IDPERTURB: Enhancing Variation in Synthetic Face Generation via Angular Perturbations

Fadi Boutros[1], Eduarda Caldeira[1, 2], Tahar Chettaoui[1], Naser Damer[1, 2]
[1]Fraunhofer Institute for Computer Graphics Research IGD, Germany
[2] TU Darmstadt, Germany
fadi.boutros@igd.fraunhofer.de

1. Supplementary Material

This supplementary material complements the main submission by providing:

1. Appendix A (Sec. 2, Table 1): Theoretical justification of IDPERTURB
2. Appendix B (Sec. 3, Table 2): Bias assessment in synthetic-based face recognition (FR)
3. Appendix C (Sec. 4, Table 5): Face recognition performance using different loss functions and network architectures
4. Appendix D (Sec 5, Table 3): Intra-class diversity and consistency on top of the Baseline(FFHQ) reported as age entropy, facial expression (Exp.) entropy, head-pose STD, D_{intra} and C_{intra}. These results are complementary to those in Table 2 (main submission). The results are reported on datasets generated with $lb \in [0.9, 0.8, 0.7, 0.6, 0.5, 0.4]$.
5. Appendix E (Sec. 6, Table 4): Intra-class diversity and consistency of our IDPERTURB compared to several State-of-the-art (SOTA) approaches
6. Appendix F (Sec. 7, Figure 1, Figure 2 and Figure 3): Histogram of Genuine-Imposter score distributions are shown in Figure 1, Figure 2 and Figure 3. These plots correspond to the results presented in the main submission in Tables 1 (Figure 1, Figure 2) and 3 Figure 3, respectively.
7. Appendix G (Sec. 8): Use of existing assets and face recognition evaluation benchmarks.
8. Appendix H (Sec. 9): Social Impact
9. Samples of the generated data in Figure 4, 5, 6 and 7. These samples complement the one provided in Figure 2 of the main submission.
10. The code to reproduce the results of the main submission is provided as a zip file CVPR2026_code.zip

2. Appendix A: Theoretical justification of IDPERTURB

Theoretical Justification for Angular Sampling Let $\mathbf{v} \in \mathbb{R}^d$ be a normalized identity embedding, i.e., $\|\mathbf{v}\|_2 = 1$. Modern face recognition (FR) models rely on angular distance between such embeddings to measure identity similarity. Given two embeddings $\mathbf{v}$ and $\tilde{\mathbf{v}}$, their cosine similarity is given by:

$$\cos(\theta_{v,\tilde{v}}) = \mathbf{v}^\top \tilde{\mathbf{v}}. \tag{1}$$

We define the angular neighborhood of $\mathbf{v}$ as the spherical cap:

$$\mathcal{C}_{lb}(\mathbf{v}) = \left\{ \tilde{\mathbf{v}} \in \mathbb{R}^d \;\middle|\; \mathbf{v}^\top \tilde{\mathbf{v}} \geq lb \right\}, \tag{2}$$

where $lb \in [0, 1]$ is a lower bound on cosine similarity (i.e., an upper bound on angular deviation $\theta_{\max} = \arccos(lb)$). Sampling within $\mathcal{C}_{lb}(\mathbf{v})$ ensures that perturbed embeddings remain semantically close to the original identity.

This geometric constraint:

- Maintains the unit norm of embeddings, preserving their location on the identity hypersphere.
- Ensures consistent and interpretable control over identity similarity, as cosine similarity correlates with FR decision boundaries.

This theoretical alignment explains the empirical success of angular perturbation, as shown in the main submission. It leverages the intrinsic structure of FR models' embedding space to inject meaningful diversity without significantly degrading identity coherence.

Comparison with Random Noise Perturbation. An alternative approach would be to add random noise in the Euclidean space:

$$\tilde{\mathbf{v}} = \mathbf{v} + \epsilon, \quad \epsilon \sim \mathcal{N}(0, \mathbf{I}), \tag{3}$$

followed by normalization. However, this method introduces limitations such as inconsistent angular similarity, as

Euclidean perturbations do not provide a guaranteed bound on angular deviation, leading to unpredictable identity similarity. This is supported by the results presented in Table 1, which reveal the poor FR recognition performance achieved when using datasets generated by models trained with the addition of random noise in comparison with IDiff-Face (baseline) and our IDPERTURB. It can also be observed that adding noise as context perturbation resulted in reduced identity-separability (high EER), which is coherent with the absence of a bounded context variation associated with this method.

Theoretical Properties of Angular Sampling

Lemma 1 (Norm Preservation)**.** *Let* $\tilde{\mathbf{v}} = \cos(\theta)\mathbf{v} + \sin(\theta)\mathbf{u}$*, where* $\|\mathbf{v}\| = \|\mathbf{u}\| = 1$*, and* $\langle \mathbf{v}, \mathbf{u} \rangle = 0$*. Then* $\tilde{\mathbf{v}}$ *satisfies:*

$$\|\tilde{\mathbf{v}}\| = 1$$

Proof.

$$\begin{aligned}\|\tilde{\mathbf{v}}\|^2 &= \|\cos(\theta)\mathbf{v} + \sin(\theta)\mathbf{u}\|^2 \\ &= \cos^2(\theta)\|\mathbf{v}\|^2 + \sin^2(\theta)\|\mathbf{u}\|^2 + 2\cos(\theta)\sin(\theta)\langle \mathbf{v}, \mathbf{u}\rangle \\ &= \cos^2(\theta) + \sin^2(\theta) + 0 = 1\end{aligned}$$

□

Lemma 2 (Cosine Similarity Control)**.** *Given* $\tilde{\mathbf{v}} = \cos(\theta)\mathbf{v} + \sin(\theta)\mathbf{u}$ *with* $\mathbf{u} \perp \mathbf{v}$*, it holds that:*

$$\langle \tilde{\mathbf{v}}, \mathbf{v} \rangle = \cos(\theta)$$

Proof.

$$\begin{aligned}\langle \tilde{\mathbf{v}}, \mathbf{v} \rangle &= \langle \cos(\theta)\mathbf{v} + \sin(\theta)\mathbf{u}, \mathbf{v} \rangle \\ &= \cos(\theta)\langle \mathbf{v}, \mathbf{v} \rangle + \sin(\theta)\langle \mathbf{u}, \mathbf{v} \rangle = \cos(\theta)\end{aligned}$$

□

Avoiding Identity Overlap

Theorem 1 Let $\mathcal{I} = \{\mathbf{v}_1, \ldots, \mathbf{v}_k\} \subset \mathbb{R}^d$ be a set of unit-normalized identity vectors. To ensure that a perturbed vector $\tilde{\mathbf{v}}$ sampled from $\mathcal{C}_{lb}(\mathbf{v}_i)$ remains closer to $\mathbf{v}_i$ than any $\mathbf{v}_l, l \neq i$, it suffices that:

$$\mathbf{lb} \leftarrow \max\left(\mathbf{lb}, \max_{l \neq i} \cos\left(\frac{\angle(\mathbf{v_i}, \mathbf{v_l})}{2}\right)\right). \tag{4}$$

Proof. Let v_j be the closest identity context to v_i in $\mathbb{R}^d$, that is $\angle(v_i, v_j) \leq \angle(v_i, v_l), v_l \in \mathcal{I} \wedge l \neq i$. Equation 4 can be rewritten as:

$$\mathbf{lb} \geq \cos\left(\frac{\angle(\mathbf{v_i}, \mathbf{v_j})}{2}\right) \leftrightarrow \theta_{\max} \leq \angle\frac{(\mathbf{v_i}, \mathbf{v_j})}{2}.$$

Considering that $\angle(v_i, \hat{v}) \leq \theta_{\max}$, it can be further derived that

$$\angle(v_i, v_j) \geq 2 \times \angle(v_i, \hat{v})$$

Taking the triangle inequality for angular distances into account:

$$\begin{aligned}\angle(v_i, v_j) &\leq \angle(v_i, \hat{v}) + \angle(\hat{v}, v_j) \leftrightarrow \\ 2 \times \angle(v_i, \hat{v}) &\leq \angle(v_i, \hat{v}) + \angle(\hat{v}, v_j) \leftrightarrow \\ \angle(v_i, \hat{v}) &\leq \angle(v_j, \hat{v}),\end{aligned}$$

which concludes the proof. Given the validity of the theorem on the most extreme scenario where $\angle(v_i, v_j) \leq \angle(v_i, v_l), v_l \in \mathcal{I} \wedge l \neq i$, it is trivial to derive its validity for all $l \neq i$.

□

3. Appendix B: Racial Bias in synthetic-based FR

Our FR models trained on the IDPERTURB dataset were evaluated on the Racial Faces in the Wild (RFW) dataset [21] to assess model bias and performance across different demographic groups. The RFW dataset contains four testing subsets corresponding to African, Asian, Caucasian, and Indian groups and includes approximately 80k images from 12k identities. We report the results as verification accuracies in (%) on each subset and as average accuracies to evaluate general recognition performance on the considered evaluation benchmarks. To evaluate the bias, we report the STD between all subsets and the SER, which is given by $\frac{max_g Error_g}{min_g Error_g}$, where g represents the demographic group, as reported in [20, 22]. A higher STD value indicates more bias across demographic groups and vice versa. For SER, models that achieved values closer to 1 are less biased. From the reported results in Table 2, comparing the bias assessment of FR models trained on the IDPERTURB dataset and SOTA data generation methods, we observe that IDPERTURB achieves lower STD and SER compared to the other considered base diffusion models trained on CasiaWebFace, while maintaining comparable STD and SER to FFHQ-based methods. Additionally, IDPERTURB achieves notably higher average accuracy on RFW compared to the other considered approaches.

4. Appendix C: Face recognition performance using different loss functions and network architectures

In the main submission, we followed established protocols from prior synthetic-based FR works and employed

Table 1. IDPERTURB vs. Noise perturbations. We compare IDPERTURB with a naive baseline in which identity embeddings are perturbed by adding Gaussian noise directly, without geometric constraints. For each of the base diffusion models trained on FFHQ and Casia-WebFace, we report results for three settings: (i) the unperturbed baseline, (ii) our proposed IDPERTURB, and (iii) random noise perturbations. We evaluate both the identity separability of the generated datasets and the face verification accuracy (%) of FR models trained on them. The results clearly show that while IDPERTURB introduces diversity without sacrificing identity semantics, random noise perturbations severely degrade identity consistency, as evidenced by the poor separability metrics and significantly lower verification performance. These findings highlight the importance of controlled, geometry-aware perturbations for generating effective synthetic training data in face recognition. These empirical results support our second justification in Appendix A.

	Identity-Separability of Training Data							**FR Verification**					
	Operation Metrics		**Score Distributions**					**Verification Benchmarks**					
			genuine		**imposter**								
Data Sampling Method	EER ↓	FMR100 ↓	mean	std	mean	std	FDR ↑	LFW	AgeDB-30	CFP-FP	CALFW	CP-LFW	Average ↑
Baseline (FFHQ)	0.005	0.004	0.509	0.104	0.023	0.081	13.680	97.60	84.10	83.36	89.08	78.77	86.58
IDPERTURB	0.124	0.332	0.245	0.133	0.023	0.060	2.316	98.55	88.85	84.27	91.42	80.85	**88.79**
Noise	0.498	0.949	0.009	0.060	0.008	0.060	0	57.48	50.97	56.76	53.07	51.45	53.94
Baseline (C-WF)	0.003	0.001	0.670	0.107	0.010	0.060	29.116	98.75	88.85	91.61	90.90	86.15	91.25
IDPERTURB	0.063	0.134	0.338	0.162	0.008	0.059	3.664	99.40	93.20	93.61	93.50	88.37	**93.62**
Noise	0.498	0.990	0.009	0.060	0.008	0.060	0	64.58	56.35	63.56	54.40	54.55	58.69

Table 2. Bias assessment of FR models trained on IDPERTURB dataset and SOTA data generation methods. The ethnicity-specific results correspond to verification accuracies in % on each ethnicity subset of RFW [21]. STD and SER are the bias metrics that can be calculated based on these verification accuracies. The best results for each metric are highlighted in bold. The results in the first row are obtained by training FR on authentic CASIA-WebFace, followed by FR model trained on digital rendering (DigiFace). We then present the results of the FR models trained on synthetically generated data, where the base generative model is trained on FFHQ, followed by a base generative model trained on CASIA-WebFace.

Models	Indian	Caucasian	Asian	African	Avg.	STD	SER
CASIA-WebFace [24]	88.50	93.38	85.38	86.62	88.47	3.52	2.21
DigiFace1M [1]	70.33	72.18	69.97	65.55	69.51	2.81	1.23
ExFaceGAN (FFHQ) [4]	71.58	73.80	71.07	62.20	69.66	5.11	1.44
IDiff-Face (FFHQ) [3]	81.50	85.37	79.77	75.78	80.61	3.98	1.66
IDperturb (FFHQ) (Ours)	84.25	86.78	81.18	77.17	82.35	4.14	1.73
SFace (Casia-WebFace) [2]	67.07	73.15	68.07	63.57	67.97	3.96	**1.36**
IDNet (Casia-WebFace) [14]	71.93	76.17	70.85	64.07	70.76	5.01	1.51
DCFace (Casia-WebFace) [12]	81.63	86.23	79.05	74.05	80.24	5.08	1.88
UIFace (Casia-WebFace) [15]	88.58	**91.68**	85.32	83.57	87.29	3.59	1.97
IDperturb (Casia-WebFace) (Ours)	**88.83**	91.60	**86.23**	**85.80**	**88.12**	**2.68**	1.69

Table 3. Intra-class diversity and consistency on top of the Baseline(FFHQ) reported as age entropy, facial expression (Exp.) entropy, head-pose STD, D_{intra} and C_{intra}. Higher age and expression entropy, and pose STD, and D_{intra} indicate greater intra-class diversity. Baseline (FFHQ) is generated from the DM without IDPERTURB. These results are complementary to those reported in the main submission Table 2.

		Attributes					D-Intra	C-Intra	FR-Perf.
Dataset	lb	Age ↑	Pose ↑			Exp. ↑			
			Yaw	Pitch	Roll				
Baseline (FFHQ)	-	0.252	15.261	6.369	2.051	0.379	0.329	0.999	86.58
IDPERTURB	0.9	0.282	15.32	6.432	2.055	0.425	0.345	0.998	87.51
	0.8	0.321	15.435	6.479	2.058	0.466	0.357	0.993	88.35
	0.7	0.361	15.490	6.565	2.081	0.496	0.367	0.978	88.73
	0.6	0.402	15.552	6.628	2.087	0.519	0.376	0.946	88.54
	0.5	0.446	15.590	6.709	2.106	0.539	0.384	0.894	88.79
	0.4	0.488	15.655	6.792	2.134	0.557	0.391	0.825	87.58

ResNet-50 [5, 7] as the backbone architecture, trained using the CosFace loss [19] function. To provide a more compre-

Table 4. Intra-class diversity and consistency of recent competitors and ours IDPERTURB on top of the Baseline(C-WF) reported as age entropy, facial expression (Exp.) entropy, head-pose STD, D_{intra} and C_{intra}. Higher age and expression entropy, and pose STD, and D_{intra} indicate greater intra-class diversity. The first row shows authentic C-WF results.

		Attributes					D-Intra	C-Intra	FR-Perf.
Dataset		Age ↑	Pose ↑			Exp. ↑			
	lb		Yaw	Pitch	Roll				
C-WF	-	0.354	23.479	9.069	5.154	0.589	0.423	0.948	94.63
IDPERTURB	0.9	0.325	19.907	8.384	3.606	0.492	0.393	0.998	92.68
	0.8	0.369	20.899	8.727	4.029	0.538	0.417	0.994	93.31
	0.7	0.416	21.816	9.117	4.551	0.574	0.439	0.978	93.44
	0.6	0.461	22.735	9.467	5.124	0.603	0.458	0.939	93.62
	0.5	0.501	23.611	9.724	5.602	0.621	0.474	0.875	93.56
	0.4	0.538	24.279	9.957	6.040	0.636	0.487	0.790	93.36
DCFace	-	0.462	17.926	7.257	3.059	0.583	0.368	0.965	89.56
SFace-60	-	0.426	24.567	9.715	5.402	0.627	0.405	0.777	77.71
DigiFace-6k	-	0.286	26.411	11.133	4.689	0.658	0.427	0.995	83.45
UIFace	-	0.386	21.478	8.593	4.103	0.608	0.414	0.988	93.27

hensive evaluation, we extend our experiments by training face recognition models on datasets generated with our proposed IDPERTURB (using a cosine lower bound $\mathbf{lb} = 0.6$ and classifier-free guidance scale $\omega = 1$) based on the IDiff-Face model originally trained on Casia-WebFace. We report performance across multiple training losses, including CosFace [19], ArcFace [5], and AdaFace [11], all utilizing ResNet-50. In addition, we assess two alternative network architectures—ResNet-100 [5, 7] and Vision Transformer (ViT) [6, 13], Vit-Small, to further examine the generalizability of IDPERTURB generated data across diverse FR model designs. The achieved results are presented in Table 5

5. Appendix D: Intra-Class Diversity and Consistency of IDPERTURB on top of Baseline (FFHQ)

In the main submission, we reported the age entropy, facial-expression entropy, standard deviation of head pose, intra-class perceptual diversity, and intra-class consistency of IDPERTURB applied on top of the Baseline (C-WF). For completeness, Table 3 additionally reports intra-class diversity and identity consistency, together with face-recognition performance, when IDPERTURB is applied on the Baseline (FFHQ). Consistent with the observations from Table 2 of the main submission, decreasing the lower bound lb from 0.9 to 0.4 results in a gradual increase in intra-class age and facial-expression variation without compromising identity coherence. The intra-class diversity D_{intra} increases monotonically as lb decreases, confirming that IDPERTURB effectively expands the perceptual manifold of each identity. These results further support the analysis presented in the main submission.

6. Appendix E: Intra-Class Diversity and Consistency of IDPERTURB compared with several SOTA approaches

We report in Table 4 the age entropy, facial expression entropy, STD of head-pose, intra-class perceptual diversity, and intra-class consistency. The results are reported on our IDPERTURB ($lb \in [0.9, 0.8, 0.7, 0.6, 0.5, 0.4]$) and compared with DCFace [12] introduces dual conditioning with identity and style inputs to control appearance variation, where diversity is achieved by sampling styles from a learned style bank. SFace [2] is a class-conditional GAN built on StyleGAN2-ADA [10] without explicit variation conditions. DigiFace [1] generates identities through physically based rendering using a computer graphics pipeline. UIFace [15] employs a two-phase sampling strategy, generating the early denoising steps unconditionally and introducing the identity condition only in the later steps, which promotes greater stochastic variation across samples.

Overall, IDPERTURB demonstrates a superior balance between intra-class diversity and identity consistency across all metrics. As the cosine lower bound lb decreases, we observe a consistent increase in diversity indicators—age and expression entropy, head-pose STD, and D_{intra}, while maintaining high identity coherence (C_{intra}) and strong face recognition (FR) performance. This confirms that our angular perturbation strategy effectively introduces variations in facial appearance without compromising identity coherence.

In contrast, DCFace, which achieves diversity through a style bank, exhibits lower expression entropy and pose variation. This could be attributed to a lack of specific variations in the style bank. SFace, while showing higher variation, suffers from poor identity consistency, reflected by its

significantly lower C_{intra} and FR accuracy. DigiFace, based on physically-based rendering, achieves high pose diversity but lacks the realistic appearance and semantic richness of diffusion-based approaches. UIFace, which introduces the identity condition only in the later denoising stages, attains moderate diversity with strong identity consistency, achieving closed, slightly lower, FR performance to our IDPERTURB.

7. Appendix F: Identity-Separability synthetic-based FR

Figures 1, 2 and 3 present the genuine and impostor score distributions of the data generated by our IDPERTURB trained on different datasets (FFHQ and CASIA) for different sampling hyperparameter values (lb and ω). These results correspond to the analysis of the identity separability metrics displayed in Tables 1 and 3 of the main submission.

8. Appendix G: Use of existing assets and face recognition evaluation benchmarks

8.1. Use of existing assets

The results of the SOTA face recognition models (in the main submission) are taken directly from their corresponding works. The pretrained IDiff-Face [3, 15]diffusion models on FFHQ and Casia-WebFace are provided under Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) license and Apache-2.0 license, respectively. Pretrained IDiff-Face on FFHQ is provided under https://github.com/fdbtrs/IDiff-Face and the one trained on Casia-WebFace is provided under https://github.com/Tencent/TFace/. Both models are trained using the same architectures and training setups described in the paper [3]. We provide the following details on the datasets used to train these models. **CASIA-WebFace [11, 15]:** CASIA-Webface consists of 494,141 face images from 10,757 different identities. A prepossessed (aligned and cropped) version of CASIA-WebFace is publicly available in the InsightFace (https://insightface.ai/) repository under Dataset-Zoo https://github.com/deepinsight/insightface/tree/master/recognition/_datasets_. The code and the databases of InsightFace are under MIT licence (https://github.com/deepinsight/insightface/blob/master/LICENSE).

FFHQ [9] Flickr-Faces-HQ Dataset (FFHQ) consists of 70,000 high-quality PNG images. The data is available for research under https://github.com/NVlabs/ffhq-dataset.

8.2. Evaluation benchmarks

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

AgeDB-30 [17]: 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 [17] and the details on the license are presented in https://ibug.doc.ic.ac.uk/resources/agedb/.

LFW [8]: 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 [8] and http://vis-www.cs.umass.edu/lfw/.

CFP-FP [18]: Celebrities in Frontal-Profile in the Wild (CFP-FP) [18] 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 [18] and http://www.cfpw.io/.

CALFW [26]: The Cross-age LFW (CALFW) dataset [26] 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 [26]. 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 [26] and http://whdeng.cn/CALFW/.

CPLFW [25]: The Cross-Pose LFW (CPLFW) dataset [25] 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 [25] and http://whdeng.cn/CPLFW/.

IJB-C [16]: The IARPA Janus Benchmark–C (IJB-C) [16] 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 [16] 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 [16] and https://www.nist.gov/programs-projects/face-challenges.

8.3. Similarity Threshold

A similarity threshold of 0.3 in Equation 8 (main submission) is selected following [11]. This choice corresponds to the operating point of the underlying FR evaluation model,

Table 5. Verification accuracies of FRs trained on IDPERTURB using CosFace, ArcFace and AdaFace loss functions as well as using different backbones, ResNet50, ResNet100 and Vit-S.

Network architecture	Loss function	Verfication accuracies (%)					
		LFW	AgeDB	CFP-FP	CA-LFW	CP-LFW	Avg.
ResNet50	CosFace	99.40	93.20	93.61	93.50	88.37	**93.62**
	ArcFace	99.27	92.17	93.06	93.00	87.73	93.05
	AdaFace	99.38	92.92	93.97	93.40	88.20	93.60
ResNet50	CosFace	99.40	93.20	93.61	93.50	88.37	93.62
ResNet100	CosFace	99.37	93.69	94.09	93.57	89.02	**93.95**
Vit-Small	CosFace	99.12	90.71	90.32	92.70	86.22	91,81

which reports a verification accuracy of 97.17% at a threshold of 0.3080 using $TPR@FPR = 0.01\%$ on the IJB-B benchmark [23]

Computing Infrastructure

This paper utilizes a computing infrastructure comprising 2 NVIDIA RTX A6000 GPUs to conduct the experiments. The code was tested on a Linux-based operating system, using PyTorch version 1.13.1. The versions of all relevant software libraries and frameworks are listed in the requirements file available in the code.zip folder.

9. Appendix H: Social Impact

This work is part of an ongoing research direction aimed at enabling the development of face recognition systems using fully synthetic data, thereby mitigating potential legal, ethical, and societal concerns associated with collecting and utilizing authentic biometric data. We emphasize that the advancement of synthetic-based face recognition serves to enhance the security, convenience, and overall quality of life for individuals—for instance, by facilitating secure access to financial or healthcare services, or improving the reliability of border control systems, all within well-defined legal and regulatory frameworks. Nonetheless, we firmly acknowledge and reject any malicious, unlawful, or discriminatory uses of this and related machine learning technologies.

References

[1] Gwangbin Bae, M. D. L. Gorce, Tadas Baltrušaitis, Charlie Hewitt, Dong Chen, Julien P. C. Valentin, Roberto Cipolla, and JingJing Shen. Digiface-1m: 1 million digital face images for face recognition. *2023 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pages 3515–3524, 2022. 3, 4

[2] Fadi Boutros, Marco Huber, Patrick Siebke, Tim Rieber, and Naser Damer. Sface: Privacy-friendly and accurate face recognition using synthetic data. *2022 IEEE International Joint Conference on Biometrics (IJCB)*, pages 1–11, 2022. 3, 4

[3] Fadi Boutros, Jonas Henry Grebe, Arjan Kuijper, and Naser Damer. Idiff-face: Synthetic-based face recognition through fizzy identity-conditioned diffusion models. In *IEEE/CVF International Conference on Computer Vision, ICCV 2023, Paris, France, October 1-6, 2023*, pages 19593–19604. IEEE, 2023. 3, 5, 7, 8

[4] Fadi Boutros, Marcel Klemt, Meiling Fang, Arjan Kuijper, and Naser Damer. Exfacegan: Exploring identity directions in gan's learned latent space for synthetic identity generation. In *IEEE International Joint Conference on Biometrics, IJCB 2023*, 2023. 3

[5] Jiankang Deng, Jia Guo, Jing Yang, Niannan Xue, Irene Kotsia, and Stefanos Zafeiriou. Arcface: Additive angular margin loss for deep face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(10):5962–5979, 2022. 3, 4

[6] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, and Neil Houlsby. An image is worth 16x16 words: Transformers for image recognition at scale. In *9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021*. OpenReview.net, 2021. 4

[7] Kaiming He, X. Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, 2015. 3, 4

[8] 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, 2007. 5

[9] Tero Karras, Samuli Laine, and Timo Aila. A style-based generator architecture for generative adversarial networks. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2019, Long Beach, CA, USA, June 16-20, 2019*, pages 4401–4410. Computer Vision Foundation / IEEE, 2019. 5

[10] Tero Karras, Miika Aittala, Janne Hellsten, Samuli Laine, Jaakko Lehtinen, and Timo Aila. Training generative adversarial networks with limited data. In *Advances in Neural Information Processing Systems 33: Annual Conference*

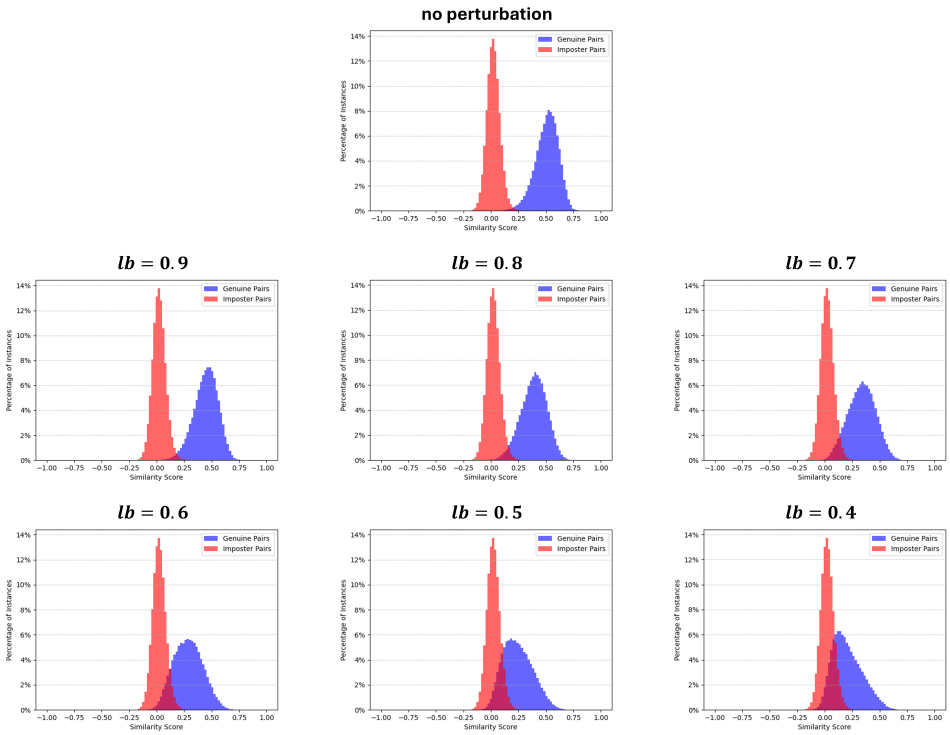

Figure 1. Histograms of the genuine and impostor score distributions for IDiff-Face [3] (baseline, no perturbation), and the proposed IDPERTURB method trained on the FFHQ dataset. The IDPERTURB results correspond to data generated by varying the d-dimensional cone boundary $\mathbf{lb} \in [0.4, 0.9]$. One can notice a decrease in identity-separability when the value of data $\mathbf{lb}$ decreases.These results correspond to a visual representation of the values presented in the first section of Table 1 of the main document.

on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual, 2020. 4

[11] Minchul Kim, Anil K. Jain, and Xiaoming Liu. Adaface: Quality adaptive margin for face recognition. In *CVPR*, pages 18729–18738. IEEE, 2022. 4, 5

[12] Minchul Kim, Feng Liu, Anil K. Jain, and Xiaoming Liu. Dcface: Synthetic face generation with dual condition diffusion model. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2023, Vancouver, BC, Canada, June 17-24, 2023*, pages 12715–12725. IEEE, 2023. 3, 4

[13] Minchul Kim, Yiyang Su, Feng Liu, Anil Jain, and Xiaoming Liu. Keypoint relative position encoding for face recognition. In *CVPR*, pages 244–255. IEEE, 2024. 4

[14] Jan Niklas Kolf, Tim Rieber, Jurek Elliesen, Fadi Boutros, Arjan Kuijper, and Naser Damer. Identity-driven three-player generative adversarial network for synthetic-based face recognition. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2023 - Workshops, Vancouver, BC, Canada, June 17-24, 2023*, pages 806–816. IEEE, 2023. 3

[15] Xiao Lin, Yuge Huang, Jianqing Xu, Yuxi Mi, Shuigeng Zhou, and Shouhong Ding. Uiface: Unleashing inherent model capabilities to enhance intra-class diversity in synthetic face recognition. In *ICLR*. OpenReview.net, 2025. 3, 4, 5

[16] 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 *ICB*, pages 158–165. IEEE, 2018. 5

[17] 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 Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 1997–2005, 2017. 5

[18] 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. *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 1–9, 2016. 5

[19] Hao Wang, Yitong Wang, Zheng Zhou, Xing Ji, Dihong Gong, Jingchao Zhou, Zhifeng Li, and Wei Liu. Cosface: Large margin cosine loss for deep face recognition. In *CVPR*, pages 5265–5274. Computer Vision Foundation / IEEE Computer Society, 2018. 3, 4

[20] Mei Wang and Weihong Deng. Mitigating bias in face recognition using skewness-aware reinforcement learning. In *Pro-*

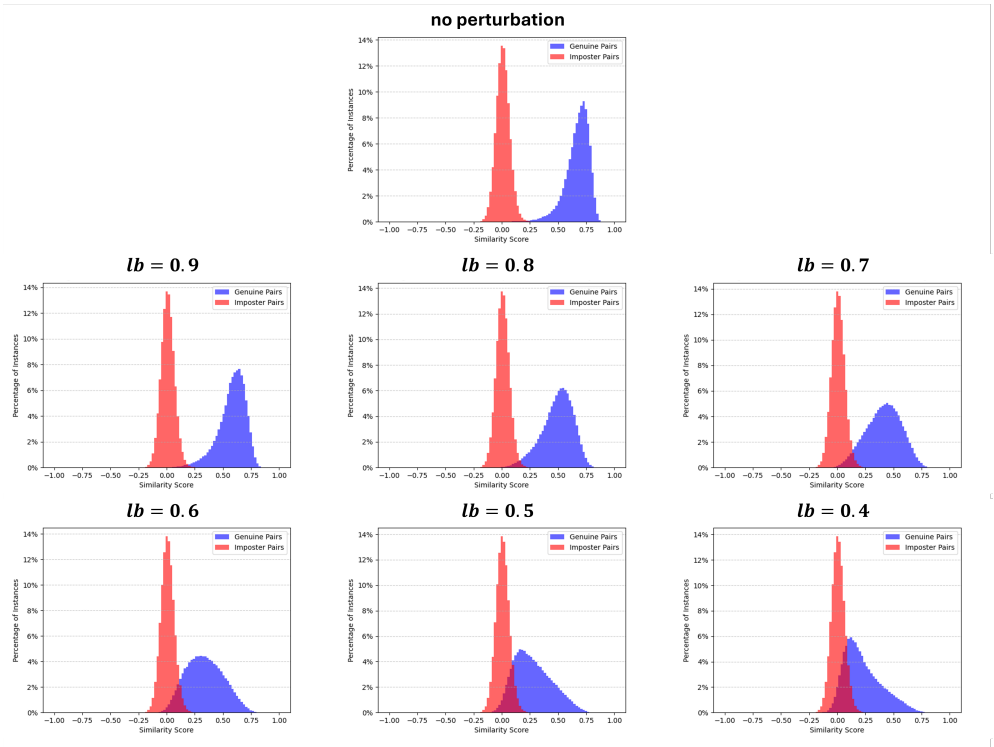

Figure 2. Histograms of the genuine and impostor score distributions for IDiff-Face [3] (baseline, no perturbation), and the proposed IDPERTURB method trained on the CASIA dataset. The IDPERTURB results correspond to data generated by varying the d-dimensional cone boundary $\mathbf{lb} \in [0.4, 0.9]$. One can notice a decrease in identity-separability when the value of data $\mathbf{lb}$ decreases.These results correspond to a visual representation of the values presented in the second section of Table 1 of the main document.

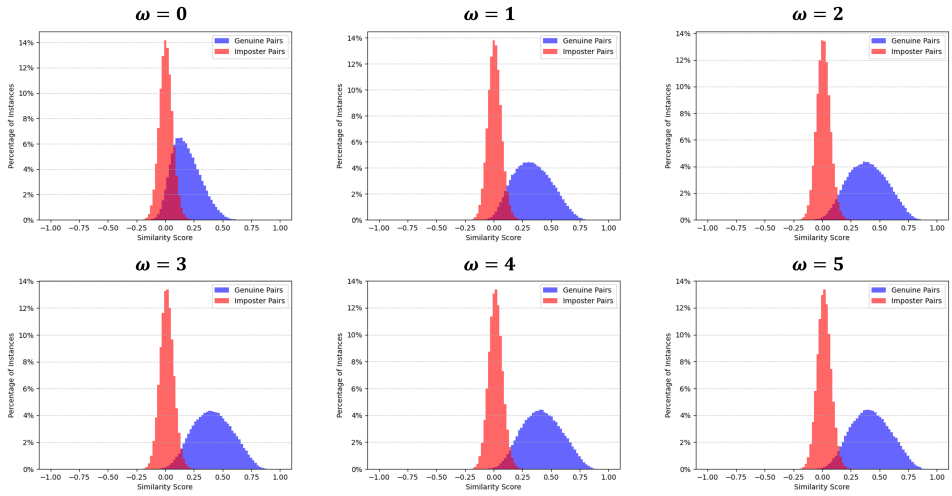

Figure 3. Histograms of the genuine and impostor score distributions for the proposed IDPERTURB method trained on the CASIA dataset. The IDPERTURB results correspond to data generated by varying scale $\omega \in [0, 5]$ and fixing the d-dimentional cone boundary as $\mathbf{lb} = 0.6$. One can notice a increase in identity-separability when the value of data ω increases.These results correspond to a visual representation of the values presented in Table 2 of the main document.

ceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 9322–9331, 2020. 2

[21] Mei Wang, Weihong Deng, Jiani Hu, Xunqiang Tao, and Yaohai Huang. Racial faces in the wild: Reducing racial bias by information maximization adaptation network. In *2019 IEEE/CVF International Conference on Computer Vi-*

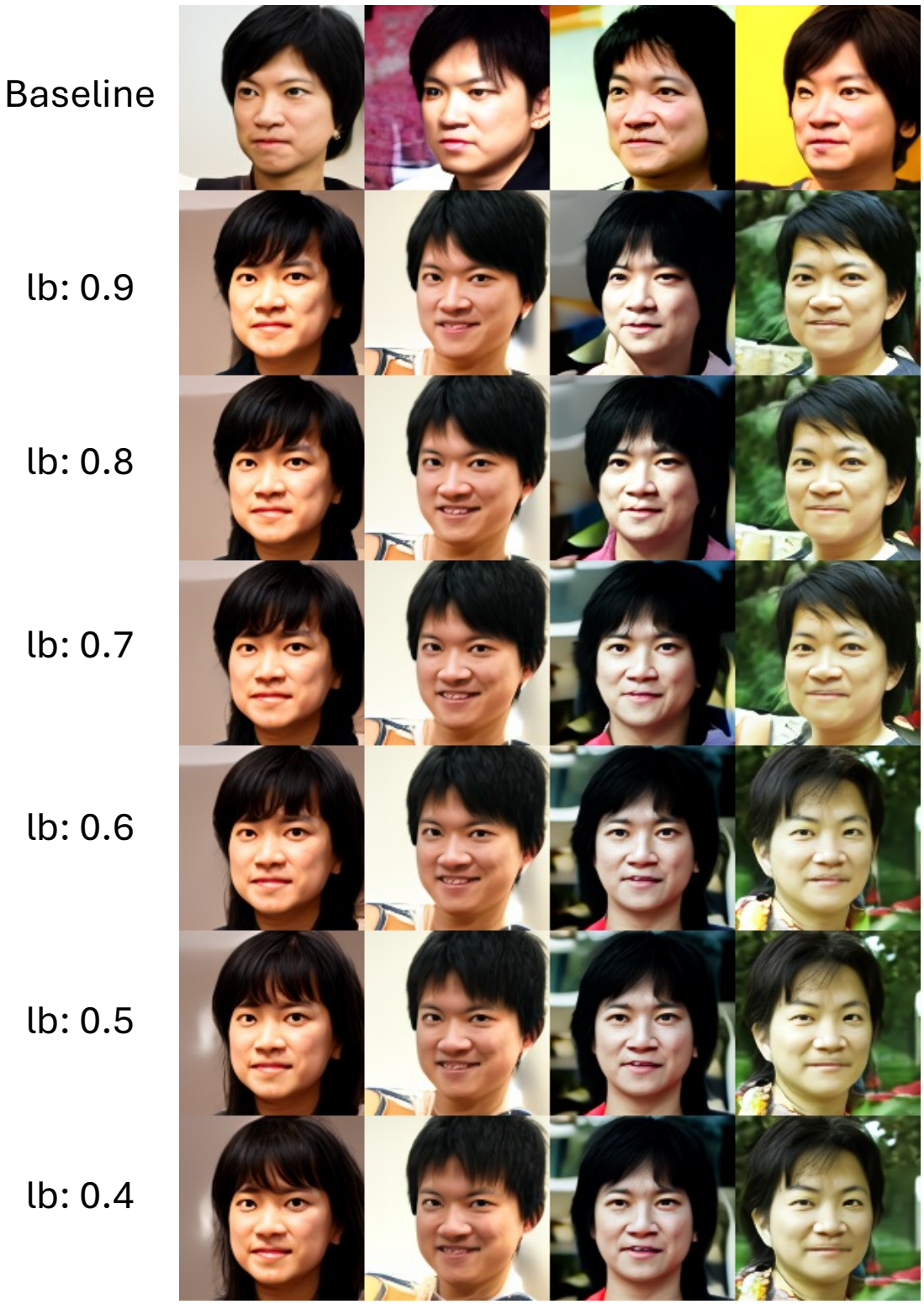

Figure 4. Samples generated from baseline model (without perturbation) and the ones generated using IDPERTURB with $lb \in [0.4, 0.9]$

sion, ICCV 2019, Seoul, Korea (South), October 27 - Novem- ber 2, 2019, pages 692–702. IEEE, 2019. 2, 3

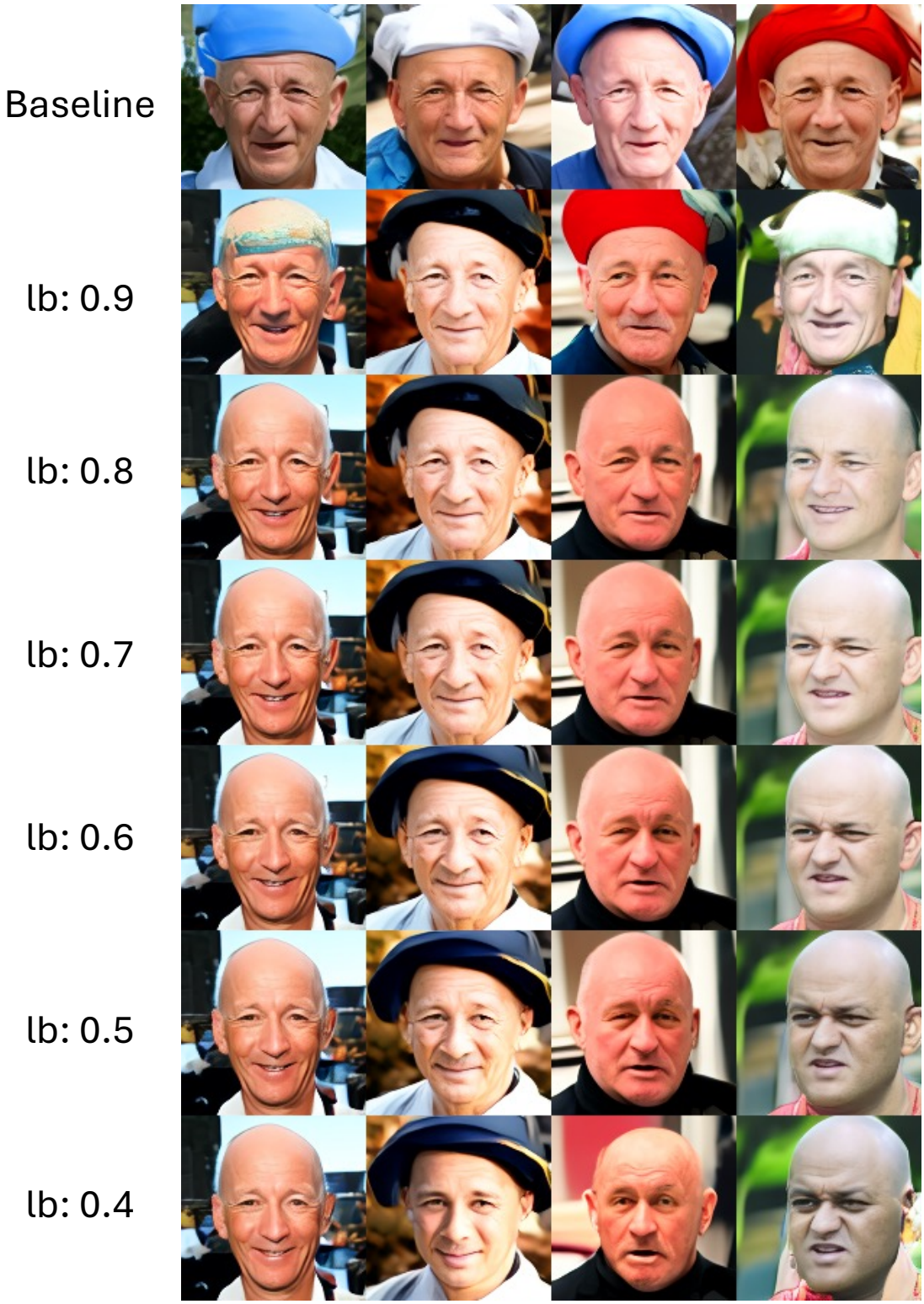

Figure 5. Samples generated from baseline model (without perturbation) and the ones generated using IDPERTURB with $lb \in [0.4, 0.9]$

[22] Mei Wang, Yaobin Zhang, and Weihong Deng. Meta balanced network for fair face recognition. *IEEE Trans. Pattern*

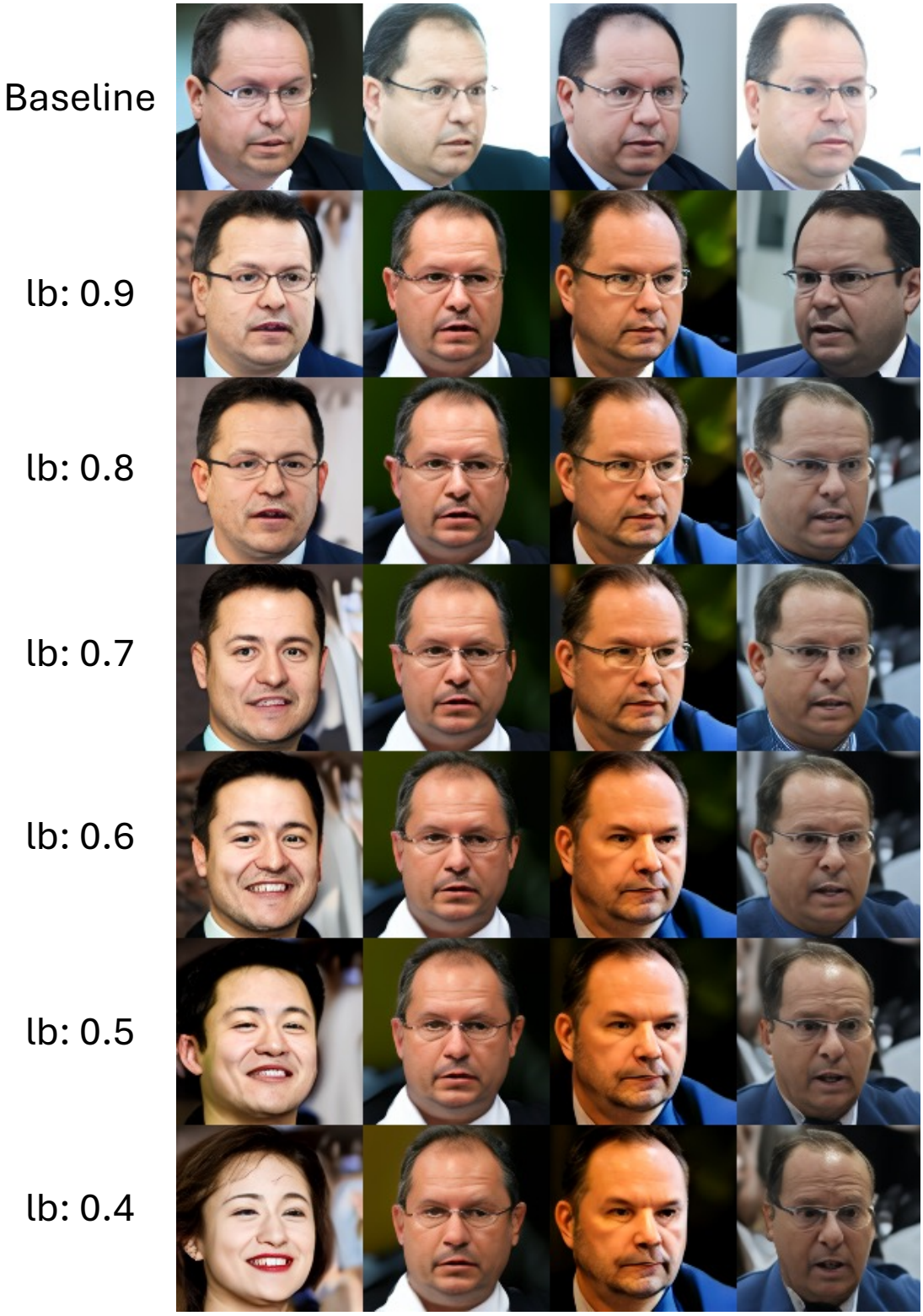

Figure 6. Samples generated from baseline model (without perturbation) and the ones generated using IDPERTURB with $lb \in [0.4, 0.9]$

Figure 7. Samples generated from baseline model (without perturbation) and the ones generated using IDPERTURB with $lb \in [0.4, 0.9]$

Anal. Mach. Intell., 44(11):8433–8448, 2022. 2

[23] 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 Ch-

eney, and Patrick Grother. IARPA janus benchmark-b face dataset. In *CVPR Workshops*, pages 592–600. IEEE Computer Society, 2017. 6

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

[25] 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, 2018. 5

[26] 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. 5