

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

Anonymous CVPR submission

068
069

Paper ID 12988

070

Abstract

071
072
073
074
075
076
077

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.

078
079
080
081
082
083
084
085
086
087
088
089
090
091
092
093
094
095
096
097
098
099
100

1. Introduction

101
102
103
104
105
106
107

Facial Image Quality Assessment (FIQA) is integral to the enhancement of face recognition (FR) systems, especially when dealing with the diversity of image quality often encountered in real-world scenarios. Traditional FIQA approaches primarily assess the standalone biometric utility of images. However, in the context of FR, this method faces a conceptual challenge known as the "Quality Paradox," as discussed by Schlett et al. [22]. This paradox highlights the need to accurately reflect the reliability of comparison

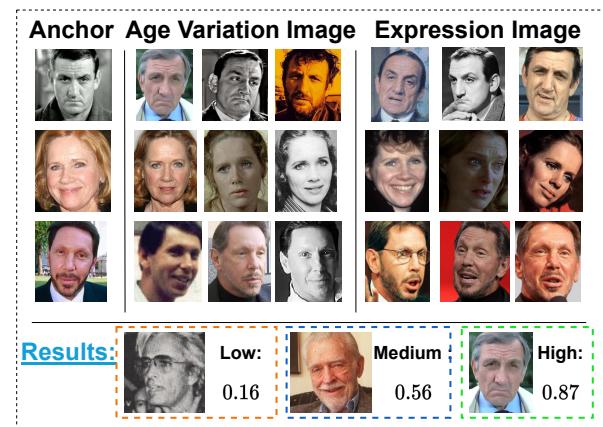


Figure 1. Displaying a range of facial images, each grouped by the set includes an anchor image and its age and expression variants. Results images categorized into low, medium, and high-quality segments, highlighting the model's proficiency in assessing quality variations.

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 on comparison scores across matched image pairs or similarity scores between probe samples and reference images.

108 On the other hand, model-based [18, 26] FIQA techniques
109 integrate quality assessment directly within the FR model,
110 evaluating the quality based on the certainty or statistics
111 from the generated facial features or embeddings.
112

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

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

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

145 2. Related Works

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

155 2.1. Innovative Approaches in FIQA

156 Recent innovations in face recognition have been driven
157 by a blend of sophisticated unsupervised and semi-

158 supervised learning methods, fundamentally aimed at en-
159 hancing the recognizability of advanced face recognition
160 systems. These methods, exemplified by seminal works
161 such as SER-FIQ [26], SDD-FIQA [20], PCNet [29], and
162 [4], have demonstrated the effectiveness of leveraging in-
163 trinsic data characteristics and a robust combination of both
164 annotated and unannotated data. They underscore the po-
165 tential of using advanced embedding variability analysis
166 and similarity distribution distancing strategies to com-
167 prehensively assess facial image quality.

168 In line with these developments, our U3FQ metric ex-
169 tends the principles of these innovative learning approaches.
170 It diverges from traditional methods by not solely relying
171 on pseudo quality labels or embedding uncertainty. Instead,
172 U3FQ incorporates additional biometric data, refining the
173 quality assessment process and addressing biases inherent
174 in label-dependent methods.

175 2.2. Integrated Biometric Analysis in FIQA

176 Drawing on the strengths of both advanced computa-
177 tional techniques and human-perceivable facial attributes,
178 the Unified Tri-Feature Quality (U3FQ) metric represents
179 a sophisticated amalgamation of the finest elements found
180 in FIQA methodologies. U3FQ, while sharing conceptual
181 similarities with notable works like CR-FIQA [6], FaceQnet
182 [14], and FaceQAN [3], distinctively pushes the boundaries
183 of conventional FIQA approaches. It incorporates a deeper,
184 more nuanced integration of biometric analysis, trans-
185 cending traditional computational assessments.

186 What distinctly sets the Unified Tri-Feature Quality
187 (U3FQ) metric apart is its meticulous attention to the sub-
188 tleties of facial biometrics, an aspect often underempha-
189 sized in other models. This robust integration ensures that
190 U3FQ not only aligns with but also significantly enhances
191 the practical applications of face recognition systems. Ac-
192 knowledging the importance of facial expressions, as under-
193 scored in seminal studies [7, 16, 24], U3FQ integrates these
194 critical aspects into its framework. Likewise, it draws upon
195 the biometric significance of facial age features, as detailed
196 in pivotal research works [2, 10, 12, 25], demonstrating how
197 age characteristics can profoundly impact recognition tasks.

198 By factoring in the intricacies of facial expressions and
199 age disparities, U3FQ emerges as an unparalleled tool, res-
200 onating profoundly with the real-world demands and com-
201 plexities of contemporary face recognition technology. It
202 transcends traditional FIQA approaches by offering a more
203 contextually enriched and biometrically informed perspec-
204 tive. This pioneering integration not only elevates U3FQ
205 within the FIQA domain but also paves the way for more
206 contextually aware and accurate face recognition systems,
207 setting a new benchmark for future innovations. U3FQ,
208 therefore, stands as a significant advancement in FIQA, of-
209 fering a comprehensive solution well-suited to the evolving

216 demands of face recognition systems. Its ability to synthe-
 217 size various learning approaches and incorporate key bio-
 218 metric features positions it as a groundbreaking tool.
 219

220 3. Methodology

221 Our work introduces the Unified Tri-Feature Quality
 222 Metric (U3FQ) for Contextual Facial Image Quality Assess-
 223 ment (FIQA), a key advancement for the precision of bio-
 224 metric identification systems. This section outlines our
 225 integrated methodology for developing and refining the U3FQ
 226 metric, which accounts for match scores, age disparities,
 227 and facial expressions. We also discuss the machine learn-
 228 ing and deep learning frameworks applied in our analysis.
 229

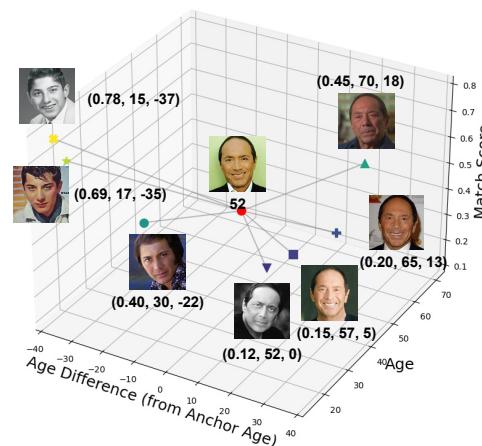
230 - In Subsection 3.1, we provide a succinct overview of
 231 the theoretical foundations of U3FQ and its significance
 232 within the field of FIQA. - Subsection 3.2 details the
 233 operational aspects of our model, including how age-related
 234 match score adjustments and expression-based calibrations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

239 3.1. Theoretical Background

240 **Facial Age Difference:** The efficacy of face matching sys-
 241 tems is significantly influenced by the age difference be-
 242 tween the anchor image and the comparison image, as il-
 243 lustrated in Figure 2. This influence varies notably with the
 244 anchor's age, necessitating a nuanced approach to model-
 245 ing age difference penalties. For anchors aged between 20
 246 and 30 years, negative age differences typically correlate
 247 with child images, which present a considerable challenge
 248 due to the substantial facial feature changes that occur dur-
 249 ing maturation. Conversely, for anchors over 35 years of
 250 age, negative age differences represent younger adult im-
 251 ages, where changes in facial features are less pronounced.
 252

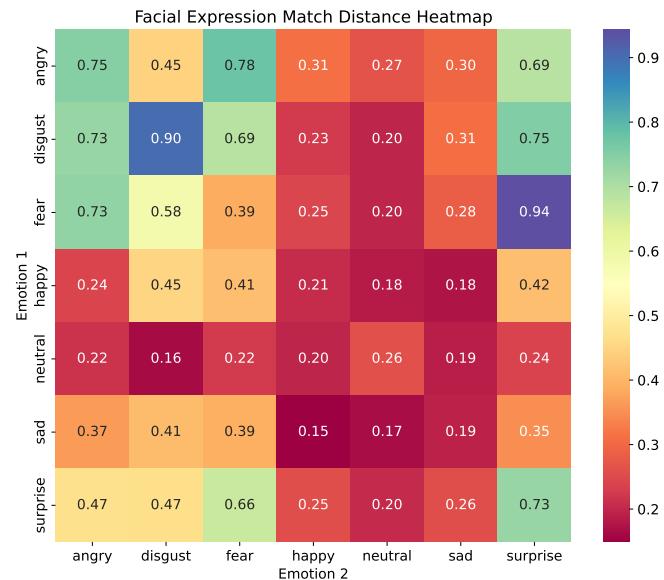
253 To empirically underpin this approach, we present Detec-
 254 tion Error Tradeoff (DET) plots that demonstrate the
 255 variance in performance with different age groups for all
 256 four models: VGG-Face [27], OpenFace [1], ArcFace [9],
 257 and FaceNet [23]. Due to page limitations, these plots are
 258 included in the supplementary material. Here, we have
 259 added the DET plots from VGG-Face in Figure 4, which
 260 show the False Non-Match Rate (FNMR) for different age
 261 groups. These plots highlight the impact of age difference
 262 on the efficacy of face-matching systems, for different an-
 263 chor age where there is a pronounced increase in FNMR
 264 as the age difference becomes more negative. The trend
 265 gradually inverts with increasing anchor age, reflecting the
 266 maturation and stabilization of facial features over time.
 267

268 **The Influence of Facial Expressions:** The similarity in
 269 facial expressions between two images significantly influ-



270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323

Figure 2. 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.



311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323

Figure 3. 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.

ences 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 3 illustrates the impact that discrepancies in facial expressions have on matching performance, as evidenced by the average match scores across diverse expression pairs.

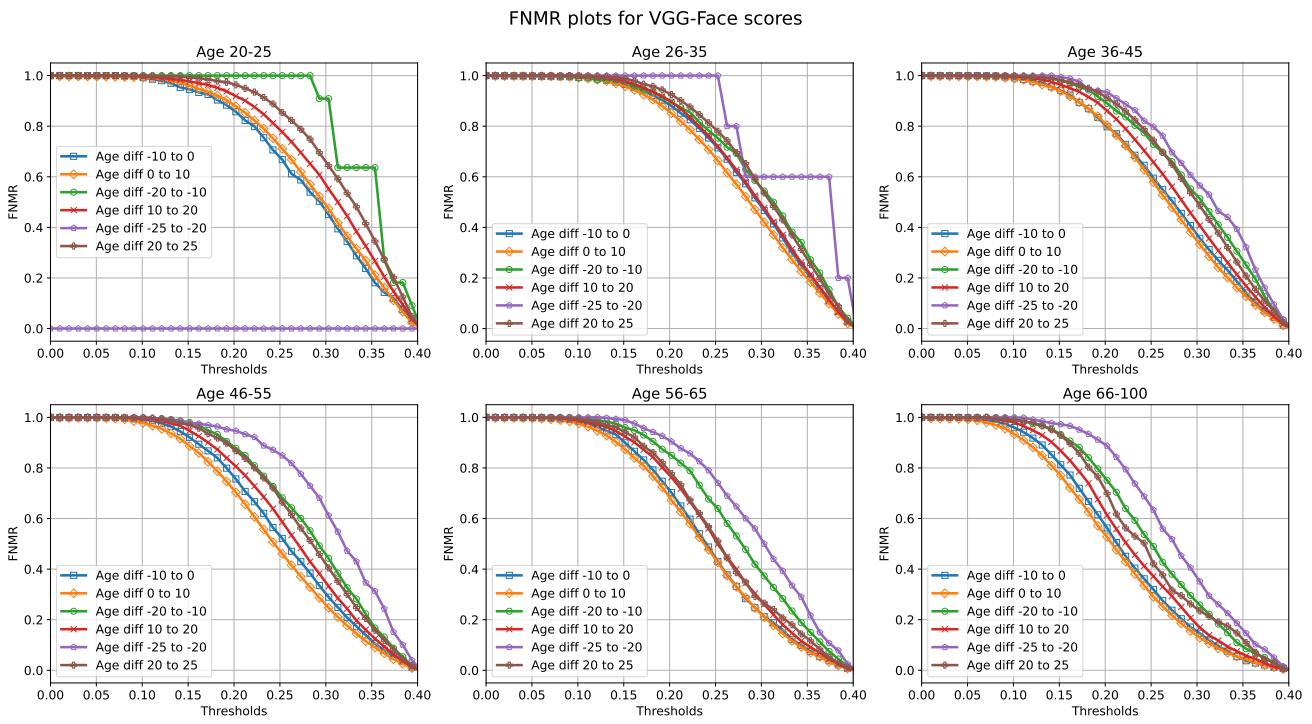


Figure 4. The VGG-Face DET plots, displaying the False Non-Match Rate (FNMR) across various age groups and age difference categories, reveal significant insights about the effects of anchor age on facial recognition accuracy. Notably, the age group of 35-45 years aligns closely with facial images across 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.

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 and the comparison image, a denotes the anchor's age, and α , β , and γ are parameters dictating the function's shape. The factor θ 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 5 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 be-

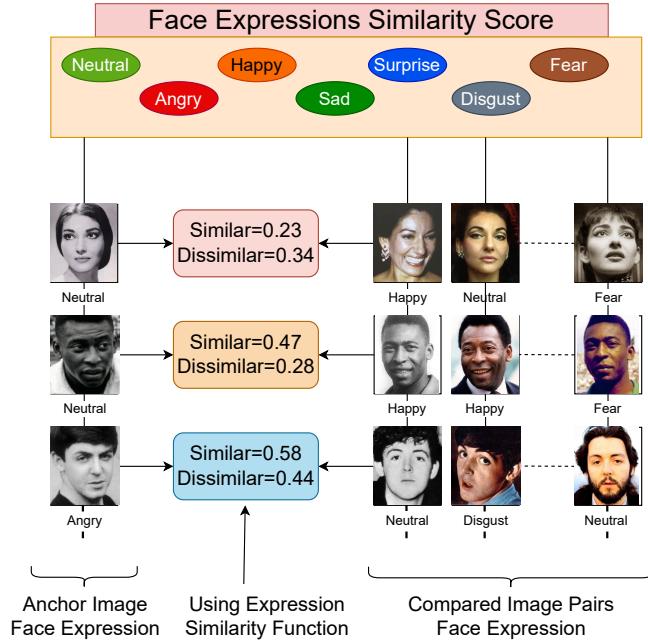


Figure 5. The methodical process of extracting similarity scores for individual images based on their matching counterparts

ing 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 $M = \{M_1, M_2, M_3, M_4\}$. 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

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:    $\mathcal{A} \leftarrow f(d, a)$ 
7:    $A \leftarrow A + [\mathcal{A}]$ 
8:   if ExpressionsAreSame( $I, model$ ) then
9:      $\mathcal{E} \leftarrow c$ 
10:   else
11:      $\mathcal{E} \leftarrow d * \text{EXPR\_SCORE}(e)$ 
12:   end if
13:    $E \leftarrow E + [\mathcal{E}]$ 
14:    $U \leftarrow U + d * \mathcal{A} * \mathcal{E}$ 
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  $finalQuality$ 
24: end procedure

```

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

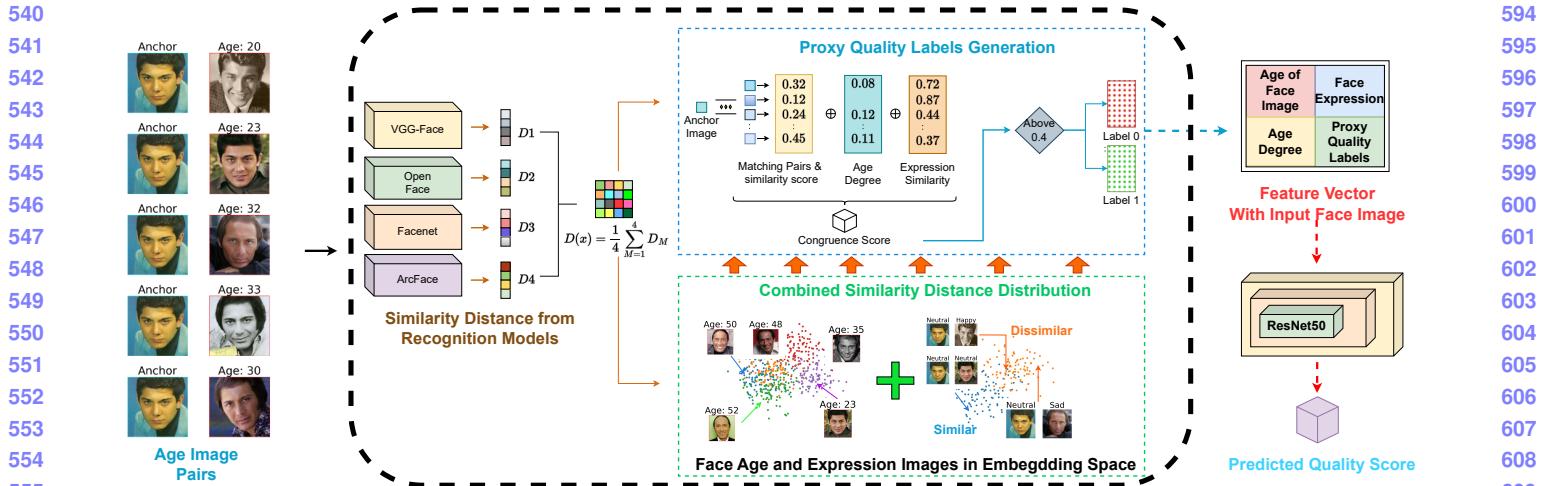


Figure 6. The figure presents a method for generating pseudo ground truth labels in face recognition by assessing age-related variations and expression similarity. It starts by calculating similarity distances between images of the same individuals at different ages using face recognition models. These distances are then normalized and combined with age and expression data to produce labels. These labels are crucial for fine-tuning regression model, leading to a comprehensive quality score that encapsulates recognition accuracy, age differences, and expression similarities.

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.

3.3.1 Regression Network and Quality Estimation

We have advanced and refined an existing Convolutional Neural Network (CNN), originally pre-trained for face recognition tasks, through a process of fine-tuning. This approach of adapting deep learning models to tasks akin to their initial training has been demonstrated effectively in numerous studies. Such networks have been repurposed for detecting facial attributes distinct from identity, including gender, age, and race. In the context of face quality assessment, it is posited that a feature vector containing discriminative facial information should inherently encapsulate aspects of image quality.

For our specific adaptation, we selected the ResNet50 architecture as the foundational network. During the fine-tuning process, we removed the classification layers and augmented the network with fully connected layers, which were then fused with the existing feature vector. This amalgamation was subjected to a sigmoid activation function, designed to yield a quality score.

Crucially, we implemented a training strategy where the weights of the pre-existing layers were frozen, ensuring that only the newly integrated layers were subject to training. This training utilized the pseudo ground truth quality labels generated in the preceding step. The outcome of this refined model is a quality score, ranging from 0 to 1, which correlates with the performance of face recognition, offering a robust measure of the quality of facial images in terms of recognition efficacy.

4. Experiments

4.1. Datasets

In our study, we utilize the AgeDB dataset [19], which comprises approximately 16,488 images representing a wide spectrum of age variations across different identities. This dataset serves as the cornerstone for our evaluations. For comprehensive assessment and benchmarking, we employ several esteemed datasets: LFW [15], ColorFeret [21], and Adience [11]. These datasets provide a diverse range of facial imagery that enables us to rigorously compare and analyze the performance of our proposed methodologies. Results are meticulously evaluated, showcasing the robustness of our approach in handling age-variant facial recognition tasks.

540	594
541	595
542	596
543	597
544	598
545	599
546	600
547	601
548	602
549	603
550	604
551	605
552	606
553	607
554	608
555	609
561	610
562	611
563	612
564	613
565	614
566	615
567	616
568	617
569	618
570	619
571	620
572	621
573	622
574	623
575	624
576	625
577	626
578	627
579	628
580	629
581	630
582	631
583	632
584	633
585	634
586	635
587	636
588	637
589	638
590	639
591	640
592	641
593	642
594	643
595	644
596	645
597	646
598	647

- 756 [7] Andrew J Calder and Andrew W Young. Understanding the
757 recognition of facial identity and facial expression. *Facial*
758 *Expression Recognition*, pages 41–64, 2016. 2
- 759 [8] Kai Chen, Taihe Yi, and Qi Lv. Lightqnet: Lightweight deep
760 face quality assessment for risk-controlled face recognition.
761 *IEEE Signal Processing Letters*, 28:1878–1882, 2021. 1
- 762 [9] Jiankang Deng, Jia Guo, Niannan Xue, and Stefanos
763 Zafeiriou. Arcface: Additive angular margin loss for deep
764 face recognition. In *Proceedings of the IEEE/CVF conference*
765 *on computer vision and pattern recognition*, pages
766 4690–4699, 2019. 3
- 767 [10] Natalie C Ebner. Age of face matters: Age-group differences
768 in ratings of young and old faces. *Behavior research methods*,
769 40:130–136, 2008. 2
- 770 [11] Eran Eidinger, Roei Enbar, and Tal Hassner. Age and gender
771 estimation of unfiltered faces. *IEEE Transactions on information*
772 *forensics and security*, 9(12):2170–2179, 2014. 6
- 773 [12] Yun Fu, Guodong Guo, and Thomas S Huang. Age synthesis
774 and estimation via faces: A survey. *IEEE transactions on pattern*
775 *analysis and machine intelligence*, 32(11):1955–1976, 2010. 2
- 776 [13] Klemen Grm and Vitomir Struc. Deep face recognition for
777 surveillance applications, 2018. 1
- 778 [14] Javier Hernandez-Ortega, Javier Galbally, Julian Fierrez,
779 Rudolf Haraksim, and Laurent Beslay. Faceqnet: Quality
780 assessment for face recognition based on deep learning. In
781 *2019 International Conference on Biometrics (ICB)*, pages
782 1–8. IEEE, 2019. 1, 2
- 783 [15] Gary B Huang, Marwan Mattar, Tamara Berg, and Eric
784 Learned-Miller. Labeled faces in the wild: A database
785 forstudying face recognition in unconstrained environments.
786 In *Workshop on faces in'Real-Life'Images: detection,*
787 *alignment, and recognition*, 2008. 6
- 788 [16] Shan Li and Weihong Deng. Deep facial expression recogni-
789 tion: A survey. *IEEE transactions on affective computing*,
790 13(3):1195–1215, 2020. 2
- 791 [17] Blaž Meden, Peter Rot, Philipp Terhörst, Naser Damer, Ar-
792 jan Kuijper, Walter J Scheirer, Arun Ross, Peter Peer, and
793 Vitomir Štruc. Privacy-enhancing face biometrics: A com-
794 prehensive survey. *IEEE Transactions on Information Foren-*
795 *sics and Security*, 16:4147–4183, 2021. 1
- 796 [18] Qiang Meng, Shichao Zhao, Zhida Huang, and Feng Zhou.
797 Magface: A universal representation for face recognition and
798 quality assessment. In *Proceedings of the IEEE/CVF conference*
799 *on computer vision and pattern recognition*, pages
800 14225–14234, 2021. 2
- 801 [19] Stylianos Moschoglou, Athanasios Papaioannou, Chris-
802 tos Sagonas, Jiankang Deng, Irene Kotsia, and Stefanos
803 Zafeiriou. Agedb: the first manually collected, in-the-wild
804 age database. In *proceedings of the IEEE conference on*
805 *computer vision and pattern recognition workshops*, pages
806 51–59, 2017. 3, 4, 6
- 807 [20] Fu-Zhao Ou, Xingyu Chen, Ruixin Zhang, Yuge Huang,
808 Shaoxin Li, Jilin Li, Yong Li, Liujuan Cao, and Yuan-Gen
809 Wang. Sdd-fqa: unsupervised face image quality assess-
810 ment with similarity distribution distance. In *Proceedings*
811 *of the IEEE/CVF conference on computer vision and pattern*
812 *recognition*, pages 7670–7679, 2021. 2
- 813 [21] P Jonathon Phillips, Hyeonjoon Moon, Syed A Rizvi, and
814 Patrick J Rauss. The feret evaluation methodology for face-
815 recognition algorithms. *IEEE Transactions on pattern analysis*
816 *and machine intelligence*, 22(10):1090–1104, 2000. 6
- 817 [22] Torsten Schlett, Christian Rathgeb, Olaf Henniger, Javier
818 Galbally, Julian Fierrez, and Christoph Busch. Face image
819 quality assessment: A literature survey. *ACM Computing*
820 *Surveys (CSUR)*, 54(10s):1–49, 2022. 1
- 821 [23] Florian Schroff, Dmitry Kalenichenko, and James Philbin.
822 Facenet: A unified embedding for face recognition and clus-
823 tering. In *Proceedings of the IEEE conference on computer*
824 *vision and pattern recognition*, pages 815–823, 2015. 3
- 825 [24] Caifeng Shan, Shaogang Gong, and Peter W McOwan. Fa-
826 cial expression recognition based on local binary patterns:
827 A comprehensive study. *Image and vision Computing*,
828 27(6):803–816, 2009. 2
- 829 [25] Gillian Slessor, Deborah M Riby, and Ailbhe N Finnerty.
830 Age-related differences in processing face configuration:
831 The importance of the eye region. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*,
832 68(2):228–231, 2013. 2
- 833 [26] Philipp Terhorst, Jan Niklas Kolf, Naser Damer, Florian
834 Kirchbuchner, and Arjan Kuijper. Ser-fiq: Unsupervised esti-
835 mation of face image quality based on stochastic embedding
836 robustness. In *Proceedings of the IEEE/CVF conference on*
837 *computer vision and pattern recognition*, pages 5651–5660,
838 2020. 2
- 839 [27] VGG. The visual geometry group (vgg) at the university of
840 oxford. 3
- 841 [28] Mei Wang and Weihong Deng. Deep face recognition: A
842 survey. *Neurocomputing*, 429:215–244, 2021. 1
- 843 [29] Weidi Xie, Jeffrey Byrne, and Andrew Zisserman. Inducing
844 predictive uncertainty estimation for face recognition. *arXiv*
845 *preprint arXiv:2009.00603*, 2020. 1, 2