

Supplementary Material - Facial Hair Area in Face Recognition Across Demographics: Small Size, Big Effect

This supplementary material provides: a) The genuine and impostor distributions of BA-test and MORPH3, shown in Figure 1 and Figure 2. b) Accuracy disparity measurement, $\Delta d'$, between CS-MI and the other two beard areas for all demographic groups in the BA-test dataset, shown in Table 1; c) Genuine and impostor distributions of five pixel color augmentations for both MORPH3 and BA-test dataset, where the features are extracted by ArcFace-R100 trained with two datasets and MagFace-R100 trained with MS1MV2, shown in Figure 3, Figure 4, Figure 5, Figure 6, Figure 7, Figure 8

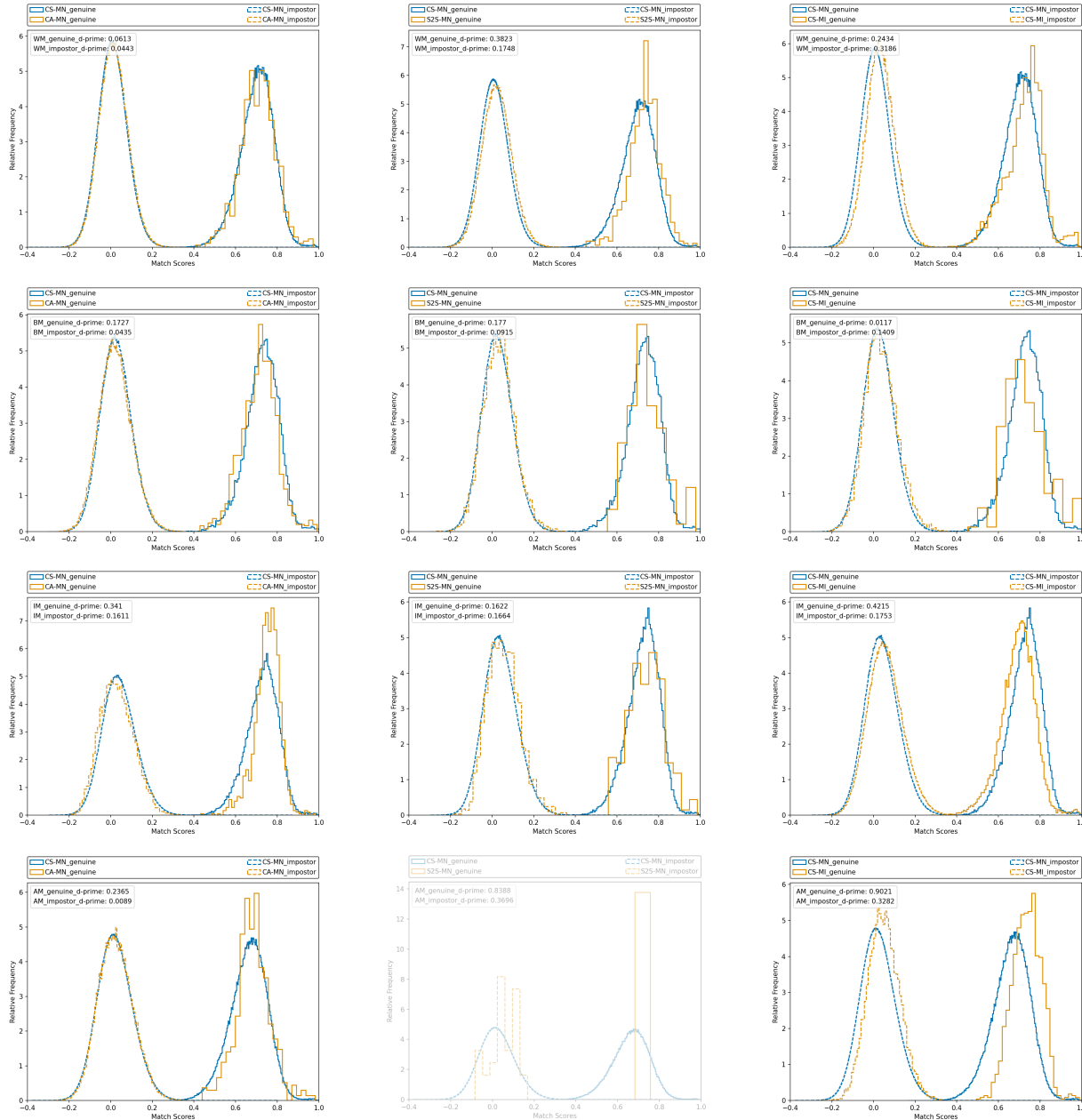


Figure 1. The genuine and impostor distributions comparisons between CS-MN and {CA-MN, S2S-MN, and CS-MI} in BA-test dataset. The face matcher is ArcFace-R100 trained with Glint360K. Note that results derived from insufficient data are rendered in a semi-transparent format.

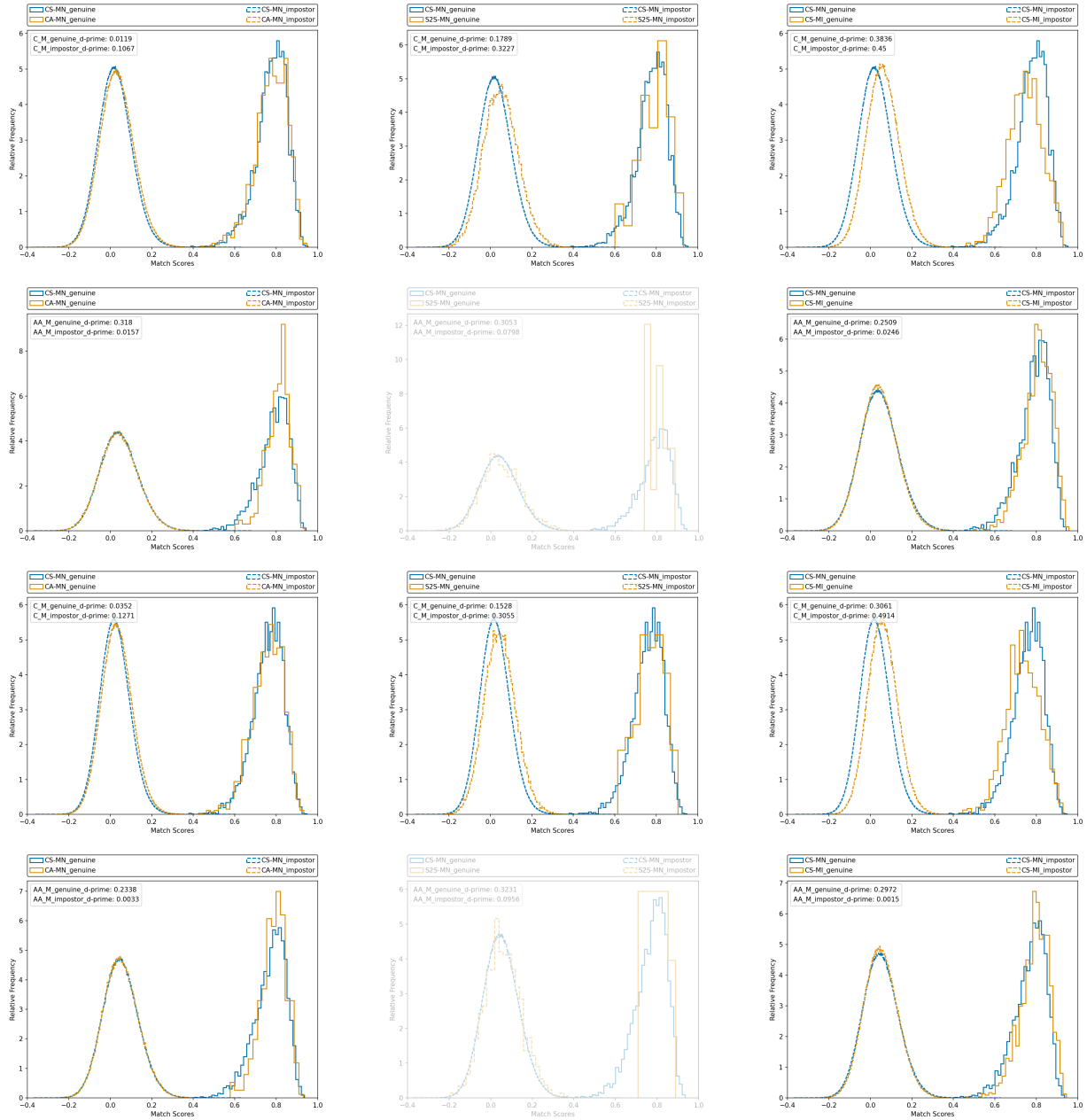
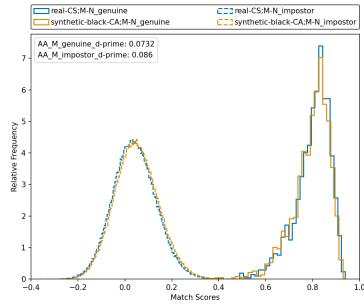


Figure 2. The genuine and impostor distributions comparisons between CS-MN and {CA-MN, S2S-MN, and CS-MI} in MORPH dataset. The face matcher is ArcFace-R100 trained with MS1MV2 (Top two rows) Glint360K (Bottom two rows). Note that results derived from insufficient data are rendered in a semi-transparent format.

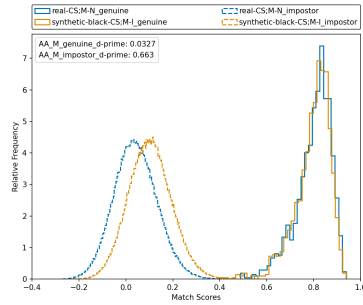
$\Delta d'$	Dataset	Strategies	AM		WM		BM		IM	
			CA-MN	S2S-MN	CA-MN	S2S-MN	CA-MN	S2S-MN	CA-MN	S2S-MN
Gen	MSV2	Original	0.3668	-	0.2259	0.0121	0.3243	0.0987	0.7327	0.6294
		ZM	0.0603	-	0.1314	0.1356	0.086	0.269	0.8038	0.7408
		HMR-Random	0.3071	-	0.25	0.0227	0.1978	0.2603	0.6883	0.6185
		HMR-ID	0.276	-	0.2136	0.0061	0.2374	0.2075	0.8299	0.6347
	Glint	Original	0.6756	-	0.1833	0.1093	0.142	0.1693	0.7929	0.5599
		ZM	0.2695	-	0.0969	0.2422	0.1418	0.242	0.9512	0.7539
		HMR-Random	0.5717	-	0.2481	0.1725	0.3342	0.1691	0.7066	0.6402
		HMR-ID	0.5926	-	0.0577	0.0543	0.2545	0.2498	0.3593	0.5178
Imp	MSV2	Original	0.3579	-	0.4015	0.2754	0.2819	0.1327	0.4446	0.1363
		ZM	0.3395	-	0.5166	0.2961	0.3596	0.2286	0.8316	0.4928
		HMR-Random	0.2846	-	0.3613	0.216	0.2226	0.0799	0.3148	0.0424
		HMR-ID	0.2658	-	0.3146	0.183	0.1802	0.0939	0.3434	0.1083
	Glint	Original	0.322	-	0.2734	0.1388	0.1804	0.05	0.3342	0.0146
		ZM	0.3563	-	0.4547	0.2377	0.2867	0.1369	0.7322	0.2839
		HMR-Random	0.2782	-	0.3155	0.3544	0.209	0.1173	0.3678	0.3532
		HMR-ID	0.235	-	0.1907	0.2089	0.2208	0.0633	0.112	0.0717

Table 1. $\Delta d'$ measurement between *CS-MI* and *listed attributes* on both genuine and impostor distributions of the BA-test dataset. Eight models trained with four training set manipulation strategies and two training sets are used to calculate the values. The worst and the best performances are shown in Red and Green.

(CS-MN, CS-MN) vs. (CA-MN, CA-MN)



(CS-MN, CS-MN) vs. (CS-MI, CS-MI)



(CS-MN, CS-MN) vs. (S2S-MN, S2S-MN)

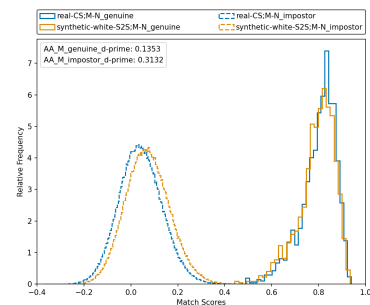
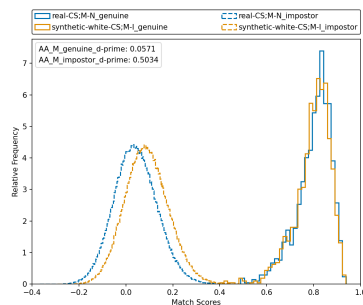
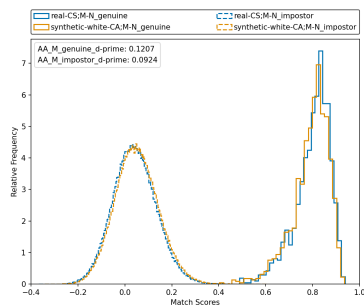
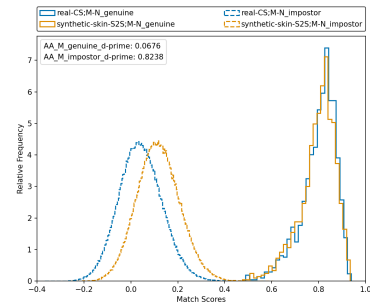
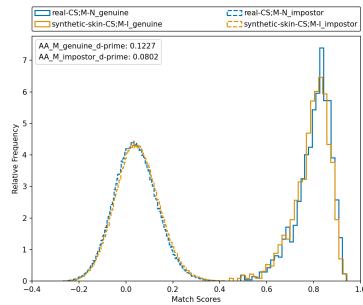
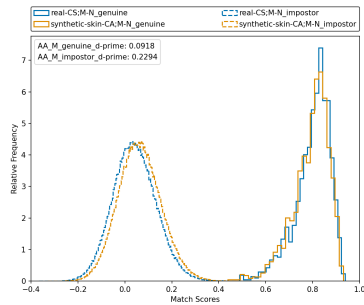
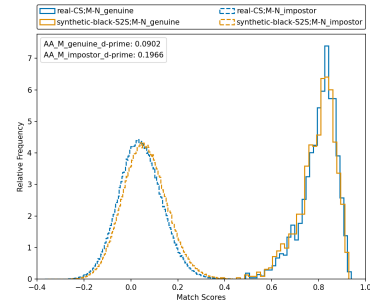
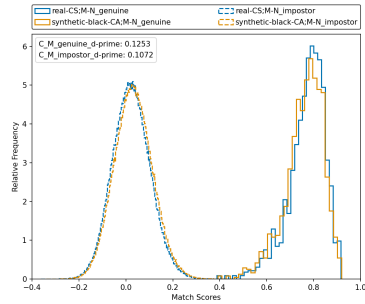
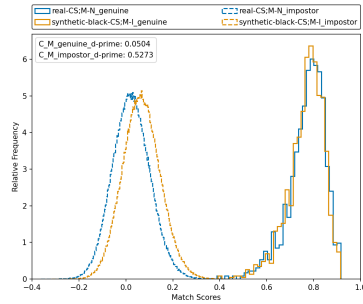


Figure 3. Comparison of genuine and impostor distributions upon augmenting images with black pixels (first row), average skin pixels (second row), and white pixels (third row), in the chin area (CA-MN), from side to side (S2S-MN), and the mustache area (CS-MI). The face matcher is ArcFace-R100, trained on the MS1MV2 dataset. The samples are picked from African-American males in the MORPH dataset.

(CS-MN, CS-MN) vs. (CA-MN, CA-MN)



(CS-MN, CS-MN) vs. (CS-MI, CS-MI)



(CS-MN, CS-MN) vs. (S2S-MN, S2S-MN)

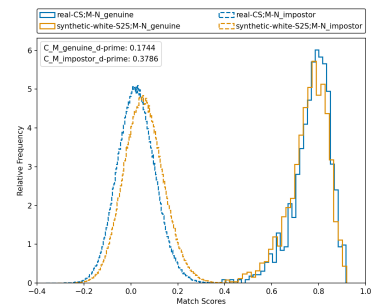
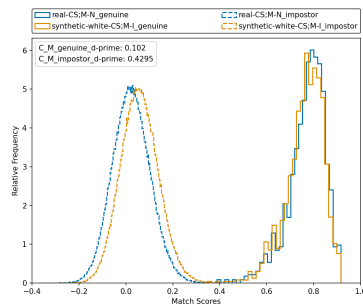
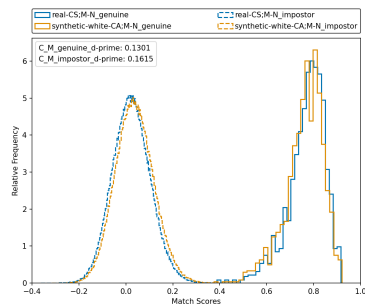
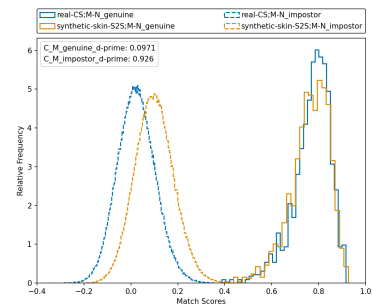
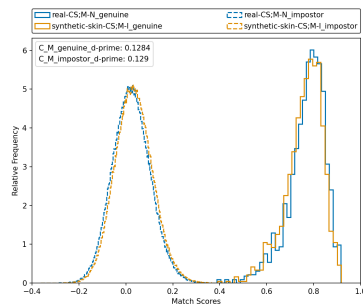
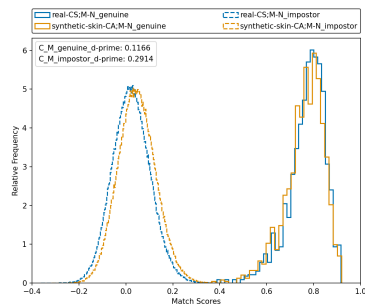
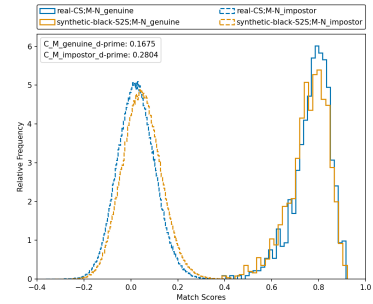
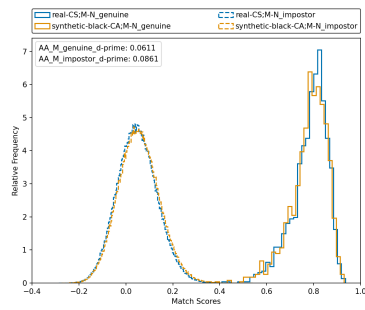
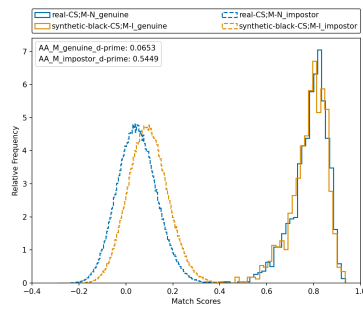


Figure 4. Comparison of genuine and impostor distributions upon augmenting images with black pixels (first row), average skin pixels (second row), and white pixels (third row), in the chin area (CA-MN), from side to side (S2S-MN), and the mustache area (CS-MI). The face matcher is ArcFace-R100, trained on the MS1MV2 dataset. The samples are picked from **Caucasian males** in the MORPH dataset.

(CS-MN, CS-MN) vs. (CA-MN, CA-MN)



(CS-MN, CS-MN) vs. (CS-MI, CS-MI)



(CS-MN, CS-MN) vs. (S2S-MN, S2S-MN)

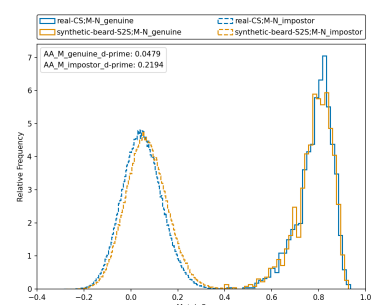
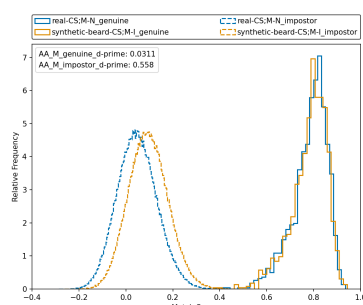
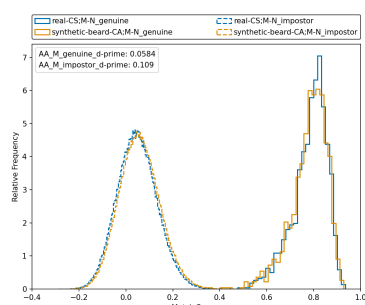
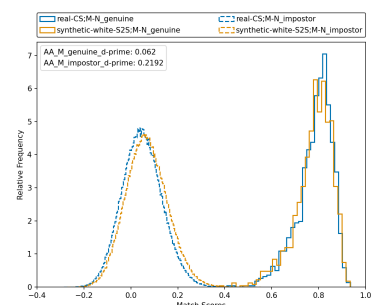
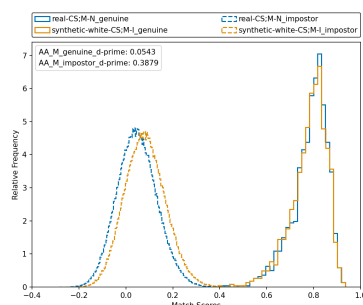
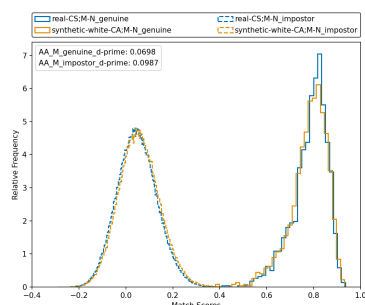
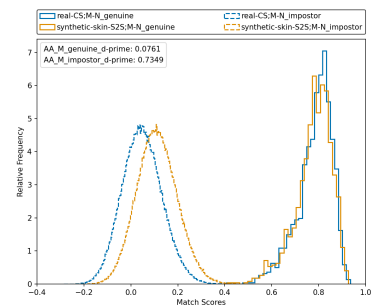
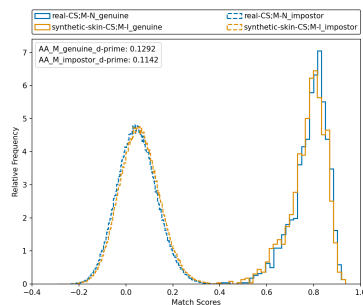
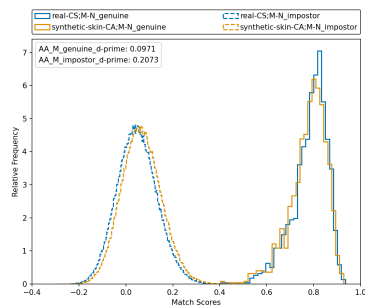
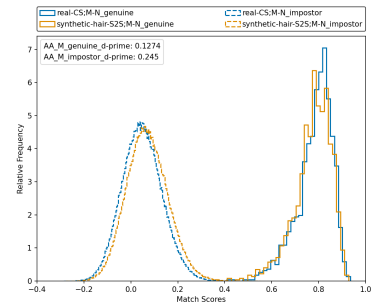
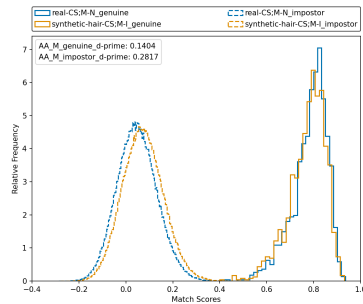
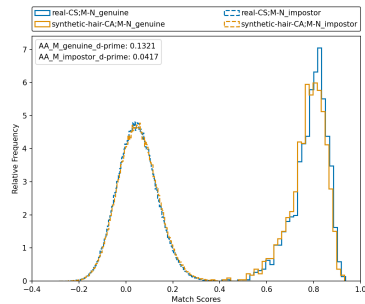
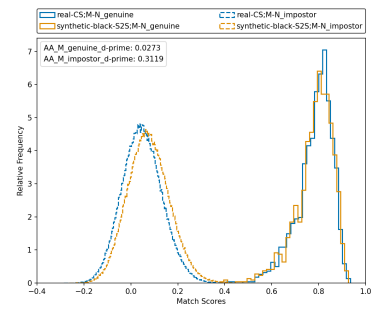
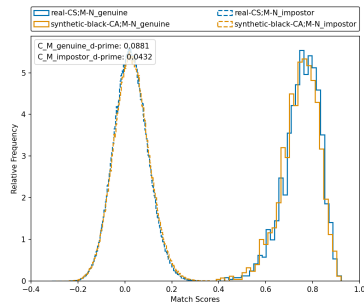
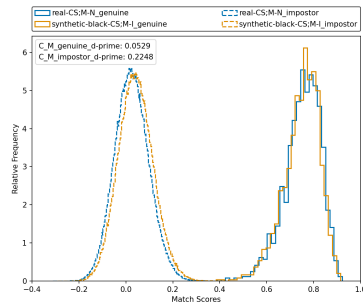


Figure 5. Comparison of genuine and impostor distributions upon augmenting images with black pixels (first row), average hair pixels (second row), average skin pixels (third row), white pixels (fourth row), and cropped beard pixels (fifth row), in the chin area (CA-MN), from side to side (S2S-MN), and the mustache area (CS-MI). The face matcher is ArcFace-R100, trained on the **Glint360K** dataset. The samples are picked from **African-American** males in the MORPH dataset.

(CS-MN, CS-MN) vs. (CA-MN, CA-MN)



(CS-MN, CS-MN) vs. (CS-MI, CS-MI)



(CS-MN, CS-MN) vs. (S2S-MN, S2S-MN)

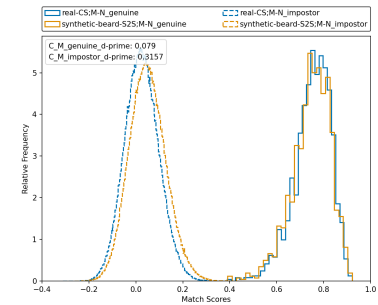
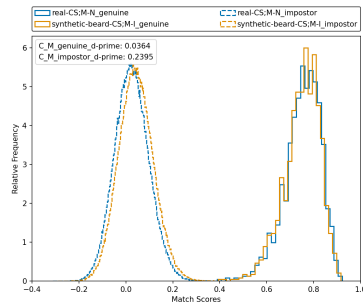
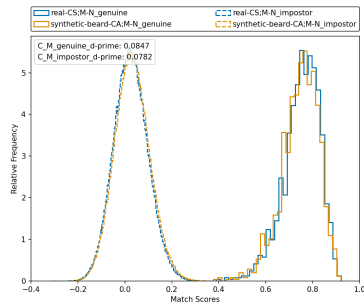
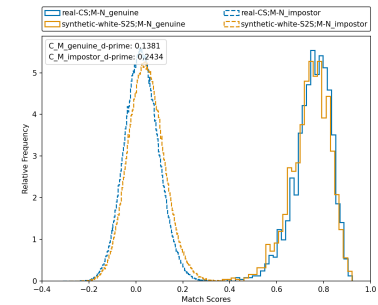
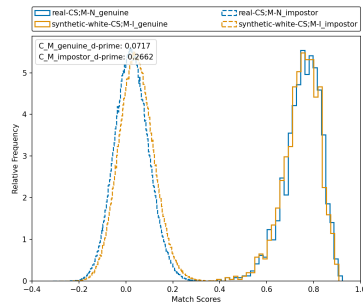
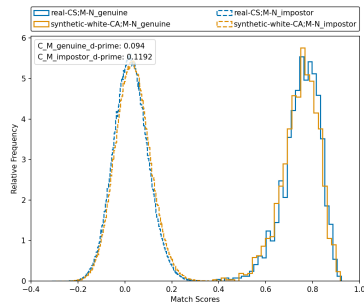
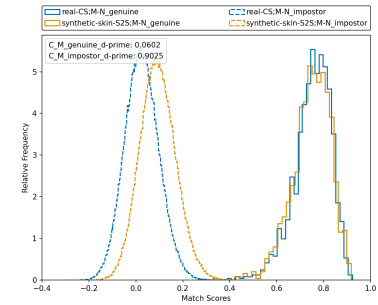
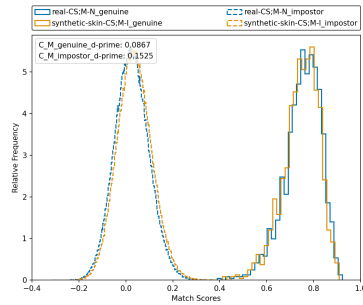
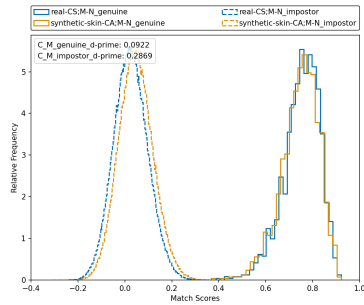
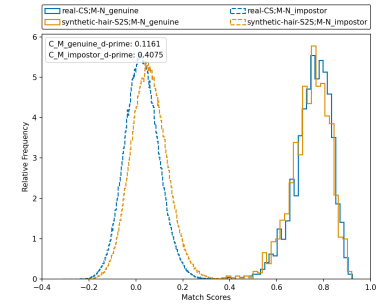
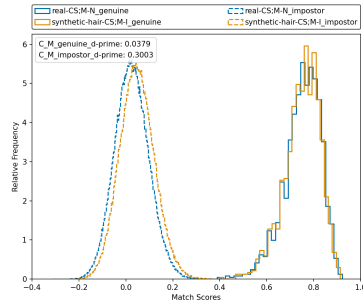
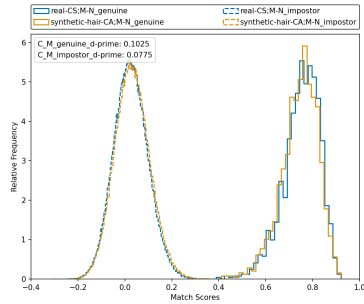
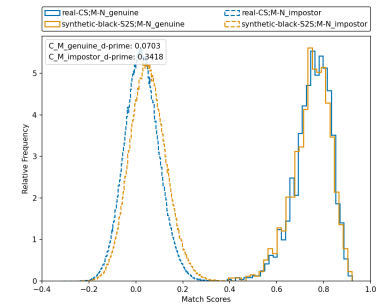
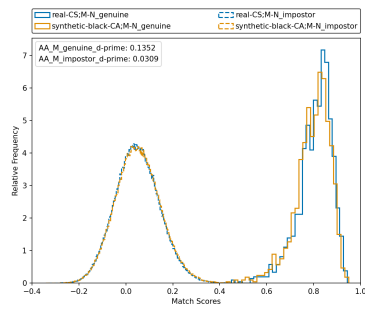
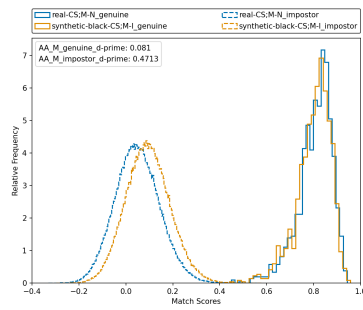


Figure 6. Comparison of genuine and impostor distributions upon augmenting images with black pixels (first row), average hair pixels (second row), average skin pixels (third row), white pixels (fourth row), and cropped beard pixels (fifth row), in the chin area (CA-MN), from side to side (S2S-MN), and the mustache area (CS-MI). The face matcher is ArcFace-R100, trained on the **Glint360K** dataset. The samples are picked from **Caucasian males** in the MORPH dataset.

(CS-MN, CS-MN) vs. (CA-MN, CA-MN)



(CS-MN, CS-MN) vs. (CS-MI, CS-MI)



(CS-MN, CS-MN) vs. (S2S-MN, S2S-MN)

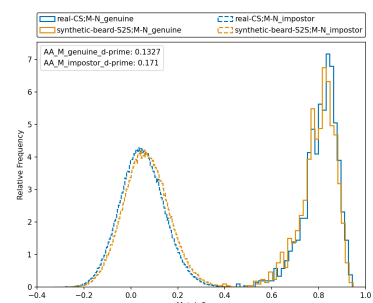
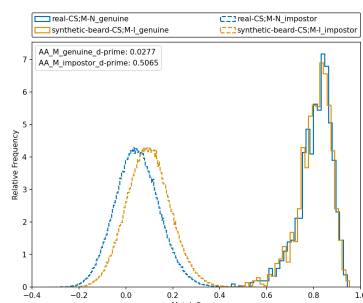
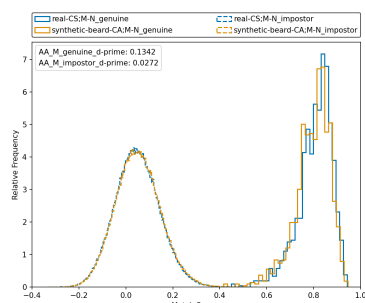
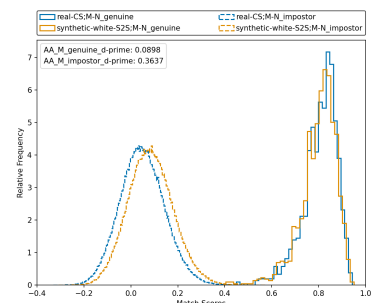
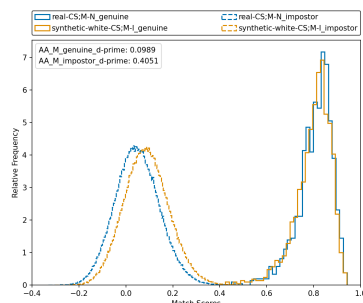
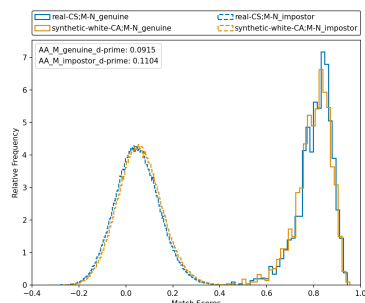
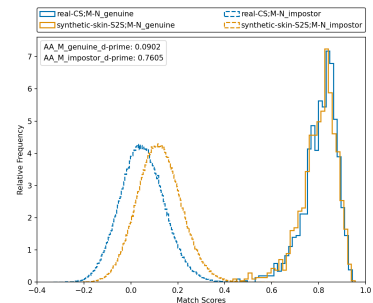
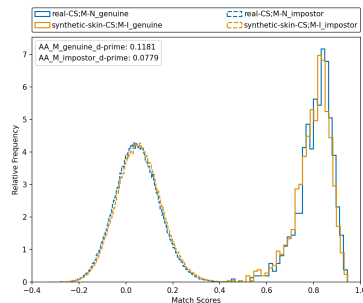
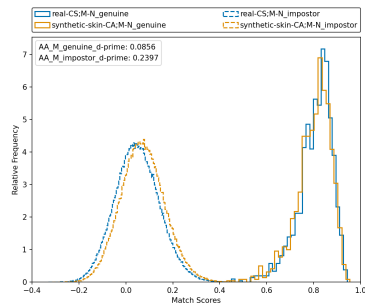
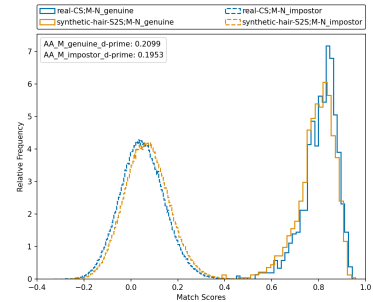
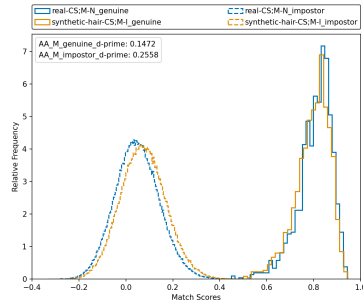
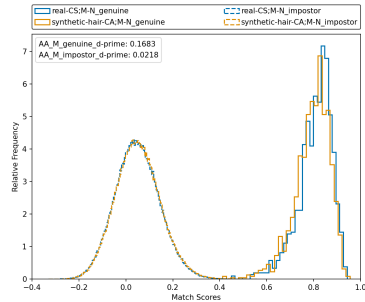
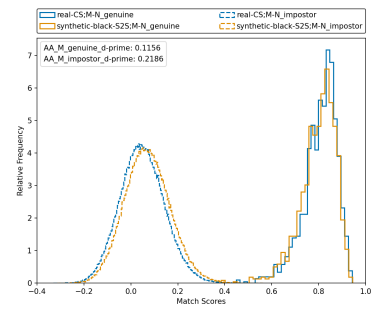
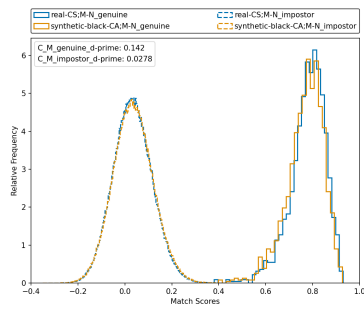
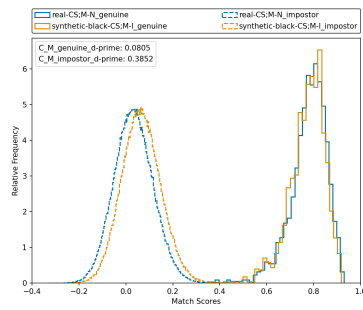


Figure 7. Comparison of genuine and impostor distributions upon augmenting images with black pixels (first row), average hair pixels (second row), average skin pixels (third row), white pixels (fourth row), and cropped beard pixels (fifth row), in the chin area (CA-MN), from side to side (S2S-MN), and the mustache area (CS-MI). The face matcher is MagFace-R100, trained on the MS1MV2 dataset. The samples are picked from African-American males in the MORPH dataset.

(CS-MN, CS-MN) vs. (CA-MN, CA-MN)



(CS-MN, CS-MN) vs. (CS-MI, CS-MI)



(CS-MN, CS-MN) vs. (S2S-MN, S2S-MN)

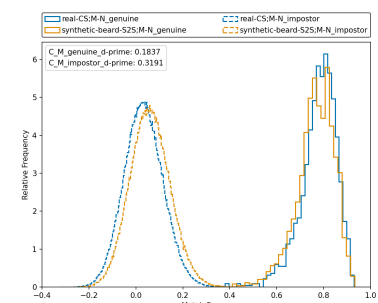
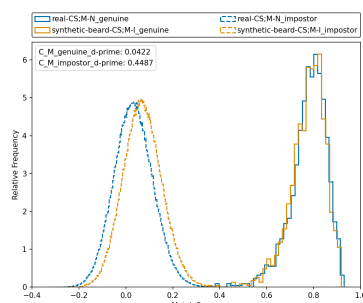
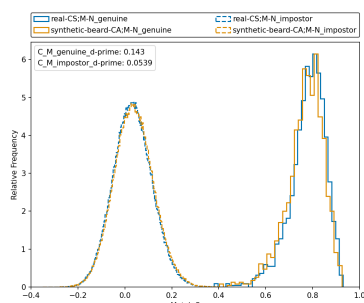
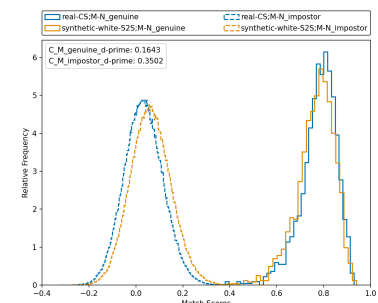
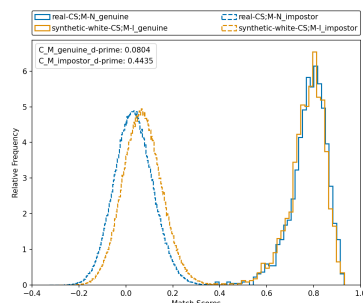
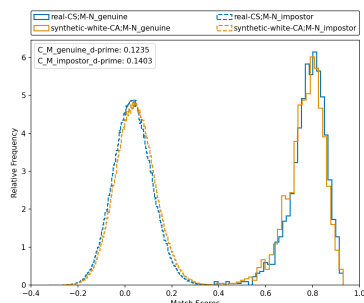
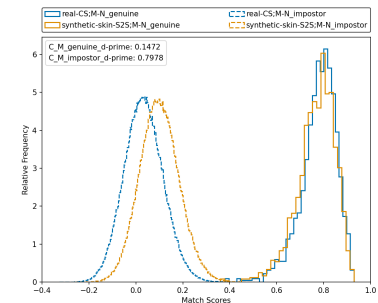
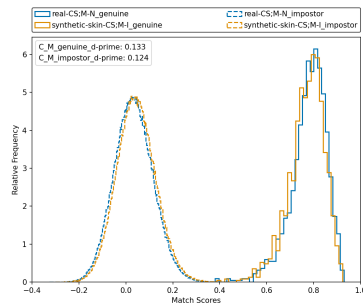
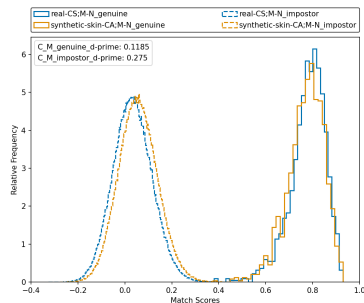
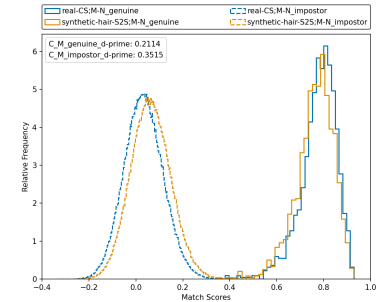
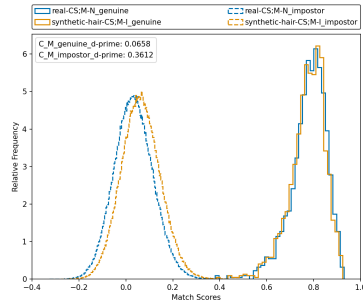
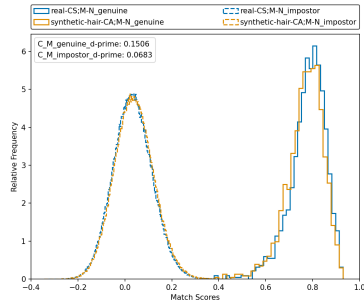
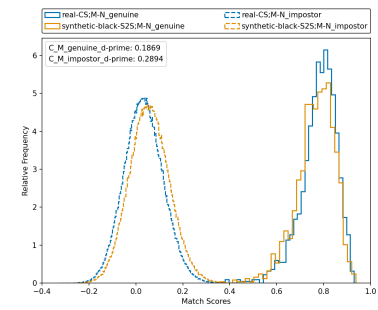


Figure 8. Comparison of genuine and impostor distributions upon augmenting images with black pixels (first row), average hair pixels (second row), average skin pixels (third row), white pixels (fourth row), and cropped beard pixels (fifth row), in the chin area (CA-MN), from side to side (S2S-MN), and the mustache area (CS-MI). The face matcher is MagFace-R100, trained on the MS1MV2 dataset. The samples are picked from **Caucasian males** in the MORPH dataset.