Supplementary Material to the Paper Stating Comparison Score Uncertainty and Verification Decision Confidence Towards Transparent Face Recognition

Marco Huber[1,2]
marco.huber@igd.fraunhofer.de
Philipp Terhörst[1,3,4]
philipp.terhoerst@uni-paderborn.de
Florian Kirchbuchner[1]
florian.kirchbuchner@igd.fraunhofer,de
Naser Damer[1,2]
naser.damer@igd.fraunhofer.de
Arjan Kuijper[1,2]
arjan.kuijper@igd.fraunhofer.de

[1] Fraunhofer Institute for Computer Graphics Research IGD
Darmstadt, Germany
[2] Technical University of Darmstadt
Darmstadt, Germany
[3] Norwegian University of Science and Technology
Gjøvik, Norway
[4] Paderborn University
Paderborn, Germany

1 Supplementary Materials

This is the supplementary material to the paper *Stating Comparison Score Uncertainty and Verification Decision Confidence Towards Transparent Face Recognition*. More details about the pre-trained models and their licences are provided in Section 2. In Section 3 we further investigate the influence of the α parameter that influences the trade-off between uncertainty and intuitive confidence for the proposed decision confidence.

2 Information on the Used Models

This section provides additional information about the used models during the experiments to increase the reproducability of the experiments.

The ArcFace [1] model was trained by the authors and provided at their official GitHub repository: `https://github.com/deepinsight/insightface`. The model can be downloaded at: `https://github.com/deepinsight/insightface/tree/master/model_zoo`. We utilized the R100 model trained on the MS1MV2 dataset. All models are available for non-commercial research purposes only and the code is released under MIT License.

The MagFace [3] model was trained by the authors and provided at their official GitHub repository: `https://github.com/IrvingMeng/MagFace`. We used the model with

© 2022. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

the iResNet100 architecture and trained on the MS1MV2 dataset. The code is released under the Apache-2.0 license.

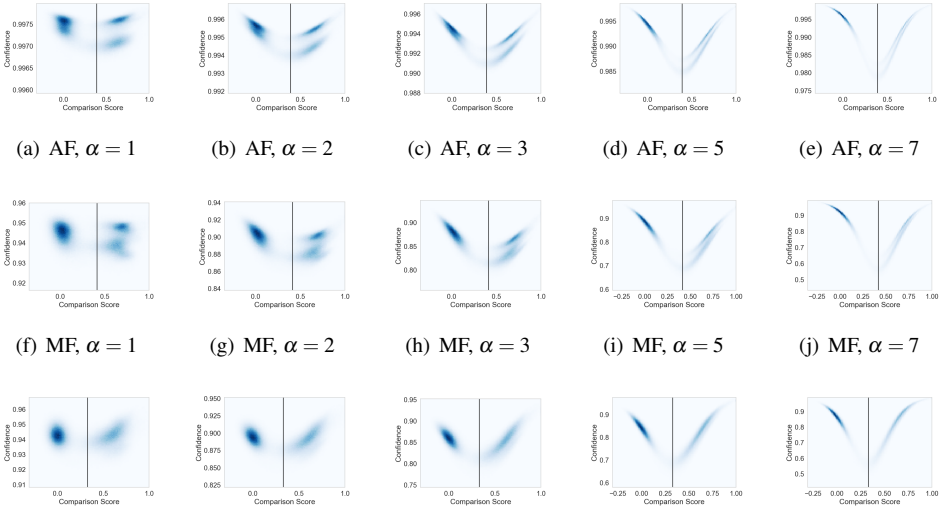
The CurricularFace [7] model was trained by the authors and provided at their official GitHub repository: https://github.com/HuangYG123/CurricularFace. The code is released under the MIT License.

3 Influence of α on the Decision Confidence

In this section, we briefly revisit the α-parameter used in the decision confidence function (Equation 5 and 6 in the paper) for the trade-off between intuitive decision confidence and uncertainty. Figure 1 shows the heatmaps for different α values on the Adience dataset. With an increase in α, the range of decision confidence values increases, and the distribution gets more shaped in a way that assigns lower confidence values near the decision threshold (black asymptote). The two different confidence clusters that can be determined on the genuine side of the decision threshold are probably due to biased groups, such as children, that are present in the Adience dataset.

Similar heatmaps are shown in Figure 2 for the experiments performed on the XQLFW dataset. Due to the cross-quality pairs, the distributions are less separated than in the Adience dataset, which can be also observed in the heatmaps. Furthermore, we also investigated low-quality pairs that lead to high variation in the confidence for small comparison scores.

The influence of the α-parameter on the error-vs-reject curves is shown in Figure 3 on the Adience dataset and in Figure 4 on the XQLFW dataset. Since at the selected threshold of FNMR@FMR=0.001 the genuine and imposter comparisons are unevenly distributed around the threshold, the recognition error increased when the ratio of unconsidered comparisons become higher. However, in most cases, the proposed decision confidence outperforms the threshold-based intuitive confidence baseline.

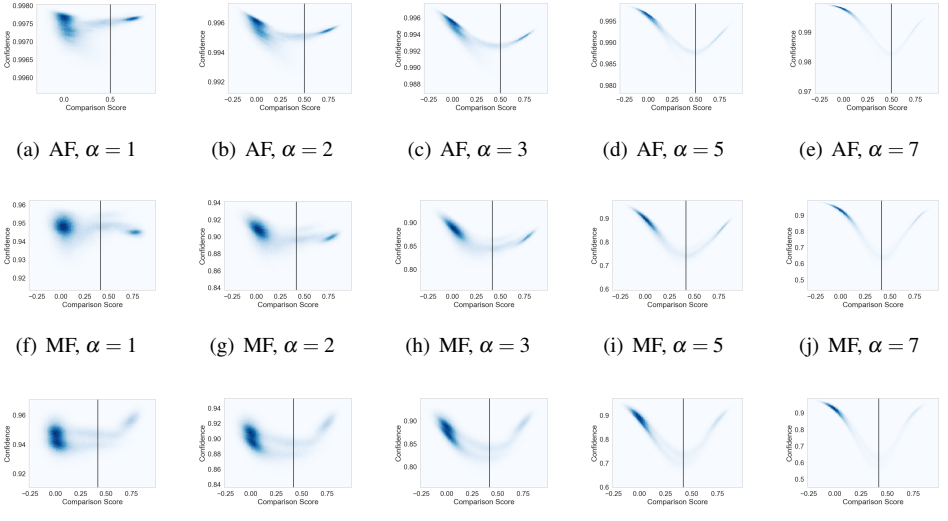


(a) AF, $\alpha = 1$ (b) AF, $\alpha = 2$ (c) AF, $\alpha = 3$ (d) AF, $\alpha = 5$ (e) AF, $\alpha = 7$

(f) MF, $\alpha = 1$ (g) MF, $\alpha = 2$ (h) MF, $\alpha = 3$ (i) MF, $\alpha = 5$ (j) MF, $\alpha = 7$

(k) CF, $\alpha = 1$ (l) CF, $\alpha = 2$ (m) CF, $\alpha = 3$ (n) CF, $\alpha = 5$ (o) CF, $\alpha = 7$

Figure 1: **Influence of the α-parameter on the Decision Confidence on the Adience dataset:** Heatmaps based on the decision confidence and comparison score. With a higher α-value, the variety in confidence increases, and the distribution of the values gets more shaped leading to a lower confidence around the decision threshold (black asymptote).

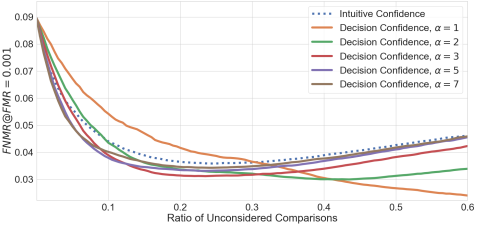


(a) AF, $\alpha = 1$ (b) AF, $\alpha = 2$ (c) AF, $\alpha = 3$ (d) AF, $\alpha = 5$ (e) AF, $\alpha = 7$

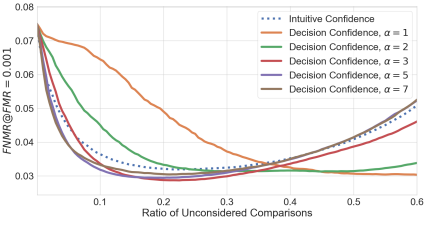
(f) MF, $\alpha = 1$ (g) MF, $\alpha = 2$ (h) MF, $\alpha = 3$ (i) MF, $\alpha = 5$ (j) MF, $\alpha = 7$

(k) CF, $\alpha = 1$ (l) CF, $\alpha = 2$ (m) CF, $\alpha = 3$ (n) CF, $\alpha = 5$ (o) CF, $\alpha = 7$

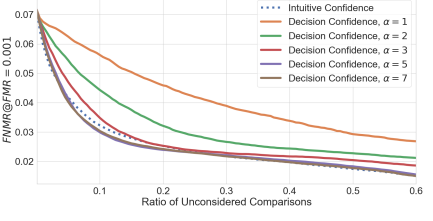
Figure 2: **Influence of the α-parameter on the Decision Confidence on the XQLFW dataset:** Heatmaps based on the confidence and comparison score. Similarly to the experiments on the Adience dataset, it can be observed that with a higher α-value, the variety in confidence increases. With an increase in α, the lowest decision confidence is more and more shifted to the decision threshold.



(a) ArcFace

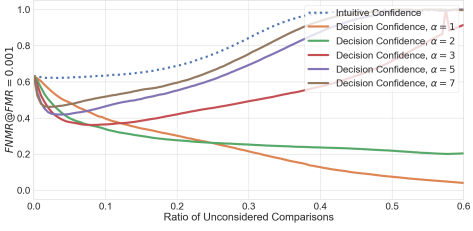


(b) MagFace

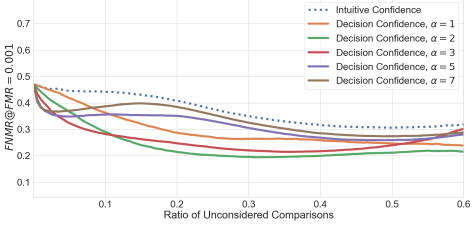


(c) CurricularFace

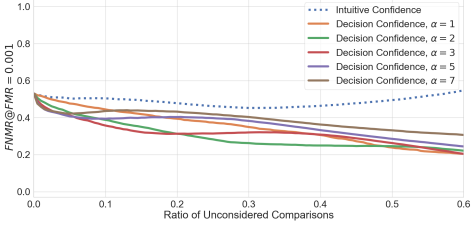
Figure 3: **Influence of the α-parameter on the Error-vs-Reject curve on the Adience dataset:** Comparisons are removed depending on their calculated confidence and the performance is evaluated on the remaining comparisons. The α-parameter determines the trade-off between intuitive confidence and uncertainty. $\alpha = 3$ and $\alpha = 5$ outperform the intuitive confidence baseline on all models.



(a) ArcFace



(b) MagFace



(c) CurricularFace

Figure 4: **Influence of the α-parameter on the Error-vs-Reject curve on the XQLFW dataset:** Comparisons are removed depending on their calculated confidence and the performance is evaluated on the remaining comparisons. The proposed decision confidence outperforms the intuitive confidence baseline on the difficult XQLFW dataset. The decrease in performance (the increase of the error) is due to the imbalance of the data at the operational point of FNMR@FMR=0.001.

References

[1] Jiankang Deng, Jia Guo, Niannan Xue, and Stefanos Zafeiriou. Arcface: Additive angular margin loss for deep face recognition. In *CVPR*, pages 4690–4699. Computer Vision Foundation / IEEE, 2019.

[2] Yuge Huang, Yuhan Wang, Ying Tai, Xiaoming Liu, Pengcheng Shen, Shaoxin Li, Jilin Li, and Feiyue Huang. Curricularface: Adaptive curriculum learning loss for deep face recognition. In *CVPR*, pages 5900–5909. Computer Vision Foundation / IEEE, 2020.

[3] Qiang Meng, Shichao Zhao, Zhida Huang, and Feng Zhou. Magface: A universal representation for face recognition and quality assessment. In *CVPR*, pages 14225–14234. Computer Vision Foundation / IEEE, 2021.