Chellappa, and David W. Jacobs. Frontal to profile face verification in the wild. In *2016 IEEE Winter Conference on Applications of Computer Vision, WACV 2016, Lake Placid, NY, USA, March 7-10, 2016*, pages 1–9. IEEE Computer Society, 2016. 1, 2

[21] Yichun Shi and Anil K. Jain. Probabilistic face embeddings. In *2019 IEEE/CVF International Conference on Computer Vision, ICCV 2019, Seoul, Korea (South), October 27 - November 2, 2019*, pages 6901–6910. IEEE, 2019. 4

[22] Philipp Terhörst, Jan Niklas Kolf, Naser Damer, Florian Kirchbuchner, and Arjan Kuijper. SER-FIQ: unsupervised estimation of face image quality based on stochastic embedding robustness. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2020, Seattle, WA, USA, June 13-19, 2020*, pages 5650–5659. Computer Vision Foundation / IEEE, 2020. 4

[23] Paul Voigt and Axel von dem Bussche. *The EU General Data Protection Regulation (GDPR): A Practical Guide*. 1st edition, 2017. 3

[24] Cameron Whitelam, Emma Taborsky, Austin Blanton, Brianna Maze, Jocelyn C. Adams, Tim Miller, Nathan D. Kalka, Anil K. Jain, James A. Duncan, Kristen Allen, Jordan Cheney, and Patrick Grother. IARPA janus benchmark-b face dataset. In *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops, CVPR Workshops 2017, Honolulu, HI, USA, July 21-26, 2017*, pages 592–600. IEEE Computer Society, 2017. 3

[25] Dong Yi, Zhen Lei, Shengcai Liao, and Stan Z. Li. Learning face representation from scratch. *CoRR*, abs/1411.7923, 2014. 1, 3

[26] T. Zheng and W. Deng. Cross-pose lfw: A database for studying cross-pose face recognition in unconstrained environments. Technical Report 18-01, Beijing University of Posts and Telecommunications, February 2018. 2, 3

[27] Tianyue Zheng, Weihong Deng, and Jiani Hu. Cross-age LFW: A database for studying cross-age face recognition in unconstrained environments. *CoRR*, abs/1708.08197, 2017. 1, 2

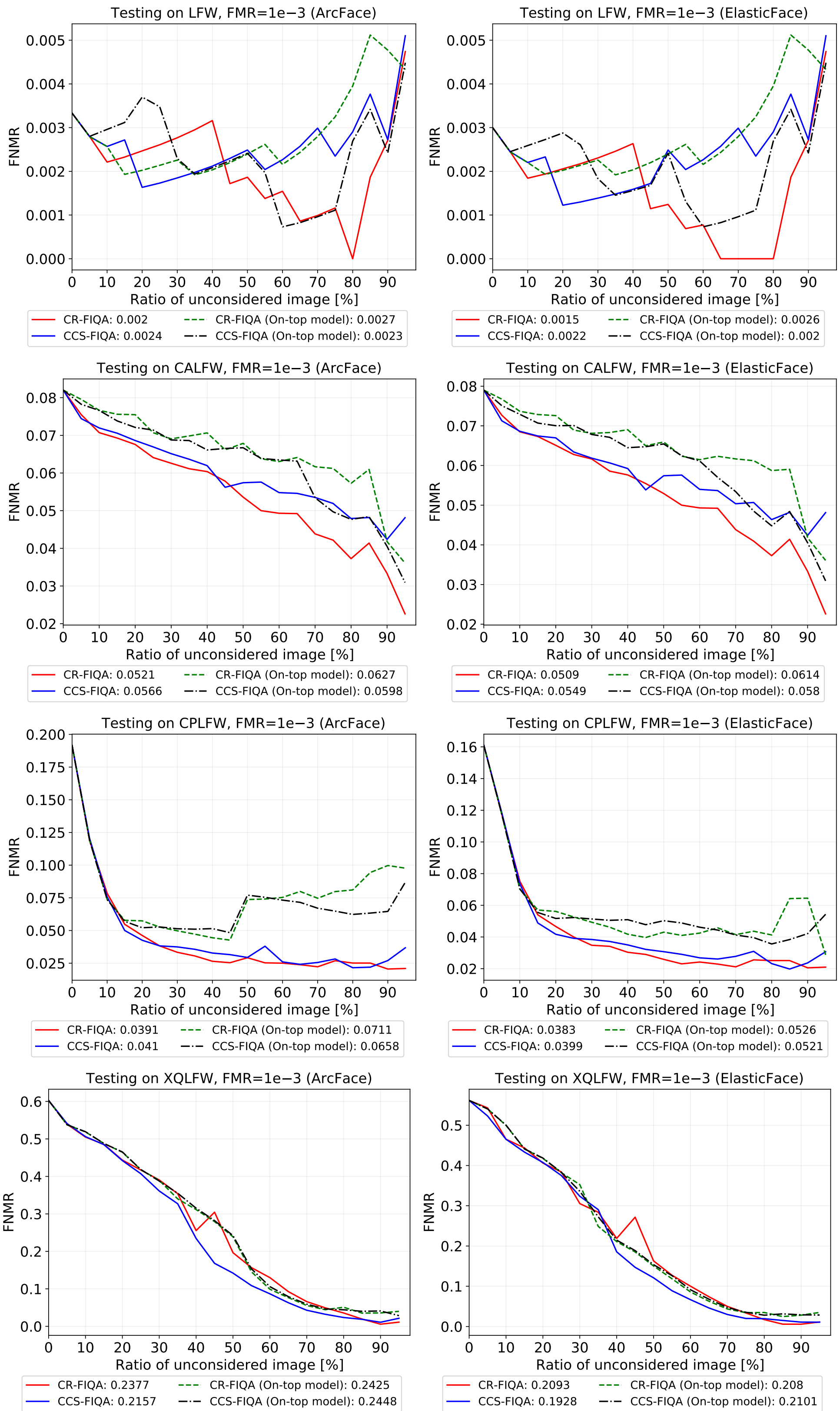


Figure 2. ERC comparison between CR-FIQA(S), CCS-FIQA(S), CR-FIQA(S) (On top) and CCS-FIQA(S) (On top). The plots show the effect of rejecting samples of lowest quality, on the verification error (FNMR at FMR1e-3) using ArcFace and ElasticFace models on LFW, CALFW, CPLFW and XQLFW benchmarks . CR-FIQA(S) and CCS-FIQA(S) outperformed the on-top solutions, and CR-FIQA(S) performs generally better than CCS-FIQA(S) (curve decays faster with more rejected samples). AUC values are mentioned under the plots.

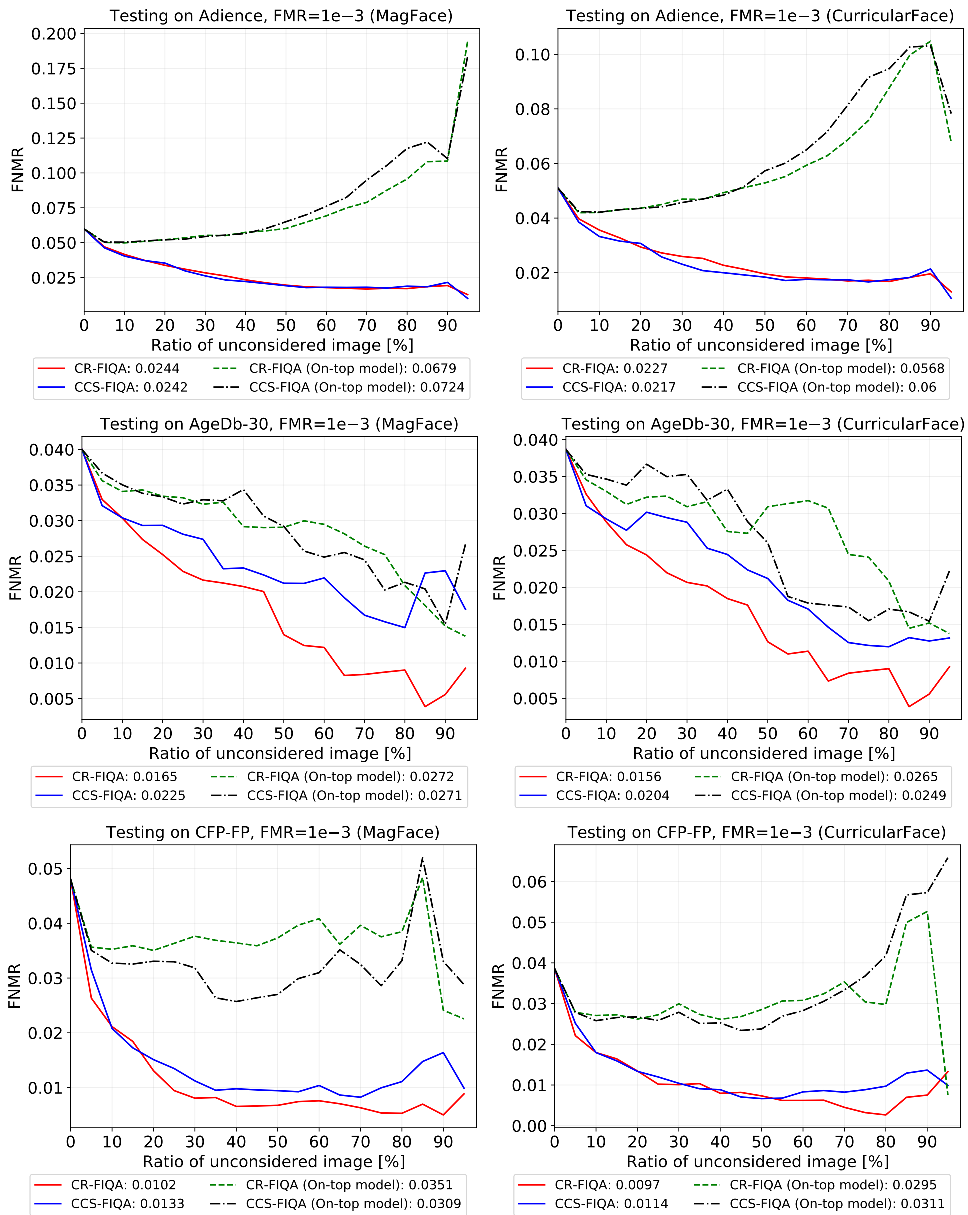


Figure 3. ERC comparison between CR-FIQA(S), CCS-FIQA(S), CR-FIQA(S) (On top) and CCS-FIQA(S) (On top). The plots show the effect of rejecting samples of lowest quality, on the verification error (FNMR at FMR1e-3) using MagFace and CurricularFace models on Adience, AgeDb-30 and CFP-FP benchmarks . CR-FIQA(S) and CCS-FIQA(S) outperformed the on-top solutions, and CR-FIQA(S) performs generally better than CCS-FIQA(S) (curve decays faster with more rejected samples). AUC values are mentioned under the plots.

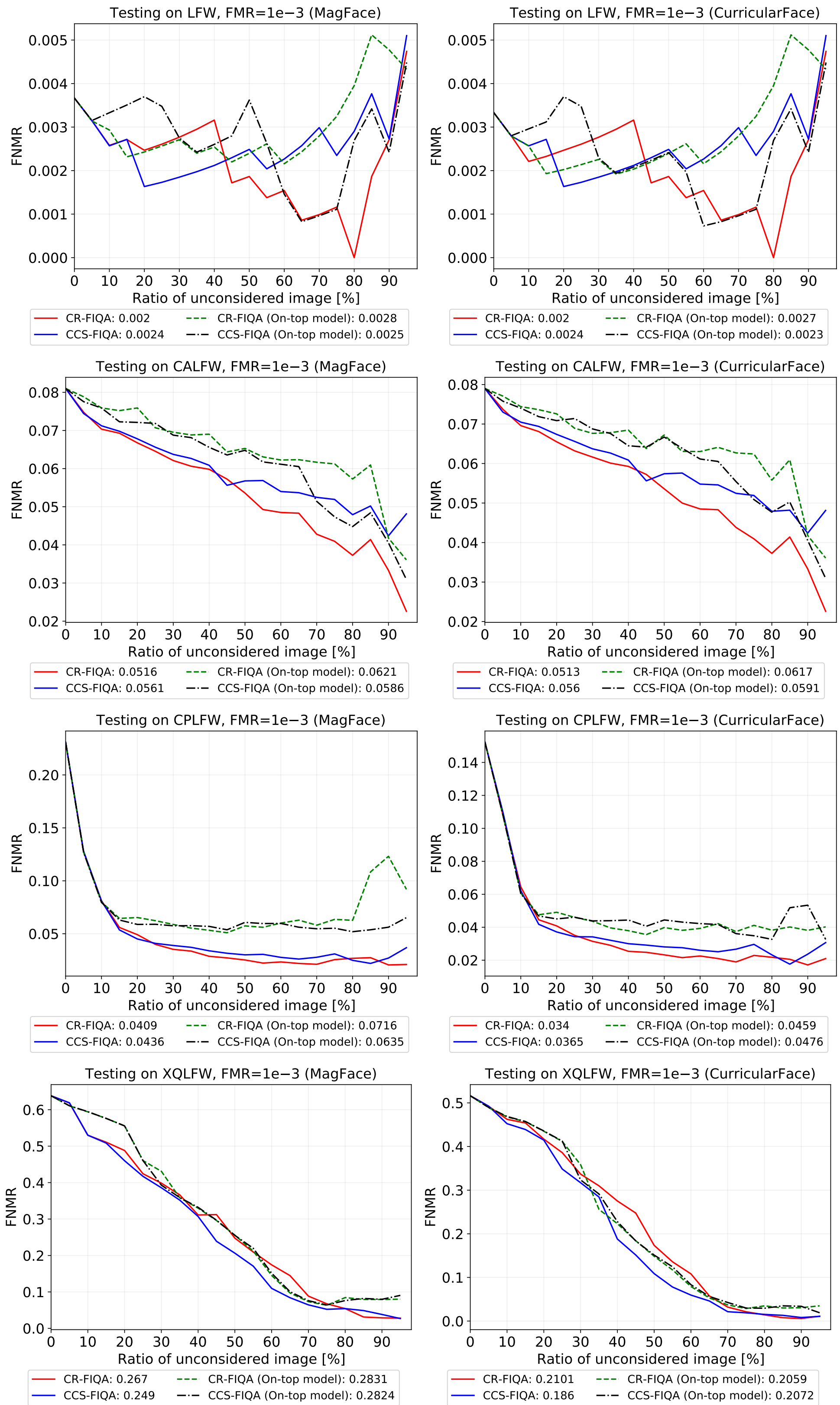


Figure 4. ERC comparison between CR-FIQA(S), CCS-FIQA(S), CR-FIQA(S) (On top) and CCS-FIQA(S) (On top). The plots show the effect of rejecting samples of lowest quality, on the verification error (FNMR at FMR1e-3) using MagFace and CurricularFace models on LFW, CALFW, CPLFW and XQLFW benchmarks . CR-FIQA(S) and CCS-FIQA(S) outperformed the on-top solutions, and CR-FIQA(S) performs generally better than CCS-FIQA(S) (curve decays faster with more rejected samples). AUC values are mentioned under the plots.

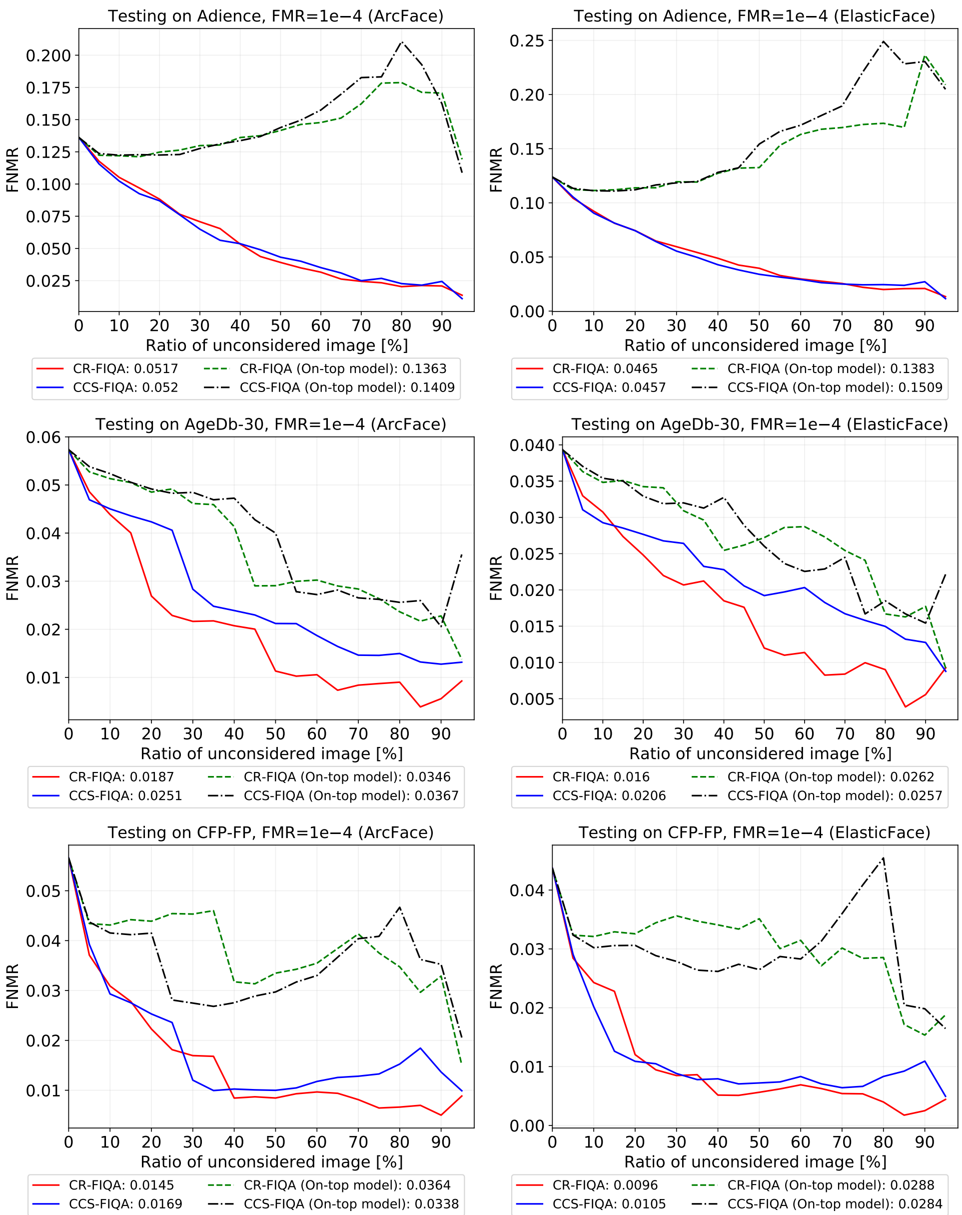


Figure 5. ERC comparison between CR-FIQA(S), CCS-FIQA(S), CR-FIQA(S) (On top) and CCS-FIQA(S) (On top). The plots show the effect of rejecting samples of lowest quality, on the verification error (FNMR at FMR1e-4) using ArcFace and ElasticFace models on Adience, AgeDb-30 and CFP-FP benchmarks . CR-FIQA(S) and CCS-FIQA(S) outperformed the on-top solutions, and CR-FIQA(S) performs generally better than CCS-FIQA(S) (curve decays faster with more rejected samples). AUC values are mentioned under the plots.

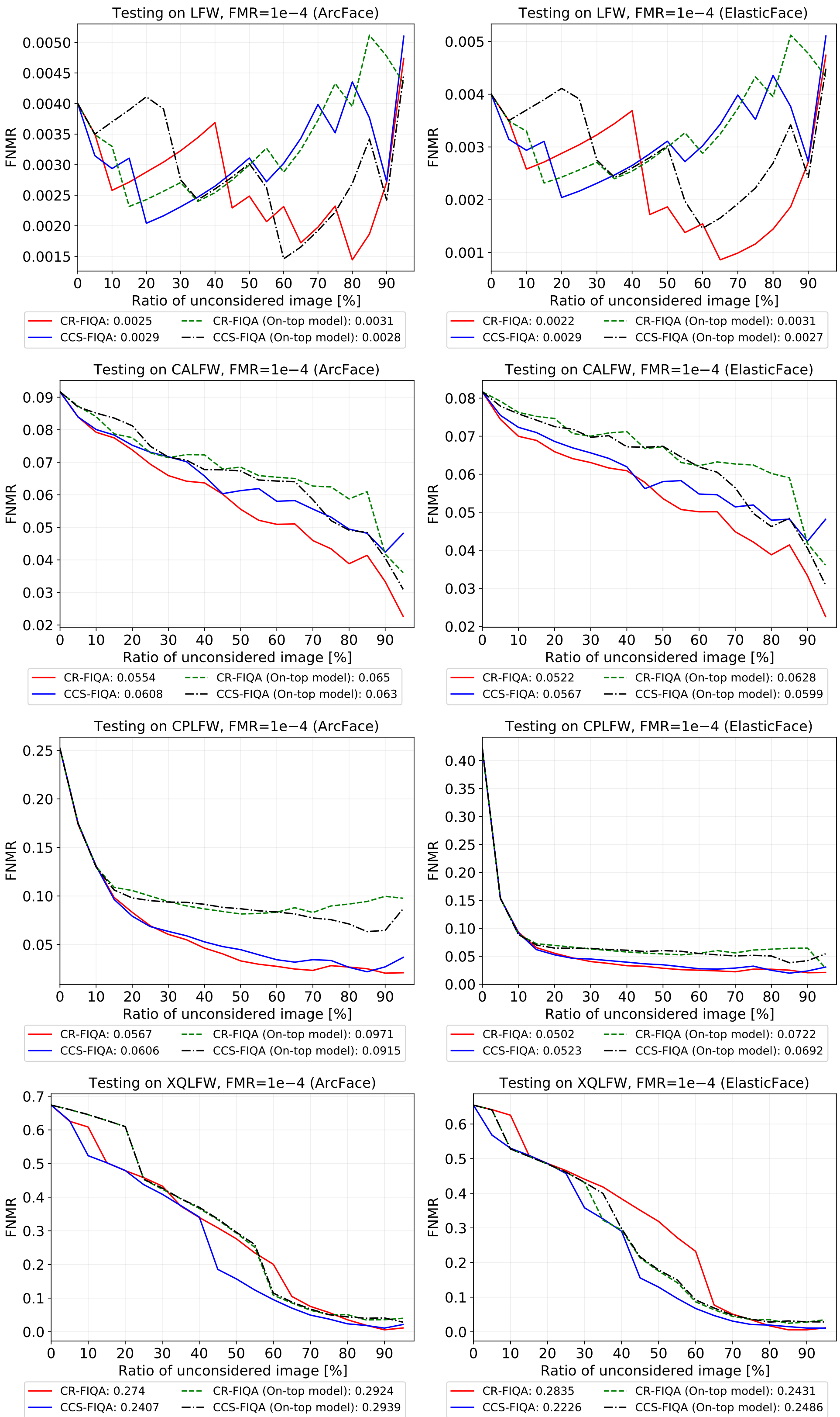


Figure 6. ERC comparison between CR-FIQA(S), CCS-FIQA(S), CR-FIQA(S) (On top) and CCS-FIQA(S) (On top). The plots show the effect of rejecting samples of lowest quality, on the verification error (FNMR at FMR1e-4) using ArcFace and ElasticFace models on LFW, CALFW, CPLFW and XQLFW benchmarks . CR-FIQA(S) and CCS-FIQA(S) outperformed the on-top solutions, and CR-FIQA(S) performs generally better than CCS-FIQA(S) (curve decays faster with more rejected samples). AUC values are mentioned under the plots.

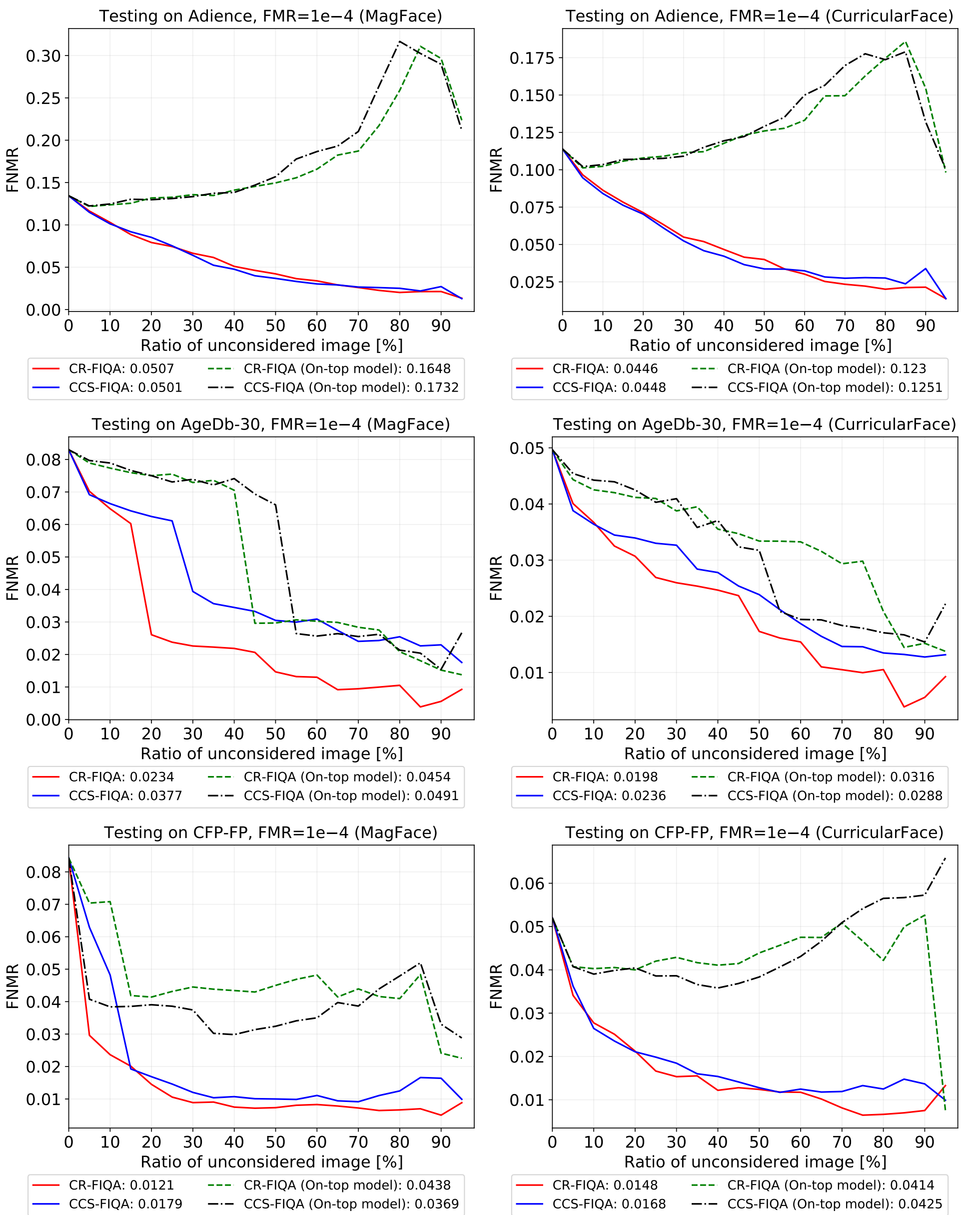


Figure 7. ERC comparison between CR-FIQA(S), CCS-FIQA(S), CR-FIQA(S) (On top) and CCS-FIQA(S) (On top). The plots show the effect of rejecting samples of lowest quality, on the verification error (FNMR at FMR1e-4) using MagFace and CurricularFace on Adience, AgeDb-30 and CFP-FP benchmarks . CR-FIQA(S) and CCS-FIQA(S) outperformed the on-top solutions, and CR-FIQA(S) performs generally better than CCS-FIQA(S) (curve decays faster with more rejected samples). AUC values are mentioned under the plots.

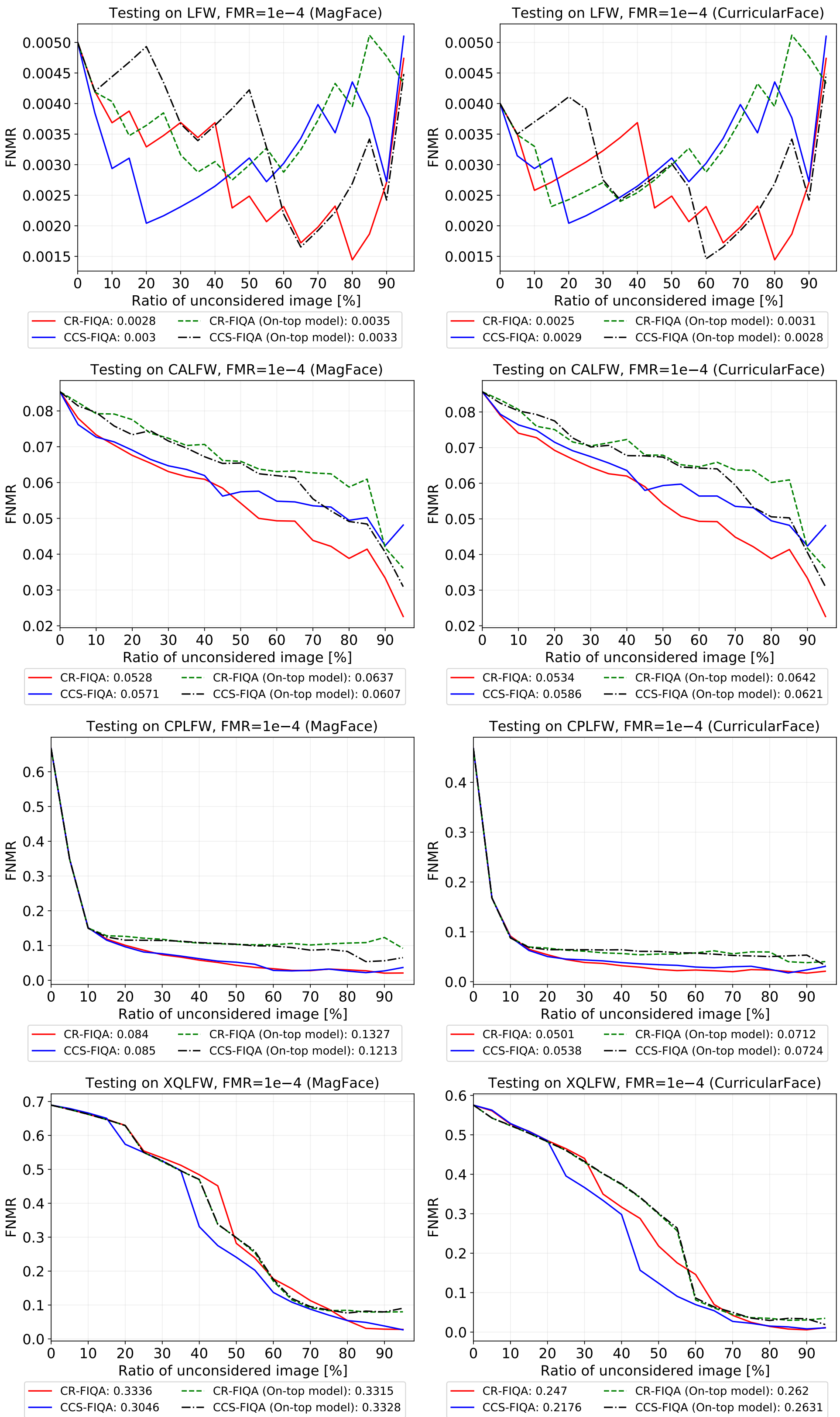
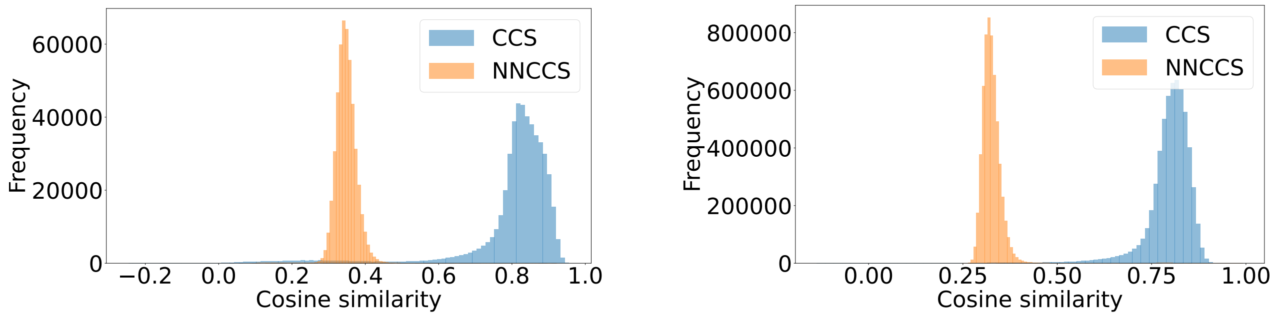
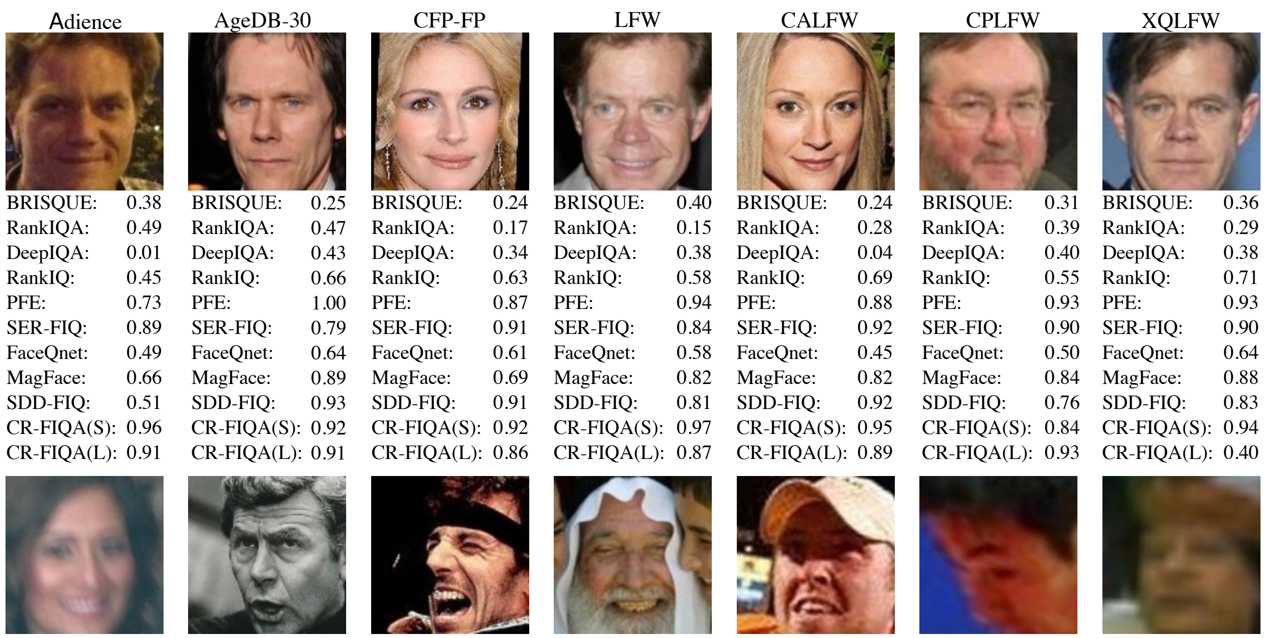


Figure 8. ERC comparison between CR-FIQA(S), CCS-FIQA(S), CR-FIQA(S) (On top) and CCS-FIQA(S) (On top). The plots show the effect of rejecting samples of lowest quality, on the verification error (FNMR at FMR1e-4) using MagFace and CurricularFace on LFW, CALFW, CPLFW and XQLFW benchmarks . CR-FIQA(S) and CCS-FIQA(S) outperformed the on-top solutions, and CR-FIQA(S) performs generally better than CCS-FIQA(S) (curve decays faster with more rejected samples). AUC values are mentioned under the plots.



(a) (b)

Figure 9. Histogram of the cosine similarity between training samples and their class centers (CCS) and nearest negative class centers (NNCCS). Similarity values in plot 9a are obtained from ResNet-50 trained on CASIA-WebFace (R50(CASIA)) and the ones in plot 9b are obtained from ResNet-100 trained on MS1MV2 (R100(MS1MV2)). In both models/databases, the values of CCS and NNCCS vary between different samples.



	Adience	AgeDB-30	CFP-FP	LFW	CALFW	CPLFW	XQLFW
BRISQUE:	0.38	0.25	0.24	0.40	0.24	0.31	0.36
RankIQA:	0.49	0.47	0.17	0.15	0.28	0.39	0.29
DeepIQA:	0.01	0.43	0.34	0.38	0.04	0.40	0.38
RankIQ:	0.45	0.66	0.63	0.58	0.69	0.55	0.71
PFE:	0.73	1.00	0.87	0.94	0.88	0.93	0.93
SER-FIQ:	0.89	0.79	0.91	0.84	0.92	0.90	0.90
FaceQnet:	0.49	0.64	0.61	0.58	0.45	0.50	0.64
MagFace:	0.66	0.89	0.69	0.82	0.82	0.84	0.88
SDD-FIQ:	0.51	0.93	0.91	0.81	0.92	0.76	0.83
CR-FIQA(S):	0.96	0.92	0.92	0.97	0.95	0.84	0.94
CR-FIQA(L):	0.91	0.91	0.86	0.87	0.89	0.93	0.40

	Adience	AgeDB-30	CFP-FP	LFW	CALFW	CPLFW	XQLFW
BRISQUE:	0.44	0.32	0.29	0.44	0.33	0.41	0.77
RankIQA:	0.40	0.31	0.43	0.22	0.57	0.43	0.24
DeepIQA:	0.31	0.16	0.76	0.39	0.53	0.48	0.62
RankIQ:	0.28	0.41	0.23	0.44	0.28	0.18	0.19
PFE:	0.60	0.65	0.49	0.74	0.36	0.43	0.53
SER-FIQ:	0.77	0.69	0.76	0.61	0.76	0.42	0.80
FaceQnet:	0.53	0.28	0.23	0.31	0.16	0.56	0.42
MagFace:	0.31	0.42	0.39	0.41	0.30	0.50	0.41
SDD-FIQ:	0.33	0.67	0.41	0.52	0.42	0.18	0.22
CR-FIQA(S):	0.60	0.34	0.30	0.38	0.36	0.25	0.31
CR-FIQA(L):	0.69	0.77	0.58	0.69	0.57	0.24	0.59

Figure 10. Samples image of the evaluation benchmarks with quality score values obtained from our CR-FIQA the SOTA methods. Noting that this figure only reflects samples with quality scores and do not necessary reflect overall performance.

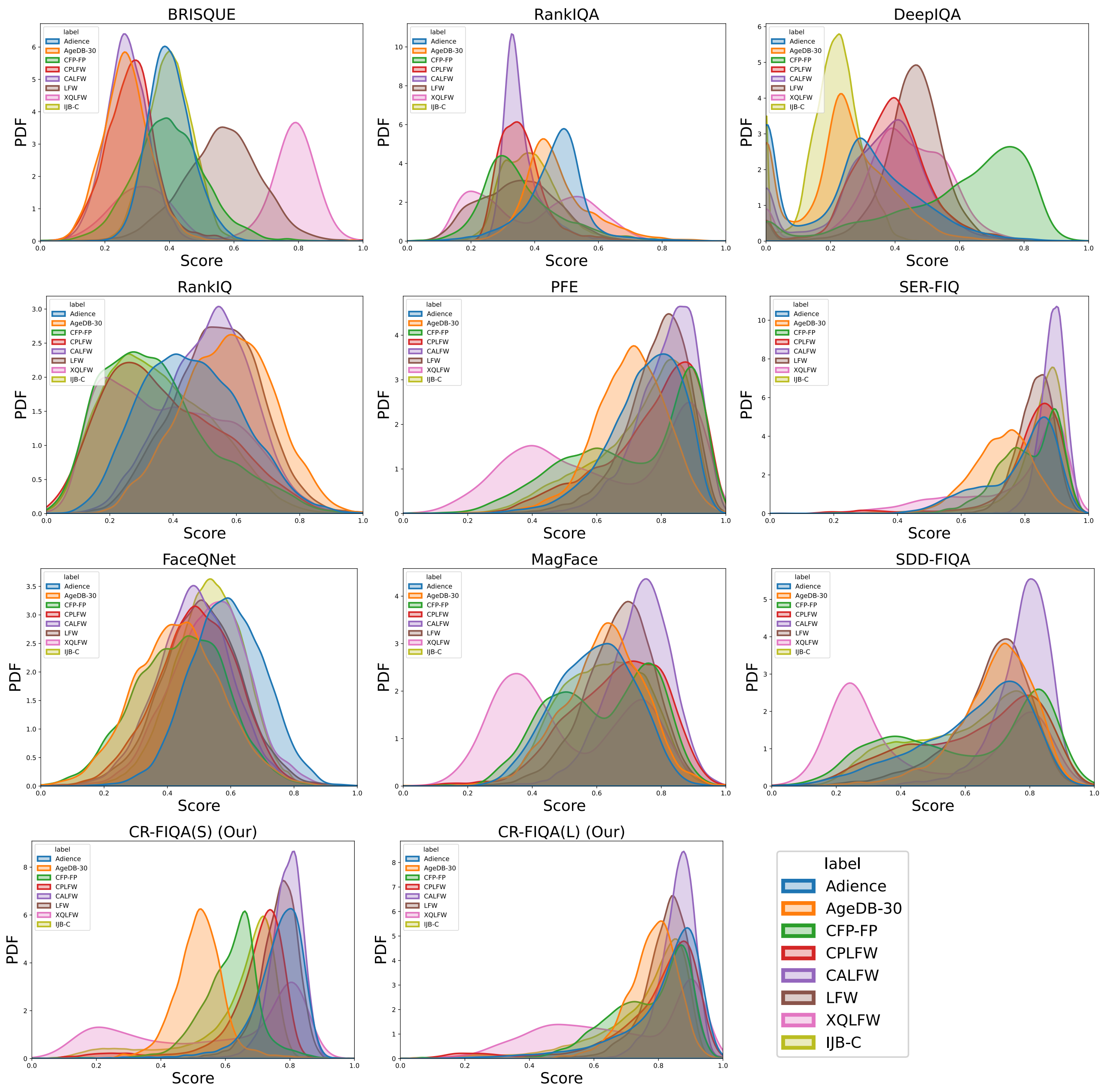


Figure 11. Quality score distribution of the evaluation benchmarks achieved by our CR-FIQA and the SOTA methods (all normalized to have values between 0 and 1).

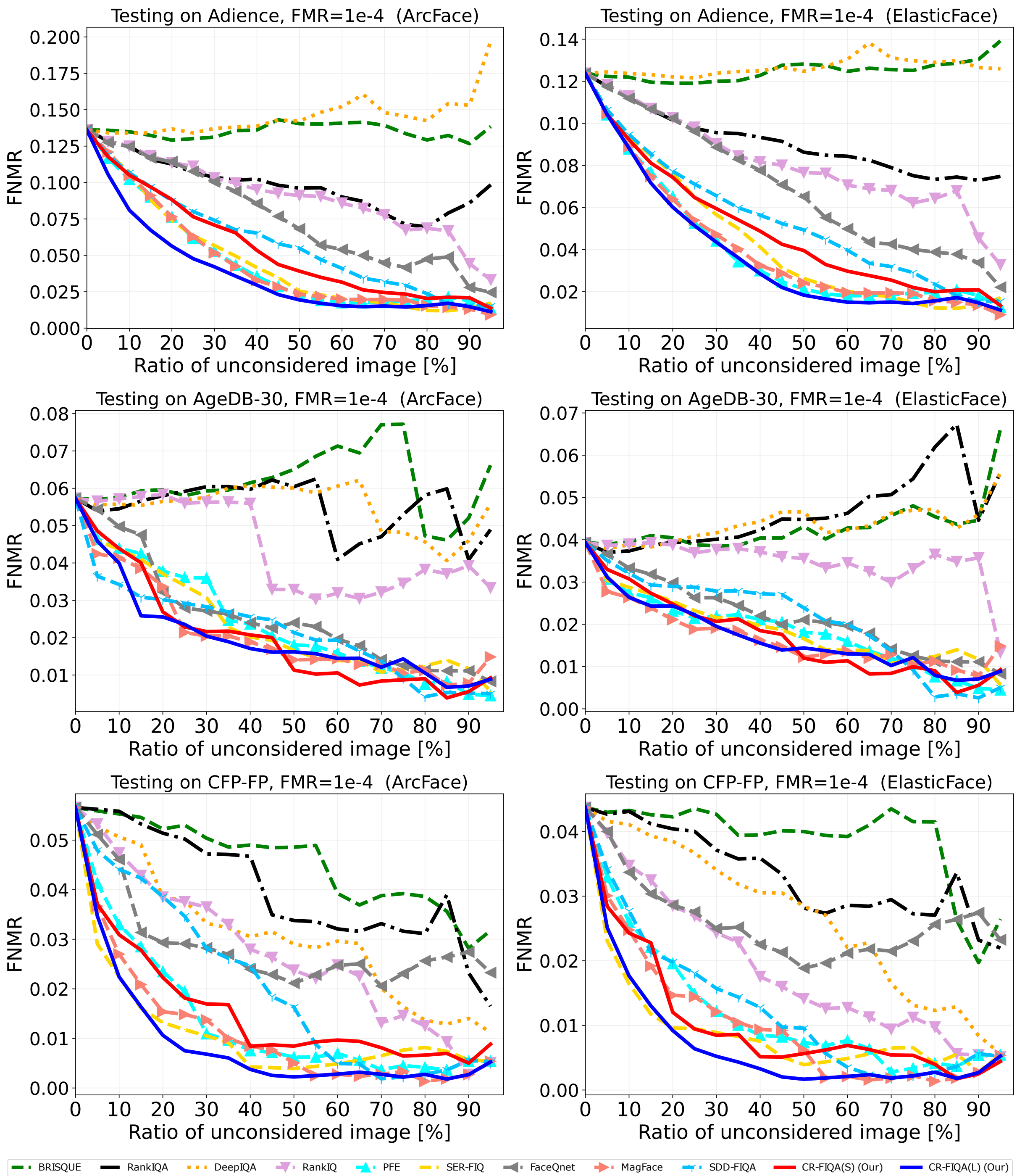


Figure 12. ERC (FNMR at FMR1e-4 vs reject) curves for ArcFace and ElasticFace on Adience, AgeDB-30 and CFP-FP benchmarks. The proposed CR-FIQA(L) and CR-FIQA(S) are marked with solid blue and red lines, respectively. CR-FIQA leads to lower verification error, when rejecting a fraction of images, of the lowest quality, in comparison to SOTA methods (faster decaying curve) under most experimental settings.

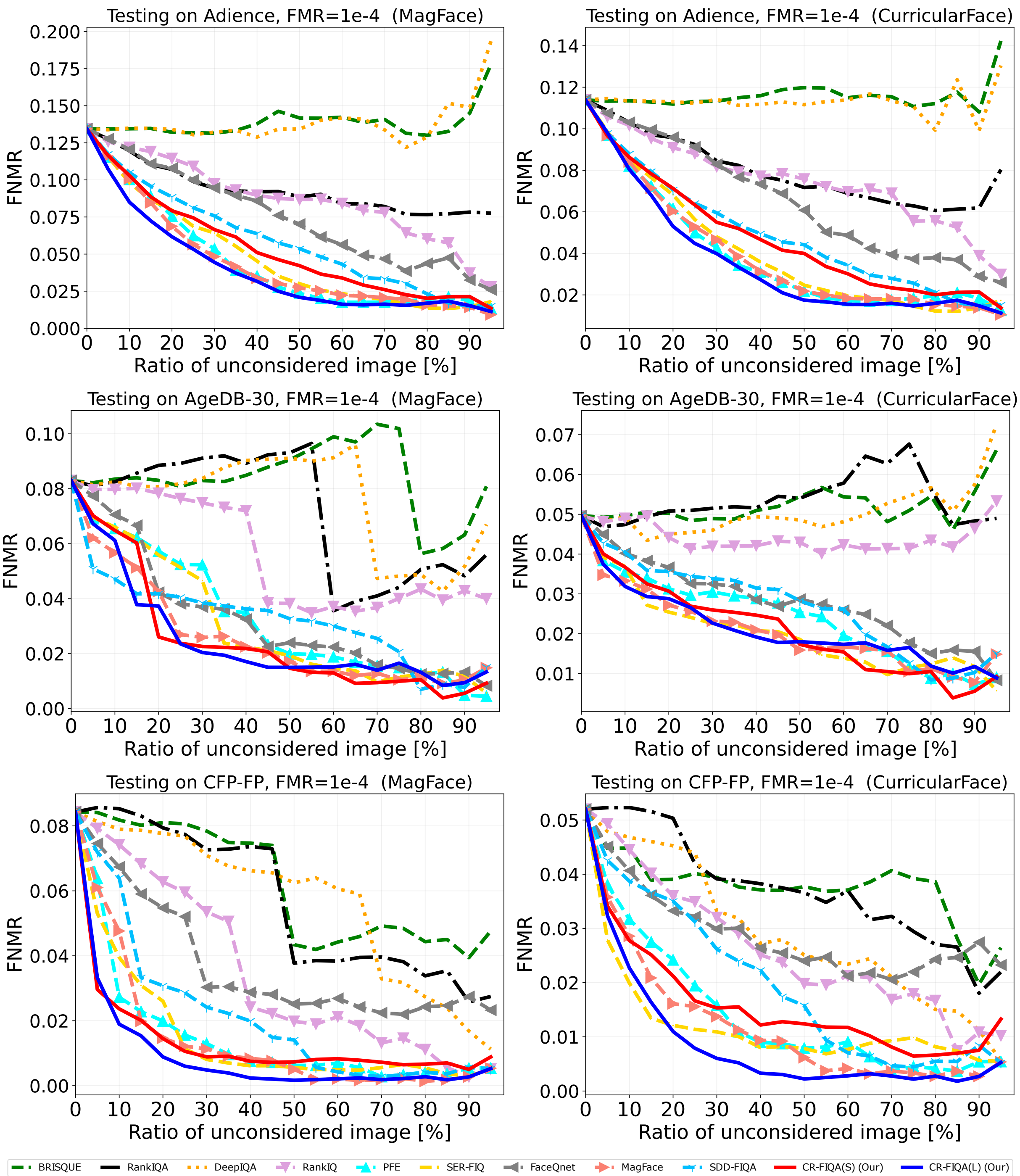


Figure 13. ERC (FNMR at FMR1e-4 vs reject) curves for MagFace and CurricularFace on Adience, AgeDB-30 and CFP-FP benchmarks. The proposed CR-FIQA(L) and CR-FIQA(S) are marked with solid blue and red lines, respectively. CR-FIQA leads to lower verification error, when rejecting a fraction of images, of the lowest quality, in comparison to SOTA methods (faster decaying curve) under most experimental settings.

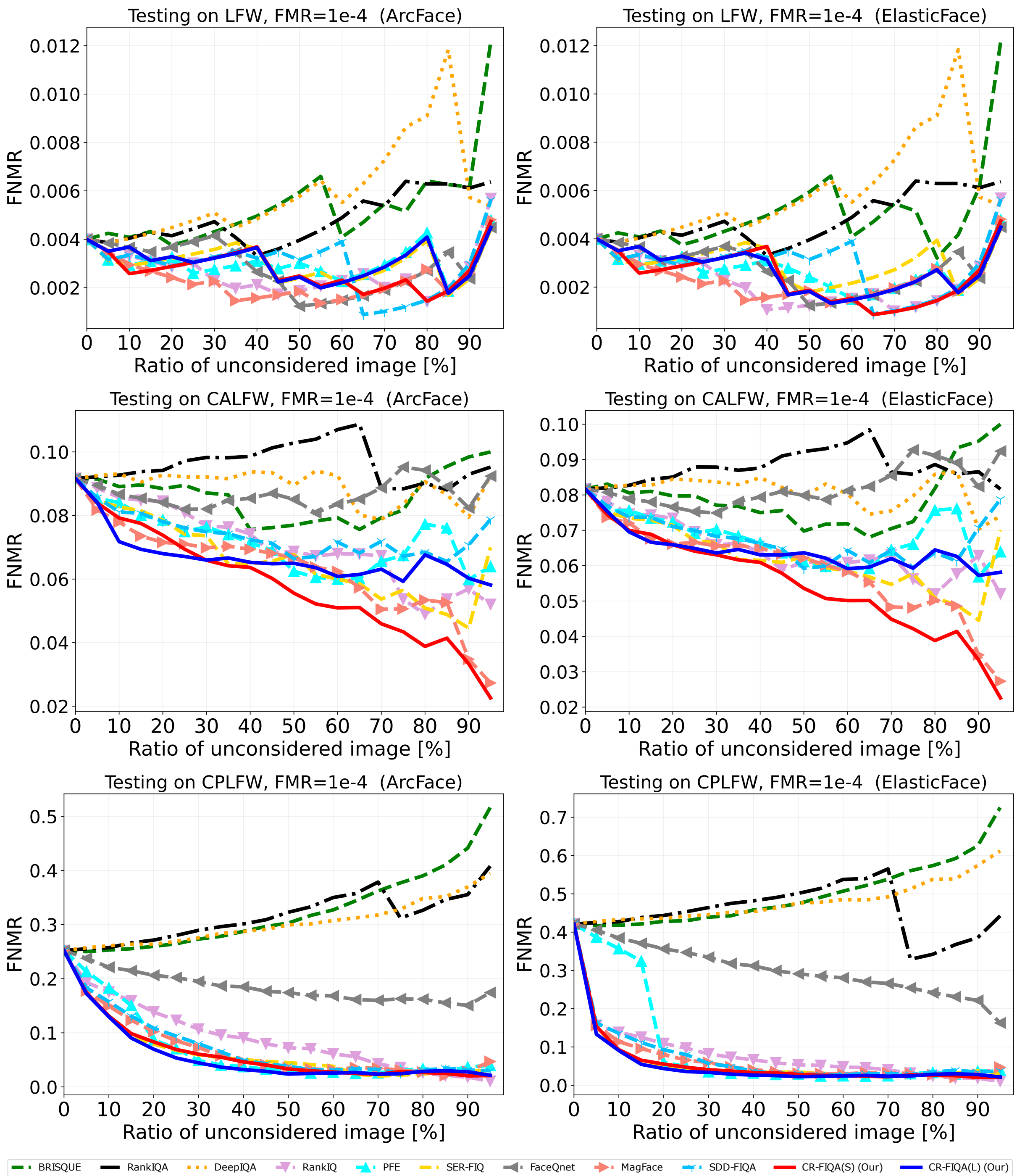


Figure 14. ERC (FNMR at FMR1e-4 vs reject) curves for ArcFace and ElasticFace on LFW, CALFW and CPLFW benchmarks. The proposed CR-FIQA(L) and CR-FIQA(S) are marked with solid blue and red lines, respectively. CR-FIQA leads to lower verification error, when rejecting a fraction of images, of the lowest quality, in comparison to SOTA methods (faster decaying curve) under most experimental settings.

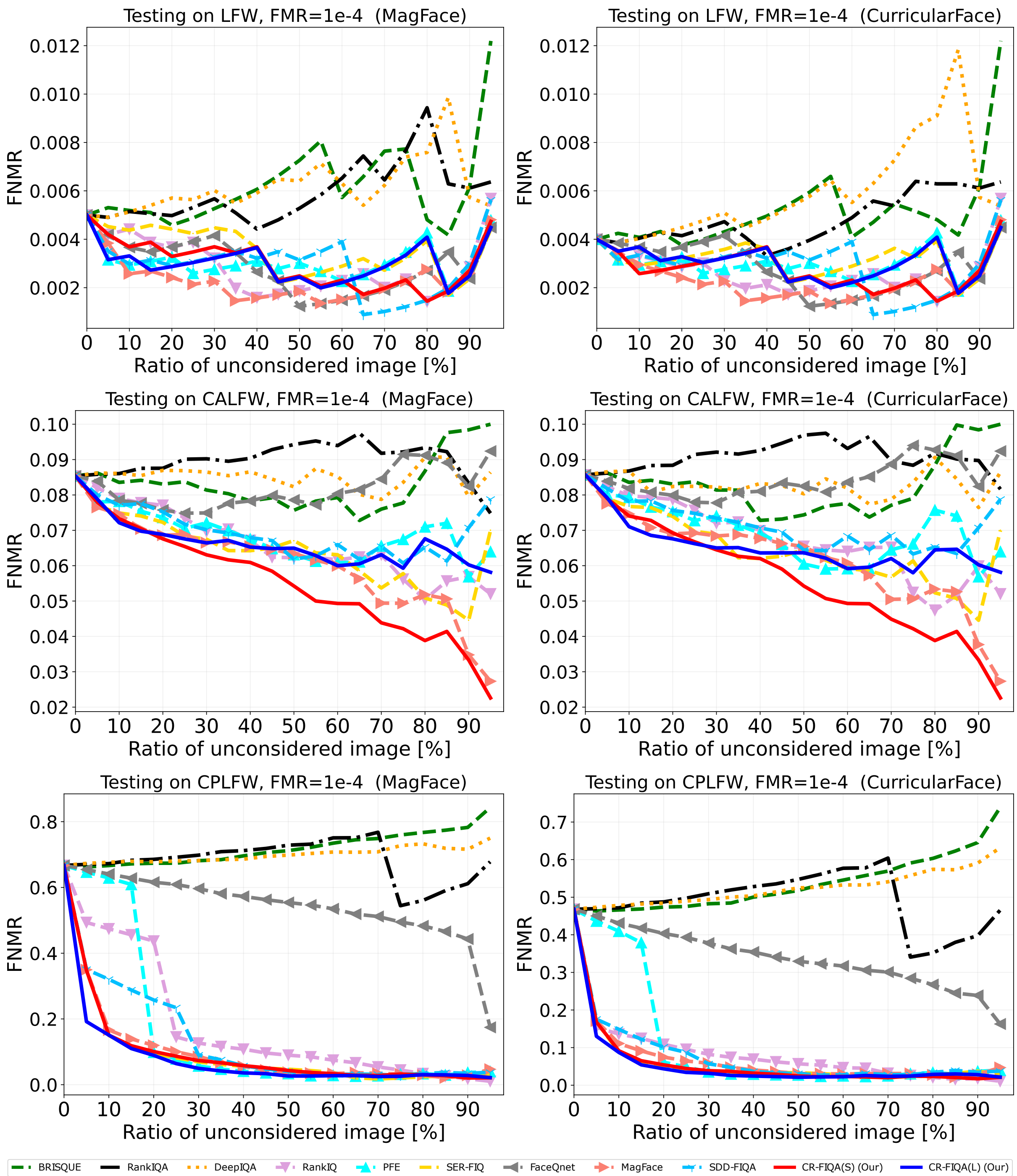


Figure 15. ERC (FNMR at FMR1e-4 vs reject) curves for MagFace and CurricularFace on LFW, CALFW and CPLFW benchmarks. The proposed CR-FIQA(L) and CR-FIQA(S) are marked with solid blue and red lines, respectively. CR-FIQA leads to lower verification error, when rejecting a fraction of images, of the lowest quality, in comparison to SOTA methods (faster decaying curve) under most experimental settings.

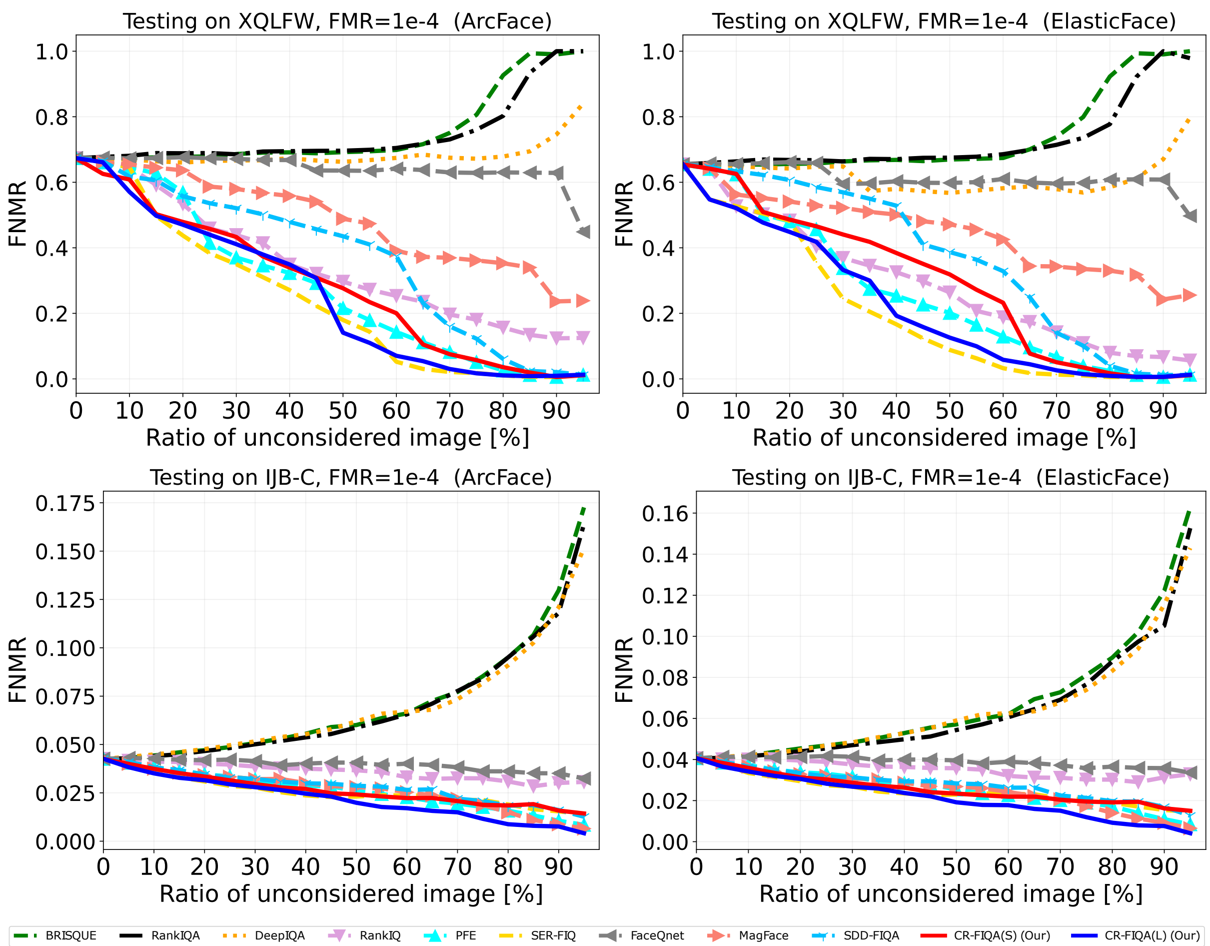


Figure 16. ERC (FNMR at FMR1e-4 vs reject) curves for ArcFace and ElasticFace on XQLFW and IJB-C benchmarks. The proposed CR-FIQA(L) and CR-FIQA(S) are marked with solid blue and red lines, respectively. CR-FIQA leads to lower verification error, when rejecting a fraction of images, of the lowest quality, in comparison to SOTA methods (faster decaying curve) under most experimental settings.

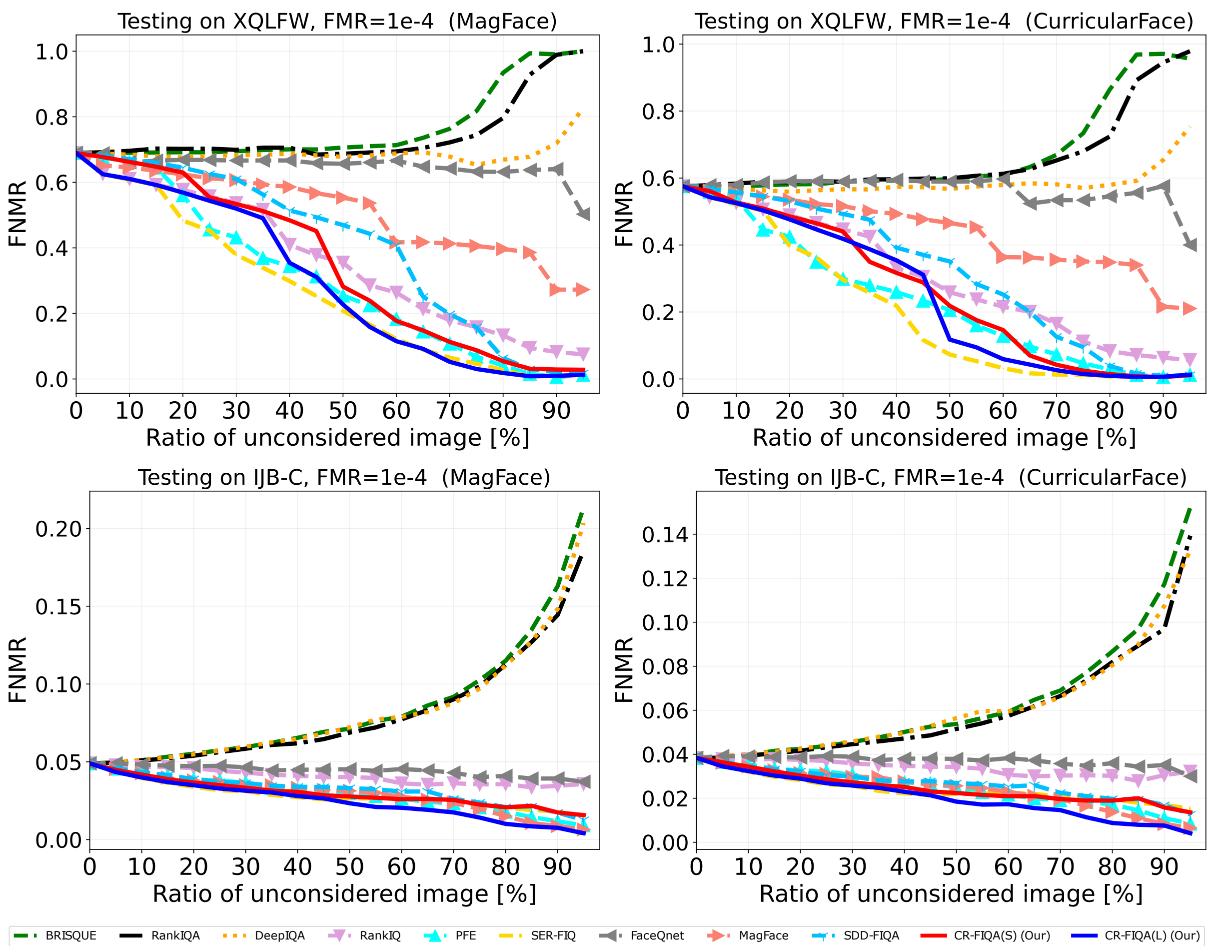


Figure 17. ERC (FNMR at FMR1e-4 vs reject) curves for MagFace and CurricularFace on XQLFW and IJB-C benchmarks. The proposed CR-FIQA(L) and CR-FIQA(S) are marked with solid blue and red lines, respectively. CR-FIQA leads to lower verification error, when rejecting a fraction of images, of the lowest quality, in comparison to SOTA methods (faster decaying curve) under most experimental settings.

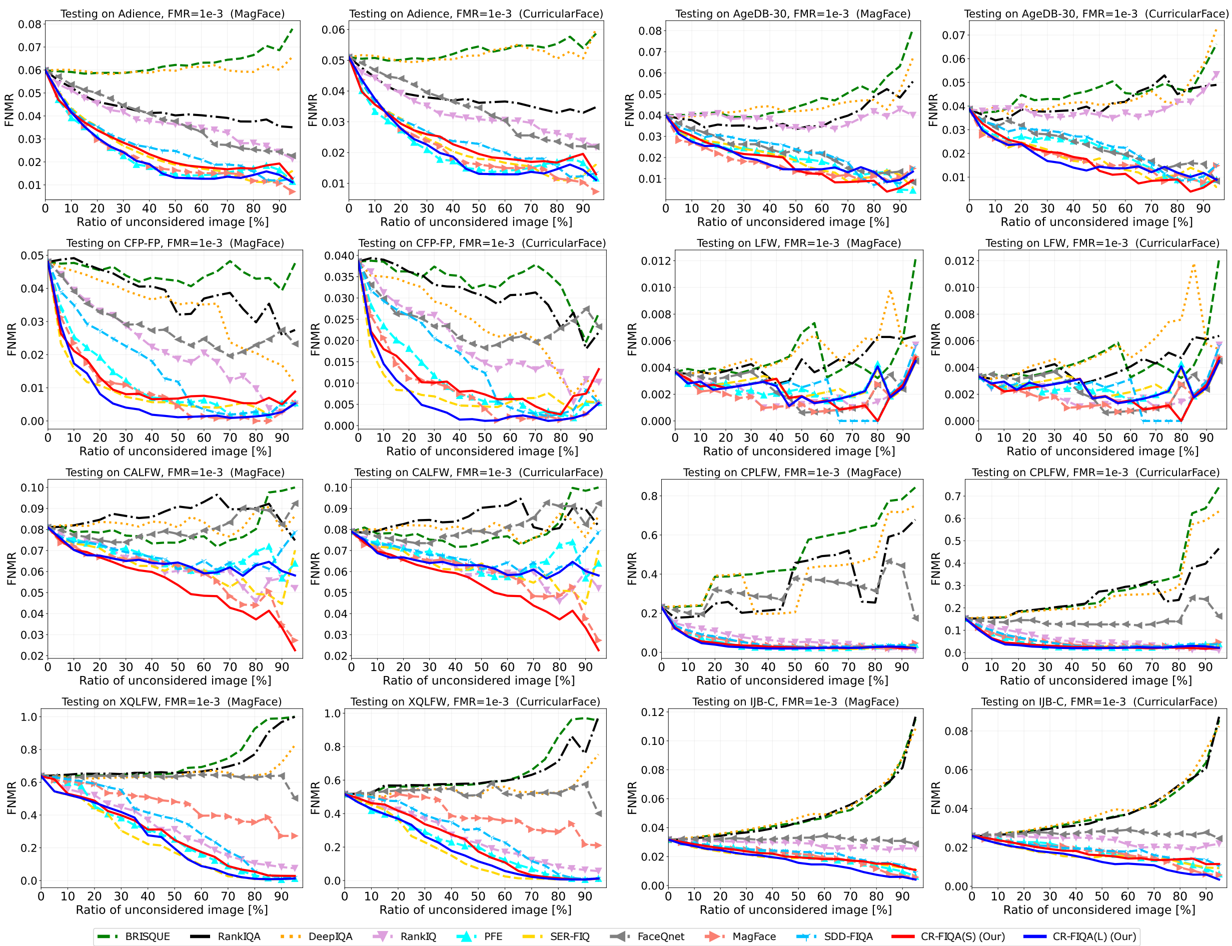


Figure 18. ERC (FNMR at FMR1e-3 vs reject) curves for all evaluated benchmarks using MagFace and CurricularFace FR models corresponding to Table 1 and complementary to the ERC curves in Figure 4 in main submission results. The proposed CR-FIQA(L) and CR-FIQA(S) are marked with solid blue and red lines, respectively. CR-FIQA leads to lower verification error, when rejecting a fraction of images, of the lowest quality, in comparison to SOTA methods (faster decaying curve) under most experimental settings.