



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ

FACULTY OF INFORMATION TECHNOLOGY

ÚSTAV POČÍTAČOVÉ GRAFIKY A MULTIMÉDIÍ

DEPARTMENT OF COMPUTER GRAPHICS AND MULTIMEDIA

ODHAD KVALITY SNÍMKŮ DUHOVKY PRO IDENTIFIKACI OSOB

IRIS IMAGE QUALITY ASSESSMENT

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. MAREK VAŠKO

VEDOUcí PRÁCE

SUPERVISOR

Ing. MICHAL HRADIŠ, Ph.D.

BRNO 2022

Zadání diplomové práce



Student: **Vaško Marek, Bc.**
Program: Informační technologie a umělá inteligence
Specializace: Inteligentní systémy
Název: **Odhad kvality snímků duhovky pro identifikaci osob**
Iris Image Quality Assessment
Kategorie: Zpracování obrazu
Zadání:

1. Prostudujte oblast měření kvality obrazů duhovky.
2. Prozkoumejte aktuálně používané metody, některé vyberte zkuste je aplikovat.
3. Aplikujte metodu měření kvality na metodu rozpoznání duhovek (algoritmus dodá Innovatrics).
4. Vyberte datové sady vhodné na experimenty.
5. Porovnejte některé metody odhadu kvality z pohledu jejich vlivu na schopnost rozpoznávání.
6. Interpretujte dosažené výsledky a diskutujte možnosti budoucího vývoje.
7. Vytvořte stručné video prezentující vaši práci, její cíle a výsledky.

Literatura:

1. Kalka, Nathan & Zuo, Jinyu & Schmid, Natalia & Cukic, Bojan. Image quality assessment for iris biometric. Proceedings of SPIE - The International Society for Optical Engineering. 6202., 2006.
2. L Wang, K Zhang, M Ren, Y Wang. Recognition Oriented Iris Image Quality Assessment in the Feature Space. IEEE International Joint Conference on Biometrics (IJCB), 2020.
3. Aditya Abhyankar, Stephanie Schuckers. Iris Quality Assessment and Bi-orthogonal Wavelet Based Encoding for Recognition, Pattern Recognition, 2009.
4. Philipp Terhörst, Jan Niklas Kolf, Naser Damer, Florian Kirchbuchner, Arjan Kuijper. SER-FIQ: Unsupervised Estimation of Face Image Quality Based on Stochastic Embedding Robustness, CVPR, 2020.
5. NIST Iris Quality Calibration and Evaluation, <https://nvlpubs.nist.gov/nistpubs/Legacy/IR/nistir7820.pdf>
6. Lubos Omelina, Jozef Goga, Jarmila Pavlovicova, Milos Oravec, Bart Jansena, A survey of iris datasets, Image and Vision Computing, 2021.

Při obhajobě semestrální části projektu je požadováno:

- Body zadání 1-3.

Podrobné závazné pokyny pro vypracování práce viz <https://www.fit.vut.cz/study/theses/>

Vedoucí práce: **Hradiš Michal, Ing., Ph.D.**
Konzultant: Beneš Radek, Ing., FINGERA
Vedoucí ústavu: Černocký Jan, doc. Dr. Ing.
Datum zadání: 1. listopadu 2021
Datum odevzdání: 18. května 2022
Datum schválení: 1. listopadu 2021

Abstrakt

Identifikácia osôb pomocou dúhovky predstavuje jeden z najpresnejších spôsobov biometrickej identifikácie. Pri všetkých biometrických systémoch je možné zapríčiniť rôzne chyby verifikácie nekvalitným vstupom. Odhadom kvality je možné odstrániť práve nekvalitné vstupy. V oblasti odhadu kvality snímok dúhovky je pomerne nevýrazný vývoj a veľa metód, ktoré by potenciálne mohli byť využité zatiaľ nie sú v tejto oblasti vyskúšané. Táto práca sa zameriava na použitie rôznych metód odhadu kvality využívaných pri rozpoznávaní tváří. Tieto metódy odhadu kvality sú následne v tejto práci aplikované na oblasť identifikácie pomocou dúhovky. V riešení sú využité verifikačné systémy založené na rôznych konfiguráciách architektúr iResNet a MobileNetV3. Na jednotlivé systémy sú aplikované vybrané metódy odhadu kvality. Jednotlivé metódy odhadu kvality buď to systém priamo trénujú alebo využívajú jeho výstupy pre získanie informácií o kvalite. Aplikované metódy odhadu kvality sú porovnané na štyroch rôznych dátových sadách. Na niektorých dátových sadách ako napríklad University of Notre Dame Iris CrossSensor 2013 dosahuje výsledný systém s aplikovaním najlepšej metódy odhadu kvality pri 10% odmietnutých vstupných pároch a false match rate 0.1% zníženie verifikačnej chyby false non-match rate až o 56% na hodnotu 0.5% pri iResNet50 a pri systéme založenom na MobileNetV3 až o 22% na hodnotu 6.4%.

Abstract

Iris image recognition is one of the most accurate ways of biometric identification. Various verification errors can be caused if the biometric system receives poor input. By assessing the image quality it is possible to eliminate inputs causing such errors. There is a relatively insignificant development in the field of iris quality assessment and many methods that could potentially be used have not been tested in this area yet. This work focuses on different quality assessment methods used in face recognition. These quality assessment methods are then applied to the area of iris identification. The solution uses verification systems based on various iResNet and MobileNetV3 architectures. Selected quality assessment methods are applied to individual systems. Different quality assessment methods train either the system directly or use its outputs to obtain information about quality. The resulting system achieves a reduction of false non-match rate by up to 56% with the absolute value of 0.5% for iResNet50 and up to 22% with the absolute value of 6.4% for MobileNetV3 when using the best quality assessment method. The results are given for the data set University of Notre Dame Iris CrossSensor 2013 with an input reject rate of 10% and a false match rate of 0.1%.

Kľúčové slová

biometria, CR-FIQA, MagFace, neurónové siete, odhad kvality snímok dúhovky, PyTorch, PFE, rozpoznávanie snímok dúhovky, SDD-FIQA, SER-FIQ, strojové učenie

Keywords

biometrics, CR-FIQA, MagFace, neural networks, iris image quality assessment, PyTorch, PFE, iris image recognition, SDD-FIQA, SER-FIQ, machine learning

Citácia

VÁŠKO, Marek. *Odhad kvality snímku dúhovky pro identifikaci osob*. Brno, 2022. Diplomová práce. Vysoké učení technické v Brně, Fakulta informačních technologií. Vedoucí práce Ing. Michal Hradiš, Ph.D.

Odhad kvality snímků duhovky pro identifikaci osob

Prehlásenie

Prehlasujem, že som túto magisterskú prácu vypracoval samostatne pod vedením pána Ing. Michala Hradiša Ph.D. Ďalšie informácie mi poskytol Ing. Radek Beneš z firmy Innovatrics. Uviedol som všetky literárne pramene, publikácie a ďalšie zdroje, z ktorých som čerpal.

.....
Marek Vaško
18. mája 2022

Poďakovanie

Špeciálne poďakovanie patrí vedúcemu práce Ing. Michalovi Hradišovi Ph.D., Ing. Radkovi Benešovi a firme Innovatrics s.r.o. za vedenie a odbornú pomoc s vypracovaním práce. Zároveň poďakovanie patrí aj MetaCentru za poskytnutie výpočtových zdrojov.

Obsah

1	Úvod	3
2	Odhad kvality snímok dúhovky	4
2.1	Štandardné kvalitatívne komponenty	4
2.2	Určenie kvality snímky dúhovky strojovým učením	6
2.3	Metriky vyhodnocovania biometrického systému	7
3	Identifikácia pomocou snímky dúhovky	10
3.1	Architektúry konvolučných neurónových sietí	10
3.2	Tréning výstupu siete pre vektor vlastností	12
4	Všeobecný odhad kvality snímok	13
4.1	Pravdepodobnostná reprezentácia vektora príznakov	13
4.2	Odhad kvality z distribučných funkcií podobností	15
4.3	Odhad kvality pomocou robustnosti vektorov príznakov	16
4.4	Odhad kvality podľa relatívnej klasifikovateľnosti	18
4.5	MagFace	20
5	Dátové sady	21
5.1	Trénovacia dátová sada	23
5.2	Verifikačné dátové sady	23
5.3	Označenie pozície dúhovky	24
6	Návrh riešenia a experimentov	25
6.1	Transformácia snímok	27
6.2	Modul odhadu kvality snímok	28
6.3	Určenie kvality páru snímok	31
6.4	Modul verifikačných neurónových sietí	31
6.5	Určenie rozličnosti páru snímok	32
6.6	Augmentácia trénovacích dát	32
7	Implementácia	33
7.1	Štruktúra kódu	33
7.2	Sledovanie priebehu tréningu	35
7.3	Nástroj pre zobrazenie výsledkov	36
8	Výsledky experimentov	37
8.1	Výsledky verifikačných neurónových sietí	37
8.2	Výsledky aplikovania metód odhadu kvality	46

8.3 Porovnanie metód odhadu kvality	59
9 Záver	65
Literatúra	67

Kapitola 1

Úvod

V poslednej dobe je biometrická identifikácia osôb využívaná v rôznych zariadeniach pre umožnenie jednoduchého prístupu do rôznych systémov. V tomto smere patrí snímka dúhovky k jednému z najbezpečnejších spôsobov identifikácie osôb. Detaily a unikátnosť dúhovky vedú k veľmi presným výsledkom rozpoznávacích algoritmov. Problémy týchto systémov sú spôsobené predovšetkým nekvalitnými vstupnými dátami. Chyby pri zlej identifikácii osôb môžu následne viesť ku povoleniu prístupu nesprávnej osoby alebo prípadným neprijatím osoby, ktorá má byť identifikovaná. Pri určení kvality je však možné vstupy spôsobujúce tieto chyby odfiltrovať.

Historicky sa objavovali metódy, ktoré popisovali kvalitu snímky dúhovky rôznymi vizuálnymi charakteristikami ako ostrosť, pomer šumu, prípadne aj informáciami ako tvar a pozícia dúhovky. Tieto metódy však vo svojej aplikácii nedosahujú výraznejšiu presnosť na novodobých systémoch a preto s príchodom a využitím neurónových sietí pre rôzne biometrické úlohy prišli aj novšie metódy odhadu kvality. Podobný vývoj je aj v iných oblastiach, z ktorých je oblasť identifikácie a odhadu kvality snímok tváre tá najvýraznejšia.

V tejto práci ukážem, že aplikáciou moderných metód využívaných pri odhade kvality snímok tváří je možné dosiahnuť ich použitím v oblasti odhadu kvality snímok dúhovky výrazne zlepšenie presnosti verifikačného systému. Zároveň demonštrujem, že tento prístup dosahuje lepšie výsledky ako aktuálne používaná moderná metóda odhadu kvality snímok dúhovky. Vybrané metódy odhadu kvality nemajú aktuálne v oblasti biometrie pomocou dúhoviek preskúmané výsledky. Z tohto pohľadu je ich samotná aplikácia výrazným prínosom a umožňuje možný smer pre ďalší vývoj.

Pre jednotlivé metódy v tejto práci vysvetľujem ich princíp a štýl aplikovania. Práca je zameraná na viacero rôznych metód pre určovanie kvality využívajúce rôzne prístupy. Hlavným zámerom výberu jednotlivých metód je využívanie v čo najväčšej miere výstupov a vlastností verifikačných systémov. V práci sú z vybraných metód odhadu kvality používané metódy ako je generovanie označení kvality pomocou podobnosti výstupov verifikačnej siete, vzorkovanie vektorov z rôznych častí sietí pre určenie ich rozptylu, tréning sietí s informáciou o kvalite ako je magnitúda alebo pomer istoty danej triedy a možnosti pravdepodobnostnej reprezentácie vektorov vlastností. Metódy sú následne vzájomne porovnávané na dvoch rozličných verifikačných systémoch a na štyroch dátových sadách, čím sa zistí aké zlepšenie presnosti je možné ich využitím dosiahnuť.

Kapitola 2

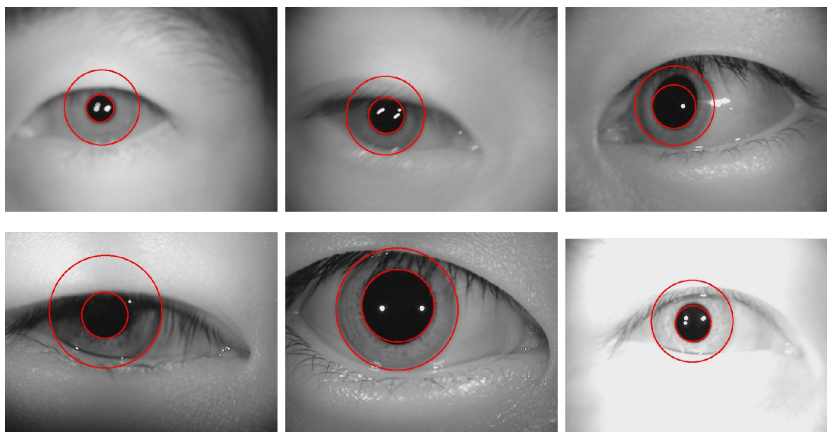
Odhad kvality snímok dúhovky

Problematika určovania kvality snímok je jednou zo súčasťou riešení biometrických systémov. Dnešné rozpoznávacie algoritmy dosahujú výsledky s vysokou presnosťou aj bez využitia odhadu kvality snímok [6, 21]. Vždy je však možné vylepšenie presnosti týchto systémov odmietnutím vstupu, ktorý môže spôsobovať chyby pri identifikácii alebo pri verifikácii. Chyby kvality vstupných snímok teda aj v týchto prípadoch môžu spôsobovať problémy s presnosťou výsledného systému.

2.1 Štandardné kvalitatívne komponenty

Dokument NIST Interagency Report 7820 [34] špecifikuje pre problematiku odhadu kvality snímky dúhovky viacero štandardných komponentov. Štandardné komponenty definujú rôzne vlastnosti snímky dúhovky. Pre tieto vlastnosti platí, že môžu byť príčinou niektorých chýb identifikačného systému. To či daná snímka obsahuje niektoré z jednotlivých komponentov a v akej miere ho obsahuje je možné zistiť ešte pred použitím snímky v systéme. Týmto je následne možné danú snímku odstrániť. Jednotlivé definované komponenty sú popísané v zozname.

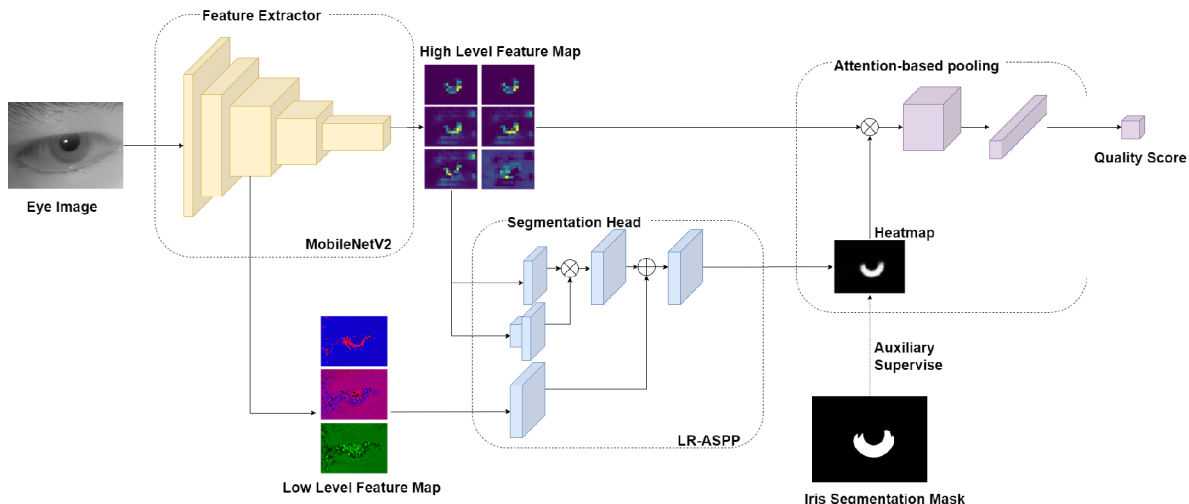
- *Využitelná plocha dúhovky* vyjadruje percento plochy dúhovky, ktorá je použiteľná pre detekciu a nie je zakrytá viečkom.
- *Kontrast zornice* meria charakteristiku snímky dúhovky na hrane dúhovky a zornice. Malý kontrast vie zhoršiť presnosť algoritmov, ktoré detekujú túto hranicu.
- *Tvar zornice* meria nerovnomernosť v tvare hraniciach zornice.
- *Kontrast skléry* je metrika, ktorá meria kontrast na hranici medzi dúhovkou a sklérou.
- *Uhol pohľadu oka* meria odchýlku smeru ktorým oko na snímku pozerá.
- *Ostrosť* je definovaná tým, že obraz neobsahuje rozmazanie, ktoré môže byť spôsobené zaostrením kamery.
- *Dilatácia* meria pomer polomeru zornice ku polomeru dúhovky.
- *Prepletávanie obrazu* je spôsobené metódou získavania dát z kamery, čím sa môžu pri využití starších zariadení prejavovať artefaktami prepletenia obrazu.



Obr. 2.1: Príklad rôznych chýb kvality. (Zdroj [22])

- *Rozpätie šedej škály* ktoré je pri obraze výrazné je indikácia obrazu s vysokou kvalitou. Problémy môžu nastať pri zlom nastavení expozície.
- *Tvar dúhovky* je definovaný hranicou medzi dúhovkou a sklérou.
- *Veľkosť dúhovky* vyjadruje počet pixlov obrazu ktoré dúhovka zaberá. Metrika je ovplyvnená hlavne vzorkovacou frekvenciou obrazu.
- *Pohybové rozmazanie* je definované ako rozmazanie spôsobené pohybom kamery alebo subjektu.
- *Pomer signálu a šumu* meria vlastnosť obrazu, ktorá špecifikuje množstvo informácie. Hlavný zdroj šumu v obraze je snímací prístroj.

Niektoré z chýb na snímkach popísaných kvalitatívnymi komponentami sú zobrazené na obrázku 2.1. Jednotlivé chyby predstavujú z prava do ľava problémy s ostrosťou, pohybovým rozmazaním, uhlom pohľadu oka, malej využiteľnej plochy dúhovky, dilatáciou a rozpätím šedej škály. Existujúce metódy [22, 1] vedia využívať tieto komponenty pre určenie skóre kvality jednotlivých snímkov. Metóda [22] využíva ostrosť, pohybové rozmazanie a uhol pohľadu. Implementácia využíva rôzne kombinácie obrazových filtrov. Týmto sa získa skóre pre každý vybraný komponent kvality. Následne sa skóre spojí do jednej hodnoty fúziou. Druhá citovaná metóda odhadu kvality [1] pridáva okrem týchto komponentov aj informáciu o kontraste zornice a využíva priemer hodnôt jednotlivých metrík kvality pre fúziu.



Obr. 2.2: Architektúra siete pre odhad kvality metódou vzdialenosti v priestore vlastností. (Zdroj [38])

2.2 Určenie kvality snímky dúhovky strojovým učením

Určovanie kvality snímky dúhovky v posledných rokoch prešlo vývojom a z domény ručne definovaných filtrov prešlo na strojové učenie. Jedna z metód, ktorá sa dá považovať za state-of-the-art v oblasti určovania kvality snímky dúhovky je odhad kvality založený na vzdialenosti v latentnom priestore (DFS) [38]. Autori tejto metódy sa zameriavajú na to, ako nekvalitný vstup ovplyvňuje výstup rozpoznávacej neurónovej siete. Skóre kvality je následne určené analýzou výstupu pre rôzne snímky. Predpokladá sa teda vzťah medzi kvalitou snímky dúhovky a vlastnosťami vektoru extrahovaných vlastností snímky.

Pri metóde sa predpokladá existencia ideálnej snímky konkrétneho subjektu označovaného ako x_0 . Za ideálne snímky sa považujú snímky zachytené pri kontrolovaných podmienkach. Tieto podmienky sa zameriavajú na kvalitu osvetlenia, kooperatívny subjekt a iné faktory ako je napríklad správne zaostrenie kamery. K tejto snímke zároveň existuje aj ideálny vektor vlastností z_0 . Trénovacia množina snímok je následne zaznamenaná pri nekontrolovaných podmienkach. Takto sa vytvoria páry snímok pre každý subjekt obsahujúce snímku x_0 ktorá predstavuje ideálnu snímku a potom snímku x_i , ktorá pochádza z množiny snímok zaznamenaných pri nekontrolovaných podmienkach. Pre obidve snímky sa následne získajú vektory vlastností s využitím pred-trénovanej rozpoznávacej neurónovej siete. Týmto sa získa pár z_0 a z_i . Kvalita snímky x_i potom predstavuje vzdialenosť vektorov vlastností v priestore $q_i = \|z_0 - z_i\|$. Predpokladá sa, že snímky s malou vzdialenosťou od ideálnej snímky sú lepšie rozpoznávané a teda majú lepšiu kvalitu. Snímky vzdialene ďalej od ideálnej snímky majú horšiu kvalitu.

Metóda trénuje neurónovú sieť pre získavanie skóre kvality q_i . Architektúra využíva ne-náročnú konvolučnú sieť založenú na MobileNetV2 [28]. K sieti je pripojená segmentačná hlava LR-ASPP [13] použitá pre získanie segmentačnej mapy dúhovky. Architektúru siete, jej vstupy a výstupy je možné vidieť na obrázku 2.2. Okrem iného je možné na obrázku pozorovať aj pripojenie zdržovacej vrstvy, ktorá umožňuje získanie jedného vektoru z výstupu hodnôt siete

$$A_i^{(\hat{z})} = \frac{\sum_{\hat{y}=0}^h \sum_{\hat{x}=0}^w S(x_i)^{(\hat{x})(\hat{y})} * F(x_i)^{(\hat{x})(\hat{y})(\hat{z})}}{\sum_{\hat{y}=0}^h \sum_{\hat{x}=0}^w S(x_i)^{(\hat{x})(\hat{y})}}. \quad (2.1)$$

Pre túto metódu združovania, predstavuje $A^{(\hat{z})}$ hodnotu výstupného vektora v dimenzii \hat{z} . Počet dimenzii vektora závisí na počte kanálov poslednej konvolučnej vrstvy neurónovej siete, $S(\cdot)$ predstavuje výstup segmentačnej mapy dúhovky a $F(\cdot)$ predstavuje výstup konvolučnej siete. Združovanie sa vykonáva cez plochu výstupu konvolučnej siete. Hodnota w predstavuje šírku a h výšku výstupu, kanál výstupu je špecifikovaný hodnotou \hat{z} . Segmentačná mapa získaná z $S(\cdot)$ je transformovaná do rovnakých rozmerov h a w . Výstupný vektor A_i je následne použitý ako vstup do plne prepojenej vrstvy, ktorej výsledkom je jedna hodnota skóre kvality. Pre optimalizáciu parametrov siete sa využíva stratová funkcia

$$Loss_i = \lambda Loss_{mask_i} + (1 - \lambda) Loss_{MSE_i}. \quad (2.2)$$

Funkcia kombinuje stratu segmentačnej masky vyjadrenej ako $Loss_{mask_i}$, a stratu predikovaného skóre kvality $Loss_{MSE_i}$ voči referencii q_i . Segmentačná časť siete sa trénuje zároveň s určovaním kvality a dátová sada okrem hodnôt kvality obsahuje aj referenčnú masku. Hodnota λ predstavuje podiel týchto dvoch metrík na výslednej stratovej funkcii. Strata segmentačnej masky je vyjadrená dvoj-dimenzionálnou stratovou funkciou krížovej entropie funkciou. Hodnota $Loss_{MSE_i}$ predstavuje štvorcovú chybu.

2.3 Metriky vyhodnocovania biometrického systému

Pri porovnaní viacerých metód určovania kvality je dôležité získanie merateľných porovnávacích parametrov. Vhodné parametre umožňujú porovnanie viacerých metód, čím je možné určiť, ktorá z metód je pre danú úlohu lepšie použiteľná.

V predošlej sekcii zmienený dokument NIST Interagency Report 7820 [34] poskytuje okrem štandardných kvalitatívnych komponentov aj prehľad metrík, pre porovnanie metód určujúcich kvalitu snímky dúhovky. V celkovom rozsahu sú z tohto pohľadu zaujímavé metriky pre hodnotenie presnosti verifikačného systému a metriky pre vyhodnotenie metódy určenia kvality.

Formálne sa bude využívať dátová sada označovaná ako $\mathcal{P} = \{n | n \in 1 \dots N_P\}$, kde N_P predstavuje počet párov snímok dúhovky, ktoré sa v danej dátovej sade nachádzajú. Pár p_i je definovaný dvojicou snímok dúhovky $p_i = (x_j, x_k)$, kde $i \in P$. Páry sú rozdelené na súhlasné a nesúhlasné. Množina súhlasných párov bude označovaná ako $\mathcal{G} \subset P$ a množina nesúhlasných párov bude označovaná ako $\mathcal{I} \subset P$. Pre množiny párov \mathcal{I} a \mathcal{G} platí podmienka prázdneho prieniku $\mathcal{G} \cap \mathcal{I} = \emptyset$ a podmienka $\mathcal{G} \cup \mathcal{I} = P$. Pre každý pár je následne možné vypočítať skóre odlišnosti vyjadrené funkciou $s(i)$. Funkcia vyhodnotí verifikačný algoritmus pre obe snímky z páru i a následne vypočíta odlišnosť výstupov. Pre skóre odlišnosti musí platiť, že menšie hodnoty znamenajú väčšiu podobnosť a väčšie hodnoty zase menšiu podobnosť páru. Pre každý pár snímok zároveň existuje aj pár skóre kvality označené ako $q_{p_i} = (q_j, q_k)$. Hodnoty skóre kvality pre jednotlivé snímky sú reprezentované monotónne rastúcou funkciou. Platí, že snímky s vysokým skóre kvality vrátia pri porovnaní s rovnakou identitou menšie hodnoty odlišnosti. V jednoduchosti je možné povedať, že vyššie skóre je priamo úmerné kvalite vstupu.

Hodnoty skóre kvality reprezentované dvojicou nie sú vhodné pre ďalšie vyhodnocovanie. Z tohto dôvodu je vhodné vyjadriť jedno skóre kvality, ktoré charakterizuje obe hodnoty označované ako \hat{q}_{p_i} . Spoločné skóre kvality môže byť vypočítané minimálnou hodnotou

$$\hat{q}_{p_i} = \min(q_{p_i}^{(1)}, q_{p_i}^{(2)}). \quad (2.3)$$

Hodnoty ⁽¹⁾ a ⁽²⁾ predstavujú prvý, prípadne druhý prvok z dvojice. Majme množinu \mathcal{R} predstavujúcu množinu párov, ktoré boli odmietnuté pre príliš nízke skóre kvality

$$\mathcal{R}(q_\tau) = \{i | i \in P \wedge \hat{q}_{p_i} < q_\tau\}, \quad (2.4)$$

kde hodnota q_t predstavuje zvolený prah kvality. Potom $\overline{\mathcal{R}}$ predstavuje komplement ku množine odmietnutých párov. Následne je možné vyjadriť počet súhlasných párov prijatých rozpoznávacím algoritmom pre rozhodovací prah väčší ako hodnota τ a prah kvality väčší ako q_t

$$M(\tau, q_\tau) = \sum_{i \in \mathcal{G} \cap \overline{\mathcal{R}(q_\tau)}} H(s(i) - \tau). \quad (2.5)$$

Funkcia $H(\cdot)$ predstavuje skokovú funkciu pre ktorú platí

$$H(x) = \begin{cases} 0 & x \leq 0 \\ 1 & x > 0 \end{cases}. \quad (2.6)$$

Metrika pomeru nesprávne neprijatých párov anglicky False Non-match Rate (FNMR) sa dá následne vyjadriť ako podiel hodnôt

$$FNMR(\tau, q_\tau) = \frac{M(\tau, q_\tau)}{M(-\infty, q_\tau)}. \quad (2.7)$$

Podobným spôsobom je možné vyjadriť aj počet nesúhlasných párov prijatých pod zvolený rozhodovací prah

$$N(\tau, q_\tau) = \sum_{i \in \mathcal{I} \cap \overline{\mathcal{R}(q_\tau)}} 1 - H(s(i) - \tau). \quad (2.8)$$

Metrika pomeru nesprávne prijatých párov anglicky False Match Rate (FMR) sa dá následne vyjadriť ako podiel hodnôt

$$FMR(\tau, q_\tau) = \frac{N(\tau, q_\tau)}{N(\infty, q_\tau)}. \quad (2.9)$$

Samotný popis pomocou hodnôt FMR a FNMR nieje dostatočný pre vyjadrenie ich vzťahu a pre porovnávanie rôznych metód. Z tohto pohľadu je vhodné tieto metriky dať do formy grafov. Následujúce časti tejto sekcie popisujú práve možnosti vizualizácie metrick defínovaných vyššie.

Receiver Operating Characteristic (ROC) graf je jeden z nástrojov pre vizualizáciu vzťahu metrick FNMR a FMR. V princípe tento graf vyjadruje jednu metriku ako funkciu druhej metriky. Hodnoty FNMR a FMR sú takto zobrazované pri rovnakej hodnote prahu. Výslednú krivku je potom možné využiť pri porovnaní rôznych metód. Z praktického hľadiska sú lepšie metódy asymptotický bližšie pri jednotlivých osiach grafu.

Detection Error Tradeof (DET) [23] graf prináša vylepšenú reprezentáciu údajov ROC krivky. Táto metóda využíva odlišnú mierku jednotlivých osí, pre ktorú je možné lepšie rozlíšiť hodnoty jednotlivých grafov. Priebeh vykreslených kriviek je viac lineárny kde v porovnaní s ROC pripomína viac exponenciálnu krivku. Pre iné ako vizuálne porovnávanie jednotlivých grafov je možné použiť hodnoty plochy pod DET krivkou, týmto sa získa jedna hodnota charakterizujúca túto krivku.

Error vs. reject graf využíva porovnania chyby identifikačného systému vzhľadom k pomeru odmietnutých párov obrázkov (IRR). Tento spôsob vyjadrenia môže byť využitý pre porovnanie metód určovania kvality. Pre získanie grafu sa využíva hodnota prahu kvality q , pre ktorú platí, že pri aplikovaní na dátovú sadu sa odmietne zvolený pomer porovnávaných párov. Pomery odmietnutých párov sú následne vyžité ako hodnoty grafu na osi x . Pre pomer odmietnutých párov snímkov sa nájde hodnota FNMR so zvolenou hodnotou FMR. Príkladom je napríklad hodnota $FMR = 0.1\%$. Pre efektívny algoritmus odhadu kvality potom platí, že hodnota chyby FNMR klesá relatívne rýchlo so zvyšujúcim sa počtom odmietnutých párov. V práci budem požívať aj graf pre relatívne hodnoty FNMR. Jediný rozdiel je použitie podielu originálnych hodnôt FNMR s hodnotou FNMR pri 0% odmietnutých párov.

Kapitola 3

Identifikácia pomocou snímky dúhovky

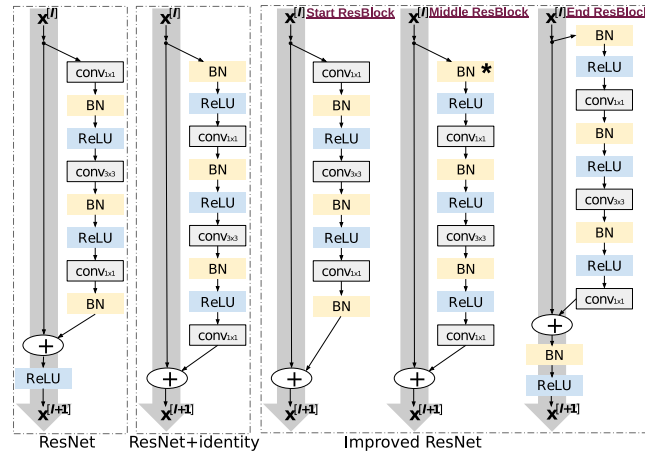
Tak ako bolo ukázané v sekcii 2.2, moderné metódy určovania kvality snímkov dúhovky využívajú pri odhade kvality rozpoznávací systém. Úloha tohto systému je získať unikátnu reprezentáciu snímku dúhovky. Jednotlivé svaly a cievy viditeľné na dúhovke tvoria vzor, ktorý môže byť prevedený do kódu. Táto reprezentácia je následne porovnateľná a môže sa zistiť, či na jednotlivých snímkach je rovnaká dúhovka. Systém s týmito vlastnosťami je následne možné využiť pre verifikáciu, alebo pri identifikácii osôb.

Historicky boli rozpoznávacie systémy založené na metódach ako Daugmanov algoritmus [7], ktorý využíva normalizáciu snímok a sadu Gaborových filtrov, prípadne na metódach založených na vlnkových transformáciách [1, 27] a iných [2]. Tieto metódy podobne ako metódy využívané v iných oblastiach biometrie, pomaly zanikajú a začínajú byť nahradzované neurónovými sieťami.

3.1 Architektúry konvolučných neurónových sietí

Hĺbkové konvolučné neurónové siete umožňujú riešenie rôznych úloh počítačového videnia. Ich využitie je počnúc detekciou, segmentáciou až po klasifikáciu alebo rozpoznávanie. Rovnako môžu byť použité aj pri rôznych biometrických systémoch založených na snímkach dúhovky [6, 21]. V posledných rokoch sú príkladmi konvolučných sietí AlexNet [19] alebo VGG [31]. S tréningom týchto sietí však spočívajú aj problémy. Jedným z hlavných problémov je miznutie gradientov pri spätnej propagácii chyby. Tento problém je spôsobený zvyšovaním počtu vrstiev, a obmedzuje možnosti tréningu hlbších sietí. Existuje množstvo architektúr, ktoré riešia tento problém ako napríklad DenseNet [15] a Inception [33]. Pre účely bude práca zameraná na siete ResNet [10] iResNet[9] a MobileNet [28, 13].

ResNet je jednou z najviac využívaných konvolučných neurónových sietí. Hlavná myšlienka tejto architektúry je založená na využití reziduálnych prepojení. Reziduálne prepojenie predstavuje jednoduché pripočítanie vstupnej hodnoty bloku k jeho výstupu. Týmto sa pri spätnej propagácii chyby vytvorí priama cesta sieťou bez aplikácie konvolučných jadier. ResNet definuje stavebný reziduálny blok, ktorý je možné vidieť na obrázku 3.1 v ľavo. Samotný reziduálny blok využíva kombináciu 1×1 a 3×3 konvolučných jadier, ktoré vytvárajú úzke hrdlo. Prvé jadro 1×1 zníži počet kanálov vstupujúcich do jadra 3×3 , posledné jadro 1×1 následne zväčší počet kanálov na originálny počet. Takto je znížený po-



Obr. 3.1: Porovnanie reziduálnych blokov. (Zdroj [9])

čet inak potrebných parametrov. Za jednotlivými jadrami sa využíva nelineárna aktivačná funkcia ReLU [25] a batch-normalizácia [16] pre tréning. Z týchto stavebných blokov sú tvorené vrstvy siete. Jednotlivé vrstvy používajú rôzne počty konvolučných blokov a každá vrstva ma za úlohu znížiť rozmery vstupného obrázku na polovicu a zväčšiť počet kanálov na dvojnásobok. Na začiatku celej siete je jedno konvolučné jadro 7×7 , ktoré spoločne s average-pooling vytvára vstup do jednotlivých vrstiev. Na konci siete je aplikovaná metóda average-poolig v kombinácii s plne prepojenou vrstvou, alebo len plne prepojená vrstva, pre získanie jedného vektoru hodnôt.

iResNet stavia nad reziduálnou sieťou ResNet. Architektúra sa snaží zlepšiť tréning a presnosť. Toto sa dosahuje zlepšením spätnej propagácii chyby. Pre tento účel sú navrhnuté tri typy blokov ktoré sú zobrazené v pravej časti na obrázku 3.1. Bloky sú rozdelené podľa toho, v akej časti konvolučnej vrstvy sa aplikujú. Týmto sa zabezpečuje prechod reziduálnych prepojení rovnakým počtom nelineárnych aktivácií. Klasický blok ResNet pri porovnaní aplikuje nelineárne aktivačné funkcie pre každé reziduálne spojenie. Metóda s využitím identity [11] na druhej strane neaplikuje na žiadne reziduálne prepojenia aktivačné funkcie. Pre vylepšené siete ResNet je aktivačná funkcia aplikovaná jedine na posledné reziduálne spojenie v konvolučnej vrstve. Podobne ako pri ResNet sa využívajú celkovo štyri vrstvy.

MobileNet v porovnaní s architektúrou ResNet sa zameriava na optimalizáciu počtu parametrov. Týmto sa dá využiť na zariadeniach ako napríklad kamery a mobilné telefóny. Samotná architektúra siete používa mierne odlišný prístup pre konvolučné bloky. Pre účel zmenšenia počtu parametrov je definovaný invertovaný reziduálny blok. Blok sa podobne ako pri ResNet skladá z troch konvolučných vrstiev. Rozdielom je využitie konvolúcií cez hĺbku. Pri štandardnom konvolučnom jadre má každý kanál výstupu počet filtrov rovný počtu kanálom vstupu. Konvolúcia cez hĺbku používa pre každý kanál výstupu len jeden konvolučný filter. Je však potrebné využiť väčší počet kanálov prechádzajúcich cez túto vrstvu. Prvé konvolučné jadro bloku o veľkosti 1×1 plní práve úlohu zväčšenia počtu kanálov. Okrem iného sú rozdiely aj vo využití aktivačných funkcií ako RELU6 [14] a Hard Swish. Novšia verzia MobileNetV3 zároveň pridáva aj nenáročnú metódu pozornosti squeeze-and-excite aplikovanú interne v reziduálnych blokoch.

3.2 Tréning výstupu siete pre vektor vlastností

Jednou z možnou unikátnou reprezentáciou snímku je získanie jeho vektoru vlastností. Tréning pre získanie vektoru vlastností prináša rôzne výzvy. Jednou z týchto výziev je dosiahnutie čo najlepšej separácie vektorov v latentnom priestore. Ideálne je dosiahnuť to, aby vektory patriace jednej identite boli čo najbližšie pri sebe a zároveň, aby ich vzdialenosť od iných identít bola čo najväčšia. Jednou z metód, ktorá rieši tento problém je stratová funkcia triplet [29]. Táto forma tréningu má však svoje nevýhody a to hlavne v počte vytvorených tréningových trojíc. Alternatívna možnosť je použitie kontrastného učenia s učiteľom [17]. Alternatívy sú taktiež aj vo využití metód upravujúcich stratovú funkciu krížovej entropie. Tréning je v týchto prípadoch ekvivalentný klasifikácii, rozdiel je však v následnom odstránení klasifikačnej vrstvy. Ukážky metód sú CosFace [37] a AM-SoftMax [36].

ArcFace [8] podobne ako predošlé spomenuté metódy sa pokúša vylepšiť separáciu vektorov zmenou výpočtu stratovej funkcie krížovej entropie. V tomto prípade sa používa x_i ako vstupný snímok, ktorý patrí identite označenej ako y_i . Pre y_i platí, že pochádza z oboru hodnôt $y_i \in 1 \dots n$, kde n predstavuje počet tried v tréningovej množine. Následne sa predpokladá vektor príznakov z_i , ktorý je výstupom neurónovej siete pre vstup x_i . Na poslednej vrstve neurónovej siete sa trénujú vektory váh w_j pre $j \in 1 \dots n$. Tieto hodnoty reprezentujú natrénované body stredov jednotlivých tried. Stratová funkcia krížovej entropie má následne formu

$$Loss_{CrossEntropy} = -\frac{1}{N} \sum_{i=1}^N \log \frac{e^{z_i w_{y_i}^T + b_{y_i}}}{\sum_{j=1}^n e^{z_i w_j^T + b_j}}. \quad (3.1)$$

Hodnota N v tomto prípade predstavuje počet tréningových vzorov, hodnoty b predstavujú posuv hodnôt. Využitie krížovej entropie explicitne nevynucuje zníženie vzdialenosti medzi stredmi rovnakej triedy. Čo má za následok, že vektory jednotlivých identít majú veľký rozptyl a hranice medzi jednotlivými identitami sú blízko seba. Autori ArcFace riešia práve získanie kompaktnejšej reprezentácie jednotlivých tried. Pre ďalší popis je teda potrebná iná reprezentácia vyššie popísanej rovnici krížovej entropie. Hodnota exponentu $z_i w_j^T$ sa dá prepísať do podoby kosínusovej podobnosti vektorov z_i a w_j a to vzťahom $z_i w_j^T = \|z_i\| \|w_j\| \cos \theta_j$. Ak sa následne predpokladá že, veľkosti vektorov z_i a w_j sú 1, čo je možné zabezpečiť v oboch prípadoch pomocou l2 normalizácie, tak je možné získať nový vzťah $z_i w_j^T = \cos \theta_j$. Krížovú entropiu je možné následovne reprezentovať ako

$$Loss'_{CrossEntropy} = -\frac{1}{N} \sum_{i=1}^N \log \frac{e^{s \cos(\theta_{y_i})}}{\sum_{j=1}^n e^{s \cos(\theta_{y_i})}}, \quad (3.2)$$

v tejto rovnici sa zanedbáva posunutie, ktoré je prítomné v rovnici 3.1. Vektory reprezentujúce jednotlivé identity w sú rozložené na hyper-guli. Získaný vektor vlastností z_i je následne reprezentovaný jeho uhlovou vzdialenosťou od bodu z . Autori následne ukazujú, že zníženie rozptylu jednotlivých vektorov vlastností je možné dosiahnuť pridaním okraju m k uhlu θ_{y_i} . Výsledná stratová funkcia je následne reprezentovaná vzťahom

$$Loss_{ArcFace} = -\frac{1}{N} \sum_{i=1}^N \log \frac{e^{s \cos(\theta_{y_i} + m)}}{e^{s \cos(\theta_{y_i})} + \sum_{j \neq y_i} e^{s \cos(\theta_{y_i})}}. \quad (3.3)$$

Kapitola 4

Všeobecný odhad kvality snímok

V sekcii 2.1 bol popísaný prehľad kvalitatívnych komponentov a prehľad metód pre určenie kvality, ktoré ich využívajú. Zároveň bola v sekcii 2.2 ukázaná metóda využívajúca natrénované hodnoty kvality, ktorá ukázala lepšie výsledky pri určovaní kvality. V posledných rokoch prišlo s využitím konvolučných sietí viacero metód pre určovanie kvality. Metódy prišli hlavne s vývojom rozpoznávania pomocou tvári [30, 26, 35, 4, 24]. Niektoré z týchto metód sú však vo svojom využití univerzálne a môžu byť potenciálne aplikované aj na problematiku určenia kvality snímok dúhovky. Okrem iného tieto metódy obvykle využívajú v nejakej forme rozpoznávaciu sieť, ktoré boli popísané v kapitole 3. V tejto kapitole sú predstavené niektoré z state-of-the-art metód v oblasti odhadu kvality.

4.1 Pravdepodobnostná reprezentácia vektora príznakov

Klasické konvolučné siete využívané pre rozpoznávanie reprezentujú svoj výstup práve jednou hodnotou vektora. Táto metóda prichádza s myšlienkou pravdepodobnostnej reprezentácie vektora príznakov (PFE) [30]. Okrem odhadu vektoru v latentnom priestore sa uvažuje aj s odhadom pravdepodobnostného rozloženia jeho hodnôt definovaného rozptylom v jednotlivých dimenziách. Pri určovaní kvality snímok je následne možné odmietnuť vstup do rozpoznávacieho systému na základe hodnôt rozptylu. Predpokladá sa, že vstupy s nižšou kvalitou majú väčší rozptyl ako vstupy s vyššou kvalitou.

Metóda pracuje s vektormi príznakov generovanými rozpoznávacou sieťou. Vektory predstavujú stredné hodnoty rozloženia pravdepodobnosti. Pre získanie vektoru určujúceho rozptyl pravdepodobnostného rozloženia je autormi definovaný nasledujúci postup. Predpokladajme, že X je priestor vstupných snímok a Z je latentný priestor vektorov extrahovaných vlastností o dimenzionalite D . Pre konkrétnu identitu potom platí, že má reprezentáciu v latentnom priestore označenú ako $z \in Z$. Následne existuje snímka patriaca tejto identite pre ktorú platí $x \in X$. Pre takúto snímku je potom možné zapísať pravdepodobnosť toho, že pochádza od konkrétnej identity v latentnom priestore ako podmienenú pravdepodobnosť $p(x|z)$. Výsledkom tréningu rozpoznávacej siete je také rozloženie v priestore Z , ktoré predstavuje podmienenú pravdepodobnosť $p(z|x)$, a teda neformálne povedané, hľadá rozloženie pravdepodobnosti identít, ktorým táto snímka môže patriť. Pre konkrétnu vstupnú snímku x_i je potom pravdepodobnostné rozloženie $p(z|x_i)$ aproximované viacrozmerným Gaussovým rozložením pravdepodobnosti

$$p(z|x_i) = \mathcal{N}(z; \mu_i, \sigma_i^2 I), \quad (4.1)$$

x_i predstavuje i -tu snímku z množiny snímok. Rozloženie pravdepodobnosti má následne vektor stredných hodnôt μ_i a diagonálnou kovariančnej matice reprezentovanou σ_i^2 . K samotnému tréningu sa využíva vždy pár vstupných snímok x_i a x_j ktoré musia pochádzať od rovnakej identity. Následne autori vyjadrujú, že pre pravdepodobnosť toho, či snímky pochádzajú od rovnakej identity sa dá vyjadriť ako

$$p(z_i = z_j) = \int \int p(z_i|x_i)p(z_j|x_j)\delta(z_i - z_j)dz_idz_j, \quad (4.2)$$

v tomto vzťahu predstavuje funkcia $\delta(\cdot)$ Diracovú δ -funkciu, jej integrál má hodnotu 1, zároveň maximum tejto funkcie je v bode 0 a v iných bodoch je rovná 0. Funkcia je rozšírená na viac dimenzií a jeden integrál vždy prechádza cez všetky dimenzie používaných vektorov. Autori ukazujú, že optimalizácia parametrov siete pre získanie vektoru rozptylov, vzhľadom na získanie maximálnej hodnoty tejto pravdepodobnosti, umožní trénovať sieť práve pre odhad rozptylov pre reprezentáciu pravdepodobnostného rozloženia vektoru vlastností. Autori pre tento účel uvádzajú nasledujúci vzťah

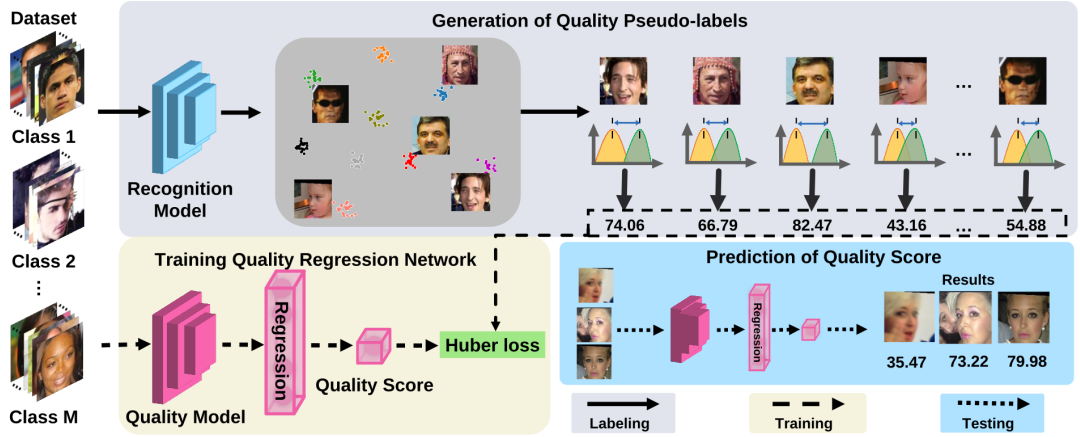
$$\begin{aligned} s(x_i, x_j) &= \log(p(z_i = z_j)) \\ &= -\frac{1}{2} \sum_{l=1}^D \left(\frac{(\mu_i^{(l)} - \mu_j^{(l)})^2}{\sigma_i^{2(l)} + \sigma_j^{2(l)}} + \log(\sigma_i^{2(l)} + \sigma_j^{2(l)}) \right) - \frac{D}{2} \log(2\pi), \end{aligned} \quad (4.3)$$

vzťah vyjadruje podobnosť dvoch získaných rozložení. Rozloženia sú definované hodnotami $\sigma_i^2, \sigma_j^2, \mu_i$ a μ_j . Hodnoty $\sigma_i^{2(l)}, \sigma_j^{2(l)}, \mu_i^{2(l)}$ a $\mu_j^{2(l)}$ predstavujú hodnoty vektorov v dimenzii l . Dôkaz tejto rovnice je dostupný z citovanej literatúry. Pre účely tréningu siete je následne definovaná stratová funkcia

$$Loss_{PFE} = \frac{1}{|P|} \sum_{(i,j) \in P} -s(x_i, x_j), \quad (4.4)$$

P predstavuje množinu párov vstupov využívaných pre tréning neurónovej siete, ktoré patria rovnakej identite.

Trénuje sa samostatný modul pre určenie rozptylov. Modul predstavuje sieť skladajúcu sa z plne prepojenej vrstvy, batch-normalizácie, ReLU, druhej plne prepojenej vrstvy a batch-normalizácia, Výstup modulu predstavuje logaritmus hodnôt vektoru rozptylov. Rozmer výstupu všetkých vrstiev je rovnaký ako rozmer výstupu rozpoznávacej siete. Samostatný modul je pripojený za rozpoznávaciu sieť. Pri tréningu sa optimalizujú jedine váhy tohto modulu.



Obr. 4.1: Znárodnenie procesu tréovania a vytvárania tréovacej sady pre SDD-FIQA. (Zdroj [26])

4.2 Odhad kvality z distribučných funkcií podobností

Jednou z možností reprezentácie kvality rozpoznávanej snímky je aj zisťovanie rôznych štatistických vlastností vektorov príznakov. Existuje možnosť využiť podobnosti vektorov príznakov rovnakej triedy v porovnaní s vektormi rozdielnych tried. Táto možnosť je využitá práve pri metóde určovania kvality pomocou vzdialenosti distribučných funkcií podobností vektorov SDD-FIQA [26].

Základná myšlienka tejto metódy využíva vlastnosti distribúcie podobností pozitívnych a negatívnych príkladov ku vektoru príznakov vybranej snímky. Pozitívne príklady predstavujú vektory príznakov, ktoré pochádzajú od tej istej triedy a negatívne príklady sú od iných tried. Pre každý takýto pár sa vypočíta podobnosť vektorov v latentnom priestore. Týmto sa získajú dve sady hodnôt $S_{x_i}^P$ pre pozitívne príklady a $S_{x_i}^N$ pre negatívne. Predpokladá sa, že ak je vybraná snímka s vysokou kvalitou, tak distribúcia podobností pozitívnych párov a distribúcia podobností negatívnych párov bude mať väčšiu vzdialenosť. Pre menej kvalitný vstup budú tieto dve distribúcie bližšie pri sebe. Výpočet skóre kvality je potom možné previesť ako

$$q_i = \mathbb{W}\mathbb{D}(S_{x_i}^P || S_{x_i}^N). \quad (4.5)$$

Funkcia $\mathbb{W}\mathbb{D}(\cdot)$ predstavuje Wasserstein metriku inak známu ako Earth Mover's Distance (EMD). Táto metrika neformálne definuje minimálnu cenu presunu jednej pravdepodobnostnej distribučnej funkcie na miesto druhej. Formálna definícia tejto metriky pre rozloženia pravdepodobnosti je vyjadrená ako

$$\mathbb{W}\mathbb{D}(S_{x_i}^P || S_{x_i}^N) = \inf_{\gamma \in \Pi(S_{x_i}^P, S_{x_i}^N)} \mathbb{E}_{(s_{x_i}^P, s_{x_i}^N) \sim \gamma} (\|s_{x_i}^P - s_{x_i}^N\|), \quad (4.6)$$

kde $\Pi(S_{x_i}^P, S_{x_i}^N)$ predstavuje množinu pravdepodobnostných distribúcií $\gamma(s_{x_i}^P, s_{x_i}^N)$, ktoré majú marginálne pravdepodobnostné distribúcie $S_{x_i}^P$ a $S_{x_i}^N$.

Pri vyhodnotení $\mathbb{W}\mathbb{D}$ pre prípad tejto metódy určovania kvality je však problém s veľkosťou množín $S_{x_i}^P$ a $S_{x_i}^N$. Pre zvýšenie efektivity výpočtu autori využívajú len náhodné vzorky z pozitívnych a negatívnych párov o predom definovanej veľkosti. Následne sa prevedie

vyhodnotenie metriky WD pre viacero takto vybraných množín, čím sa zvyšuje presnosť výpočtu. Výpočet výsledného skóre kvality je

$$\tilde{q}_i = \frac{\sum_{k=1}^K \text{WD}(S_{x_i^k}^{P_m} || S_{x_i^k}^{N_m})}{K}. \quad (4.7)$$

Pri tomto zápise predstavujú množiny $S_{x_i^k}^{P_m}$ a $S_{x_i^k}^{N_m}$ k -te náhodné vzorky z množín vzdialeností pozitívnych a negatívnych párov pre snímok x_i . Veľkosť týchto množín je m . K predstavuje počet náhodne vzorkovaných množín.

Samotná architektúra použitej siete pre určenie kvality a vizualizácia toho, čo skóre kvality predstavuje je zobrazené na obrázku 4.1. Autori uvažujú tréning s pred-trénovanými sieťami ResNet. Pre tréningovanie regresného modelu sa následne posledná plne prepojená vrstva nahradí, plne prepojenou vrstvou s jednou výstupnou hodnotou. Pre optimalizáciu parametrov siete sa využíva Huber stratová funkcia

$$Loss_{Huber} = \frac{1}{N} \sum_{i=1}^N \begin{cases} \frac{1}{2}(\tilde{q}_i - F(x_i))^2 & \text{ak } |\tilde{q}_i - F(x_i)| \leq \delta \\ \delta|\tilde{q}_i - F(x_i)| - \frac{1}{2}\delta^2 & \text{inak} \end{cases}. \quad (4.8)$$

V tejto rovnici predstavuje $F(\cdot)$ funkciu výstupu konvolučnej neurónovej siete ktorá sa trénuje pre odhad kvality snímky x_i . Výstupom tejto siete je predikovaná hodnota skóre kvality. Parameter δ predstavuje prah, pre ktorý sa stratová funkcia Huber mení z kvadratickej chyby na absolútnu chybu.

4.3 Odhad kvality pomocou robustnosti vektorov príznakov

Metóda určenia kvality pomocou robustnosti vektorov vlastností (SER-FIQ) [35] využíva narozdiel od predošlých metód odhad kvality snímky bez učiteľa. Predpokladá sa, že kvalitný vstup by mal byť zároveň viac robustný. Táto vlastnosť sa určuje využitím porovnania viacerých výstupov rozpoznávacej siete. Získané vzorky sú vždy z rôznych častí siete. Očakáva sa menší rozptyl výstupov pri vstupoch s lepšou kvalitou.

Metóda využíva existujúci rozpoznávací model a vstupnú snímku x_i . Pre tento vstup sa získa zo siete množina stochastických vzorov vektorov v latentnom priestore $S_i = \{z_i^j\}_{j \in \{1,2,\dots,m\}}$. Vzorky sú získané využitím rôznych vzorov drop-out [32] vrstvy. Počet vzorov $m = 100$ je autormi zvolený ako kompromis medzi rýchlosťou výpočtu a presnosťou určenej kvality. Pre získané vzorky vektorov vlastností sa následne dá vypočítať skóre kvality vzťahom

$$q_i = \frac{2}{m^2} \sum_{j < k} \|z_i^j - z_i^k\|. \quad (4.9)$$

Využitie euklidovskej vzdialenosti stochastických vzorov je podľa autorov vhodná metrika pre určenie podobnosti. Väčšia variácia v pozíciách v latentnom priestore vedie ku väčším hodnotám euklidovskej vzdialenosti a tým väčším hodnotám skóre kvality. Týmto teda indikuje malú robustnosť vstupných dát a teda horšiu kvalitu. Celý proces získavania stochastických vzorov a výpočtu kvality je popísaný algoritmom 1.

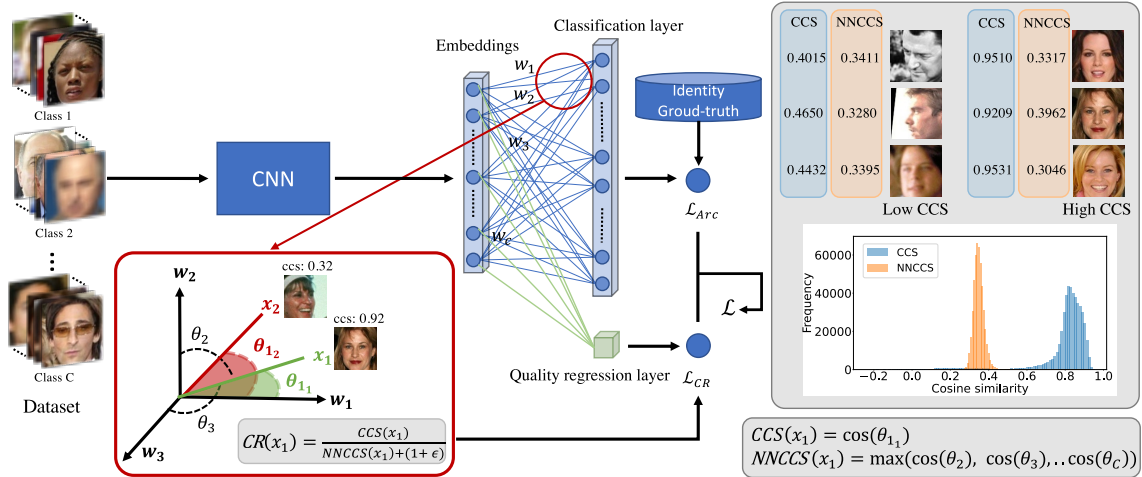
Pre implementáciu sú autormi poskytnuté dva rôzne spôsoby. Je možné využiť existujúci model, za ktorý sa pridajú ďalšie vrstvy obsahujúce drop-out vrstvu. Tieto vrstvy sa trénujú rovnakým spôsobom ako rozpoznávacia sieť. Metóda je autormi nazývaná on-top model. Druhý spôsob je využitie existujúceho rozpoznávacieho modelu. V tomto prípade musí byť

Algoritmus 1: Stochastic Embedding Robustness. (Zdroj [35])

Input: Rozpoznávací model F a snímok x_i
Output: Odhadnutá hodnota skóre kvality q_i
for j **in** $1 \dots m$ **do**
 | $z_i^j \leftarrow F.predict(x_i, dropout = True);$
end
return q_i

model trévaný s drop-out vrstvou. Model je nazývaný ako same-model. Experimentálne výsledky autorov ukazujú, že metóda, ktorá využíva same-model je lepšia pre určovanie kvality.

Pre zvýšenie efektivity výpočtu náhodných vzorov v latentnom priestore je možné využiť optimalizáciu. Konvolučné neurónové siete využívajú drop-out jedine na poslednej vrstve. Výstup predošlých vrstiev sa teda nemení a predpokladá sa, že pri viacnásobnom vyhodnotení rovnakého vstupu bude rovnaký. Je teda vhodné metódu implementovať tak, aby bol prechod konvolučnými vrstvami len jeden a výstup poslednej konvolučnej vrstvy pred drop-out bol nekopírovaný m -krát. Následne sa s m vzorkami prechádza len poslednou vrstvou a nie celým modelom.



Obr. 4.2: Prehľad architektúry siete a tréning metódy CR-FIQA. (Zdroj [4])

4.4 Odhad kvality podľa relatívnej klasifikovateľnosti

Metódy určovania kvality snímky spomínané v predchádzajúcich sekciách očakávali prítomnosť existujúcej rozpoznávacej siete, ktorá môže byť využitá pre generovanie skóre kvality. Metóda určovania kvality snímku pomocou učenia relatívnej klasifikovateľnosti vzorky [4] (CR-FIQA), trénuje zároveň sieť pre extrakciu príznakov a skóre kvality.

Pre popis metódy predpokladám vstupný snímok x_i , pre ktorý existuje vektor vlastností v latentnom priestore z_i . Pre klasifikačnú vrstvu potom existuje vektor váh w_{y_i} , kde y_i predstavuje triedu do ktorej patrí vstupná snímka. Trieda y_i patrí do množiny všetkých tried $1 \dots n$. Vektor w_{y_i} predstavuje stred triedy, ktorej patrí snímka x_i , podobne ako pri sekcii 3.2 popisujúcej ArcFace [8]. Uhol θ_{y_i} , predstavuje uhol medzi stredom triedy a výstupným vektorom v latentnom priestore

$$z_i w_{y_i}^T = \|z_i\| \|w_{y_i}\| \cos(\theta_{y_i}). \quad (4.10)$$

Metóda následne využíva hodnotu $\cos(\theta_{y_i})$ ako jeden z parametrov pre určenie skóre kvality. Autori nazývajú tento parameter anglicky ako Class Center Angular Similarity (CCS)

$$CCS_i = \cos(\theta_{y_i}). \quad (4.11)$$

Následne sa určí ešte podobnosť s najbližším stredom negatívnej triedy, anglicky Nearest Negative Class Center Angular Similarity (NNCCS), ktorý sa nájde z výstupov uhlových podobností pre všetky ostatné stredy tried

$$NNCCS_i = \max_{j=1, j \neq y_i}^n (\cos(\theta_j)). \quad (4.12)$$

Metóda predpokladá, že snímky s vysokou kvalitou budú mať väčšiu hodnotu CCS a teda sú bližšie pri bode reprezentujúcom ich triedu. Zároveň sa predpokladá, že hodnota $NNCCS$ bude menšia, keďže takto klasifikovaný vstup bude čo najďalej od stredy ostatných tried. Na druhej strane, vstupy s nízkou kvalitou budú viac vzdialené od stredy svojej triedy a bližšie pri stredy iných tried. Z týchto dvoch hodnôt určiť pomer istoty, Certainty Ratio (CR)

$$q_i = CR_i = \frac{CCS_i}{NNCCS_i + (1 + \epsilon)}. \quad (4.13)$$

Hodnota ϵ je malá kladná hodnota, ktorá vo výraze $(1 + \epsilon)$ uisťuje, že hodnota menovateľa bude rôzna od nuly a výsledok bude v maximálne 1. Týmto obíde problém s delením nulou. V experimentoch autori využili hodnotu $\epsilon = 1e - 9$.

Pri tréovaní sa zároveň kombinuje tréovanie regresného modelu určujúceho kvalitu snímky pomocou *CR* a tréovanie rozpoznávacej siete pomocou *ArcFace*. Autori využívajú ako základ sietí, ku ktorej poslednej vrstve je okrem klasifikačnej vrstvy pripojená plne prepojená vrstva s jedným výstupom určená pre predikciu skóre kvality. Celková architektúra siete je znázornená na Obrázku 4.2. Ako stratová funkcia pre hodnoty kvality sa využíva *Smooth L1-Loss*, ktorá je vyjadrená ako

$$Loss_{SmoothL1} = \frac{1}{N} \sum_{i=1}^N \begin{cases} \frac{(q_i - F(x_i))^2}{2\beta} & \text{ak } |q_i - F(x_i)| \leq \beta \\ |q_i - F(x_i)| - \frac{1}{2}\beta & \text{inak} \end{cases}. \quad (4.14)$$

$F(x_i)$ predstavuje hodnotu predikovaného skóre kvality, a teda len výstup časti siete určujúcej kvalitu pre vstup x_i . Hodnota β predstavuje prah pri ktorom sa vyberie medzi stratovými funkciami L1 alebo L2. Stratová funkcia funkcia celého modelu sa následne definuje ako

$$Loss_{CRFIQA} = Loss_{ArcFace} + \lambda Loss_{SmoothL1}. \quad (4.15)$$

Parameter λ predstavuje pomer, ktorým sa podieľajú jednotlivé stratové funkcie na výslednej hodnote stratovej funkcie podľa ktorej sa optimalizujú parametre modelu. Autori odporúčajú hodnotu tohto parametru vyladiť, keďže hodnoty $Loss_{SmoothL1}$ sú hlavne na začiatku tréovanie výrazne menšie ako hodnoty $Loss_{ArcFace}$. Toto môže viesť pri tréovaní ku zameraniu modelu len na hodnotu $Loss_{ArcFace}$. Hodnota parametru $\lambda = 10$ je využívaná autormi pri experimentoch.

4.5 MagFace

Kvalitu snímky je zároveň možné zistiť aj zo samostatného vektora príznakov. Hlavným faktorom je práve magnitúda tohto vektora. V predošlej kapitole bola spomenutá jedna z možných metód pre učenie s využitím metrík. Metrika popísaná v tejto sekcii sa snaží okrem zmenšenia rozptylu vektorov príznakov danej identity rovnako optimalizovať aj magnitúdu vektora pre určenie kvality, čím má výstižný názov MagFace [24]. Metóda narozdiel od ArcFace využíva informácie o magnitúde vektorov v latentnom priestore počas tréningu, podľa ktorej adaptívne mení parametre stratovej funkcie. Cieľom tejto zmeny je získanie reprezentácie vektorov zároveň reprezentujúcich svoju kvalitu. Jeden výstup siete môže byť využitý pre rozpoznávanie a zároveň aj pre odhad kvality. Výsledné vektory pre vstupy s nízkou kvalitou budú bližšie pri strede priestoru a vstupy s vyššou kvalitou budú od stredu ďalej.

Pre popis algoritmu využijem podobné označenie ako pri popise ArcFace zo sekcie 3.2. Predpokladám teda vektor príznakov z_i , v latentnom priestore s cieľovou triedou y_i . Následne predpokladám vektory váh w , predstavujúce stredy jednotlivých tried. Metóda rovnako zisťuje hodnotu $a_i = \|z_i\|$, ktorá reprezentuje magnitúdu vektora v latentnom priestore. Naďalej sa využíva metrika kosínusovej vzdialenosti $\cos(\theta_{y_i}) = \frac{z_i w_{y_i}^T}{\|z_i\| \|w_{y_i}\|}$. Pre túto metódu sa však nevyužíva L2 normalizácia, ktorá bola prítomná pri ArcFace. Hodnotu stratovej funkcie funkcie je možné vypočítať ako

$$Loss_{Magface} = -\frac{1}{N} \sum_{i=1}^N \log \frac{e^{s \cos(\theta_{y_i} + m(a_i))}}{e^{s \cos(\theta_{y_i} + m(a_i))} + \sum_{j \neq y_i} e^{s \cos(\theta_j)}} + \lambda_g g(a_i). \quad (4.16)$$

Pre účely adaptívnych zmien hodnôt stratovej funkcie vzhľadom na magnitúdu vstupu sú v rovnici pridané dve funkcie. Funkcia $m(\cdot)$ vyjadruje hodnotu adaptívneho okraja, ktorý je vyhodnotený pre hodnotu magnitúdy vektoru príznakov. Funkcia $g(\cdot)$ má následne úlohu adaptívnej regularizácie. Parameter λ_g je využitý pre ladenie efektu regularizácie voči adaptívnemu okraju.

Pre funkciu $m(\cdot)$ musí platiť, že je to rastúca konvexná funkcia. Funkcia $g(\cdot)$ musí byť klesajúca konvexná funkcia. Autori uvádzajú ako príklady možných funkcií

$$m(a_i) = \frac{u_m - l_m}{u_a - l_a} (a_i - l_a) + l_m, \quad (4.17)$$

$$g(a_i) = \frac{1}{a_i} + \frac{1}{u_a^2} a_i, \quad (4.18)$$

v týchto príkladoch je funkcia $m(\cdot)$ reprezentovaná rastúcou lineárnou funkciou a funkcia $g(\cdot)$ predstavuje hyperbolu. Následne sú definované dva parametre u_a a l_a . Musí platiť, že vstup $a_i \in [u_a, l_a]$, a teda definičný obor oboch funkcií je považovaný za tento interval. Pre hodnoty l_m a u_m platí vzťah $m(u_a) = u_m$ a $m(l_a) = l_m$.

Hodnoty $u_a, l_a, u_m, l_m, \lambda_g$ môžu byť zvolené pri tréningu tak, aby spĺňali predom spomenuté podmienky. Pre experimenty autori zvolili $u_a = 110$, $l_a = 10$, $\lambda_g = 35$. Hodnoty parametrov u_m a l_m sú stanovené s viacerými pokusmi, pri ktorých je potrebné tieto parametre vyladiť. Autori dosiahli najlepšie výsledky s nastavením $u_m = 0.8$ a $l_m = 0.4$.

Kapitola 5

Dátové sady

Výber a vytvorenie vhodnej dátovej sady je základným krokom k správne vyhodnoteniu biometrických systémov. Zároveň je pre tréning neurónových sietí potrebné veľké množstvo anotovaných dát. Pre účely tejto práce je potrebné mať dátové sady obsahujúce snímky dúhovky a informácie o osobe, od ktorej bol snímok získaný. Zároveň je potrebné vedieť, z ktorej strany tváre bol snímok dúhovky získaný. Pravá a ľavá dúhovka sú natoľko odlišné, že sa nedajú porovnať ich vektory príznakov ako pochádzajúce od jednej osoby. Z tohto pohľadu je vhodné, aby databázy mali pravé a ľavé snímky oddelené.

V oblasti databáz obsahujúcich snímky dúhoviek je dostupných mnoho zdrojov z ktorých nie všetky sú ideálne pre využitie v tejto práci. Dátové sady sa odlišujú v tom, či využívajú viditeľné spektrum, alebo prípadne infračervené. Rozdiely môžu byť aj vo veľkosti símanej oblasti tváre. Jeden snímok môže takto obsahovať napríklad aj celú tvár. Preto je vhodné zamerať sa na jeden typ databáz. Zároveň je vhodné zameranie len na infračervené spektrum. Toto spektrum svetla je štandardné pre rôzne komerčne dostupné snímacie zariadenia.

Okrem dátových sád využívaných pre tréning je rovnako potrebné získať sady pre overenie verifikácie. Ako bolo spomenuté v predošlých kapitolách, tak pre tento štýl vyhodnotenia je potrebné aby dátová sada obsahovala okrem snímok aj informácie o pároch. Za týmto účelom je možné využiť rovnaký typ dátových sád ako pri tréningu. Dátová sada musí spĺňať podmienku, že neobsahuje rovnaké identity ako v trénovacej sade. Páry jednotlivých pozitívnych a negatívnych príkladov sa následne dajú vytvoriť zo známych anotácii snímok.

V nasledujúcich častiach kapitoly sa popíšu parametre vybraných dátových sád zároveň aj s krátkym zhrnutím informácii o nich. Na záver bude popísaný proces vytvorenia trénovacej a verifikačnej dátovej sady. Zároveň budú popísané niektoré dodatočné informácie ku dátovým sadám. Bolo vybraných celkovo 6 dátových sád, z ktorých boli všetky poskytnuté firmou Innovatrics s.r.o.

CASIA IrisV4 [3] je rozšírenie dátovej sady IrisV3 a v tejto verzii obsahuje sady rôznych reálnych a virtuálnych subjektov. Sada obsahuje zároveň šesť rôznych podskupín.

- **CASIA-Iris-Interval** je dátová sada obsahujúca snímky z infračervenej kamery s využitím kruhového osvetlenia.
- **CASIA-Iris-Lamp** obsahuje snímky snímané ručným snímačom.
- **CASIA-Iris-Twins** obsahuje snímky 100 párov dvojčiek.

- **CASIA-Iris-Distance** obsahuje snímky snímané diaľkovým snímačom pre multimoďálnu biometriu.
- **CASIA-Iris-Thousand** obsahuje snímky snímané kamerou IKEMB-100.
- **CASIA-Iris-Syn** obsahuje synteticky generované snímky dúhovky v infračervenom spektre.

IITD [20] je dátová sada obsahujúca snímky dúhovky študentov a pracovníkom na univerzite Indian Institute of Technology, Delhi. Dátová sada je pomerne malá. Vytvorená bola v roku 2007 a obsahuje zastúpenie osôb oboch pohlaví vo vekovom rozmedzí od 14 do 55 rokov. Snímky boli získavané kamerou JIRIS JPC1000.

UTIRIS [12] táto dátová sada obsahuje okrem iného zároveň snímky dúhoviek vo viditeľnom a infračervenom svetle. Ako bolo spomenuté, tak pre účely tejto práce sa využijú jedine infračervené snímky. Snímky pochádzajú z roku 2007 a bola vytvorená v University of Tehran. Snímky boli vytvorené kamerou ISG Lightwise LW. Databáza obsahuje pomerne malé množstvo subjektov a celkovo je v databáze 542 snímok.

University of Notre Dame Iris CrossSensor 2013 ¹ obsahuje snímky dúhoviek zosnímaných počas 27 rôznych sedení. Celkovo táto sada obsahuje 29986 snímok zo senzoru LG4000 a 116564 snímok zo senzoru LG2200. Každý subjekt sa nachádza aspoň v dvoch snímkach. Dáta boli zbierané v rokoch 2008 až 2010. Väčšina obrázkov v tejto dátovej sade má veľkosť 640×480 pixelov.

University of Notre Dame Iris 0405 [5] je podobne ako predošlá dátová sada snímaná kamerou LG2200. Snímky v tejto sade pochádzajú z rokov 2004-2005. Pre každý subjekt je prítomný snímok pravej a ľavej dúhovky. Subjekty sú vo vekovom rozmedzí 18 až 75 rokov, zároveň táto dátová sada pokrýva obidva pohlavia a viacero etník.

Proprietárna dátová sada Innovatrics Okrem predchádzajúcich dátových sád je vďaka spolupráci na diplomovej práci s Innovatrics s.r.o. sprístupnená aj proprietárna dátová sada. Pri získavaní tejto dátovej sady bol kladený dôraz na to, aby bola podchytená variabilita, ku ktorej môže dochádzať pri snímaní dát. Dátová sada takto pokrýva viacero scenárov ktoré obsahujú zmeny v pomere zakrytia dúhovky a rozťahnutia zornice. Zároveň pre snímanie bolo využitých viacero rôznych snímačov ako TD100, Tascent, BMT-20, Crossmatch a IriMagic.

¹<https://cvrl.nd.edu/projects/data/#nd-crosssensor-iris-2013-data-set>

Tabuľka 5.1: Rozdelenie snímok v dátových sadách

Dátová sada	Trénovanie		Verifikácia			
	Snímky	Triedy	Sní.	Tr.	Páry	
					Súhl.	Nesúh.
CASIA	15890	1600	-	-	-	-
IITD	-	-	2240	435	4338	43380
Innovatrics	-	-	14639	390	2730	26826
N. D. 0405	51880	570	6071	78	2832	56543
N. D. 2013	101968	940	37552	404	4035	39776
UTIRIS	542	108	-	-	-	-

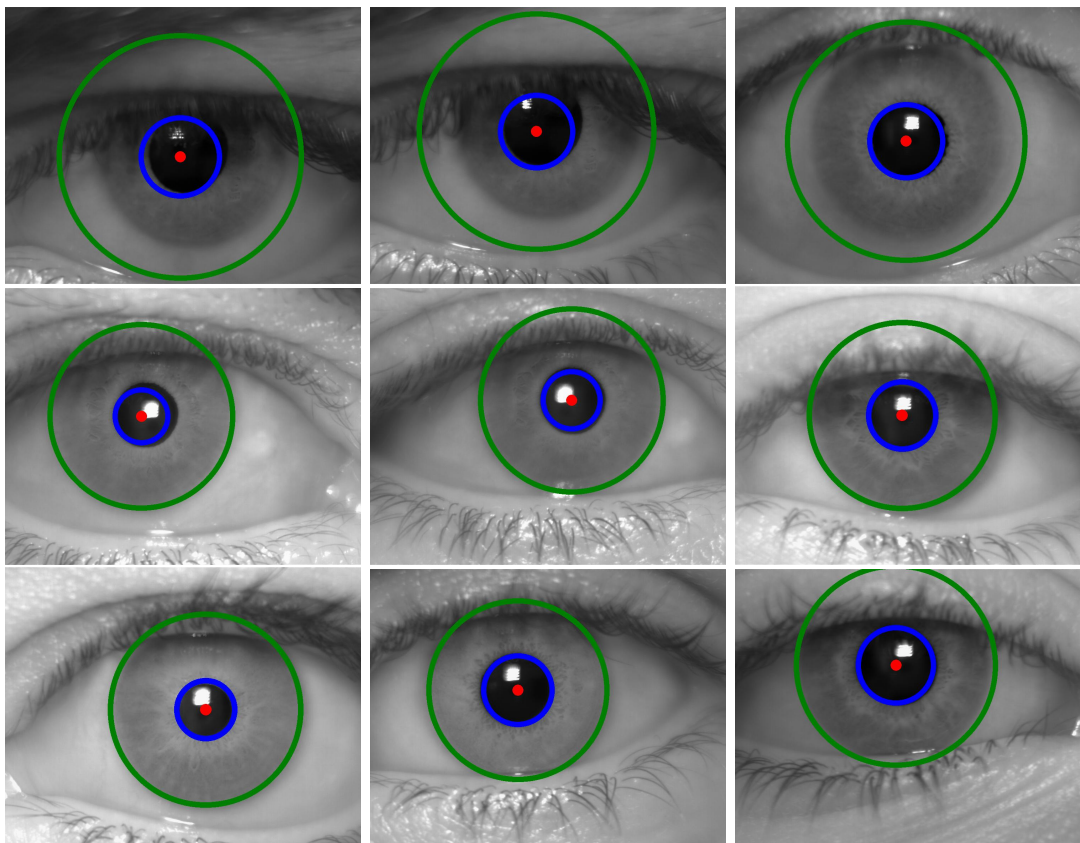
5.1 Trénovacia dátová sada

Samotnú trénovaciu dátovú sadu je potrebné mať čo ideálne najväčšiu možnú. Z tohto dôvodu sa pre tréning sietí bude využívať kombinácia viacerých dátových sád. Tento prístup zároveň zaručí väčšie množstvo snímok, tried, pokrytie rôznych etník a rôznych typov kamier. Celkové zloženie dátovej sady pozostáva zo štyroch menších dátových sád. Rozpis používaných dátových sád a počty snímok, ktoré sa v nich nachádzajú je možné vidieť na tabuľke 5.1. Pre tréning sa vyberali hlavne dátové sady ktoré sú relatívne veľké, ako sú dátové sady od Notre Dame a CASIA. Celkovo kombinovaná sada obsahuje 3218 tried, kde jedna trieda predstavuje kombináciu človeka a pravej alebo ľavej dúhovky. Celkovo sa v dátovej sade nachádza až 170280 snímok dúhoviek.

5.2 Verifikačné dátové sady

Pre verifikáciu sú vybrané štyri dátové sady. Ako je možné vidieť v tabuľke 5.1, tak niektoré dátové sady sú rovnaké ako v trénovacej sade. Dáta pre verifikáciu boli z týchto dátových sád získané výberom disjunktných identít. V princípe sa jedná o proces, pri ktorom sa dátová sada rozdelí na dve takým spôsobom, aby v oboch častiach neboli rovnaké identity. Tento princíp je využitý pri sadách od University of Notre Dame. Zvyšné dátové sady sú využité celé pre verifikáciu.

Okrem klasického označenia identity sú pre všetky verifikačné dátové sady vygenerované páry súhlasných a nesúhlasných snímok. S týmito informáciami je následne možné porovnávať metódy verifikácia a efektu odhadu kvality na verifikáciu podľa metodiky popísanej v sekcii 2.3.



Obr. 5.1: Ukážka snímok z dátovej sady UTIRIS [12] s označením pozície dúhovky

5.3 Označenie pozície dúhovky

Dátové sady pridávajú k jednotlivým snímkam aj viacero informácií, ktoré môžu byť potenciálne využité aj na iné účely ako je verifikácia a tréning klasifikácie. Dátové sady obsahujú aj informácie o zafarbení oka, prípadne o etniku osoby. Alternatívne je možné do dátovej sady pridať aj informácie využiteľné pre pred-spracovanie snímok.

Firma Innovatrics s.r.o. poskytuje k jednotlivým dátovým sadám viacero rozširujúcich informácií. Jednou z týchto informácií je aj lokalizácia dúhovky. Dúhovka je lokalizovaná nasledujúcimi parametrami, ktoré budú používané v ďalších kapitolách, (c_{i_x}, c_{i_y}) označuje stred dúhovky, r_{iris_i} polomer dúhovky a r_{pupil_i} polomer zornice.

Údaje pozície sú v reálnom systéme generované neurónovou sieťou zo vstupných snímok. Pre každú snímku dátovej sady je na tento účel použitá detekčná neurónová sieť od Innovatrics. Údaje sú staticky uložené v metadátoch danej snímky. Z pohľadu tejto práce umožnia tieto informácie experimentovať s rôznou transformáciou vstupných snímok.

Ukážka takto lokalizovaných dúhoviek na snímkach je zobrazená na obrázku 5.1. Je možné vidieť menšiu kružnicu reprezentujúcu zornicu a bod v jej strede označujúci stredný bod. Väčšia kružnica zobrazuje hranice dúhovky. Táto reprezentácia je menej presná tým, že pre jednotlivé hranice využíva kružnice. Samotná dúhovka môže byť lepšie charakterizovaná ako elipsa. Z obrázku je možné vidieť aj možné nepresnosti detektoru a to hlavne na jednotlivých snímkoch v prvom riadku. Z pohľadu tejto práce a experimentov s jednotlivými sieťami je však reprezentácia pomocou kružníc detekovaných týmto spôsobom postačujúca.

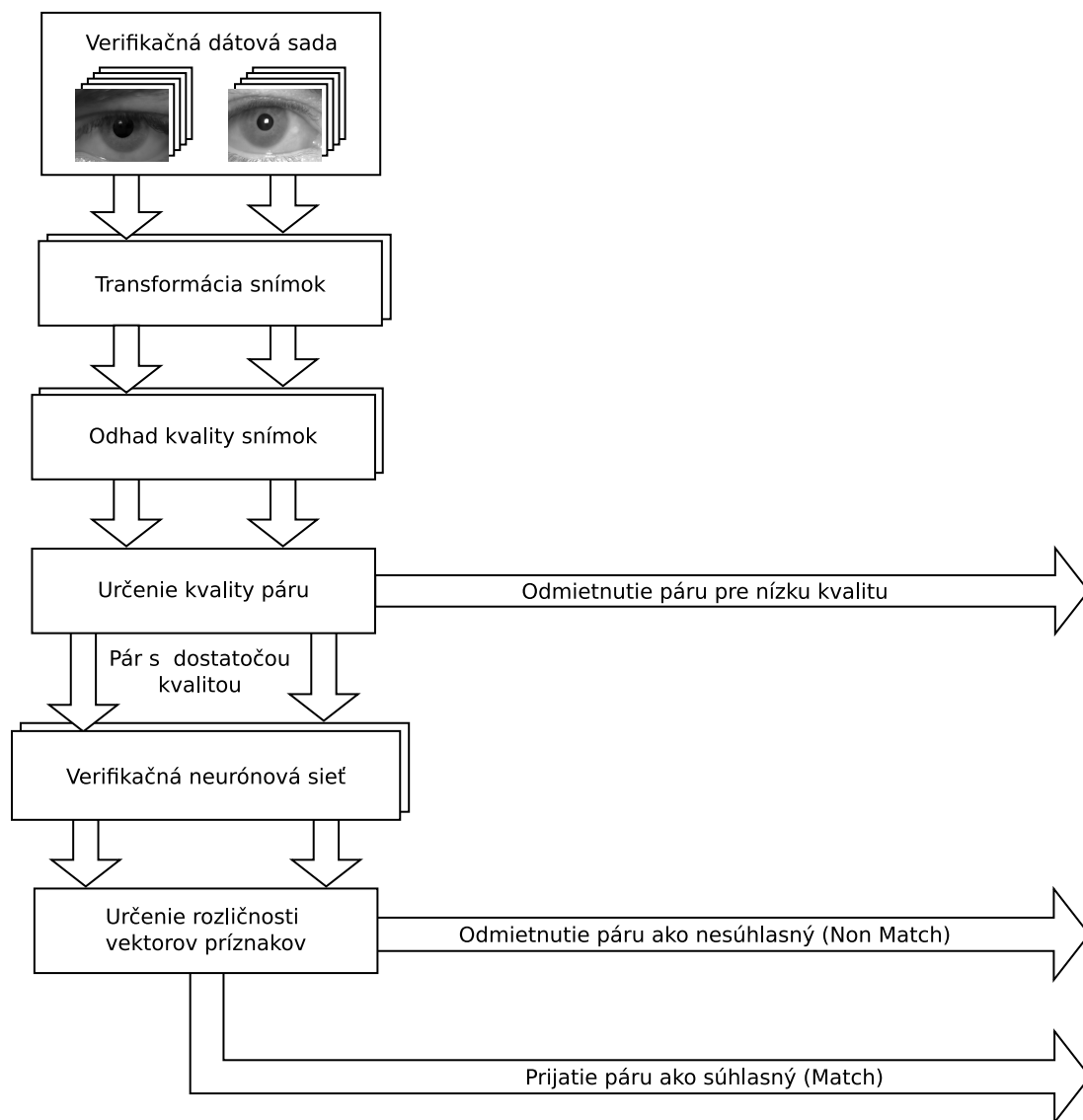
Kapitola 6

Návrh riešenia a experimentov

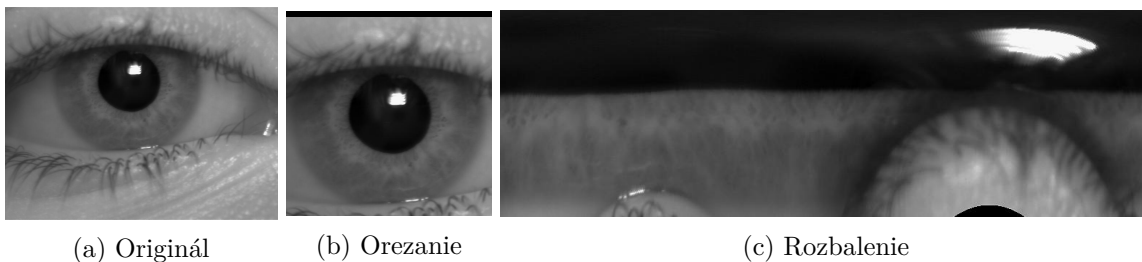
V svojej diplomovej práci som sa rozhodol experimentálne vyskúšať aplikovanie vybraných metód odhadu kvality využívaných pri tvárach na dúhovky. Podľa mojich aktuálnych znalostí tieto metódy neboli ešte na oblasť rozpoznávania osôb podľa dúhovky aplikované, očakávam teda, že dosiahnuté výsledky z mojej práce budú mať informačný prínos. V navrhovaných experimentoch aplikujem metódy MagFace [24], CR-FIQA [4], SDD-FIQA [26], SER-FIQ [35] a PFE [30]. V oblasti rozpoznávania tvári tieto metódy dosahujú state-of-the-art výsledky. Kde použitie týchto metód ukázalo zlepšenie chyby verifikačného systému naprieč rôznymi dátovými sadami. Príkladom môže aplikácia na dátovej sade dátovej sade LFW, kde zlepšenie FNMR pri $FMR=1 \times 10^{-3}$ a pri 10% odmietnutých pároch je z približne 2.9×10^{-3} na približné hodnoty MagFace 1.7×10^{-3} , CR-FIQA 2.1×10^{-3} , SDD-FIQA 1.4×10^{-3} , SER-FIQ 1.6×10^{-3} a PFE 2×10^{-3} . V oblasti identifikácie pomocou snímkov dúhovky majú tieto metódy zatiaľ nepreskúmané výsledky. Výberom jednotlivé metódy pokrývajú rôzne prístupy k určovaniu skóre kvality, čím môžem dosiahnuť zaujímavé vzájomné porovnanie. Metódy zároveň porovnávam voči odhadu kvality snímky dúhovky pomocou DFS [38]. DFS najskôr aplikujem na jednotlivé verifikačné sady pre získanie porovnávacích hodnôt, ktoré považujem za aktuálny state-of-the-art.

Metódy odhadu kvality sú využívané v navrhnutom experimentálnom biometrickom systéme. Pri tomto návrhu som pracoval s predpokladom, že systém pre odhad kvality a systém pre verifikáciu musí byť vzájomne nezávislý. Toto mi umožňuje výmenu rôznych verifikačných neurónových sietí a ich kombináciu s rôznymi systémami pre odhad kvality. Zároveň mi to umožní krížové porovnania jednotlivých metód. Krížové porovnanie môžem následne použiť pri metódach MagFace, CR-FIQA, SER-FIQ a SDD-FIQA, kde vytváram systém v ktorom jedna verifikačná neurónová sieť môže získavať skóre kvality pre úplne inú sieť.

Samotný návrh obsahuje päť hlavných častí. Celkovú schému verifikačného systému je možné vidieť na obrázku 6.1. Schéma systému popisuje priebeh spracovania a jednotlivé kroky, ktoré ovplyvňujú výsledky systému. Systém rieši transformáciu vstupných dát, následné vyhodnotenie kvality a verifikáciu. Pracuje s verifikačnými dátovými sadami popísanými v sekcii 5.2. Vo výsledku vie systém vyhodnotiť, či daný pár snímkov z verifikačnej dátovej sady bude odmietnutý na základe príliš nízkej kvality jednej snímky, či patrí rovnakej dúhovke a teda je súhlasný, alebo či bude vyhodnotený ako nesúhlasný. Podľa jednotlivých výstupov systému pre celú dátovú sadu spolu so zvolenými prahmi odlišnosti a kvality sa s informáciou o skutočnom vzťahu snímkov v páre, môže vyhodnocovať úspešnosť verifikácie. Výstupy tohto systému zároveň umožňujú následne sledovať metriky FNMR a FMR pre rôzne konfigurácie zvolených prahov.



Obr. 6.1: Schéma experimentálneho verifikačného systému



Obr. 6.2: Ukážka orezania a rozbalenia snímky dúhovky

6.1 Transformácia snímok

Z pohľadu identifikácie podľa snímok dúhovky som sa rozhodol pracovať s dvoma možnosťami transformácie snímok. Snímky, ktoré pochádzajú priamo z dátovej sady obsahujú z veľkej časti oblasti tváre, ktoré nie sú využiteľné pre identifikáciu podľa dúhovky. Tieto časti sú napríklad oko, časti nosa, obočie a podobné. Okrem iného, tieto časti tváre môžu byť premenlivé a to aj zo dňa na deň. Príkladom môže byť úprava obočia prípadne použitie rôznych kozmetických úprav. Pre tento prípad sa budem snažiť minimalizovať plochu samotnej tváre a brať do úvahy čo najviac oblasť dúhovky.

Prvý navrhnutý spôsob transformácie využíva informácie o pozícii a veľkosti dúhovky popísaných v sekcii 5.3. V tomto prípade využijem hodnoty c_{i_x} , c_{i_y} a r_{iris_i} . Následne podľa týchto hodnôt orežem vstupnú snímku z dátovej sady. Orezaná snímka má tvar štvorca s dĺžkou hrany $2r_{iris_i}$, čo predstavuje priemer dúhovky. Geometrický stred vybraného štvorca sa je na pozícii (c_{i_x}, c_{i_y}) . Hrany štvorca sú paralelné s originálnymi hranami obrázka. Na obrázku 6.2b je možné vidieť, že po aplikácii tejto transformácie obsahuje snímka z väčšej časti len dúhovku.

Druhý navrhnutý spôsob transformácie využíva rozbalenie. V princípe je táto metóda projekcia z kartézkych do polárnych súradníc. Podobne ako pri orezaní, tak aj tu využívam hodnoty c_{i_x} , c_{i_y} a r_{iris_i} získané z dátovej sady. Pre samotnú transformáciu využívam rovnicu pre rozbalenie

$$x_{i_{unrolled}}(\theta, r) = x_i \left(c_{i_x} + r \cdot \sin \left(\frac{\theta}{r_{iris_i}} \right), c_{i_y} + r \cdot \cos \left(\frac{\theta}{r_{iris_i}} \right) \right), \quad (6.1)$$

kde x_i , originálny predstavuje vstupný snímok a $x_{i_{unrolled}}$ jeho rozbalenú podobu. Parametre týchto funkcií predstavujú pozície v smere osi x a y , následne vrátia hodnotu snímky v tomto vybranom bode. Samotná výsledná snímka ma veľkosť $2\pi r_{iris_i} \times r_{iris_i}$. Výsledok rozbalenia dúhovky je možné vidieť na obrázku 6.2c. Avšak aj v tomto prípade je možné vidieť nedokonalosť tejto metódy, napríklad rozbalenie viečka a zornice. Pre experimenty sa bude využívať táto jednoduchšia varianta s tým, že experimentujem jedine s odstránením rotačnej zložky zo vzoru dúhovky.

Následne v experimentoch budem porovnávať efekt orezania snímku s jednoduchým rozbalením. V oboch prípadoch sa zmenení veľkosť výsledných snímok na 224×224 . Z tohto pohľadu ma zaujíma hlavne porovnanie, či rozbalenie zlepši celkovú presnosť verifikačného systému bez určenia kvality. Typ transformácie, ktorý bude dávať lepšie výsledky následne použijem aj pre systémy určovania kvality.

6.2 Modul odhadu kvality snímok

Modul v navrhovanom systéme získava skóre kvality. Reprezentuje sa ním všetkých 6 vybraných metód odhadu kvality. Pre konkrétny experiment sa vždy využije práve jedna metóda. Jednotlivé metódy využívajú rôzne konfigurácie sietí. Pre state-of-the-art metódu zo sekcie 2.2 využívam referenčnú konfiguráciu a pred-trénovaný model poskytnutý autormi. Pri zvyšných metódach som zvolil konfigurácie sietí, ktoré plynú z výsledkov s verifikačnou neurónovou sieťou, alebo boli využívané v referenčných experimentoch od autorov. V tejto sekcii následne popisujem rôzne prístupy, ktoré som využil pri získavaní informácií o kvalite pre jednotlivé metódy.

Výpočet kvality pre MagFace Niektoré metódy určenia kvality vyžadujú previesť výsledné hodnoty inej formy ako je výstup samotnej siete. Pre metódu MagFace je potrebné stanoviť skóre kvality z vektora príznakov. Pre skóre kvality využívam dĺžku vektora

$$q_{i_{MagFace}} = \|z_i\|, \quad (6.2)$$

podobne ako v predošlých sekciách, tak z_i predstavuje vektor príznakov získaný z verifikačnej neurónovej siete.

Určenie kvality pomocou PFE Výstupom metódy PFE popísanej v sekcii 4.1 je vektor rozptylov pravdepodobnostného rozloženia. Podobne ako pri MagFace, tak aj v tomto prípade potrebujem získať skóre kvality z vektora. Využívam na to harmonický priemer hodnôt z vektora rozptylov

$$q_{i_{PFE}} = \frac{D}{\sum_{l=1}^D \frac{1}{\sigma_i^{2(l)}}}, \quad (6.3)$$

rovnako ako pri sekcii 4.1 hodnota σ_i^2 predstavuje vektor rozptylov odhadnutého pravdepodobnostného rozloženia, $\cdot^{(l)}$ predstavuje prístup ku hodnote z dimenzie l a D je počet dimenzii vektoru. Tento spôsob získania kvality využívam z dôvodu, že ho popisujú autori metódy vo svojich experimentoch.

Z pohľadu tréningu modulu neistoty, ktorý počíta vektor rozptylov, som najskôr potreboval navrhnúť úpravu tréningovej dátovej sady, tak aby bolo možné použiť efektívny výpočet stratovej funkcie popísanej v rovnici 4.4. Nutnosť úpravy vychádza priamo z prevzatej implementácie. Implementácia vyžaduje, aby sa v jednom batch nachádzalo viacero snímok z rovnakej triedy. Páry sa následne vytvoria podľa označenia toho, ku ktorej triede jednotlivé snímky patria. Štandardne je tréningová dátová sada navrhnutá tak, aby jeden batch obsahoval rôzne snímky, pri ktorých výbere nezáleží na ich konkrétnej triede a nie je garantované, že v jednom batch bude rovnaká trieda viac-krát. Pre vytvorenie tréningovej dátovej som navrhol úpravu poradia snímok v tréningovej sade, ktorá je popísaná algoritmom 2. V novej dátovej sade sa nachádza vždy v 5 snímok v skupinách z rovnakej triedy. Týmto je garantované, že pri výbere bude batch obsahovať viacero snímok z rovnakých tried.

Algoritmus 2: Úprava dátovej sady pre tréovanie pravdepodobnostnej reprezentácie vektorov vlastností

Input: Originálna dátová sada D_{old}
Output: Upravená dátová sada D
 $D \leftarrow$ prázdny zoznam;
while $D_{old} \neq \emptyset$ **do**
 $cls \leftarrow$ náhodne vybraná trieda z dátovej sady D ;
 for j *in* $1 \dots 5$ **do**
 $x \leftarrow$ náhodne vybraný snímok z dátovej sady D patriaci triede cls ;
 $D.append(x)$;
 end
end
return D

Určenie kvality pomocou SER-FIQ Pre určenie kvality pomocou metódy SER-FIQ popísanej v sekcii 4.3 využívam kombináciu verifikačných neurónových sietí s rôznymi architektúrami. Na tieto siete aplikujem algoritmus 1. Z konkrétnych architektúr používam iResNet, čo je sieť využívaná priamo autormi. Implementácia autorov využíva sieť, ktorá na výstup posledných vrstiev konvolučných sietí aplikuje drop-out a následne plne prepojenú vrstvu. Experimentálne vyskúšam, aký efekt má táto metóda, ak využijem kombináciu priemeru výstupu poslednej konvolučnej vrstvy, za ním aplikovaný drop-out a následne posledná plne prepojená vrstva. Zároveň z pohľadu experimentov skúsím aplikáciu SER-FIQ na verifikačnú sieť MobileNet, čoho efekt nebol ku môjmu vedomiu zatiaľ zistený.

Určenie kvality pomocou SDD-FIQA Metóda SDD-FIQA zo sekcie 4.2 trénuje sieť pomocou generovaných označení kvality snímok. Ako architektúru tejto siete som zvolil iResNet50 s jedným výstupom. Sieť je väčšia verzia neurónovej siete iResNet18, ktorú používajú autori. Väčšiu sieť som zvolil z dôvodu toho aby počet vrstiev nijako neovplyvnil výsledky pre odhad kvality. Sieť je pred-tréovaná pre verifikáciu. Týmto prístupom zrýchlim tréning. Sieť je následne do-tréovaná na tréovacej sade snímok ku ktorým som pridal informácie o určenej kvalite. Pre vytvorenie označení kvality snímok z tréovacej dátovej sady som vytvoril algoritmus 3. Z hľadiska možných experimentov som sa rozhodol používať pre generovanie označení kvality rôzne verifikačné neurónové siete. Experimentálne následne skúsím efekt, ktorý majú výstupy menších sietí ako iResNet18 a MobileNetV3 v určovaní kvality pre väčšiu sieť iResNet50 a naopak.

Určenie kvality pomocou CR-FIQA Táto metóda popísaná v sekcii 4.4 je mierne odlišná od predchádzajúcich metód, pretože sa okrem iného spoločne so skóre kvality trénuje aj samotná rozpoznávací sieť. Pomocou tejto metódy trénujem siete iResNet50 a iResNet18 s plne prepojenou vrstvou a s priemerovaním hodnôt poslednej vrstvy. Zároveň používam aj sieť MobileNetV3-Large. Siete sú pred-tréované pre rozpoznávanie. Pri experimentovaní následne využijem jedine výstup, ktorý tieto siete dávajú pre skóre kvality, keďže z originálnych sietí môžem mať výsledky priamo.

Algoritmus 3: Generovanie SDD-FIQA skóre pre snímok x_i

Input: Rozpoznávací model F , snímok x_i , množina snímkov od rovnakej identity P_{x_i} a od nesúhlasných identít N_{x_i}

Output: Skóre kvality q_i

$z_i \leftarrow F(x_i)$;

$W_i \leftarrow$ prázdny zoznam;

for k **in** $1 \dots K$ **do**

$S_{x_i^k}^{P_m} \leftarrow$ prázdny zoznam;

$P_m \leftarrow$ náhodný výber m snímkov z množiny P_{x_i} ;

for x_j **in** P_m **do**

$z_j \leftarrow F(x_j)$;

$S_{x_i^k}^{P_m}.$ *append* $(\text{CosineSimilarity}(z_i, z_j))$;

end

$S_{x_i^k}^N \leftarrow$ prázdny zoznam;

$N_m \leftarrow$ náhodný výber m snímkov z množiny N_{x_i} ;

for x_j **in** $N_{x_i^k}$ **do**

$z_j \leftarrow F(x_j)$;

$S_{x_i^k}^N.$ *append* $(\text{CosineSimilarity}(z_i, z_j))$;

end

$W_i.$ *append* $(\mathbb{W}\mathbb{D}(S_{x_i^k}^{P_m} || S_{x_i^k}^N))$;

end

$q_i \leftarrow \text{Average}(W_i)$;

return q_i

6.3 Určenie kvality páru snímok

Samotné metódy určenia kvality dávajú rozdielne výstupy s rozličnými usporiadaniami hodnôt. Pre niektoré metódy znamená menší výstup horšiu kvalitu a pre iné lepšiu. Tento blok som navrhol tak aby sa pre zvolenú metódu určenia kvality upravilo výsledné skóre tak aby ich bolo možné jednotne porovnávať. Skóre, ktoré dávajú metódy SER-FIQ, PFE a metóda DFS, sú negatívne korelované so skóre kvality. To znamená, že vstup s lepšou kvalitou dáva nižšie hodnoty a naopak vstup s horšou kvalitou zase vyššie. Metódy ako určenie magnitudy vektoru, SDD-FIQA a CR-FIQA sú na druhej strane pozitívne korelované s kvalitou snímok. Z pohľadu porovnávania popísaného v sekcii 2.3 je potrebné používať práve pozitívnu koreláciu skóre kvality. Pre metódu SER-FIQ využívam priamo metódu používanú autormi, ktorá tieto hodnoty prevedie do požadovanej formy vzťahom

$$q_{i_{new}} = \sigma(\alpha \cdot (\sigma(-q_i) - r)), \quad (6.4)$$

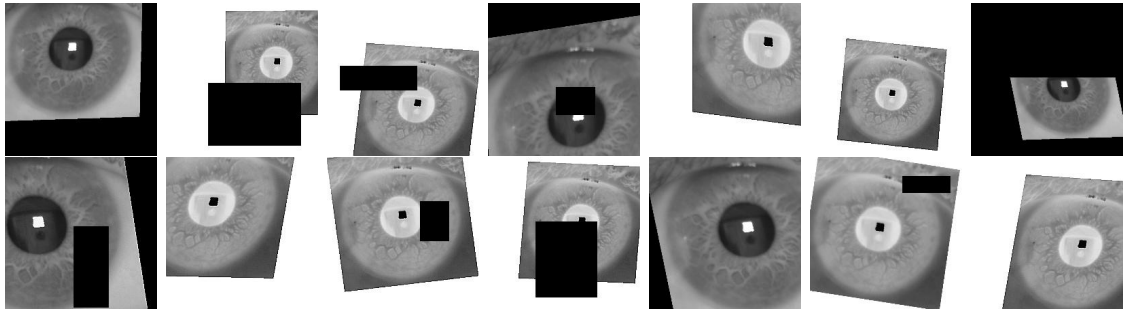
kde q_i predstavuje skóre kvality získané algoritmom 1, parameter α predstavuje škálovací faktor, r predstavuje posuv skóre a funkcie $\sigma(\cdot)$ predstavujú funkciu sigmoidy. Pri metódach PFE a DFS používam zápornú hodnotu originálnej návratovej hodnoty skóre

$$q_{i_{new}} = -q_i. \quad (6.5)$$

Skóre kvality páru snímok predstavuje minimálnu hodnotu zo skóre kvality získaných pre jednotlivé snímky. Zvolený prah následne určuje, či je pár prijatý na ďalšie spracovanie alebo nie. Páry s hodnotami skóre pod prahom sú odmietnuté a hodnoty nad prahom prijaté.

6.4 Modul verifikačných neurónových sietí

Pre účely verifikácie využívam konvolučnú neurónovú sieť pre rozpoznávanie. Uvažujem využitie architektúr konvolučných sietí iResNet[9] a MobileNetV3[13], ktoré boli popísané v sekcii 3.1. Siete trénujem pre rozpoznávanie pomocou ArcFace [8] popísaného v sekcii 3.2 a MagFace zo sekcie 4.5. Niektoré metódy extrakcie príznakov pre dúhovky [6] využívajú architektúru sietí ResNet [10], s touto myšlienkou využívam práve vylepšené verzie tejto architektúry iResNet. Zároveň bol demonštrovaný tréning siete MobileNetV3 práve pre extrakciu príznakov dúhovky pomocou ArcFace [21]. V práci navrhujem použiť aj iné stratové funkcie ako MagFace a tým môžem zistiť, ako sú tieto siete porovnateľné pri verifikácii. Motiváciou je taktiež ukázať efekt a porovnanie na presnej avšak veľkej sieti iResNet50, s podobnou sieťou s menším počtom parametrov iResNet18 a s na rýchlu sieťou MobileNetV3-Large. Pri sieťach iResNet zároveň preskúvam možnosti použitia rôznych konfigurácií posledných vrstiev využívaných pre získanie vektoru príznakov. Dve uvažované možnosti sú buď plné prepojenie z predchádzajúcej konvolučnej vrstvy, alebo priemer hodnôt poslednej vrstvy a následná aplikácia plne prepojenej vrstvy na vektor priemerných hodnôt. Týmto ukážem, či je pre identifikáciu podľa dúhovky potrebné využívať len časť vstupnej snímky alebo snímku ako celok. Tieto siete následne používam, ako základ rôznych metód určovania kvality alebo kvalitu získavam priamo z ich výstupu. Z tohto dôvodu som zvolil tento postup v porovnaní s postupom zo zadania.



Obr. 6.3: Ukážka rôznych augmentácií aplikovaných na jeden snímok dúhovky

6.5 Určenie rozličnosti páru snímok

Majme dva rôzne vstupné snímky x_i a x_j , pre ktoré existujú dva vektory príznačkov z_i a z_j . Tieto vektory sú výsledkom verifikačnej neurónovej siete. Pre určenie ich podobnosti viem zaviesť viacero možných porovnaní. Pre použitie v experimentálnom systéme som zvolil metódu porovnania pomocou kosínusovej podobnosti,

$$\text{CosineSimilarity}(z_i, z_j) = \frac{z_i z_j^T}{\|z_i\| \|z_j\|} \quad (6.6)$$

Táto metóda porovnania je z pohľadu akým sa trénujú vektory príznačkov najlepšia. Metódy ArcFace a MagFace rozmiestňujú vektory na hyper-guli, tým pádom je porovnanie vektorov podľa ich uhlovej vzdialenosti najlepšia reprezentácia podobnosti. Zároveň keďže táto hodnota určuje podobnosť a nie rozličnosť, tak využívam zápornú hodnotu výsledku. Takto zachovávam usporiadanie nad hodnotami a teda menšie hodnoty predstavujú menšiu rozličnosť a väčšie hodnoty zase väčšiu.

6.6 Augmentácia trénovacích dát

Jedna z posledných informácií, ktorá je súčasťou návrhu je aj augmentácia vstupných dát. Táto časť ide mimo samotný verifikačný systém, ale je dôležitou súčasťou pri tréningu neurónových sietí. Pre účeli tejto práce používam kombináciu viacerých náhodných úprav vstupných snímok. Účel týchto úprav je zvýšenie počtu trénovacích vstupov čím sa môže dosiahnuť väčšia presnosť pri rôznych metódach. Augmentácia je aplikovaná pre tréning rozpoznávacích sietí a tréning metódou CR-FIQA. Pre samotnú augmentáciu som si zvolil náhodnú afinnú transformáciu (posuv, škálovanie, rotácia, skosenie), náhodnú úpravu ostrosti, náhodnú automatickú úpravu kontrastu, náhodnú inverziu farieb a náhodné vymazanie časti snímku. To, či táto augmentácia zlepšuje presnosť, zistím experimentálne pri vyhodnotení verifikačných neurónových sietí. Na obrázku 6.3 je možné vidieť aplikáciu týchto rôznych kombinácií augmentácií na jeden orezaný snímok dúhovky.

Kapitola 7

Implementácia



Obr. 7.1: Prehľad využitých technológií

Pre vykonanie experimentov a ich vyhodnotenie je potrebné jednotlivé metódy implementovať, prípadne integrovať do navrhnutého riešenia. Táto kapitola približuje niektoré implementačné detaily a popisuje princípy, ktoré viedli implementáciu v tejto forme. Pri implementácii využívam technológie, ktorých prehľad je možné vidieť na obrázku 7.1. V implementácii využívam programovací jazyk Python vo verzii 3.10. Používam aj viacero rôznych knižníc pre funkcie ako tréning neurónových sietí (PyTorch, PyTorch Lightning), načítanie dát (Pandas), výpočty s maticami (NumPy) a vykresľovanie grafov (Plotly, Matplotlib).

7.1 Štruktúra kódu

V tejto sekcii popisujem logické rozloženie kódu, ktorý používam v diplomovej práci. Toto rozloženie je vedené hlavne využitím knižnice PyTorch Lightning. Knižnica je postavená nad PyTorch a prichádza s novými úrovňami abstrakcie, ktoré sú využiteľné pre tréning. Kód používaný na tréning a všetky pomocné funkcie sa nachádzajú v Python module, ktorý som implementoval pod zložkou označenou ako *lightning*. Následné sekcie popisujú rôzne pod-moduly, ktoré bolo potrebné implementovať a prípadne použiť z iných riešení.

Načítanie dátových sád je zabezpečené v module *dataset*. Tento modul implementuje načítanie jednotlivých dátových sád spoločne s ich meta-dátami ako sú trieda, ktorej daná snímka patrí, pozícia dúhovky a databáza súhlasných/nesúhlasných párov. Načítavajú sa dátové sady popísané v kapitole 5. Pre samotné načítanie dátových sád bol použitý PyTorch modul *VisionDataset* ako základ. Nad ním je postavená jedna základná trieda dátovej sady pre načítanie snímok dúhovky v ktorej zároveň implementujem metódy pre rozbalenie a orezanie snímky. Následne implementujem dátovú sadu pre načítanie a spojenie tréningových sád, a dve triedy pre rôzne verzie verifikačných dátových sád. Pri načítaní tréningovej

dátovej sady môže aplikovať algoritmus 2, pre získanie dátovej sady vo formáte využívanom pri tréningu PFE [30]. Taktiež je pre trénovanie možné načítať alternatívne označenie jednotlivých trénovacích snímok, využité pri trénovaní pomocou SDD-FIQA [26].

Konvolučné neurónové siete predstavujú Python modul *models*. V tomto module sa nachádzajú implementácie konvolučných sietí iResNet [9] a MobileNetV3 [13], ktoré sa používajú v experimentoch. Jednotlivé siete majú implementované rovnaké rozhranie. Toto rozhranie je zabezpečené využitím rovnakých inicializačných parametroch a použitím rovnakých typov vstupných/výstupných dát. Implementácie sietí boli prevzaté od autorov iResNet¹, MagFace² a z repozitára reimplementujúceho MobileNetV3 do PyTorch³. Jednotlivé siete som následne mierne upravil, tak aby bola umožnená efektívna implementácia metódy SER-FIQ [35]. Úprava spočívala v pridaní možnosti kopírovania výstupu poslednej konvulčnej vrstvy so zvoleným počtom kópií.

Stratové funkcie sú definované v module *metrics*. Nachádzajú sa tu hlavne implementácie stratových funkcií ArcFace [8]⁴, MagFace [24], PFE⁵ a CR-FIQA [4]⁶. Jednotlivé implementácie stratových funkcií sú prevzaté priamo od autorov. Pre lepšie možnosti výmeny jednotlivých funkcií som prevzaté implementácie integroval do samostatných modulov. Implementované moduly stratových funkcií MagFace, ArcFace a CR-FIQA obsahujú aj lineárnu vrstvu, ktorá sa používa len pri tréningu.

Trénovacie moduly sa používajú na vytvorenie komplexnejších sietí. Kombinujú zároveň modely konvolučných sietí, moduly stratových funkcií používaných pre tréning a konfiguráciu optimalizátorov. Okrem iného umožňujú definovať rôzne výstupy pre inferenciu, trénovanie a validáciu. Je to jedna z abstrakcií z knižnice PyTorch Lightning. Pomocou nej implementujem moduly pre trénovanie rozpoznávacích sietí, CR-FIQA, PFE a SDD-FIQA. Zároveň implementujem aj modul pre inferenciu pomocou referenčného modelu určovania kvality DFS [38] prevzatého od autorov⁷.

Trénovanie všetkých modulov je implementované v jednom Python skripte. Trénovací skript sa nakonfiguruje pomocou YAML súboru. Konfiguráciou sa vyberie trénovací modul, konfigurácia optimalizátorov, konfigurácia trénovacej sady a konfigurácia rôznych transformácií a argumentácii. Týmto spôsobom sa zredukoval počet duplicitného kódu, zároveň sa vytvorila prehľadnejšia a udržateľnejšia implementácia s jednoduchosťou prípadného rozšírenia. Skript zároveň umožňuje export výsledku inferencii sietí a export netrénovaných modulov konvolučných sietí. Inferencia prebieha priamo na verifikačných dátových sádach. Pri tréningu sa zároveň využívajú rôzne možnosti knižnice PyTorch Lightning ako ukládanie check-pointov a logovanie.

¹<https://github.com/iduta/iresnet>

²<https://github.com/IrvingMeng/MagFace>

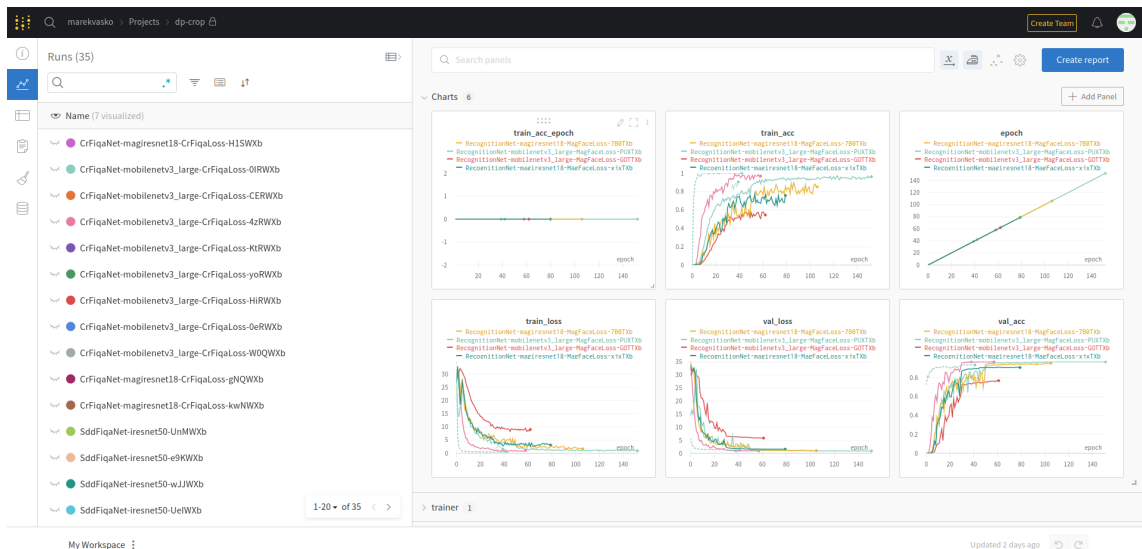
³<https://github.com/d-li14/mobilenetv3.pytorch>

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

⁵<https://github.com/0ntheway361/pfe-pytorch>

⁶<https://github.com/fdbtrs/CR-FIQA>

⁷<https://github.com/Debatrrix/DFSNet>



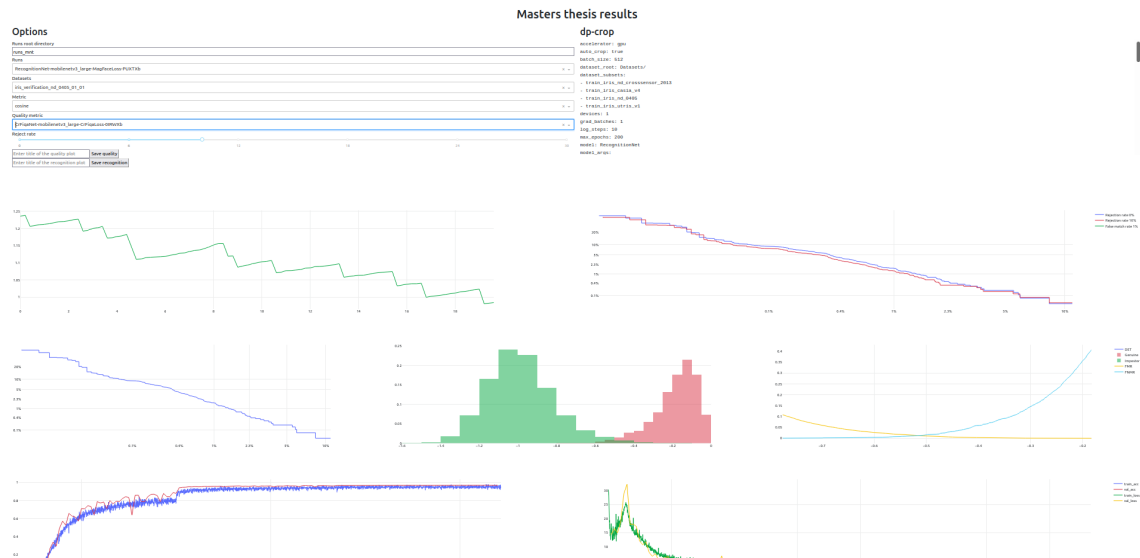
Obr. 7.2: Ukážka sledovania priebehu tréningu pomocou Weights & Biases

Pomocné aplikácie predstavujú Python skripty, ktoré som implementoval pre metódy SDD-FIQA a SER-FIQA. Pre SDD-FIQA som implementoval skript pre vytvorenie trénovacích označení skóre kvality pomocou algoritmu 3. Pre zvolenú rozpoznávaciu neurónovú sieť skript nájde uložené vektory vlastností získaných z inferencie na trénujúcej dátovej sade, a následne pre snímky vypočíta vzdialenosti distribučných funkcií. Získané skóre sa uloží do súboru, ktorý sa následne môže načítať spoločne s dátovou sadou. Druhý pomocný skript implementuje metódu SER-FIQA popísanú algoritmom 4.9. Pre tento skript som využil referenciu poskytnutú autormi napísanú v MXNet⁸, ktorú som re-implementoval v PyTorch. Pri implementácii zároveň využívam úpravu konvolučných sietí, ktorú som spomenul v predošlých sekciách.

7.2 Sledovanie priebehu tréningu

Pri trénovaní je taktiež dôležité sledovanie metrík ako hodnoty stratovej funkcie prípadne presnosť výsledkov modelu. Takto sa zistí, či sa model trénuje správne alebo nie. Zároveň sa môžu sledovať aj iné metriky ako stav hardware, čím sa môže efektívne využitie výkonu. V tejto práci som použil systém Weights & Biases. Tento systém poskytuje cloudovú platformu s vizualizáciou priebehu tréningu pomocou grafov. Zároveň sa ku jednotlivým trénujúcim behom ukládajú zvolené konfigurácie sietí, optimalizátorov a iné parametre. Výhodou je aj jednoduchá integrácia s Pytorch Lightning. Ukážka užívateľského rozhrania tejto služby so zobrazenými sledovanými trénujúcimi metrikami je na obrázku 7.2.

⁸<https://github.com/pterhoer/FaceImageQuality>



Obr. 7.3: Ukážka implementovaného nástroja pre zobrazovanie výsledkov

7.3 Nástroj pre zobrazenie výsledkov

Vzhľadom na množstvo implementovaných metód som sa pre účely práce rozhodol vytvoriť aj malý nástroj pre prehľadávanie výsledkov. Nástroj je implementovaný v Python skripte pomocou knižnice Plotly Dash. Je interaktívny a umožňuje prehľadávať výsledky natrénovaných modelov. Pomocou webového rozhrania je možné si zvoliť konkrétnu natrénovanú sieť pre rozpoznávanie, k nej vybrať spôsob určenia podobnosti vektorov, dátovú sadu a metódu odhadu kvality. Nástroj následne načíta potrebné dáta pre danú sieť a vykreslí jednotlivé grafy. Vykresľujú sa grafy chyby, DET krivky a krivky FNMR vzhľadom na pomer odmietnutých párov. Jednotlivé krivky je možné vykresliť s rôznymi zvolenými parametrami ako percento odmietnutých párov alebo hodnota FMR. Pri vykreslení DET kriviek po aplikovaní metódy kvality sa zároveň zobrazuje aj krivka samostatnej verifikačnej siete. Okrem týchto grafov sa pre verifikačnú sieť vykresľuje aj priebeh tréningu a rozloženie hodnôt porovnávacej funkcie pri súhlasných a nesúhlasných pároch. Uživatelské rozhranie tejto aplikácie je ukázané na obrázku 7.3.

Kapitola 8

Výsledky experimentov

Pre získanie výsledkov experimentov som aplikoval jednotlivé metódy určovania kvality v navrhnutom verifikačnom systéme. Takto určím, ktoré metódy najviac zlepšujú metriky FNMR a FMR. Pre postup som najskôr zvolil natrénovanie verifikačnej neurónovej siete. Experimentálne zisťujem, aká je najlepšia konfigurácia pre rozpoznávanie podľa dúhovky. Experimentujem s rôznymi metódami rozbalenia, prepojení siete, architektúr, augmentácie a stratových funkcií. Podľa výsledkov následne vyberám najlepšie konfigurácie a s tými pokračujem pri metódach určovania kvality. Pri aplikovaní jednotlivých metód určovania kvality pracujem s tréňovaním rozpoznávacích sietí metódami, ktoré zároveň poskytujú možnosti pre získanie skóre kvality, pracujem so získavaním skóre z pred-tréňovaných sietí pre verifikáciu a skúšam tréňovať siete nad snímkami s označením skóre kvality. Metódy odhadu kvality zároveň aplikujem na rôzne verifikačné siete. Keďže niektoré metódy využívajú práve verifikačné siete tak viem dosiahnuť krížové porovnanie a teda zistiť či menej presné siete vedú pri určovaní skóre kvality zlepšiť presnosť presnejších sietí a opačne.

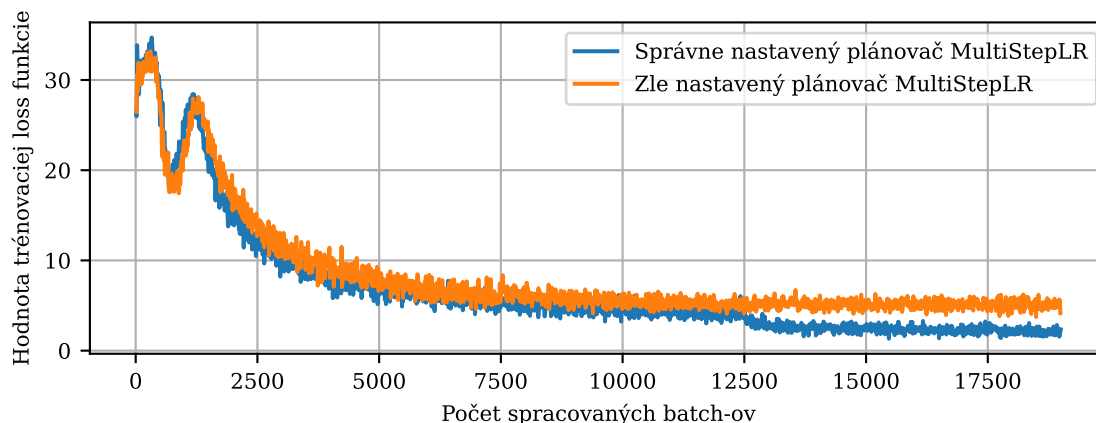
Hlavnou úlohou tejto kapitoly je ukázať výsledky jednotlivých experimentov, bližšie priblížiť jednotlivé parametre, ktoré sú využívané pri jednotlivých experimentoch a ukázať výsledky, podľa ktorých som sa v priebehu experimentovania rozhodoval. Metódy odhadu kvality budú porovnávané s state-of-the-art metódou DFS popísanej v sekcii 2.2. Tréňovanie jednotlivých sietí prebieha nad dátovými sadami popísanými v sekcii 5.1. Metriky FNMR a FMR vyhodnocujem na verifikačných dátových sadoch, University of Notre Dame Iris CrossSensor 2013 (N. D. 2013), University of Notre Dame Iris 0405 (N. D. 0405), IITD a Innovatrics popísaných v sekcii 5.2. Všetky dátové sady používané pri tréňovaní a verifikácii sú poskytnuté firmou Innovatrics s.r.o.

8.1 Výsledky verifikačných neurónových sietí

Samotným cieľom práce je získanie prehľadu, toho ako dané metódy vplývajú na presnosť verifikačnej neurónovej siete. Väčšina metód zároveň využíva sieť ktorá je veľmi podobná ak nie identická s touto sieťou. A preto som som na úvod zvolil experimenty s cieľom získať čo najlepšie verifikačné siete. Pri tréningu jednotlivých sietí využívam metódu optimalizácie parametrov pomocou stochastického gradientného zostupu (SGD). K optimalizátoru som zvolil aj plánovač miery učenia Multi-step learning rate scheduler. Výber som založil na využití jednotlivých metód autormi ArcFace [8]. Pri tréningu používam konfigurácie optimalizátoru s learning rate 0.1, parametrom weight-decay 0.00001 a momentum 0.9. Parametre optimalizátoru preberám od autorov ArcFace. Nastavenie plánovača learning rate

Tabuľka 8.1: Konfigurácia MultiStepLR pre využívané siete

Sieť	Mílniky			Gamma
iResNet50	30	45	65	0.1
iResNet18	50	100	150	0.1
MobileNetV3-Large	50	100	150	0.1



Obr. 8.1: Ukážka efektu nastavenia plánovača learning rate

pre siete založené na iResNet18, iResNet50 a MobileNetV3-Large je popísané v tabuľke 8.1. Konfiguráciu plánovača som určil experimentovaním s tréningom pri rôznych nastaveniach mílnikov plánovača. Úloha plánovača je vynásobiť mieru učenia po počte epoch určenými mílnikmi hodnotou gamma. Pre zistenie vhodných mílnikov som najskôr skúsil tréning s prvotným nastavením a sledoval som priebeh hodnôt stratovej funkcie. Nastavenie som následne upravil, tak aby sa jednotlivé mílniky dosiahli vždy po ustálení hodnôt stratovej funkcie. Ukážka experimentu je zobrazená na obrázku 8.1. Tento experiment som vykonal pre každú sieť, avšak pri sieťach MobileNetV3-Large a iResNet18 som dosiahol s rovnakým nastavením podobný priebeh hodnôt stratovej funkcie a teda ďalej nebolo potrebné pokračovať s experimentmi. Nastavenie plánovača a optimalizátora je pri všetkých nasledujúcich experimentoch rovnaké.

Experimentálne zisťujem efekt tréningu s augmentáciou a bez. Pri tréningu s augmentáciou dat využívam nastavenie je zobrazené na tabuľke 8.2. Pre každú sieť využívam normalizáciu hodnôt vstupných snímok do rozloženia so strednou hodnotou 0.5 a smerodajnou odchýlkou 0.25. Normalizáciu aplikujem pri tréningu pred náhodným vymazávaním. Transformácie popísané v sekcii 6.1 aplikujem pred augmentáciou. Po augmentácii sa výsledný obrázok transformuje do veľkosti 224×224 . Pre tréning sa následne vytvára batch snímok s veľkosťou 512. Pre niektoré siete je však táto veľkosť príliš veľká vzhľadom na využívaný hardvér. Zvolil som teda možnosť akumulácie gradientov cez viacero batch-ov. Gradienty sa vždy akumulujú tak, aby vo výsledku bolo pred spustením optimalizátora spracovaných 512 snímok. Výsledná sieť je trénovaná pre extrakciu vektoru príznačkov s veľkosťou 512 hodnôt.

V procese trénovania zároveň využívam rozdelenie trénovacej dátovej sady do troch častí. Pred týmto rozdelením sú jednotlivé snímky náhodne zamiešané. Prvá časť dátovej sady obsahuje 75% snímok je využívaná pre samotný tréning. Druhú časť dátovej sady s

Tabuľka 8.2: Parametre augmentácie dát

Typ augmentácie	Parameter	Hodnota
Náhodná úprava ostrosti	Pravdepodobnosť	50%
	Faktor ostrosti	2
Náhodná rotácia	Rozpätie uhlov otočenia	$\pm 10^\circ$
Náhodná zmena veľkosti	Percentuálne rozpätie náhodnej veľkosti	[30%, 130%]
Náhodné skosenie	Rozpätie uhlov skosenia	$\pm 5^\circ$
Náhodný posuv	Rozpätie posuvu v smere osi x	$\pm 10\%$
	Rozpätie posuvu v smere osi y	$\pm 10\%$
Náhodný automatický kontrast	Pravdepodobnosť úpravy	50%
Náhodné vymazávanie	Pravdepodobnosť vymazania	50%
	Rozsah podielu vymazanej plochy	[2%, 33%]
	Rozsah pomeru strán vymazanej plochy	[0.33, 3]
Náhodná inverzia farieb	Pravdepodobnosť inverzie	50%

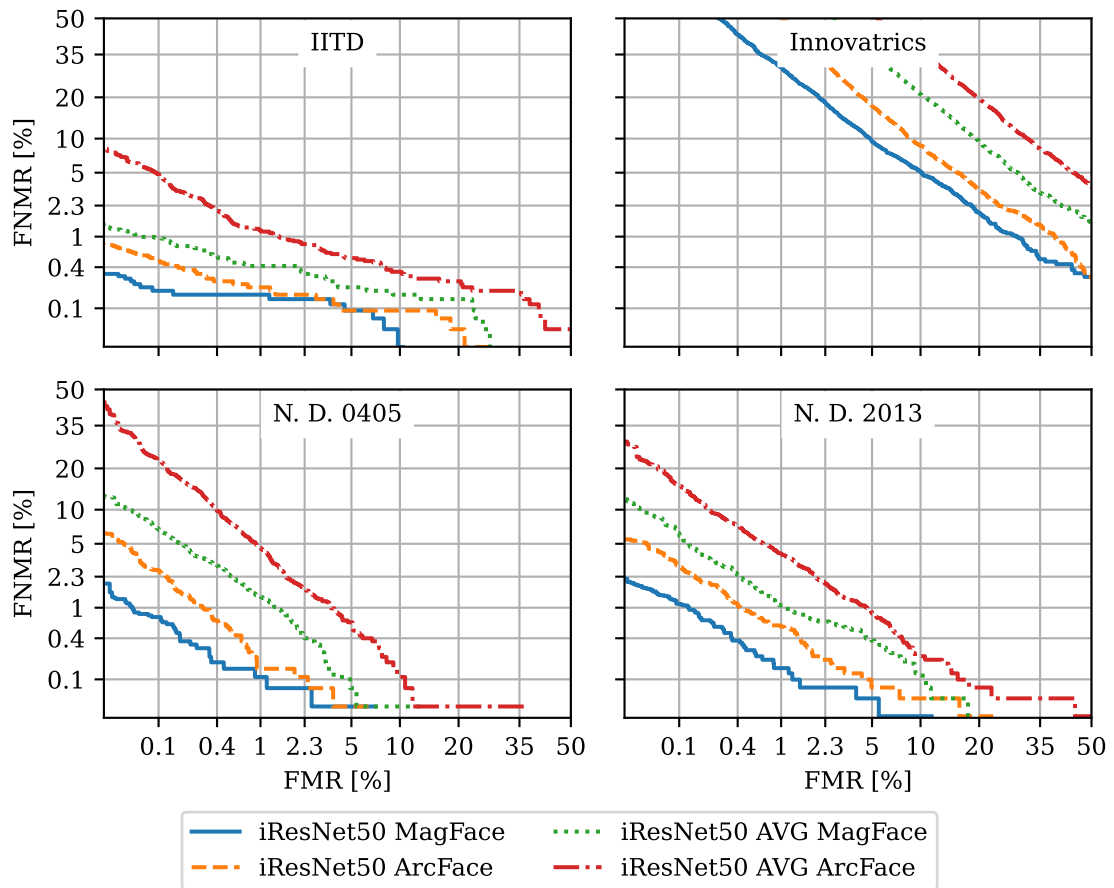
20% snímkami používam na konci epochy ku vyhodnoteniu hodnôt stratovej funkcie. Posledná časť obsahuje 5% snímkov a používam ju na začiatku k overeniu platnosti konfigurácie jednotlivých sietí, čo je súčasťou knižnice využívannej pre tréning. Toto rozdelenie mi umožňuje sledovať priebeh tréningu na dátach pre ktoré sieť nebola optimalizovaná a tak môžem odhaliť, či sa sieť nepretrénuje.

Siete následne trénujem po dobu maximálne 200 epoch pri iResNet18 a MobileNetV3-Large. Pri sieti iResNet50 trénujem maximálne na 80 epochách. z experimentov pri tréningu som zistil, že menšie siete potrebujú viac epoch na dosiahnutie nemennej hodnoty stratovej funkcie na 20% vyhodnocovacích dát a teda podľa toho som zvolil aj tieto maximálne počty epoch.

Vplyv rôznych stratových funkcií a konfigurácií posledných vrstiev sietí V prvom experimente som sa rozhodol zistiť ako ovplyvňuje výsledky verifikácie použitie stratových funkcií ArcFace a MagFace [24]. Zároveň overujem, či je pre vytvorenie vektora príznakov potrebné použitie plne prepojenej vrstvy z poslednej konvolučnej vrstvy siete, alebo či postačuje využitie priemernej hodnoty a následné aplikovanie plne prepojenej vrstvy nad vektorom. Vo výsledkoch sledujem efekt tejto konfigurácie ako na tréning pomocou MagFace tak aj ArcFace. Pri tomto prvotnom experimente som sa rozhodol používať ako základ konvolučnú sieť iResNet50. Vo výsledkoch následne označujem jednotlivé siete ako iResNet50 ArcFace, iResNet50 MagFace, iResNet50 AVG ArcFace, iResNet50 AVG MagFace, kde symbol AVG predstavuje použitie priemerných hodnôt poslednej vrstvy. Všetky siete sa trénujú s využitím drop-out [32] s pravdepodobnosťou 0.5. Zároveň pracujem s orezanými snímkami dúhovky.

Tabuľka 8.3: Plochy pod DET krivkami pre všetky verifikačné dátové sady s využitím siete iResNet50 s rôznym prepojením posledných vrstiev, trénovanej pomocou ArcFace a MagFace.

$[\times 10^{-4}]$	IITD	Innovatrics	N. D. 0405	N. D. 2013
iResNet50 AVG ArcFace	14.60	1164.07	20.66	20.79
iResNet50 AVG MagFace	4.92	707.23	5.50	7.10
iResNet50 ArcFace	2.31	349.15	1.48	3.10
iResNet50 MagFace	1.06	205.37	0.63	0.98



Obr. 8.2: Verifikačná chyba v závislosti na použitých stratových funkciách a konfiguráciách posledných vrstiev. Obrázok ukazuje porovnanie DET kriviek na všetkých verifikačných dátových sadoch pre sieť iResNet50 trénovanou pomocou ArcFace, MagFace pri konfigurácii s plne prepojenou vrstvou a s priemerovaním hodnôt poslednej vrstvy.

Výsledky experimentov nad verifikačnými dátovými sadmi je možné vidieť v tabuľke 8.3 a na obrázku 8.2. Z hodnôt pod DET krivkami jednotlivých sietí vyplýva, že sieť iResNet50 s plne prepojenou poslednou vrstvou, trénovaná pomocou MagFace má najnižšiu verifikačnú chybu naprieč všetkými dátovými sadami. Zároveň konštatujem, že všeobecne dáva trénovanie sietí pomocou MagFace lepšiu verifikačnú chybu ako pri ArcFace. Taktiež môžem povedať, že využitie plne prepojenej vrstvy je rovnako lepšie ako použitie priemeru výstupu. Z hľadiska experimentov mám odpovede na otázku, či pri identifikácii postačuje len časť snímky dúhovky v porovnaní s celou snímkou. Keďže plne siete splne prepojenou vrstvou majú nižšiu verifikačnú chybu ako siete využívajúce priemer hodnôt poslednej vrstvy tak je samozrejme lepšie použiť plne prepojené vrstvy a teda celú snímku dúhovky. Z pohľadu ďalšieho experimentovania mám informácie o štýle tréningu a typu sietí od ktorých môžem očakávať lepšie výsledky. Pri nasledujúcich experimentoch budem využívať hlavne siete trénované pomocou MagFace a kde bude možné, tak aj s plne prepojenou poslednou vrstvou.

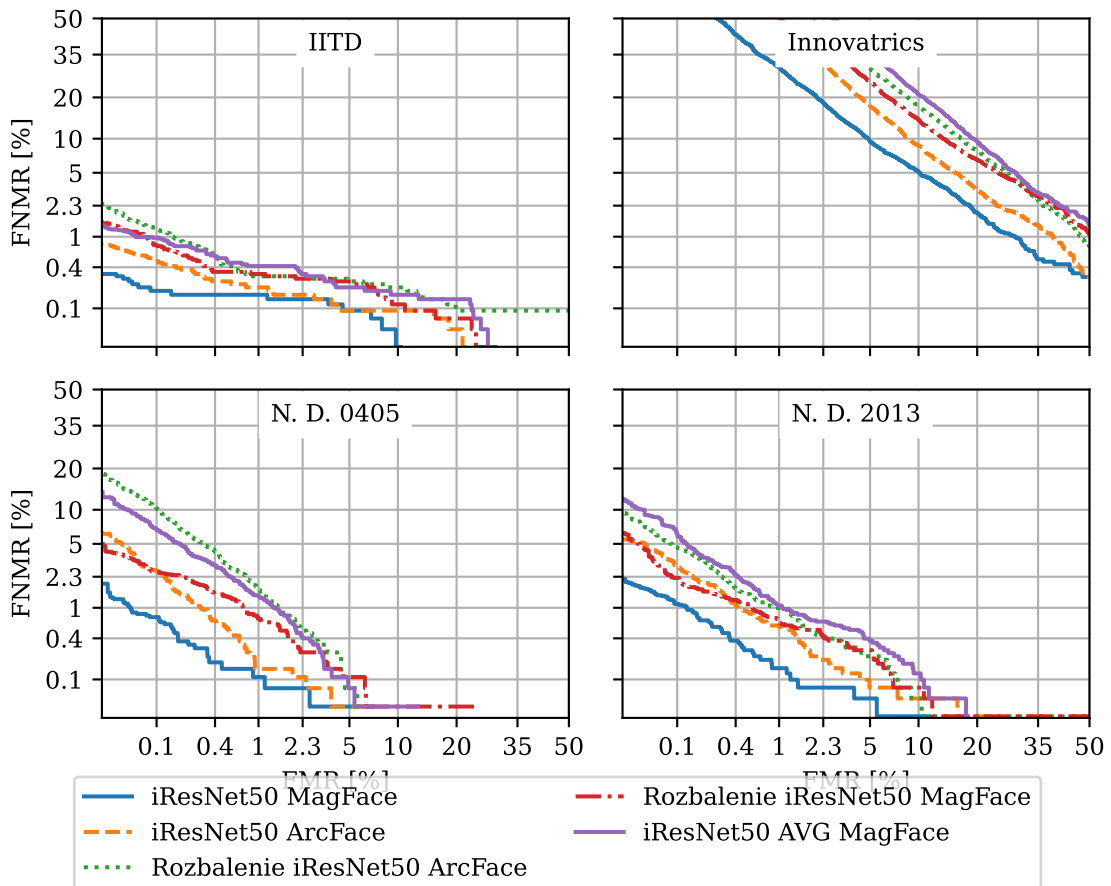
Vplyv rozbalenia snímky dúhovky V sekcii 6.1 som popísal možnosti rôznych transformácií snímok dúhovky. Pre experiment som sa rozhodol porovnať, či rozbalenie v porovnaní s orezaním snímku zvýši presnosť verifikačnej siete. Z výsledkov predchádzajúceho experimentu som použil sieť iResNet50 s plne prepojenou poslednou vrstvou. Sieť je v tomto experimente trénovaná pomocou MagFace, skúsil som aj tréning pomocou ArcFace. Využitie iného typu vstupu pre neurónovú sieť takto umožní dodatočne porovnať, či zlepšenie presnosti spôsobené využitím inej stratovej funkcie je nezávislé od vstupu siete. V oboch prípadoch používam augmentácie popísané v tabuľke 8.2.

Nové získané siete označené ako Rozbalenie iResNet50 ArcFace a Rozbalenie iResNet50 MagFace porovnávam spolu so sieťami iResNet50 ArcFace a iResNet50 MagFace z predošlého experimentu. Na obrázku 8.3 a v tabuľke 8.4 je následne možné vidieť, že siete trénované na rozbalených snímkach dosahujú nižšiu presnosť pri verifikácii. Toto platí naprieč všetkými dátovými sadami. Okrem iného by som podotkol, že toto platí pri využití oboch metód trénovania a to ArcFace a MagFace. V niektorých prípadoch je plocha pod DET krivkou až dvojnásobne väčšia v porovnaní s ekvivalentnou sieťou trénovanou na orezaných snímkach. Týmto teda konštatujem, že použitie rozbalenia neuvažujem v ďalších experimentoch.

Z pohľadu vylepšenia výsledkov pri rozbalení sa do budúcnosti ponúkajú aj iné možnosti transformácie, ktoré môžu byť viac sofistikované a zabezpečiť tým lepšie výsledky siete. Mnou navrhnutá transformácia je pomerne jednoduchá a nedokonalá a existuje tu priestor na možné zlepšenia ako odstránenie zornice, odstránenie viečka, segmentácia plochy dúhovky, lepšia detekcia pozície a využitie elíps pre reprezentáciu tvaru dúhovky. Z tohto pohľadu sa poskytuje aj možnosť nového vývoja.

Tabuľka 8.4: Plochy pod DET krivkami pre všetky verifikačné dátové sady s využitím siete iResNet50, trénovanej pomocou ArcFace a MagFace na rozbalených a orezaných snímkach.

$[\times 10^{-4}]$	IITD	Innovatrics	N. D. 0405	N. D. 2013
Rozbalenie iResNet50 ArcFace	9.05	577.04	7.06	5.98
Rozbalenie iResNet50 MagFace	3.78	497.62	3.67	5.03
iResNet50 ArcFace	2.31	349.15	1.48	3.10
iResNet50 MagFace	1.06	205.37	0.63	0.98



Obr. 8.3: Verifikačná chyba v závislosti na použití rôznych transformácií snímkov. Obrázok ukazuje porovnanie DET kriviek na všetkých verifikačných dátových sady pre sieť iResNet50 trénovanou pomocou ArcFace, MagFace na rozbalených a orezaných snímkach.

Vplyv augmentácie na verifikáciu Pri predchádzajúcich experimentoch som skúšal tréning verifikačnej siete na augmentovaných snímkoch. Z pohľadu oblasti rozpoznávania podľa dúhovky nie je určený vplyv vybranej augmentácie na presnosť verifikácii. Z tohto dôvodu som sa rozhodol experimentálne porovnať aký rozdiel v presnosti môžem očakávať ak augmentáciu nepoužijem a prípadne ak aplikujem jedine čiastočnú augmentáciu. V experimente trénujem sieť iResNet50 s plne prepojenou poslednou vrstvou pomocou MagFace. Ako referenciu používam sieť iResNet50 MagFace, ktorá je výsledkom predchádzajúcich experimentov. Táto sieť používa augmentáciu popísanú na tabuľke 8.2. Následne ju porovnávam s ďalšími dvoma sieťami. Prvá sieť nevyužíva žiadnu augmentáciu a druhá používa náhodné škálovanie v rozmedzí 33% až 133%, náhodný posuv o 20% v smere šírky a výšky a skosenie $\pm 10^\circ$. Výber týchto konkrétnych augmentácií nebol špecifický a vznikol skôr náhodne, avšak porovnaním je výsledok zaujímavý.

Na obrázku 8.5 a v tabuľke 8.5 je možné vidieť, že augmentácia zvyšuje presnosť verifikačného systému naprieč všetkými dátovými sadami. Zároveň môžem podotknúť, že aplikovanie čiastočnej augmentácie zvyšuje presnosť na všetkých dátových sadoch okrem IITD, čo je možné vidieť pri porovnaní plôch pod DET krivkami. Plná augmentácia má však stále najlepšiu lepšiu presnosť. Zaujímavé je, že na dátovej sade Innovatrics je výsledná DET krivka pre sieť trénovanú bez augmentácie pod iným uhlom, kde v hodnotách FMR od približne 4% sú hodnoty FNMR nižšie ako pri sieťach trénovanými pomocou augmentácie, celková plocha krivky je však stále o 15% vyššia ako pri plnej augmentácii. Ako záver tohto experimentu budem pri zvyšných experimentoch s verifikačnými sieťami a s metódou tréningu CR-FIQA [4] budem využívať siete trénované pomocou augmentácie .

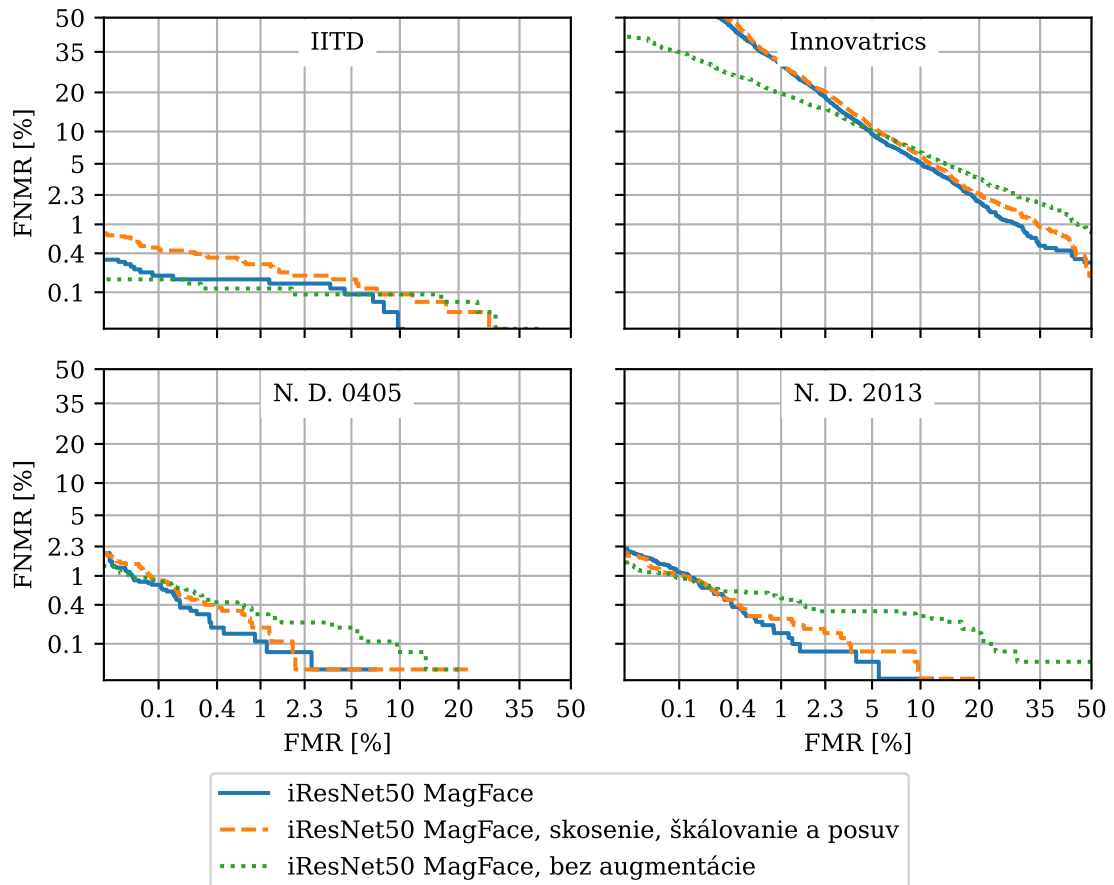
Výsledky menších verifikačných sietí Z pohľadu praktického využitia v biometrickom systéme ma zaujíma aj to, aké výsledky dosahujú menšie neurónové siete. Pre Experimenty som si zvolil MobileNetV3-Large a iResNet18. Pri iResNet18 zároveň skúmam aj efekt použitia priemerných hodnôt poslednej konvolučnej vrstvy v porovnaní s plne prepojenou vrstvou. Siete trénujem nad augmentovanými dátami a konfiguráciou z tabuľky 8.2. Pri sieťach iResNet18 používam drop-out s pravdepodobnosťou 0.5, pri MobileNetV3-Large využívam drop-out s pravdepodobnosťou 0.2, táto pravdepodobnosť je používaná autormi MobileNetV3.

Výsledky pre jednotlivé experimenty sú zobrazené na obrázku 8.5 a v tabuľke 8.6. Je možné vidieť, že vo všeobecnosti sú menšie siete pri verifikácii menej presné ako referenčná sieť iResNet50. Z hodnôt plôch pod DET krivkami vyplýva, že menšie siete majú zároveň aj nižšiu presnosť pri verifikácii. Okrem toho je však dôležité brať do úvahy aj veľkosť jednotlivých modelov, 4.8M pre MobileNetV3-Large, 62.5M pre iResNet18 s plne prepojenou poslednou vrstvou, 82.1M pre iResNet50 s plne prepojenou poslednou vrstvou a 11.4M pre iResNet18 s priemerovaním hodnôt poslednej vrstvy. V tomto kontexte je sieť MobileNetV3-Large pri niektorých dátových sadoch lepšia ako iResNet18 s viac ako dvojnásobným počtom parametrov. Zároveň je aj veľmi blízko pri sieti iResNet18 s plne prepojenou poslednou vrstvou s takmer 13-násobným počtom parametrov.

Z tohto dôvodu pre nasledujúce experimenty využívam sieť MobileNetV3-Large ako reprezentanta rýchlej malej siete. Sieť iResNet50 s plne prepojenou vrstvou považujem za presnú sieť. Sieť iResNet18 s plne prepojenou vrstvou využívam experimentálne aby som zistil, či skóre kvality získané rôznymi metódami z menej presnej siete ovplyvní presnosť väčších sietí.

Tabuľka 8.5: Plochy pod DET krivkami pre všetky verifikačné dátové sady s využitím siete iResNet50, trénovanej pomocou MagFace s použitím rôznych augmentácií.

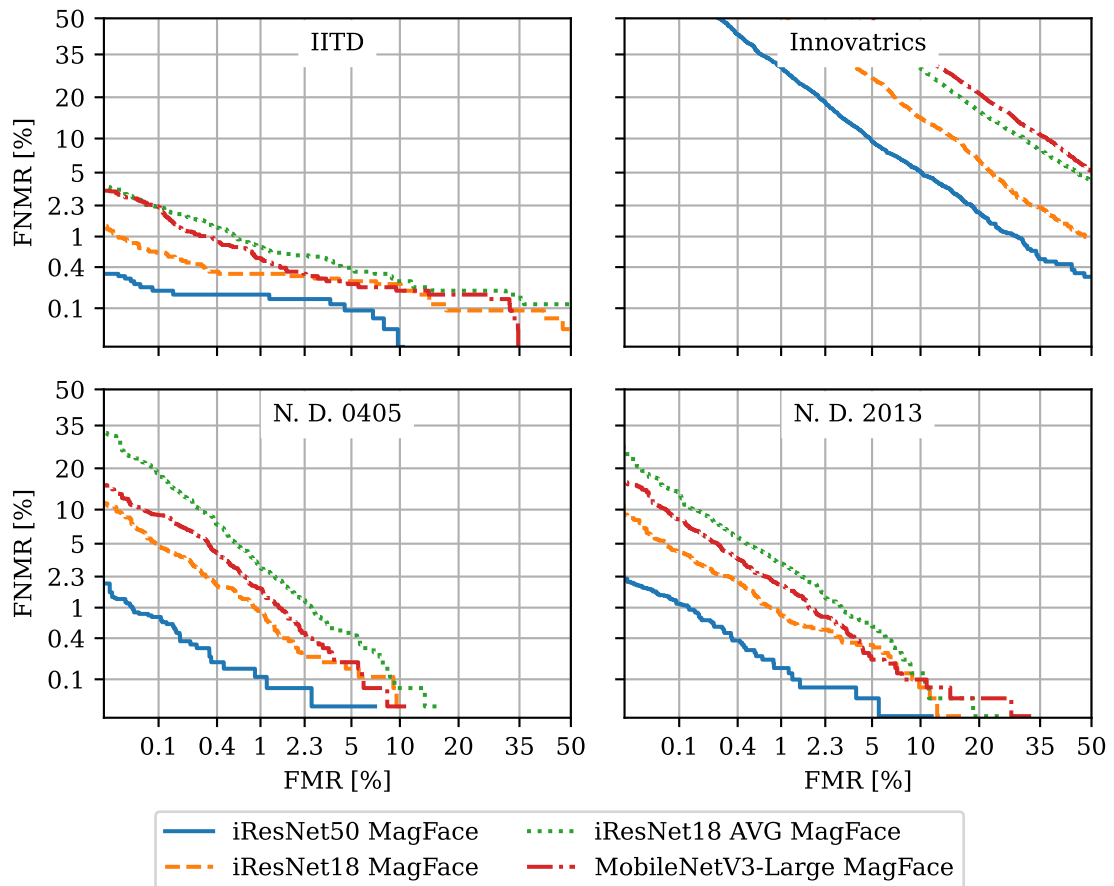
$[\times 10^{-4}]$	IITD	Innovatrics	N. D. 0405	N. D. 2013
iResNet50 MagFace	1.06	205.37	0.63	0.98
iResNet50 MagFace, bez augmentácie	2.59	238.03	2.46	8.26
iResNet50 MagFace, skosenie, škálovanie a posuv	2.78	229.56	1.33	1.63



Obr. 8.4: Verifikačná chyba v závislosti na použití rôznych augmentácií snímok. Obrázok ukazuje porovnanie DET kriviek na všetkých verifikačných dátových sadoch pre sieť iResNet50 trénovanou pomocou MagFace na augmentovaním, bez augmentovania a s aplikáciou augmentácie skosením, škálovaním a posuvom.

Tabuľka 8.6: Plochy pod DET krivkami pre všetky verifikačné dátové sady s využitím architektúr siete iResNet50, iResNet18 a MobileNetV3-Large, trénovanej pomocou MagFace. Pri sieti iResNet18 sa porovnáva aj efekt plne prepojenej vrstvy v porovnaní s priemernou hodnotou výstupu siete.

$[\times 10^{-4}]$	IITD	Innovatrics	N. D. 0405	N. D. 2013
MobileNetV3-Large MagFace	7.10	1231.49	7.18	9.07
iResNet18 AVG MagFace	14.35	1031.90	14.68	14.29
iResNet18 MagFace	7.09	497.25	4.30	5.31
iResNet50 MagFace	1.06	205.37	0.63	0.98



Obr. 8.5: Verifikačná chyba v závislosti na použití rôznych architektúr sietí. Obrázok ukazuje porovnanie DET kriviek na všetkých verifikačných dátových sadách pre siete iResNet50, iResNet18 a MobileNetV3-Large trénované pomocou MagFace. Pri sieti iResNet18 je ukázaný aj efekt plne prepojenej poslednej vrstvy v porovnaní s priemerom hodnôt a vrstvy pre vytvorenie vektoru vlastností.

Tabuľka 8.7: Hodnoty verifikačnej chyby FNMR pri FMR=0.1% na každej dátovej sade pre vybrané siete iResNet a MobileNetV3, trénované pomocou MagFace alebo ArcFace. Siete sú trénované na orezaných snímkach dúhovky s augmentáciou.

FNMR@FMR=0.1% [%]	IITD	Innovatrics	N. D. 0405	N. D. 2013
MobileNetV3-Large MagFace	2.19	93.74	9.01	8.30
iResNet18 MagFace	0.62	80.73	4.69	4.11
iResNet50 AVG MagFace	0.97	97.14	6.69	5.90
iResNet50 ArcFace	0.48	89.78	2.69	3.05
iResNet50 MagFace	0.18	59.85	0.76	1.12

8.2 Výsledky aplikovania metód odhadu kvality

Pre jednotlivé experimenty so získavaním kvality snímok dúhovky som využíval ako verifikačné siete iResNet50, iResNet18 a MobileNetV3-Large. Tieto siete sú výsledkami predošlých experimentov. Siete som vybral pre reprezentáciu rôznych systémov v ktorých môže byť získavanie kvality aplikované. Kde iResNet50 predstavuje reprezentanta presného verifikačného systému a sieť MobileNetV3-Large rýchly systém pre verifikáciu. Zo sieťou iResNet18 experimentujem pri získavaní kvality, kde sa snažím určiť efekt toho ako vplyvajú výsledky menej presnej siete na určenie kvality pre iné siete. Pri všetkých jednotlivých experimentoch následne porovnávam aký efekt má odmietnutie podielu snímok na verifikačnú chybu FNMR. Hodnotu chyby získavam pri FMR=0.1%. V experimentoch následne porovnávam relatívne hodnoty tejto chyby vzhľadom na hodnoty, ktoré dávajú jednotlivé siete bez aplikovania kvality. Relatívnu mierku som zvolil pre lepšiu vizualizáciu výsledkov. Pre využitie tohto porovnania je najskôr potrebné definovať k akej absolútnej hodnote je táto hodnota relatívna. Tabuľka 8.7 zobrazuje prehľad hodnôt FNMR pri FMR=0.1% pre verifikačné siete na všetkých verifikačných dátových sadách. Je však zjavné, že na verifikačnej dátovej sade od Innovatrics je táto hodnota pre všetky siete pomerne vysoká približne od 59% do 93%. Táto dátová sada obsahuje veľa nekvalitných vstupov, je teda zaujímavé zistiť či aj pre tieto vysoké hodnoty je však možné dosiahnuť relatívne zlepšenie presnosti verifikácie. Z výsledkov porovnávam hodnoty relatívnych hodnôt chyby FNMR pre miery odmietnutia párov od 0% po 20%. Hranica 20% je zvolená ako maximálna miera odmietnutia párov snímok, čo predstavuje pri predpoklade výberu náhodných párov každý piaty pár. V referenčnej literatúre považujú maximálnu použiteľnú mieru 15% [34], 20% je mierne nad touto mierou, čo však umožní lepšiu vizualizáciu vplyvu rôznych mier odmietnutia párov na hodnoty verifikačnej chyby FNMR.

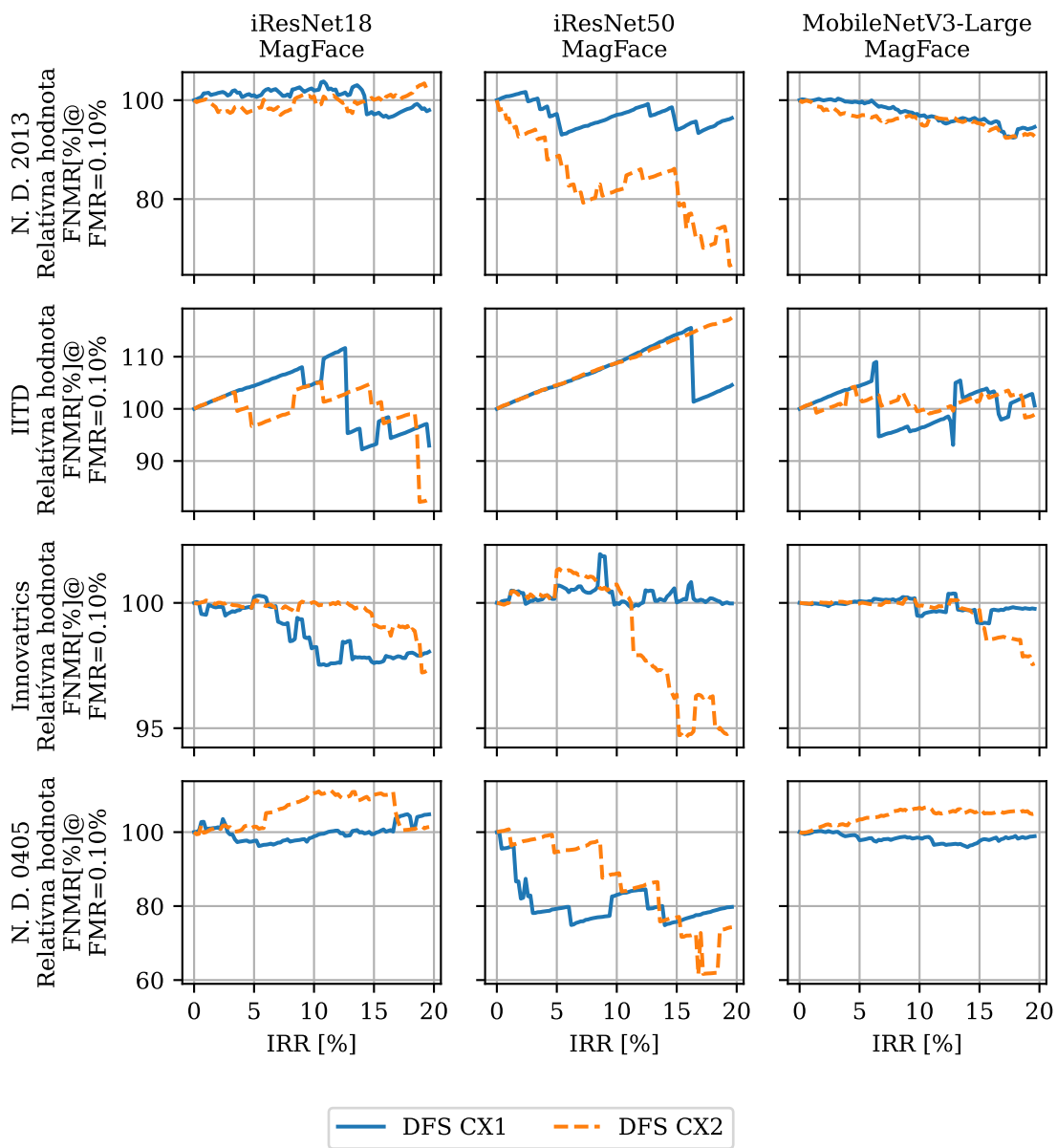
Pri experimentoch sledujem hlavne to, ako rôzne konfigurácie jednotlivých metód určovania kvality vplyvajú na verifikačný systém. Z tohto pohľadu pracujem pri jednotlivých metódach s rôznou kombináciou využívaných pred-trénovaných sietí. Z tohto pohľadu nedbám na to, či výsledný systém má vyššie výpočetné nároky ako verifikačná sieť. Zaujímá ma jedine to, akým spôsobom viem získať čo najlepšie zlepšenie presnosti pri verifikácii. Pri následnom porovnaní môžem vybrať najlepšie metódy odhadu kvality pre konkrétnu mieru odmietnutia párov.

Výsledky referenčnej metódy určovania kvality Pre referenčnú metódu DFS používam pred-trénované siete poskytnuté autormi. Autori poskytujú dve siete trénované na dátových sadách CASIA Iris CX1 a CASIA Iris CX2. Pre obe siete som získal referenčné skóre kvality pre všetky verifikačné dátové sady. Pri získavaní skóre kvality som hodnoty snímok normalizoval so strednou hodnotou 0.480 a smerodajnou odchýlkou 0.2. Snímky som následne transformoval do rozmeru 640×480 . Využívam zároveň celé snímky dúhoviek bez orezania a bez rozbalenia.

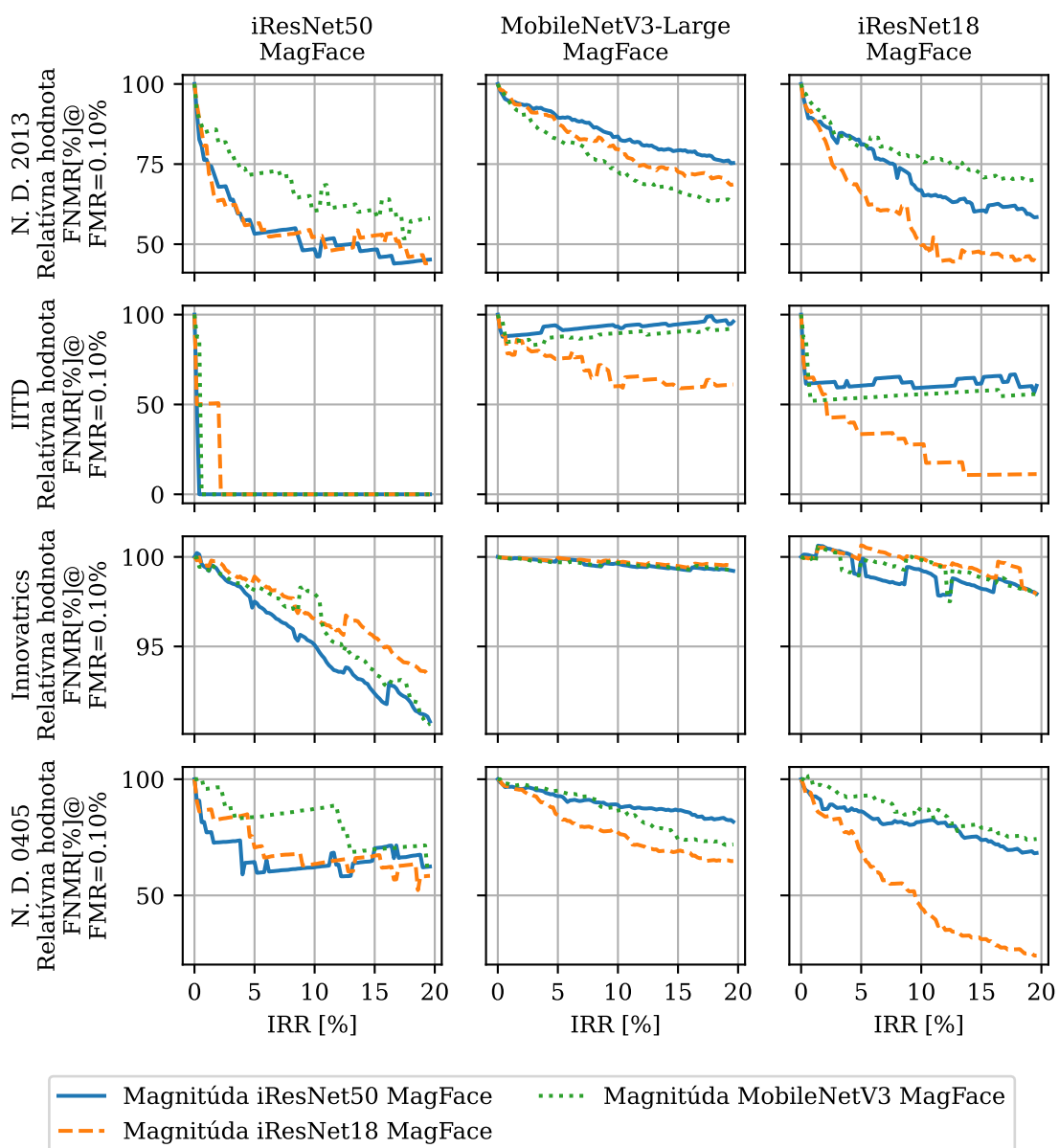
Obrázok 8.6 zobrazuje výslednú presnosť verifikačného systému pre dve referenčné siete určovania kvality. Metódu určenia kvality som aplikoval na všetky vybrané verifikačné siete. Z výsledkov aplikácie tejto metódy môžem konštatovať, že najlepšie zlepšenie dosahuje jedine na sieti iResNet50. Pri zvyšných sieťach táto metóda zlepšuje výsledky určením kvality no nie až tak výrazne. Pre sieť MobileNetV3-Large sieť určovania kvality trénovaná na dátovej sade CASIA Iris CX2 zhoršuje výslednú presnosť a to pri verifikačnej sade IITD a N. D. 0405.

Získanie skóre kvality z magnitúdy Pre experimenty zo získaním skóre z magnitúdy vektorov vlastností využívam siete trénované pomocou MagFace. Následne porovnávam efekt rôznych architektúr sietí pri krížovom porovnaní aplikácie získanej kvality. V tomto experimente porovnávam aký efekt má magnitúda získaná zo sietí iResNet50, iResNet18 a MobileNetV3-Large na verifikačné siete využívajúce rovnaké a rozličné architektúry.

Obrázok 8.7 ukazuje efekt po aplikovaní skóre kvality získaného z magnitúdy vektoru príznakov na výslednú presnosť jednotlivých verifikačných systémov. Je možné vidieť, že pre verifikačnú sieť iResNet50 je najlepšie aplikovanie magnitúdy, ktorá je výsledkom priamo tejto siete. Čo je zaujímavé tak je, že pri aplikovaní kvality získanej pomocou siete iResNet18 vylepší presnosť aj pri verifikačnej sieti MobileNetV3-Large, čo pozorujem na dátových sadách IITD, N. D. 0405 a Innovatrics. Kde na takmer celom vybranom intervale pomeru odmietnutých párov je hodnota chyby FNMR nižšia ako pre všetky ostatné získané skóre kvality.



Obr. 8.6: Relatívna hodnota verifikačnej chyby v závislosti na pomere odstránených párov pre metódu DFS. Obrázok ukazuje porovnanie kriviek na všetkých verifikačných dátových sádach pre verifikačné siete iResNet50, iResNet18 a MobileNetV3-Large.

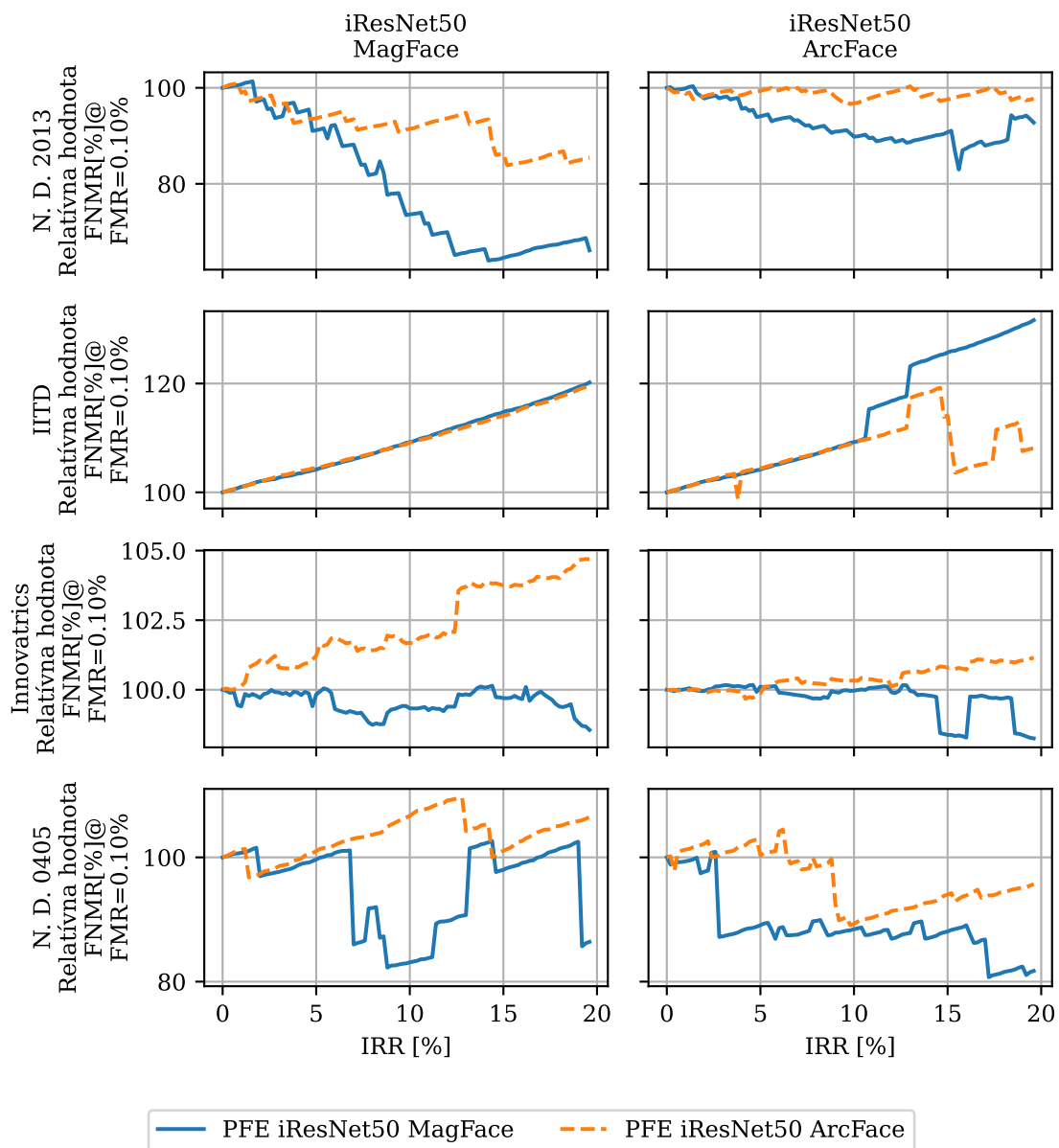


Obr. 8.7: Relatívna hodnota verifikačnej chyby v závislosti na pomere odstránených párov pre metódu určovania kvality pomocou magnitúdy vektoru príznakov. Obrázok ukazuje porovnanie kriviek na všetkých verifikačných dátových sadách pre verifikačné siete iResNet50, iResNet18 a MobileNetV3-Large. Verifikačné siete iResNet18 a iResNet50 používajú plne prepojenú poslednú vrstvu. Magnitúda je získaná z rovnakých sietí, následne je ukázaný efekt aplikácie týchto metrík kvality na všetky spomenuté verifikačné siete.

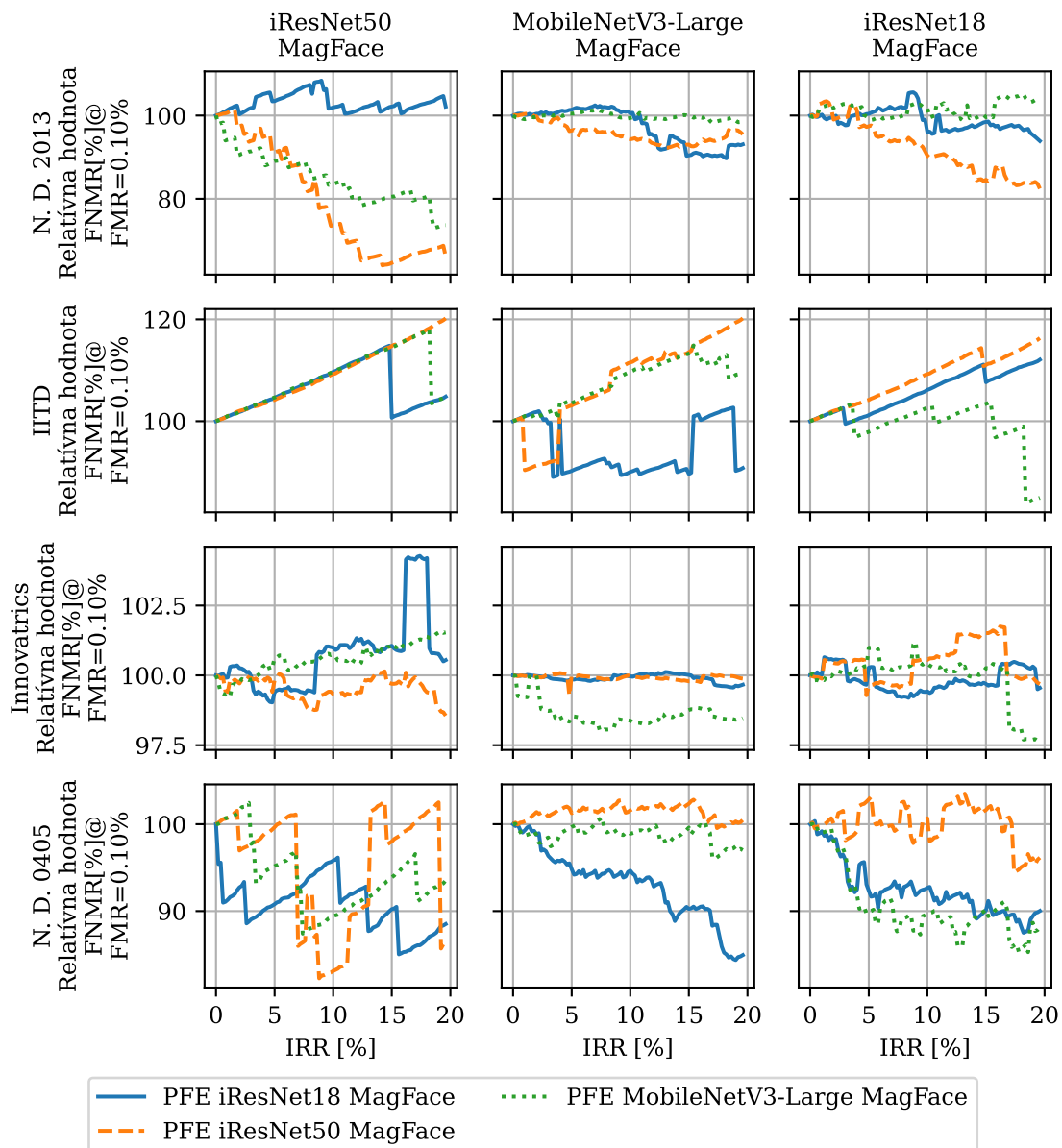
Trénovanie pravdepodobnostnej reprezentácie vektoru príznakov Pri experimentovaní s trénovaním siete pre pravdepodobnostnú reprezentáciu vektoru príznakov (PFE) [30] som sa rozhodol pracovať s myšlienkou, že informácia obsiahnutá vo vektore príznakov môže ovplyvniť získané skóre kvality. Pri tomto vychádzam z toho, že metóda netrénuje konvolučnú sieť ale jedine modul pre odhad rozptylov pravdepodobnostného rozloženia. Tento modul má jediný vstup a to práve vektor príznakov. Z pohľadu experimentov je zaujímavé použiť siete, ktorými výstupom sú vektory príznakov s dodatočnými informáciami. Jedna z možností je použitie siete pred-trénovanej pomocou MagFace a porovnávať ju so sieťou pred-trénovanou pomocou ArcFace. Dodatočne trénujem modul pripojený ku sieti pomocou optimalizátora SGD s learning rate 0.001, weight decay 0.0001 a momentum 0.9. Pri tréningu nevyužívam žiadne augmentácie ani drop-out. Týmto nastavením sledujem postup autorov. Trénuje sa maximálne na 6 epochách s batch o veľkosti 256 a s orezanými snímkami dúhovky.

Efekt použitia sietí pred-trénovaných na MagFace a ArcFace testujem s využitím verifikačnej siete iResNet50. Táto sieť je využívaná zároveň pre vytváranie vektoru príznakov ktorý sa využíva ako vstup pri trénovaní. Využívam teda dve siete, ktoré sú trénované pomocou ArcFace a MagFace. Výsledky tohto experimentu sú na obrázku 8.8. Je možné vidieť, že využitie pred-trénovanej siete pomocou MagFace dáva na všetkých dátových sadách okrem IITD po aplikácii na verifikačnú sieť lepšie výsledky ako sieť pred-trénovaná pomocou ArcFace. Pri dátovej sade IITD môže byť chyba spôsobená malým počtom chybných snímkov, ktoré metóda nevie hneď odhaliť.

V ďalších experimentoch s touto metódou som sa teda rozhodol využívať siete pred-trénované pomocou MagFace. Pri týchto experimentoch využívam výstup verifikačných sietí iResNet50, iResNet18 a MobileNetV3-Large. Výsledky tohto experimentu je možné vidieť na obrázku 8.9. Výsledky však nie úplne jasne ukazujú, aké zlepšenie môžem dosiahnuť na rôznych dátových sadách, keďže sú málo konzistentné. Je stále možné zistiť pre jednotlivé siete to, že v niektorých prípadoch dochádza k zlepšeniu relatívnej hodnoty verifikačnej chyby až pri 10% až 15% odmietnutých párov. Maximálne zlepšenie je však na dátovej sade N. D. 2013 a to pri sieti iResNet50 v kombinácii s modulom získania rozptylov založenom na výstupe rovnakej siete. Z výsledkov je tiež zaujímavé, že pri sieti MobileNetV3-Large sa dosahuje výraznejšie zlepšenia ak sa využívam modul pre určovanie kvality založený na iResNet18. Toto je pozorovateľné na dátových sadách IITD a N. D. 0405. Pri dátovej sade N. D. 2013 táto konfigurácia dosahuje lepšie výsledky až po odmietnutí viac ako 12% párov. Na dátovej sade Innovetrics sa dosahujú najlepšie výsledky s modelmi určujúcimi pravdepodobnostné rozloženie založenými na rovnakej sieti ako je verifikačná sieť.



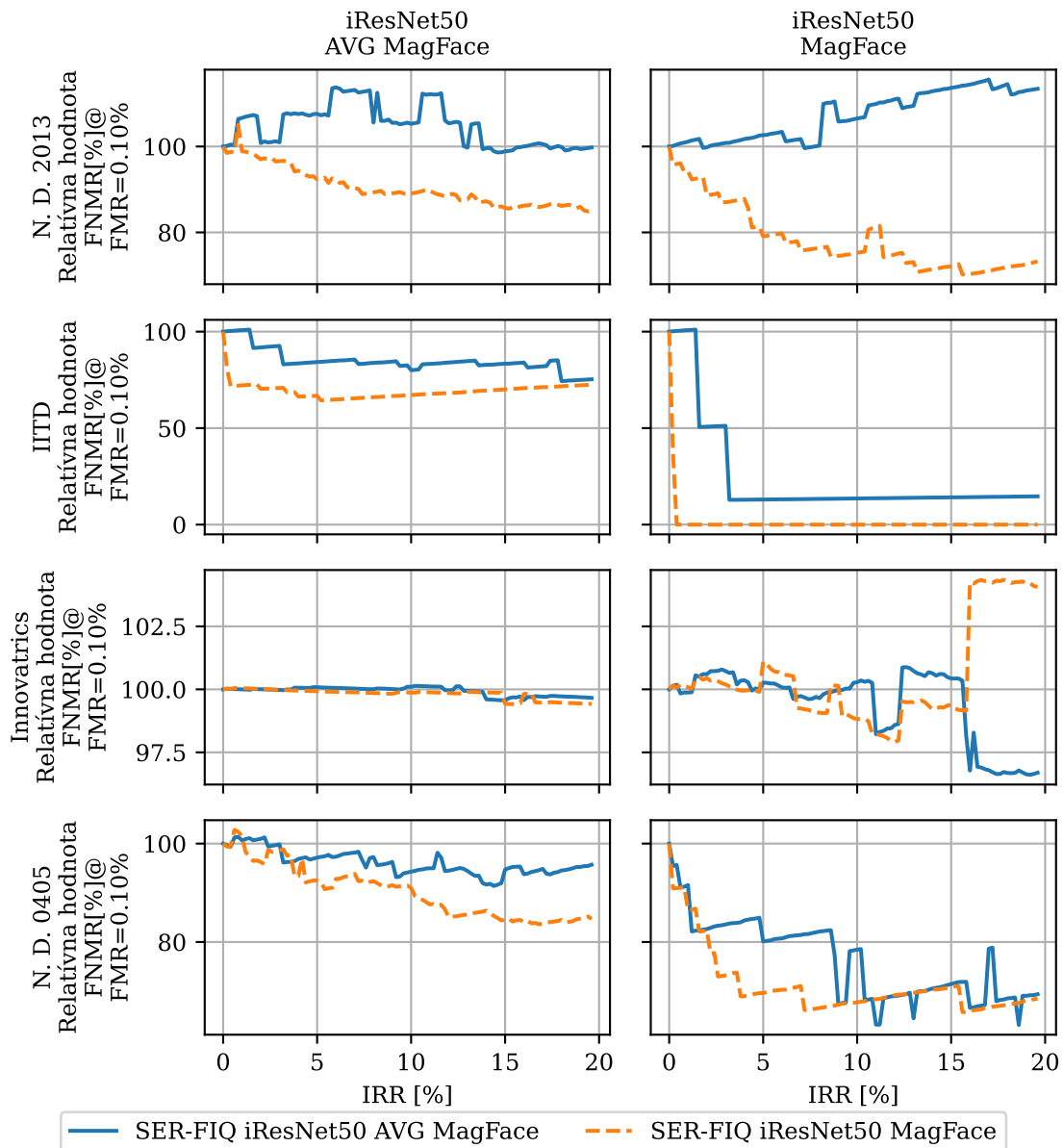
Obr. 8.8: Relatívna hodnota verifikačnej chyby v závislosti na pomere odstránených párov pre metódu určovania kvality pomocou pravdepodobnostnej reprezentácie príznakov. Porovnáva sa efekt použitia pred-trénovanej siete s ArcFace a MagFace, a aplikácie tejto metódy na verifikačné siete trénované pomocou ArcFace a MagFace.



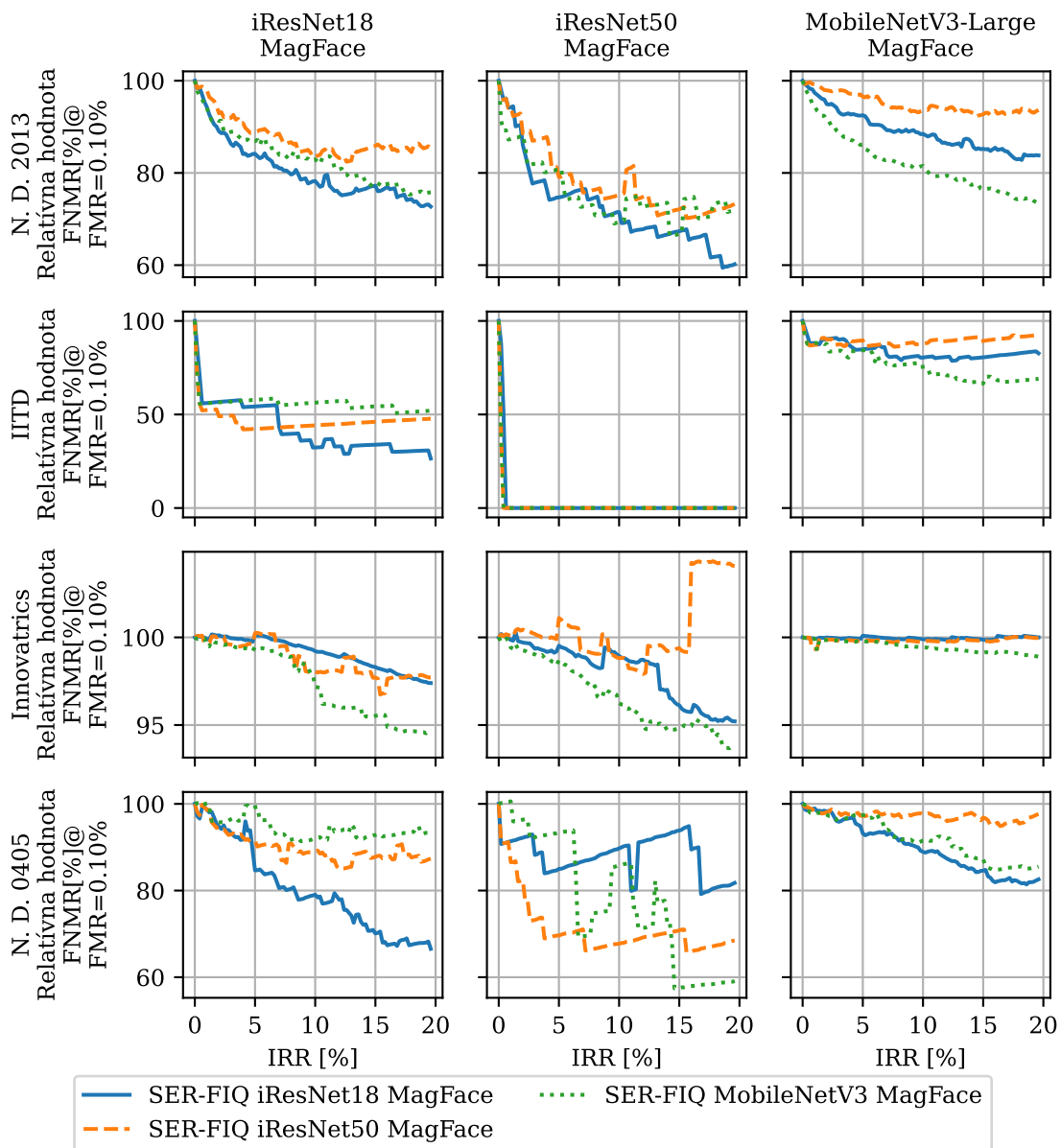
Obr. 8.9: Relatívna hodnota verifikačnej chyby v závislosti na pomere odstránených párov pre metódu určovania kvality pomocou pravdepodobnostnej reprezentácie príznakov. Ukazuje sa efekt využitia pred-trénovaných sietí iResNet50, iResNet18 a MobileNetV3-Large trénovaných pomocou MagFace. Následne je zobrazené aplikovanie tejto metódy určenia kvality na verifikačné siete iResNet18, iResNet50 a MobileNetV3-Large trénované pomocou MagFace.

Výsledky aplikovania metódy SER-FIQ [35] Pre experimentovanie s metódou SER-FIQ som sa v prvom rade rozhodol zistiť, či je touto metódou získanie skóre kvality aj z verifikačných sietí, ktoré využívajú priemernú hodnotu posledných konvolučných vrstiev v porovnaní s plne prepojenou vrstvou. Pri plne prepojenej vrstve sa aplikuje drop-out na celý výstup konvolučnej siete a tým pre túto metódu bude vo výsledku vzor tvorený kombináciou hodnôt z rôznych častí siete ako v ploche obrazu tak aj v hĺbke kanálov. Pri priemere hodnôt poslednej konvolučnej vrstvy je zase drop-out aplikovaný nad vektorom priemerných hodnôt a teda vzor je tvorený jedine kombináciou rôznych kanálov. Výsledky tohto experimentu sú na obrázku 8.10. V tomto prípade som experimentoval jedine s aplikovaním tejto metódy na siete iResNet50, ktoré boli výsledkami predošlých experimentov s verifikačnými sietami. Z výsledkov je možné vidieť, že pri použití siete s plne prepojenou poslednou vrstvou verifikačná chyba rýchlejšie naprieč všetkými ostatnými dátovými sadami. Výnimkou je dátová sada od Innovatives avšak aj tu je relatívna verifikačná chyba pre využitie plne prepojenej vrstvy zväčša pod metódou využívajúcu priemer do 15% odmietnutých párov. Pri aplikácii metódy na sieť využívajúcu priemernú hodnotu poslednej konvolučnej vrstvy sa pri dátovej sade N. D. 2013 je dokonca možné pozorovať zhoršenie presnosti verifikácie. Konfigurácia s plným prepojením dáva v tomto prípade lepšie výsledky a presnosť dokonca zvyšuje.

Pri ďalších experimentoch následne využívam pre siete iResNet18 a iResNet50 konfigurácie s plne prepojenými poslednými vrstvami. Experimentujem aj s krížovou aplikáciou metódy určenia kvality pre kde beriem do úvahy aj sieť MobileNetV3-Large. Odhad kvality pomocou tejto metódy aplikujem na verifikačné siete iResNet18, iResNet50 a MobileNetV3-Large. Výsledky pre tento experiment je možné vidieť na obrázku 8.11. Zaujímavé zistenie je, že aj pre sieť MobileNet dáva aplikácia metódy SER-FIQ zlepšenie verifikačnej chyby naprieč všetkými dátovými sadami a naprieč všetkými verifikačnými sietami. Konfigurácia vrstiev siete na ktoré sa aplikuje drop-out pri MobileNetV3 pracuje nad vektorom priemerných hodnôt z poslednej konvolučnej vrstvy. Aplikácia metódy SER-FIQ je teda podobná ako pri sieťach iResNet avšak výsledky sú iné. Toto môže byť spôsobené aj odlišným nastavením drop-out voči tréningu, kde pri tréningu využívam pravdepodobnosť 0.2 a pri používaní metódy využívam pravdepodobnosť 0.5, prípadne to môže byť spôsobené inou konfiguráciou posledných vrstiev siete, ktorá obsahuje plne prepojenú vrstvu pred a po aplikácii drop-out. Aplikácia metódy na menšie siete zároveň zlepšuje aj verifikačnú chybu pri väčších sieťach, výrazný efekt je možné vidieť na dátovej sade N. D. 2013, kde skóre získané z iResNet18 dosahuje lepšie výsledky pre sieť iResNet50. N iných dátovej sade Innovatives zase pre túto sieť dáva lepšie výsledky skóre získané z MobileNetV3. Avšak na iných dátových sadách tento fakt neviem pozorovať a teda neviem s istotou určiť, či neistota získaná z horšie natrénovaných sietí vie lepšie reprezentovať menej kvalitné snímky.



Obr. 8.10: Relatívna hodnota verifikačnej chyby v závislosti na pomere odstránených párov pre metódu určovania kvality SER-FIQ. Porovnáva sa získavania kvality z pred-trénovanej siete pomocou a MagFace, splne prepojenou poslednou vrstvou a s využitím priemernej hodnoty poslednej konvolučnej vrstvy a následnou aplikáciou plne prepojenej vrstvy. Porovnanie je na verifikačných sieťach s rovnakými konfiguráciami.



Obr. 8.11: Relatívna hodnota verifikačnej chyby v závislosti na pomere odstránených párov pre metódu určovania kvality SER-FIQ. Metóda je aplikovaná na siete iResNet50, iResNet18 a MobileNetV3-Large trénovaných pomocou MagFace, pri sieťach iResNet je používaná plne prepojená posledná vrstva. Pre porovnanie sú použité verifikačné siete iResNet50, iResNet18 a MobileNetV3-Large s rovnakou konfiguráciou poslednej vrstvy.

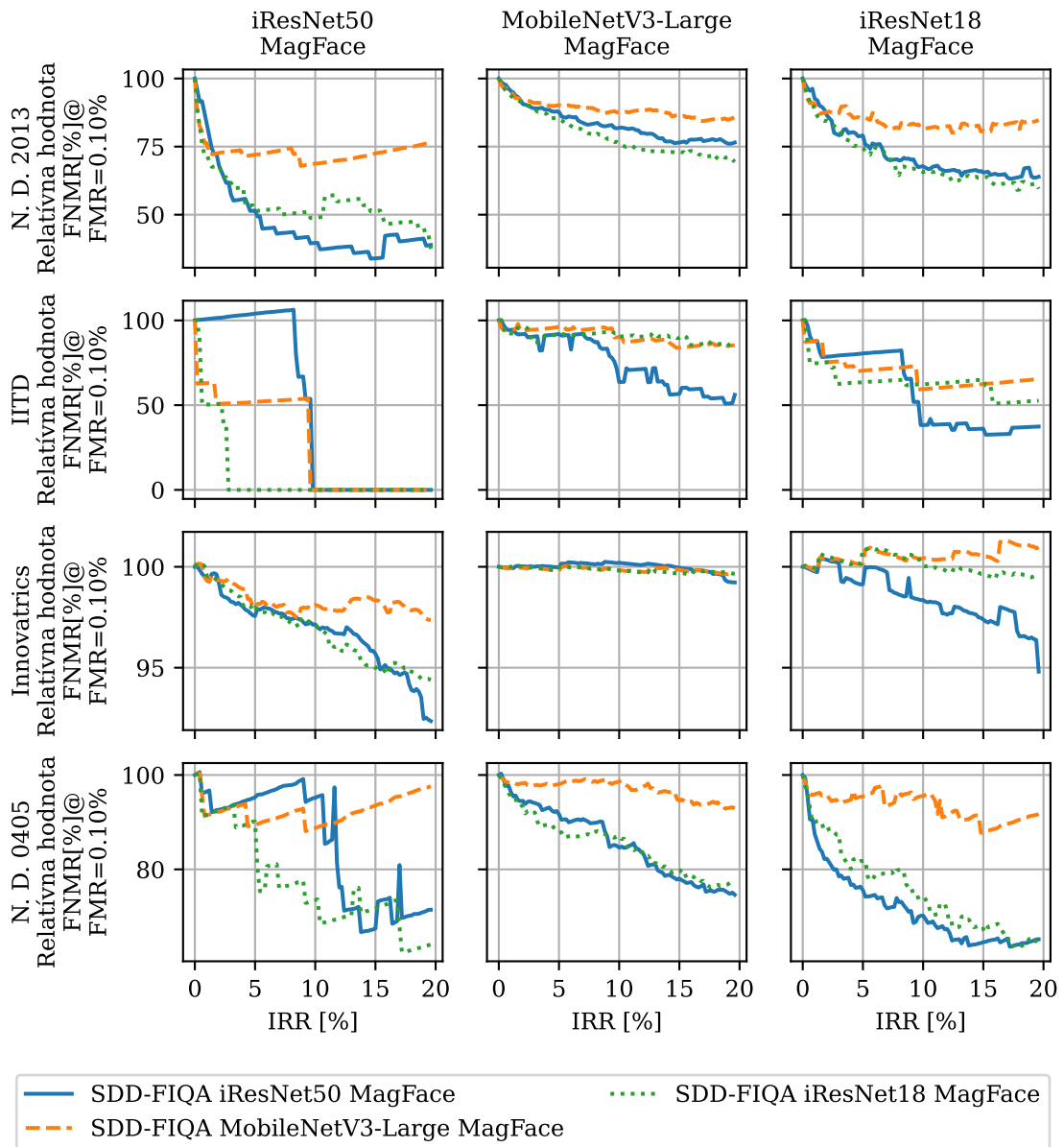
Trénovanie siete pre odhad kvality s generovanými označeniami snímok Pre experimenty s metódou tréningu siete pre predikciu vzdialenosti distribučných funkcií podobnosti SDD-FIQA [26] som zvolil pre generovanie nových označení dát trénovanej sady verifikačné siete iResNet50, iResNet18 a MobileNetV3-Large trénované s MagFace. Siete iResNet používajú plne prepojenú poslednú vrstvu a teda siete es najlepšimi výsledkami pri experimentoch s verifikáciou. Metódou trénujem s novým označením dát neurónovú sieť iResNet50. Pre tréning siete nevyužívam žiadnu augmentáciu snímok. Používam optimalizátor Adam [18] s konfiguráciou learning rate 0.001 a weight decay 0.0005. Sieť trénujem maximálne 10 epoch s drop-out 0.5, zároveň je po 5 epochách aplikované násobenie learning rate hodnotou 0.1. Využívaná sieť iResNet50 je prvotne pred-trénovaná pre verifikáciu pomocou MagFace.

Z výsledkov na obrázku 8.12 vyplýva, že metóda zlepšuje presnosť pre každú verifikačnú sieť. Zároveň sa zlepšuje presnosť aj pre rôzne kombinácie generovaných označení a verifikačných sietí. Najvýraznejšie zlepšenie verifikačnej chyby zároveň môžem pozorovať hlavne pri použití označení trénovacej sady generovanými pomocou siete iResNet18 a siete iResNet50. Označenia generované pomocou MobileNetV3-Large nedosahujú ani pri aplikácii na rovnakú verifikačnú sieť lepšie výsledky ako pri použití označení generovanými sieťami iResNet, keďže ostatné výsledky sú menej jasné, tak výber ideálnych označení pre trénovanie siete pre predikciu kvality použitú v jednotlivých verifikačných systémoch bude potrebné vybrať pre konkrétny prah.

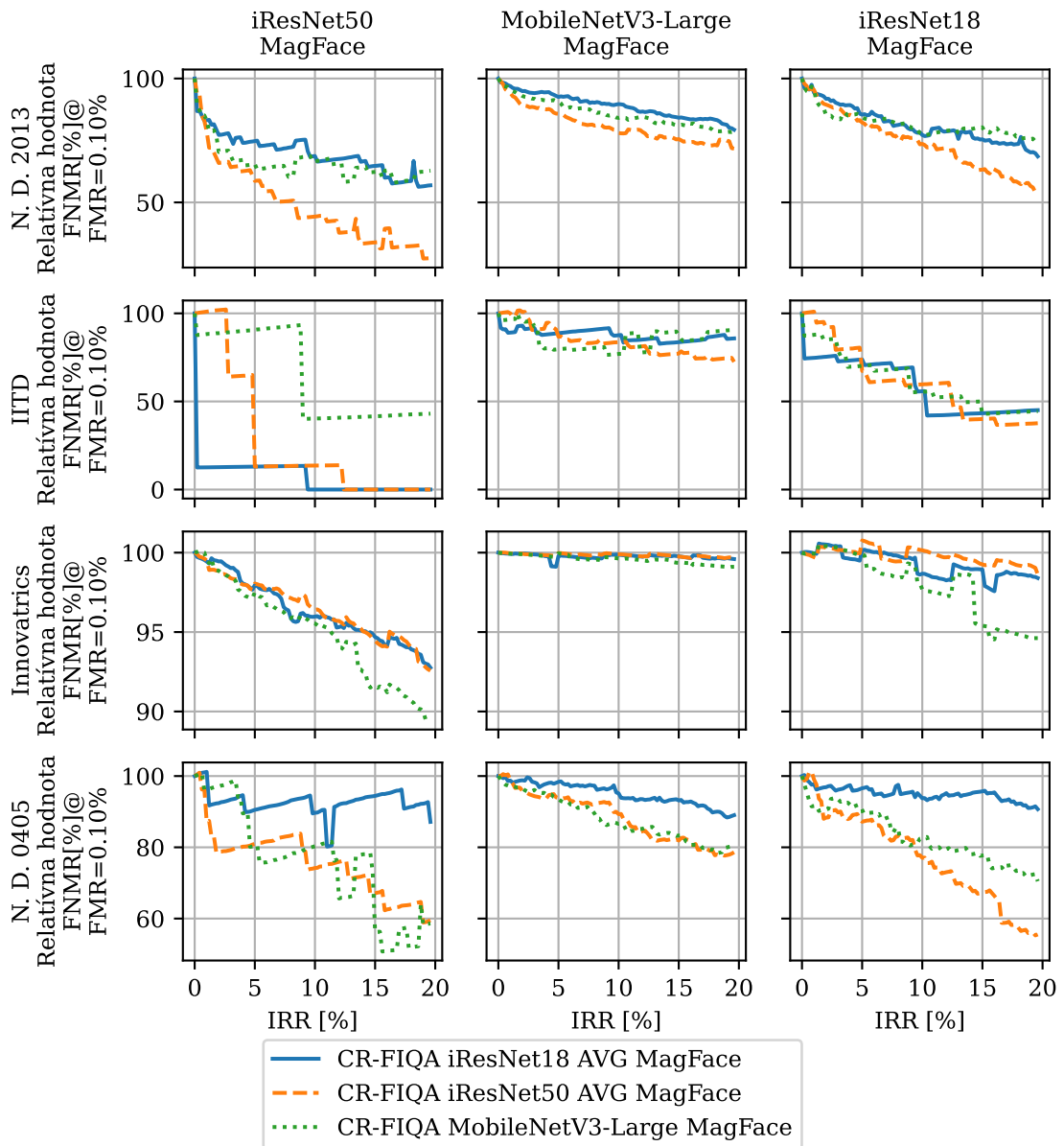
Trénovanie relatívnej klasifikovateľnosti [4] Posledná metóda určenia kvality s ktorou som experimentoval je metóda odhadu pomocou relatívnej klasifikovateľnosti CR-FIQA. Pomocou tejto metódy do-trénujem verifikačné neurónové siete s novou stratovou funkciou. Využívam neurónové siete pred-trénované pomocou MagFace. Pre optimalizáciu váh modelu využívam optimalizátor SGD s použitím learning rate 0.1, momentum 0.9 a weight decay 0.0001. Zároveň využívam rovnakú augmentáciu trénovacích dát ako pri tréningu sietí pre verifikáciu. Pomocou tejto metódy trénujem sieť iResNet s plne prepojenou poslednou vrstvou a s priemerovaním hodnôt poslednej konvolučnej vrstvy. Využívam siete iResNet18 a iResNet50. Zároveň pomocou tejto metódy trénujem aj sieť MobileNetV3-Large. Keďže siete sú pred-trénované rozhodol som sa využiť plánovač s mílnikmi 10, 20 a 25 pri a hodnotou gamma 0.1. Siete sa trénujú po maximálnu dobu 30 epoch.

Pri prvotnom experimente som vyskúšal sieť iResNet50 s plne prepojenou poslednou vrstvou, avšak po tréňovaní sa neznižovala hodnota stratovej funkcie. Tréning mal v priebehu konštantné hodnoty a po ôsmich epochách sa následne začala hodnota zvyšovať do rozmedzí 10^{12} . Tento výsledok sa opakoval aj pri tréningu iResNet18 s plne prepojenou poslednou vrstvou. Z tohto dôvodu som sa rozhodol do-tréňovať touto metódou iResNet18 a iResNet50 bez plne prepojenej poslednej vrstvy. Pre tieto siete a aj sieť MobileNetV3-Large bol tréning úspešný.

Keďže zväčša pri porovnávaní využívam verifikačné siete využívajúce plne prepojenú poslednú vrstvu tak aj v tomto prípade zvolím túto možnosť. Obrázok 8.13 zobrazuje porovnanie pri aplikácii získaného skóre kvality na jednotlivé verifikačné siete. Z týchto výsledkov vyplýva, že pri získaní skóre kvality touto metódou dochádza ku zlepšeniu verifikačnej chyby naprieč všetkými dátovými sadami. Najviac sa vylepšuje verifikačná chyba siete iResNet50 a to hlavne pri aplikovaní skóre získaných zo siete iResNet50 trénovanej pomocou CR-FIQA. Pre zvyšné verifikačné siete rovnako dáva sieť iResNet50 trénovaná pomocou CR-FIQA najlepšie zlepšenie verifikačnej chyby na väčšine dátových sádach. Jedine na dátovej sade Innovatrics nie je táto sieť najlepšie aplikovateľná.



Obr. 8.12: Relatívna hodnota verifikačnej chyby v závislosti na pomere odstránených párov pre metódu určovania kvality SDD-FIQA. Sieť využívaná pri určovaní kvality je trénovaná pomocou označení generovanými sieťami iResNet50, iResNet18 a MobileNetV3-Large trénovanými pomocou MagFace. Pre porovnanie sú použité verifikačné siete iResNet50, iResNet18 a MobileNetV3-Large trénované pomocou MagFace.



Obr. 8.13: Relatívna hodnota verifikačnej chyby v závislosti na pomere odstránených párov pre metódu určovania kvality CR-FIQA. Siete využívané pri používaní architektúry iResNet50, iResNet18 a MobileNetV3-Large. Pre porovnanie sú použité verifikačné siete iResNet50, iResNet18 a MobileNetV3-Large tréované pomocou MagFace.

Tabuľka 8.8: Výber konfigurácii jednotlivých metód odhadu kvality pre verifikačné siete iResNet50 a MobileNetv3-Large tréované pomocou MagFace. Konfigurácie metód pri miere odmietnutých párov 10% a FMR 0.1% majú najnižšie hodnoty FNMR pre čo najviac dátových sád. V ľavom stĺpci tabuľky sú skratky a názvy jednotlivých metód, v tabuľke je následne názov konfigurácie pre danú metódu.

	Verifikačné siete	
	iResNet50 MagFace	MobileNetV3-Large MagFace
DFS	CX2	CX1
Magnitúda	iResNet50 MagFace	iResNet18 MagFace
PFE	iResNet50 MagFace	iResNet18 MagFace
SER-FIQ	MobileNetV3-Large MagFace	MobileNetV3-Large MagFace
SDD-FIQA	iResNet18 MagFace	iResNet50 MagFace
CR-FIQA	iResNet50 AVG MagFace	iResNet50 AVG MagFace

8.3 Porovnanie metód odhadu kvality

Pre vzájomné porovnanie jednotlivých metód a následný výber toho, ktorá metóda odhadu kvality je najlepšia som prvotne potreboval zvoliť vhodné kombinácie verifikačných sietí a konfigurácie jednotlivých metód odhadu kvality. Rozhodol som sa pracovať s porovnaním na dvoch verifikačných systémoch a to na systéme s verifikačnou sieťou iResNet50 a MobileNetV3-Large. Sieť iResNet18 nevyužívam, keďže predstavuje pri počte využívajúcich parametrov len veľmi malé zlepšenie presnosti. Pre účel výberu konfigurácii jednotlivých metód odhadu kvality som si zvolil prah odmietnutých snímok na 10%. Tento prah predstavuje pri viacerých metódach hranicu, pre ktorú je možné pozorovať výraznú zmenu v efekte na presnosť verifikačného systému. Týmto je zvolený prah kompromis medzi množstvom odmietnutých vstupov, a získanou presnosťou. Pri tomto prahu následne pre každú metódu odhadu kvality snímky a každú verifikačnú sieť hľadám takú konfiguráciu, pre ktorú platí, že pre čo najviac dátových sadách dáva najnižšiu hodnotu FNMR. Výber je z výsledkov z predošlej sekcie kde pre jednotlivé priebehy hodnôt FNMR využívam fixnú hodnotu FMR 0.1%. Jednotlivé konfigurácie metód odhadu kvality následne aplikujem na verifikačné neurónové siete, pre ktoré boli vybrané. Výber jednotlivých aplikovaných konfigurácii metód odhadu kvality je zobrazený na tabuľke 8.8. V následujúcich vizualizáciách využívam jedine názvy jednotlivých metód spolu s informáciou o percente odmietnutých párov. Vyhodnotenie porovnania rozdelím podľa využívaných verifikačných neurónových sietí.

Výsledok DET kriviek pre sieť iResNet50 je zobrazený na tabuľke 8.9 a obrázku 8.17. DET krivky boli vytvorené pre prah odmietnutých párov 10%. Z jednotlivých výsledkov vyplýva, že pre túto sieť záleží najlepšia konfigurácia metód odhadu kvality od konkrétnej dátovej sady. Pri jednotlivých dátových sadách sú minimálne plochy pod DET krivkami rôzne. Z jednotlivých metód pri tomto výbere dosahujú najlepšie výsledky magnitúda vektoru, SER-FIQ, CR-FIQA a SDD-FIQA. Ak však sledujem skôr priebeh FNMR pri rôznom počte odmietnutých párov a hodnote FMR 0.1%, tak z obrázku 8.17 môžem sledovať najlepšie metódy pre rôzne prahy odmietnutých párov. V absolútnych hodnotách FNMR som si následne vybral tri hodnoty pre mieru odmietnutých párov a to 5%, 10% a 15%. Pre jednotlivé prahy viem dosiahnuť zlepšenie na dátovej sade University of Notre Dame Iris CrossSensor 2013 z 1.12% na 0.57%, 0.49% a 0.37%. Kde pri prahu 5% je najlepšia metóda SDD-FIQA tréovaná na hodnotách kvality generovanými sieťou iResNet18 a pri

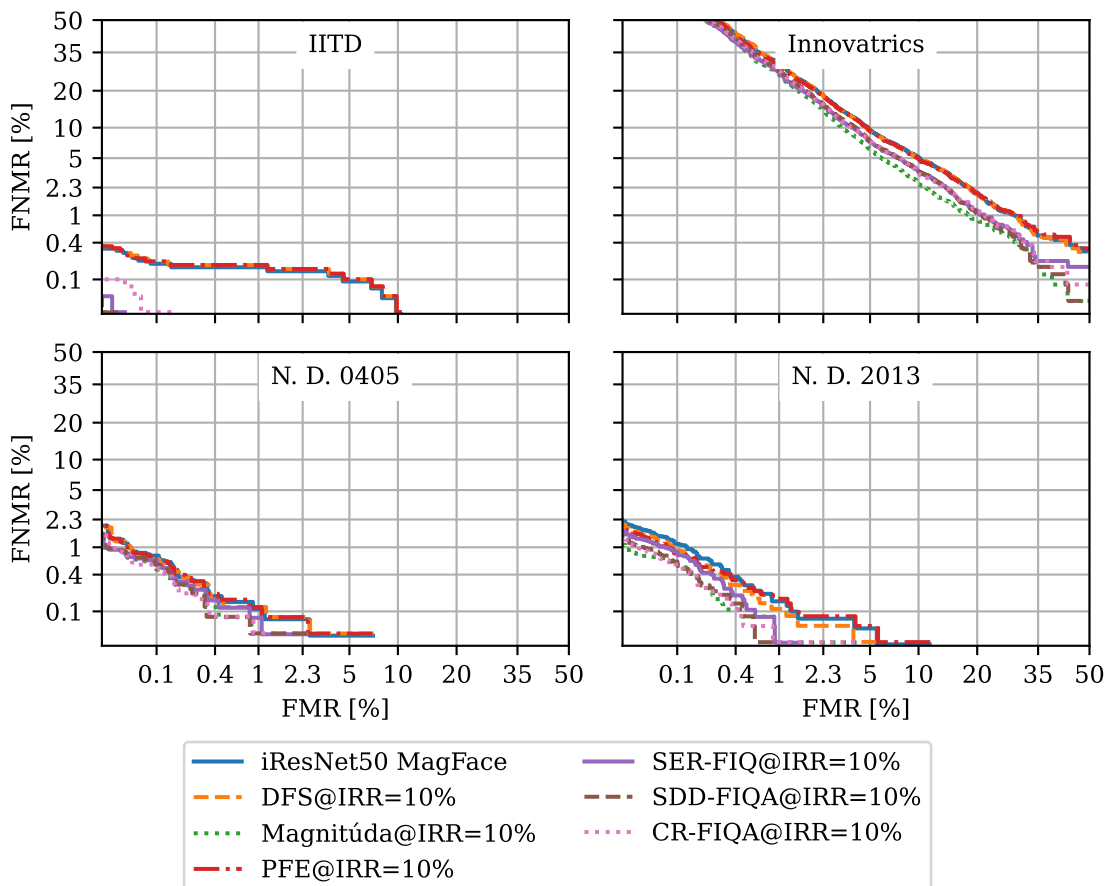
zvyšných prahoch CR-FIQA. Pri dátovej sade Innovatrics pozorujem najlepšie zlepšenie pri jednotlivých prahoch z 59.85% na 58.36%, 56.91% a 55.28%. Vo všetkých vybraných bodoch dosahuje tieto hodnoty magnitúda vektoru získaná z rovnakej siete. Pre dátovú sadu University of Notre Dame Iris 0405 dosahujem na jednotlivých bodoch zlepšenie z 0.76% na 0.49%, 0.47% a 0.43%. Pre jednotlivé prahy tieto hodnoty získala metóda určenia kvality z magnitúdy a pri prahu 10% metóda SER-FIQ spustená nad menšou sieťou MobileNetV3-Large. Pre dátovú sadu dosahujem zlepšenie na všetkých prahoch z 0.18% na 0% metódou získania skóre kvality z magnitúdy vektoru príznakov. Keďže metóda určenia skóre z magnitúdy je vo viacerých vybraných bodoch najlepšia ta môže byť považovaná za celkovo najlepšiu konfiguráciu avšak variabilita je pomerne veľká a teda je vhodnejší výber konkrétnej metódy vzhľadom na konkrétny prah a charakteristiku snímok ktoré sa očakávajú na vstupe.

Z obrázku 8.16 a tabuľky 8.16 je možné vidieť, že pre verifikačný systém založený na sieti MobileNetV3-Large dávajú rovnaké metódy podobné výsledky. A teda na jednotlivých dátových sadách sú plochy pod DET krivkou pre 10% odmietnutých je najlepšia pri dátovej University of Notre Dame Iris 0405 metóda odhadu z magnitúdy vektoru, pri Innovatrics je to metóda SER-FIQ, pri University of Notre Dame Iris CrossSensor 2013 je to metóda CR-FIQA a pri IITD to je metóda SDD-FIQA. Podobne ako v predošlom porovnaní tak aj tu som si zvolil tri prahy pre mieru odmietnutých párov a to 5%, 10% a 15%. Pre dátovú sadu N. D. 2013 dosahujem zlepšenie z 8.3% a 7.12%, 6.47% a 6.02%. Kde zlepšenie som pri prvých dvoch prahoch dosiahol metódou CR-FIQA, ktorou sa do-trénovala sieť iResNet50, a pre posledný prah to bola metóda získania skóre z magnitúdy nad sieťou iResNet18. Pre dátovú sadu IITD dosahujem najlepšie zlepšenie z 2.18% na 1.65%, 1.31% a 1.23%. V prvých dvoch prípadoch dosahujem zlepšenie pomocou magnitúdy vektoru a v poslednom prípade pomocou SDD-FIQA s tréningovými skóre získanými pomocou iResNet50. Pre dátovú sadu Innovatrics dosahujem konzistentne s metódou SER-FIQ využívajúcou sieť identickú s verifikačnou zlepšenie z 93.73% na 93.53%, 93.19% a 92.9%. Na dátovej sade N. D. 0405 dosahujem konzistentné zlepšenie s využitím metódy určovania skóre z magnitúdy a to z 8.3% na 7.28%, 6.61% a 6.02%. Pre túto sieť je podobne ako v predošlom porovnaní metóda určenia kvality z magnitúdy vektoru najlepšia vo väčšine vybraných bodoch. Avšak je vhodné doplniť, že pre tento prístup bola zvolená magnitúda získaná z inej siete a teda ak by bolo potrebné optimalizovať metódu na systém s menším výkonom tak by bolo potrebné otestovať iné metódy prípadne metódu získania skóre z magnitúdy aplikovať na výstup tej istej siete.

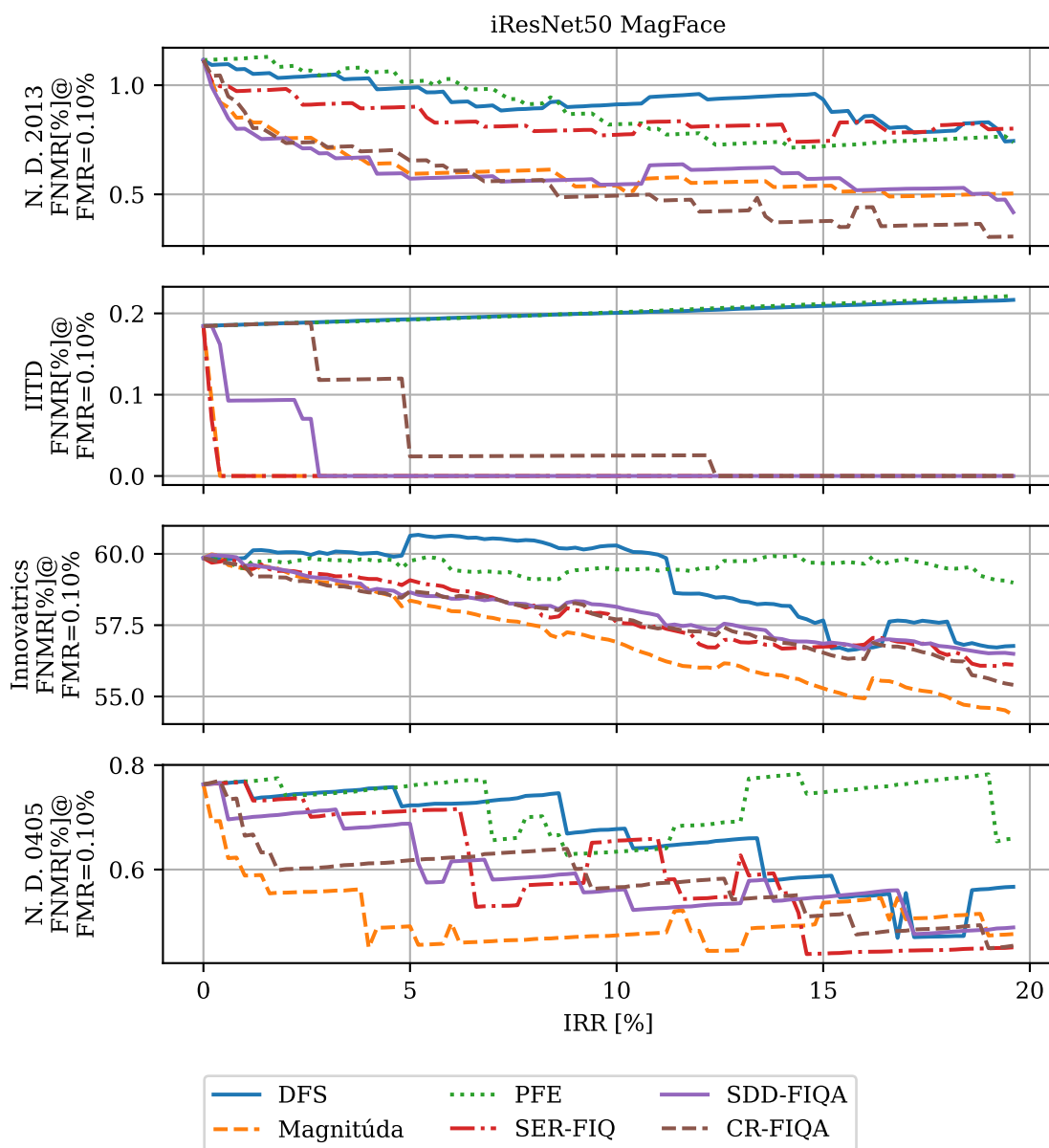
Taktiež môžeme konštatovať, že v porovnaní s referenčnou state-of-the-art metódou DFS dávajú takmer všetky metódy lepšie výsledky z pohľadu plochy pod grafom DET kriviek pri odmietnutí 10% párov. Táto metóda zároveň pre plochy pod grafmi vylepšuje takmer žiadne výsledky. Z pohľadu hodnôt FNMR však v niektorých prípadoch prekonáva metódu PFE. Pri príkladoch ako je jej aplikácia v dátovej sade Innovatrics na verifikačnej sieti MobileNetV3-Large, tak pri správnych hodnotách miery odmietnutých snímokoch vie dosiahnuť lepšie výsledky ako metóda určenia kvality z magnitúdy.

Tabuľka 8.9: Plochy pod DET krivkami pre všetky verifikačné dátové v závislosti na použití rôznych metód odhadu kvality nad verifikačnou sieťou iResNet50. Hodnoty plôch sú pri metódach určovania kvality vyhodnocované pre pomer odmietnutých párov 10%. Zvýraznené sú najmenšie hodnoty plochy pod DET kvikov pre jednotlivé dátové sady.

$[\times 10^{-4}]$	IITD	Innovatrics	N. D. 0405	N. D. 2013
iResNet50 MagFace	1.06	205.37	0.63	0.98
CR-FIQA@IRR=10%	0.03	156.39	0.23	0.29
DFS@IRR=10%	1.14	203.96	0.64	0.63
Magnitúda@IRR=10%	0.03	133.38	0.28	0.30
PFE@IRR=10%	1.15	204.21	0.65	0.96
SDD-FIQA@IRR=10%	0.03	155.35	0.29	0.22
SER-FIQ@IRR=10%	0.01	157.25	0.32	0.33



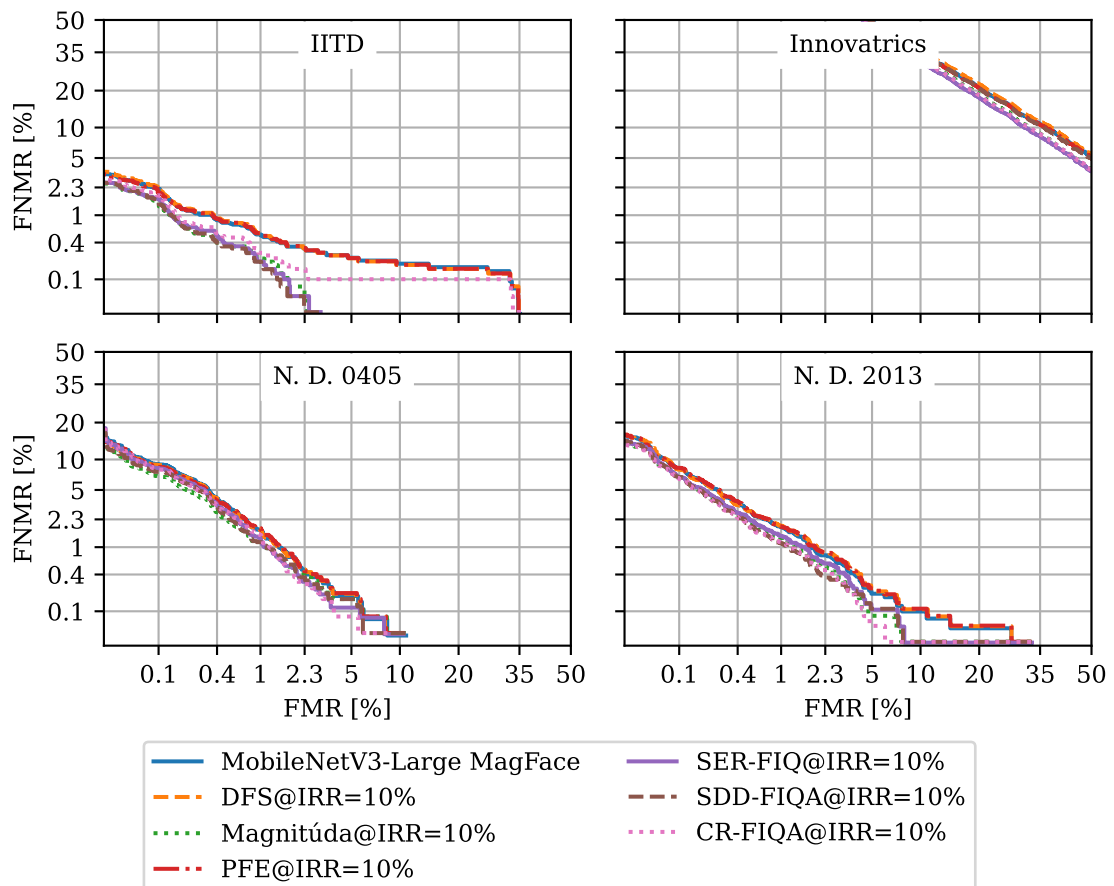
Obr. 8.14: Verifikačná chyba v závislosti na použití rôznych metód odhadu kvality nad verifikačnou sieťou iResNet50 pre mieru odmietnutia párov vstupných snímok 10%.



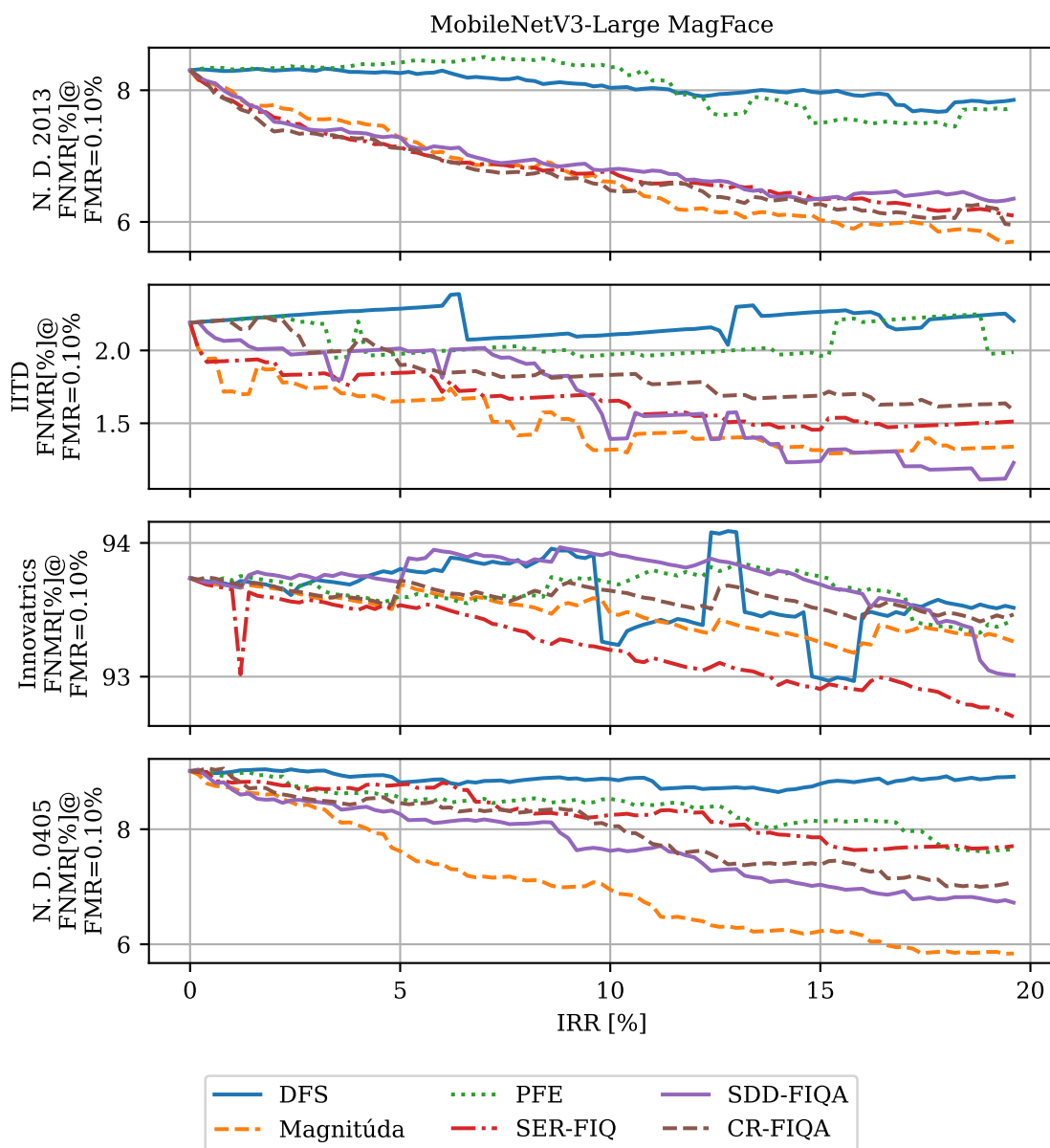
Obr. 8.15: Verifikačná chyba v závislosti na použití rôznych metód odhadu kvality nad verifikačnou sieťou iResNet50 pre rôzne miery odmietnutia párov s hodnotou FMR 0.1%.

Tabuľka 8.10: Plochy pod DET krivkami pre všetky verifikačné dátové v závislosti na použití rôznych metód odhadu kvality nad verifikačnou sieťou MobileNetV3-Large. Hodnoty plôch sú pri metódach určovania kvality vyhodnocované pre pomer odmietnutých párov 10%. Zvýraznené sú najmenšie hodnoty plochy pod DET kvikov pre jednotlivé dátové sady.

$[\times 10^{-4}]$	IITD	Innovatrics	N. D. 0405	N. D. 2013
MobileNetV3-Large MagFace	7.10	1231.49	7.18	9.07
CR-FIQA@IRR=10%	4.15	1093.24	5.67	5.74
DFS@IRR=10%	6.92	1268.85	7.03	9.59
Magnitúda@IRR=10%	0.81	1099.92	5.57	6.15
PFE@IRR=10%	6.92	1212.94	7.19	9.62
SDD-FIQA@IRR=10%	0.74	1210.93	5.77	5.94
SER-FIQ@IRR=10%	0.81	1063.58	5.97	6.79



Obr. 8.16: Verifikačná chyba v závislosti na použití rôznych metód odhadu kvality nad verifikačnou sieťou MobileNetV3-Large pre mieru odmietnutia párov vstupných snímkov 10%.



Obr. 8.17: Verifikačná chyba v závislosti na použití rôznych metód odhadu kvality nad verifikačnou sieťou MobileNetV3-Large pre rôzne miery odmietnutia párov s hodnotou FMR 0.1%.

Kapitola 9

Záver

Cieľom tejto práce bolo preskúmať oblasť merania kvality snímok dúhovky, nájsť aktuálne využívané metódy pre meranie kvality a následne ich aplikovať na verifikačný systém.

Experimentoval som s viacerými rôznymi novodobými metódami pre určovanie kvality snímok. Zameral som sa predovšetkým na spôsoby získavania skóre kvality z výstupov a rôznych vlastností získaných z rozpoznávacích sietí. V prístupe som najskôr zvolil natréňovanie rozpoznávacej neurónovej siete pre identifikáciu osôb pomocou dúhovky. Pre sieť som vyskúšal viacero možných kombinácií tréňovania, úprav dát a konfigurácii aby som získal čo najpresnejší verifikačný systém. Následne som s výslednými sieťami experimentoval využitím rôznych metód určovania kvality.

Pre jednotlivé metódy odhadu kvality som si zvolil rôzne prístupy, ktoré sú aktuálne využívané v oblasti rozpoznávania tvári a doposiaľ neboli na oblasť identifikácie pomocou dúhovky aplikované. Experimentoval som s rôznymi spôsobmi tréningu rozpoznávacích sietí, tak aby bolo možné reprezentovať ich výstupom aj kvalitu vstupu. Pri týchto metódach so využíval tréning pre optimalizáciu magnitúdy vektoru príznakov a metódu tréňovania mieri istoty siete pre daný vstup. Z iných metód som použil aj metódu generovania nových označení tréňovacích snímok, s ktorými som následne trénoval model pre odhad kvality. Pracoval som aj s možnosťou získavania skóre kvality vzorkovaním vektorov z rôznych častí rozpoznávacej siete a tréning siete pre odhad pravdepodobnostného rozloženia vektoru príznakov.

Jednotlivé metódy som porovnával na štyroch rôznych dátových sadách voči aktuálnemu state-of-the-art v oblasti odhadu kvality snímok dúhovky. Pri jednotlivých metódach som sa zameral na sledovanie zlepšenia presnosti verifikačného systému po odmietnutí pomeru párov snímok z dátovej sady. Za prínosy mojej práce považujem predovšetkým dosiahnutie lepších výsledkov pri odhade kvality snímok dúhovky v porovnaní s aktuálnym state-of-the-art. Zároveň považujem za prínos aj zistenie toho aký efekt majú jednotlivé metódy pri použití v identifikačných systémoch založených na biometrii dúhovky.

Z jednotlivých využívaných metód som dosiahol na všetkých dátových sadách zlepšenie presnosti pre dve rôzne verifikačné siete pri false match rate s hodnotou 0.1% a s pomerom odmietnutých párov 10%. Pre väčšiu verifikačnú sieť iResNet50 viem aplikáciou jednotlivých metód tejto konfigurácii dosiahnuť zlepšenie false non-match až o 100% na dátovej sade IITD čím sa odstránia všetky problematické páry snímok. Na dátových sadách od of Notre Dame dosahujem zlepšenia až 56% pri sade Iris CrossSensor 2013 a 38% sade pri Iris 0405. Na dátovej sade od Innovarics dosahujem zlepšenie až 5%. Pre verifikačnú sieť MobileNetV3-Large dosahujem na dátovej sade IITD zlepšenie false non-match rate až 40%. Pre dátové sad od University of Notre Dame je to až a takmer 22% pre sadu Iris CrossSensor 2013 a

až 20% pre sadu Iris 0405. Na dátovej sade od Innovatrics zlepšujem presnosť tejto siete o 0.5%.

Z mojich zistení je následne možné určiť aké sú najlepšie možné metódy odhadu kvality kvality snímok dúhovky. Z výsledkov jednotlivých metód vyplýva, že naprieč rôznymi systémami dosahujú výrazne zlepšenie metóda pre odhad kvalitu z magnitúdy vektoru príznakov, stochastické vzorkovanie výstupov siete metódou SER-FIQ, trénovanie s mierov istoty CR-FIQA a generovanie označení kvality dátových sád z výstupu verifikačných sietí metódou SDD-FIQA.

V oblasti budúceho vývoja vidím hlavne možnosti využitia tréningu sietí pre odhad kvality pomocou generovaných označení dátových sád. Hlavný priestor vo vývoji vidím pri použití metód určovania kvality na základe sietí trénovanými pomocou MagFace pre generovanie týchto označení. Hlavnou motiváciou tohto je, že pre jednotlivé siete dávala táto metrika v niektorých prípadoch výrazne lepšie výsledky ak bola pre generovanie skóre kvality využitá väčšia a presnejšia sieť. Z tohto dôvodu si myslím že pre optimalizáciu výpočtového výkonu by bolo možné týmito označeniami trénovať napríklad sieť s architektúrou MobileNetV3. Podobný prístup je možný aj pri metóde SER-FIQ. Keďže sa na jednotlivých dátových sádach dosahovali rôzne výsledky pre rôzne metódy tak vidím priestor aj vo vývoji smerom ku fúzii viacerých metrík kvality. Takto spojená metóda odhadu kvality by mohla získať lepšie výsledky naprieč všetkými dátovými sadami. Hlavnými kandidátmi pre fúziu sú podľa mňa metóda určenia kvality z magnitúdy, SER-FIQ, SDD-FIQA a CR-FIQA.

Literatúra

- [1] ABHYANKAR, A. a SCHUCKERS, S. Iris quality assessment and bi-orthogonal wavelet based encoding for recognition. *Pattern Recognition*. 2009, zv. 42.
- [2] ALI, S. Person Identification Technique Using Human Iris Recognition. *Journal of Communications Technology and Electronics*. December 2013, zv. 3, s. 1–5.
- [3] AUTOMATION (CASIA), C. A. of Sciences' Institute of. *CASIA Iris Image Database*. Dostupné z: <http://biometrics.idealtest.org/>.
- [4] BOUTROS, F., FANG, M., KLEMT, M., FU, B. a DAMER, N. CR-FIQA: Face Image Quality Assessment by Learning Sample Relative Classifiability. December 2021.
- [5] BOWYER, K. W. a FLYNN, P. J. The ND-IRIS-0405 Iris Image Dataset. *CoRR*. 2016.
- [6] BOYD, A., CZAJKA, A. a BOWYER, K. Deep Learning-Based Feature Extraction in Iris Recognition: Use Existing Models, Fine-tune or Train from Scratch? *2019 IEEE 10th International Conference on Biometrics Theory, Applications and Systems, BTAS 2019*. 2019.
- [7] DAUGMAN, J. High confidence visual recognition of persons by a test of statistical independence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1993, zv. 15, č. 11.
- [8] DENG, J., GUO, J., YANG, J., XUE, N., COTSIA, I. et al. ArcFace: Additive Angular Margin Loss for Deep Face Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2021.
- [9] DUTA, I. C., LIU, L., ZHU, F. a SHAO, L. Improved residual networks for image and video recognition. *Proceedings - International Conference on Pattern Recognition*. 2020.
- [10] HE, K., ZHANG, X., REN, S. a SUN, J. Deep residual learning for image recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. December 2016.
- [11] HE, K., ZHANG, X., REN, S. a SUN, J. Identity mappings in deep residual networks. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 2016, 9908 LNCS.
- [12] HOSSEINI, M., ARAABI, B. a SOLTANIAN ZADEH, H. Pigment Melanin: Pattern for Iris Recognition. *Instrumentation and Measurement, IEEE Transactions on*. April 2010, zv. 59, č. 4.

- [13] HOWARD, A., SANDLER, M., CHEN, B., WANG, W., CHEN, L. C. et al. Searching for mobileNetV3. *Proceedings of the IEEE International Conference on Computer Vision*. Október 2019.
- [14] HOWARD, A. G., ZHU, M., CHEN, B., KALENICHENKO, D., WANG, W. et al. *MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications*. 2017.
- [15] HUANG, G., LIU, Z., MAATEN, L. V. D. a WEINBERGER, K. Q. Densely connected convolutional networks. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*. Január 2017.
- [16] IOFFE, S. a SZEGEDY, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *32nd International Conference on Machine Learning, ICML 2015*. 2015, zv. 1.
- [17] KHOSLA, P., TETERWAK, P., WANG, C., SARNA, A., TIAN, Y. et al. Supervised Contrastive Learning. In: LAROCHELLE, H., RANZATO, M., HADSELL, R., BALCAN, M. a LIN, H., ed. *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2020, sv. 33.
- [18] KINGMA, D. P. a BA, J. L. Adam: A method for stochastic optimization. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*. 2015.
- [19] KRIZHEVSKY, A., SUTSKEVER, I. a HINTON, G. E. ImageNet Classification with Deep Convolutional Neural Networks. In: PEREIRA, F., BURGESS, C., BOTTOU, L. a WEINBERGER, K., ed. *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2012, sv. 25.
- [20] KUMAR, A. a PASSI, A. Comparison and combination of iris matchers for reliable personal authentication. *Pattern Recognition*. 2010, zv. 43.
- [21] LAT, R. A., DANISHVAR, S., HERAVI, H. a DANISHVAR, M. Boosting Iris Recognition by Margin-Based Loss Functions. *Algorithms*. 2022, zv. 15.
- [22] LI, X., SUN, Z. a TAN, T. Comprehensive assessment of iris image quality. *2011 18th IEEE International Conference on Image Processing*. 2011.
- [23] MARTIN, A., DODDINGTON, G., KAMM, T., ORDOWSKI, M. a PRZYBOCKI, M. *The DET curve in assessment of detection task performance*. National Inst of Standards and Technology Gaithersburg MD, 1997.
- [24] MENG, Q., ZHAO, S., HUANG, Z. a ZHOU, F. MagFace: A Universal Representation for Face Recognition and Quality Assessment. *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE. Jún 2021.
- [25] MONTÚFAR, G., PASCANU, R., CHO, K. a BENGIO, Y. On the number of linear regions of deep neural networks. *Advances in Neural Information Processing Systems*. 2014, zv. 4.

- [26] OU, F. Z., CHEN, X., ZHANG, R., HUANG, Y., LI, S. et al. SDD-FIQA: Unsupervised Face Image Quality Assessment with Similarity Distribution Distance. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 2021.
- [27] POURSAEBI, A. a ARAABI, B. Iris Recognition for Partially Occluded Images: Methodology and Sensitivity Analysis. *EURASIP J. Adv. Sig. Proc.* Január 2007, zv. 2007.
- [28] SANDLER, M., HOWARD, A., ZHU, M., ZHMOGINOV, A. a CHEN, L. C. MobileNetV2: Inverted Residuals and Linear Bottlenecks. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 2018.
- [29] SCHROFF, F., KALENICHENKO, D. a PHILBIN, J. FaceNet: A unified embedding for face recognition and clustering. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Júl 2015, zv. 07.
- [30] SHI, Y. a JAIN, A. Probabilistic face embeddings. *Proceedings of the IEEE International Conference on Computer Vision*. Október 2019.
- [31] SIMONYAN, K. a ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*. 2015.
- [32] SRIVASTAVA, N., HINTON, G., KRIZHEVSKY, A., SUTSKEVER, I. a SALAKHUTDINOV, R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*. 2014, zv. 15, č. 56.
- [33] SZEGEDY, C., LIU, W., JIA, Y., SERMANET, P., REED, S. et al. Going deeper with convolutions. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Júl 2015, zv. 07.
- [34] TABASSI, E., GROTH, P. a SALAMON, W. *Iris Quality Calibration and Evaluation (IQCE): Evaluation Report*. NIST Interagency/Internal Report (NISTIR), National Institute of Standards and Technology, Gaithersburg, MD, september 2011.
- [35] TERHORST, P., KOLF, J. N., DAMER, N., KIRCHBUCHNER, F. a KUIJPER, A. SER-FIQ: Unsupervised Estimation of Face Image Quality Based on Stochastic Embedding Robustness. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 2020.
- [36] WANG, F., CHENG, J., LIU, W. a LIU, H. Additive Margin Softmax for Face Verification. *IEEE Signal Processing Letters*. 2018, zv. 25.
- [37] WANG, H., WANG, Y., ZHOU, Z., JI, X., GONG, D. et al. CosFace: Large Margin Cosine Loss for Deep Face Recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 2018.
- [38] WANG, L., ZHANG, K., REN, M., WANG, Y. a SUN, Z. Recognition Oriented Iris Image Quality Assessment in the Feature Space. *2020 IEEE International Joint Conference on Biometrics (IJCB)*. 2020.