

Iterativna optimizacija ocen kakovosti slikovnih podatkov v sistemih za razpoznavanje obrazov

Žiga Babnik, Vitomir Štruc

Univerza v Ljubljani, Fakulteta za Elektrotehniko, Tržaška cesta 25, 1000 Ljubljana, Slovenija
{ziga.babnik, vitomir.struc}@fe.uni-lj.si

Iterative Optimization of Pseudo Ground-Truth Face Image Quality Labels

While recent face recognition (FR) systems achieve excellent results in many deployment scenarios, their performance in challenging real-world settings is still under question. For this reason, face image quality assessment (FIQA) techniques aim to support FR systems, by providing them with sample quality information that can be used to reject poor quality data unsuitable for recognition purposes. Several groups of FIQA methods relying on different concepts have been proposed in the literature, all of which can be used for generating quality scores of facial images that can serve as pseudo ground-truth (quality) labels and be exploited for training (regression-based) quality estimation models. Several FIQA approaches show that a significant amount of sample-quality information can be extracted from mated similarity-score distributions generated with some face matcher. Based on this insight, we propose in this paper a quality label optimization approach, which incorporates sample-quality information from mated-pair similarities into quality predictions of existing off-the-shelf FIQA techniques. We evaluate the proposed approach using three state-of-the-art FIQA methods over three diverse datasets. The results of our experiments show that the proposed optimization procedure heavily depends on the number of executed optimization iterations. At ten iterations, the approach seems to perform the best, consistently outperforming the base quality scores of the three FIQA methods, chosen for the experiments.

1 Uvod

Zaradi preteklega razvoja sistemov za razpoznavanje obrazov, le-ti dosegajo izvrstne rezultate [8], tudi na večjih, ter bolj težavnih podatkovnih zbirkah obraznih slik, kot je recimo zbirka IARPA Janus Benchmark-C (IJB-C) [12]. Prenos teh izjemnih rezultatov v realni svet, za naloge kot je video nadzor, je zaradi slabe kakovosti slikovnih podatkov zaenkrat še neuresničljiv. Posledično se je za lažje uresničevanje doseganja dobrih in zanesljivih rezultatov obraznih razpoznavalnikov razvilo raziskovalno področje ocene kakovosti obraznih slik, ki skuša

oceniti biometrično kakovost vzorca za namene razpoznavanja [10]. Pri tem se biometrična kakovost pogosto definirana kot *koristnost oz. primernost* vzorca za namene obdelave v sodobnih sistemih za razpoznavanje obrazov [14]. Biometrična kakovost je torej tesno povezana z vizualno kakovostjo vzorca, a ji ni povsem enaka.

Obstaja več skupin pristopov ocenjevanja (biometrične) kakovosti obraznih slik (angl. Face Image Quality Assessment - FIQA). Najbolj razširjena skupina se poslužuje ustvarjanja psevdo referenčnih vrednosti kakovosti večjega nabora vzorcev, ustvarjene referenčne vrednosti pa kasneje uporabi za učenje regresijskih modelov, sposobnih samostojnega napovedovanja ocen kakovosti [9, 6, 17, 4]. V zadnjem času se pojavljajo tudi metode, ki združujejo nalogo razpoznavanja obrazov in ocenjevanja kakovosti [13, 15], ter analitični postopki, ki za oceno kakovosti uporabijo zgolj karakteristike vhodnega vzorca in lastnosti izbranega razpoznavalnika [16, 1].

Pristope iz katerekoli zgoraj omenjene skupine je mogoče uporabiti tudi za razvoj FIQA metod z nadzorovanim učenjem. V tem primeru za večji nabor obraznih slik s pomočjo izbrane FIQA metode pridobimo ocene (psevdo) referenčnih oznak kakovosti, le-te pa nato uporabimo za učenje modela za ocenjevanje kakovosti. Pri tem lahko pridobljene ocene kakovosti pred učenjem obogatimo in izboljšamo s pomočjo dodatnih zunanjih informacij, kot je recimo porazdelitev podobnosti med ujemajočimi se pari obraznih slik (v smislu identitete). Mnogi obstoječi FIQA pristopi namreč kažejo na dejstvo, da podobnost takšnih parov vsebuje precejšnjo količino informacij o sami kakovosti posameznih vzorcev [6, 4]. Na podlagi predstavljene ideje v tem članku predstavimo pristop za (iterativno) izboljševanje začetnih ocen kakovosti obraznih slik, ustvarjenih z izbranim FIQA modelom, ki temelji na vključevanju dodatnih informacij, pridobljenih iz primerjav ujemajočih se slik obrazov.

2 Pregled področja

V tem razdelku predstavimo tri glavne skupine FIQA metod, ki jih lahko razdelimo v: analitične, regresijske ter mrežne pristope. Podrobnejši pregled področja je predstavljen v nedavnem preglednem članku [14].

Analitični pristopi temeljijo na izločanju ocene kakovosti iz informacije prisotne v samem vzorcu. Zaradi tega se pretežno osredotočajo na vizualno kakovost in

Podprto s strani ARRS raziskovalnega programa P2-0250 (B), ter ARRS programom mladih raziskovalcev.

pogostjo ne dosegajo konkurenčnih rezultatov v primerjavi z najnaprednejšimi rešitvami iz literature. Starješi pristop, ki so ga predstavili Gao *et al.* [7] poskuša pridobiti oceno kakovosti slikovnih podatkov z ocenjevanjem obrazne simetrije. Pred kratkimi sta se pojavila dva pristopa, ki poleg karakteristik slik hkrati upoštevata tudi informacijo izbranega razvrščevalnika, in tako dosegata vrhunske rezultate. Prvi pristop, ki so ga predstavili Terhörst *et al.* [16], se zanaša na uporabo izpustnih slojev mrež, medtem ko drugi, predlagan s strani Babnika *et al.* [1], za oceno kakovosti izrablja nasprotniške pristope.

Regressijski pristopi predstavljajo najštevilčnejšo skupino FIQA metod, ki temeljijo na učenju regresijskih modelov za ocenjevanje kakovosti iz pridelanih psevdoreferenčnih oznak kakovosti vzorcev. Eden izmed začetnih postopkov iz te skupine, Ortege *et al.* [9], uči regresijsko nevronska mrežo, z uporabo oznak, pridobljenih s primerjavo vložitev najkakovostnejši sliki vsakega posameznika. Pri tem se najkakovostnejše slike posameznikov izloči s pomočjo zunanega orodja za preverjanje kakovosti. Naprednejši pristop, ki so ga predlagali Ou *et al.* [6], za napoved referenčnih oznak kakovosti uporabi podobnosti tako ujemajočih kot tudi neujemajočih se parov obraznih slik in s tem pridela koristnejše ocene kakovosti za učenje regresijskega FIQA modela.

Mrežni pristopi največkrat združijo nalogi razpoznavanja ter ocenjevanja kakovosti in se naučijo preslikave vhodnih slik v posebne vložitve, iz katerih je možno izločiti informacijo o identiteti, kot tudi informacijo o kakovosti vzorca. Starejši pristop avtorjev Shi in Jain [15] se nauči napovedati dvojice vektorjev, kjer prvi – povprečni vektor vsebuje informacijo o identiteti, drugi – vektor odklona, iz katerega je možno izraziti kakovost vzorca, pa se navezuje na nedoločeno povprečnega vektorja. Novejši pristop, predlagan s strani Meng *et al.* [13], za potrebe učenja razpoznavalnika obrazov uporabi prirejeno ArcFace funkcijo izgube, ki vključuje tudi informacijo o kakovosti vhodnega vzorca.

3 Metodologija

Poglavitna naloga metod za oceno kakovosti obraznih slik je zajetje čim večje količine informacij o koristnosti oz. primernosti vzorca za biometrično razpoznavanje v končno oceno kakovosti. Mnoge raziskave kažejo na dejstvo, da podobnosti ujemajočih se parov obraznih slik vsebujejo velike količine informacij o kakovosti vzorca, zato v tem razdelku predstavimo pristop, ki na podlagi teh informacij, iterativno optimizira predhodne napovedi kakovosti slikovnih vzorcev.

3.1 Pregled pristopa

Predpostavimo poljuben FIQA pristop za oceno kakovosti obraznih slik F_q , ki za podani obrazni vzorec I_i vrne oceno kakovosti $q_{I_i} = F_q(I_i)$, večjo podatkovno zbirko obraznih slik $\mathcal{I} = \{I_i\}_{i=1}^n$, ki vsebuje slike $N \leq n$ različnih posameznikov, ter ciljni obrazni razpoznavalnik M . Cilj našega pristopa je pridobiti optimizirane ocene kakovosti $\mathcal{Q}^* = \{q_{I_i}^*\}_{i=1}^n$, z iterativnim posodabljanjem osnovnih ocen kakovosti $\mathcal{Q} = \{q_{I_i}\}_{i=1}^n$, ki nam jih vrne

pristop za oceno kakovosti F_q , na podlagi informacij pridobljenih iz porazdelitve podobnosti ujemajočih se slik obrazov $\mathcal{D} = \{d_i\}_{i=1}^m$, kjer je m število upoštevanih primerjav slik.

3.2 Inicializacija pristopa

Preden lahko izvedemo optimizacijo ocen, jih je potrebno najprej pridobiti. Zato uporabimo izbran pristop F_q in ga izvedemo nad celotno zbirko podatkov \mathcal{I} . Tako pridobimo osnovne ocene $\mathcal{Q} = F_q(\mathcal{I})$, ki jih nato normiramo na interval $[0, 1]$. Prav tako potrebujemo vnaprej ustvarjene vložitve celotne zbirke \mathcal{I} , tj. $\mathcal{E} = M(\mathcal{I})$, ki jih pridobimo s pomočjo razpoznavalnika obrazov M .

Na podlagi izbranega nabora \mathcal{I} in pripadajočega seznama identitet $\{C_i\}_{i=1}^N$, zgradimo seznam ujemajočih se slik, tako da za vsak vzorec iz zbirke izberemo k ključnih vzorcev iste identitete. Za posamezen pristni par vzorcev (I_i, I_j) , kjer velja $I_i, I_j \in C_k$, nato izračunamo podobnost z uporabo kosinusne podobnosti:

$$s_{cos}(I_i, I_j) = \frac{e_i \cdot e_j}{\|e_i\| \cdot \|e_j\|}, \quad (1)$$

kjer sta $e_i = M(I_i)$ ter $e_j = M(I_j)$ vložitvi prej omejenih vzorcev. Izračunane podobnosti ležijo na intervalu $[-1, 1]$, zato jih podobno kot ocene kakovosti normiramo na interval $[0, 1]$. Z izračunom podobnosti med vsemi pristnimi dvojicami lahko ustvarimo porazdelitev podobnosti pristnih dvojic \mathcal{D} . Za lažje razumevanje nadaljnjih korakov je potrebno izpostaviti še definicijo kakovosti para slik (I_i, I_j) , ki je pogosto uporabljena na področju ocenjevanja kakovosti:

$$q_p(I_i, I_j) = \min(q_{I_i}, q_{I_j}), \quad (2)$$

ki pravi, da je kakovost para kar enaka kakovosti slabše slike v paru.

3.3 Iterativna optimizacija

Po inicializaciji podobnosti \mathcal{D} pristnih dvojic lahko izvedemo posamezno iteracijo predlaganega pristopa. Prvi korak vključuje izračun korekcijskega faktorja θ_{I_i} vzorca I_i :

$$\theta_{I_i} = \frac{1}{k} \sum_{j=0}^k d(I_i, I_j) \cdot s_{cos}(I_i, I_j), \quad (3)$$

kjer je (I_i, I_j) pristni par, $d(I_i, I_j)$ pa predstavlja preprosto indikatorsko funkcijo

$$d(I_i, I_j) = \begin{cases} 1, & q_{I_i} \leq q_{I_j} - \lambda \\ 0, & \text{else} \end{cases}, \quad (4)$$

ki vrača 1 v primeru, ko je kakovost prvega vzorca manjša kot kakovosti drugega in je razlika med njima večja od vrednosti vhodnega parametra metode λ . Kot prikazuje Enačba (2), pri izračunu kakovosti dvojice uporabimo le vrednost manj kakovostne slike, zato v paru upoštevamo samo nižje kakovosten vzorec pri izračunu korekcijskega faktorja θ_{I_i} . S pomočjo izračunanega korekcijskega faktorja lahko posodobimo trenutno kakovost vzorca posamezne slike I_i kot

$$q_{I_i}^{j+1} = q_{I_i}^j + \epsilon \cdot (\theta_{I_i} - q_{I_i}^j), \quad (5)$$

kjer je j trenutna iteracija, ter ϵ vhodni parameter pristopa, ki določa velikost premika posamezne iteracije. Po izvedbi L iteracij, pridobimo nove, optimizirane ocene kakovosti $\mathcal{Q}^* = \{q_{I_i}^*\}_{i=1}^n$, kjer je $q_{I_i}^* = q_{I_i}^L$.

Ker lahko z naključnim izbiranjem pristnih dvojic v pristop vnesemo pristranskost, celoten postopek ponovimo v T ponovitvah. Tako končno oceno kakovosti izračunamo kot

$$\mathcal{Q}_{I_i} = \frac{1}{T} \sum_{t=0}^T \mathcal{Q}_{I_i}^{t*}, \quad (6)$$

kjer je $\mathcal{Q}_{I_i}^{t*}$ množica optimiziranih ocen kakovosti t -te ponovitve postopka.

4 Eksperimenti in rezultati

4.1 Zasnova eksperimentov

Modeli in podatki. Za začetne ocene kakovosti obraznih slik smo uporabili tri napredne FIQA pristope: CR-FIQA¹ [2], FaceQAN² [1] ter SDD-FIQA³ [6]. Za potrebe učenja in testiranja smo izbrali obrazni razpoznavnik ArcFace⁴ [5], ki temelji na ResNet100 arhitekturi, učen z izgubo kotnega razmaka. Za učenje smo izbrali zbirko VGGFace2 [3], ki vsebuje približno 3 milijone slik, 9131 različnih identitet, za testiranje pa zbirke: CPLFW [18], CALFW [19] ter XQLFW [11]. Podatkovna zbirka CPLFW se osredotoča na razlike v obrazni pozi, CALFW na razlike v starosti, XQLFW pa vsebuje obraze s širokim razponom v (vizualni) kakovosti slik.

Ovrednotenje zmogljivosti. Omenjeni razpoznavnik ArcFace smo priredili za potrebe ocenjevanja kakovosti tako, da smo dodali regresijsko glavo in ga učili s pomočjo L_1 funkcije izgube. Učenje smo izvedli dvakrat, prvič na osnovnih/začetnih ocenah kakovosti in drugič na optimiziranih ocenah kakovosti vseh pristopov. Za potrebe primerjave zmogljivosti mrež učenih iz osnovnih in optimiziranih ocen smo uporabili ustaljen pristop krivulj zmote zoper zavrnitve, ki merijo stopnjo lažno negativnih parov pri vnaprej določeni vrednosti stopnje lažno pozitivnih parov kot funkcijo različnih stopenj zavrnitve slik zaradi nizke kakovosti [1]. Za pridobljene krivulje izračunamo ploščino pod krivuljo, kjer manjša vrednost namiguje na boljše ocene kakovosti. Za namene ovrednotenja optimizacijskega postopka, v članku prikažemo ploščino pod krivuljo pri različnih stopnjah zavrnitve.

Podrobnosti o implementaciji. Za zagotavljanje robustnih ocen kakovosti, smo s pomočjo preliminarne eksperimentov določili $T = 10$ za število ponovitev, pri čemer za vsako sliko v posamezni ponovitvi zgradimo $N = 10$ pristnih parov. Hkrati tako zagotovimo robustnost ter omejimo časovno kompleksnost pristopa. Cilj samega pristopa je le natančnejša prilagoditev začetnih ocen kakovosti, zato izberemo $\epsilon = 0.01$ in tako močno omejimo korak posamezne iteracije metode. Določimo še $\lambda = 0.05$ in tako dodatno kaznujemo pare slik, kjer sta

kvaliteti pod zastavljeno mejo. Za število iteracij L izberemo tri različne vrednosti $L \in [5, 10, 15]$, saj nas zanima učinek povečevanja števila iteracij na končne rezultate.

4.2 Primerjava dobljenih rezultatov

Tabela 1 prikazuje rezultate eksperimentov, natančneje, vrednosti površine pod krivuljo za vse tri izbrane pristope ocenjevanja kakovosti obraznih slik. Za realne aplikacije so bolj pomembni rezultati pri manjših deležih zavrnjenih slik, saj zaradi težav s kakovostjo ne želimo zavrniti prevelikega deleža slik. V tabeli 1 so zato prikazane vrednosti površin pri 10%, 20%, 40% ter 80% zavrnjenih slik. Posamezne vrstice rezultatov so označene s številom izvedenih iteracij, 0 iteracij predstavlja osnovne rezultate pristopa brez optimizacije. Za lažjo predstavitev in interpretacijo rezultatov predstavimo tudi povprečno vrednost za vse stopnje izpuščenih slik – označeno z μ .

FaceQAN. Opazimo lahko, da FaceQAN najboljše rezultate v povprečju doseže pri desetih iteracijah, kjer pridobimo najboljše rezultate za zbirkah CPLFW ter CALFW. Na zbirki XQLFW pa boljše rezultate pristop doseže le še pri petih iteracijah. Kljub rezultatom na bazi XQLFW, metoda v povprečju preko vseh podatkovnih zbirk pridela najslabše rezultate prav pri petih iteracijah, ki so slabši tudi od osnovnih ocen kakovosti metode.

CR-FIQA. Rezultati pristopa CR-FIQA preko različnih podatkovnih zbirk kažejo zelo raznolike rezultate, ki so močno odvisni od karakteristik testnih podatkov. Za bazo CPLFW dosegajo najboljše rezultate ocene pri petnajstih iteracijah, za CALFW pri desetih in za XQLFW pri petih iteracijah. Z izjemo desetih iteracij na bazi CPLFW, dosegajo ocene desetih in petnajstih iteracij dosledno boljše rezultate kot osnovne ocene. Kljub izjemi na bazi CPLFW lahko opazimo, da najboljše rezultate v povprečju preko vseh zbirk dosegajo ocene desetih iteracij.

SDD-FIQA. Rezultati pristopa SDD-FIQA kažejo, da predstavljena metoda optimizacije uspe izboljšati ocene kakovosti v vseh primerih razen na podatkovni bazi CPLFW v primeru petnajstih iteracij, kjer je rezultat slabši od osnovnih ocen pristopa. Kljub temu, da v primeru desetih iteracij ne dosežemo najboljšega rezultata za nobeno izmed podatkovnih zbirk, so rezultati v tem primeru najboljše. Tesno sledijo rezultati petih in petnajstih iteracij, najslabše rezultate pa dosegajo osnovne ocene metode.

Primerjava med pristopi. Iz rezultatov je razvidno, da lahko z izbiro pravega števila iteracij pridobimo boljše ocene kakovosti na posamezni podatkovni zbirki, kot jih zagotavljajo osnovni FIQA pristopi. Izboljšava je sicer majhna, kar ni nepričakovano, saj je namen pristopa le optimizacija predhodnih ocen kakovosti vzorcev, a je lahko pomembna za realne sisteme razpoznavanja obrazov. V primeru izbire desetih iteracij, dobimo za skoraj vse kombinacije pristopov in podatkovnih zbirk boljše rezultate v primerjavi z osnovnimi ocenami. Medtem, ko so rezultati pri petih in petnajstih iteracijah močno odvisni od izbrane zbirke. V primeru zbirke XQLFW je očitna boljše izbira manjšega števila iteracij, za zbirko CALFW pa večje število iteracij.

¹<https://github.com/fdbtrs/CR-FIQA>

²<https://github.com/LSIbabrikz/FaceQAN>

³<https://github.com/Tencent/TFace/tree/quality>

⁴<https://github.com/deepinsight/insightface>

Tabela 1: Primerjava vrednosti površine pod krivuljo (\downarrow) za vse izbrane metode na treh izbranih podatkovnih zbirkah. Najslabši rezultat znotraj posamezne podatkovne zbirke, metode in deleža zavrnjenih slik je obarvan rdeče, najboljši zeleno.

M [†]	I [‡]	CPLFW					CALFW					XQLFW				
		10%	20%	40%	80%	μ	10%	20%	40%	80%	μ	10%	20%	40%	80%	μ
FaceQAN	0	1.417	2.226	3.511	5.155	3.077	1.512	2.737	4.594	7.939	4.195	5.148	8.882	13.2	16.325	10.888
	5	1.424	2.229	3.555	5.303	3.127	1.524	2.782	4.845	8.139	4.322	5.14	8.56	12.998	16.046	10.686
	10	1.407	2.203	3.492	4.788	2.972	1.523	2.72	4.587	7.941	4.192	5.076	8.75	13.083	16.246	10.788
	15	1.414	2.201	3.489	4.891	2.998	1.532	2.847	4.69	8.113	4.295	5.169	8.883	13.143	16.199	10.848
CR-FIQA	0	1.408	2.193	3.131	4.373	2.776	1.485	2.773	5.034	8.327	4.404	5.136	9.171	14.874	18.333	11.878
	5	1.41	2.197	3.133	4.408	2.787	1.486	2.77	5.052	8.328	4.409	5.135	9.169	14.074	17.543	11.480
	10	1.404	2.191	3.327	4.577	2.874	1.473	2.754	4.916	8.008	4.287	5.266	9.297	14.184	17.435	11.545
	15	1.408	2.192	3.105	4.334	2.759	1.477	2.74	4.873	8.086	4.294	5.277	9.436	14.456	17.928	11.774
SDD-FIQA	0	1.517	2.467	3.47	4.615	3.017	1.465	2.719	4.922	8.248	4.338	5.214	9.414	16.695	22.968	13.572
	5	1.505	2.431	3.404	4.586	2.981	1.45	2.705	4.803	8.375	4.333	5.165	9.406	16.19	21.711	13.118
	10	1.482	2.372	3.544	4.671	3.017	1.466	2.715	4.874	8.122	4.294	5.43	9.756	16.07	21.498	13.188
	15	1.483	2.382	3.587	4.756	3.052	1.461	2.705	4.725	8.08	4.242	5.439	9.813	15.965	21.748	13.241

M[†] - FIQA metoda, I[‡] - število iteracij

- najboljši rezultat, - najslabši rezultat

5 Zaključek

V članku smo predstavili izviren postopek optimizacije predhodno izračunanih ocen kakovosti obraznih slik, ki skuša v začetni oceni kakovosti vnesti dodatno informacijo, pridobljeno iz distribucije podobnost ujemačih se parov slik. Rezultati eksperimentov s tremi metodami za oceno kakovosti in tremi podatkovnimi zbirkami (CPLFW, CALFW in XQLFW) kažejo na močno odvisnost uspešnosti postopka od izbire števila iteracij optimizacije. V primeru manjšega števila iteracij, so rezultati za podatkovni zbirki CPLFW ter CALFW slabši od osnovnih ocen kakovosti. V primeru prevelikega števila iteracij se zgodba ponovi za zbirko XQLFW. Najboljše rezultate dosledno ustvari pristop pri desetih iteracijah, kjer so rezultati večinoma boljši od osnovnih ocen izbranih pristopov.

Literatura

- [1] Ž. Babnik, P. Peer, and V. Štruc. FaceQAN: Face Image Quality Assessment Through Adversarial Noise Exploration. In *ICPR*, 2022.
- [2] F. Boutros, M. Fang, M. Klemt, B. Fu, and N. Damer. Crfiqa: Face image quality assessment by learning sample relative classifiability, 2021.
- [3] Q. Cao, L. Shen, W. Xie, O. M. Parkhi, and A. Zisserman. Vggface2: A dataset for recognising faces across pose and age. In *IEEE international conference on automatic face & gesture recognition (FG 2018)*, pages 67–74, 2018.
- [4] K. Chen, T. Yi, and Q. Lv. Lightqnet: Lightweight deep face quality assessment for risk-controlled face recognition. *IEEE Signal Processing Letters*, 28, 2021.
- [5] J. Deng, J. Guo, N. Xue, and S. Zafeiriou. Arcface: Additive Angular Margin Loss for Deep Face Recognition. In *CVF/IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4690–4699, 2019.
- [6] O. Fu-Zhao, X. Chen, R. Zhang, Y. Huang, S. Li, J. Li, Y. Li, L. Cao, and W. Yuan-Gen. SDD-FIQA: Unsupervised Face Image Quality Assessment with Similarity Distribution Distance. In *CVF/IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021.
- [7] X. Gao, S. Z. Li, R. Liu, and P. Zhang. Standardization of face image sample quality. In *International Conference on Biometrics*, pages 242–251. Springer, 2007.
- [8] K. Grm, V. Štruc, A. Artiges, M. Caron, and H. K. Ekel. Strengths and weaknesses of deep learning models for face recognition against image degradations. *IET Biometrics*, 7(1):81–89, 2018.
- [9] J. Hernandez-Ortega, J. Galbally, J. Fierrez, R. Haraksim, and L. Beslay. Faceqnet: Quality assessment for face recognition based on deep learning. In *Proceedings of the IEEE International Conference on Biometrics (ICB)*, pages 1–8, 2019.
- [10] ISO/IEC JTC 1/SC 37 Biometrics. Information Technology - Biometric Sample Quality - Part 1: Framework. Standard ISO/IEC 29794-1:2016, International Organization for Standardization, 2016.
- [11] M. Knoche, S. Hoermann, and G. Rigoll. Cross-quality lfw: A database for analyzing cross-resolution image face recognition in unconstrained environments. In *2021 16th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2021)*, pages 1–5, 2021.
- [12] B. Maze, J. Adams, J. A. Duncan, N. Kalka, T. Miller, C. Otto, A. K. Jain, W. T. Niggel, J. Anderson, J. Cheney, et al. Iarpa janus benchmark-c: Face dataset and protocol. In *International Conference on Biometrics (ICB)*, pages 158–165, 2018.
- [13] Q. Meng, S. Zhao, Z. Huang, and F. Zhou. Magface: A universal representation for face recognition and quality assessment. In *CVF/IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 14225–14234, 2021.
- [14] T. Schlett, C. Rathgeb, O. Henniger, J. Galbally, J. Fierrez, and C. Busch. Face image quality assessment: A literature survey. *ACM Computing Surveys*, 2022.
- [15] Y. Shi and A. K. Jain. Probabilistic face embeddings. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pages 6902–6911, 2019.
- [16] P. Terhorst, J. N. Kolf, N. Damer, F. Kirchbuchner, and A. Kuijper. SER-FIQ: Unsupervised Estimation of Face Image Quality Based on Stochastic Embedding Robustness. In *CVF/IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5651–5660, 2020.
- [17] W. Xie, J. Byrne, and A. Zisserman. Inducing predictive uncertainty estimation for face verification. In *British Machine Vision Conference (BMVC)*, 2020.
- [18] T. Zheng and W. Deng. Cross-pose lfw: A database for studying cross-pose face recognition in unconstrained environments. Technical Report 18-01, Beijing University of Posts and Telecommunications, February 2018.
- [19] T. Zheng, W. Deng, and J. Hu. Cross-age LFW: A database for studying cross-age face recognition in unconstrained environments. *CoRR*, abs/1708.08197, 2017.