Supplementary

In the main paper, we limited information on the databases to essential parts. At this point, we provide additional information in the context of the used databases to enhance the understanding of this work. More precisely, this supplementary material consists of three parts. First, we provide information on the quality distributions of the databases to better understand the choice of the training dataset. Second, we provide additional information on the database licenses and their creation processes. Third, demonstrate the effect of using different training databases for QMagFace to support the reasoning concerning the generalizability of QMagFace from Section 5.4 in a more direct manner.

Quality Distributions of MagFace

The proposed QMagFace approach makes use of MagFace qualities and includes these in the decision process. To get a better understanding of the quality distributions of the different used databases, Figure 3 shows these distributions for the three MagFace backbones. For all backbones, LFW and Morph consist of the highest FIQ values and share a similar distribution due to the fact that both databases consist of mostly frontal and well-illuminated images with high image quality. For MagFace-50 and MagFace-100, the quality distribution of ColorFeret shows the widest range of FIQ values. ColorFeret consists of high-quality images that were taken under controlled capturing conditions. The high variety of FIQ values origin from the head pose variations and the lowest FIQ values come from the profile face images since these prove to have a very low utility for recognition [41]. The Adience database consists of face images with a wide variety of quality-decreasing factors such as variations in image quality, occlusions, expressions, and head poses. However, it does not contain many full profile images and thus, ColorFeret consists of images with lower FIQ values. It should be noted that the quality estimation performance of MagFace is dependent on its FR performance. Consequently, for MagFace-18, this leads to many wrongly assessed qualities and thus, to a lower performance of QMagFace-18. This also explains why for MagFace-18 the quality distributions are similar while for MagFace-50 and MagFace-100 the distributions show strong differences.

Additional Information on the Utilized Databases

After discussing the additional properties of the databases themselves, this section provides additional information about the licenses and the creation process of these databases. Since the amount of information about the used datasets is restricted by the page limit, the paper focuses on the most important aspects to make the experiments understandable and reproducible.

LFW [17] is licensed under CC-BY-4.0. It is based on the Faces in the Wild database [4] collected by Tamara Berg at Berkeley and consists of face captioned from news images. More details can be found in [17] or under `http://vis-www.cs.umass.edu/lfw/index.html`.

AgeDB [32] is available for non-commercial research purposes only and consists of images manually collected from the internet. More details on the collection process can be found in [32] and the details on the license are presented in `https://ibug.doc.ic.ac.uk/resources/agedb/`.

CFP-FP [39] consist of manually collected images of celebrities in frontal and profile views. More information can be found in [39] and `http://www.cfpw.io/`. To get more information on license and consent, we reached out to the first author via mail.

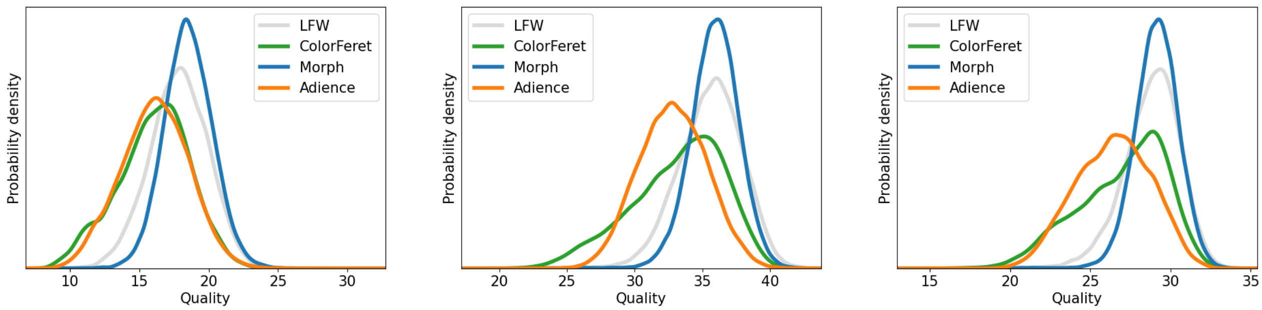
XQLFW [25] is licensed under the MIT License and is based on the modified images of the LFW dataset [17] (CC-BY-4.0). Detailed information can be found in [25], `https://martlgap.github.io/xqlfw/pages/citation.html`, and `https://github.com/Martlgap/xqlfw`.

The images of the IJB-B [46] and the IJB-C [30] databases from the National Institute for Standards and Technology (NIST) are made available under different Creative Commons license variants. Details on the collection process and corresponding informationcan be found in [46] and IJB-C [30]. More information on the license are shown under `https://nigos.nist.gov/datasets/ijbc/request` and `https://nigos.nist.gov/facechallenges/data/IJBC/IJBC_LICENSES.TXT`.

Adience [12] is a database that includes a compilation of individual images which were uploaded to the internet and tagged as publicly available by the original author. It is limited to research purposes only. More information can be found in [12] and under `https://talhassner.github.io/home/projects/Adience/LICENSE.txt`.

In this work, the academic version of the Morph dataset [22] is used. This is restricted to for research purposes only. The legacy photographs associated with these records were taken 1962 and 1998. Digital scans of these photographs were collected with legal considerations and IRB approval. More information can be found in [22] and under `https://uncw.edu/oic/tech/morph.html`.

ColorFeret [35] database is restricted to face recognition research. During the data collection, the different subjects were photographed in 15 sessions over three years under controlled conditions. Detailed license information can be found under `https://www.nist.gov/system/files/documents/2019/11/25/colorferet_release_agreement.pdf`. More



(a) MagFace-18 (b) MagFace-50 (c) MagFace-100

Figure 3: Quality distributions for the four FR datasets for the three MagFace backbones based on iResNet-18/50/100.

details can be found in [35] and `https://www.nist.gov/itl/products-and-services/color-feret-database`.

Additional Experiments on Various Training Data

After finalizing the discussion on the databases, this section aims to emphasize the high generalizability of the proposed approach against various training data. In Section 5.4, this was already demonstrated indirectly by comparing the optimal quality-weighting functions for each database. In this section, we show this in a more direct way by iteratively using the different FR databases for training.

In Table 5, the effect of the different training databases on the single-image FR benchmarks are shown. For QMagFace-18, the performance does not improve in all cases due to the limited quality estimation performance of MagFace-18 as discussed above. In contrast, adding the quality-awareness to the MagFace-100 model improves the recognition performance independent of the utilized training data. Moreover, it seems that the choice in the paper to use Adience for training was wrong since the performance when using the other databases for training is higher. However, the choice for Adience was done to its large variety in quality-decreasing factors, such as occlusions, head poses, illuminations, and image qualities. When it comes to smaller FMRs, these factors become more important and Adience might be the better choice for stable improvements in the recognition performance.

In Tables 6, 7, 8, and 9, the effect of different training databases is analysed over a wide range of FMRs. For low FMRs, such as 10^{-5}, a larger variety of quality factors play important roles in enhancing the recognition performance and thus, using Adience as the training database leads to very stable performance improvements in all cases. However, including the quality-awareness leads to strong performance improvements for most FMRs and the different training datasets generally leads to similar performances demonstrating the high generalizability of the proposed QMagFace approach.

Table 5: The effect of the different training databases on the single-image FR benchmarks. The performance is reported in terms of benchmark accuracy (%). For comparison, the performance of the QMagFace variants is shown against the MagFace models without quality-awareness. It turns out that choosing Adience, as done in the paper, leads to the weakest performance on these benchmarks. Consequently, the proposed approach, QMagFace-100, achieves state-of-the-art face recognition performance independent of the training data, especially in cross-age (AgeDB), cross-pose (CFP-FP), and cross-quality (XQLFW) scenarios.

Trained on	Model	AgeDB	CFP-FP	LFW	XQLFW
	MagFace-18	93.37	93.11	99.22	69.55
	MagFace-50	97.60	97.33	99.72	80.60
	MagFace-100	98.18	98.36	99.73	83.90
Adience	QMagFace-18	92.98	94.00	99.30	68.60
	QMagFace-50	97.88	97.74	99.73	80.63
	QMagFace-100	98.50	98.74	99.80	83.97
ColorFeret	QMagFace-18	92.90	94.03	99.33	68.68
	QMagFace-50	97.88	97.80	99.73	80.63
	QMagFace-100	98.48	98.76	99.80	84.03
Morph	QMagFace-18	93.02	94.06	99.33	68.67
	QMagFace-50	97.95	97.86	99.73	80.57
	QMagFace-100	98.55	98.77	99.82	83.82
LFW	QMagFace-18	92.92	94.07	99.35	68.72
	QMagFace-50	97.88	97.86	99.72	80.58
	QMagFace-100	98.60	98.77	99.83	83.97

Table 6: Evaluation on Adience based on different training datasets - The performance [%] is reported in terms of FNMR at different FMRs and EER. Three MagFace variants are compared against QMagFace variants that are trained on different training sources.

			FNMR at FMR				
Training database	Model	EER	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}
	MagFace-18	4.505	2.665	10.639	28.935	49.982	74.662
	MagFace-50	2.432	1.334	3.463	8.818	18.396	44.821
	MagFace-100	2.291	1.395	2.926	5.478	11.211	30.331
ColorFeret	QMagFace-18	3.798	2.110	8.466	26.547	49.403	74.881
	QMagFace-50	2.371	1.276	3.310	8.604	18.386	44.232
	QMagFace-100	2.255	1.368	2.818	5.336	11.098	30.188
LFW	QMagFace-18	3.794	2.099	8.360	25.556	48.337	73.807
	QMagFace-50	2.369	1.286	3.280	8.479	19.001	45.802
	QMagFace-100	2.266	1.381	2.810	5.277	11.765	30.967
Morph	QMagFace-18	3.792	2.120	8.396	26.397	49.208	74.669
	QMagFace-50	2.368	1.289	3.282	8.412	19.405	45.381
	QMagFace-100	2.264	1.369	2.817	5.288	11.443	30.704

Table 7: Evaluation on ColorFeret based on different training datasets - The performance [%] is reported in terms of FNMR at different FMRs and EER. Three MagFace variants are compared against QMagFace variants that are trained on different training sources

			FNMR at FMR				
Training database	Model	EER	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}
	MagFace-18	5.312	4.067	9.193	22.094	83.968	97.517
	MagFace-50	3.635	2.553	5.056	7.560	12.393	22.416
	MagFace-100	2.629	1.789	3.297	4.791	7.523	16.909
Adience	QMagFace-18	4.232	3.068	6.902	22.951	82.531	97.723
	QMagFace-50	2.941	1.464	4.173	6.832	12.247	23.426
	QMagFace-100	2.060	0.950	2.616	4.409	7.145	16.454
LFW	QMagFace-18	4.282	3.082	6.917	24.382	82.466	96.127
	QMagFace-50	2.961	1.572	4.194	6.815	12.045	22.677
	QMagFace-100	2.031	1.033	2.565	4.275	7.144	18.318
Morph	QMagFace-18	4.245	3.085	6.901	23.900	80.080	94.926
	QMagFace-50	2.939	1.566	4.141	6.733	12.130	24.999
	QMagFace-100	2.051	1.055	2.596	4.302	7.109	17.213

Table 8: Evaluation on LFW based on different training datasets - The performance [%] is reported in terms of FNMR at different FMRs and EER. Three MagFace variants are compared against QMagFace variants that are trained on different training sources

			FNMR at FMR				
Training database	Model	EER	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}
	MagFace-18	1.057	0.324	1.096	3.710	9.613	20.163
	MagFace-50	0.349	0.110	0.290	0.462	0.586	0.821
	MagFace-100	0.277	0.159	0.255	0.297	0.441	0.621
Adience	QMagFace-18	0.724	0.186	0.607	2.324	7.282	14.377
	QMagFace-50	0.332	0.035	0.172	0.407	0.517	0.752
	QMagFace-100	0.195	0.145	0.172	0.221	0.331	0.517
ColorFeret	QMagFace-18	0.793	0.204	0.721	2.516	7.438	18.768
	QMagFace-50	0.304	0.086	0.199	0.379	0.516	0.797
	QMagFace-100	0.195	0.091	0.138	0.212	0.335	0.556
Morph	QMagFace-18	0.778	0.204	0.716	2.505	7.307	18.655
	QMagFace-50	0.295	0.086	0.192	0.357	0.455	0.662
	QMagFace-100	0.184	0.090	0.141	0.208	0.314	0.477

Table 9: Evaluation on Morph based on different training datasets - The performance [%] is reported in terms of FNMR at different FMRs and EER. Three MagFace variants are compared against QMagFace variants that are trained on different training sources

			FNMR at FMR				
Training database	Model	EER	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}
	MagFace-18	0.883	0.813	0.873	1.185	2.189	50.892
	MagFace-50	0.788	0.784	0.825	0.832	0.843	0.894
	MagFace-100	0.848	0.777	0.814	0.824	0.834	0.848
Adience	QMagFace-18	0.843	0.779	0.834	1.036	1.908	40.070
	QMagFace-50	0.821	0.473	0.812	0.826	0.835	0.880
	QMagFace-100	0.773	0.363	0.760	0.817	0.829	0.840
ColorFeret	QMagFace-18	0.846	0.790	0.841	1.059	1.969	44.750
	QMagFace-50	0.790	0.475	0.784	0.798	0.814	0.860
	QMagFace-100	0.763	0.371	0.747	0.802	0.813	0.824
LFW	QMagFace-18	0.838	0.797	0.841	1.032	1.871	57.971
	QMagFace-50	0.761	0.477	0.775	0.795	0.808	0.845
	QMagFace-100	0.737	0.417	0.729	0.794	0.804	0.818