VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
FACULTY OF INFORMATION TECHNOLOGY

ÚSTAV INTELIGENTNÍCH SYSTÉMŮ
DEPARTMENT OF INTELLIGENT SYSTEMS

DETEKCE PREZENTAČNÍHO ÚTOKU NA TECHNOLOGII SNÍMÁNÍ LIDSKÉ RUKY V INFRAČERVENÉ OBLASTI

PRESENTATION ATTACK DETECTION ON HAND SENSING TECHNOLOGY IN INFRARED AREA

BAKALÁŘSKÁ PRÁCE
BACHELOR'S THESIS

AUTOR PRÁCE **JAKUB RICHTARIK**
AUTHOR

VEDOUCÍ PRÁCE **Prof. Ing. MARTIN DRAHANSKÝ, Ph.D.**
SUPERVISOR

BRNO 2021

Vysoké učení technické v Brně
Fakulta informačních technologií

Ústav inteligentních systémů (UITS) Akademický rok 2020/2021

Zadání bakalářské práce

Student: **Richtarik Jakub**
Program: Informační technologie
Název: **Detekce prezentačního útoku na technologii snímání lidské ruky v infračervené oblasti**
Presentation Attack Detection on Hand Sensing Technology in Infrared Area
Kategorie: Vestavěné systémy

Zadání:

1. Prostudujte technologii multispektrální detekce živosti pro otisky prstů a celé ruky. Seznamte se s technologiemi snímačů otisků prstů se zabudovanou detekcí živosti, příp. i snímači ruky.
2. Navrhněte mechanismus snímání celých rukou v infračervené oblasti.
3. Realizujte snímací zařízení z bodu 2 a nasnímejte databázi alespoň 100 uživatelů (každá ruka minimálně 3 krát), přičemž nejméně 10 % musí tvořit osoby s výrazně odlišným množstvím melaninu v kůži (tyto hodnoty u každé osoby zaznamenejte). Zároveň nasnímejte databázi odpovídajícího rozsahu s falzifikáty rukou rozličných materiálů a povrchových úprav.
4. Navrhněte algoritmus rozlišení reálné a živé ruky od jakéhokoliv druhu prezentačního útoku.
5. Algoritmus z bodu 4 implementujte.
6. Proveďte experimentální ověření jak hardwarového, tak i softwarového řešení, a dosažené výsledky shrňte. Diskutujte možná rozšíření.

Literatura:

- GOMEZ-BARRERO, Marta; BUSCH, Christoph. Multi-Spectral Convolutional Neural Networks for Biometric Presentation Attack Detection. *NISK Journal*, 2019, 12.
- HEIDARI Mona, GOLDMANN Tomáš, DVOŘÁK Michal a DRAHANSKÝ Martin. Antispoofing and multispectral (optical) methods in hand-based biometrics. *Hand-Based Biometrics: Methods and Technology*. IET Book Series on Advances in Biometrics. London: The Institution of Engineering and Technology, 2018, s. 337-365. ISBN 978-1-78561-224-4.
- ROBISON, Charles D.; ANDREWS, Maxwell S. *System and method of fingerprint anti-spoofing protection using multi-spectral optical sensor array*. U.S. Patent No 10,242,245, 2019.

Pro udělení zápočtu za první semestr je požadováno:

- Body 1 a 2.

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

Vedoucí práce: **Drahanský Martin, prof. Ing., Dipl.-Ing., Ph.D.**
Vedoucí ústavu: Hanáček Petr, doc. Dr. Ing.
Datum zadání: 1. listopadu 2020
Datum odevzdání: 30. července 2021
Datum schválení: 11. listopadu 2020

Abstrakt

Pri verifikácii odtlačku prsta môže útočník použiť falzifikát, vyrobený z umelého materiálu. Zabrániť sa tomu dá napríklad využitím multispektrálnej analýzy, kedy rôzne materiály majú pri určitých vlnových dĺžkach inú odrazivosť. Existuje niekoľko štúdií, ktoré sa týmto zaoberali, avšak zamerané boli vždy na jeden prst. Cieľom tejto práce je detekcia živosti na celej dlani a prstoch, ako na väčšom objekte, čo prispeje k ešte väčšiemu zabezpečeniu. Vo výslednom riešení bola využitá NIR kamera na nasnímanie datasetu. Ten je použitý na natrénovanie konvolučnej siete, ktorá vie následne určiť, či sa jedná o živú ruku, alebo falzifikát.

Abstract

When verifying a fingerprint, an attacker can use a counterfeit made of synthetic material. This can be prevented, for example, by using multispectral analysis, when various materials have different reflectance for certain wavelengths. There are several studies that have addressed this, but have always focused on one finger. The aim of this work is liveness detection on the whole palm and fingers, as on a larger object, which will contribute to even higher level of security. In the final solution, a NIR camera was used to capture the dataset, which is used to train a convolutional network to determine whether it is a living hand or a counterfeit.

Kľúčové slová

detekcia prezentačného útoku, detekcia živosti, odtlačky prstov, biometria, multispektrálna analýza, infračervená oblasť, snímanie ľudskej ruky, konvolučné neurónové siete

Keywords

presentation attack detection, liveness detection, fingerprints, biometrics, multispectral analysis, infrared area, human hand imaging, convolutional neural networks

Citácia

RICHTARIK, Jakub. *Detekce prezentačního útoku na technologii snímání lidské ruky v infračervené oblasti*. Brno, 2021. Bakalářská práce. Vysoké učení technické v Brně, Fakulta informačních technologií. Vedoucí práce Prof. Ing. Martin Drahanský, Ph.D.

Detekce prezentačního útoku na technologii snímání lidské ruky v infračervené oblasti

Prehlásenie

Prehlasujem, že som túto bakalárskou prácu vypracoval samostatne pod vedením pána Prof. Ing., Dipl.-Ing. Martina Drahanského Ph.D. Uviedol som všetky literárne pramene, publikácie a ďalšie zdroje, z ktorých som čerpal.

.....
Jakub Richtarik
15. júla 2021

Poďakovanie

Chcel by som poďakovať pánu Prof. Ing., Dipl.-Ing. Martinovi Drahanskému Ph.D. za vedenie tejto bakalárskej práce a odborné rady. Ďakujem tiež svojej rodine a kamarátom za pomoc a podporu. Takisto som vďačný všetkým, ktorý sa zúčastnili snímania datasetu.

Obsah

Kapitola 1

Úvod

S neustálym vývojom nových technológií sa kladie čoraz väčší dôraz na bezpečnosť. Nielen kvôli zabezpečeniu vojenských objektov, vládnych budov, ale aj rôznych firiem ako takých. Najrýchlejším a zároveň najpresnejším spôsobom, ako overiť, či ide o danú osobu, je využiť jej špecifické a jedinečné znaky. Patrí sem napríklad rozpoznávanie tváre a tiež rozpoznávanie sietnice. Najrozšírenejšia je však detekcia odtlačkov prstov, kedysi využívaná len v kriminalistike.

Tak, ako existuje veľké množstvo rôznych metód detekcie odtlačkov prstov, existuje aj nespočetne veľa spôsobov, ako vytvoriť umelé odtlačky. Naskytuje sa preto otázka, ako správne detekovať živosť snímaného prstu. Realizovaných už bolo viacero prístupov, avšak s postupom času a vývojom nových technológií pribudli vždy aj nedostatky a zraniteľnosti. Ukazuje sa, že žiadne z doterajších riešení nie je úplne spoľahlivé.

Cieľom tejto práce bolo zmapovať rôzne, už existujúce riešenia v oblasti detekcie živosti, zhodnotiť ich prínosy a negatíva. Výsledkom je návrh a realizácia zariadenia na snímanie ruky v infračervenej oblasti, spolu s implementovaním navrhnutého algoritmu pre rozlíšenie živej ruky od prezentačného útoku, s ohľadom na už existujúce riešenia v oblasti snímania prstov.

Druhá kapitola rozoberá súčasný stav v oblasti detekcie živosti. Pre lepšie pochopenie je v jej úvode rozobratá stručná história využitia odtlačkov prstov a tiež falzifikátov. Väčší dôraz je kladený na 19. a 20. storočie. Značná časť je ďalej venovaná metódam detekcie živosti na prstoch a multispektrálnej analýze s detekciou živosti, založenou na jej princípe. V rámci metód detekcie živosti sa práca venuje najmä ich výhodám a nedostatkom, s ohľadom na už preukázané spôsoby, ako sa dajú oklamať. Následne sa táto kapitola venuje umelým neurónovým sieťam vo všeobecnosti. Popisuje ich základné princípy, stavbu, využitie v praxi, výhody a nevýhody. V skratke sú popísané základné druhy neurónových sietí a záverečná časť pojednáva o konvolučnej sieti, ktorá bola použitá pri implementovaní výsledného riešenia, vrátane vysvetlenia jej vrstiev.

Tretia kapitola je venovaná návrhu zariadenia, spolu s výberom kamier a objektívu. Taktiež hovorí o tom, ktoré vlnové dĺžky boli zvolené pre realizáciu snímania ruky, s ohľadom na zistené fakty v 2. kapitole. Načrtne zber datasetu, jeho rozdelenie, či prípadné problémy pri snímaní. Popisuje architektúru CNN, ktorá bola využitá pri prvotnom trénovaní a niekoľkých ďalších pokusoch.

V kapitole Realizácia je popísané výsledné skonštruované zariadenie. Rozobraté je snímanie datasetu, aj so stručným popisom implementácie, avšak nie zbytočne do hĺbky. Prehľadne je popísaná výsledná štruktúra datasetu. Nasleduje popis základnej konfigurácie

neurónovej siete a metrika, na základe ktorej boli porovnávané výstupy, spolu s experimentami a ich výsledkami.

Na záver je zhodnotený výsledok práce, ako po softwarovej, tak i po hardwarovej stránke. Rozanalyzované sú výhody a nedostatky. Prínosom je tiež úvaha nad budúcimi možnými vylepšeniami.

Kapitola 2

Prehľad súčasného stavu v oblasti detekcie živosti

Biometrické rozpoznávanie sa stalo každodennou súčasťou života. Či už ide o vstup do budov, firiem, pasovú kontrolu, alebo odomykanie mobilu. V súčasnosti sa síce do popredia dostáva rozpoznávanie tváre a sietnice, no najpoužívanejšou metódou je stále detekcia odtlačkov prstov.

2.1 História

Jedinečnosť odtlačkov prstov bola známa už medzi rokmi 206 a 201 pred Kristom v Egypte, a tiež v Číne, kde sa používali pri overovaní dokumentov. Aj v starovekom Grécku ľudia vedeli, že sú jedinečné a že existuje len mizivá šanca, že by stretli niekoho s veľmi podobným odtlačkom. Zanechávali ich napríklad na keramike, či iných výrobkoch, ako svoj autorský podpis.

Odtlačky prstov

Ich vlastnosti začal ako prvý podrobne skúmať Marcello Malpighi, taliansky profesor anatómie na University of Bologna. Bolo to v roku 1686 a podľa neho dostala aj pomenovanie malpighiho vrstva kože. Neskôr, v roku 1823, rozobral český profesor Jan Evangelista Purkyně, vo svojej práci, papilárne línie na dlani a prstoch a ako prvý popísal 9 základných vzorov, z ktorých sa v budúcnosti vychádzalo [1].

V roku 1882 použil geológ Gilbert Thompson odtlačky prstov na dokumente, aby zabránil jeho sfalšovaniu. Išlo o ich prvé známe využitie v Spojených štátoch amerických.

Rok nato, vychádza kniha *Life on the Mississippi* od známeho spisovateľa Marka Twaina, v ktorej bol vrah identifikovaný práve na základe odtlačkov. V tom čase však šlo zatiaľ len o fikciu.

Ich jedinečnosť sa v kriminalistike využila prvýkrát až v roku 1892, v Argentíne, pri identifikácii páchateľa vraždy [6].

V dvadsiatom storočí to nabralo rýchlejší spád. Od roku 1901 začali v Anglicku a všetkých anglicky hovoriacich krajinách používať Henryho klasifikačný systém a využívajú ho až do dnes. V Amerike sa začínajú používať vo väzniciach a tiež na policajných staniciach. V roku 1924 vzniká špeciálne oddelenie FBI, zaoberajúce sa identifikáciou a v roku 1946 už majú spracovaných prvých sto miliónov hárkov, každý pre jedného človeka, obsahujúci až 10 odtlačkov [3].

Falzifikáty

Prvé pokusy sa datujú až niekedy do roku 1924, kedy sa Alertovi Wehdemu, politickému radikálovi, podarilo na základe jeho skúseností s fotografovaním a gravírovaním vytvoriť prvý falošný odtlačok prsta. Najprv vytvoril kontrastnú fotografiu odtlačku, ktorú následne preniesol na tenký medený pliešok. Pomocou neho mohol na predmetoch zanechávať odtlačky iného človeka [19]. Na podobnom princípe fungujú falzifikáty aj dodnes, kedy sa takýto pliešok, či iný tvrdý materiál odtlačí do mäkkého materiálu, ktorý sa následne stane akýmsi umelým prstom.

V roku 1998 David Willis a Mike Lee oklamali skener odtlačkov prstov. Podarilo sa im to v štyroch prípadoch zo šiestich testovaných, kedy prístroje nevedeli odlíšiť reálny prst a jeho umelú kópiu. Odtlačok vyrobili zo silikónu vo voskovej forme a svoje zistenia aj publikovali v časopise Network Computing [41].

Ďalší veľkým míľnikom boli závery, ku ktorým došiel Tsutomu Matsumoto so svojou výskumnou skupinou na Yokohama National University, v januári roku 2002. Najprv testovali rôzne silikónové odtlačky a zistili, že systémy vybavené kapacitnými snímačmi a niektoré s optickými senzormi ich dokážu správne identifikovať ako falošné. Išli však ešte ďalej a vyrobili falzifikáty zo želatíny, ktorá obsahuje kolagén nachádzajúci sa v kostiach a tiež v rôznych tkanivách. Tým docielili presnejšiu napodobeninu ľudského prstu, ako z pohľadu vlhkosti, tak aj textúry a elektrického odporu. Testovaných bolo 11 zariadení, no tentokrát všetky nesprávne vyhodnotili odtlačky ako skutočné, ľudské [24].

V roku 2002 prebehli aj ďalšie experimenty, napríklad L. Thalheim, J. Krissler, a P. Ziegler, ktorým sa podarilo oklamať viaceré skenery. Kapacitné s použitím miniatúrnej "špongie", alebo dýchnutím na senzor, kedy sa tam dostali kvapôčky z dychu a falzifikát prstu sa zdal byť živým. Fungovalo aj využitie grafitového prášku spolu s adhéznou fóliou, ktorá bola tlačená na plochu skenera. Spomenutý prášok tiež fungoval na optické senzory za použitia halogénovej lampy a termálne senzory sa dali oklamať silikónovými odtlačkami, ktoré sa museli zahriať na určitú teplotu [39].

V súčasnosti sa využíva celá škála materiálov na výrobu falzifikátov, od vyššie spomínaných, latexu, cez tenké plastické, či papierové odtlačky, až po odtlačky z živice a rôzne materiály používané v potravinárskom priemysle.

2.2 Metódy detekcie živosti na prstoch

Na to, aby sme zamedzili, respektíve čo najviac znemožnili útočníkovi používať falzifikáty na oklamanie snímačov odtlačkov prstov, sa využívajú rôzne doplňujúce metódy. Mnohé zo súčasných existujúcich systémov na skenovanie odtlačkov prstov sú totiž sami o sebe zraniteľné. Buď sa môžeme zamerať na konkrétneho jedinca a pre každú osobu kontrolovať aj iné dáta, napríklad štruktúru ciev v prste [22], alebo môžeme využiť rôzne formy detekcie živosti. Hlavným cieľom je, aby sme dokázali určiť, či ide o odtlačok prsta ľudskej bytosti, alebo umelého prstu, nezávisle na tom, o koho ide.

Existujú dva prístupy. Prvým je detekovanie živosti počas snímania odtlačku prsta, využívajúc extra hardware. Nevýhodou tohto riešenia je zväčša vyššia cena, väčšie rozmery ako aj fakt, že riešenie je menej užívateľsky prívetivé. V niektorých prípadoch je dokonca možné použiť živý prst pre detekciu živosti a súčasne oklamať senzor falošným odtlačkom (nalepený falzifikát na živom prste). Druhý prístup je trochu náročnejší. Ide o odlíšenie neživých prípadov spracovávaním a vyhodnotením získaných dát, najmä pomocou neurónových sietí. Ide skôr o softwarové riešenia, sú lacnejšie a nie sú až tak rozmerné.

Detekcia s využitím prídavného hardwaru

Jednou z najjednoduchších metód na určenie živosti je meranie teploty na povrchu prsta. Ide o low-cost riešenie, ktoré sa dá ľahko oklamať. Keďže senzor na meranie teploty musí brať do úvahy niekoľko stupňovú odchýlku (v prípade vonkajšieho snímača ešte väčšiu), tak podľa [29] postačí, ak útočník použije tenkú vrstvu silikónu alebo želatíny s falošným odtlačkom, prilepenú na vlastnom prste.

Na teplotu, avšak nie ako statický ukazovateľ, sa zameriava aj ďalšia metóda. Telo totiž môže na vhodný podnet zareagovať zmenou toku krvi a reguláciou teploty v prste. Detekciou tejto zmeny a rozlíšením primeranej reakcie tak vieme identifikovať, či šlo o reakciu ľudského tkaniva. Nevýhodou tohto postupu je ale reakčný čas, ktorý bude rádovo v sekundách, keďže stimul nemôže byť nebezpečný, či nekomfortný pre užívateľa [16].

Inou možnosťou je detekovať pulz, napríklad meraním množstva svetla, ktoré prejde cez prst na druhú stranu. Hemoglobín má totiž rozličnú absorpciu svetla vzhľadom na aktuálne množstvo kyslíka v krvi. Tu sa však stretávame hneď s niekoľkými problémami. Ak opäť použijeme na prste veľmi tenký falzifikát, pulz zdetekujeme rovnako, ako keby tam daný falzifikát nebol. Musíme tiež vziať do úvahy, že človek má veľmi variabilný tep. V pokoji to môže byť pod 60 úderov za minútu, no pri námahe či strese aj cez 120. Navyše, ak by mal človek príliš nízky tep, musel by držať prst na senzore aj 4 sekundy, aby ho zmeral. Podobným spôsobom sa dá využiť oximetria (z medicíny), ktorá vie určiť koncentráciu kyslíka v krvi. Podľa [33] ale pri použití želatíny odmeriame hodnoty z prstu a opäť nezdetekujeme útok.

Na určenie živosti sa dá tiež použiť meranie tlaku krvi. Väčšina takýchto senzorov však musí merať tlak priamo zo žily, alebo dokonca na dvoch miestach naraz (obidve ruky). Opäť sa ale dostávame k problému, že pri použití super tenkej vrstvy silikónu by sme odmerali tlak vo vnútri prsta a umelý odtlačok označili za živý.

Kedysi sa efektívnou zdala detekcia na základe rôzneho elektrického odporu pre ľudský prst a iné materiály. Ten však veľmi závisí od vlhkosti. Keďže už vyššie spomínaný Matsumoto zistil [24], že vlhkosť na živom prste je na úrovni 16% a pre želatínový odtlačok okolo 23%, táto metóda je v praxi taktiež nepoužiteľná. Je to príliš malý rozdiel. Na vlhkosť má navyše veľký vplyv počasie, emócie a tiež hladina stresu, ktorá zvyšuje potenie, a preto by sa tento rozdiel poľahky stratil. Navyše existujú ľudia, ktorý sa potia až príliš (hyperhydróza).

V minulosti boli pokusy skombinovať viaceré metódy, ako napríklad meranie teploty na prste, spolu s oximetriou a pridaním EKG (elektrokardiogram) do rovnice. Teplota a oximetria sa ale dá obísť už spomínanou želatínou, pričom pri EKG by musel byť prst držaný bez pohnutia nejakých 6 až 8 sekúnd.

Zaujímavé je i využitie vlastnosti, že vzor odtlačku prsta sa nachádza nie len na jeho povrchu, ale aj pod jeho vrchnou vrstvou, epidermou. Na jeho zachytenie sa dá využiť napríklad ultrazvukový senzor, vďaka ktorému by bol tento systém náročnejší na prekonanie útočníkom. Musel by totiž vedieť, aká metóda bola použitá na detekciu živosti, a následne by musel vzor odtlačku preniesť do vnútra falošného odtlačku, v rovnakej pozícii (to by vedel dosiahnuť napríklad zubný technik) [29]. Sila tohto systému spočíva najmä v nevedomosti útočníka, aká metóda detekcie je použitá. Ak by sa ale použil takýto senzor na porovnávanie štruktúry ciev konkrétnej osoby, tak by bol takýto systém takmer neprekonateľný.

K detekcii živosti patria aj metódy zaoberajúce sa optickými charakteristikami, ktoré sú založené na rôznej absorpcii, či odrazivosti umelých materiálov vzhľadom ku koži. Viaceré z nich sa síce dajú oklamať želatínou, ktorá má podobné vlastnosti, no do veľkej miery to

závisí aj od použitého žiarenia (R/G/B, infračervené, laser). Nemusíme sa ale zamerať len na povrch prsta. Ako uvádza štúdia [16], pomocou zdroja NIR svetla a CCD kamery dokážeme detekovať cievy v prste, či už jeho presvietením naskrz, kedy je prst medzi zdrojom a kamerou, alebo aj pomocou zachytenia odrazu. Takto vieme získať celý vzor ciev v rámci prstu, čo ešte viac komplikuje použitie falzifikátov.

Mnohé firmy si zakladajú na princípe, že svoje riešenia len tak nezverejňujú, takže dopátrať sa dá len k tým z verejných zdrojov. Riadia sa heslom *Security through obscurity*, alebo tiež *bezpečnosť skrze neznalosť*, čo zabezpečí, že ich systém bude ťažšie prelomiteľným. Teda, aspoň na začiatku.

Detekcia vyhodnocovaním získaných informácií

Tieto metódy sú založené na špecifických informáciách, ktoré boli zistené o koži a odtlačkoch prstov. Prvou z nich je tá, ktorá využíva deformitu. Zameriava sa na to, ako je koža deformovaná, keď sa prst pritlačí na povrch senzoru, keďže zakaždým dôjde k inej nelineárnej deformite. To isté môžeme pozorovať pri posunutí prstu po povrchu senzora do strán. Dnes je ale dokázané, že aj tieto zmeny sa už dajú napodobniť [23].

K ďalšej metóde je potrebný senzor s veľmi vysokým rozlíšením. Je to preto, aby sme mohli zachytiť veľmi malé detaily, napríklad zakončenia potných žliaz, ktoré je náročné imitovať. Podľa Matsumota [24] je aj to možné so želatínou. Ďalej však tvrdí, že je dosť komplikované, takmer až nemožné, presne zreplikovať veľkosť, tvar a pozíciu týchto pórov.

Podľa Marie Sandstrom [32] sa potením zaoberali aj na Biomedical Signal Analysis Laboratory v USA, kedy skúmali priebeh potenia v čase na videu. Len veľmi komplikovane sa totiž dá napodobniť priebeh postupného uvoľňovania potu z pórov, a práve preto sú dynamické analýzy častokrát lepšie. Výsledky boli sľubné, no o ďalšom priebehu projektu sa toho mnoho nedá dohľadať. Veľmi dobré výsledky dosiahli aj autori práce [27], v ktorej porovnávali snímače založené na kapacitnej, elektrooptickej a optickej technológii.

V máji tohto roku (2020) vyšla štúdia [17], kde sa výskumníci snažili odlíšiť živý prst od falzifikátu z videa, s využitím optického a termálneho senzoru. V závere práce zhodnotili, že prst je veľmi bohatým zdrojom informácií (potenie, teplota, elasticita), než len statickým vzorom odtlačku.

Jedna z metód využíva na určenie živosti analýzu šumu. Ľudský prst má drážky hladké, no umelý ich má pri pohľade zblízka zrnité. Výsledky ukázali, že pre kapacitné, optické a elektrooptické senzory je tento prístup veľmi efektívny (90%-100% úspešnosť) [2].

Na podobnom princípe je založený aj systém, ktorý sa zameriava na čistotu, hrúbku a spojitosť reliéfov papilárnych línií, pričom v štúdii [9] dosiahol úspešnosť cez 90%. Takýto postup je navyše veľmi rýchly, pri porovnaní s inými.

Ďalšou možnosťou je využiť sledovanie reakcie na určitý podnet. Ako príklad je možné uviesť svalovú reakciu na elektrický stimul, alebo zmenu farby prstu po pritlačení na povrch senzoru [33]. Nevýhodou ale je, že je to nepraktické. Stačí si predstaviť, že by mal užívateľ zakaždým dostať do prsta výboj elektriny.

Ďalším spôsobom je sledovanie zmien na povrchu kože, v dôsledku rozširovania a zužovania ciev spôsobených tepom. Môžeme tak pozorovať periodickú zmenu pozície kože, ako aj zmenu jej farby, z dôvodu nárastu a poklesu hladiny kyslíka v krvi [16].

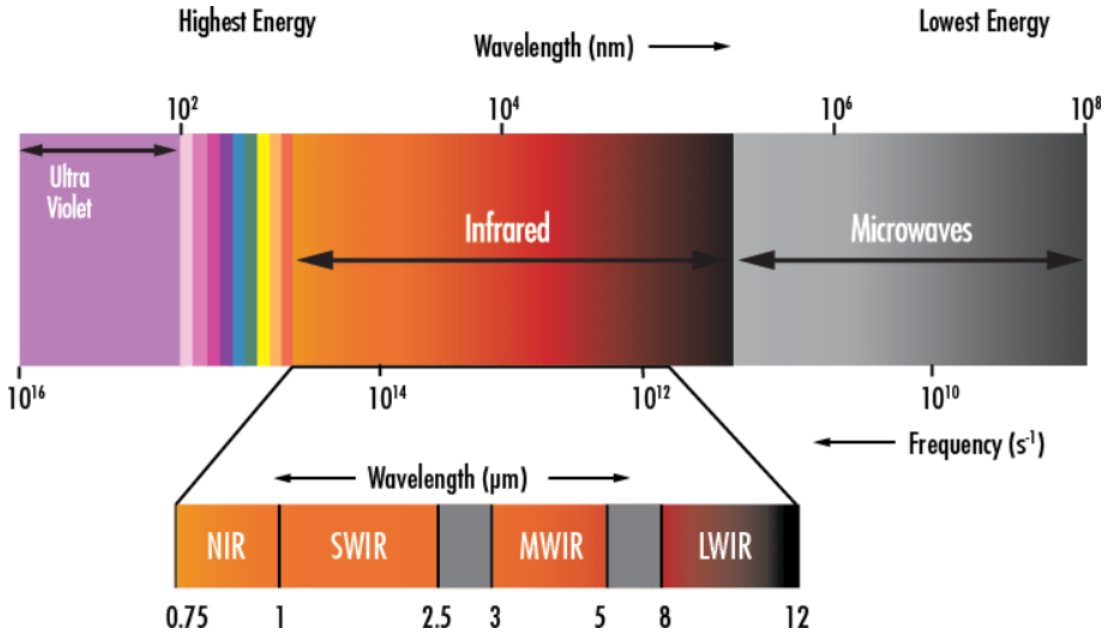
V súčasnosti sa často využívajú najmä riešenia, založené na konvolučných neurónových sieťach. Prvé, od R.F. Nogueira, datujeme do roku 2014 [8]. Ich výsledkom je zväčša veľmi rýchle a komplexnejšie určenie živosti, avšak, ako sa ukázalo v [7] z januára 2020, môžu byť zraniteľné. V angličtine *adversarial attack* [10], je druh útoku na konvolučnú sieť, kedy

pri určitých úpravách na vstupe, vieme na výstupe dosiahnuť výsledok, aký chceme. Práve týmto útokom dokázali J. Fei a jeho tím, že neurónová sieť môže za istých podmienok vyhodnotiť snímok umelého odtlačku po jeho úprave, ako skutočný odtlačok živého prsta. Otvorili tak diskusiu o bezpečnosti týchto systémov a o potrebe tvorby robustnejších algoritmov.

2.3 Multispektrálna analýza

V dnešnom svete zohráva dôležitú úlohu spracovávanie obrazu, no najmä jeho rozpoznávanie, spolu s celou škálou potrebných matematických algoritmov. Či už ide o rozpoznávanie písaného textu, rôznych objektov, alebo zvierat. Veľký význam má hlavne v čoraz viac preferovaných autonómnych autách, ktoré musia byť schopné detekovať chodcov na vozovke, aby včas zabrzdili, jazdné pruhy a dopravné značky pri krajnici. Ďalšou veľkou oblasťou je rozpoznávanie tváre. Najmä na letiskách (nežiadúce osoby, teroristi), a vo vládnych inštitúciách. V Číne sú napríklad na systém rozpoznávania tvárí napojené takmer všetky kamery, a preto ich databáza obsahuje záznamy o viac ako 2 miliardách ľudí.

Niekedy obyčajné rozpoznávanie na základe tvaru, farby, alebo pohybu vo viditeľnom spektre nestačí. Ako by sme dokázali z obrazu odlíšiť dva rôzne prášky rovnakej farby? Tento problém sa snaží odstrániť multispektrálna analýza. Okrem viditeľného svetla, sa totiž elektromagnetické spektrum, podľa rozličných vlnových dĺžok, skladá aj z rádiových vĺn, mikrovlnného žiarenia, infračerveného a ultrafialového svetla, röntgenového a gama žiarenia. Zároveň platí, že rozličné materiály, či dokonca samotné látky v nich, reagujú pri určitej vlnovej dĺžke inak. Majú rôznu mieru odrazivosti a to najmä u NIR a SWIR skupiny (viď obrázok 2.1), ktoré spadajú do rozpätia infračerveného svetla. Na tomto fakte je založená multispektrálna analýza, kde sa vyhotovujú snímky pri rôznych vlnových dĺžkach a rôznych zložkách elektromagnetického spektra a to pomocou filtrov, alebo senzorov citlivých len na určité dĺžky.

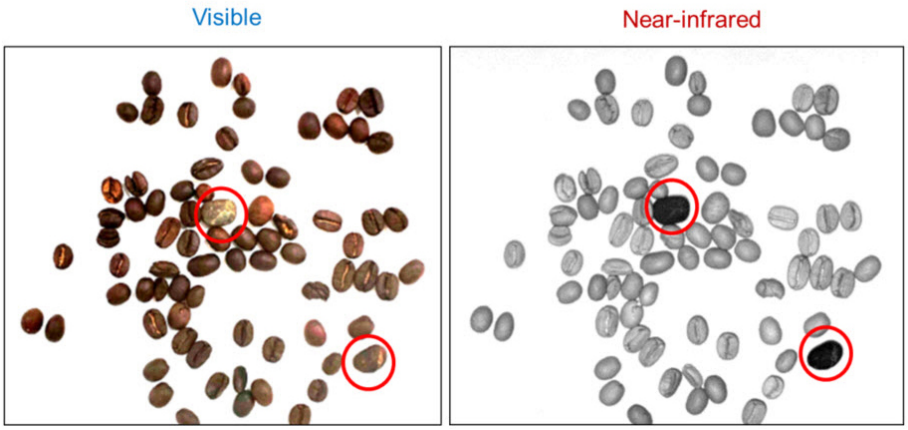


Obr. 2.1: Detail zložiek infračerveného svetla. Zdroj: [18]

Rôzne publikácie udávajú iné hraničné vlnové dĺžky. Jeden z výrobcov kamier uvádza, že označenie NIR sa používa pre vlnové dĺžky 780-1400 nm, SWIR pre 1400-3000 nm, ba dokonca, že niekedy je jedno z označení použité na celý rozsah 780-3000 nm [14].

Využitie

Multispektrálna analýza má v posledných rokoch čoraz väčšie uplatnenie. Najmä v potravinárskom priemysle pri pásovom dopravníku. Pomocou nej dokážeme napríklad odlíšiť obité, prípadne nahnité jablká, alebo iné ovocie, zdetekovať kamienky medzi zrnkami kávy (obr. 2.2), či obilnín. Tiež je používaná, pri farmárčení (minimalizovanie postrekov podľa škôd), mapovaní deforestácie, ale aj pri prepisoch starých dokumentov, kedy má atrament a podklad po toľkých rokoch rovnakú farbu.



Obr. 2.2: Detekcia kamienkov medzi zrnkami kávy. Zdroj: [14]

Veľkým prínosom sa javí pri triedení odpadu, ako je plast, pretože niektoré plasty sú nerecyklovateľné, prípadne vyžadujú iné postupy. Pomocou SWIR kamier napríklad vieme odlíšiť PVC, PS, PET a akryl rovnakej farby [15]. Takisto vieme určiť množstvá tekutiny v nepriehľadných nádobách, alebo miniatúrne trhliny v obale. Vo fotovoltaike je napríklad možné pomocou tejto technológie detekovať nielen fyzické praskliny, ale aj oslabené a mŕtve oblasti. Ďalším využitím je mapovanie priestorového rozloženia minerálov na exponovaných územiach a tiež sledovanie banských činností a ich vplyvov [26]. S tým súvisí aj to, že takto dokážeme zisťovať prítomnosť určitých minerálov aj na iných planétach, mesiacoch a iných vesmírnych telesách.

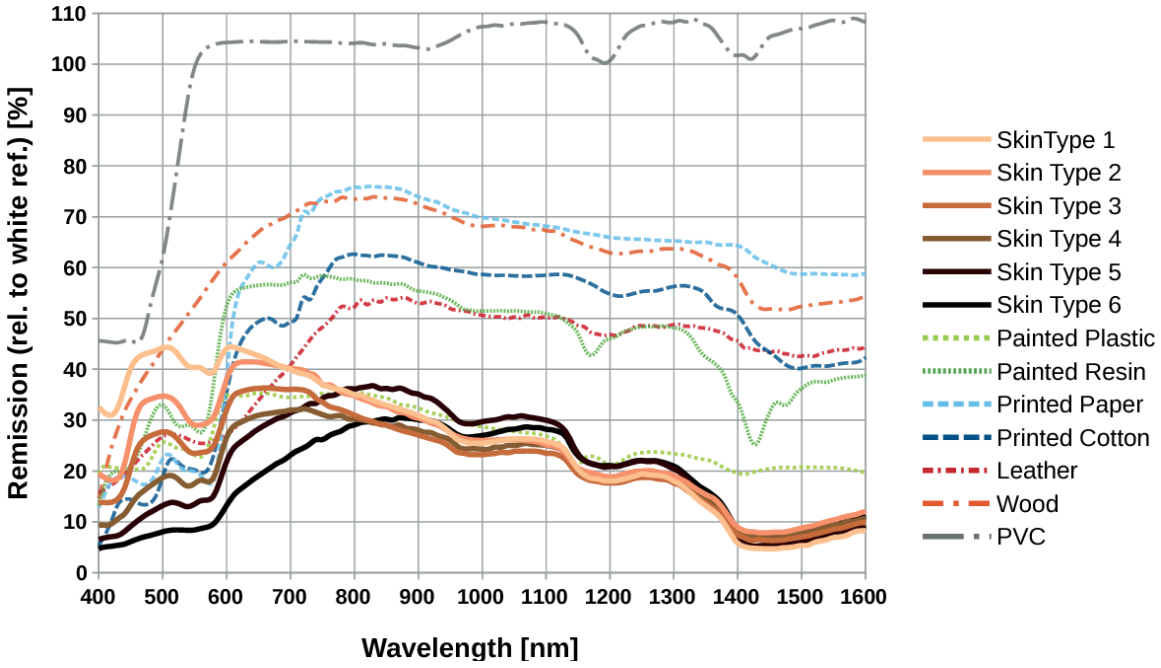
V armáde nachádza využitie pri vyhľadávaní výbušnín, detekcii nepriateľských jednotiek, nočnom či termálnom videní. Veľmi dôležitým komponentom je v obranných systémoch proti balistickým strelám, kedy vie zdetekovať telo rakety pomocou strednovlnného infračerveného svetla a jej chvost pomocou dlhovlnného.

V najbližších rokoch bude mať obrovský dopad aj na medicínu, pri liečení rôznych chorôb. V štúdii [35] bolo dokázané, že zo snímkov multispektrálneho zobrazenia je možné rozoznať nádorové a zdravé tkanivá. V budúcnosti by mohli tieto poznatky, v kombinácii s umelou inteligenciou, prispieť k rýchlejšej a presnejšej identifikácii rakoviny a tiež prevencii vzniku ťažšieho štádia. Pomocou takéhoto systému by sa zároveň dalo overiť, či boli u pacienta, pri prípadnej operácii, úspešne odstránené všetky rakovinotvorné tkanivá.

Keďže je multispektrálna analýza veľmi efektívna pri odlišovaní rôznych látok, dá sa využiť aj pri detekovaní ľudskej kože, čiže pri detekcii živosti.

2.4 Detekcia živosti pomocou multispektrálnej analýzy

Ako už bolo spomenuté, multispektrálna analýza sa dá využiť pre rozlíšenie rôznych materiálov a látok, z čoho vyplýva, že je využiteľná aj pri detekcii kože. Práve týmto smerom sa v súčasnosti začína uberať detekcia živosti, či už pri skenovaní tváre, alebo odtlačkov prstov, kedy neurónová sieť po vyhodnotení SWIR snímkov určí, či ide o ľudskú kožu [40], [38]. Ak sa bližšie pozrieme na to, koľko žiarenia pri danej vlnovej dĺžke odrazí koža v porovnaní s inými materiálmi (percentuálne k bielej farbe), zistíme, že pri vyšších vlnových dĺžkach sa rôzne typy kože správajú rovnako (viď graf na obrázku 2.3).



Obr. 2.3: Remisia (odrazivosť v pomere k odrazivosti bielej farby) kože a iných materiálov v %. Zdroj: [36]

Na uvedenom grafe môžeme pozorovať, že pri vlnovej dĺžke menšej ako 1300 nm, by bolo náročné, ba priam až nemožné, odlíšiť kožu od zafarbeného plastu. Z tohto dôvodu musia byť vo výslednom riešení, ak má byť spoľahlivé, zahrnuté aj vyššie vlnové dĺžky, avšak s ohľadom na odrazivosť iných materiálov, ktoré nie sú zobrazené na grafe. Vhodným príkladom je mokré drevo, ktorého odrazivosť je pri vlnových dĺžkach do 700 nm veľmi podobná koži, a pri rozsahu približne 1400-1500 nm, môže byť dokonca identická [37].

Skvelým príkladom, ako sa dajú spojiť poznatky o multispektrálnych vlastnostiach kože s detekciou živosti, je práca [11], ktorú vypracovali Gomez-Barrero a Busch. V nej sa snažili vytvoriť spoľahlivý mechanizmus, na odlíšenie živých a neživých odtlačkov prstov vo SWIR oblasti svetla. Pre každý prst využili čiernobielu štvoricu snímkov (pri 1200, 1300, 1450 a 1550 nm), ktoré vždy spracovali na jeden RGB obrázok. Ten bol následne vstupom pre jednu z vybratých konvolučných neurónových sietí, ktorých výstupy navzájom porovnávali.

Neskôr testovali aj fúziu viacerých sietí. Dosiahli značný úspech a potvrdili tak, že detekcia živosti by sa mala uberať smerom k neurónovým sieťam.

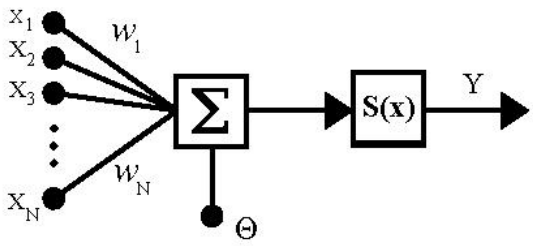
2.5 Umelé neurónové siete

Pre správnu detekciu prezentačného útoku na technológiu snímania ľudskej ruky budeme využívať konvolučnú neurónovú sieť (ich fúziu). Najprv si však vysvetlíme, čo sú to umelé neurónové siete a pozrieme sa na typy sietí. Keďže ich existuje veľké množstvo, spomenieme si len niektoré. Nebudeme ich dopodrobna rozoberať, keďže to nie je cieľom tejto práce.

Úvod

Umelá neurónová sieť je paralelný výpočtový model, alebo tiež systém, ktorý vznikol ako abstrakcia vlastností odpovedajúcich biologických neurónových systémov. Tak ako existuje biologický neurón, tak umelá neurónová sieť pozostáva z umelých neurónov, obsahujúcich ľubovoľný počet vstupov a jeden výstup. Tieto neuróny sú základnou stavebnou jednotkou siete a môžu si navzájom predávať signály v podobe reálnych čísel. Spojenia s neurónmi nazývame hranami, pričom každá hrana má svoju váhu.

Existuje celý rad modelov umelého neurónu, no jeden z najpoužívanejších, tak ako ho definovali McCulloch a Pitts v roku 1943 (prvý výpočtový model), môžeme vidieť na nasledujúcom obrázku (2.4).



Obr. 2.4: Základný model neurónu. Zdroj: [20]

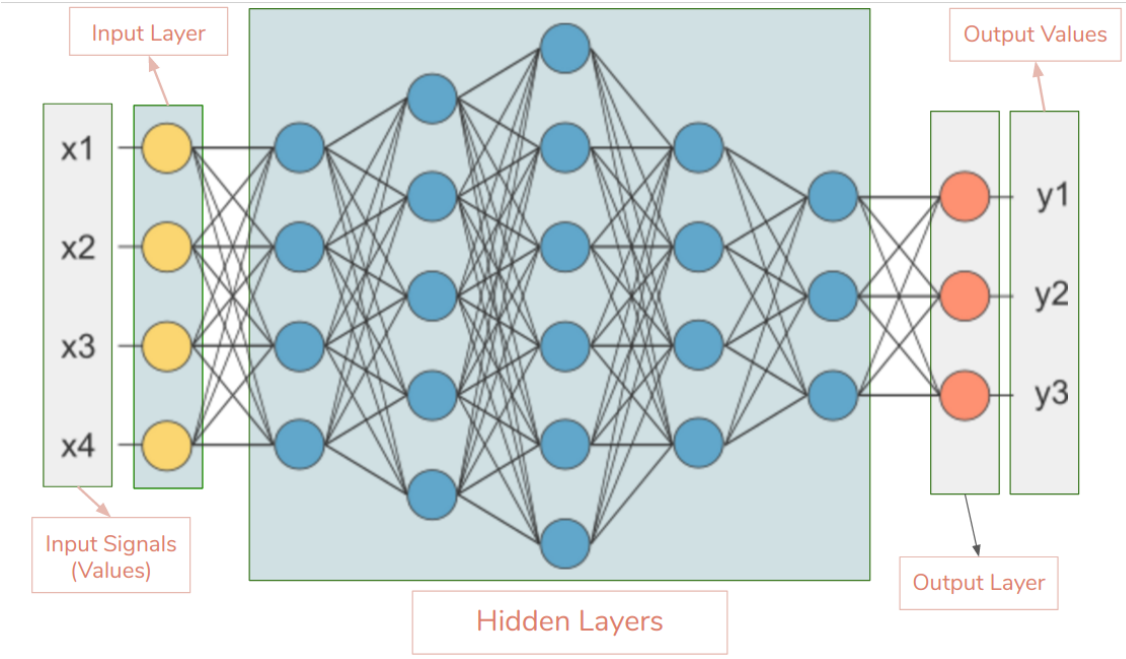
Tento obrázok odpovedá rovnici:

$$Y = S(\sum_{i=1}^{N}(w_i x_i) + \Theta)$$

kde hodnoty x_1 až x_N reprezentujú vstupy neurónu a w_i sú ich príslušné váhy. Vyššia váha znamená vyššiu dôležitosť daného vstupu, respektíve veľkosť jeho vplyvu na výsledok. Θ je prah, inak nazývaný bias a ovplyvňuje aktivačnú funkciu. $S(x)$ je prenosovou (aktivačnou) funkciou neurónu a Y je jeho samotný výstup. Používajú sa rôzne aktivačné funkcie, pričom každá má iné výhody a nevýhody a používa sa v iných prípadoch.

Ako môžeme vidieť na obrázku 2.5 na ďalšej strane, viaceré neuróny je možné spájať do väčšej neurónovej siete, ktorá je pri hĺbkovom učení tvorená vrstvami. Hĺbka modelu značí počet vrstiev, pričom výstupy jednej vrstvy, sú zároveň vstupmi tej ďalšej. Platí, že prvá vrstva, cez ktorú dáta vstupujú do siete, sa nazýva vstupná, posledná výstupná produkuje výsledky a medzi nimi sa nachádzajú vrstvy skryté [28]. Existujú však aj siete s jednou vrstvou, prípadne bez vrstiev. V prípade, že neuróny posielajú informáciu len neurónom

v nasledujúcej vrstve (v smere od vstupnej k výstupnej), hovoríme o dopredných sieťach, no v prípade, že sa v sieti nachádzajú aj spätné väzby, ide o rekurentné siete. Na modeli z obrázka 2.5 by boli prípadné smery väzieb naznačené šípkami.



Obr. 2.5: Vrstvy neurónovej siete. Zdroj: [28]

Aby sa neurónová sieť správala tak ako chceme, musí prejsť procesom učenia. Počas tohto procesu sa upravujú hodnoty jednotlivých váh. Existuje učenie s učiteľom, kedy pre trénovacie dáta poznáme požadované výsledky, a bez učiteľa, kedy sieti poskytneme meradlo kvality reprezentácie, ktorú sa má naučiť. Kým učenie bez učiteľa môže byť častokrát nepredvídateľné, dá sa použiť na komplexnejšie problémy. Tie rieši pomocou zhlukovania a asociácií. Pri zhlukovaní zoskupuje prvky s podobnými vlastnosťami do skupín a pri asociáciách hľadá určité pravidlá, ktoré opisujú vzťahy medzi väčšinou prvkov dát [13].

Výhody a limity

Medzi hlavné výhody umelých neurónových sietí, patrí rýchlosť spracovania informácií, najmä ak ide o prípady bez učiaceho sa algoritmu. Ich veľkou výhodou je tiež redukcia rozmeru dát a schopnosť abstrakcie riadiacich pravidiel nejakého regulátora (s pomalším výpočtom) a jeho nahradenie, napríklad u človeka. Sú vhodné pre úlohy zahrňujúce klasifikáciu, identifikáciu, či aproximáciu, pričom aproximovať dokážu ľubovoľnú spojitú funkciu s akoukoľvek presnosťou.

Nevýhodou však je, že pri ich implementácii je jediným riešením metóda pokusu a omylu, keďže neexistujú žiadne zadefinované postupy ako navrhnúť správnu architektúru siete, alebo podľa čoho zvoliť funkciu opisujúcu neurón. Nedokážeme tiež simulovať podobný paralelizmus ako u ľudských neurónov (ľudské neuróny v mozgu obsahujú viac ako 10^3 synapsií). Zároveň platí, že učenie je zväčša časovo náročné, a získavame riešenie, ktoré nemusí byť stopercentne presné.

Využitie

- Detekcia rozličných objektov, zvierat, či človeka,
- rozpoznávanie textu, tváre a reči,
- analyzovanie a vyhodnocovanie dát (predpovede počasia), optimalizácia,
- inteligentné domácnosti, robotika,
- osobní asistenti, napríklad od Googlu, alebo Amazon Alexa, či Siri,
- autonómne autá (Tesla),
- návrhy (eshopy, či iné služby), na základe predikcie,
- finančné trhy - obchodovanie, risk management a iné [4].

Typy sietí

Rôzne druhy sietí rozlišujeme podľa ich topológie. Tým najzákladnejším je Perceptrón (P). Ide o sieť s jediným neurónom, ktorá slúži na dichotomickú klasifikáciu (2 skupiny). Existuje aj viacvrstvová perceptronová sieť, obsahujúca 3 a viac vrstiev perceptrónov, ktorá sa využíva pri strojovom preklade rôznych textov, či rozpoznávaní reči.

Dopredná, feed-forward (FF), neurónová sieť je umelou neurónovou sieťou, v ktorej spojenia medzi uzlami netvoria cykly. Zároveň sa v nej informácie šíria len jedným smerom, smerom od vstupnej vrstvy k výstupnej. Faktom je, že jednovrstvová perceptronová sieť je poddruhom doprednej siete [12].

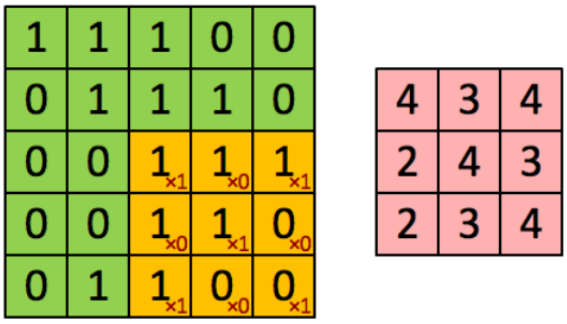
Veľmi známa je tiež rekurentná sieť (RNN), ktorá je poddruhom doprednej siete. Táto sieť sa využíva v prípade, že niektorý z neurónov, potrebuje poznať predchádzajúcu informáciu v aktuálnej iterácii (napríklad predchádzajúce slovo pri predikcii nasledujúceho slova vo vete). Nevýhodou siete je pomalšia rýchlosť a neschopnosť pamätať si aj veľmi staré informácie [34].

Medzi známe siete ešte patrí autoenkodér (AE), kde učenie prebieha bez učiteľa. Pri tomto type je počet skrytých neurónov menší, ako tých vo vstupnej a výstupnej vrstve. Pozostáva z 2 častí, enkodéru a dekodéru.

2.6 Konvulučná neurónová sieť

Konvolučná neurónová sieť (CNN) bola inšpirovaná vizuálnym kortexom mozgu. Ide o algoritmus hlbokého učenia, ktorý vo všeobecnosti na vstupe očakáva obrázok, čiže maticu s hodnotami pixelov, a dokáže v ňom rozlišovať jednotlivé objekty. Práve kvôli tomu nachádza využitie najmä v oblasti počítačového videnia. V prípade, že ide o farebný RGB obrázok, tak je vstup reprezentovaný 3 kanálmi, čiže 3 maticami. Táto sieť obsahuje 3 základné typy vrstiev.

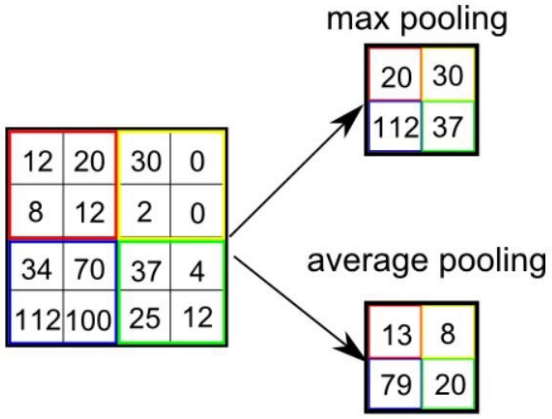
Konvolučná vrstva vykonáva konvolúciu, ktorej jadro sa nazýva Kernel, K, alebo tiež Filter a má vždy rovnakú hĺbku, ako počet kanálov vstupného obrázku. Cieľom konvolučných operácií je extrahovať hrany, farbu, a podobne. Na obrázku 2.6 môžeme pozorovať, ako prebehol posledný krok konvolúcie, kedy sme zo vstupnej matici o veľkosti 5x5 (matica vľavo) s Kernelom o veľkosti 3x3 (červený text), získali výstupnú 3x3 maticu. Niekedy



Obr. 2.6: Posledný krok konvolúcie. Zdroj: [31]

sa používa aj dekonvolučná (transponovaná konvolučná) vrstva, ktorá vykonáva opačnú operáciu, teda sa snaží odhadnúť maticu na vstupe zo znalosti tej výstupnej [31].

Ďalšou vrstvou je **pooling vrstva**, ktorej cieľom je znížiť priestorovú veľkosť vstupu, a teda nepriamo zlepšiť distribúciu informácie. Taktiež znižuje výpočetnú náročnosť učenia. Opäť si zadefinujeme Kernel o určitej veľkosti, a podľa neho môžeme vyberať maximálne hodnoty (max pooling), alebo priemerné hodnoty (average pooling) [31]. Na obrázku 2.7 môžeme vidieť obidva výstupy pre vstupnú maticu o veľkosti 4x4 a Kernel o veľkosti 2x2.



Obr. 2.7: Typy poolingu. Zdroj: [31]

Poslednou je **fully connected (FC) vrstva**, nazývaná tiež lineárna, alebo dense vrstva, ktorá zabezpečí finálny výsledok. Tá vytvára spojenia medzi každým vstupným a výstupným prvkom, čím napomáha distribúcii informácie. Pre každú vrstvu je potrebné definovať rozmer vstupu a výstupu. Keďže očakáva na vstupe vektor, musíme vstupnú maticu vektorizovať [42].

Existuje nespočetné množstvo rôznych architektúr a vznikajú stále ďalšie. Medzi tie najznámejšie však patria LeNet, AlexNet, VGGNet, ResNet, či MobileNet.

Kapitola 3

Návrh

V tejto kapitole si rozoberieme návrh zariadenia na snímanie ruky, ktoré bude neskôr s prípadnými úpravami realizované. Využije sa pri tvorbe datasetu, ktorý poslúži pre natrénovanie vybranej konvolučnej neurónovej siete, aby vedel celý mechanizmus spoľahlivo rozlišovať živú ruku od akéhokoľvek druhu prezentačného útoku, čiže od ruky umelej. Pod pojmom ruka je potrebné si predstaviť oblasť prstov a dlane až po predlaktie.

3.1 Zariadenie na snímanie ruky

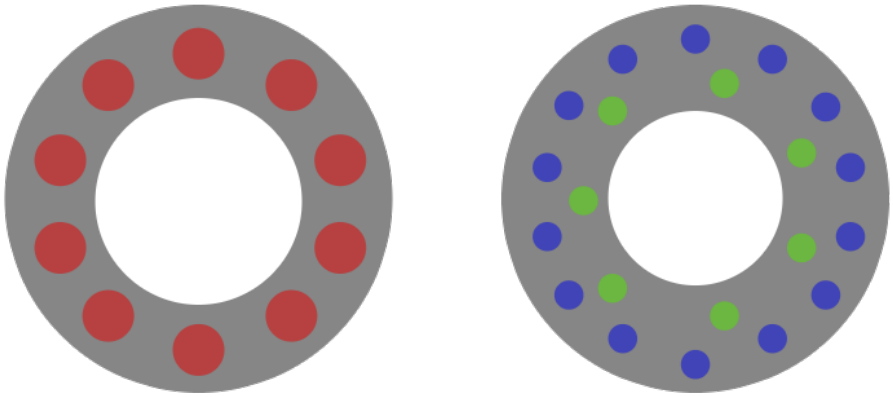
V nadväznosti na sekcie 2.3 a 2.4 je potrebná kamera, respektíve nejaký senzor, ktorý by bol funkčný pri určitých vlnových dĺžkach od približne 700 po 1600 nm, a rovnako aj zdroj svetla.

Na začiatku bolo nutné preskúmať trh s kamerami, čo zahŕňalo kontaktovanie mnohých zahraničných firiem. Po zvážení všetkých možností, prichádzali do úvahy len kamery InGaAs, alebo kamery so senzorom na báze germánia (kvôli funkčnosti v potrebnom vlnovom rozsahu). V oboch prípadoch však šlo o neakceptovateľne drahé zariadenia, v desiatkach tisíc eur. Nakoniec som ale narazil na iný druh kamery, a tá bola vo výsledku aj zakúpená. Šlo o kameru Digital CamIR 1550 USB 2.0, ktorá v porovnaní s najlacnejšími InGaAs kamerami stála len približne štvrtinu ceny a mala dvojnásobne lepšie rozlíšenie (1296x964). Podľa vzdialenosti medzi miestom inštalácie kamery a snímanou rukou bol ešte následne vybraný potrebný objektív. Zvolený bol Kowa LM8HC-SW 8mm. Nevýhodou kamery je však malý rozsah vlnovej dĺžky (1500-1600 nm), kvôli čomu bola potrebná ešte jedna. Vybratá bola IDS UI-3360CP-NIR-GL R, ktorú mala fakulta k dispozícii spolu s objektívom.

Vychádzal som zo štúdie [11], ktorej autori realizovali detekciu živosti na prstoch a zvolili si 4 vlnové dĺžky, pri ktorých vyhotovovali snímky. Na rozdiel od nich som zvolil iba 3, a to kvôli použitiu 2 kamier, ktoré nepokrývali celý rozsah vlnových dĺžok. Sken ruky bude vo výsledku predstavovať vyhotovenie 3 snímkov, kedy senzory zachytia odraz svetla od kože pri 3 rôznych vlnových dĺžkach.

V závislosti od citlivosti prvej kamery sa použijú LED diódy EOLD-1550-525, ktorých špičková vlnová dĺžka je 1550 nm. S ohľadom na ich cenu predpokladáme v návrhu použitie 10 kusov. Pre druhú kameru, s ohľadom na jej efektivitu a na graf na obrázku 2.3, zvolíme 2 vlnové dĺžky, 850 nm a 740 nm. Tým prislúcha 7 kusov s označením 15400585A3590 a 14 kusov OIS-330-740-X-T (SMD LED). Pri výbere diód bol kladený dôraz najmä na intenzitu vyžarovania. Tá mala vplyv na počet použitých kusov.

V zariadení by kamery mali byť namontované bezprostredne pri sebe. V ich okolí sa predpokladá zdroj svetla, čiže vybrané LED diódy. Tu bola inšpiráciou práca [36], v ktorej okrem iného testovali efektívnosť rôznych rozložení diód okolo kamery. Výsledný návrh rozloženia diód okolo kamier je popísaný na nasledujúcom obrázku 3.1.



Obr. 3.1: Návrh rozloženia LED diód okolo kamier. ■1550 nm, ■850 nm, ■740 nm

V prípade, že by zachytené snímky ruky neboli dostatočne osvetlené, je možné použiť biely box, aký sa používa pri fotení, alebo presunúť diódy bližšie k snímanej ruke. V takomto prípade by však pravdepodobne bolo potrebné zvoliť iný spôsob uchytenia, napríklad na konštrukciu.

3.2 Dataset

V tejto práci je dataset tvorený súborom RGB snímkov rúk. Tie vzniknú spojením grayscale obrázkov, vyhotovených pri osvetlení s konkrétnymi vlnovými dĺžkami. Potreba tvorby vlastného datasetu vyplýva z faktu, že detekcia živosti s využitím multispektrálnej analýzy ešte nie je veľmi rozšírená. Skúmala sa pri snímaní tváre alebo prstov, no nie celej ruky, a preto takýto dataset voľne dostupný neexistuje.

Snímaná bude vždy aspoň jedna ruka subjektu pre tri zvolené vlnové dĺžky. Takáto trojica snímkov však bude vždy vyhotovená aspoň trikrát pre jeden subjekt. Veľkou komplikáciou pri tvorbe datasetu je prítomnosť pandémie COVID-19 a nutnosť dodržiavania s tým spojených hygienických opatrení. Z toho dôvodu bude potrebné pristúpiť k augmentácii dát, a to napríklad v podobe horizontálneho preklopenia obrazu, čím vznikne zo snímku ľavej ruky snímok pravej a naopak. Takisto sa môže použiť augmentácia pomocou náhodného posunu po x-ovej a y-ovej osi, či augmentácia otočením. Zo zadania vyplýva, že bude potrebné nasnímať aspoň 100 ľudí, pričom aspoň u 10 z nich by mala byť koža zreteľne odlišne sfarbená, teda by mala obsahovať výrazne odlišné hodnoty melanínu.

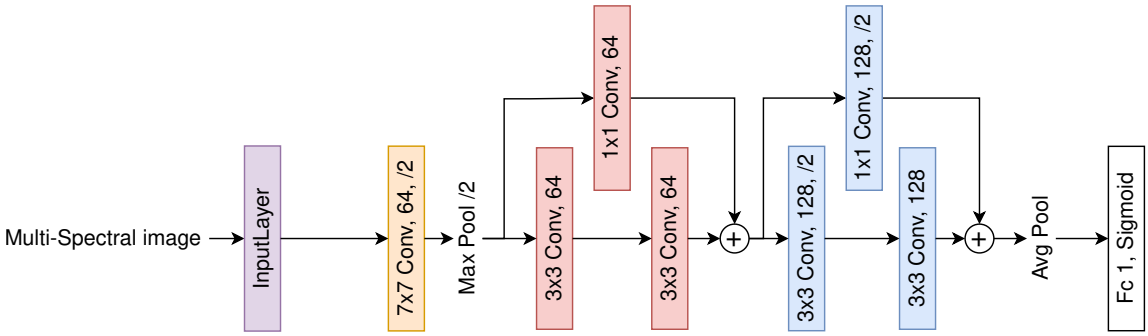
Výsledný dataset by mal byť rozdelený na 3 časti: trénovacie, validačné a testovacie dáta. Trénovacie spolu s validačnými slúžia pre natrénovanie váh neurónovej siete a testovacími sa overí výsledné riešenie. Vytvorený dataset však pravdepodobne nebude príliš veľký. V takýchto prípadoch sa na validáciu a testovanie často používajú tie isté dáta, a dataset je teda rozdelený len do 2 častí. Nevýhodou môže byť vo výsledku nižšia robustnosť riešenia.

3.3 Neurónová sieť

Pre odlíšenie živej ruky od tej falošnej je vhodné použiť konvolučnú neurónovú sieť, ktorá v praxi dosahuje veľmi dobré výsledky pri spracovaní obrazu. Vychádzam zo štúdie Marty Gomez-Barrero[11] a zvolil som štruktúru ResNet. Jedná sa o reziduálnu sieť. To značí, že využíva reziduálne, teda zbytkové učenie, a teda kombinuje informácie z minulých vrstiev.

Pri podrobnejšom preštudovaní nákresu modelu som odhalil chybu, ktorú som po konzultácii s autorkou odstránil. Výsledný návrh štruktúry tohto modelu môžeme vidieť na obrázku 3.2. Úprava spočívala v tom, že v konvolučnej vrstve znázornenej modrou farbou, s kernelom o veľkosti 1x1, sa musí použiť krok 2. Inak by v ďalšom kroku nemohlo prebehnúť sčítanie kvôli odlišným rozmerom matíc. Je nutné podotknúť, že chyba bola len v nákrese a nie vo výslednom riešení.

Ďalším rozdielom je spôsob vytvorenia RGB obrázka. Kým v citovanej štúdii sú vstupom do siete 4 snímky a RGB obrázok sa tvorí priamo v nej, v našom prípade bude tvorba RGB obrázka zahrnutá už do procesu snímania datasetu.



Obr. 3.2: Architektúra použitej reziduálnej kovolúčnej siete ResNet.

Pre schému 3.2 platí, že snímky sú zahrnuté v kanáloch vstupného multispektrálneho obrázka. Vrstva *Input* definuje rozmery vstupu. Pre zápis *AxB Conv, M, /N* platí, že *AxB* sú rozmery použitého kernelu, *M* je počet filtrov a *N* reprezentuje krok. Pred *Fc* vrstvou je umiestnená vrstva *Flatten* (konverzia na 1 riadok), ktorá sa nezvykne zakresľovať v prípade popisu architektúry.

Kapitola 4

Realizácia

Nasledujúce strany pojednávajú o spôsobe, akým bolo prevedené zariadenie a ako bol vo výsledku snímaný dataset. Táto kapitola je tiež venovaná implementácii neurónovej siete a na konci popisuje jednotlivé experimenty.

4.1 Zariadenie

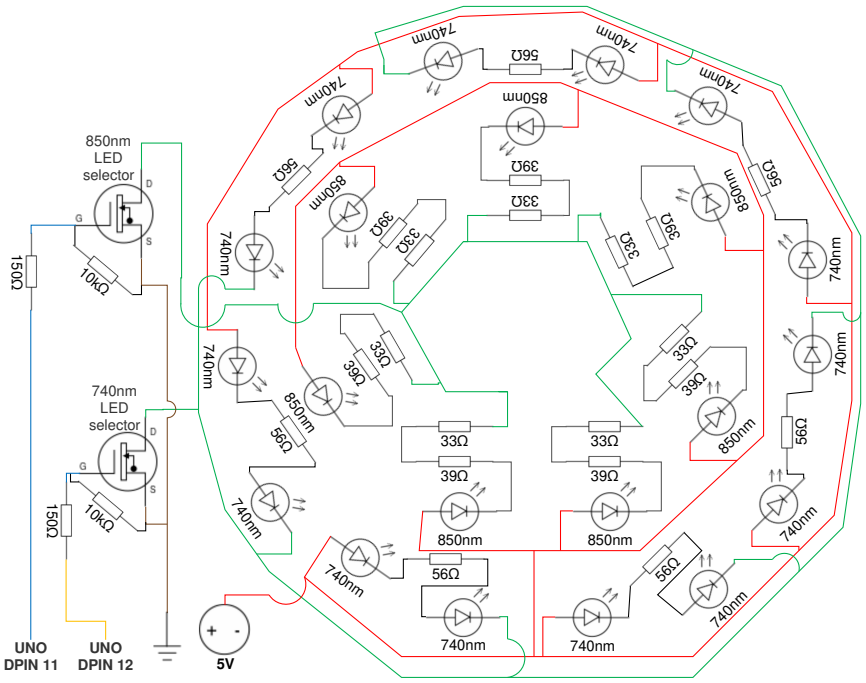
Ako prvé boli na SMD LED díody s označením OIS-330-740-X-T prispájkované vodiče. Inak by práca s nimi bola príliš komplikovaná. Takýmto spôsob vznikli diódy s predĺženou katódou a anódou, pričom s nimi bolo možné manipulovať ako s klasickými ledkami.

Ďalším krokom bolo otestovanie všetkých diód. Pri ledkách s vrcholovou vlnovou dĺžkou 850 nm bolo možné pozorovať jas pomocou kamery zabudovanej v mobile a pri 740 nm dokonca voľným okom. Čo sa týka diód EOLD-1550-525, tak pri nich nastala komplikácia. Mnou zvolená kamera (Digital CamIR 1550) nebola schopná zaznamenať ich vyžarovanie do okolia, a teda ledky boli na obrazovke viditeľné len pri priamom nasmerovaní kamery na nich, a len ako malé svetlé body. Dôvodov môže byť niekoľko. Mohlo to byť spôsobené nižšou citlivosťou kamery, ako aj nižším počtom a menšou svietivosťou použitých lediek. V datasheete, ktorý bol pribalený ku kamere, boli zároveň protichodné informácie, podľa ktorých je senzor uspôsobený pre vlnové dĺžky 1550-1600 nm, avšak scéna má byť nasvietená pri 1000-1185 nm. Z uvedeného dôvodu bolo nutné pokračovať bez využitia tejto kamery a vlnovej dĺžky 1550 nm.

Podľa návrhu rozloženia lediek 3.1 (varianta pre 740 a 850 nm) a s ohľadom na špecifikáciu zvoleného ovládacieho zariadenia (arduino Freaduino UNO 2017), bol navrhnutý elektrický obvod, ktorý odpovedá obrázku 4.1. Vo výsledku sú diódy, rezistory a tranzistory na jednej strane tenkej dosky (obr. 4.2), prepojené vodičmi na druhej strane. Táto doska tvorí vrchnú stenu zásuvného modulu, ktorý sa vloží do kartónovej krabice, kde budú snímané ruky. Do modulu sa ešte umiestni kamera, pripojená pevne skrutkou k podstavci. Pre lepšiu predstavu je pripravená fotodokumentácia v prílohe B.

K bočnej stene modulu je upevnená doska arduina, ovládajúca zapínanie LED diód pomocou MOSFET tranzistorov. Tie sú pripojené k dátovým pinom 11 a 12 na doske arduina. Zdrojom je 5V zdrojový pin a uzemnenie je pripojené na GND pin. V strede kružnice z lediek sa nachádza objektív s kamerou.

Pri pohľade na zapojenie v 4.1 si môžeme všimnúť, že medzi elektródami G (brána) a pinom GND sa nachádza 10kΩ pull-down rezistor. Ten bol použitý, aby sa zamedzilo preblikávaniu diód, napr. pri zapnutí či vypnutí arduina.



Obr. 4.1: Výsledné zapojenie jednotlivých LED diód a tranzistorov, spolu s odpormi.

Obr. 4.2: Doska s ledkami a objektívom. Ide o vrchnú časť modulu, ktorý sa zasúva do kartónovej krabice.

Pulzne šírková modulácia PWM

V práci bola otestovaná pulzne šírkovú modulácia. Ide o diskrétnu moduláciu signálu zmenou striedy, teda šírky impulzu. Signál tak nadobúda len hodnoty logickej 1, alebo 0. Podľa [36] je totiž možné s využitím PWM zvýšiť svietivosť lediek, a to vďaka využitiu vyššej maximálnej hodnoty prúdu (použijú sa rezistory s menším odporom), v závislosti od špecifikácie danej ledky.

Pri diódach s označením OLD-1550-525, bolo najprv potrebné upraviť frekvenciu na pine arduina, a to na aspoň 20kHz, čomu najbližšie zodpovedalo 31372.55 Hz. Podľa datasheetu bol pred diódu zvolený vhodný rezistor, tak aby cez ňu pretekal prúd 200mA (dvojnásobok

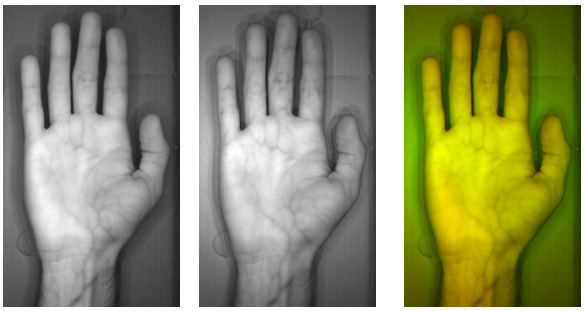
oproti realizácii bez PWM). Následne sa ešte upravil PWM pomer pinu na 1/2. Ani to však nepomohlo a zábery s kamerou CamIR pri 1550 nm neboli použiteľné.

U lediek s označením OIS-330-740-X-T sa dokonca pri použití PWM znížila svietivosť, aj napriek tomu, že bolo možné použiť 5-násobne vyššiu hodnotu prúdu (150mA). Nakoniec bola PWM predsa len využitá pri ledkách 15400585A3590, no pri štandardnej hodnote prúdu 100mA, ktorý vyžadovala LED. Dôvodom k tejto úprave bolo to, že svietivosť týchto lediek bola omnoho vyššia ako tých s vlnovou dĺžkou 740 nm. Clona objektívu sa preto upravila pre snímanie so 740 nm diódami a jas 850 nm diód bol znižovaný pomocou PWM až na prijateľnú hodnotu. Výsledkom bol pomer svietenia 110:255, ktorý bol v tomto prípade pre arduino nastavený volaním funkcie *analogWrite(11,110)*, kde 11 reprezentuje číslo digitálneho pinu, ktorým bolo ovládané spínanie tranzistora pre vlnovú dĺžku 850 nm.

4.2 Snímanie datasetu

Na začiatku, ešte pred trénovaním neurónovej siete, bolo potrebné vytvoriť dataset a vymyslieť čo najefektívnejší spôsob jeho nasnímania. Pre tento účel bola navrhnutá a implementovaná konzolová aplikácia `capture_hand` v jazyku C++. Tá je schopná pomocou knižníc *FlyCapture2* a *uEye* komunikovať s oboma kamerami z návrhu. V prípade, že je aplikácia spustená s prepínačom `-d` (default), použije sa prednastavený cieľ pre ukladanie snímkov. V opačnom prípade je nutné argumentom definovať cestu. V danom adresári sa vytvoria podadresáre pre snímky živých rúk a falzifikátov (ak ešte neexistujú), pričom kód automaticky pomenúva ukladané súbory. LED diódy sú ovládané pomocou signálov posielaných prostredníctvom sériového portu, ktoré spracuje program `LedController.ino` v arduine a zopne príslušný tranzistor pre rozsvietenie lediek. Programovo je zároveň ošetrené, aby mohol svietiť vždy len 1 druh lediek, inak by to zdroj arduina nemusel zvládnuť.

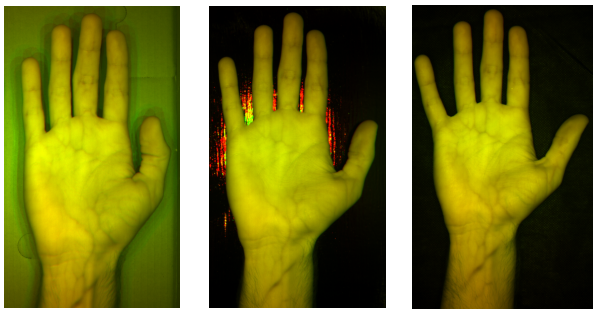
Z dôvodov už spomenutých v predchádzajúcej kapitole sa pristúpilo k snímaniu len s jednou kamerou. Konzolová aplikácia teda zabezpečí rozsvietenie 740 nm LED diód, vyhotovenie fotografie a rovnaké zachytenie obrazu ruky pri 850 nm. Oba snímky sú spojené do RGB obrázka (viď obr. 4.3), ktorý slúži pre trénovanie neurónovej siete. 740 nm snímok tvorí R-kanál a 850 nm je G-kanál. Kanál B tvorí nulová matica s odpovedajúcimi rozmermi. Je potrebné podotknúť, že aplikácia, ako aj kód arduina, sú pripravené aj na snímanie s druhou kamerou pri 1550 nm. Tu by však bolo potrebné doplniť kód o spracovanie tohto snímku a jeho použitie vo výsledku ako B-kanál.



Obr. 4.3: Spojenie 740 nm (vľavo) a 850 nm (v strede) snímku do jedného RGB obrázka (napravo).

Pri spojení snímkov do RGB obrázka sa vyskytol nežiaduci efekt. Za rukou sa vytváral viacnásobný tieň, spôsobený nedokonalým rozptylom svetla. Otestovalo sa preto natretie

plochy za rukou na čierno, čo však zapríčinilo odraz svetla od povrchu do kamery. Ako vhodné sa ukázalo až využitie čiernej netkanej textílie, ktorou sa pokryla stena za rukou. Porovnanie jednotlivých verzií je možné vidieť na obrázku 4.4.



Obr. 4.4: Porovnanie úprav povrchu steny krabice za snímanou rukou. Smerom zľava: pôvodný kartónový povrch, čierny náter a čierna netkaná textília.

Snímky z 4.3 a 4.4 slúžili len pre ilustráciu a boli orezané. Na obrázku 4.5 je však možné vidieť vzorový snímok, ktorý bol vytvorený programom `capture_hand`. To, ako ďaleko, či pod akým uhlom bola ruka vložená do zariadenia bolo ponechané na účastníkoch zberu datasetu. Podmienkou bolo len to, aby bolo na snímku vidieť aj zápästie, a aby bola zachovaná vzdialenosť od kamery (kvôli prednastavenej clone a zaostreniu objektívu).

Obr. 4.5: Vzorový snímok, ktorý je súčasťou datasetu.

Celkovo sa vytvoril súbor 381 RGB snímkov živých rúk, ktoré patria 114 rôznym subjektom. Desiati z nich boli cudzinci z 9 rôznych krajín. Zopár ľudí (už však nie cudzincov), teda bolo nasnímaných nad rámec bodu 3 zadania. Zastúpené boli vekové kategórie od 17 rokov až po dôchodkový vek, no najpočetnejšia bola skupina tínedžerov. Snímaná bola zväčša pravá ruka, u niektorých však ľavá, či prípadne obe (v závislosti od faktorov ako bola prítomnosť tetovania, hodiniek či prsteňov, zranenia, alebo časová vyťaženosť).

Najväčším problémom bola prebiehajúca pandémia COVID-19. Sfarbenie RGB snímkov rúk mohlo byť teoreticky ovplyvnené aplikovaním dezinfekčných prostriedkov. Náročné tiež bolo nasnímať ľudí s výrazne odlišnými hodnotami melanínu (so zariadením sa bolo potrebné presúvať kvôli každému subjektu). Zároveň nebolo možné úplne naplniť tretí bod zadania, kedy neboli zaznamenávané hodnoty melanínu. Prístroj na tento účel si totiž nebolo možné zapožičať, nachádzal sa na inej univerzite, a preto bola aspoň zaznamenávaná národnosť osôb.. Dáta boli zberané anonymne, a tak nie je možné spojiť jednotlivých účastníkov s ich snímkami.

Vyhotovených bolo aj 218 snímkov 25 rôznych typov falzifikátov. Umelé ruky boli vyrobené po zápästie, alebo aj s predlaktím, no použili sa aj syntetické násady prstov na živej

ruke. Z materiálov stojí za zmienku plast, silikón, guma, papier, živica, plastelína, sadra, či latex. Použili sa aj už hotové napodobeniny vyrobené na fakulte, pri ktorých nebol uvedený materiál. Niektoré látky boli povrchovo nanášané na ruky vyrobené 3d tlačiarňou.

Testovací (20% snímkov) a trénovací dataset (80%) boli reprezentované *NumPy* poľami, do ktorých boli postupne vložené jednotlivé obrázky a ich identifikátory (1 pre živú ruku, 0 pre falzifikát). Falzifikáty boli v datasetoch zastúpené v rovnakom pomere, pričom testovací dataset sa využil aj na validáciu. Obrázky boli zmenšované na 409x217px, čo predstavuje 20% z pôvodných 2048x1088px. Využitá bola aj augmentácia horizontálnym preklopením a rotáciou s rozdielom maximálne 20 stupňov, kedy však bola pamäť RAM počas trénovania neurónovej siete naplno vyťažená. Stavba datasetu je popísaná v nasledujúcej tabuľke 4.1.

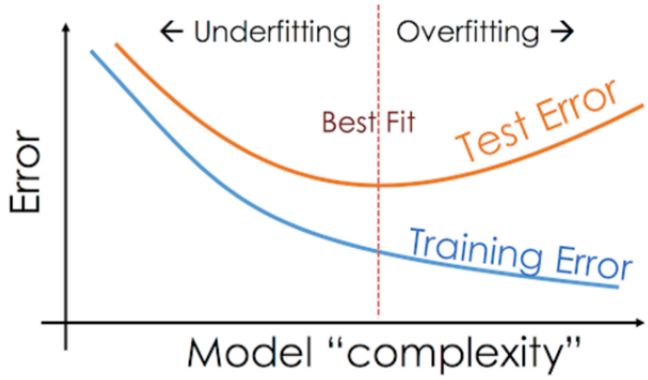
	Bez augmentácie			**S augmentáciou**		
	Počet snímkov	**BF snímky**	**PA snímky**	**Počet snímkov**	**BF snímky**	**PA snímky**
Trénovací dataset	480	305	175	1918	1220	698
Testovací dataset (= validačný)	119	76	43	478	304	174
Dataset celkovo	599	381	218	2396	1524	872

Tabuľka 4.1: Štruktúra trénovacieho a testovacieho datasetu, ktorý bol použitý aj pre validáciu. Výraz BF (bona fide) bol použitý pre snímky živých rúk a PA (presentation attack) pre falzifikáty.

4.3 Trénovanie CNN

Trénovanie neurónovej siete, ako aj všetky ďalšie pomocné skripty boli realizované v prostredí Colaboratory ("Colab") od firmy Google. Prístup síce nebol spoplatnený, no zdroje boli obmedzené. Konvolučná neurónová sieť bola implementovaná podľa návrhu v podkapitole 3.3, využívajúc knižnice *TensorFlow* a *Keras*. Model bol vytvorený s optimalizátorom Adam a mierou učenia 0.0001, tak ako v práci [11] od Gomez-Barrero a Busch. Pri jednotlivých snímkoch boli ešte pred trénovaním normalizované pixely (hodnota pixelu vydelená 255). Skript bežal na grafickej karte NVIDIA Tesla T4 GPU.

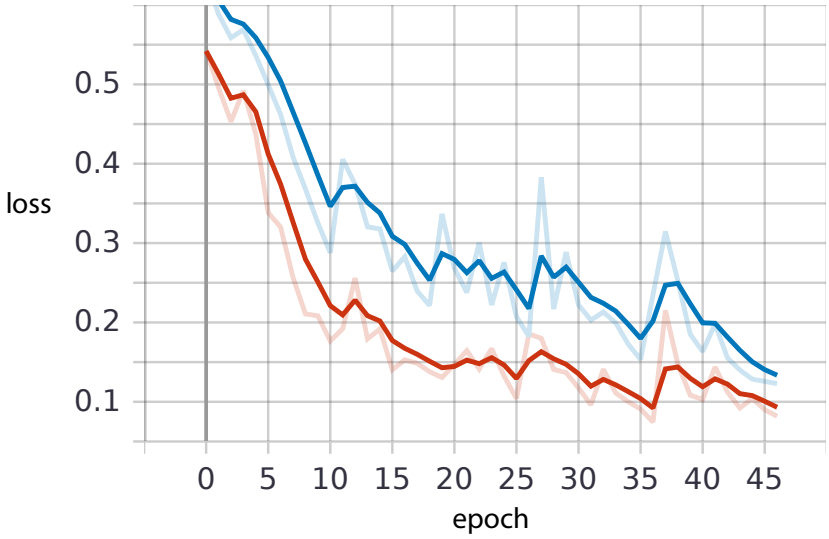
Pri trénovaní môže nastať nežiaduci efekt, a to pretrénovanie, čiže overfitting. Výsledkom je to, že kým miera chybovosti na trénovacích snímkoch stále klesá a sieť sa ich učí naspamäť, validačná miera chybovosti (v tomto prípade chybovosť na testovacích dátach) začne od určitej epochy stúpať, pretože neurónová sieť už nevie dobre generalizovať. Aby sa tento efekt eliminoval, bola využitá technika včasného ukončenia (*early stopping*) [5]. Keďže by sa však mohlo stať, že chybovosť by v jednej epoche vzrástla, no v ďalších by mohla klesnúť ešte nižšie, využila sa aj vlastnosť *patience* s nastavením pre 10 epoch. Znamená to, že trénovanie sa ukončí až vtedy, keď sa chybovosť nezníži v 10 ďalších epochách. Zároveň sa však neurónová sieť musí trénovať dostatočne dlho, aby nebola podtrénovaná (*underfitting*), preto bol maximálny počet epoch stanovený na 100. Pre zapamätanie si natrénovaných váh bol využitý callback *ModelCheckpoint*, kde sa parametrom stanoví uloženie len najlepšieho výsledku. Problematika je vyobrazená na grafe 4.6.



Obr. 4.6: Priebeh chybovosti počas trénovania. Cieľom je nájsť bod (konkrétnu epochu učenia), kedy je testovacia, respektíve validačná miera chybovosti najnižšia. Zdroj: [21]

Základná konfigurácia

Ako prvé prebehlo trénovanie modelu popísaného v 3.3 bez augmentácie dát a bez začlenenia subjektov s výrazne odlišnými hodnotami melanínu v datasete. Výsledok trénovania je možné vidieť na nasledujúcom obrázku 4.7. Na y-ovej osi je znázornená miera chybovosti, ktorá je závislá od zvolenej chybovej funkcie (*loss function*, v tejto práci ako *binary_crossentropy* v *Kerase*).



Obr. 4.7: Priebeh miery chybovosti. Červená znázorňuje validačnú chybovosť a modrá trénovaciu, pričom tmavšou je exponenciálny kĺzavý priemer a svetlejšou presné hodnoty. Graf bol vytvorený vizualizačným nástrojom TensorBoard. Ide o sieť Resnet z 3.2

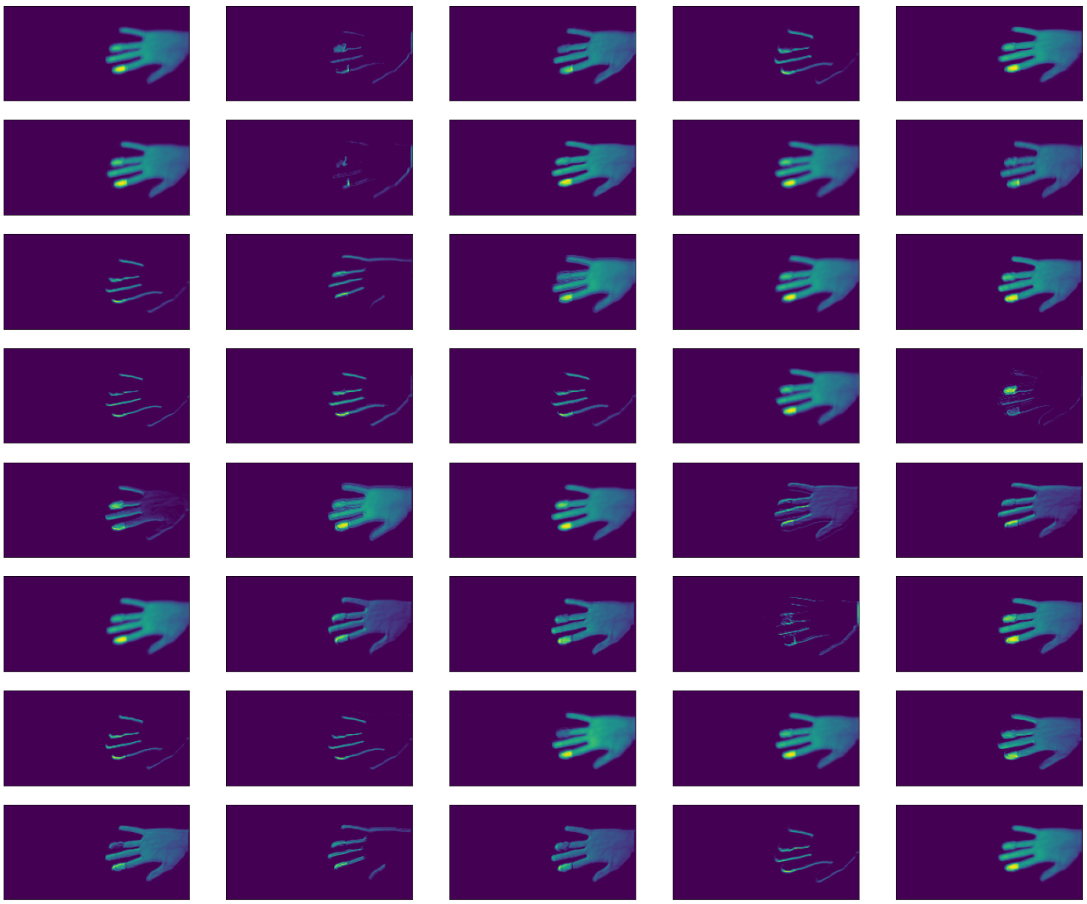
Pri určovaní úspešnosti jednotlivých konfigurácií boli vyhodnocované:

- **Attack Presentation Classification Error Rate (APCER)** - percento falzifikátov (AP snímky) vyhodnotených nesprávne ako živá ruka
- **Bona Fide Presentation Classification Error Rate (BPCER)** - percento živých rúk (BF snímky) vyhodnotených nesprávne ako falzifikát

Pri falzifikátoch bola dosiahnutá 95.35% úspešnosť a pri detekcii živých rúk bola na úrovni 97.06%. Inak povedané chybovosť na falzifikátoch, čiže Attack Presentation Classification Error Rate (APCER) bol 4.65% a na živých rukách, Bona Fide Presentation Classification Error Rate (BPCER), bol 2.94%. Percento správne určených živých rúk aj falzifikátov bolo na úrovni 96.4%. Treba však podotknúť, že ak by bol testovací a validačný dataset odlišný, úspešnosť by bola pravdepodobne nižšia. Zároveň platí, že pred trénovaním sú váhy nastavené náhodne a pri opakovaných pokusoch o natrénovanie kolísala úroveň úspešnosti v jednotkách percent.

Vizualizácia

Pomocou mapy vlastností (*feature map*) je možné vizualizovať, oblasti, na ktoré sa neurónová sieť pri určovaní konkrétneho snímku zamerala. Takýmto spôsobom sa dalo overiť, či sa neurónová sieť netrénovala na nepodstatných elementoch snímku. Na obrázku 4.8 je možné vidieť výsledok pre živú ruku s dvoma syntetickými odtlačkami prstov (prezentačný útok). Mapa bola vygenerovaná z prvej konvolučnej vrstvy, pretože z ostatných by bola menej detailnejšia (snímok sa prechodom vrstvami zjednodušuje). Miesta, na ktoré sa CNN zamerala, čiže ruka a ešte jasnejšie prsty s falzifikátmi, sú zobrazené žiarivejšou farbou.



Obr. 4.8: Mapa vlastností (*feature map*) z prvej konvolučnej vrstvy v poradí. Použitý bol snímok živej ruky s 2 odtlačkami vyrobenými z 2 tvarovacích hmôt s odlišnou mierou odrazivosti v infra svetle. Zobrazených je 40 z celkových 64 snímkov, ktoré tvoria mapu.

4.4 Experimenty

V predchádzajúcej kapitole boli okrem iného popísané základné parametre, s ktorými bola neurónová sieť trénovaná. Keďže však bola snaha o čo najlepší výsledok, pristúpilo sa k viacerým experimentom, ktorých cieľom bolo nájsť optimálne riešenie. Tieto experimenty sú postupne popísané na nasledujúcich stranách.

4.4.1 Využitie augmentácie

Ako už bolo spomenuté v podkapitole 4.2, pristúpilo sa ku augmentácii horizontálnym preklopením a rotáciou (-10°až 10°). Pri rotácii však nastal problém, kedy bolo navyše potrebné aplikovať posun, aby sa skryl výrez, ktorý by vznikol na fotografiách v mieste predlaktia.

Po natrénovaní neurónovej siete bola dosiahnutá celková úspešnosť na úrovni 92.54%, čo je mierne zhoršenie oproti verzii bez augmentácie. Pokles APCER na 0.77% sa dá vysvetliť väčším množstvom dát. Pri živých rukách však štvornásobne vzrástol BPCER, až na 11.81%. Znázornenie mapy vlastností naznačuje, že to mohlo byť spôsobené vzorom textílie na pozadí za rukou (viď. obr. 4.9). Tento vzor bol síce nepatrný, no predsa len ovplyvnil samotné učenie siete. Bez augmentácie totiž bolo pozadie z pohľadu neurónovej siete na všetkých snímkoch zhodné, no pri použití náhodnej rotácie už nie.

Obr. 4.9: Detail z feature mapy - rozžiarené je aj pozadie a teda sa CNN trénovala aj na jeho vzore. Zároveň je vpravo dole viditeľná línia po augmentácii otočením.

Odstránenie šumu a prahovanie

Aby sa problém popísaný vyššie eliminoval, boli otestované rôzne techniky z knižnice *OpenCV* a ich kombinácie. Medzi nimi bilaterálny, mediánový, či Gaussov filter, taktiež 2D konvolúcia a rôzne metódy na odstránenie šumu. Ako najefektívnejšie sa ukázalo využitie funkcie *fastNlMeansDenoising()* a následná aplikácia prahovania (thresholdu). Hranica pre funkciu prahovania bola stanovená experimentálne.

Po zapracovaní tejto úpravy bola zaznamenaná hodnota APCER na úrovni 3.84% a BPCER už len 2.93%. To značí 96.15% úspešnosť na falzifikátoch a 97.07% na živých rukách. Celková percentuálna úspešnosť sa zlepšila o približne 3% na výsledných 96.72%. Na všetkých mapách vlastností už pozadie nežiarilo, teda sa naň CNN nezameriavala.

Výsledok

Zo zistených informácii vyplýva, že augmentácia môže zlepšiť kvalitu celkového riešenia, teda schopnosť detekovať falzifikáty a rozpoznávať živé ruky. Je však potrebné zabezpečiť, aby sa neurónová sieť netrénovala na artefaktoch vznikajúcich na pozadí. To je možné dosiahnuť jednofarebným pozadím bez štruktúry na povrchu (zároveň nesmie byť lesklý), alebo odstránením vzorky zo snímku, napríklad ako v tomto prípade, pomocou prahovania a odstránenia šumu.

4.4.2 Zaradenie subjektov s odlišnými hodnotami melanínu

V tomto experimente boli do procesu zaradené už aj snímky ľudí z výrazne odlišným sfarbením pokožky. Ako prvé bolo otestované využitie týchto snímkov len v testovacom (validačnom) datasete, s cieľom overiť, aký vplyv bude mať rozdielna pigmentácia na úspešnosť siete. V ďalšom experimente boli tieto snímky zaradené aj do procesu trénovania, pričom k zvýšeniu úspešnosti prispelo použitie augmentácie.

Zaradenie do testovacieho datasetu

Ukázal sa výrazný pokles úspešnosti, najmä u BF snímkoch, a to až na 78.95%, čo predstavuje až 21.05% BPCER. Pri manuálnej kontrole bolo zistené, že problematickými boli snímky s rukami tmavšieho pigmentu. 740 nm snímok sa totiž použil ako R-kanál a 850 nm ako G-kanál, a teda tmavšia pokožka je na RGB obrázku sfarbená viac dozelena (najmä línie na dlani). Toto zistenie odpovedá grafu na obrázku 2.3 v kapitole 2.4. Čo sa týka úspešnosti na falzifikátoch, tá bola 83.72% (16.28% APCER). Zvýšenie tejto hodnoty je spôsobené tým, že na rozdiel od predchádzajúcich experimentov, sa neurónová sieť prestala učiť už v 5. epoche a chybovosť začala prudko rásť.

Zaradenie aj do procesu trénovania

Všetky zozbierané snímky boli náhodne premiešané a rozdelené medzi trénovacie a testovacie dáta (overilo sa, že medzi testovacími boli aj snímky s rozličnými hodnotami melanínu). Otestovaná bola verzia bez a s využitím augmentácie. Nasledujúce výsledky boli dosiahnuté s augmentáciou, pričom sa tak úspešnosť zvýšila ešte o ďalších niekoľko percent.

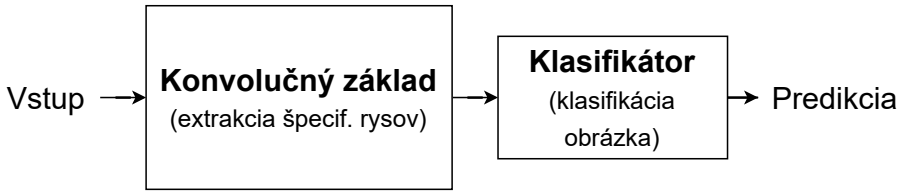
Snímky s tmavšou pokožkou už tentokrát neboli problémom. BF úspešnosť bola na úrovni 95.30%, čo predstavuje 4.69% BPCER, pri APCER na úrovni 7.94% (úspešnosť 92.06%). Aj u živých rúk a aj u falzifikátov teda došlo k značnému zlepšeniu. Problémy boli zaznamenané opäť pri plastoch, odlíšení chirurgickej rukavice a pri silikónovej rukavici.

Výsledok

Z uvedeného vyplýva, že pri súčasnom stave zariadenia je nutné, aby bola neurónová sieť trénovaná aj na snímkoch rúk s inými hodnotami melanínu. V prípade, že by bolo takéto zariadenie využívané len občanmi jednej krajiny, tak by nemuseli nastať komplikácie, keďže by šlo o verziu z 4.3 a schopnosť rozlišovať falzifikáty by bola relatívne vysoká. Za zváženie tiež stojí využitie ďalšej kamery v oblasti 1550 nm, ktorej snímky by dopomohli k väčšej úspešnosti. Jednak pri rôznych hodnotách melanínu (v tejto oblasti totiž neovplyvní odrazivosť pokožky), ale aj pri rozpoznávaní falzifikátov.

4.4.3 Porovnanie s inými modelmi

V prípade klasifikácie obrázkov platí, že model založený na CNN sa skladá z konvolučného základu (inak báza, alebo extraktor) a z klasifikátora. Konvolučná báza slúži pre extrakciu charakteristických rysov zo snímkov a klasifikátor zabezpečí zatriedenie snímku do určitej skupiny, viď obrázok 4.10.



Obr. 4.10: Zjednodušená stavba modelu založeného na CNN.

V tomto experimente boli porovnávané výsledky ResNetu z predchádzajúceho experimentu (prípad, v ktorom boli snímky cudzincov zaradené aj do trénovania) s ďalšími vytvorenými modelmi. Pri ich tvorbe sa využila metóda *transfer learning*, teda využitie znalostí z už predtrénovanej siete. Ako konvolučná báza sa teda použil základ z takejto už existujúcej predtrénovanej siete, pričom klasifikátor (koncová časť siete) sa vytvoril nový, tak aby vo výsledku sieť rozlišovala živé ruky a falzifikáty. Váhy klasifikátora sa v procese učenia trénovali, avšak trénovanie už natrénovaných váh v konvolučnej báze bolo pozastavené [25].

Použité boli predtrénované modely z knižnice *Keras*, konkrétne VGG16, VGG19, MobileNetV2, DenseNet169 a InceptionResnetV2, pričom boli načítané s parametrom *include_top=False* (bez klasifikátora). Vo všetkých prípadoch sa ako klasifikátor použila vrstva *Flatten* nasledovaná 2 *Dense* vrstvami (alebo *Fully connected*, *FC*). Pri prvej z nich bol stanovený počet filtrov 256 a aktivačná funkcia *ReLU*. Posledná vrsta mala len 1 filter a aktivačnú funkciu *sigmoid*, tak ako v pôvodnom modeli 3.3. Pri tvorbe klasifikátora bola inšpiráciou práca od Gomez-Barrero a Busch [11]. Výsledné architektúry v prípadoch VGG16 a VGG19, sú zobrazené v prílohe C. Ostatné modely síce boli príliš rozsiahle pre grafické znázornenie, avšak princíp je rovnaký.

	APCER (%)	**BPCER (%)**	**Celková úspešnosť (%)**
ResNet	*7.94*	*4.69*	*94.10*
Základ z **VGG16**	*2.38*	*1.88*	*97.94*
Základ z **VGG19**	*4.76*	*2.35*	*96.76*
Základ z **MobileNetV2**	**1.59**	**0.47**	**99.12**
Základ z **DenseNet169**	*3.17*	*0.94*	*98.23*
Základ z **InceptionResnetV2**	*4.76*	*1.88*	*97.05*

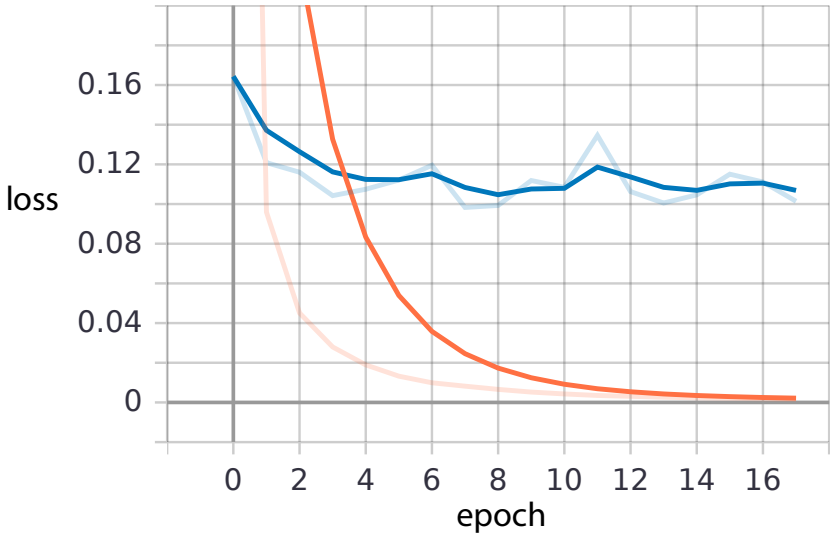
Tabuľka 4.2: Porovnanie výsledkov jednotlivých modelov. Šlo o prípad so snímkami pokožky s výrazne odlišným sfarbením aj v trénovacom datasete, pri použití augmentácie. ResNet bola trénovaná od nuly, zatiaľ čo ostatné modely mali základ v už natrénovaných sieťach (Transfer learning) a počas učenia bolo povolené trénovanie len pre pridané vrstvy pre klasifikáciu (klasifikátor).

V tabuľke 4.2 je možné vidieť porovnanie chybovosti a celkovej úspešnosti jednotlivých modelov. Zo všetkých najlepšie obstál model založený na sieti MobileNetV2. Po natrénovaní klasifikátora dosiahol model APCER len na úrovni 1.59%, čo je o takmer 8% lepší výsledok ako s pôvodnou reziduálnou CNN, BPCER len na úrovni 0.47%. V oboch kategóriach tak šlo o najnižšiu chybovosť.

Snímky s odlišným sfarbením pokožky len v testovacom datasete

Na základe výsledkov vyššie rozobratého porovnania vyšiel model s extraktorom z CNN MobileNetV2 ako najlepší. Naskytla sa preto otázka, ako veľmi by bol úspešný v prípade, že by sa snímky subjektov s odlišnými hodnotami melanínu nachádzali len vo validačnom, respektíve testovacom datasete (zopakovanie prvej časti experimentu 4.4.2).

Po natrénovaní siete bol viditeľný výrazný pokles chybovosti oproti reziduálnej CNN ResNet. Navyše, bod, v ktorom bola validačná chybovosť najnižšia, a teda sieť bola natrénovaná, bol dosiahnutý za výrazne kratší čas, ako tomu bolo pri trénovaní Resnetu, a to už v 7. epoche (obr. 4.11).



Obr. 4.11: Priebeh miery chybovosti počas trénovania po jednotlivých epochách. Oranžovou je znázornená validačná chybovosť a modrou trénovacia, pričom tmavšou je exponenciálny kĺzavý priemer a svetlejšou presné hodnoty. Graf bol vygenerovaný prostredníctvom vizualizačného nástroja TensorBoard. Ide o model založený na MobileNetV2.

V tabuľke 4.3 je možné vidieť porovnanie s výsledkami z experimentu 4.4.2. Vo výsledku je model pomerne úspešný aj napriek tomu, že sieť sa nestretla s výrazne iným sfarbením kože v procese trénovania. Na znížení BPCER (5.16%) by ešte bolo potrebné popracovať, ale APCER sa drží len na 0.79%.

Výsledok

Týmto experimentom bolo dokázané, že využitie techniky transfer learning, môže priniesť lepšie výsledky ako pôvodná sieť. V prípade tejto práce to môže byť spôsobené najmä malým datasetom, na ktorom sa trénovala pôvodná reziduálna sieť, ktorá následne nevedela extra-

	APCER (%)	BPCER (%)	Celková úspešnosť (%)
ResNet	*16.27*	*21.05*	*80.67*
Základ z **MobileNetV2**	**0.79**	**5.16**	**96.46**

Tabuľka 4.3: Porovnanie výsledkov Resnetu a modelu založenom na predtrénovanej sieti MobileNetV2. Šlo o prípad s pokožkou s výrazne odlišným sfarbením len vo validačnom (testovacom) datasete, pri použití augmentácie.

hovať rysy z obrázka tak dobre, ako predtrénované modely. Tie mali totiž váhy nastavené podľa trénovania na datasete *imagenet*, ktorý obsahuje viac než 14 miliónov obrázkov.

4.5 Zhodnotenie stavu a návrhy vylepšení

Niektoré z vylepšení už boli načrtnuté v práci, no v tejto podkapitole budú rozobraté podrobnejšie spolu s ďalšími. Popísané budú nedostatky a postupy, ako ich eliminovať.

Zariadenie

V prvom rade je potrebné povedať, že zostrojené zariadenie bolo prototypom. Veľkou výhodou bolo použitie Freaduina UNO 2017, a to kvôli štvornásobnej hodnote maximálneho prúdu oproti bežne dostupným verziám arduina. Čo sa týka spájkovaných spojov, tak podľa správnosti by ešte mali byť dodatočne zaizolované, aby nenastal prípadný skrat. V budúcnosti by mal byť box vyrobený z pevnejších, no stále ľahkých materiálov, prípadne v teleskopickej verzii, kedy by ho pri transporte bolo možné zložiť na menšiu veľkosť.

Pri problematike rozličného sfarbenia pokožky je potrebné podotknúť, že síce bola s pomocou MobileNetV2 (technika transfer learning) dosiahnutá BPCER len na úrovni 5.16%, no väčšina chybne rozlíšených snímkov bola práve s najtmavšou pokožkou. Tento problém môže byť eliminovaný zaradením takýchto snímkov do trénovacieho datasetu, kedy bola BPCER len na úrovni 0.47%. Celkovo by však bolo potrebné zapojiť viacero takýchto subjektov do zberu datasetu a problém ako taký by sa neodstránil. Lepšie by bolo riešenie, kedy by bola do zariadenia upevnená ešte jedna kamera pre vlnovú dĺžku 1550 nm aj s príslušnými ledkami, prípadne by sa kamera IDS nahradila inou, ktorá by bola dostatočne citlivá v celom rozsahu od 740 nm, až po 1550 nm. Netreba ale zabudnúť, že vo vyšších vlnových dĺžkach majú diódy nižšiu svietivosť, a teda je ich potrebné väčšie množstvo. Napríklad Steiner vo svojej práci [36] použil 30, po revízii až 80ks EOLD-1550-525.

Za úvahu tiež stojí zmena usporiadania LED diód. Ukázalo sa totiž, že diódy pre 850 nm mali pomerne malý rozptyl. Vo výsledku sú teda snímky v strede sfarbené viac dozelena, zatiaľ čo v okrajových oblastiach viac dočervena, kvôli spájaniu kanálov do RGB obrázka (viď obr. 4.5 v podkapitole 4.2).

Čo sa týka steny zariadenia za rukou, tak jej povrch je potrebné ošetriť lepším spôsobom, tak aby neobsahovala vzory, a teda sa mohol použiť jemnejší filter a threshold. Tým nestratíme niektoré detaily ruky na snímkoch. Niekedy sa stávalo, že sa kamera v zariadení mierne posunula a na kraji snímku bol viditeľný kartón bez textílie, čo malo negatívny efekt pri augmentácii otočením. Preto je nutné upevniť kameru lepším spôsobom.

Proces snímania bol po softwarovej stránke bezproblémový a pri presune bola príprava zariadenia aj s pripojením k notebooku veľmi rýchla. Dal by sa však vyladiť priebeh vyhotovenia multispektrálneho obrázka. V prípade, že subjekt pohol rukou, kanály spojeného

obrázka boli posunuté. Obrázok tak niekedy vyzeral ako 3d anaglyf a snímanie bolo potrebné zopakovať. Tento problém by sa dal odstrániť softwarovo, alebo aj mechanicky, kedy by do boxu boli nainštalované zarážky, ktoré by presne nasmerovali ruku do požadovanej polohy. Navyše by sa tak CNN zameriavala viac na odrazivosť ako na tvar, pričom by tiež bolo zamedzené použitiu niektorých falzifikátov, keďže útočník by musel poznať požadovanú polohu prstov.

Dataset

Na dosiahnutie lepších výsledkov je potrebné nasnímať výrazne rozsiahlejší dataset, najmä s väčším počtom snímkov rúk s výrazne odlišnými hodnotami melanínu (v desiatkách až stovkách, miesto jednotiek a pre rôzne vekové skupiny). Spolu s následným rozdelením datasetu až na 3 časti, kedy testovací a validačný nebudú totožnými, bude výsledok hodnovernejší a zabezpečí vyššiu robustnosť. V práci [11], ktorá bola zameraná na detekciu živosti na prstoch, sa pre porovnanie do snímania zapojilo až 562 subjektov a použili asi dvojnásobok snímkov falzifikátov (443).

Takisto by bolo vhodné rozšíriť augmentáciu, napríklad pomocou neurónovej siete, ktorá by zabezpečila ďalšie dáta navyše. Prínosná by bola optimalizácia procesu trénovania, a to kvôli vysokým požiadavkám na operačnú pamäť RAM. S tým je tiež spojená náhrada metódy *fit()* za *fit_generator()*, pre ktorú by bolo potrebné vytvoriť generátor, ktorý by model napĺňal dátami postupne, miesto predania celého poľa naraz.

Fine-tuning

Pri fine-tuningu je postup podobný ako pri technike transfer learning, ktorá bola použitá v experimente s porovnávaním modelov 4.4.3. Líši sa len tým, že nie všetky vrstvy predtrénovaných modelov ostanú zmrazené (neprebieha trénovanie), ale na niektorých, učenie prebieha. Teraz sa totiž menili len váhy klasifikátora. Týmto postupom je niekedy možné dosiahnuť ešte lepšie výsledky, ako pri metóde transfer learning, najmä ak je použitý dataset výrazne odlišný od toho, na ktorom boli siete predtrénované. Sú prípady, kedy pomôže aj pretrénovanie celej siete, avšak s počiatočnými váhami nastavenými z predtrénovania a pri použití nízkej miery učenia (hodnota learning rate) [30].

Kapitola 5

Záver

Táto bakalárska práca bola venovaná tvorbe zariadenia na snímanie datasetu, spolu s implementáciou systému na detekciu falzifikátov v rámci technológie snímania ľudskej ruky v infra svetle. Rozoberá rôzne už zaužívané postupy detekcie živosti, ako aj dôvody, prečo sú na rozdiel od tohto systému v súčasnosti nevyhovujúce. Výsledný systém je prvý svojho druhu a v budúcnosti môže prispieť k väčšej bezpečnosti v oblasti biometrie.

Pri návrhu zariadenia sa vychádzalo z už existujúcich poznatkov o infračervenom svetle a jeho odrazivosti od kože. Samotná konštrukcia, ako aj zapojenie a proces snímania boli výsledkom viacerých pokusov. Cieľom bola spoľahlivosť, ako aj efektívnosť riešenia.

Podarilo sa nasnímať 381 snímkov živých rúk a ďalších 218 snímkov falzifikátov. Do zberu dát bolo zapojených 114 subjektov a použitých bolo 25 rôznych typov falzifikátov. V datasete sú zahrnuté aj snímky ľudí s odlišným pigmentom.

Pre rozlíšenie živých rúk a falzifikátov bola použitá neurónova sieť Resnet. Jej architektúra bola inšpirovaná prácou, venujúcou sa detekcii živosti na prstoch. Najprv bola sieť natrénovaná a vyhodnotená v základnej konfigurácii. Následne sa pristúpilo k experimentom, ktorých cieľom bolo zlepšovať výsledky. Využitá bola augmentácia, ako aj odstránenie šumu a prahovanie. Pomocou techniky transfer learning bolo možné dosiahnuť s inými, už predtrénovanými modelmi ešte lepšie výsledky, ako s pôvodným modelom. Okrem reziduálnej siete bolo do porovnania zahrnutých ešte 5 ďalších modelov.

Vyhodnocovaná bola chybovosť detekovania živej ruky BPCER a prezentačného útoku APCER. Ako najlepší obstál model odvodený od MobileNetV2, pri ktorom prebiehalo trénovanie len na klasifikátore. V prípade, že sa snímky odlišne sfarbených rúk nachádzali aj v trénovacom datasete, dosiahla neurónová sieť APCER na úrovni 1.59% pri 0.47% BPCER, čo zodpovedalo 99.12% celkovej úspešnosti. Ak sa tieto snímky nachádzali len vo validačnom datasete, a teda sa s nimi sieť nestretla v procese učenia, dosiahol model APCER len na úrovni 0.79%, avšak s BPCER 5.16%. Treba podotknúť, že použité boli len nižšie vlnové dĺžky, pri ktorých je zmena pigmentu na snímkoch výrazne viditeľná. Kamera pre rozsah 1500-1600 nm totiž nebola dostatočne citlivá.

V závere poslednej kapitoly bolo spomenutých niekoľko vylepšení do budúcna. Asi najpotrebnejším je doplnenie využitia vyšších vlnových dĺžok, kedy sa očakáva vyššia úspešnosť pri detekcii ľudskej ruky kvôli rovnakej odrazivosti pri rôznych hodnotách melanínu. Takisto by mohla klesnúť APCER, najmä v prípade plastov. Veľmi potrebné je tiež nasnímať rozsiahlejší dataset s ešte dôveryhodnejšími falzifikátmi, ktoré by boli ťažšie odhaliteľné. Otázne je aj to, aký efekt by mala na úspešnosť technika fine-tuning.

Literatúra

[1] Andrzej Grzybowski, K. P. Jan Evangelista Purkynje (1787–1869): First to describe fingerprints. *Clinics in Dermatology*. 2015, zv. 33, č. 1, s. 117 – 121. DOI: https://doi.org/10.1016/j.clindermatol.2014.07.011. ISSN 0738-081X. Leprosy: I. Dostupné z: http://www.sciencedirect.com/science/article/pii/S0738081X14001539.

[2] Bozhao, T. a Stephanie, S. New approach for liveness detection in fingerprint scanners based on valley noise analysis. *J. Electronic Imaging*. Január 2008, zv. 17, s. 011009. DOI: https://doi.org/10.1117/1.2885133.

[3] Bradley, J. N., Brislaw, C. M. a Hopper, T. The FBI WaveletScalar Quantization Standard for grayscale fingerprint image compression. *Proc. SPIE 1961*. 1993. DOI: https://doi.org/10.1117/12.150973.

[4] Branke, J. a Motahari, M. *How is AI being used in finance and stock markets?* [online]. 2020 [cit. 2020-12-26]. Dostupné z: https://londonlovesbusiness.com/how-is-ai-being-used-in-finance-and-stock-markets/.

[5] Chen, B. *Early Stopping in Practice: an example with Keras and TensorFlow 2.0* [online]. 2020 [cit. 2021-06-06]. Dostupné z: https://towardsdatascience.com/a-practical-introduction-to-early-stopping-in-machine-learning-550ac88bc8fd.

[6] Editors, H. *A bloody fingerprint elicits a mother's evil tale in Argentina* [online]. A & E Television Networks, 2019 [cit. 2020-11-02]. Dostupné z: https://www.history.com/this-day-in-history/a-bloody-fingerprint-elicits-a-mothers-evil-tale-in-argentina.

[7] Fei, J., Xia, Z., Peipeng, Y. a Xiao, F. Adversarial attacks on fingerprint liveness detection. *EURASIP Journal on Image and Video Processing*. Január 2020. DOI: https://doi.org/10.1186/s13640-020-0490-z.

[8] Frassetto Nogueira, R., de Alencar Lotufo, R. a Campos Machado, R. Evaluating software-based fingerprint liveness detection using Convolutional Networks and Local Binary Patterns. In: *2014 IEEE Workshop on Biometric Measurements and Systems for Security and Medical Applications (BIOMS) Proceedings*. 2014, s. 22–29. DOI: https://doi.org/10.1109/BIOMS.2014.6951531.

[9] Galbally, J., Alonso Fernandez, F., Fierrez, J. a Ortega Garcia, J. A high performance fingerprint liveness detection method based on quality related features. *Future Generation Computer Systems*. 2012, zv. 28, č. 1, s. 311 – 321. DOI: https://doi.org/10.1016/j.future.2010.11.024. ISSN 0167-739X. Dostupné z: http://www.sciencedirect.com/science/article/pii/S0167739X1000244X.

[10] Gomes, J. *Adversarial Attacks and Defences for Convolutional Neural Networks* [online]. 2018 [cit. 2020-11-30]. Dostupné z: https://medium.com/onfido-tech/adversarial-attacks-and-defences-for-convolutional-neural-networks-66915ece52e7.

[11] Gomez Barrero, M. a Busch, C. Multi-Spectral Convolutional Neural Networks for Biometric Presentation Attack Detection. *Proceedings of the 12th Norwegian Information Security Conference*. November 2019, zv. 12.

[12] Gupta, T. *Deep Learning: Feedforward Neural Network* [online]. 2017 [cit. 2020-12-26]. Dostupné z: https://towardsdatascience.com/deep-learning-feedforward-neural-network-26a6705dbdc7.

[13] Guru99. *Supervised vs Unsupervised Learning: Key Differences* [online]. [cit. 2020-12-23]. Dostupné z: https://www.guru99.com/supervised-vs-unsupervised-learning.html.

[14] Hamamatsu. *NIR and SWIR Questions and Answers* [online]. [cit. 2020-11-30]. Dostupné z: https://hub.hamamatsu.com/us/en/ask-engineer/nir-and-swir-questions-and-answers/index.html.

[15] Hamamatsu. *See Beyond Visible: Short Wavelength Infrared Introduction and Applications* [online]. [cit. 2020-11-30]. Dostupné z: https://hub.hamamatsu.com/sp/hc/resources/webinars/SWIR_intro_V103.pdf.

[16] Heidari, M., Goldmann, T., Dvořák, M. a Drahanský, M. Antispoofing and multispectral (optical) methods in hand-based biometrics. In: *Hand-Based Biometrics: Methods and Technology*. The Institution of Engineering and Technology, 2018, s. 337–365. IET Book Series on Advances in Biometrics. ISBN 978-1-78561-224-4. Dostupné z: https://www.fit.vut.cz/research/publication/11698.

[17] Husseis, A., Liu Jimenez, J., Goicoechea Telleria, I. a Sanchez Reillo, R. Dynamic Fingerprint Statistics: Application in Presentation Attack Detection. *IEEE Access*. Máj 2020, PP, s. 1–1. DOI: https://doi.org/10.1109/ACCESS.2020.2995829.

[18] Inc., E. O. *What is SWIR?* [online]. [cit. 2020-11-30]. Dostupné z: https://www.edmundoptics.com/knowledge-center/application-notes/imaging/what-is-swir/.

[19] Jain, A. K., Flynn, P. a Ross, A. A. *Handbook of Biometrics*. Springer, 2008. ISBN 978-0-387-71040-2.

[20] Jeanlagi. *Draw by own force, CC BY 3.0* [online]. 2008 [cit. 2020-11-02]. Dostupné z: https://commons.wikimedia.org/w/index.php?curid=3910972.

[21] Kasirajan, A. *UNDERFIT and OVERFIT Explained* [online]. 2020 [cit. 2021-06-10]. Dostupné z: https://medium.com/@minions.k/underfit-and-overfit-explained-8161559b37db.

[22] Lee, J., Moon, S. a Lim, J. Imaging of the Finger Vein and Blood Flow for Anti-Spoofing Authentication Using a Laser and a MEMS Scanner. *Sensors*. Apríl 2017. DOI: https://doi.org/10.3390/s17040925.

[23] Maltoni, D., Maio, D. a Jain, A. K. *Handbook of Fingerprint Recognition*. Springer, 2003. ISBN 978-1-84882-253-5.

[24] Matsumoto, T., Matsumoto, H., Yamada, K. a Hoshino, S. Impact of Artificial "Gummy"Fingers on Fingerprint Systems. *Proceedings of SPIE*. Január 2002, zv. 4677. DOI: https://doi.org/10.1117/12.462719.

[25] Mwiti, D. *Transfer Learning Guide: A Practical Tutorial With Examples for Images and Text in Keras* [online]. 2021 [cit. 2021-06-25]. Dostupné z: https://neptune.ai/blog/transfer-learning-guide-examples-for-images-and-text-in-keras.

[26] Notesco, G., Kopackova Strnadova, V., Rojík, P., Schwartz, G., Livne, I. et al. Mineral Classification of Land Surface Using Multispectral LWIR and Hyperspectral SWIR Remote-Sensing Data. A Case Study over the Sokolov Lignite Open-Pit Mines, the Czech Republic. *Remote Sensing*. Júl 2014, zv. 6, s. 7005–7025. DOI: https://doi.org/10.3390/rs6087005.

[27] Parthasaradhi, S. T. V., Derakhshani, R., Hornak, L. A. a Schuckers, S. A. C. Time-series detection of perspiration as a liveness test in fingerprint devices. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*. 2005, zv. 35, č. 3, s. 335–343. DOI: https://doi.org/10.1109/TSMCC.2005.848192.

[28] Protocol, R. *Everything you need to know about Neural Networks* [online]. 2017 [cit. 2020-12-23]. Dostupné z: https://medium.com/ravenprotocol/everything-you-need-to-know-about-neural-networks-6fcc7a15cb4.

[29] Putte, T. van der, Keuning, J. a Prabhakar, S. Biometrical fingerprint recognition: don't get your fingers burned. *Proceedings of IFIP TC8/WG8.8 Fourth Working Conference on Smart Card Research and Advanced Applications*. Kluwer Academic Publishers. September 2000, s. 289–303.

[30] Roman, V. *CNN Transfer Learning & Fine Tuning* [online]. 2020 [cit. 2021-06-27]. Dostupné z: https://towardsdatascience.com/cnn-transfer-learning-fine-tuning-9f3e7c5806b2.

[31] Saha, S. *A Comprehensive Guide to Convolutional Neural Networks* [online]. 2018 [cit. 2020-12-27]. Dostupné z: https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53.

[32] Sandstrom, M. Liveness Detection in Fingerprint Recognition Systems. Jún 2004, s. 38.

[33] Schuckers, S. A. C. Spoofing and Anti-Spoofing Measures. *Information Security Technical Report*. Kluwer Academic Publishers. December 2002, zv. 7, č. 4, s. 61.

[34] Shukla, P. a Iriondo, R. *Main Types of Neural Networks and its Applications* [online]. 2020 [cit. 2020-12-26]. Dostupné z: https://medium.com/towards-artificial-intelligence/main-types-of-neural-networks-and-its-applications-tutorial-734480d7ec8e.

[35] Spreinat, A., Selvaggio, G., Erpenbeck, L. a Kruss, S. Multispectral near infrared absorption imaging for histology of skin cancer. *Journal of Biophotonics*. 2020, zv. 13, č. 1, s. e201960080. DOI: https://doi.org/10.1002/jbio.201960080. Dostupné z: https://onlinelibrary.wiley.com/doi/abs/10.1002/jbio.201960080.

[36] Steiner, H. *Active Multispectral SWIR Imaging for Reliable Skin Detection and Face Verification*. Január 2017. ISBN 9783736994508.

[37] Steiner, H., Schwaneberg, O. a Jung, N. Advances in active near-infrared sensor systems for material classification and skin detection for safety applications. *Safety Science Monitor*. Január 2013, zv. 17, s. 3–4.

[38] Steiner, H., Sporrer, S., Kolb, A. a Jung, N. Design of an Active Multispectral SWIR Camera System for Skin Detection and Face Verification. *Journal of Sensors*. Január 2016, zv. 2016, s. 1–16. DOI: https://doi.org/10.1155/2016/9682453.

[39] Thalheim, L., Krissler, J. a Ziegler, P.-M. Biometric Access Protection Devices and their Programs Put to the Test. *C'T Magazine*. November 2002, s. 114.

[40] Tolosana, R., Gomez Barrero, M., Busch, C. a Ortega Garcia, J. Biometric Presentation Attack Detection: Beyond the Visible Spectrum. *IEEE Transactions on Information Forensics and Security*. August 2019, PP, s. 1. DOI: https://doi.org/10.1109/TIFS.2019.2934867.

[41] Willis, D. a Lee, M. Six Biometric Devices Point the Finger and Security. *Network Computing*. Jún 1998, zv. 9, č. 10, s. 84–96.

[42] Zacha, J. *Konvoluční neuronové sítě pro klasifikaci objektů z LiDARových dat*. 2019. 9 s. Diplomová práca. České vysoké učení technické v Praze.

Príloha A

Obsah priloženého pamäťového média

Priložené DVD obsahuje súbory v nasledujúcej hierarchii:

- Sablona2019 - adresár s LaTeXsúbormi pre vygenerovanie tohto dokumentu
- *bakalarskapraca.pdf* - elektronická verzia tejto bakalárskej práce
- *cameras.xlsx* - Excelova tabuľka porovnania kamier, ktorá vznikla pri prieskume trhu (užitočná pri pokračovaní práce s vyššími vlnovými dĺžkami)
- Codes - adresár obsahujúci podadresáre so zdrojovými kódmi a príslušnou dokumentáciou:
  - capture_hand - podadresár s programom na snímanie datasetu a popisom v readme
  - capture_hand_libs - knižnice k programu capture_hand, so sprievodným readme
  - led_controller - podadresár obsahujúci program pre arduino (Freaduino Uno 2017), popis v readme a schému zapojenia
  - liveness_detection - podadresár s kódmi pre Google Colab (načítanie a úprava dát, neurónové siete a ich trénovanie, testovanie, vizualizácia feature máp) + readme
- trained_models.zip - archív obsahujúci jednotlivé natrénované modely, spolu s popísanými výsledkami v readme súbore
- dataset_readme.md - informácie k zozbieraným dátam (dataset u autora a vedúceho práce)

Príloha B

Zariadenie na snímanie datasetu

Na nasledujúcich snímkoch je zobrazené výsledné zariadenie. Zásuvný modul bol uspôsobený na to, že vonkajšia krabica mohla byť položená vertikálne alebo horizontálne. Kvôli potrebe snímania oboch rúk sa ukázal prvý spôsob ako vhodnejší. Ruka bola vložená do krabice vo vodorovnej polohe, dlaňou smerom nadol.

Obr. B.1: Výsledné zariadenie - kartónová krabica, obsahujúca zásuvný modul s kamerou a arduinom, pripojenými k PC. Štrbina v uzávere krabice nijako neovplyvňovala snímanie, keďže senzor kamery nebol na toľko citlivý.

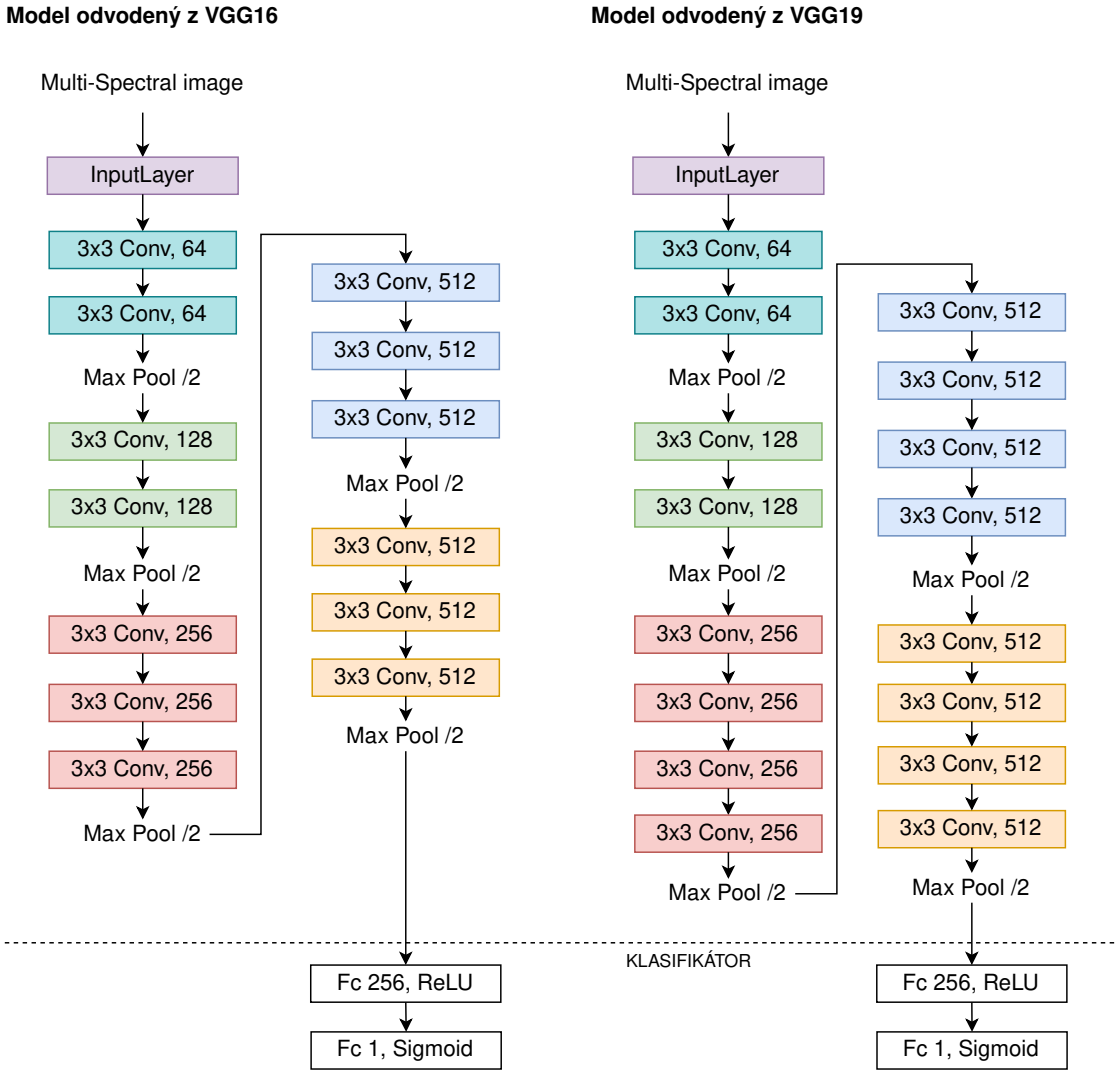
Obr. B.2: Naľavo pohľad do vnútra zásuvného modulu, kde arduino sprostredkuje uzemnenie a zdroj pre napájanie LED. 2 digitálne piny slúžia na spínanie MOSFET tranzistorov, riadiacich rozsvietenie jednotlivých druhov LED. Kamera je skrutkou upevnená k podstavci. Na obrázku vpravo je už podstavec umiestnený v krabici na výšku, tak aby kamera spolu s LED diódami smerovali nahor. Horná vnútorná stena krabice je vystlaná netkanou textíliou, tvoriacou pozadie na snímkach ruky.

Obr. B.3: Ruka zobrazená na displeji PC počas snímania, pri 850nm (s funkciou prisvietenia vo vytvorenej aplikácii).

Príloha C

Architektúry modelov z porovnania

Využitá bola technika transfer learning, pri ktorej sa pre každý model vytvoril nový klasifikátor a trénovanie prebiehalo len na ňom. Klasifikátor bol u všetkých modelov totožný. Váhy ostatných vrstiev boli stanovené predtrénovaním na datasete Imagenet. Na obrázku nižšie je znázornená výsledná architektúra pre modely odvodené od VGG16 a VGG19. Čo sa týka MobileNetV2, DenseNet169 a InceptionResnetV2, princíp bol rovnaký. Ich architektúry však boli príliš rozsiahle pre grafické znázornenie.



Obr. C.1: Architektúra modelov odvodených z predtrénovaných modelov VGG16 a VGG19, dostupných v knižnici *Keras*. Trénovanie prebiehalo na vrstvách znázornených bielou farbou, teda na klasifikátore.