Learning to Attract and Repel: Dual Quality Margin Learning for Face Recognition (DQM-Face)

El Ouanas Belabbaci[1], Bhavesh Wani[2], and Philipp Terhörst[2]

[1] Laboratoire d'Ingénierie des Systèmes Intelligents et Communicants (LISIC), Faculty of Electrical Engineering, University of Sciences and Technology Houari Boumediene (USTHB), Algiers, Algeria
[2] Johannes Gutenberg University Mainz, Germany

Abstract. Face recognition in unconstrained environments remains highly challenging due to diverse and extreme variations encountered in real-world scenarios. To mitigate these effects, existing margin-based approaches model sample quality through feature magnitude. However, magnitude-based modeling alone is susceptible to identity-agnostic noise, which can degrade the reliability and discriminative power of learned representations. In this paper, we propose Dual Quality Margin Learning for Face Recognition (DQM-Face), a novel framework that enables refined attraction and repulsion dynamics during representation learning. Our approach unifies conventional magnitude-based quality estimation with a newly introduced semantic quality learning mechanism, realized via squeeze-and-excitation semantic attention. By jointly leveraging magnitude and semantic cues, we construct enhanced quality-aware margins that adaptively strengthen intra-class compactness through improved attraction during learning. To further enhance inter-class discrimination, we introduce a repulsion margin formulation that explicitly enlarges inter-class separation. The unified integration of semantic quality modeling with dual attraction–repulsion margin optimization results in a more structured and discriminative feature geometry. Extensive experiments on multiple challenging benchmarks demonstrate that DQM-Face consistently outperforms state-of-the-art face recognition methods. Moreover, we show that the quality learned for margin optimization is highly effective for face image quality assessment within the proposed framework, demonstrating that the learned quality signal is intrinsically aligned with the recognition objective. The code is publicly available: `https://github.com/RAIB-group/DQM-Face`

Keywords: Face Recognition, Face Image Quality Assessment, Quality-Aware Face Recognition

1 Introduction

Face recognition (FR) remains one of the most widely deployed biometric technologies due to the collectability, universality, and high discriminative potential

of the facial trait [1]. With the proliferation of applications in security, surveillance, and authentication, these systems are increasingly used in unconstrained real-world scenarios [44]. In such environments, recognition robustness is frequently challenged by extreme variations, such as in pose, illumination, occlusion, expression, or age, often resulting in degraded representations and a higher likelihood of false matches.

To analyze and mitigate the effects of these variations, face image quality assessment (FIQA) has emerged as a key research direction. FIQA methods aim to estimate the utility of a face image for recognition, enabling systems to quantify image reliability and improve matching performance by rejecting low-quality samples. Parallel research in representation learning has focused on developing margin-based loss functions [8,29,43], which improve feature discriminability by enforcing intra-class compactness and inter-class separation. More recently, approaches that integrate quality information directly into training objectives [22,31,39] have shown promising advances, allowing networks to adapt decision making and learning behavior based on sample difficulty. However, current quality-aware margin formulations predominantly rely on feature magnitude as a proxy for image quality. While effective to some extent, magnitude-based estimation remains identity-agnostic: factors such as blur, occlusion, or illumination can artificially inflate embedding norms, leading to inaccurate quality predictions. Furthermore, existing methods primarily focus on intra-class compactness without explicitly enforcing inter-class repulsion, leaving a residual risk of feature overlap across identities in challenging conditions [28,38].

To address these limitations, we propose **Dual Quality Margin Learning for Face Recognition (DQM-Face)**, a unified framework that incorporates dual quality modeling and dual margin optimization. Our method combines conventional magnitude-based quality estimation with a novel semantic quality branch implemented via squeeze-and-excitation attention. This dual formulation enables identity-aware, context-dependent quality assessment that better reflects the true recognizability of each sample. Based on these quality cues, DQM-Face introduces a dual-margin optimization mechanism that adaptively scales the target (attractive) margin according to sample quality while simultaneously applying an explicit inter-class repulsion margin. This synergy ensures compact clustering of high-quality samples while maintaining sufficient angular separation between different identities.

Extensive experiments on both image-based and video-based benchmarks demonstrate that DQM-Face consistently achieves state-of-the-art performance. The proposed approach delivers robust recognition under severe variations in pose, illumination, and image quality, ranking first or second across all major protocols, including LFW, CFP-FP, AgeDB-30, CPLFW, IJB-B, and IJB-C.

In summary, the main contributions of this work are as follows: (1) We propose a unified Dual Quality Margin Learning (DQM-Face) framework that integrates sample quality modeling and margin-based optimization for more robust face representation learning. (2) We introduce a dual quality modeling mechanism that fuses magnitude-based and semantic attention–based quality cues, produc-

ing more reliable and identity-sensitive quality estimates. These prove to be highly effective for face image quality assessment within the proposed framework and face recognition in general. (3) We design a dual-margin optimization strategy combining an adaptive attraction margin and a curriculum-based repulsion margin to enhance both intra-class compactness and inter-class separability. (4) We achieve consistent state-of-the-art results across multiple challenging benchmarks, validating the effectiveness and generalization ability of the proposed approach.

2 Related Work

2.1 Face Image Quality

Face Image Quality Assessment (FIQA) aims to estimate the utility of a face image for recognition, thereby predicting the expected impact on recognition performance. Early FIQA methods were based on predefined image quality metrics, whereas recent methods learn quality predictors aligned with face recognition performance. Uncertainty-driven methods, including SER-FIQ [40], estimate quality by measuring embedding stability under stochastic perturbations. Recognition-aware learning-based methods, including CR-FIQA [7], CLIB-FIQA [33], GraFIQs [25], FROQ [2], and FaceQAN [3], learn quality predictors directly from face recognition supervision, such as genuine–impostor score separability or recognition score distributions, in supervised or semi-supervised settings. Diffusion-based approaches, such as DiffIQA [4] and eDiffIQA [5], use diffusion models to estimate image utility based on reconstruction consistency or denoising stability. While these methods provide valuable indicators of facial image quality, they are typically used as external assessors and not integrated into the feature learning process itself.

2.2 Face Recognition

The introduction of margin-based softmax losses such as SphereFace [29], CosFace [43], and ArcFace [8] established the foundation for modern FR by enforcing margins in the normalized embedding space. These formulations improve intra-class compactness and inter-class separability, making them standard practice for identity-discriminative feature learning. However, fixed-margin formulations treat all samples equally, ignoring variations in sample difficulty and image quality. To address these limitations, several methods have explored adaptive margin mechanisms and optimization curricula. CurricularFace [19] applied curriculum learning to progressively increase margin difficulty during training, while ElasticFace [6] introduced stochastic margin sampling from a Gaussian distribution to improve regularization and model generalization. GroupFace [23] extended margin-based learning by incorporating group-aware supervision to capture structured intra-class relations across identities. More recently,

CoReFace [38] employed contrastive regularization to explicitly enhance inter-class separation and stabilize hard-sample optimization. QCFace [9] further revisited quality-driven representation learning by integrating content consistency constraints within the margin optimization process. Despite these advances, most methods remain agnostic to explicit image quality characteristics, leaving robustness under unconstrained conditions an open challenge.

2.3 Quality-Aware Face Recognition

Integrating face image quality directly into the feature learning and decision-making process has proven to be an effective strategy for achieving robust performance in unconstrained environments. MagFace [31] introduced a unified learning framework in which the embedding magnitude serves as an implicit indicator of image quality, enabling adaptive margin scaling during training and resulting in more discriminative representations. QMagFace [39] extended this concept by incorporating the magnitude-derived quality information into the decision-making process, improving robustness under challenging conditions. AdaFace [22] further refined MagFace's quality–embedding principle by learning the relationship between feature magnitude and verification performance through adaptive quality weighting. Although these methods achieve strong performance, their magnitude-only quality proxies remain sensitive to identity-agnostic noise such as blur, occlusion, or illumination variations, which can distort the underlying quality estimation.

In contrast, our proposed DQM-Face integrates both magnitude-based and semantic quality modeling within the learning objective itself. By combining feature magnitude with semantic attention–based cues and jointly optimizing attraction and repulsion margins, DQM-Face produces identity-aware, quality-adaptive embeddings that maintain high discriminability even under severe data variations.

3 Methodology

In this section, we present the proposed Dual Quality Margin Learning framework (DQM-Face) designed to jointly account for sample quality and class separability in face representation learning. Building upon the standard margin-based softmax formulation, our method extends traditional magnitude-aware approaches by integrating dual quality modeling and dual margin optimization. Specifically, we combine magnitude- and semantic-based quality estimation to derive an adaptive attraction margin that promotes intra-class compactness, while introducing an explicit repulsion margin that enhances inter-class discriminability. The resulting Dual Quality Margin (DQM) loss unifies these components within a single optimization objective, enabling the model to learn structured and quality-aware embeddings suitable for robust recognition under unconstrained conditions.

Figure 1 shows an overview of the effect of different loss functions on decision boundaries in the feature space. While the standard softmax leads to an overlap zone (a) and ArcFace works with an unflexible and fixed margin (b), DQM-Face leads to a highly structured feature space with a wide clearance zone and inter-class compactness (e). This is achieved through the adaptive attraction (c) and the explicit repulsion margin (d).

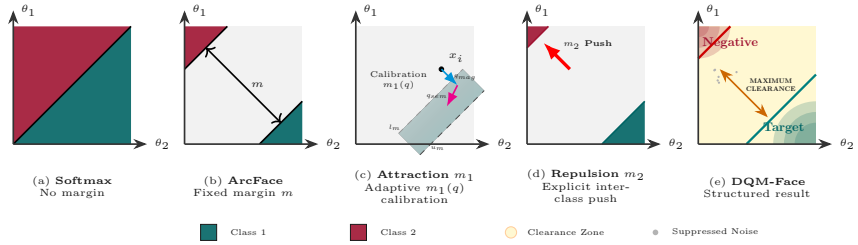


Fig. 1: Geometrical interpretation - Feature space and decision boundaries are shown for (a) Softmax baseline with inter-class overlap, (b) ArcFace with a fixed margin m, (c) our adaptive attraction margin m_1 calibrated by dual-quality fusion, (d) explicit inter-class repulsion m_2 pushing non-target features away, and (e) the resulting structured feature space characterized by maximized intra-class compactness and a wide Clearance Zone.

3.1 Dual Quality Margin Loss

In face recognition, the most commonly used loss formulation is the margin-based softmax loss [8]

$$\mathcal{L}_{base} = -\frac{1}{N}\sum_{i=1}^{N}\log\frac{e^{s\cdot\phi(\theta_{y_i},m)}}{e^{s\cdot\phi(\theta_{y_i},m)}+\sum_{j\neq y_i}e^{s\cos\theta_j}}, \tag{1}$$

where $\mathbf{x}_i \in \mathbb{R}^d$ denotes the face representation (template) of the i-th sample belonging to class y_i, $\theta_j = \arccos(W_j^T\mathbf{x}_i/s)$ describes the angle between the representation $\mathbf{x}_i$ and the class center $W_j \in \mathbb{R}^d$ for class j and $\phi(\theta_{y_i}, m)$ describes a margin function. For $\phi(\theta_{y_i}, m) = \cos(\theta_{y_i} + m)$, this leads to the ArcFace loss [8] and for $\phi(\theta_{y_i}, m) = \cos\theta_{y_i} - m$, this leads to the CosFace loss [43]. These formulations enforce intra-class compactness by increasing the margin between samples and their ground-truth centers, but they lack mechanisms to account for sample quality or explicitly separate inter-class features.

In this work, we propose a Quality Magnitude Combination (QMC) loss

$$\mathcal{L}_{QMC} = \mathcal{L}_{DQM}(x_i, m_1(q_{mag}, q_{sem}), m_2 + \lambda_g\mathcal{L}_{mag}(\|\mathbf{x}_i\|), \tag{2}$$

consisting of our novel Dual Quality Margin (DQM) loss and MagFace [31] loss. The MagFace loss $\mathcal{L}_{mag}(\|\mathbf{x}\|) = \frac{1}{\|\mathbf{x}\|} + \frac{\|\mathbf{x}\|}{u_a^2}$, is a regularization term that enforces magnitude monotonicity and that the length of the representation encodes its quality. The Dual Quality Margin (DQM) loss

$$\mathcal{L}_{DQM} = -\frac{1}{N}\sum_{i=1}^{N}\log\frac{e^{s\cdot\cos(\theta_{y_i}+m_1(q_{mag,i},q_{sem,i}))}}{e^{s\cdot\cos(\theta_{y_i}+m_1(q_{mag,i},q_{sem,i}))}+\sum_{j\neq y_i}e^{s\cdot\cos(\theta_j-m_2)}} \tag{3}$$

consists of the attractive quality margin $m_1(q_{mag,i}, q_{sem,i})$ based on magnitude and semantic qualities, as well as the explicit repulsion margin m_2. The gradient of the combined loss $\mathcal{L}_{QMC}$ demonstrates its effect well during training

$$\frac{\partial\mathcal{L}_{DQM}}{\partial\mathbf{x}_i} = \underbrace{\frac{\partial\mathcal{L}_{DQM}}{\partial\cos\theta_{y_i}}\cdot\frac{\partial\cos\theta_{y_i}}{\partial\mathbf{x}_i}}_{\text{intra-class pull}} + \underbrace{\sum_{j\neq y_i}\frac{\partial\mathcal{L}_{DQM}}{\partial\cos\theta_j}\cdot\frac{\partial\cos\theta_j}{\partial\mathbf{x}_i}}_{\text{inter-class push}} + \underbrace{\lambda_g\frac{\partial\mathcal{L}_{mag}}{\partial\|\mathbf{x}_i\|}\cdot\frac{\partial\|\mathbf{x}_i\|}{\partial\mathbf{x}_i}}_{\text{quality regularization}}$$

The first term pulls the representation toward its ground-truth center with strength modulated by the adaptive margin m_1. The second term pushes it away from all non-target centers with a repulsive force determined by m_2. The third term enforces the quality-magnitude correlation, ensuring that the feature norm accurately reflects the recognizability of the sample. In the following, we describe the contributions of both margins in more detail.

3.2 Attraction Margin Learning with Quality Awareness

Magnitude-based Quality Learning For the magnitude-based quality learning, following [31], the feature norm $\|\mathbf{x}_i\|$ is used as the magnitude-based quality indicator. Given the operational bounds, the quality score is defined as

$$q_{mag} = \frac{\min(\max(\|\mathbf{x}_i\|, l_a), u_a) - l_a}{u_a - l_a}, \tag{4}$$

where l_a and u_a are the lower and upper bounds of the feature norm, respectively. The operation $\min(\max(\|\mathbf{x}_i\|, l_a), u_a)$ ensures that the feature norm is projected into the interval $[l_a, u_a]$, preventing extreme values from dominating the quality estimation. This linear normalization maps the bounded feature norm to a quality score. However, this does not cover identity-agnostic noise, which can degrade the discriminativeness of the representations.

Semantic Quality Learning As shown in [9], embedding magnitude, such as [22, 31], can be significantly inflated by identity-agnostic artifacts such as motion blur, occlusion, or illumination variations, leading the model to assign high quality scores to unidentifiable samples. To overcome this issue of identity-agnostic noise, we introduce a semantic quality learning approach based on squeeze-and-excitation semantic attention. Given a face representation $\mathbf{x}$, a

squeeze-and-excitation gating mechanism [17] is utilized to create an attended representation

$$\mathbf{x}_{att} = \mathbf{x} \odot \sigma(\mathbf{W}_2 \cdot \delta(\mathbf{W}_1 \mathbf{x})), \tag{5}$$

with $\odot$ denotes element-wise multiplication, $\sigma(\cdot)$ is the sigmoid activation, $\delta(\cdot)$ is the GELU activation [16], and $\mathbf{W}_1 \in \mathbb{R}^{\frac{d}{r} \times d}$, $\mathbf{W}_2 \in \mathbb{R}^{d \times \frac{d}{r}}$ are learnable parameters. The reduction ratio r is a hyperparameter that controls the bottleneck dimension. The attended representation $\mathbf{x}_{att}$ is then passed through a small multi-layer perceptron to produce a semantic quality score $q_{sem} \in [0, 1]$

$$q_{sem} = \sigma\left(\mathbf{W}_4 \cdot \mathrm{BN}\left(\delta(\mathbf{W}_3 \mathbf{x}_{att})\right)\right), \tag{6}$$

where $\mathbf{W}_3 \in \mathbb{R}^{d \times d}$, $\mathbf{W}_4 \in \mathbb{R}^{1 \times d}$, and BN denotes batch normalization. This score represents the estimated "identity-utility" of the feature, ensuring that only facial structures genuinely contributing to recognizability are prioritized during margin adaptation.

Magnitude and Semantic Quality-based Margin Learning Lastly, an unified quality score q

$$q = \alpha \cdot q_{sem} + (1 - \alpha) \cdot q_{mag}, \tag{7}$$

is obtained through a simple combination of the semantic and magnitude-based quality scores. Here, $\alpha \in [0, 1]$ is a trade-off parameter balancing the physical reliability of the embedding norm with semantic consistency. As we will show later, this quality score q is highly effective for FIQA within the DQM-Face framework. The fused quality score q is then used to adaptively scale the target-class margin m_1:

$$m_1(q) = (u_m - l_m) \cdot q + l_m, \tag{8}$$

where l_m and u_m are the lower and upper bounds of the angular margin. This linear mapping ensures that higher-quality samples receive larger margins, pulling them closer to their class centers, while lower-quality samples are subject to weaker constraints, preventing overfitting on ambiguous features.

3.3 Explicit Repulsion Margin Learning

Most margin-based losses focus on intra-class compactness by enforcing a margin between samples and their ground-truth centers. However, this one-sided optimization can increase the risks of increase false matches as noted in [28, 38]. Therefore, we explicitly enforce a repulsive margin m_2 to maximize discriminability. In Eq 3, the logit is defined as

$$z_j = s \cdot \cos(\theta_j - m_2), \tag{9}$$

with the repulsion margin $m_2 > 0$. This formulation creates an asymmetric decision boundary by actively pushing features away from incorrect class centers.

Geometrically, it enforces a stricter angular "clearance zone" around each class, ensuring that samples not only cluster tightly around their own centers but also maintain distance from all other centers.

To improve optimization stability, we adopt a curriculum-based schedule for the inter-class repulsion margin m_2. Specifically, m_2 is progressively increased in three stages throughout training following:

$$m_2^{(t)} = \begin{cases} 0.05, & 1 \leq t \leq 10 \\ 0.10, & 11 \leq t \leq 18 \\ 0.15, & 19 \leq t \leq 25 \end{cases} \tag{10}$$

where t is the current epoch. This staged strategy allows stable cluster formation during early epochs ($m_2 = 0.05$), followed by moderate inter-class separation in the intermediate stage ($m_2 = 0.10$), before enforcing a stronger repulsive margin ($m_2 = 0.15$) in the final stage. This progressive schedule prevents gradient instability while maximizing global separability in the embedding space.

4 Experiments

4.1 Datasets, Training, and Evaluation

Model training is conducted on the refined MS1MV2 dataset [8], which contains approximately 5.8 million face images spanning 85,742 unique identities. Following standard protocols, all identities overlapping with evaluation benchmarks are removed to ensure unbiased open-set evaluation. Training is performed using Stochastic Gradient Descent (SGD) with a momentum of 0.9 and a weight decay of 5×10^{-4}. The initial learning rate is set to 0.1 and reduced by a factor of 10 at epochs 10, 18, and 22, with training terminating at epoch 25. The feature scale is fixed at $s = 64$, following standard practice in margin-based losses [8,43]. All models are trained with a batch size of 512.

For evaluating face recognition, we conduct experiments on widely adopted face verification benchmarks that cover diverse variations in pose, age, and image quality: LFW [18], CFP-FP [36], AgeDB-30 [32], and CPLFW [46]. These benchmarks provide predefined face pairs that enable model comparison using the accuracy metric. For more challenging large-scale evaluations, we further experiment with the IJB-B [45] and IJB-C [30] benchmarks, which include both still images and video frames with significant variations in resolution, occlusion, motion blur, and illumination. On these datasets, we follow the official 1:1 verification protocol and report the True Accept Rate (TAR) at different False Accept Rate (FAR) operating points, ranging from 10^{-3} to 10^{-5}. Note that these are system-level metrics that account for preprocessing errors (e.g., detection and alignment), consistent with standard face recognition evaluation practices [20].

To evaluate Face Image Quality Assessment (FIQA) performance, experiments are conducted on four widely used face verification benchmarks: LFW [18], Adience [11], CPLFW [46], and XQLFW [24]. These datasets represent diverse

challenges, including age variations (Adience), pose variations (CPLFW), severe image quality degradations (XQLFW), and an unconstrained in-the-wild setting (LFW). Performance is compared against nine state-of-the-art quality estimation methods, including eDifFIQA [5], DifFIQA [4], CR-FIQA [7], CLIB-FIQA [33], GraFIQs [25], FROQ [2], MagFace [31], SER-FIQA [40], and FaceQAN [3]. Evaluation is conducted using the Error-vs-Discard (EvD) Characteristics [21], also known as Error-vs-Reject curves [41]. In this procedure, images with the lowest predicted quality scores are progressively removed, and recognition performance is evaluated on the remaining subset. The resulting curves quantify how effectively a quality assessment method identifies and filters low-quality samples to reduce recognition errors. For the error rate metric, we follow the recommendations of the European Border Control Agency (Frontex) [13], computing the False Non-Match Rate (FNMR) at a fixed False Match Rate (FMR) of 10^{-3}. To enable concise and comparable evaluation across methods, we additionally compute the partial Area Under the Curve (pAUC) of the EvD up to a 30% rejection rate [34]. The pAUC represents the area under the EvD between 0% and 30% rejection; lower pAUC values indicate better FIQA performance. All evaluations are performed using face embeddings extracted from the proposed DQM-Face model, since the predicted quality is model-specific and this ensures consistent comparisons across different FIQA methods.

4.2 Architecture and Preprocessing

For the backbone network, a iResNet-100 architecture [10] is used, which extracts 512-dimensional face embeddings. The semantic-attention branch is implemented as a lightweight module inserted after the global average pooling layer, consisting of a Squeeze-and-Excitation (SE) block with a reduction ratio of 16, followed by a two-layer MLP with hidden dimension 512. For preprocessing, we follow the InsightFace framework [14].

4.3 Magnitude Hyperparameter Settings

The main hyperparameters used in DQM-Face follow standard practices from margin-based face recognition and refinements by [31]. The quality-dependent target margin is linearly interpolated between a lower bound of 0.35 and an upper bound of 0.8. The feature magnitude bounds are set to 10 and 110, with the magnitude regularization weight fixed at 35. Otherwise, the quality fusion weight is set to 0.5, balancing the contributions of magnitude- and semantic-based quality components. The maximum inter-class repulsive margin is fixed at 0.15 and follows a curriculum scheduling strategy during training. The choice of these is justified in the ablation studies.

Table 1: Recognition Performance on Single-Image Benchmarks - The highest verification accuracy (%) is highlighted in bold, and the second-best result is underlined. The proposed DQM-Face solution performs well across the different benchmarks.

Method	LFW	AgeDB-30	CFP-FP	CPLFW	Avg.
SphereFace [29]	99.67	97.05	96.84	91.27	96.21
CosFace [43]	99.81	98.12	98.12	92.48	97.13
ArcFace [8]	<u>99.83</u>	98.15	98.40	92.72	97.28
GroupFace [23]	**99.85**	98.28	98.63	93.17	97.48
CurricularFace [19]	99.80	98.32	98.37	93.13	97.41
MagFace [31]	<u>99.83</u>	98.17	98.46	92.87	97.33
AdaFace [22]	99.82	98.05	98.49	**93.53**	97.47
ElasticFace-Arc+ [6]	99.82	98.35	98.67	93.28	<u>97.53</u>
ElasticFace-Cos+ [6]	99.80	98.28	98.73	93.23	97.51
QMagFace [39]	<u>99.83</u>	**98.50**	<u>98.74</u>	-	-
CoReFace [38]	-	<u>98.37</u>	98.60	93.27	-
QCFace [9]	99.80	98.18	98.46	92.83	97.31
DQM-Face $\alpha_{0.4}$ (Ours)	<u>99.83</u>	98.32	98.61	93.22	97.50
DQM-Face $\alpha_{0.5}$ (Ours)	<u>99.83</u>	98.35	**98.83**	<u>93.30</u>	**97.58**

5 Results

5.1 Recognition Performance on Single-Image Benchmarks

To assess the effectiveness of the proposed DQM-Face method, Table 1 presents its verification performance in comparison with state-of-the-art face recognition approaches on standard image-to-image benchmarks. These evaluations cover challenging scenarios such as cross-pose and cross-age matching. All models were trained on the MS1MV2 dataset [15] using an iResNet-100 backbone [10]. The results indicate that DQM-Face consistently achieves superior performance, ranking first or second across all evaluated benchmarks, demonstrating the effectiveness of the quality-guided dual margin. We attribute the consistent performance observed across these challenging data pairs to the integration of quality estimation into the learning process, which facilitates context-dependent representation learning.

However, as shown in [12], these benchmarks are relatively small in size, and the resulting performance differences are so minor that they cannot be considered statistically significant. Therefore, we conduct additional experiments on the large-scale IJB-B and IJB-C benchmarks to provide a more robust evaluation.

5.2 Recognition Performance on Video-based Benchmarks

To evaluate the performance of the proposed DQM-Face method on large-scale datasets, Table 2 reports the TAR at several FARs on the widely used IJB-B and IJB-C benchmarks. The results, compared with those of state-of-the-art approaches, demonstrate that DQM-Face consistently achieves the best or

Table 2: Recognition Performance on Video-based Benchmarks - Results are reported as True Acceptance Rate (TAR, %) at various False Acceptance Rate (FAR) thresholds. The best and second-best results are highlighted in bold and underlined, respectively. The proposed DQM-Face solution achieves state-of-the-art performance across different FARs on both, IJB-B and IJB-C.

	IJB-B			**IJB-C**		
Method	10^{-3}	10^{-4}	10^{-5}	10^{-3}	10^{-4}	10^{-5}
CosFace [43]	-	94.01	89.25	-	95.56	92.68
ArcFace [8]	-	94.09	88.50	-	95.74	92.69
PFE [37]	-	-	-	95.49	93.25	89.64
GroupFace [23]	-	94.93	-	-	96.26	-
CurricularFace [19]	-	94.80	-	-	96.10	-
EQFace [27]	96.48	94.51	89.62	97.39	95.84	93.01
MagFace [31]	96.19	94.50	90.36	97.24	95.97	94.08
SCF-ArcFace [26]	-	94.74	90.68	-	96.09	94.04
ElasticFace-Cos+ [6]	-	<u>95.30</u>	-	-	<u>96.57</u>	-
SphereFace-R v2 [28]	-	94.51	86.55	-	95.96	93.01
DDC [42]	-	94.70	-	-	96.10	-
QMagFace [39]	96.48	94.70	90.29	97.62	96.19	94.27
Face-TB [47]	-	94.37	-	-	95.72	-
CoReFace [38]	-	95.09	<u>91.33</u>	-	96.43	94.73
QCFace [9]	-	94.30	89.39	-	95.84	93.82
DQM-Face $\alpha_{0.4}$ (Ours)	**96.71**	**95.38**	**91.35**	**97.65**	**96.58**	**94.83**
DQM-Face $\alpha_{0.5}$ (Ours)	<u>96.65</u>	95.20	90.68	<u>97.64</u>	96.53	<u>94.78</u>

second-best performance across all evaluated FARs. We attribute the strong performance in higher FAR regions (10^{-3}) to the quality-guided learning, as face image quality factors have a pronounced impact in this regime [39]. Similarly, the high performance in lower FAR regions (10^{-5}) may be attributed to the dual-margin learning, as enhanced identity separability is expected to be beneficial when stricter discrimination is required.

5.3 Face Image Quality Assessment Performance

To measure the effectiveness of the proposed quality estimate utilized for the margin learning, the FIQA performance is investigated using EvD characteristics and pAUC values on four standard datasets, each posing unique challenges: LFW (unconstrained), Adience (age variations), CPLFW (pose variations), and XQLFW (image quality variations). Figure 2 illustrates the EvD characteristics, while Table 3 summarizes the pAUC values computed at a 30% discard rate. For a comprehensive evaluation, the proposed method is compared against nine state-of-the-art FIQA approaches.

The DQM-Face model demonstrates consistently strong performance, ranking among the top-performing methods across all datasets. It achieves the best

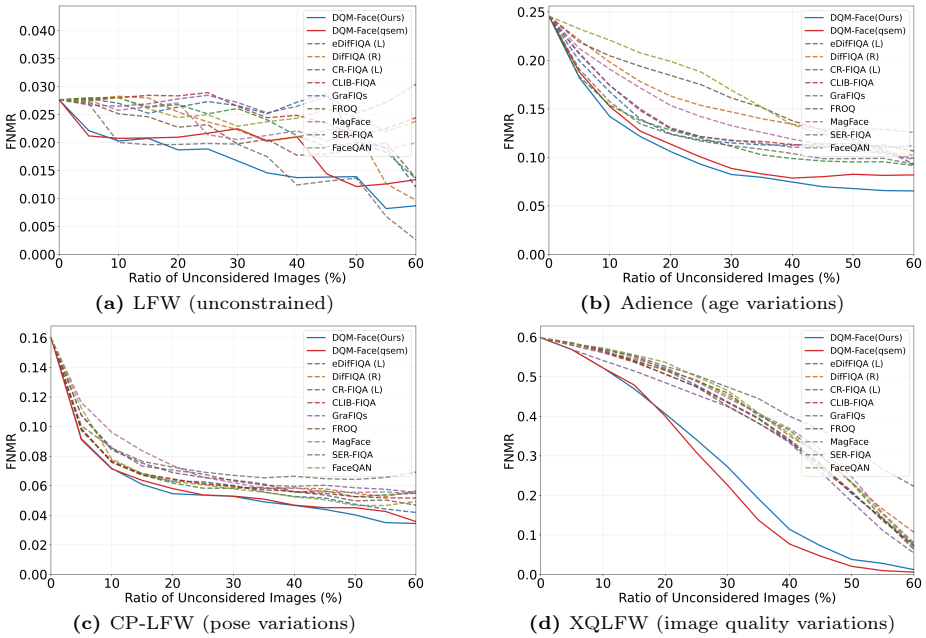
Table 3: Face Image Quality Assessment Evaluation - is evaluated using pAUC, where lower values indicate better performance. The best result is highlighted in bold, and the second-best is underlined. Evaluations are conducted on widely used FIQA datasets. The proposed approach is compared against nine state-of-the-art methods using the DQM-Face embeddings, demonstrating that its predicted quality scores effectively capture its underlying sample utility.

Method	LFW	Adience	CPLFW	XQLFW
DQM-Face (Ours)	**0.0061**	**0.0404**	**0.0220**	<u>0.1373</u>
DQM-Face (qsem)	<u>0.0065</u>	<u>0.0424</u>	<u>0.0222</u>	**0.1346**
eDifFIQA (L) [5]	0.0080	0.0465	0.0239	0.1603
DifFIQA (R) [4]	0.0079	0.0555	0.0239	0.1612
CR-FIQA (L) [7]	0.0081	0.0451	0.0235	0.1622
CLIB-FIQA [33]	0.0084	0.0481	0.0239	0.1594
GraFIQs [25]	0.0082	0.0483	0.0258	0.1538
FROQ [2]	0.0073	0.0590	0.0257	0.1586
MagFace [31]	0.0076	0.0530	0.0273	0.1615
SER-FIQA [40]	0.0065	0.0451	0.0259	0.1638
FaceQAN [3]	0.0078	0.0628	0.0246	0.1641

results on LFW, Adience, and CPLFW, and remains highly competitive on XQLFW. To further analyze the role of semantic quality, we introduce the DQM-Face (qsem) variant, trained with $\alpha = 1$ to capture only the semantic quality component q_{sem}. This variant achieves the highest performance on XQLFW and maintains competitive results on the remaining datasets, indicating that the semantic quality contributes to robustness under severe quality variations. Although the FIQA performance is strong, the quality learning mechanism is deeply integrated into the face recognition framework to capture subtle, identity-preserving variations. Consequently, the learned quality is inherently model-specific, relying on architecture-dependent cues rather than aiming for universal generalization, yet it still achieves strong overall performance.

5.4 Visual Quality Analysis

To provide qualitative insight into the different quality branches, we employ Grad-CAM [35] to visualize the spatial regions that contribute most to the recognition decision. Figure 3 compares Grad-CAM heatmaps for the three model variants: q_{mag} (magnitude-only, $\alpha = 0$), q_{sem} (semantic-only, $\alpha = 1$), and the proposed DQM-Face (fused quality, $\alpha = 0.5$). The heatmaps are overlaid on the input face images to illustrate the regions emphasized by each model. The q_{mag} variant exhibits attention that is occasionally dispersed toward non-discriminative regions, such as background textures or facial boundaries, particularly when feature magnitudes are affected by blur, occlusion, or low image quality. This observation supports the hypothesis that feature magnitude alone cannot reliably represent identity quality. In contrast, the q_{sem} variant



(a) LFW (unconstrained) **(b)** Adience (age variations)

(c) CP-LFW (pose variations) **(d)** XQLFW (image quality variations)

Fig. 2: Face Image Quality Assessment Performance - The FIQA performance is evaluated using Error-vs-Discard curves on four popular FIQA datasets with specific challenges. The proposed method is compared against nine state-of-the-art approaches on the DQM-Face embeddings, demonstrating that its predicted quality scores effectively reflect the underlying sample utility for the DQM-Face recognition model.

attends to identity-discriminative facial regions more consistently, including the eyes, nose, and mouth.. This behavior demonstrates that the semantic quality branch successfully learns identity-aware quality representations that are more closely aligned with the face recognition objective. The proposed DQM-Face model ($\alpha = 0.5$) combines the strengths of both quality cues. Its attention maps exhibit well-localized and stable activation over the most informative facial regions while remaining robust to challenging imaging conditions. Compared with either individual branch, the fused model produces more focused and semantically meaningful attention, indicating that magnitude-based and semantic quality provide complementary information.

5.5 Complexity Analysis

The complexity of DQM-Face is comparable to similar FR solutions, such as ArcFace. The SE branch adds +0.099M parameters (+0.15%) and less than 0.0001 GFLOPs on the backbone (of 12.149 GFLOPs, 65.156M parameters). This overhead is negligible compared to ArcFace [8] ($< 0.2\%$). The inference time is dominated by the used backbone, as m_2 adds no cost and quality estimation is based directly on the extracted features.

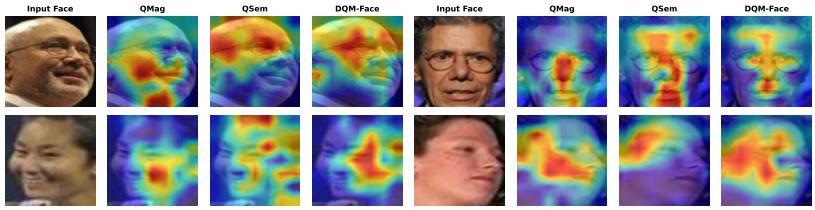


Fig. 3: Component-wise Visual Attribution Analysis - Comparison of Grad-CAM heatmaps for the three model variants: q_{mag} (magnitude-only, $\alpha = 0$), q_{sem} (semantic-only, $\alpha = 1$), and DQM-Face (fused quality, $\alpha = 0.5$). The heatmaps highlight the facial regions that contribute most to identity recognition.

Table 4: Ablation study on quality margin – The effect of the quality margin m_1 through the variations of the fusion weight α is studied. $\alpha = 0$ corresponds to magnitude-based quality only (q_{mag}), while $\alpha = 1$ corresponds to semantic quality only (q_{sem}).

	Verification Accuracy (%)					IJB-B (TAR@FAR)			IJB-C (TAR@FAR)		
α	LFW	CFP-FP	AgeDB	CPLFW	Avg.	10^{-3}	10^{-4}	10^{-5}	10^{-3}	10^{-4}	10^{-5}
0.0	99.81	98.41	98.31	93.00	97.38	96.49	95.01	90.58	97.61	96.40	94.45
0.1	99.83	98.24	98.20	93.08	97.34	96.56	94.99	91.19	97.54	96.31	94.64
0.2	99.83	98.50	98.27	93.02	97.41	**96.77**	95.25	90.97	97.63	96.43	94.49
0.3	99.83	98.44	98.20	93.05	97.38	96.48	94.80	89.70	97.43	96.20	94.34
0.4	99.83	98.61	98.32	93.22	97.50	96.71	**95.38**	**91.35**	**97.65**	**96.58**	**94.83**
0.5	**99.83**	**98.83**	**98.35**	**93.30**	**97.58**	96.65	95.20	90.68	97.64	96.53	94.78
0.6	99.82	98.67	98.25	93.02	97.44	96.57	94.82	89.95	97.55	96.21	94.14
0.7	99.80	98.57	98.25	93.20	97.46	96.67	94.97	89.28	97.56	96.32	94.33
0.8	99.82	98.34	98.20	93.10	97.37	96.54	94.81	90.46	97.55	96.23	94.22
0.9	99.83	98.36	98.28	93.07	97.39	96.48	94.95	89.20	97.52	96.38	94.24
1.0	99.82	98.69	98.22	93.17	97.48	96.67	95.22	90.55	97.62	96.51	94.67

5.6 Ablation Studies

To validate the choice of hyperparameters, we conducted an ablation study investigating different configurations. Since the face recognition performance in our approach depends on the quality margin $m_1(q(\alpha))$, which is influenced by the hyperparameter and the quality fusion weight α, we analyzed how varying α affects recognition accuracy. Table 4 summarizes the results for α values across the full range $[0, 1]$. The model achieves its highest accuracy with $\alpha = 0.5$, while performance gradually decreases toward the extreme values.

Based on these observations, we adopt $\alpha = 0.5$ as the default setting for the proposed method. Overall, these findings support our hypothesis that fusing complementary quality cues yields a more robust assessment of sample utility than relying on either component in isolation.

Table 5 analyzes the effect of the repulsion margin hyperparameter m_2. For training DQM-Face, a curriculum-based scheduling strategy for m_2 was introduced in Section 3.3. To evaluate its impact, Table 5 reports the performance obtained with different fixed margin values, ranging from $m_2 = 0$ (no inter-

class repulsion) to $m_2 = 0.2$ (strong fixed repulsion). We additionally compare against linear $m_2(t) = 0.15 \cdot t/T$ and cosine $m_2(t) = 0.15 \cdot (1 - \cos(\pi t/T))/2$ schedules. While both achieve competitive performance, they underperform the staged schedule. The staged design maintains constant m_2 during each phase, providing stable gradients and better alignment with curriculum learning. The results indicate that the proposed curriculum-based schedule yields the best performance, as it enables the model to first form compact identity clusters and then progressively enforce stronger inter-class separation. This gradual adaptation stabilizes the gradient, and thus its training process, and promotes higher global separability among the face representations.

Table 5: Ablation study on the repulsion margin - The effect of the repulsion m_2, and specifically its scheduled learning procedure, is studied.

	Verification Accuracy (%)					IJB-B (TAR@FAR)			IJB-C (TAR@FAR)		
m_2	LFW	CFP-FP	AgeDB	CPLFW	Avg.	10^{-3}	10^{-4}	10^{-5}	10^{-3}	10^{-4}	10^{-5}
0	99.83	98.43	98.13	92.92	97.32	96.63	94.87	87.21	97.58	96.19	93.75
0.05	99.81	98.64	98.18	93.26	97.47	96.54	94.94	89.85	97.57	96.32	94.31
0.10	99.80	98.71	98.13	93.23	97.46	**96.78**	95.05	90.04	97.64	96.43	94.69
0.15	99.83	98.57	98.22	93.20	97.46	96.67	95.08	88.80	**97.66**	96.52	94.53
0.20	99.33	85.96	95.05	85.63	91.49	91.60	86.91	76.69	93.38	89.28	83.37
Linear	99.83	98.35	98.28	93.06	97.38	96.77	**95.25**	**90.97**	97.63	96.43	94.49
Cosine	99.83	98.50	98.26	93.01	97.40	96.48	94.95	89.20	97.52	96.38	94.24
Staged	**99.83**	**98.83**	**98.35**	**93.30**	**97.58**	96.65	95.20	90.68	97.64	**96.53**	**94.78**

6 Conclusion

In this work, we introduced DQM-Face, a Dual Quality Margin Learning framework designed to address the challenges of face recognition in unconstrained environments. By jointly modeling magnitude-based and semantic quality cues through a unified attraction–repulsion margin optimization, the proposed approach enables the learning of more discriminative and structured feature representations. The quality derived from margin optimization achieves high FIQA performance within the proposed framework, demonstrating its intrinsic alignment with the underlying recognition objective. Comprehensive evaluations across diverse image-to-image and video benchmarks demonstrate that DQM-Face consistently achieves the best or second-best face recognition performance compared to state-of-the-art methods. These results confirm the effectiveness of integrating semantic quality modeling into margin-based learning.

Acknowledgement. This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Grant 544631027.

References

1. Al-Refai, R., Cabarcos, P.A., Terhörst, P.: A comprehensive re-evaluation of biometric system properties in the modern era. In: IEEE International Joint Conference on Biometrics, IJCB 2025, Osaka, Japan, September 8-11, 2025. pp. 1–11. IEEE (2025)
2. Babnik, Ž., Jain, D.K., Peer, P., Štruc, V.: Froq: Observing face recognition models for efficient quality assessment. arXiv preprint arXiv:2509.17689 (2025), https://arxiv.org/abs/2509.17689
3. Babnik, Ž., Peer, P., Štruc, V.: Faceqan: Face image quality assessment through adversarial noise exploration. In: 2022 26th International Conference on Pattern Recognition (ICPR). pp. 748–754. IEEE (2022)
4. Babnik, Ž., Peer, P., Štruc, V.: Diffiqa: Face image quality assessment using denoising diffusion probabilistic models. In: 2023 IEEE International Joint Conference on Biometrics (IJCB). pp. 1–10. IEEE (2023)
5. Babnik, Ž., Peer, P., Štruc, V.: ediffiqa: Towards efficient face image quality assessment based on denoising diffusion probabilistic models. IEEE Transactions on Biometrics, Behavior, and Identity Science **6**(4), 458–474 (2024). https://doi.org/10.1109/TBIOM.2024.3376236
6. Boutros, F., Damer, N., Kirchbuchner, F., Kuijper, A.: Elasticface: Elastic margin loss for deep face recognition. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 1578–1587. IEEE (2022)
7. Boutros, F., Fang, M., Klemt, M., Fu, B., Damer, N.: Cr-fiqa: Face image quality assessment by learning sample relative classifiability. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 5836–5845. IEEE (2023)
8. Deng, J., Guo, J., Xue, N., Zafeiriou, S.: Arcface: Additive angular margin loss for deep face recognition. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 4690–4699. IEEE (2019)
9. Doan-Ngo, D.P., Diep, T.D., Nguyen-Duc, T., LE, T.S., Thoai, N.: Qcface: Image quality control for boosting face representation and recognition. arXiv preprint arXiv:2510.15289 (2025), https://arxiv.org/abs/2510.15289
10. Duta, I.C., Liu, L., Zhu, F., Shao, L.: Improved residual networks for image and video recognition. In: Proceedings of the 25th International Conference on Pattern Recognition (ICPR). pp. 9415–9422. IEEE (2020)
11. Eidinger, E., Enbar, R., Hassner, T.: Age and gender estimation of unfiltered faces. IEEE Transactions on Information Forensics and Security **9**(12), 2170–2179 (2014). https://doi.org/10.1109/TIFS.2014.2359646
12. Fallahi, M., Ramesh, R., Ramasamy, P.P., Cabarcos, P.A., Strufe, T., Terhörst, P.: On the reliability of biometric datasets: How much test data ensures reliability? arXiv preprint arXiv:2501.06504 (2025), https://arxiv.org/abs/2501.06504
13. Frontex: Best practice technical guidelines for automated border control (abc) systems. Tech. rep., European Agency for the Management of Operational Cooperation at the External Borders of the Member States of the European Union (2015), version 4
14. Guo, J., Deng, J.: Insightface: 2d and 3d face analysis library. https://github.com/deepinsight/insightface (2019)
15. Guo, Y., Zhang, L., Hu, Y., He, X., Gao, J.: Ms-celeb-1m: A dataset and benchmark for large-scale face recognition. In: European Conference on Computer Vision. pp. 87–102. Springer (2016)

16. Hendrycks, D., Gimpel, K.: Gaussian error linear units (gelus). arXiv preprint arXiv:1606.08415 (2016), https://arxiv.org/abs/1606.08415
17. Hu, J., Shen, L., Sun, G.: Squeeze-and-excitation networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 7132–7141. IEEE (2018)
18. Huang, G.B., Ramesh, M., Mattar, T., Berg, T., Learned-Miller, E.: Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Tech. Rep. 07-49, University of Massachusetts, Amherst (2007)
19. Huang, Y., Wang, Y., Tai, Y., Liu, X., Shen, P., Li, S., Li, J., Huang, F.: Curricularface: Adaptive curriculum learning loss for deep face recognition. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 5901–5910. IEEE (2020)
20. International Organization for Standardization: ISO/IEC 19795-1:2006 – Information Technology — Biometric Performance Testing and Reporting (2006), international Standard
21. International Organization for Standardization: ISO/IEC 29794-5:2025 – Face Image Quality. https://www.iso.org/standard/92541.html (2025), international Standard
22. Kim, M., Jain, A.K., Liu, X.: Adaface: Quality adaptive margin for face recognition. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 18750–18759. IEEE (2022)
23. Kim, Y., Park, W., Roh, M.C., Shin, J.: Groupface: Learning latent groups and constructing group-based representations for face recognition. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 5621–5630. IEEE (2020)
24. Knoche, M., Hormann, S., Rigoll, G.: Cross-quality lfw: A database for analyzing cross-resolution image face recognition in unconstrained environments. In: 2021 16th IEEE International Conference on Automatic Face and Gesture Recognition (FG). pp. 1–5. IEEE (2021). https://doi.org/10.1109/FG52635.2021.9666960
25. Kolf, J.N., Damer, N., Boutros, F.: Grafiqs: Face image quality assessment using gradient magnitudes. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 1490–1499. IEEE (2024)
26. Li, S., Xu, J., Xu, X., Shen, P., Li, S., Hooi, B.: Spherical confidence learning for face recognition. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 15629–15637. IEEE (2021)
27. Liu, R., Tan, W.: Eqface: A simple explicit quality network for face recognition. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 1482–1490. IEEE (2021)
28. Liu, W., Wen, Y., Raj, B., Singh, R., Weller, A.: Sphereface revived: Unifying hyperspherical face recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence **45**(2), 2458–2474 (2022)
29. Liu, W., Wen, Y., Yu, Z., Li, M., Raj, B., Song, L.: Sphereface: Deep hypersphere embedding for face recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 212–220. IEEE (2017)
30. Maze, B., Adams, J., Duncan, J.A., Kalka, N., Miller, C., Otto, A., Niggel, W., Anderson, J., Cheney, J., et al.: Iarpa janus benchmark-c: Face dataset and protocol. In: International Conference on Biometrics (ICB). pp. 158–165. IEEE (2018)
31. Meng, Q., Zhao, S., Huang, Z., Zhou, F.: Magface: A universal representation for face recognition and quality assessment. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 14225–14234. IEEE (2021)

32. Moschoglou, S., Papaioannou, A., Sagonas, C., Deng, J., Kotsia, I., Zafeiriou, S.: Agedb: The first manually collected, in-the-wild age database. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. pp. 51–59. IEEE (2017)
33. Ou, F.Z., Li, C., Wang, S., Kwong, S.: Clib-fiqa: Face image quality assessment with confidence calibration. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 1694–1704. IEEE (June 2024)
34. Schlett, T., Rathgeb, C., Tapia, J.E., Busch, C.: Considerations on the evaluation of biometric quality assessment algorithms. IEEE Trans. Biom. Behav. Identity Sci. **6**(1), 54–67 (2024). https://doi.org/10.1109/TBIOM.2023.3336513, https://doi.org/10.1109/TBIOM.2023.3336513
35. Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., Batra, D.: Grad-cam: Visual explanations from deep networks via gradient-based localization. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 618–626. IEEE (2017)
36. Sengupta, S., Chen, J.C., Castillo, C., Patel, V.M., Chellappa, R., Jacobs, D.W.: Frontal to profile face verification in the wild. In: IEEE Winter Conference on Applications of Computer Vision (WACV). pp. 1–9. IEEE (2016)
37. Shi, Y., Jain, A.K.: Probabilistic face embeddings. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 6902–6911. IEEE (2019)
38. Song, Y., Wang, F.: Coreface: Sample-guided contrastive regularization for deep face recognition. Pattern Recognition **152**, 110483 (2024)
39. Terhorst, P., Ihlefeld, M., Huber, M., Damer, N., Kirchbuchner, F., Raja, K., Kuijper, A.: Qmagface: Simple and accurate quality-aware face recognition. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. pp. 3484–3494. IEEE (2023)
40. Terhorst, P., Kolf, J.N., Damer, N., Kirchbuchner, F., Kuijper, A.: Ser-fiq: Unsupervised estimation of face image quality based on stochastic embedding robustness. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 5650–5659. IEEE (2020)
41. Terhörst, P., Kolf, J.N., Damer, N., Kirchbuchner, F., Kuijper, A.: Face quality estimation and its correlation to demographic and non-demographic bias in face recognition. In: 2020 IEEE International Joint Conference on Biometrics, IJCB 2020, Houston, TX, USA, September 28 - October 1, 2020. pp. 1–11. IEEE (2020). https://doi.org/10.1109/IJCB48548.2020.9304865, https://doi.org/10.1109/IJCB48548.2020.9304865
42. Uzun, B., Cevikalp, H., Saribas, H.: Deep discriminative feature models (ddfms) for set based face recognition and distance metric learning. IEEE Transactions on Pattern Analysis and Machine Intelligence **45**(5), 5594–5608 (2022)
43. Wang, H., Wang, Y., Zhou, Z., Ji, X., Gong, D., Zhou, J., Li, Z., Liu, W.: Cosface: Large margin cosine loss for deep face recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 5265–5274. IEEE (2018)
44. Wang, M., Deng, W.: Deep face recognition: A survey. Neurocomputing **429**, 215–244 (2021). https://doi.org/10.1016/J.NEUCOM.2020.10.081, https://doi.org/10.1016/j.neucom.2020.10.081
45. Whitelam, C., Taborsky, E., Blanton, A., Maze, B., Adams, J., Miller, N., Kalka, N., Jain, A.K., Duncan, J.A., et al.: Iarpa janus benchmark-b face dataset. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. pp. 90–98. IEEE (2017)

46. Zheng, T., Deng, W.: Cross-pose lfw: A database for studying cross-pose face recognition in unconstrained environments. Beijing University of Posts and Telecommunications, Tech. Rep **5** (2018)
47. Zhu, Y., Ren, M., Jing, H., Dai, L., Sun, Z., Li, P.: Joint holistic and masked face recognition. IEEE Transactions on Information Forensics and Security **18**, 3388–3400 (2023)