

000  
001  
002054  
055  
056

# Advancing FIQA with Intrinsic Age Features: Introducing the U3FQ (Unified Tri-Feature Quality) Metric

057  
058  
059003  
004  
005  
006  
007  
008  
009  
010  
011  
012  
013  
014  
015060  
061  
062  
063  
064  
065  
066  
067  
068  
069

Anonymous CVPR submission

070  
071  
072  
073  
074  
075  
076  
077  
078  
079  
080  
081  
082  
083

Paper ID 12988

084  
085  
086  
087  
088  
089  
090  
091  
092  
093  
094  
095  
096  
097  
098

## Abstract

099  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153

Facial Image Quality Assessment (FIQA) is crucial in enhancing face matching and recognition systems. Traditional FIQA metrics often focus on subjective human visibility, which may not correspond with the features crucial for accurate recognition. To address this issue, we propose the Unified Tri-Feature Quality Metric (U3FQ), a novel assessment framework that integrates three critical elements: age variance, facial expression impact, and congruence scores from four state-of-the-art recognition models. U3FQ utilizes an advanced learning paradigm, employing a ResNet model designed for facial image quality assessment. This approach diverges from conventional metrics by focusing on aspects directly influencing recognition accuracy, such as expression intensities and facial features' congruence with recognition models. Our method enhances congruence scores with quantitative modifiers that account for expression intensities, ensuring a more accurate quality metric for predicting recognition success likelihood. U3FQ was rigorously evaluated against general IQA techniques—BRISQUE, NIQE, and PIQE—and specialized FIQA methodologies like FaceQnet, SER-FIQ, and MagFace. The results demonstrate that U3FQ represents a significant advancement in FIQA, offering a holistic and theoretically robust assessment tool that is highly relevant for various facial recognition scenarios.

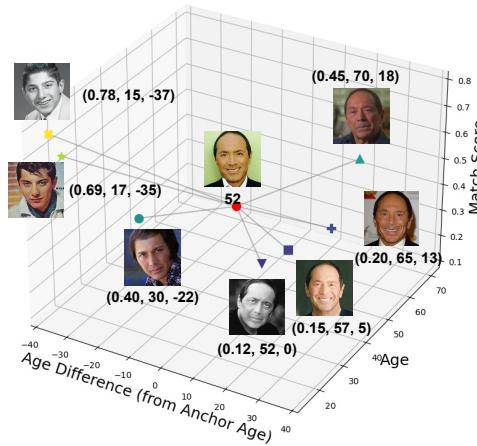


Figure 1. The efficacy of face matching systems is significantly impacted by the age variation between the images being compared. The triplet representation emphasizes the similarity distance, the age of the compared image, and the age difference in relation to the anchor image, with Image 6 as the reference.

scores for image pairs that include the assessed image, thus adding a layer of complexity to FIQA's role in face recognition performance. In recent advancements, FR techniques have shown remarkable results with high-quality frontal images and those of varying quality [13, 17]. However, they still face significant hurdles in completely unconstrained environments [5, 28] where the quality of captured facial images cannot be guaranteed. FIQA methods strive to enhance the performance of FR systems in such settings by offering critical insights into the quality of input images. This input allows FR models to discern and possibly discard images of inferior quality that could lead to erroneous non-matches.

Modern FIQA methods are generally categorized into two distinct styles: regression-based and model-based approaches. Regression-based methods [8, 14, 29] develop a direct mapping from the image space to quality labels generated in a semi-automated manner. These labels often draw

108 on comparison scores across matched image pairs or similarity scores between probe samples and reference images.  
109 On the other hand, model-based [18, 26] FIQA techniques  
110 integrate quality assessment directly within the FR model,  
111 evaluating the quality based on the certainty or statistics  
112 from the generated facial features or embeddings.  
113

114 In this paper, we introduce the Unified Tri-Feature Quality  
115 (U3FQ) metric, a novel approach in Facial Image Quality  
116 Assessment (FIQA). U3FQ redefines FIQA by integrating  
117 recognizability and quality estimation through a unique,  
118 learning-based methodology. Our method diverges from  
119 conventional paradigms by employing match scores in a  
120 weakly supervised manner, serving as our primary quality.  
121 The salient contributions of our work are:  
122

- 123 • U3FQ is the first to systematically analyze facial age  
124 and expression similarity within the FIQA framework,  
125 offering new insights into age-related dynamics and  
126 expression variance in face recognition systems.  
127
- 128 • Our method demonstrates enhanced superior accuracy  
129 and robust generalization across various benchmark  
130 datasets, marking a substantial significant advance-  
131 ment over current FIQA and IQA methodologies.  
132

133 The notable enhancement in U3FQ is the adoption of a  
134 binary classification scheme for facial expressions. This so-  
135 phisticated approach acknowledges the crucial role of ex-  
136 pressions in determining face recognition accuracy. By dis-  
137 tinctly categorizing expressions as congruent or disparate,  
138 also highlights the substantial impact that facial expressions  
139 wield in the overall quality assessment process. Further, as  
140 illustrated in Figure 1, U3FQ integrates age disparity as a  
141 fundamental component, offering an insightful visualiza-  
142 tion of how age variations influence matching similarity.  
143 This integration is pivotal, as it profoundly acknowledges  
144 the impact of age on the authenticity and reliability of face  
145 recognition systems. By doing so, U3FQ captures the true  
146 essence of biometric image quality, utility and fidelity in  
147 face recognition scenarios, transcending the conventional  
148 boundaries of FIQA in Face Recognition Systems.  
149

## 150 2. Related Works 151

152 In this paper, we contextualize our work within the  
153 broader FIQA landscape, in which we are introducing the  
154 Unified Tri-Feature Quality (U3FQ) metric as a novel per-  
155 spective in FIQA. Our approach, inspired by the latest  
156 trends in unsupervised, semi-supervised, and regression-  
157 based learning, uniquely integrates facial biometrics fea-  
158 tures such as age and facial expressions. This integration  
159 enriches the conventional FIQA framework, steering it to-  
160 wards more nuanced and holistic assessments.  
161

### 162 2.1. Innovative Approaches in FIQA 163

164 Recent innovations in face recognition have been driven  
165 by a blend of sophisticated unsupervised and semi-  
166 supervised learning methods, fundamentally aimed at en-  
167 hancing the recognizability of advanced face recognition  
168 systems. These methods, exemplified by seminal works  
169 such as SER-FIQ [26], SDD-FIQA [20], PCNet [29], and  
170 [4], have demonstrated the effectiveness of leveraging in-  
171trinsic data characteristics and a robust combination of both  
172 annotated and unannotated data. They underscore the po-  
173 tential of using advanced embedding variability analysis  
174 and similarity distribution distancing strategies to comprehen-  
175 sively assess facial image quality.  
176

177 In line with these developments, our U3FQ metric ex-  
178 tends the principles of these innovative learning approaches.  
179 It diverges from traditional methods by not solely relying  
180 on pseudo quality labels or embedding uncertainty. Instead,  
181 U3FQ incorporates additional biometric data, refining the  
182 quality assessment process and addressing biases inherent  
183 in label-dependent methods.  
184

### 185 2.2. Integrated Biometric Analysis in FIQA 186

187 Drawing on the strengths of both advanced computa-  
188 tional techniques and human-perceivable facial attributes,  
189 the Unified Tri-Feature Quality (U3FQ) metric represents  
190 a sophisticated amalgamation of the finest elements found  
191 in FIQA methodologies. U3FQ, while sharing conceptual  
192 similarities with notable works like CR-FIQA [6], FaceQnet  
193 [14], and FaceQAN [3], distinctively pushes the boundaries  
194 of conventional FIQA approaches. It incorporates a deeper,  
195 more nuanced integration of biometric analysis, trans-  
196 cending traditional computational assessments.  
197

198 What distinctly sets the Unified Tri-Feature Quality  
199 (U3FQ) metric apart is its meticulous attention to the sub-  
200 tleties of facial biometrics, an aspect often underempha-  
201 sized in other models. This robust integration ensures that  
202 U3FQ not only aligns with but also significantly enhances  
203 the practical applications of face recognition systems. Ac-  
204 knowledging the importance of facial expressions, as under-  
205 scored in seminal studies [7, 16, 24], U3FQ integrates these  
206 critical aspects into its framework. Likewise, it draws upon  
207 the biometric significance of facial age features, as detailed  
208 in pivotal research works [2, 10, 12, 25], demonstrating how  
209 age characteristics can profoundly impact recognition tasks.  
210

211 By factoring in the intricacies of facial expressions and  
212 age disparities, U3FQ emerges as an unparalleled tool, res-  
213 onating profoundly with the real-world demands and com-  
214 plexities of contemporary face recognition technology. It  
215 transcends traditional FIQA approaches by offering a more  
216 contextually enriched and biometrically informed perspec-  
217 tive. This pioneering integration not only elevates U3FQ  
218 within the FIQA domain but also paves the way for more  
219 contextually aware and accurate face recognition systems,  
220

216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
setting a new benchmark for future innovations. U3FQ, therefore, stands as a significant advancement in FIQA, offering a comprehensive solution well-suited to the evolving demands of face recognition systems. Its ability to synthesize various learning approaches and incorporate key biometric features positions it as a groundbreaking tool.

### 3. Methodology

Our work introduces the Unified Tri-Feature Quality Metric (U3FQ) for Contextual Facial Image Quality Assessment (FIQA), a key advancement for the precision of biometric identification systems. This section outlines our integrated methodology for developing and refining the U3FQ metric, which accounts for match scores, age disparities, and facial expressions. We also discuss the machine learning and deep learning frameworks applied in our analysis.

- In Subsection 3.1, we provide a succinct overview of the theoretical foundations of U3FQ and its significance within the field of FIQA. - Subsection 3.2 details the operational aspects of our model, including how age-related match score adjustments and expression-based calibrations are integrated using the AgeDB [19] dataset. - The design of our computational framework, which utilizes both Random Forest and ResNet models for quality score prediction, is presented in Subsection 3.3.

#### 3.1. Theoretical Background

**Facial Age Difference:** The efficacy of face matching systems is significantly influenced by the age difference between the anchor image and the comparison image, as illustrated in Figure 1. This influence varies notably with the anchor's age, necessitating a nuanced approach to modeling age difference penalties. For anchors aged between 20 and 30 years, negative age differences typically correlate with child images, which present a considerable challenge due to the substantial facial feature changes that occur during maturation. Conversely, for anchors over 35 years of age, negative age differences represent younger adult images, where changes in facial features are less pronounced.

To empirically underpin this approach, we present Detection Error Tradeoff (DET) plots that demonstrate the variance in performance with different age groups for all four models: VGG-Face [27], OpenFace [1], ArcFace [9], and FaceNet [23]. Due to page limitations, these plots are included in the supplementary material. Here, we have added the DET plots from VGG-Face in Figure 3, which show the False Non-Match Rate (FNMR) for different age groups. These plots highlight the impact of age difference on the efficacy of face-matching systems, for different anchor age where there is a pronounced increase in FNMR as the age difference becomes more negative. The trend gradually inverts with increasing anchor age, reflecting the maturation and stabilization of facial features over time.

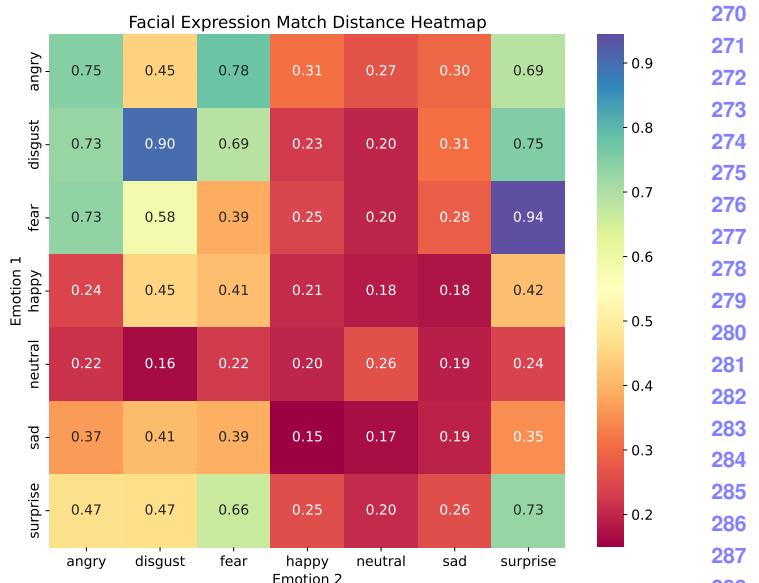


Figure 2. The differential impact of facial expressions on the match score, with weak emotions having a constant effect and strong emotions modifying the score proportionally to their intensity.

**The Influence of Facial Expressions:** The similarity in facial expressions between two images significantly influences recognition performance, as variations in expressions can distort critical facial features used in establishing a match. Consequently, this also affects the overall quality of recognition. Figure 2 illustrates the impact that discrepancies in facial expressions have on matching performance, as evidenced by the average match scores across diverse expression pairs.

Our methodology ensures a more refined and context-sensitive assessment of facial similarity, taking into account not just the physical resemblance but also the expressive context of each face. This leads to a more accurate and realistic evaluation of facial images, particularly relevant in dynamic real-world scenarios where facial expressions can vary significantly.

#### 3.2. Formulations and Optimization

Building on the observations from empirical evidence, we proceed to formulate the mathematical model that incorporates age difference penalties into the facial match score. The age difference penalty function is adapted as follows:

$$f(d, a) = \begin{cases} e^{\alpha(d+\beta)} & \text{if } d < 0 \text{ and } a \leq 30, \\ e^{\alpha(d+\beta)/\theta} & \text{if } d < 0 \text{ and } a > 30, \\ \gamma \cdot d & \text{if } d \geq 0, \end{cases} \quad (1)$$

where  $d$  represents the age difference between the anchor

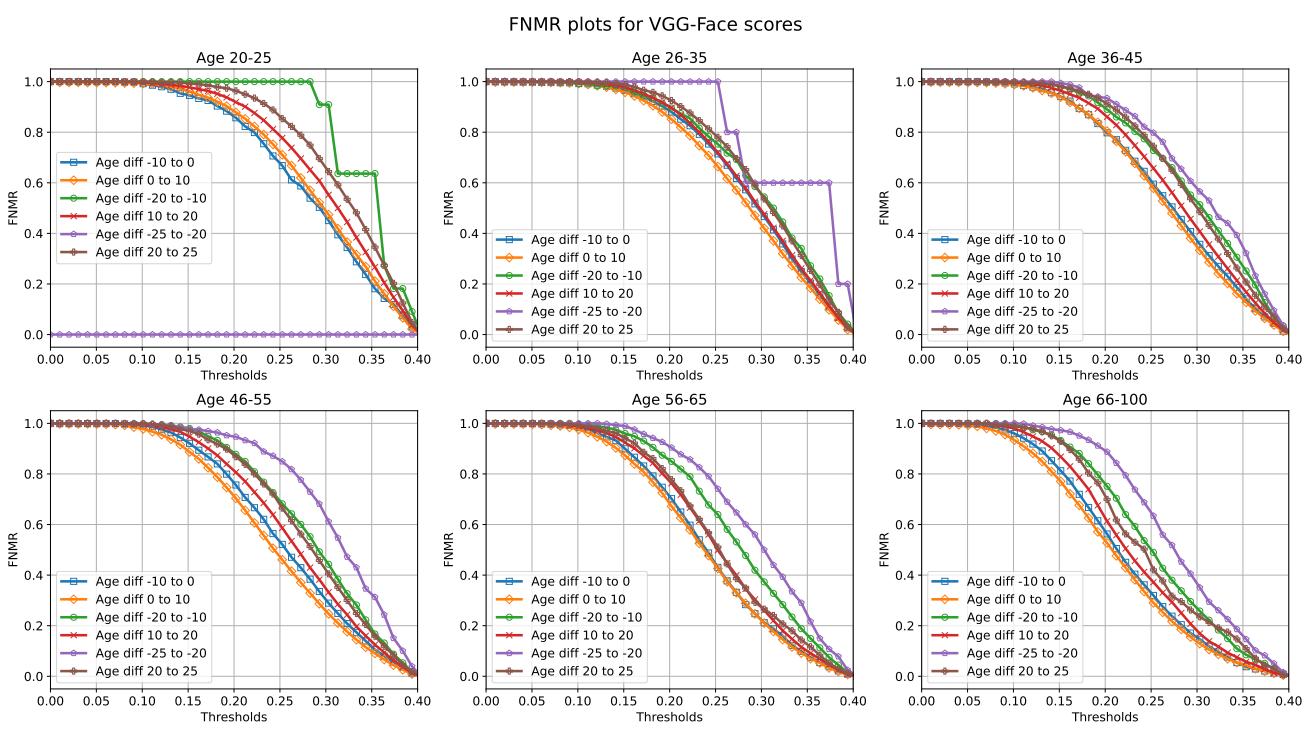


Figure 3. The VGG-Face DET plots, displaying the False Non-Match Rate (FNMR) across various age groups and age difference categories, reveal significant insights about the effects of anchor age on facial recognition accuracy. Notably, the age group of 35-45 years aligns closely with facial images across a wide age range, suggesting enhanced feature consistency within this demographic. For individuals over 60 years, a broad age difference (-30 to 30 years) exhibits minimal impact on FNMR, indicating a decreased variation in facial features with age. These observations, drawn from the AgeDB dataset [19] using the deepface face verification library.

and the comparison image,  $a$  denotes the anchor's age, and  $\alpha$ ,  $\beta$ , and  $\gamma$  are parameters dictating the function's shape. The factor  $\theta$  serves as a damping parameter that reduces the penalty for older anchors.

Our methodology also accounts for the subtle yet significant influence of facial expressions on the match score. This is achieved through the facial expression impact function  $g(e)$ , which distinguishes between 'weak' and 'strong' emotions, as detailed below:

$$g(e) = \begin{cases} c & \text{if } e \text{ is a weak emotion,} \\ d \cdot \text{EXPR\_SCORE}(e) & \text{if } e \text{ is a strong emotion,} \end{cases} \quad (2)$$

where  $c$  is a constant factor for weak emotions, and  $d$  scales the expression score  $\text{EXPR\_SCORE}(e)$  for strong emotions.

In Figure 4 utilizing the equation 2 designed for face expression similarity function. Our function is calibrated to assign higher scores to faces that are similar, effectively distinguishing them from dissimilar ones.

A key feature of our approach is the nuanced consideration of facial expressions in determining these scores.

For instance, neutral expressions, which are generally more predictable and consistent for recognition purposes, are assigned the highest scores. In contrast, faces exhibiting strong emotions such as surprise or happiness, despite being similar, receive comparatively lower scores. This adjustment acknowledges the impact of expressive variability on the recognizability of faces.

These formulations, alongside the empirical insights, collectively enhance the fidelity of the FIQA model's predictions. By incorporating the dynamics of human aging and expressions, we ensure that our facial recognition system is not only secure but also user-friendly, accommodating the complexities of human features and behaviors.

The algorithm detailed below outlines the process for computing the contextual quality score and estimating the age for a given input image using a ResNet model. The procedure leverages a feature vector that encompasses age, expression, and congruence score, which are derived from the input image and used to predict the quality score.

### 3.3. Architecture

The U3FQ algorithm initiates with the computation of the match score distance from an ensemble of models  $M =$

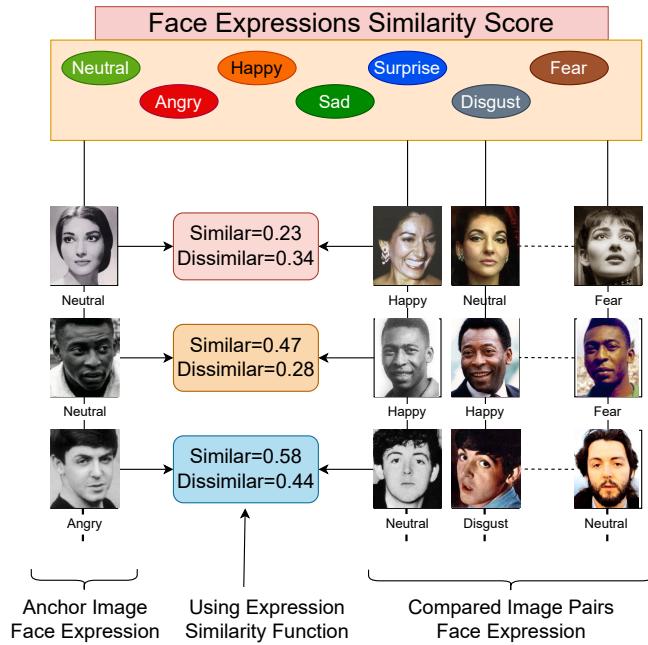


Figure 4. The methodical process of extracting similarity scores for individual images based on their matching counterparts

$\{M_1, M_2, M_3, M_4\}$ . The distance metric, denoted as  $d$ , is derived from the pairwise discrepancies in the features extracted by each model for the given image  $I$ . The age difference function  $f(d, a)$  is applied to adjust  $d$  based on the age of the anchor image  $a$ .

Simultaneously, the expression impact function  $g(e)$  adjusts the congruence score depending on the facial expression  $e$ , where  $c$  is a constant factor for weak emotions and  $d$  is a scaling factor for strong emotions, coupled with the expression score  $\text{EXPR\_SCORE}(e)$ .

These functions are crucial as they capture the dynamic nature of facial recognition where age and emotional expression significantly impact the quality of facial features captured in an image.

The algorithm then constructs a feature vector  $F$ , encompassing the age  $a$ , expression  $e$ , and the aggregate congruence score  $U$ , calculated as a weighted sum of the match score distances adjusted by the age and expression multipliers. This vector is integral in capturing the nuanced elements that contribute to the facial image quality.

$$U \leftarrow \sum_{model \in M} d \cdot f(d, a) \cdot g(e)$$

$$F \leftarrow [a, e, U]$$

A set of stochastic embeddings are generated through the ResNet model  $RN$  across  $m$  iterations to provide robust estimates of the image quality  $Q$  and the subject's age. The

### Algorithm 1 U3FQ: Unified Tri-Feature Quality Assessment for Contextual Facial Image Quality

```

Input: Single input image  $I$ , ResNet model  $RN$ , age  $a$ , expression  $e$ , match score distance models  $M = \{M_1, M_2, M_3, M_4\}$ 
Output: U3FQ Score or Quality Score
1:  $U \leftarrow 0$ 
2:  $A \leftarrow []$ 
3:  $E \leftarrow []$ 
4: for all  $model \in M$  do
5:    $d \leftarrow \text{ComputeMatchScoreDistance}(I, model)$ 
6:    $A \leftarrow f(d, a)$ 
7:    $A \leftarrow A + [A]$ 
8:   if  $\text{ExpressionsAreSame}(I, model)$  then
9:      $E \leftarrow c$ 
10:    else
11:       $E \leftarrow d * \text{EXPR\_SCORE}(e)$ 
12:    end if
13:     $E \leftarrow E + [E]$ 
14:   $U \leftarrow U + d * A * E$ 
15: end for
16:  $F \leftarrow [a, e, U]$ 
17: procedure U3FQ_ASSESSMENT( $I, F, RN, m = 100$ )
18:    $QualityScores \leftarrow []$ 
19:   for  $i \leftarrow 1$  to  $m$  do
20:      $quality \leftarrow RN.\text{Predict}(I, F)$ 
21:      $QualityScores \leftarrow QualityScores + [quality]$ 
22:   end for
23:    $finalQuality \leftarrow \text{Average}(QualityScores)$ 
24:   return  $finalQuality$ 
end procedure

```

embeddings are processed to yield a final quality score, reflecting the stability and robustness of the features in the presence of inherent variabilities in facial images.

This mathematical and algorithmic formulation of the U3FQ model demonstrates a robust mechanism for assessing facial image quality, providing insights into the complex interplay between age, expression, and recognition robustness. The model's efficacy is further corroborated through empirical evaluations, showcasing its potential to enhance the performance of biometric systems significantly.

### 3.4. Analytical Study

This subsection embarks on a discourse surrounding alternative methodologies that hold potential for future exploration. Considerations include the grouping of facial images by expression, the differential impacts of varying age differences, and other conceivable permutations that could influence the FIQA landscape.

540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555556 Figure 5. Generation of proxy ground truth quality label with face images and their age difference for advancing FIQA.  
557558 

## 4. Experiments

559 

### 4.1. Datasets

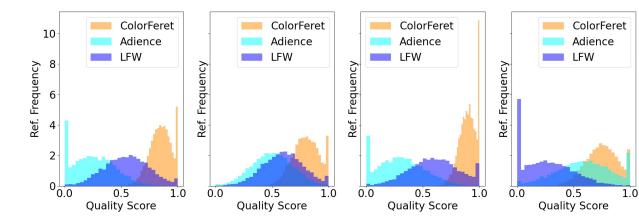
560 In our study, we utilize the AgeDB dataset [19], which  
561 comprises approximately 16,488 images representing a  
562 wide spectrum of age variations across different identities.  
563 This dataset serves as the cornerstone for our evaluations.  
564 For comprehensive assessment and benchmarking, we em-  
565 ploy several esteemed datasets: LFW [15], ColorFeret [21],  
566 and Adience [11]. These datasets provide a diverse range of  
567 facial imagery that enables us to rigorously compare and an-  
568alyze the performance of our proposed methodologies. Re-  
569 sults are meticulously evaluated, showcasing the robustness  
570 of our approach in handling age-variant facial recognition  
571 tasks.

572 

## 5. Results

573 

U3FQ Quality Distributions for each model

583 Figure 6. Distribution of scores on different datasets.  
584585  
586  
587  
588  
589  
590  
591  
592  
593594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647

648

## 6. Conclusion

The Unified Tri-Feature Quality Metric (U3FQ) signifies a pivotal advancement in the realm of Facial Image Quality Assessment (FIQA). By integrating critical elements such as age variance and facial expression impact, U3FQ presents a nuanced and comprehensive method for evaluating facial images, which closely aligns with the practical demands of face recognition systems. This research emphasizes the significance of these biometric features in enhancing the accuracy and reliability of recognition models, thereby transcending the confines of traditional FIQA metrics that predominantly rely on subjective human visibility assessments. Through rigorous evaluations on an extensive array of face quality image datasets and benchmark comparisons with state-of-the-art techniques, U3FQ has demonstrated its superiority in delivering relevant and precise quality assessments.

Looking ahead, our future work aims to augment the predictive power of U3FQ by incorporating additional factors such as illumination and pose alongside age variance and expression. This enhancement will aim to further refine the accuracy of reference quality labels, ensuring that U3FQ remains at the forefront of FIQA methodologies. By addressing these additional variables, we intend to broaden the scope and effectiveness of U3FQ, thereby making it an even more robust tool for assessing facial image quality in diverse and challenging recognition scenarios.

676

## 677 References

- [1] Brandon Amos, Bartosz Ludwiczuk, Mahadev Satyanarayanan, et al. Openface: A general-purpose face recognition library with mobile applications. *CMU School of Computer Science*, 6(2):20, 2016. 3
- [2] Raphael Angulu, Jules R Tapamo, and Aderemi O Adewumi. Age estimation via face images: a survey. *EURASIP Journal on Image and Video Processing*, 2018(1):1–35, 2018. 2
- [3] Žiga Babnik, Peter Peer, and Vítomír Štruc. Faceqan: Face image quality assessment through adversarial noise exploration. In *2022 26th International Conference on Pattern Recognition (ICPR)*, pages 748–754. IEEE, 2022. 2
- [4] Lacey Best-Rowden and Anil K Jain. Automatic face image quality prediction. *arXiv preprint arXiv:1706.09887*, 2017. 2
- [5] Fadi Boutros, Naser Damer, Jan Niklas Kolf, Kiran Raja, Florian Kirchbuchner, Raghavendra Ramachandra, Arjan Kuijper, Pengcheng Fang, Chao Zhang, Fei Wang, et al. Mfr 2021: Masked face recognition competition. In *2021 IEEE International Joint Conference on Biometrics (IJCB)*, pages 1–10. IEEE, 2021. 1
- [6] Fadi Boutros, Meiling Fang, Marcel Klemt, Biying Fu, and Naser Damer. Cr-fifa: face image quality assessment by learning sample relative classifiability. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5836–5845, 2023. 2

- [7] Andrew J Calder and Andrew W Young. Understanding the recognition of facial identity and facial expression. *Facial Expression Recognition*, pages 41–64, 2016. 2
- [8] Kai Chen, Taihe Yi, and Qi Lv. Lightqnet: Lightweight deep face quality assessment for risk-controlled face recognition. *IEEE Signal Processing Letters*, 28:1878–1882, 2021. 1
- [9] Jiankang Deng, Jia Guo, Niannan Xue, and Stefanos Zafeiriou. Arcface: Additive angular margin loss for deep face recognition. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 4690–4699, 2019. 3
- [10] Natalie C Ebner. Age of face matters: Age-group differences in ratings of young and old faces. *Behavior research methods*, 40:130–136, 2008. 2
- [11] Eran Eidinger, Roei Enbar, and Tal Hassner. Age and gender estimation of unfiltered faces. *IEEE Transactions on information forensics and security*, 9(12):2170–2179, 2014. 6
- [12] Yun Fu, Guodong Guo, and Thomas S Huang. Age synthesis and estimation via faces: A survey. *IEEE transactions on pattern analysis and machine intelligence*, 32(11):1955–1976, 2010. 2
- [13] Klemen Grm and Vítomír Štruc. Deep face recognition for surveillance applications, 2018. 1
- [14] Javier Hernandez-Ortega, Javier Galbally, Julian Fierrez, Rudolf Haraksim, and Laurent Beslay. Faceqnet: Quality assessment for face recognition based on deep learning. In *2019 International Conference on Biometrics (ICB)*, pages 1–8. IEEE, 2019. 1, 2
- [15] Gary B Huang, Marwan Mattar, Tamara Berg, and Eric Learned-Miller. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. In *Workshop on faces in 'Real-Life' Images: detection, alignment, and recognition*, 2008. 6
- [16] Shan Li and Weihong Deng. Deep facial expression recognition: A survey. *IEEE transactions on affective computing*, 13(3):1195–1215, 2020. 2
- [17] Blaž Meden, Peter Rot, Philipp Terhörst, Naser Damer, Arjan Kuijper, Walter J Scheirer, Arun Ross, Peter Peer, and Vítomír Štruc. Privacy-enhancing face biometrics: A comprehensive survey. *IEEE Transactions on Information Forensics and Security*, 16:4147–4183, 2021. 1
- [18] Qiang Meng, Shichao Zhao, Zhida Huang, and Feng Zhou. Magface: A universal representation for face recognition and quality assessment. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 14225–14234, 2021. 2
- [19] Stylianos Moschoglou, Athanasios Papaioannou, Christos Sagonas, Jiankang Deng, Irene Kotsia, and Stefanos Zafeiriou. Agedb: the first manually collected, in-the-wild age database. In *proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pages 51–59, 2017. 3, 4, 6
- [20] Fu-Zhao Ou, Xingyu Chen, Ruixin Zhang, Yuge Huang, Shaolin Li, Jilin Li, Yong Li, Liujuan Cao, and Yuan-Gen Wang. Sdd-fifa: unsupervised face image quality assessment with similarity distribution distance. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 7670–7679, 2021. 2

- 756 [21] P Jonathon Phillips, Hyeonjoon Moon, Syed A Rizvi, and 810  
757 Patrick J Rauss. The feret evaluation methodology for face- 811  
758 recognition algorithms. *IEEE Transactions on pattern anal- 812  
759 ysis and machine intelligence*, 22(10):1090–1104, 2000. 6 813  
760 [22] Torsten Schlett, Christian Rathgeb, Olaf Henniger, Javier 814  
761 Galbally, Julian Fierrez, and Christoph Busch. Face image 815  
762 quality assessment: A literature survey. *ACM Computing 816  
763 Surveys (CSUR)*, 54(10s):1–49, 2022. 1 817  
764 [23] Florian Schroff, Dmitry Kalenichenko, and James Philbin. 818  
765 Facenet: A unified embedding for face recognition and clus- 819  
766 tering. In *Proceedings of the IEEE conference on computer 820  
767 vision and pattern recognition*, pages 815–823, 2015. 3 821  
768 [24] Caifeng Shan, Shaogang Gong, and Peter W McOwan. Fa- 822  
769 cial expression recognition based on local binary patterns: 823  
770 A comprehensive study. *Image and vision Computing*, 824  
771 27(6):803–816, 2009. 2 825  
772 [25] Gillian Slessor, Deborah M Riby, and Ailbhe N Finnerty. 826  
773 Age-related differences in processing face configuration: 827  
774 The importance of the eye region. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 828  
775 68(2):228–231, 2013. 2 829  
776 [26] Philipp Terhorst, Jan Niklas Kolf, Naser Damer, Florian 830  
777 Kirchbuchner, and Arjan Kuiper. Ser-fiq: Unsupervised esti- 831  
778 mation of face image quality based on stochastic embedding 832  
779 robustness. In *Proceedings of the IEEE/CVF conference on 833  
780 computer vision and pattern recognition*, pages 5651–5660, 834  
781 2020. 2 835  
782 [27] VGG. The visual geometry group (vgg) at the university of 836  
783 oxford. 3 837  
784 [28] Mei Wang and Weihong Deng. Deep face recognition: A 838  
785 survey. *Neurocomputing*, 429:215–244, 2021. 1 839  
786 [29] Weidi Xie, Jeffrey Byrne, and Andrew Zisserman. Inducing 840  
787 predictive uncertainty estimation for face recognition. *arXiv preprint arXiv:2009.00603*, 2020. 1, 2 841  
788 842  
789 843  
790 844  
791 845  
792 846  
793 847  
794 848  
795 849  
796 850  
797 851  
798 852  
799 853  
800 854  
801 855  
802 856  
803 857  
804 858  
805 859  
806 860  
807 861  
808 862  
809 863