Facial Image Quality Assessment for Improved Face Recognition Systems: A Survey

Leonardo Rodovero[1]

Abstract: This survey reviews the importance of Face Image Quality Assessment in detecting and discarding low-quality images for improving face recognition systems. It includes the aspects of FIQA related to database construction, pixel-oriented approaches, benchmarking against state-of-the-art techniques, and alternative metrics such as Interpretable Face Quality Assessment. Relevant improvements are noticed in the development of pixel-based methods for exhaustive quality estimates and the generation of large FIQA databases. Further research in this regard is required to improve dataset diversity, model interpretability, and face recognition system robustness.

Keywords: Face Image Quality Assessment (FIQA), Facial Recognition (FR), Pixel-based methods, Database exploration, Interpretable Face Quality Assessment (IFQA)

1 Introduction

Face image quality assessment, or FIQA, is an essential step in achieving maximum performance in facial recognition systems. By detecting low-quality images, one can prevent their use. In FR, FIQA can considerably increase the accuracy and reliability of the results, a critical requirement, for example, for security-sensitive applications such as surveillance and biometric authentication [Te20b]. Recently, FIQA moved from simple scalar quality scores to more fine-grained pixel-level quality analytics [Su24]. The former advanced change provides a detailed understanding of the individual image regions on overall quality and hence makes way for developing superior FIQA methods that can cater to the challenges brought forward by real-world image variances.

This paper reviews different FIQA techniques, such as data-driven, pixel-based, state-of-the-art, and metrics-driven techniques. Notably, the major contributions noted include Liu's [Li24] exhaustive FIQA database and Su's [Su24] versatile IQA dataset (Section 3.1) that has formed a strong basis for developing and validating novel solutions.

Some of the pixel-based, more fine-grained techniques, e.g. works by Biagi [Bi23] and Huber [Hu22], emphasize that it is important to reason about how much individual regions of the image contribute towards the overall recognition performance (Section 3.2). The comparison with state-of-the-art methods like FaceQNet [He19], SDD-FIQA [Ou21], and MagFace [Me21] demonstrates the efficacy of modern FIQA approaches. Additionally, the

[1] Human-Centered-AI, Denmark Technical University, s240095@dtu.dk

survey has also opposed FIQA to a more recent and more precise human-aligned metric: Interpretable Face Quality Assessment [B2], providing more accurate human-aligned quality estimates by focusing on salient facial regions (Section 3.3). This in-depth study underlines the evolution of the field of FIQA and further confirms that it is necessary to have new innovative methods to achieve better explainability and robustness for FR systems.

2 Backgroud

Face Image Quality Assessment is the process for quality assessment of face images to determine their utility for face recognition systems [Sc22]. The methods of FIQA are important for filtering low-quality images that can degrade the performance of FR systems applied in various applications, such as surveillance and biometric authentication [Jo23]. Traditional FIQA methods often use scalar quality scores that describe the whole quality of an image; recent developments, however, have introduced pixel-level face image quality assessment (PLFIQA) methods to provide more detailed and interpretable quality maps [Hu22]. In this way, with pixel-level assessments, one can know which parts contribute to the overall quality of an image and thus offer actionable insights on how to improve image acquisition processes [Hu22].

'Error versus Discard Characteristic' [Sc23] plots are used to measure the effectiveness of these methods of FIQA. Using this approach, in EDC plots, changes in some biometric recognition error—for example, the False Non-Match Rate—can be depicted when their images are progressively discarded based on their quality score [Sc23]. Indeed, one often computes the area under these curves so called "partial Area Under Curve"[Sc23], as a means to quantitatively compare the performance of different algorithms for FIQA. The obtained pAUC values quantify how good a method of FIQA is at improving recognition performance by discarding low-quality images; the higher the pAUC, the better the performance [Sc23]. The evaluation framework is very important in developing new and benchmarking FIQA methods so that they guarantee to enhance the robustness and accuracy of FR systems.

2.1 State-of-the-art methods

In Section 3.4, it is mentioned how the state-of-the-art methods in FIQA are enhanced by more recent techniques developed in [Ba24] and [BD3]. It could be useful to summarize the principle state-of-art to understand the following sections, see Table 1. An approach mentioned in [BD3] is to categorize them into three different types of methods (see Figure 1): Analytical FIQA, in which the techniques are mostly unsupervised [Te20b], [BP2], Regression-based FIQA, in which the techniques are supervised [He19], [Ou21], [BP3] and model-based FIQA which techniques are quality aware [Me21], [Bo23]. Modern FIQA techniques assign a single, unifying numerical score that reflects the biometric quality of a face image to determine its utility for FR. Greater compatibility for FR is implied by higher scores, which often suggest higher quality images [DKG16]. Supervised methods make use of labeled data to project image quality from features relative to class representations, such as FaceQNet [He19], while unsupervised methods are reliant on the inherent properties of images and how they relate to a model of recognition without the use of labeled data, such as used in SDD-FIQA [Ou21]. Lastly, quality-aware methods, including MagFace [Me21], integrate quality assessment directly into the recognition process, using feature magnitudes or robustness as quality indicators.

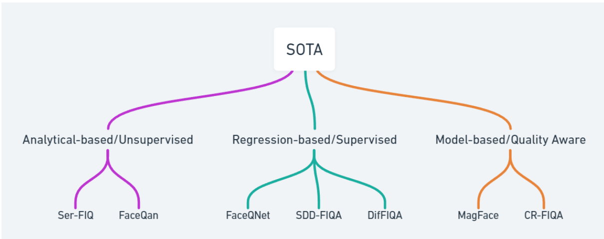

Fig. 1: Taxonomy based on state-of-the-art methods in FIQA

Method	**Description**	**Reference**
DifFIQA	Utilizes denoising diffusion probabilistic models to assess face image quality by analyzing embedding stability under perturbations.	[BP3]
MagFace	Learns is a universal feature embedding whose magnitude can measure the quality of the face.	[Me21]
SDD-FIQA	An unsupervised method that uses Similarity Distribution Distance to assess face image quality by comparing intra-class and inter-class similarity distributions.	[Ou21]
FaceQAN	Employs adversarial noise exploration to assess face image quality, linking image quality to adversarial attacks.	[BP2]
CR-FIQA	Predicts face image quality by learning sample relative classification based on the allocation of the training sample feature representation in angular space.	[Bo23]
SER-FIQ	Uses stochastic embedding robustness for FIQA, measuring the robustness of face representations against dropout variations.	[Te20b]

Tab. 1: Comparison of SOTA FIQA Methods

2.2 Bias in FIQA

The development of accurate FIQA methods has advanced significantly, but bias in these algorithms is a critical issue that is receiving more and more attention. The significance of evaluating bias in facial image quality assessment (FIQA) techniques is emphasized in [B2]. This is especially important because face recognition systems have been demonstrated to display demographic bias. Any biases in the underlying face recognition system may spread to, or even be enhanced by, the FIQA process as its goal is to predict an image's suitability for face recognition. This raises questions about how well FIQA systems work fairly and equally for various demographic groups.

Biases abound in the face image quality assessment (FIQA) sector, especially about sex and race [B2]. The strong link between how attractive a face is rated and how well it is recognized makes these biases even worse. [Te20a]. These biases also affect the state-of-the-art FIQA techniques, including FaceQNet [He19], MagFace [Me21], and SER-FIQ [Te20b]. Despite that, there are effective ways to lessen these biases. For example, Chen's L2RT-FIQA method [CY23] improves consistency, decreasing biases in FR. In addition, Ou [Ou23] offers a brand-new methodology to mitigate ethnic quality bias-related generalization issues: the Ethnic-Quality-Bias Mitigating (EQBM) framework. Maragkoudakis [Ma24] propose methods that consider visual quality and the mismatch between training and test face images, and sampling strategies to mitigate bias in face synthesis methods, respectively. Despite these advancements, there is still a need for more deep learning approaches in FIQA [Sc22]. This survey will help to visualize and get an easier perspective on how they can be improved by different methods.

3 Methods

Many techniques have been developed in recent years to evaluate the quality of facial images; in particular, deep learning-based techniques are becoming more and more popular [Sc22]. These techniques usually produce an image's unique quality score, which attempts to express the image's overall effectiveness for facial recognition tasks. FIQA techniques that provide information on how quality scores are calculated or include features in their models are of special relevance. Examples of such techniques include: Fu-Zhao [FD22] presented tools based on the behavior of face recognition models while an unsupervised FIQA technique utilizing 'Similarity Distribution Distance' was presented by Ou [Ou21]. Forensic facial recognition on a broad scale: Rodriguez [Ro23] introduced multi-task explainable quality networks.

Other important contributions include Boutros' method [Bo23] for evaluating facial image quality using sample relative classification and Kim [Ki24]'s work on addressing erroneous pseudo-labels by estimating sample class variance. Huber [Hu22] concentrated on applying deletion and insertion evaluation curves to assess pixel-level FIQA. Lastly, Babnik [BP3] introduced a potent FIQA method based on probabilistic diffusion denoising models.

To better structure this survey of FIQA systems, we first investigate database augmentation and optimization [Su24] [Li24]. Second, we explore pixel-based techniques and investigate how they might be used to fine-tune algorithms through pixel-by-pixel analysis [Hu22] [Te23] [Jo23]. Lastly, we look into how current state-of-the-art approaches can be improved by utilizing artificial intelligence and optimization techniques [Ba24] [BD3]. On the other hand, it has been analyzed another kind of metric that can be used for FR systems and it works very efficiently. The main reason is that traditional face image quality assessment metrics often fall short in accurately evaluating perceptually significant facial features crucial for human cognition. Interpretable Face Quality evaluation [Jo23] addresses this limitation by employing a pixel-level analysis focused on salient facial regions, aligning with human visual perception. For major clarification, the follow visualization Figure 2 summarizes the approaches described:

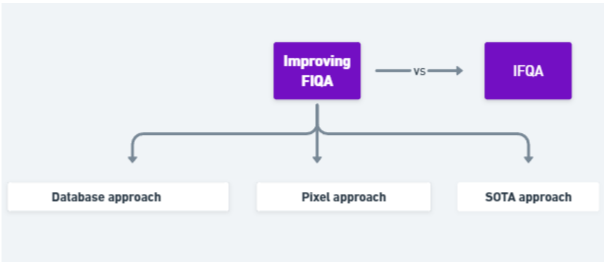

Fig. 2: Categorization of FIQA methods and illustration of another similar metric IFQA [Jo23]

3.1 Database approach

By offering extensive databases and cutting-edge techniques to enhance Face Image Quality Assessment (FIQA), Liu [Li24] and Su [Su24] both make substantial contributions to the discipline. A comprehensive FIQA database with 42,125 face photos with varying degrees of content and distortion is created by Liu [Li24]. The quantity and variety of facial components and deformities in this database make it noteworthy. Additionally, a brand-new Transformer-based technique called TransFQA is presented in the research to evaluate the quality of face photos. This approach makes use of a face component (FC)-guided Transformer network (FT-Net), which combines a novel progressive attention mechanism to integrate global context, face region, and detailed FC data. Furthermore, to reliably anticipate final quality scores and weight various distortions, a distortion-specific prediction network (DP-Net) is created.

Su [Su24], on the other hand, introduces the IQA database which includes 20,000 human faces. This database includes a diverse range of individuals in varied circumstances. Su develops a novel deep learning model that leverages generative priors for IQA (see Figure 3) [Su24] , utilizing rich statistics encoded in well-pretrained off-the-shelf generative models. By extracting and utilizing generative statistics guided by style codes, the model compares input images to latent references derived from a pristine image manifold within StyleGAN2 [Ka20]. This integration of generative priors has been very helpful in face IQA, given that the model performed well on the GFIQA-20K dataset.

Despite their great contributions toward FIQA, Su and Liu's work has many limitations. Small sample-size subjective biases in quality assessment (see Subsection 2.2), and no comparative exercise with current IQA datasets may be some of the major challenges against Su's study. Such high reliance on artificial distortions by Liu may not appropriately

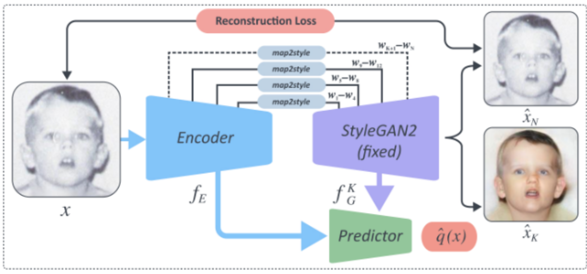

Fig. 3: The proposed face IQA model utilizing generative priors, based on [Su24]

represent real-world complexities and may also not be properly accounted for concerning the subjective nature of quality perception with the selected evaluation criteria. Real-world performance will also be more proper in evaluating the developed TransFQA approach's robustness and application.

3.2 Pixel-based methods approach

Traditional Face Image Quality Assessment has been performed by assigning a single quality score to the whole image, which can hide the influence of some facial features or regions on the final assessment. Research to improve FIQA has focused on pixel-based techniques, whose technologies are capable of evaluating face image quality in detail, imparting a completely different way of understanding the influence on recognition systems. Biagi underlines the role of understandability in FIQA and proposes a SHAP-based methodology [Bi23], while Huber introduces strategies for the evaluation of attributes at the pixel level [Hu22]. Lian [Li22] and Terhorst [Te23] separately proposed training-free approaches to estimating pixel-level attributes. The former based his approach on model-specific quality values, and the latter suggested an unsupervised approach for local contribution variance.

A very evident trend of recent research in Face Image Quality Assessment is toward pixel-level assessments to increase the effectiveness of face recognition systems. Whereas traditional approaches to FIQA have returned just one quality score for the whole image, these approaches enable analysis at the pixel level to derive granular insights into the factors impacting the accuracy of recognition, as done by Biagi, Huber, and Terhorst. This is evident in the development of methods that identify those parts of an image impacting the effectiveness of recognition, and the introduction of new schemes for evaluation, like deletion and insertion evaluation curves, for quantifying the impact of pixel-level quality assessment [Te23]. Ex batches use explainability techniques such as the Shapley Values ([Bi23], which further boosts this growing emphasis toward explainability in FIQA.

As shown by Terhorst's training-free approach and Biagi's use of Shapley Values [Wi02], pixel-level FIQA methods have significant advantages in terms of targeting image enhancement. However, they can also have drawbacks in terms of computational complexity or due to their reliance on some face recognition systems or model-specific data. While useful, Huber's [Hu22] evaluation criteria in assessing pixel-level FIQA also point to the challenge of generalizing results across a range of face recognition systems. Their unsupervised methodology in data-rich contexts might not be as accurate as supervised methods, though more adaptable.

3.3 IFQA vs FIQA

Quality evaluation of face images is critical in image processing, more so in facial identification applications. Traditional face image quality assessment measures often meet the unique requirements placed on facial images by frequently neglecting the correct features to be evaluated, which are usually deep and influential in human cognition, such as the mouth, nose, and eyes. Interpretable Face Quality Assessment (IFQA) [Jo23], a novel approach that emphasizes the specificity of evaluation in face regions, has been presented to address these constraints. A new approach called Interpretable Face Quality Assessment has been introduced, stressing evaluation specificity in face regions. In the same way, it has been analyzed a general comparison between IFQA and FIQA Table 2 and IFQA into one with a detailed face-specific metric, representing a significant advancement in face quality assessment. Ignoring traditional methods and accentuating perceptually salient facial regions at a pixel level is what makes the IFQA approach very close to human visual understanding, hence yielding more accurate assessments. That is to say, the new scheme moves ahead in performance and relevance toward facial images in an adversarial framework with a per-pixel discriminator over existing metrics. Hence, it has especially significant value in areas of applications where the quality of facial images is critical, such as security and identification systems.

3.4 Enhancing SOTA approaches in FIQA

Although state-of-the-art FIQA algorithms have taken dramatic steps toward the evaluation of face image quality for the success of face recognition, as established from comprehensive surveys like that by Schlett [Sc22], ongoing research is still increasing the capacity in this area. This section provides a general overview of new developments and state-of-the-art approaches conceived to increase the functionality and robustness of FIQA methods.

Among the optimization-based approaches is that of Babnik [BD3], where they increase the predictive power for FIQA through the incorporation of model-specific quality information to enhance the ranking accuracy, more on face images of similar quality. In this case, it outperforms the existing methods of FIQA in various models of recognition. Additionally,

Feature	IFQA [Jo23]	Traditional FIQA
Focus	Specifically focuses on facial regions, enhancing the interpretability of quality assessments.	General image quality metrics without a specific focus on facial features.
Methodology	Uses an adversarial framework with a per-pixel discriminator for detailed, face-centric evaluation.	Relies on general assessment metrics that may not consider detailed facial characteristics.
Scalability	Designed to be more scalable than human studies, which are traditionally used for face assessments.	Typically scalable but may lack the precision in facial assessments.
Primary Region Emphasis	Emphasizes primary facial regions (eyes, nose, mouth) due to their impact on human cognition.	Does not typically emphasize specific facial regions, leading to less targeted assessments.
Improvement	Can lead to improvement when used as an objective function in face restoration models.	General metrics may not directly contribute to improvements in facial image processing tasks.

Tab. 2: Comparison of IFQA and Traditional FIQA

Babnik [Ba24] addresses alignment sensitivity through an 'Alignment Invariant Knowledge Distillation' process, which utilizes knowledge from different alignment procedures to enhance the robustness of FIQA against alignment variations. addresses the sensitivity to alignment by proposing an 'Alignment Invariant Knowledge Distillation' process that involves using knowledge from the different procedures of alignment to increase the robustness of FIQA to changes in alignment. This shows significant gains in performance with both misaligned and properly aligned face images and outperforms six state-of-the-art [BD3] FIQA techniques on multiple datasets and recognition models. Traditional methods, such as SER-FIQ [Te20b] and FaceQnet [He19], have been focused on unsupervised or supervised quality estimation and hence usually disregard the sensitivity to alignment. The new approaches by Babnik provide major steps ahead in introducing into the model certain biases and misalignment problems, therefore going much further in making techniques related to FIQA both reliable and accurate in many scenarios.

The optimization method proposed in Babnik [BD3] aims to improve the performance of Face Image Quality Assessment (FIQA) techniques by incorporating model-specific quality information into the quality scores produced by existing FIQA methods. This approach was evaluated using six state-of-the-art FIQA techniques (see Section 2.1) and showed significant improvements in most cases, particularly evaluated on five commonly used benchmark databases: XQLFW [KHR21], CPLFW [ZD18], CFP-FP [Se16], CALFW [ZDH17] and LFW [Hu08]. Table 3 below compares the results of the Area Under the Error-versus-Reject Characteristic Curves (ERC) obtained by the optimization method with those obtained by the baseline FIQA methods. The AUC values are multiplied by 10^3 for readability, and lower values indicate better performance. The column labeled 'Improvement' in the table tells whether the optimization method normally fared better compared with the baseline for every method of FIQA over all tested datasets. Results show that except for the mixed results from CR-FIQA, the optimization approach has generally improved the performance of the techniques of FIQA.

FIQA Method	Improvement
CR-FIQA	Mixed
FaceQAN	Improved
MagFace	Improved
SER-FIQ	Mixed
SDD-FIQA	Improved
PCNet	Improved

Tab. 3: Results after applied Optimization methods, based on [BD3]

Finally, the authors have proved in the paper that with the AI-KD technique, there is substantial improvement in the performance of existing state-of-the-art FIQA methods across various benchmarks and recognition models, as illustrated in Table 4. The results show that applying the AI-KD approach to three different state-of-the-art Face Image Quality Assessment methods, SER-FIQ [Te20b], DifFIQA(R) [BP3], and CR-FIQA [Bo23], always yields superior performance than the base FIQA techniques while considering both Cross-Model and Same-Model scenarios. This can be reflected through the overall pAUC scores, wherein across all datasets, such extended methods of AI-KD really show higher values with respect to the baseline methods.

FIQA Method	FR Model	Baseline pAUC	AI-KD Extended pAUC	Improvement
SER-FIQ	AdaFace	0.85	0.90	+0.05
DifFIQA(R)	CosFace	0.87	0.92	+0.05
CR-FIQA	SwinFace	0.88	0.93	+0.05

Tab. 4: Improvment of pAUC scorsed with AI-KD method, based on [Ba24]

Despite significant advancements in FIQA, the papers by Babnik [BD3] acknowledge limitations that require further investigation. The dependence of the optimization-based approach on initial quality scores and specific FR models requires exploring more universal and adaptable optimization strategies [BD3]. Again, while AI-KD does enhance robustness against alignment variations, he says its dependence on a baseline alignment accuracy reflects the need to extend applicability to poorly aligned or occluded faces [Ba24].

4 Future Research

In recent years, the FIQA sector has increased to a greater degree now, importance is updated towards clarity and generalization beyond any particular FR tasks. There has been a lot of progress in terms of the covered FIQA techniques in this survey, and many such problems yet remain unsettled, which require close attention. For instance, in [Li24] based on the requirement, bigger and more diverse datasets are required, which would include more variations of distortions and demographic variations. Moreover, it is mentioned in [Su24] that another potential research direction needs to be including generative priors inherent in a few pre-trained generative models into blind IQA to enhance precision.

Modern state-of-the-art FIQA predominantly measures image quality by predicting a unified quality score, which may be incompetent in capturing such elusive image quality, and is in no position to explain how the quality of images is perceived differently by different facial recognition models-classifier [BD3]. This is particularly true because the performance of a variety of FIQA techniques is quite dependent upon the facial recognition model at hand, and different models may have different ideas about what constitutes image quality. Thus, developing more elaborate face image quality assessment techniques that give insights into the specific attributes or features affecting the quality evaluations can greatly improve their utility in real-world applications. This may involve methods that measure the quality of images as a whole but give feedback regarding specific attributes of the image, like alignment, lighting, or occlusion, which will affect recognition performance, as suggested by the alignment sensitivity issues raised in Babnik's work on AI-KD [Ba24]. On the other side, closing this gap further between FIQA output and practical consequences for facial recognition systems could be done by including mechanisms for model-specific feedback about how changes in image quality metrics impact recognition accuracy.

5 Conclusion

This research has introduced the comprehensive state of the art of FIQA methods and their important role in FR systems' performance improvement and reliability enhancement. It has also thoroughly investigated the approach with the most popular database-driven methods, pixel-based techniques, and comparisons against state-of-the-art techniques with alternative metrics.

It has been highlighted some of the noteworthy contributions that have made a significant impact on FIQA research. Liu's comprehensive database for FIQA and Su's diverse IQA dataset contributed much to developing and rigorously evaluating new ways for assessing FIQA [B2, Su24]. Novel pixel-based techniques from researchers like Biagi [Bi23] by way of SHAP-based pixel-level FIQA and Huber [Hu22] through the evaluation of pixel-level assessment, have given way to a new era of fine-grain insight into image quality and implored the importance of knowing how local parts of an image impact recognition performance.

Most of the modern approaches to FIQA, like methods sophisticated as FaceQNet, SDD-FIQA, and MagFace, have proved pretty effective in estimating the quality of face images. Furthermore, the emerging IFQA has pushed a further step forward toward more accurate and human-aligned evaluations by focusing on salient facial regions.

One of the obvious trends of research into FIQA over the past few years has been the application of deep learning and generative models to boost the accuracy of quality assessments [Sc22]. Future research efforts should be focused on re-processing available datasets, enhancing model-specific feedback mechanisms, and developing new methods that can offer clear and understandable insights about the peculiar properties of images that influence the effectiveness of recognition. Effectively solving these problems will drive FIQA along its development path and guarantee that more interpretable systems of facial recognition will be developed to meet increasing real-world application demands.

Bibliography

[Ba24] Babnik, Žiga; Boutros, Fadi; Damer, Naser; Peer, Peter; Štruc, Vitomir: , AI-KD: Towards Alignment Invariant Face Image Quality Assessment Using Knowledge Distillation, April 2024. arXiv:2404.09555 [cs].

[BD3] Babnik, Žiga; Damer, Naser; Štruc, Vitomir: Optimization-Based Improvement of Face Image Quality Assessment Techniques. In: 2023 11th International Workshop on Biometrics and Forensics (IWBF). pp. 1–6, April 2023.

[Bi23] Biagi, Clara; Rethfeld, Louis; Kuijper, Arjan; Terhörst, Philipp: Explaining Face Recognition Through SHAP-Based Pixel-Level Face Image Quality Assessment. 2023 IEEE International Joint Conference on Biometrics (IJCB), pp. 1–10, September 2023. Conference Name: 2023 IEEE International Joint Conference on Biometrics (IJCB) ISBN: 9798350337266 Place: Ljubljana, Slovenia Publisher: IEEE.

[Bo23] Boutros, Fadi; Fang, Meiling; Klemt, Marcel; Fu, Biying; Damer, Naser: CR-FIQA: Face Image Quality Assessment by Learning Sample Relative Classifiability. pp. 5836–5845, 2023.

[BP2] Babnik, Žiga; Peer, Peter; Štruc, Vitomir: FaceQAN: Face Image Quality Assessment Through Adversarial Noise Exploration. In: 2022 26th International Conference on Pattern Recognition (ICPR). pp. 748–754, August 2022. ISSN: 2831-7475.

[BP3] Babnik, Žiga; Peer, Peter; Štruc, Vitomir: DifFIQA: Face Image Quality Assessment Using Denoising Diffusion Probabilistic Models. 2023 IEEE International Joint Conference on Biometrics (IJCB), pp. 1–10, September 2023.

[B2] Babnik, Žiga; Štruc, Vitomir: , Assessing Bias in Face Image Quality Assessment, November 2022. arXiv:2211.15265 [cs].

[CY23] Chen, Zehao; Yang, Hua: L2RT-FIQA: Face Image Quality Assessment via Learning-to-Rank Transformer. volume 1766, Springer Nature Singapore, Singapore, pp. 270–285, 2023. Book Title: Digital Multimedia Communications Series Title: Communications in Computer and Information Science.

[DKG16] Darmstadt-KdNr, Normen-Download-Beuth-Hochschule; Gzndr, Id: Technologies de l'information — Qualité d'échantillon biométrique — Partie 1: Cadre. 2016.

[FD22] Fu, Biying; Damer, Naser: Explainability of the Implications of Supervised and Unsupervised Face Image Quality Estimations Through Activation Map Variation Analyses in Face Recognition Models. IEEE Computer Society, pp. 349–358, January 2022.

[He19] Hernandez-Ortega, Javier; Galbally, Javier; Fierrez, Julian; Haraksim, Rudolf; Beslay, Laurent: FaceQnet: Quality Assessment for Face Recognition based on Deep Learning. In: 2019 International Conference on Biometrics (ICB). IEEE, Crete, Greece, pp. 1–8, June 2019.

[Hu08] Huang, Gary; Mattar, Marwan; Berg, Tamara; Learned-Miller, Eric: Labeled Faces in the Wild: A Database forStudying Face Recognition in Unconstrained Environments. Tech. rep., 10 2008.

[Hu22] Huber, Marco; Terhöst, Philipp; Kirchbuchner, Florian; Damer, Naser; Kuijper, Arjan: On Evaluating Pixel-Level Face Image Quality Assessment. In: 2022 30th European Signal Processing Conference (EUSIPCO). pp. 1052–1056, August 2022. ISSN: 2076-1465.

[Jo23] Jo, Byungho; Cho, Donghyeon; Park, In Kyu; Hong, Sungeun: IFQA: Interpretable Face Quality Assessment. pp. 3444–3453, 2023.

[Ka20] Karras, Tero; Laine, Samuli; Aittala, Miika; Hellsten, Janne; Lehtinen, Jaakko; Aila, Timo: Analyzing and Improving the Image Quality of StyleGAN. pp. 8110–8119, 2020.

[KHR21] Knoche, Martin; Hörmann, Stefan; Rigoll, Gerhard: Cross-Quality LFW: A Database for Analyzing Cross- Resolution Image Face Recognition in Unconstrained Environments. 2021 16th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2021), pp. 1–5, 2021.

[Ki24] Kim, Minsoo; Nam, Gi Pyo; Kim, Haksub; Park, Haesol; Kim, Ig-Jae: IG-FIQA: Improving Face Image Quality Assessment through Intra-class Variance Guidance robust to Inaccurate Pseudo-Labels. 2024. Publisher: arXiv Version Number: 1.

[Li22] Lian, Qiye; Xie, Xiaohua; Zheng, Huicheng; Zhang, Yongdong: Variance of Local Contribution: an Unsupervised Image Quality Assessment for Face Recognition. 2022 26th International Conference on Pattern Recognition (ICPR), pp. 4665–4670, August 2022. Conference Name: 2022 26th International Conference on Pattern Recognition (ICPR) ISBN: 9781665490627 Place: Montreal, QC, Canada Publisher: IEEE.

[Li24] Liu, Tie; Li, Shengxi; Xu, Mai; Yang, Li; Wang, Xiaofei: Assessing Face Image Quality: A Large-Scale Database and a Transformer Method. IEEE Transactions on Pattern Analysis and Machine Intelligence, 46(5):3981–4000, May 2024. Scale_Database_and_a_Transformer_Method.

[Ma24] Maragkoudakis, Emmanouil; Papadopoulos, Symeon; Varlamis, Iraklis; Diou, Christos: Sampling Strategies for Mitigating Bias in Face Synthesis Methods. 2024. Publisher: arXiv Version Number: 1.

[Me21] Meng, Qiang; Zhao, Shichao; Huang, Zhida; Zhou, Feng: MagFace: A Universal Representation for Face Recognition and Quality Assessment. pp. 14225–14234, 2021.

[Ou21] Ou, Fu-Zhao; Chen, Xingyu; Zhang, Ruixin; Huang, Yuge; Li, Shaoxin; Li, Jilin; Li, Yong; Cao, Liujuan; Wang, Yuan-Gen: SDD-FIQA: Unsupervised Face Image Quality Assessment with Similarity Distribution Distance. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 7666–7675, June 2021. Conference Name: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) ISBN: 9781665445092 Place: Nashville, TN, USA Publisher: IEEE.

[Ou23] Ou, Fu-Zhao; Chen, Baoliang; Li, Chongyi; Wang, Shiqi; Kwong, Sam: Troubleshooting Ethnic Quality Bias with Curriculum Domain Adaptation for Face Image Quality Assessment. In: 2023 IEEE/CVF International Conference on Computer Vision (ICCV). pp. 20661–20672, October 2023. ISSN: 2380-7504.

[Ro23] Rodriguez, Andrea Macarulla; Unzueta, Luis; Geradts, Zeno; Worring, Marcel; Elordi, Unai: Multi-Task Explainable Quality Networks for Large-Scale Forensic Facial Recognition. IEEE Journal of Selected Topics in Signal Processing, 17(3):612–623, May 2023.

[Sc22] Schlett, Torsten; Rathgeb, Christian; Henniger, Olaf; Galbally, Javier; Fierrez, Julian; Busch, Christoph: Face Image Quality Assessment: A Literature Survey. ACM Computing Surveys, 54(10s):210:1–210:49, September 2022.

[Sc23] Schlett, Torsten; Rathgeb, Christian; Tapia, Juan; Busch, Christoph: , Considerations on the Evaluation of Biometric Quality Assessment Algorithms, October 2023. arXiv:2303.13294 [cs].

[Se16] Sengupta, Soumyadip; Chen, Jun-Cheng; Castillo, Carlos; Patel, Vishal M.; Chellappa, Rama; Jacobs, David W.: Frontal to profile face verification in the wild. In: 2016 IEEE Winter Conference on Applications of Computer Vision (WACV). pp. 1–9, 2016.

[Su24] Su, Shaolin; Lin, Hanhe; Hosu, Vlad; Wiedemann, Oliver; Sun, Jinqiu; Zhu, Yu; Liu, Hantao; Zhang, Yanning; Saupe, Dietmar: Going the Extra Mile in Face Image Quality Assessment: A Novel Database and Model. IEEE Transactions on Multimedia, 26:2671–2685, 2024.

[Te20a] Terhorst, Philipp; Kolf, Jan Niklas; Damer, Naser; Kirchbuchner, Florian; Kuijper, Arjan: Face Quality Estimation and Its Correlation to Demographic and Non-Demographic Bias in Face Recognition. 2020 IEEE International Joint Conference on Biometrics (IJCB), pp. 1–11, September 2020. Conference Name: 2020 IEEE International Joint Conference on Biometrics (IJCB) ISBN: 9781728191867 Place: Houston, TX, USA Publisher: IEEE.

[Te20b] Terhorst, Philipp; Kolf, Jan Niklas; Damer, Naser; Kirchbuchner, Florian; Kuijper, Arjan: SER-FIQ: Unsupervised Estimation of Face Image Quality Based on Stochastic Embedding Robustness. pp. 5651–5660, 2020.

[Te23] Terhörst, Philipp; Huber, Marco; Damer, Naser; Kirchbuchner, Florian; Raja, Kiran; Kuijper, Arjan: Pixel-Level Face Image Quality Assessment for Explainable Face Recognition. IEEE Transactions on Biometrics, Behavior, and Identity Science, 5(2):288–297, April 2023. Conference Name: IEEE Transactions on Biometrics, Behavior, and Identity Science.

[Wi02] Winter, Eyal: Chapter 53 The shapley value. In: Handbook of Game Theory with Economic Applications, volume 3, pp. 2025–2054. Elsevier, January 2002.

[ZD18] Zheng, Tianyue; Deng, Weihong: Cross-Pose LFW : A Database for Studying Cross-Pose Face Recognition in Unconstrained Environments. 2018.

[ZDH17] Zheng, Tianyue; Deng, Weihong; Hu, Jiani: Cross-Age LFW: A Database for Studying Cross-Age Face Recognition in Unconstrained Environments. ArXiv, abs/1708.08197, 2017.