



**VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ**  
BRNO UNIVERSITY OF TECHNOLOGY

**FAKULTA INFORMAČNÍCH TECHNOLOGIÍ**  
FACULTY OF INFORMATION TECHNOLOGY

**ÚSTAV INTELIGENTNÍCH SYSTÉMŮ**  
DEPARTMENT OF INTELLIGENT SYSTEMS

**ANALÝZA KONVOLUČNÍCH NEURONOVÝCH SÍTÍ PRO  
DETEKCI A KLASIFIKACI POŠKOZENÍ OTISKU PRSTU**  
ANALYSIS OF CONVOLUTIONAL NEURAL NETWORKS FOR DETECTION AND CLASSIFICATION  
OF DAMAGES IN FINGERPRINT IMAGES

**SEMESTRÁLNÍ PROJEKT**  
TERM PROJECT

**AUTOR PRÁCE**  
AUTHOR

**Bc. KATEŘINA FOŘTOVÁ**

**VEDOUCÍ PRÁCE**  
SUPERVISOR

**Ing. ONDŘEJ KANICH, Ph.D.**

**BRNO 2022**

## **Abstrakt**

Cílem tohoto semestrálního projektu je nastudovat literaturu ohledně biometrie a zpracování otisků prstů s důrazem na možná onemocnění, které mohou oblast prstu postihnout. Následně se práce zabývá rozpoznáváním s využitím neuronových sítí. Projekt popisuje architektury konvolučních neuronových sítí a modelů pro detekci objektů až po nejnovější přístupy. Je navrhnuuto několik metod pro detekci a klasifikaci onemocnění otisků prstů využívající moderní architektury, rozličné typy páteřních sítí a metod pro detekci.

## **Abstract**

The aim of this semester project is to study the literature on biometrics and fingerprint processing with an emphasis on possible diseases that may affect the fingertip area. Subsequently, the work deals with neural network-based recognition. The project describes the architectures of convolutional neural networks and object detection approaches up to the latest research. Several detection methods for detection and classification of skin diseases affecting fingertip are proposed with using modern architectures, different types of backbone networks and detection methods.

## **Klíčová slova**

onemocnění kůže, otisky prstů, konvoluční neuronové sítě, detekce, klasifikace

## **Keywords**

skin diseases, fingerprints, convolutional neural networks, detection, classification

## **Citace**

FOŘTOVÁ, Kateřina. *Analýza konvolučních neuronových sítí pro detekci a klasifikaci poškození otisku prstu*. Brno, 2022. Semestrální projekt. Vysoké učení technické v Brně, Fakulta informačních technologií. Vedoucí práce Ing. Ondřej Kanich, Ph.D.

# **Analýza konvolučních neuronových sítí pro detekci a klasifikaci poškození otisku prstu**

## **Prohlášení**

Prohlašuji, že jsem tento semestrální projekt vypracovala samostatně pod vedením Ing. Ondřeje Kanicha, Ph.D. Uvedla jsem všechny literární prameny, publikace a další zdroje, ze kterých jsem čerpala.

.....  
Kateřina Fořtová  
20. ledna 2022

# Obsah

<b>1</b>	<b>Úvod</b>	<b>2</b>
<b>2</b>	<b>Rozpoznávání podle otisků prstů</b>	<b>3</b>
2.1	Biometrie . . . . .	3
2.2	Struktura kůže . . . . .	4
2.3	Nemoci kůže ovlivňující otisky prstů . . . . .	6
2.4	Zpracování otisku prstů . . . . .	10
2.5	Typy senzorů . . . . .	14
2.6	Generování syntetických otisků prstů . . . . .	16
<b>3</b>	<b>Detekce a klasifikace pomocí neuronových sítí</b>	<b>18</b>
3.1	Umělý neuron a perceptron . . . . .	18
3.2	Vícevrstvý perceptron . . . . .	19
3.3	Aktivační funkce . . . . .	20
3.4	Loss funkce . . . . .	21
3.5	Trénování neuronové sítě . . . . .	22
3.6	Konvoluční neuronové sítě . . . . .	23
3.7	Existující architektury neuronových sítí pro zpracování obrazu . . . . .	24
3.8	Přístupy pro detekci objektů . . . . .	27
<b>4</b>	<b>Návrh řešení a implementace</b>	<b>32</b>
4.1	Databáze otisků prstů . . . . .	32
4.2	Anotace dat . . . . .	33
4.3	Předzpracování otisků prstů . . . . .	33
4.4	Použité metody pro detekci a klasifikaci . . . . .	35
4.5	Výstup detekce a klasifikace onemocnění otisků prstů . . . . .	36
<b>5</b>	<b>Závěr</b>	<b>39</b>
<b>Literatura</b>		<b>40</b>

# Kapitola 1

## Úvod

V současném světě je stále více rozšířeným užití biometrie jako součásti bezpečnostních mechanismů. Jednou z nejrozšířenějších biometrických charakteristik je otisk prstu. Pro snímání by měly být zaručeny ideální podmínky, avšak jedná se o nesplnitelný úkol v běžném nasazení. Snímky získané ze senzoru mohou být nekvalitní, mohou obsahovat nečistoty, které dále komplikují proces rozpoznávání. Vysoké procento populace trpí kožními onemocněními. Mnohá široce rozšířená onemocnění, například atopický ekzém nebo bradavice, postihují i oblast bříška prstu. U některých z onemocnění (např. bradavice) je jejich vliv na rozpoznávání nižší, protože papilární linie jsou zachovány v nepostiženém místě otisku prstu, u tolik rozšířeného atopického ekzému však jsou ve vážnějších podobách papilární linie těžko rozpoznatelné. Potenciální detektor a klasifikátor těchto onemocnění by mohl odhalit možná požadované uživatele využívající biometrický systém.

Problematika detekce a klasifikace pro rozličné typy objektů je stále jedním z předních témat současnosti. Techniky, které kdysi spíše využívaly ruční extrakci rysů, byly nahrazeny konvolučními neuronovými sítěmi, které dosahují vyšší přesnosti, ale i výpočetní náročnosti. Moderní detektory typů Faster R-CNN (Region Based Convolutional Neural Network) nebo SSD (Single Shot Detector) jsou využívané v reálných aplikacích a při rozpoznávání rozličných typů objektů. Cílem této práce je využít několika typů těchto moderních detektorů na úloze detekce a klasifikace různých typů onemocnění otisku prstu. Přitom je kladen důraz na experimenty s dostupnými architekturami a snahou nalézt co neuniverzálnější detektor, který by zvládl rozlišit co nejvíce typů onemocnění – pro začátek jsou pro experimenty vybrány atopický ekzém, psoriáza, dishydrozia a bradavice.

## Kapitola 2

# Rozpoznávání podle otisků prstů

Lidské otisky prstů jsou jednou z nejznámějších a nejužívanějších biometrických charakteristik. V roce 1823 bylo vytvořeno Janem Evangelistou Purkyně první klasifikační schéma otisků prstů, které rozdělovalo otisky prstů do devíti tříd na základě konfigurace hřebenů papilárních linií. V roce 1863 bylo ve Velké Británii akceptováno, že žádní dva lidé nemají totožné otisky prstu [29]. Následující kapitola se zabývá stručným popisem oboru biometrie a jejími základními pojmy. Hlavním z cílů této kapitoly je však představit onemocnění, které mohou plochu otisku prstu postihnout a popsat základní přístupy pro zpracování a rozpoznávání otisků prstů.

### 2.1 Biometrie

Cílem biometrie (biometrického rozpoznávání) je použití anatomických a behaviorálních charakteristik, které se nazývají biometrické rysy, k automatickému rozpoznávání jedinců. Biometrie se postupem času stala efektivním řešením lidské identifikace reprezentující tělesnou identitu jednotlivce [29]. Představuje jednu z technologií, které mohou umožnit naší společnosti bezpečnější a poskytnout vyšší uživatelské pohodlí [29]. Biometrický systém však může být napadnutelný a biometrická identita nemůže být v případě prozrazení anulována [12].

#### 2.1.1 Identita, identifikace, verifikace a autentizace

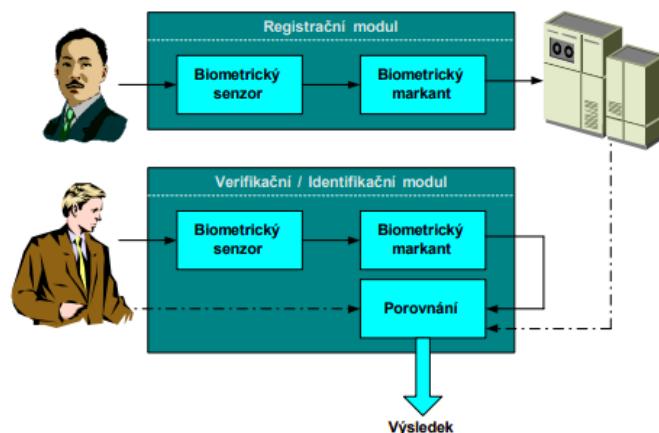
Identita, identifikace, verifikace a autentizace jsou základními pojmy, které biometrie využívá. Rozpoznání je založeno na jednoznačné identitě jedince. Identita je jednoznačnou charakteristikou každého z nás. Rozlišujeme fyzickou a elektronickou identitu. Fyzická identita je pouze jediná a je definována vzhledem a chováním. Elektronických identit můžeme mít ovšem vytvořených více (např. více účtů na sociálních sítích) [12].

Identifikace je situací, kdy daná osoba zadá svoji biometrickou vlastnost systému, ale nesdílí mu svoji identitu. Daný systém pak má za úkol rozpoznat identitu uživatele – porovnává se vzorek zadaný na vstupu s celou databází. Systém poté naleze nebo nenaleze danou identitu. Identifikaci můžeme chápat jako porovnání 1:N [12]. Při verifikaci je sdělena systému elektronická identita uživatele a dojde k ověření fyzické identity. V databázi je vyhledán daný záznam uživatele, který obsahuje biometrická data. Pokud je záznam nenalezen, dojde k zamítnutí přístupu uživatele, jinak dochází k porovnání dat – výsledkem je potvrzení či nepotvrzení identity. Proces verifikace je také jinak nazýván porovnání 1:1 [12]. Při autentizaci má systém za úkol potvrdit autentičnost (hodnověrnost) osoby

žádající o přístup. Autentizace může probíhat jak při identifikaci, tak při autentizaci. Rozhodnutí je obvykle prováděno na základě nastavené hodnoty prahu [12].

### 2.1.2 Biometrické systémy a vlastnosti

Biometrický systém, znázorněn na obrázku 2.1, je složen z registračního a verifikačního / identifikačního modulu. Součástí obou uvedených modulů je biometrický senzor a biometrický markant – extrahované rysy z biometrické informace na vstupu. Hlavním rozdílem obou modulů je princip práce s databází a biometrickým markantem. Při práci s registračním modulem je biometrický markant uložen do databáze. Verifikační / identifikační modul naopak neukládá biometrický markant do databáze, ale místo toho načítá data z databáze pro porovnání biometrického markantu na vstupu s údaji databáze. Při tomto porovnání obdržíme výsledek na základě nalezení případné shody a případného operačního módu – verifikaci nebo identifikaci [12].



Obrázek 2.1: Schéma biometrického systému [12]

Biometrické charakteristiky můžeme dělit na anatomické a dynamické vlastnosti. U anatomických vlastností je jeden pevný rys jednou biometrickou vlastností. Do anatomických vlastností spadá např. otisk prstu, obličej, duhovka a sítnice oka, dlaň nebo geometrie ruky. Dynamické vlastnosti jsou spjaty s nějakou akcí uživatele, do této kategorie spadá např. chůze, hlas, gestikulace obličeje či dynamické vlastnosti podpisu [12].

## 2.2 Struktura kůže

Kůže je největším lidským orgánem, který funguje jako počáteční obrana proti patogenům, ultrafialovému (UV) záření, chemikáliím nebo zranění. Také slouží jako regulátor teploty. Plocha kůže tvoří asi  $1,5 - 1,8 \text{ m}^2$  a má hmotnost zhruba 4,5 kg [15]. Lidská kůže je složena ze tří vrstev:

- epidermis (pokožka),
- corium (dermis – škára),
- tela subcutanea (subcutis — podkožní vazivo) [15].

**Pokožka** (epidermis) je složena z několika vrstev plochých buněk. Buňky na povrchu postupně prochází procesem odumírání, rohovatění a odlupování, kdy jsou nahrazeny buňkami z hlubší epidermální vrstvy. Zvláštní vazivové buňky v hlubších vrstvách obsahují kožní pigment melanin. Barva lidské kůže pak závisí na množství melaninu, jeho hloubce uložení a prokrvení kůže. Melanin má za funkci pohlcovat UV složku slunečního záření, která by jinak dokázala poškodit buňky hlubších vrstev [15].

**Škára** (corium) je složena z vazivových buněk, elastických vláken a tukových buněk. Elastická vlákna mají za úkol zajišťovat pružnost, roztažitelnost, pevnost a štěpitelnost kůže. Ve škáře jsou uloženy dva typy kožních žláz – mazové a potní žlázy. Také se zde vyskytují krevní, mízní cévy, vlasové kořeny a nervy. Ve výběžcích proti epidermální vrstvě se vyskytují nervová zakončení – receptory umožňující vnímání bolesti, tepla nebo chladu [15].

**Podkožní vazivo** (subcutis) je tvořené sítí kolagenních, elastických vláken a vazivovými buňkami. Slouží jako potenciální tuková tkáň, která je schopna ukládat v buňkách velké množství tukových kapének [15].

Nejdůležitějšími funkcemi kůže jsou:

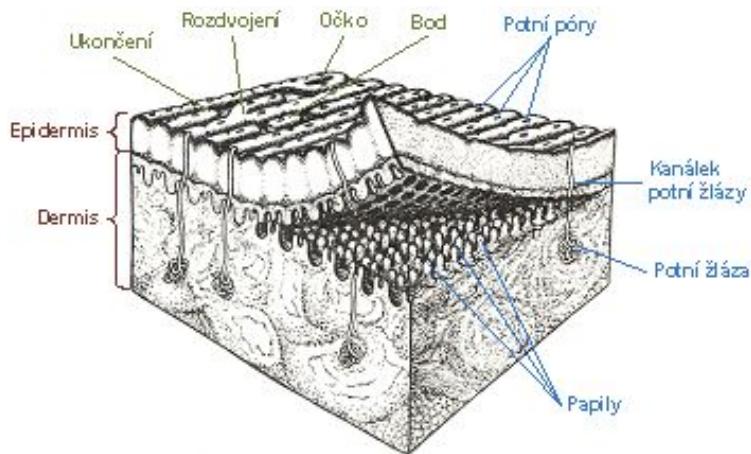
- Ochrana těla – Kůže zabraňuje vnikání škodlivých látek do vnitřního prostředí organismu a je odolnou vůči mechanickému poškození. Pigment melanin chrání proti UV záření [15].
- Smyslová funkce kůže – V kůži se vyskytuje velké množství receptorů sloužící k vnímání mechanických, tepelných a bolestivých počitků. Specializované receptory slouží k vnímání tepla, chladu a hmatu [15].
- Udržování tělesné teploty – Zrohovatělá vrstva buněk díky své špatné tepelné vodivosti chrání organismus před většími tepelnými ztrátami. Podkožní vazivo tvoří poměrně silnou vrstvu u každého jedince [15].
- Skladovací funkce kůže – Množství tuku v podkožním vazivu tvoří energetickou zásobárnu organismu. Probíhá zde uskladnění vitamínů rozpustných v tucích (A, D, E, K) [15].
- Vylučovací funkce kůže – Vyloučené sekrety mazových a potních žláz se uplatňují při ochraně kůže i celého organismu [15].

### 2.2.1 Papilární linie

Na prstech rukou a nohou jsou přítomné vyvýšené kožní útvary, které se nazývají papilární linie. Základ papilárních linií je formovaný na hranici kožních vrstev epidermis a corium, kde dochází k vzájemné interakci spodní epidermální vrstvy a horní vrstvy coria. Po určitém zvlnění epidermální vrstvy se na bříškách prstů objeví papilární linie [13].

Struktura kůže a papilárních linií je znázorněna na obrázku 2.2, na obrázku 2.3 je pak zobrazen nasnímaný otisk prstu tvořený charakteristickými papilárními liniemi. Papilární linie slouží k určení identity jedince, tvoří dermatoglyfy, které jsou pro jedince charakteristické. Přesný důvod vzniku papilárních linií však není známý – uvažuje se o hypotéze, že to jsou např. výsledky napětí a tlaků v kůži během embryonálního vývoje. Uspořádání papilárních linií je z 90 % podmíněno geneticky, z 10 % je pak závislé na vnějších podmínkách [9]. Ani jednovaječná dvojčata nemají shodné otisky prstů – shodují se sice v základních vztazích, ale odlišují se v drobných detailech rozhodujícími pro identifikaci osob [9]. Papilární linie se vyvíjejí během nitroděložního vývoje a po celý život jedince vykazují stabilitu, kdy

se s přibývajícím věkem mění rozměry plošek prstů, dlaní či chodidel, avšak struktura zůstává stejná [13].



Obrázek 2.2: Struktura kůže [13]



Obrázek 2.3: Příklad nasnímaného otisku prstu [14]

## 2.3 Nemoci kůže ovlivňující otisky prstů

Nemoci kůže představují důležitý, ale často opomíjený faktor ovlivňující snímání otisků prstů a jejich následnou analýzu. Uvádí se, že zhruba 20 – 25 % pacientů trpí některou z kožních chorob [11]. Tato kapitola uvádí výčet onemocnění, které mohou postihnout oblast prstů a způsob, jakým se projevují na zachyceném snímku na biometrickém senzoru.

### 2.3.1 Atopický ekzém

Atopický ekzém (viz obrázek 2.4), je jedním z druhů ekzému – nejčastějšího chronického zánečlivého onemocnění. Původ onemocnění zahrnují například genetické faktory, defekt kožní bariéry a vliv prostředí. Předpokládá se, že atopický ekzém je spjatý s příbuznými onemocněními jako astma, potravinová alergie nebo alergická rýma. Pokud jeden z rodičů dítěte trpí atopickým ekzémem, je více jak poloviční pravděpodobnost, že bude symptomy postihnut i jeho potomek. Při postižení obou rodičů je tato šance dokonce 80 % [27]. Atopický ekzém se projevuje svěděním, snížením hydratace a zvýšením suchosti pokožky. Postižené místo

je citlivé k alergenům a dráždivým látkám [27]. Při léčbě je nutné vyhýbat se spouštěcím faktorům a zamezit vysušování pokožky.

Při snímání se onemocnění atopickým ekzémem projevuje více způsoby. Příklad snímku atopického ekzému zachyceného biometrickým senzorem je znázorněn na obrázku 2.5. U pacientů s mírnějším průběhem se nemoc projevuje úzkými bílými liniemi přecházejícími přes papilární linie. Papilární linie zůstávají přerušené však pouze na několika místech a obvykle jsou stále uspokojivě čitelné. Ovšem otisky prstů pacientů s vážnějším průběhem nemoci jsou velmi poškozené, papilární linie jsou skoro nečitelné a vedle bílých linií se často vyskytují i tmavé nepravidelné skvrny.



Obrázek 2.4: Snímek pacienta trpící atopickým ekzémem [4]



Obrázek 2.5: Atopický ekzém zachycený biometrickým senzorem [4]

### 2.3.2 Bradavice

Bradavice (verruca vulgaris) (viz obrázek 2.6) jsou běžnými novotvary způsobené lidskými papilomaviry (HPV). Tyto viry, kterých existuje více jak 100 různých typů, dokáží vstoupit do kůže prostřednictvím malých řezných ran a způsobit zvýšený růst buněk. Kůže se tedy na daném místě postupně stává silnější – je vytvořena bradavice. Bradavice se mohou vyskytovat jednotlivě nebo ve skupinách na různých částech lidského těla. Zbavení se bradavic může být náročné zejména pro lidi se slabým imunitním systémem. U lidí se zdravým imunitním systémem obvykle bradavice zmizí postupem času [23]. Bradavice se na snímcích projevují jako bílé skvrny kulatého tvaru, občas s černými tečkami uvnitř (viz obrázek 2.7). Poškození není tolik závažné, jako např. u atopického ekzému a psoriázy a papilární linie zbytku otisku prstu jsou nepoškozené.



Obrázek 2.6: Snímek pacienta trpící bradavici [11]



Obrázek 2.7: Bradavice zachycená biometrickým senzorem [4]

### 2.3.3 Psoriáza

Psoriáza (lupénka) (viz obrázek 2.8) se vyskytuje zhruba u 1–3 % lidské populace. Původ onemocnění je neznámý, avšak je uvažováno o dědičných predispozicích. Nejčastějším projevem tohoto chronického onemocnění je odlupování pokožky. Výskyt onemocnění na dlaních a prstech se vyznačuje červenými plaky. Při zhoršeném průběhu se můžou objevit bolestivé trhliny a krvácení. Psoriáza stejně jako bradavice nemá žádnou dostupnou léčbu. Obvykle se nemoc postupně zhorší a zlepší [30, 11]. Onemocnění je často na snímcích rozpoznatelné pomocí tmavé skvrny nepravidelného tvaru, která je ohraničena bílými liniemi (viz obrázek 2.9). Papilární linie jsou obvykle čitelné pouze na menší části snímku.



Obrázek 2.8: Snímek pacienta trpící psoriázou [11]



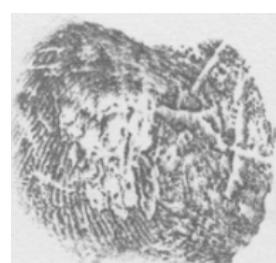
Obrázek 2.9: Psoriáza zachycená biometrickým senzorem [4]

### 2.3.4 Dyshidróza

Dyshidróza (viz obrázek 2.10) je jednou z nejčastějších kožních poruch. Původ onemocnění je neznámý, ale uvažuje se například o důvodu styku s kontaktním alergenem. Onemocnění je charakteristické výskytem drobných svědivých puchýřků na dlaních a stranách prstu. Kůže může být vlhká a zarudlá. Postupem času tyto puchýřky prasknou a objevují se zarudlé šupinaté skvrny [11]. Příklad zachycení onemocnění dyshidrózy biometrickým senzorem je zobrazen na obrázku 2.11. Při snímcích na senzoru se onemocnění projevuje více způsoby – většinou se jedná o bílé čáry a nepravidelné bílé skvrny. Při vážnějším poškození nejsou dostatečně viditelné papilární linie.



Obrázek 2.10: Snímek pacienta trpící dyshidrózou [11]



Obrázek 2.11: Dyshidróza zachycená biometrickým senzorem [4]

### 2.3.5 Acrodermatitis continua

Acrodermatitis continua (viz obrázek 2.12) je chronickým onemocněním, jedním ze zvláštních druhů psoriázy. Projevuje se výskytem pustul – dutin vyplněných hnisek na špičkách

prstů, většinou vyskytující se častěji na prstech dlaní než chodidel. Toto onemocnění nemá žádný standardizovaný postup léčby a jeho původ není známý, uvažuje se však o např. infekčním nebo genetickém původu. Při vážném průběhu onemocnění může onemocnění poškodit oblast nehtu [41]. Snímky na senzoru jsou typické nerozpoznatelnými papilárními liniemi, celý otisk prstu pokrývají tmavé menší skvrny různých tvarů (viz obrázek 2.13).



Obrázek 2.12: Snímek pacienta trpící Acrodermatitis continua [41]



Obrázek 2.13: Acrodermatitis continua zachycená biometrickým senzorem [4]

### 2.3.6 Raynaudova nemoc

Raynaudova nemoc (viz obrázek 2.14) je charakteristická křečovitými stahy drobných tepen a tepének z důsledku chladu a stresu. Dochází k poruše krevního zásobení postižené oblasti a postižená periferní oblast (např. prsty) pak postupně mění barvu na bílou, modrou a červenou. Odhad prevalence – počtu jedinců trpící nemocí vzhledem k počtu jedinců v dané populaci – odhadem činí 5 až 20 %. Nemoc se častěji vyskytuje u žen a problémy jsou obvykle dlouhodobé. Jedním z důvodů postižení tímto onemocněním je např. onemocnění pojivové tkáně nebo hematologické příčiny [11]. Toto onemocnění se projevuje pouze změnou zbarvení prstů a snímky ze senzoru se proto jeví jako nepoškozené (viz obrázek 2.15).



Obrázek 2.14: Snímek pacienta trpící Raynaudovou nemocí [31]



Obrázek 2.15: Raynaudova nemoc zachycená biometrickým senzorem [4]

### 2.3.7 Systémová sklerodermie

Systémová sklerodermie (viz obrázek 2.16) je chronickým autoimunitním onemocněním. Pro více než 90 % pacientů je typické onemocnění Raynaudovou nemocí a postižené jsou více ženy [11]. Původ nemoci je neznámý. Onemocnění se projevuje otoky prstů, malými jizvičkami na špičce prstu nebo sklerodaktylií. Při sklerodaktylii se v tkáních ruky ukládá

vazivo, čímž je narušená pohyblivost a vzniká charakteristická ztuhlá ruka. Nemoc je neléčitelná, ke tlumení projevů se mohou využít imunosupresivní léky tlumící imunitní reakce v těle – kortikoidy. Kožní postižení není pouze jediným projevem systémové sklerodermie – může dojít i k plicnímu postižení nebo postižení oběhového systému [42].



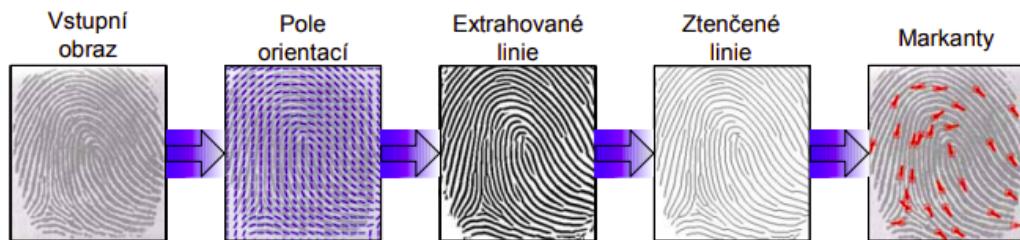
Obrázek 2.16: Snímek pacienta trpící systémovou sklerodermií [11]



Obrázek 2.17: Systémová sklerodermie zachycená biometrickým senzorem [4]

## 2.4 Zpracování otisku prstů

Nejběžnější způsob zpracování otisku prstu je znázorněn na obrázku 2.18.



Obrázek 2.18: Schéma zpracování otisku prstu [12]

Na začátku samotného zpracování otisku prstu je načten **vstupní obraz**. Snímek může být získaný z různých druhů senzorů (např. optické, kapacitní nebo bezdotykové) – vznikne tak digitální otisk prstu. Vstupní obraz obvykle obsahuje vysokou hodnotu šumu. Je nutné dbát na kontrolu živosti prstu nebo na vlivy jako je znečištění a poranění [12].

Následně probíhá výpočet **pole orientací**. Principem tohoto algoritmu je spočítání směru papilární linie pro každý pixel obrazu, kdy se využívá šedotónových hodnot. Následně je vypočtena lokální orientace celého bloku pixelů. Vznikne orientované pole, které může být namapováno na původní otisk prstu [12].

Při **extrahování linií** se používají různé úpravy vstupního obrazu. Patří k nim např. úprava histogramu, se kterou se spjata kontrola kvality obrazu. 2D Gaborovu funkce je jedním z možných způsobů filtrování. Pro samotnou detekci papilárních linií je možné použít schéma RAT nebo detekci papilárních linií dle Honga [12].

Detekce papilárních linií dle Honga vychází z předpokladu, že papilární linie probíhají paralelně vůči sobě a dosahují maximální šedotónové úrovně uprostřed samotné linie – jedná se o černé (tmavé) body. Otisk prstu je násobený dvěma maskami  $h_t$  a  $h_b$ , tyto masky mají

navzájem posunutou fázi o 180 stupňů. Po aplikaci algoritmu je nutné ve výsledném obrazu provést kontrolu a úpravu porušení papilárních linií a zrušení zlomů v průběhu papilárních linií [14].

Detectce papilárních linií pomocí schématu RAT rozděluje obrázek na bloky o velikost  $8 \times 8$  pixelů. V této oblasti je spočtená průměrná úroveň šedé a levá část  $8 \times 4$  pixelů se nastaví na tuto získanou hodnotu. Poté je posunuto okno o 4 body doprava. Pokud je dosažený pravý okraj, tak se posouvá okno o 8 bodů dolů a opět se začíná zleva [12].

Pro extrakci markantů je nutné následně **ztenčit papilární linie** na šířku 1 pixelu. V tomto případu je nutné, aby algoritmus ztenčování byl vše směrový, kdy v žádném směru nesmí ubývat papilární linie. Jednou z vhodných metod je Metoda Emyroglu, která používá dva typy bodů – RMP (Ridge Meeting Point) a RCP (Ridge Continuity Point) [12].

Obvykle se detekují dva základní **typy markantů** – ukončení papilární linie a vidlička. Typů markantů existuje mnohem více, avšak jsou kombinací těchto dvou základních typů. Ke každému markantu je pak uložena jeho pozice X, pozice Y, typ markantu (ukončení nebo vidlička) a gradient – směr pokračování papilární linie [12].

Metod pro porovnání otisků prstů existuje celá řada. Metoda založená na markantech, která je nejčastěji se vyskytující metodou, užívá generování globálního překryvu a následného hledání lokálního překryvu markantů. Také se zavádí oblast tolerance při hledání lokálního překryvu – porovnání. Dalšími způsoby porovnání otisků prstů je metoda založená na korelace, která využívá 2D korelace mezi vstupem a šablonou a metoda porovnání založená na rozpoznání vzorů, která využívá neuronových sítí a je použitelnou i při malé ploše otisků (např. senzory na mobilních telefonech) [14].

#### 2.4.1 Daktyloskopie

Platí následující daktyloskopické zákony:

- Papilární linie jsou pro každého z jedinců unikátní [12].
- Vzor, který tvoří papilární linie, zůstává po celý život jedince relativně neměnným [12].
- Dorůstáním kůže na povrchu prstů jsou papilární linie obnovovány. Papilární linie nemohou být pozměněny nebo odstraněny, pokud není poškozena dermální vrstva kůže. Pokud však k poškození dojde, pak již na tomto místě nedojde k obnově papilárních linií [12].
- Konfigurační typy se mohou individuálně měnit, ovšem jedná se pouze o malé změny, které leží v tolerančních limitech a umožňují systematickou klasifikaci [12].

#### 2.4.2 Typy otisků prstů, druhy markantů a třídy otisků prstů

Otisky prstů jsou děleny na základě způsobu pořízení následovně (viz obrázek 2.19):

- válený otisk (barvený, rolovaný),
- píchaný otisk (živý),
- latentní otisk (skrytý) [12].

Válené otisky jsou získávány způsobem rolování prstu po celé délce, kdy je snahou zachytit všechny detaily hřebenů papilárních linií. Píchané otisky vznikají přiložením bříška prstu k plochému povrchu – rolování není prováděno. Píchané otisky prstů pokrývají menší

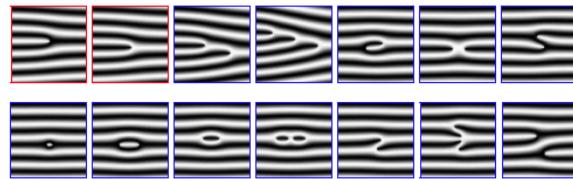
plochu, než válené otisky prstů a typicky nejsou zkresleny právě procesem rolování. Oba tyto typy jsou získávány pomocí senzoru nebo off-line způsobem snímání – např. nanesením inkoustu na bříško prstu. Latentní otisky prstu jsou obvykle získávány z objektů, kterých se daná osoba dotkla – typickým prostředím pro práci s tímto typem otisku prstů je kriminalistika [24]. Ve srovnání s válenými a píchanými otisky, které mají obvykle dobrou kvalitu a obsahují mnoho informace, jsou latentní otisky těžší na správné získání – obvykle se používají speciální fyzikální a chemické metody (např. daktyloskopický prášek, kyanoakrylát nebo ninhydrin) [35].



Obrázek 2.19: Válený, píchaný a latentní otisk [12]

Druhy markantů slouží pro základní rozpoznávání otisků prstů. Existuje mnoho typů markantů (viz obrázek 2.20), avšak se obvykle u přístupových systémů pracuje pouze s typem markantu ukončení nebo vidličky:

- ukončení (*Line Ending*),
- jednoduchá vidlička (*Simple Bifurcation*),
- dvojitá vidlička (*Double Bifurcation*),
- trojitá vidlička (*Triple Bifurcation*),
- hák (*Hook*),
- křížení (*Crossing*),
- boční kontakt (*Side Contact*),
- bod (*Point*),
- interval (*Interval*),
- jednoduchá smyčka (*Single Whorl*),
- dvojitá smyčka (*Double Whorl*),
- jednoduchý most (*Single Bridge*),
- dvojitý most (*Twin Bridge*),
- průsečná linie (*Through Line*) [12].

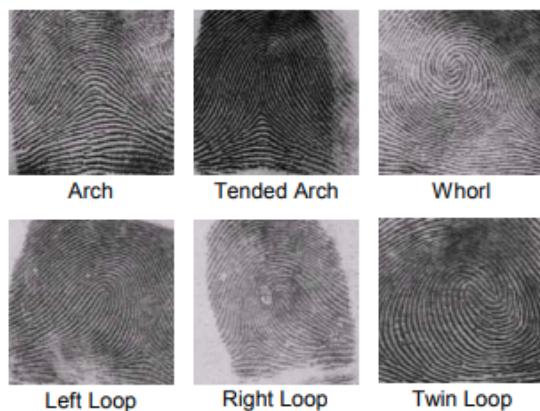


Obrázek 2.20: Typy markantů ve vyjmenovaném pořadí výše [12]

Otisky prstů mohou být také rozřazeny do následujících tříd na základě charakteristického průběhu papilárních linií (viz obrázek 2.21):

- oblouk (*Arch*),
- klenutý oblouk (*Tended Arch*),
- spirála / závit (*Whorl*),
- levá smyčka (*Left Loop*),
- pravá smyčka (*Right Loop*),
- dvojitá smyčka (*Twin Loop*) [12].

V daktyskopických (kriminalistických) systémech se obvykle používají všechny třídy otisků prstů s výjimkou dvojité smyčky [12].



Obrázek 2.21: Typy tříd otisků prstů [12]

#### 2.4.3 Měření kvality otisků prstů

Verifikace otisku prstu je často ovlivněna kvalitou snímku otisku prstu poskytnutým senzorem. Dokonce nejlepší verifikační systémy mají mnohdy problémy s zašuměnými a nekvalitními snímky. V mnoha studiích bylo zjištěno, že změna typu senzoru může mít vliv na rozpoznávání – každý z nich poskytuje snímek jiné kvality. Vlhkost / suchost otisku prstu, šum, nečistoty jak na prstу, tak na senzoru nebo onemocnění mohou mít důležitý vliv na biometrický systém. Existuje mnoho různých způsobů pro měření kvality otisků

prstu – vedle způsobů zaměřených na analýzu šírky a kontinuity papilárních linií se může využít i např. analýza na základě jasnosti papilárních linií [3].

Měření na základě šírky papilárních linií:

- úroveň jistoty orientace – měření koncentrace energie okolo dominantního směru hřebenů s využitím gradientu intenzity,
- směrodatná odchylka výstupu Gaborova filtru – snímek je filtrován Gaborovým filtrem, pro otisk prstu s kvalitními oblastmi bude jeden z mnoha výstupů filtru vyšší jak ostatní, směrodatná odchylka  $m$  výstupu filtru je použita pro ohodnocení kvality [3].

Měření na základě propojení (kontinuity) papilárních linií:

- kvalita lokální orientace – průměrný absolutní rozdíl úhlu papilární linie s okolními bloky obrazu,
- kontinuita pole orientací – u kvalitních snímků předpokládá plynulý průběh hřebenů a údolí v lokálně konstantním směru, analyzuje náhlé změny směru mezi sousedními bloky [3].

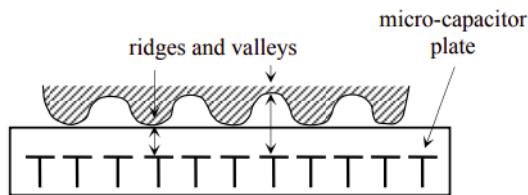
## 2.5 Typy senzorů

První techniky pro získávání otisků prstů byly typu off-line. Prsty byly očerněny inkoustem a přitisknuty nebo rolovány po papírové kartě. Tato karta byla dále naskenována a převedena tak do digitální podoby. Tato technika je stále užívanou v kriminalistice pod názvem daktoloskopická karta. Vedle této metody se však v přístupových systémech užívají techniky pro živé naskenování otisku prstu pomocí biometrického senzoru – tento způsob je mnohem více efektivním, levným a uživatelsky přívětivým. Může být tedy nasazen v široké oblasti aplikací [29]. Existuje mnoho typů senzorů pro snímání otisků prstů, tato kapitola popisuje pouze dosud nejužívanější – kapacitní a optický senzor.

### 2.5.1 Kapacitní senzor

Jedná se o jeden z nejužívanějších typů senzorů. Schéma kapacitního senzoru je znázorněno na obrázku 2.22. Kapacitní senzor je tvořen 2D polem mikrokapacitorů, které jsou umístěny na čipu. Po přiložení prstu se mezi povrchem prstu a každou z křemíkových destiček vytvoří malý elektrický náboj, jehož hodnota závisí na vzdálenosti mezi povrchem prstu a kapacitními destičkami. Hřebeny a údolí papilárních linií mají jinou hodnotu elektrického náboje a odlišnou výslednou intenzitu ve finálním snímku. Každý výrobce kapacitních senzorů má obvykle vlastní nastavenou hodnotu, jak rozlišit hřebeny a údolí papilárních linií, kterou je obvykle obtížné upravit. Kapacitní senzory nelze zneužít přiložením fotografie otisku prstu – senzor měří dané vzdálenosti a nasnímán může být pouze 3D povrch [29].

Povrchovou vrstvu kapacitního senzoru je třeba chránit před chemickými látkami, které jsou přítomné v potu (např. sodík). Povrchová vrstva senzoru nesmí být však příliš silná, protože pak je snížena rozlišovací schopnost senzoru mezi hřebeny a údolími papilárních linií. Povrchová vrstva nesmí být ani příliš tenká, protože pak není odolná vůči mechanickému poškození [29].

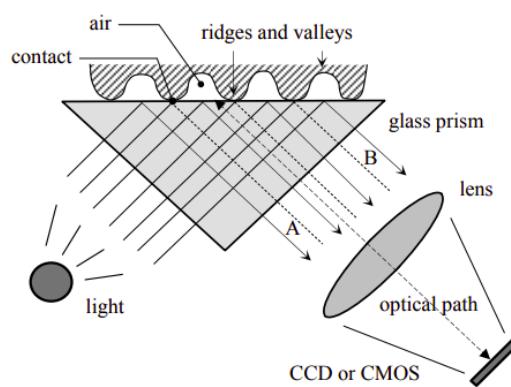


Obrázek 2.22: Schéma kapacitního senzoru [29]

### 2.5.2 Optický senzor

Optický senzor FTIR (Fourier-Transform Infrared Spectroscopy) je nejstarší a dosud nejúžívanější způsob pro získávání snímků otisků prstů. Schéma optického senzoru FTIR je znázorněno na obrázku 2.23. Prst je přiložen na povrch hranolu, vrcholy papilárních linií se povrchu přímo dotýkají a hřebeny papilárních linií zůstávají od povrchu vzdálené. Levá část hranolu je obvykle nasvícena např. LED (Light-Emitting Diode) diodami. Toto světlo, které vstupuje do hranolu je odráženo údolími a náhodně pohlcováno hřebeny papilárních linií. Hřebeny se pak díky neschopnosti odrážet světlo jeví na obrazu jako tmavé oblasti ve srovnání s bílými oblastmi údolí, které světlo odrážejí. Odražené světlo od povrchu hranolu dopadá přes čočku na CCD (Charge-Coupled Device) nebo CMOS (Complementary Metal–Oxide–Semiconductor) kameru, která nasnímá výsledný obraz. Stejně jako kapacitní senzory snímá FTIR senzor 3D povrch a proto nemůže být podvrhnout např. vytisknutou fotografií [29].

Pokud je prst velmi suchý, nemusí docházet k rovnoměrnému kontaktu s povrchem senzoru FTIR. Někteří výrobci FTIR senzorů proto používají typicky silikonovou vrstvu na povrchu hranolu, která zlepšuje optický kontakt kůže s hranolem. Často se z důvodu ceny používá na výrobu hranolu a čoček plast místo skla a CMOS kamery místo nákladnějších CCD kamer [29].



Obrázek 2.23: Schéma FTIR senzoru [29]

### 2.5.3 Daktyloskopická karta

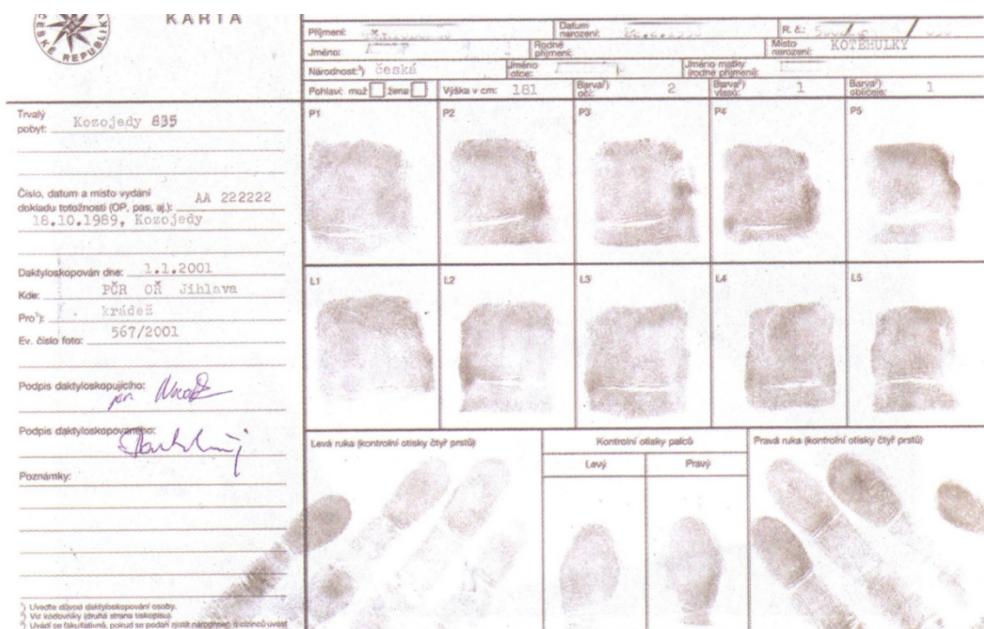
Daktyloskopická karta je off-line způsobem snímání, která se stále užívá v kriminalistice. Obecně obsahuje osobní údaje o daktyloskopované osobě, otisky prstů a další záznamy týka-

jící se daktyloskopování. Příklad daktyloskopické karty Policie ČR je zobrazen na obrázku 2.24.

Daktyloskopická karta obsahuje následující otisky prstů:

- válené otisky pro každý prst (každý umístěn a označen ve zvláštním poli),
- píchaný otisk pro levý a pravý palec (každý umístěn a označen ve zvláštním poli),
- píchané otisky ostatních prstů (otisky umístěny ve dvou polích podle levé nebo pravé ruky) [35].

Při daktyloskopování je dbáno na to, aby kresba papilárních linií na získaných otiscích byla plynulá, dobře čitelná a nepřerušovaná. Také se snažíme o zachycení co největší plochy dané otiskované části prstu [35].



Obrázek 2.24: Příklad daktyloskopické karty Policie ČR [35]

## 2.6 Generování syntetických otisků prstů

Generování syntetických otisků prstů je inverzní biometrickou úlohou. Syntetické otisky mohou být vytvořeny z náhodných bodů nebo ISO šablony a vygenerovaný otisk je obvykle dokonalým. Mohou být užity pro hodnocení výkonu nebo učení modelů neuronových sítí. Výhodou je, že nepodléhají ochraně osobních údajů a proto mohou být snadněji zakomponovány do součástí vědeckých článků a prezentací výsledků [14].

Jelikož syntetický otisk prstu je obvykle dokonalým, při generování je kláden důraz na generování poškození nebo šumu pro větší podobnost s reálným otiskem prstu. Pro generování syntetických otisků bylo využito prostředí generátoru SFinGe, syntetickým otiskům je následně vygenerováno onemocnění generátorem popsaným v bakalářské práci Generování onemocnění kůže do syntetických otisků prstů z SFinGe [43], v bakalářské práci

Detekce a klasifikace poškození otisku prstu s využitím neuronových sítí [48] byl generátor následně upraven a přidána funkce automatického anotování bounding boxů onemocnění v JSON formátu. Na obrázku 2.25 je zobrazen příklad syntetického otisku prstu generovaným generátorem SFinGe, na obrázcích 2.26 a 2.27 se pak nachází příklad generování onemocnění do syntetických otisků prstů.

Syntetické otisky prstů jsou nedílnou součástí pro učení modelů konvolučních neuronových sítí, zvláště při malé velikosti datasetu reálných otisků prstů získaných od pacientů trpících onemocněním. Model je však potřeba otestovat i na reálných otiscích, protože neuronová síť může mít tendenci se přetrénovat na pouze vygenerovaných syntetických otiscích prstů.



Obrázek 2.25: Syntetický otisk prstu vygenerovaný generátorem SFinGe

Obrázek 2.26: Generované onemocnění bradavice do syntetického otisku prstu

Obrázek 2.27: Generované onemocnění dishydroyz do syntetického otisku prstu

## Kapitola 3

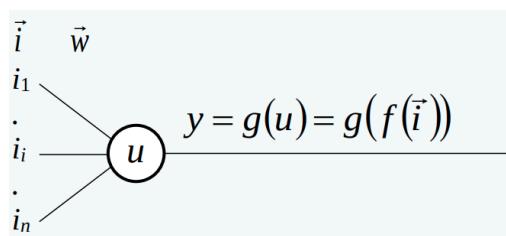
# Detekce a klasifikace pomocí neuronových sítí

Neuronová síť je abstraktním modelem, který simuluje chování lidského mozku. Od roku 1943, kdy bylo poprvé popsáno chování neuronů a implementovaná jednoduchá neuronová síť využívající elektrické obvody, prošla problematika velikým vývojem [16]. Následující kapitola popisuje základní myšlenku funkce perceptronu, způsob trénování a učení neuronové sítě. Následně se práce zaměřuje speciálně na chování konvolučních neuronových sítí. Vedle obecného principu je popsán obsáhlý výčet dnes užívaných architektur – od jedné z prvních VGG-16 až po architektury z minulých let jako je EfficientNet vyvinutá společností Google roku 2019. Následně jsou představeny modely pro detekci objektů – opět jsou kapitoly zaměřeny na popis začátků tohoto problému, až po přístupy z současných let.

### 3.1 Umělý neuron a perceptron

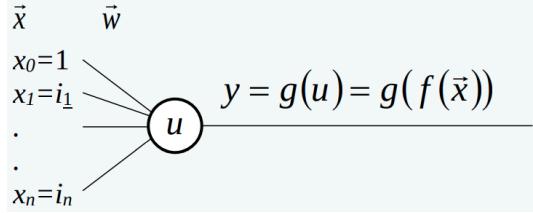
Pro umělý neuron se stal inspirací biologický neuron, který je základem nervové soustavy živých organismů. Lidský mozek obsahuje přibližně  $10^{11}$  neuronů –  $10^{14}$  synapsí, což je druh spojení dvou neuronů sloužící k předávání vztazů. Tento vysoký počet synapsí představuje paměťové schopnosti člověka a proces učení spočívá ve způsobu nastavování synaptických vaz jednotlivých neuronů [52].

Obecné schéma umělého neuronu je znázorněno na obrázku 3.1. Pro umělý neuron se používají na rozdíl od biologického neuronu pouze statické výstupní signály. Proměnná  $u$  představuje vnitřní potenciál,  $f$  bázovou funkci,  $g$  aktivační funkci, vektor  $i_1, i_2, \dots, i_n$  představuje vstup neuronu a vektor  $w_1, w_2, \dots, w_n$  je vektorem vaz [52].



Obrázek 3.1: Schéma umělého neuronu [52]

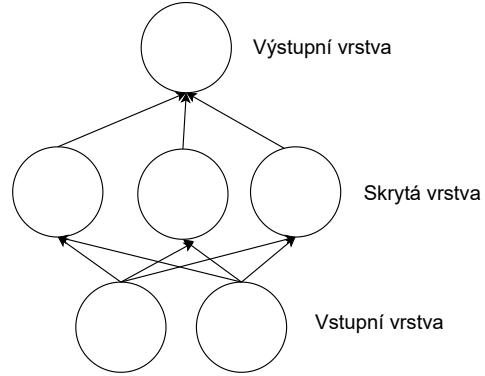
Nejznámější umělý neuron se nazývá perceptron. Schéma perceptronu je znázorněno na obrázku 3.2. Vstupní vektor  $x_0, x_1, \dots, x_n$  má fixně nastavenou nultou složku na hodnotu 1 z důvodu zjednodušení vztahů pro výpočet odezvy a pro učení. Váha  $w_0$  je pak nastavena na hodnotu záporného prahu. Vnitřní stav má následující rovnici:  $w_0x_0 + \sum_{i=1}^n w_i x_i$ , kde  $x_0$  představuje speciální stav bias a  $w_0$  jeho váhu. Na vnitřní reprezentaci je aplikována aktivační funkce a získaný výstup je předán další vrstvě. Při zpětné propagaci jsou pak upraveny váhy neuronové sítě [52].



Obrázek 3.2: Schéma perceptronu [52]

## 3.2 Vícevrstvý perceptron

Vícevrstvý perceptron (Multilayer Perceptron – MLP) je klasickým typem neuronové sítě. Obecné schéma sítě MLP je zobrazeno na obrázku 3.3. Data jsou přivedena na vstup vstupní vrstvě, následně síť obsahuje několik skrytých vrstev, která zajišťuje různé úrovně abstrakce. Výstupní vrstva poté provádí výslednou predikci. Všechny vrstvy jsou plně propojené – každý perceptron z vrstvy je propojen s každým perceptronem následující vrstvy [5].



Obrázek 3.3: Schéma sítě MLP

Sítě MLP jsou vhodnými pro využití např. pro klasifikaci na základě atributů při tabulkových datasetech [5]. Jejich využití na problémy jako je zpracování obrazu je však velmi nevhodné. MLP sítě obsahují velké množství parametrů a nejsou invariantní vůči posunutí. Pokud by byl obraz převeden na 1D vektor a zpracován MLP sítí, tak je ztracena prostorová informace [10].

### 3.3 Aktivační funkce

Aktivační funkce definuje, jak se suma vstupů bude transformovat na výstupy v průběhu průchodu neuronovou sítí. Aktivační funkce jsou obvykle diferencovatelné a různé aktivační funkce mohou být užity v různých částech neuronové sítě. Volba aktivační funkce má vliv na výkon a konečné výsledky zvoleného modelu [8].

Dobrá aktivační funkce by měla mít následující vlastnosti:

- Monotónní funkce – Aktivační funkce by měla být buď zcela nerostoucí nebo neklesající. Pokud aktivační funkce monotónní není, pak zvýšení váhy neuronu může mít menší vliv na minimalizaci hodnoty chyby – průměrné hodnoty loss (viz sekce 3.4) přes celý trénovací dataset [1].
- Diferencovatelnost – U aktivačních funkcí je naším cílem, aby byly diferencovatelné, protože chceme spočítat změnu chyby vzhledem k vahám v daném okamžiku gradientního sestupu (viz sekce 3.5) [1].
- Rychlá konvergence – Aktivační funkce by měla rychle konvergovat k požadovaným hodnotám. Pokud je naším cílem získat např. dva různé výstupy o hodnotách 0 a 1, pak by měla aktivační funkce rychle konvergovat k těmto dvěma hodnotám [1].

Skryté vrstvy obvykle užívají stejnou aktivační funkci, která se obvykle liší od aktivační funkce zvolené pro výstupní vrstvu. Aktivační funkce pro výstupní vrstvu obvykle závisí na typu úlohy, kterou má neuronová síť realizovat (např. úloha binární klasifikace, klasifikace do více tříd) [8].

#### 3.3.1 Aktivační funkce skryté vrstvy

Pro aktivační funkci skryté vrstvy se užívá diferencovatelná nelineární aktivační funkce. Model se pak dokáže naučit složitější funkce, než model užívající lineární aktivační funkci. Řešené problémy reálného světa jsou totiž nelineárními [8].

Mezi známé aktivační funkce skryté vrstvy neuronové sítě patří:

- ReLU – ReLU (Rectified Linear Unit) je nejužívanější aktivační funkci. V kladné části nabývá derivace hodnoty 1, v záporné části je pak derivace nulová (tedy konstantní). Funkce a její derivace jsou monotónní [39]. Funkci ReLU lze matematicky vyjádřit jako:

$$R(z) = \max(0, z) \quad (3.1)$$

- Sigmoida – Výstupem sigmoidy jsou hodnoty z intervalu (0, 1). Čím je výstup vyšší, tím bude hodnota blíže hodnotě 1, čím je výstup nižší, tím více bude hodnota blízká hodnotě 0. Funkce je monotónní, avšak derivace funkce monotónní není [39]. Funkce je vhodná pro predikci pravděpodobnosti díky právě jejímu definovanému oboru hodnot. S pojmem sigmoidy však souvisí problém mizejících gradientů. Jedná se o problém, kdy přidání vrstev využívající tuto aktivační funkci má za následek, že gradient loss funkce dosáhne k nule. Trénování modelu se tedy stává obtížným, což může vést k celkové nepřesnosti celé sítě. Vysoké změny vstupu funkce vedou k malým změnám na výstupu – hodnota derivace se blíží k nule. Tento problém je specifický pro sítě o více vrstvách [50]. Graf sigmoidy a její derivace je znázorněný na obrázku 3.4. Sigmoida je definována následujícím předpisem:

$$\sigma(z) = \frac{1}{1 + e^{-z}} \quad (3.2)$$

- Tanh – Rozsah výstupů je v intervalu  $(-1, 1)$ . Pro záporné vstupy je výsledek funkce negativním. Funkce je monotónní, derivace monotónní není [39]. Tato aktivační funkce je definována následovně:

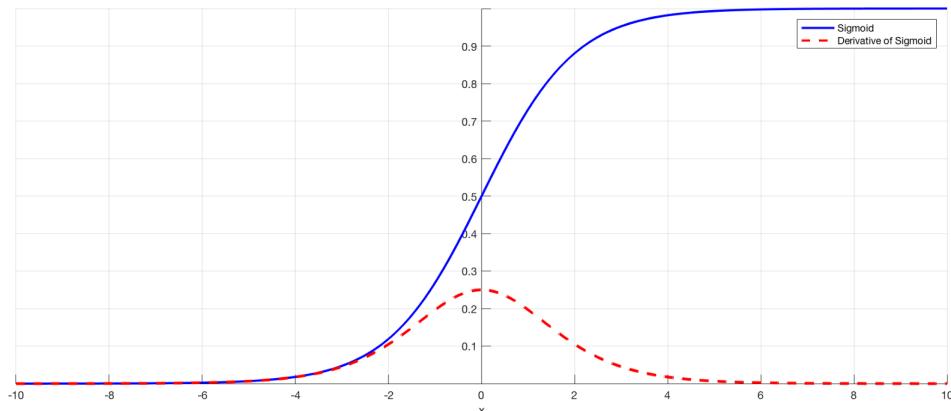
$$\sigma(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}} \quad (3.3)$$

### 3.3.2 Aktivační funkce výstupní vrstvy

Aktivační funkce výstupní vrstvy je závislá na řešeném problému. Obvykle je její výběr spjatý i s výběrem dané loss funkce (viz sekce 3.4), které jsou také charakteristické pro různé zvolené úlohy.

Aktivačními funkcemi užitými pro výstupní vrstvy neuronové sítě jsou:

- Linearita – Lineární aktivační funkce nemění váhovanou sumu vstupu a přímo vrací danou hodnotu. Běžně se užívá při regresi – predikci numerické hodnoty [8].
- Sigmoida – Sigmoida je vhodnou aktivační funkcí např. při binární klasifikaci [8].
- Softmax – Výstupem softmax je vektor hodnot, jejichž součet je jednotkový – jedná se o pravděpodobnost příslušnosti výstupu do dané třídy. Je vhodnou použít při klasifikaci do více než dvou tříd [8].



Obrázek 3.4: Sigmoid funkce a její derivace [50]

## 3.4 Loss funkce

S problematikou loss funkce souvisí pojem objektivní funkce – funkce, která má za úkol vyhodnocovat kandidátní řešení (např. množinu vah) a úkolem modelu je maximalizovat nebo minimalizovat právě tuto objektivní funkci. U neuronových sítí se typicky snažíme o minimalizaci chyby – objektivní funkce je zde právě nazývána loss funkcí [7]. Stejně jako u aktivační funkce výstupní vrstvy výběr dané loss funkce závisí na vybrané úloze, kterou řešíme:

- MAE (Mean Absolute Error) – Jedná se o nejednodušší, avšak robustní loss funkci užitou pro regresi. MAE je ideální i při operaci s odlehlymi hodnotami. Výpočet je počítán jako průměr druhých mocnin mezi predikovanými a skutečnými hodnotami [7].

- Cross-entropy loss – Tato loss funkce je užívanou při problematice binární klasifikace a klasifikace do více tříd. Při binární cross-entropy loss je predikovaná hodnota porovnána se skutečnou třídou (hodnoty 0 nebo 1) a vypočítané skóre penalizuje pravděpodobnost na základě vzdálenosti od očekávané hodnoty. Penalizace je logaritmická, kdy pro malé vzdálenosti od skutečné hodnoty nabývá malých hodnot a naopak velkých hodnot pro velké rozdíly. Binární cross-entropy loss je vypočítaná jako průměr aplikace cross-entropy přes všechny vzorky. Pro aplikaci cross-entropy loss při klasifikaci do více tříd se používá one-hot kódování pro dané třídy. Skóre penalizace je vypočítané přes každou hodnotu vektoru získaným one-hot kódováním a zprůměrováno přes všechny vzorky [7].

## 3.5 Trénování neuronové sítě

Cílem algoritmu učení je minimalizace objektivní funkce, což je suma výsledků loss funkce přes všechny prvky datasetu. Snahou je nalézt takové parametry, které minimalizují hodnotu objektivní funkce. Gradienty jsou derivacemi objektové funkce s konkrétními parametry a počítají se pro konkrétní parciální derivaci – vektor reálných čísel. Minimalizuje se objektivní funkce  $J(\theta)$  parametrizovaná parametry modelu  $\theta$  pomocí aktualizace parametrů v opačném směru gradientu objektivní funkce  $\nabla_{\theta}J(\theta)$ . Parametr  $\eta$  udává learning rate – velikost kroku pro dosažení (lokálního) minima [38]. Při zpětné propagaci se aktualizují váhy neuronové sítě.

### 3.5.1 Batch gradient descent

Batch gradient descent počítá gradient objektivní funkce s využitím celé datové sady. Pro vypočítání jedné aktualizace parametrů je tedy zapotřebí vypočítání gradientů pro celý dataset:

$$\theta = \theta - \eta \nabla_{\theta}J(\theta) \quad (3.4)$$

Tento výpočet však může být velmi pomalým a neřešitelným pro datasety, které se svým rozsahem nevezjdou do paměti. Tato metoda nám také neumožňuje online aktualizaci modelu – např. při rozšíření datasetu v průběhu trénování [38].

### 3.5.2 Stochastic gradient descent

Stochastic gradient descent počítá aktualizaci parametrů pro každý prvek z trénovací sady  $x^{(i)}$ , který spadá do třídy  $y^{(i)}$ :

$$\theta = \theta - \eta \nabla_{\theta}J(\theta; x^{(i)}; y^{(i)}) \quad (3.5)$$

Tato metoda řeší problém provádění redundantních výpočtů pro rozsáhlé datasety, který je typickým pro přístup batch gradient descent. Stochastic gradient descent je také mnohem rychlejší, umožňuje online aktualizaci modelu a vyhledávání nových možných lepších potenciálních lokálních minim. Při snižování parametru learning rate vykazuje stochastic gradient descent stejně konvergenční chování jako batch gradient descent a konverguje k lokálnímu nebo globálnímu minimu v závislosti na typu úlohy [38].

### 3.5.3 Mini-batch gradient descent

Mini-batch gradient descent provádí aktualizaci pro každý mini-batch  $n$  trénovacích vzorků:

$$\theta = \theta - \eta \nabla_{\theta} J(\theta; x^{(i:i+n)}; y^{(i:i+n)}) \quad (3.6)$$

Snižování počtu aktualizací může vést ke stabilnější konvergenci a při této metodě může být využito maticových optimalizací typických pro současné knihovny strojového učení – výpočet mini-batch gradient descent je pak efektivním. Velikost mini-batch se pohybuje v širokém rozmezí (např. mezi 50 a 256) a liší se na základě řešené úlohy i výpočetního prostředí [38].

## 3.6 Konvoluční neuronové sítě

Konvoluční neuronové sítě jsou speciálním typem neuronových sítí, které jsou v dnešní době vysokým způsobem užívané při zpracování obrazu a řeči. Jejich název je odvozen od matematické operace konvoluce, která je v jejich architektuře použita. Jelikož konvoluční neuronové sítě explicitně předpokládají, že vstupy jsou např. ve formě obrázků nebo zvuku, tak můžeme zakódovat určité vlastnosti do jejich architektury [10]. Výrazně se snižuje počet parametrů modelu oproti tzv. sítím MLP (Vícevrstvý perceptron), které byly popsány v sekci 3.2.

Obecně se konvoluční neuronová síť skládá z:

- vstupní obraz,
- konvoluční vrstvy,
- nelineární aktivační funkce,
- pooling vrstvy,
- plně propojené vrstvy [6].

Konvoluční neuronová síť dokáže zohlednit místní konektivitu – konvoluční filtr je postupně posouván po celém vstupu (v našem případě obrázku) podle určité velikosti a kroku. Dochází tedy k nalezení a porovnávání vzorů bez ohledu na to, kde se dané vzory v obrazu nachází. Váhy jsou sdílené a trénování je obecně efektivnějším. Každý uzel není propojen s každým dalším uzlem [10].

Na každý pixel obrazu je obecně aplikováno více konvolučních filtrov. První vrstvy konvoluční neuronové sítě obsahují jednoduché filtry reagující pouze na hrany, filtry v jedné z posledních vrstev nám umožňují získat z obrazu detailnější informace (např. filtry reagující na barvy). Aplikací každého filtrov je získána aktivační mapa [10].

Další nedílnou součástí konvolučních neuronových sítí je pooling vrstva. Její parametry jsou nižší, jak velikost získané aktivační mapy konvoluční vrstvou – např. velikost  $2 \times 2$  a parametr stride (krok posunu konvolučního filtrov) 2 pixely. Při tomto nastavení pooling vrstvy je aktivační mapa dvakrát zmenšena – každá dimenze je poloviční, počet pixelů v aktivační mapě je pak zmenšen čtyřikrát – např. při vstupu aktivační mapy o velikosti  $6 \times 6$  (36 pixelů) je po aplikaci pooling vrstvy nová velikost aktivační mapy  $3 \times 3$  (9 pixelů) [6]. Pooling vrstvy jsou neučitelnou částí konvolučních neuronových sítí. Nejznámějšími operací spojenou s pooling vrstvou je max pooling umožňující nalezení hodnot lokálních maxim.

Pooling vrstva efektivně snižuje prostorové dimenze vstupu a pracuje s myšlenkou, že relativní umístění prvku (vlastnosti) vzhledem k ostatním prvkům (vlastnostem) je důležitější jak jeho samotné přesné umístění [6].

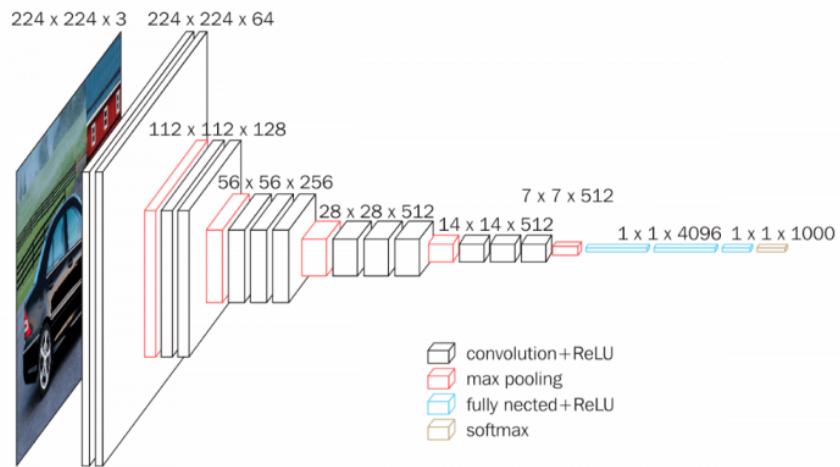
Obraz je poté po dostatečném zpracování s využitím konvolučních vrstev převeden na 1D vektor a přiveden na vstup plně propojených vrstev. Tyto plně propojené vrstvy pak tvoří výstupní část celého modelu a je určen výsledek na základě dané úlohy.

### 3.7 Existující architektury neuronových sítí pro zpracování obrazu

Páteřní konvoluční neuronová síť je základní částí každého detekčního modelu. V průběhu zpracovávání vstupu konvoluční neuronovou sítí jsou extrahovány rysy různých úrovní – v prvních vrstvách sítě se získávají rysy nízké úrovně, které mají vysoké rozlišení, avšak malý počet informací. Naopak na konci jsou získány rysy s nízkým rozlišením, ale vysokou hodnotou informace. Běžné principy detekce objektů jako Faster R-CNN nebo SSD užívají jako páteřní síť např. ResNet-50 nebo MobileNet. Jeden z nejnovějších přístupů EfficientDet užívá však speciální EfficientNet síť.

#### 3.7.1 VGG-16

VGG-16 je jednoduchou a široce užívanou konvoluční neuronovou sítí představenou roku 2014. Není však příliš výpočetně optimalizovanou. Její architektura je zobrazena na obrázku 3.5. Vstupem sítě je barevný obrázek o rozměrech  $224 \times 224$ . Všechny konvoluční vrstvy jsou výpočetně stejně složité a z rozlišení se složitost přesouvá do parametru počtu kanálů – při zmenšení rozlišení se dvakrát zvýší počet kanálů. Velikost jádra konvoluční vrstvy je  $3 \times 3$ . Síť má v hlubších vrstvách více nelinearit a nižší počet parametrů. Vstup první plně propojené vrstvy má dimenze  $7 \times 7 \times 512$ . Po plně propojených vrstvách následuje predikce s využitím aktivační funkce softmax např. do 1000 tříd pro dataset ILSVRC [36].



Obrázek 3.5: Architektura VGG16 [36]

### 3.7.2 ResNet-50

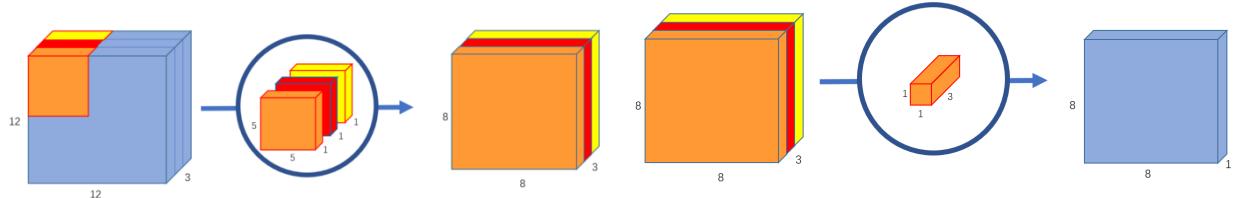
ResNet-50 je variantou modelu ResNet. Je složen z 48 konvolučních vrstev, jedné max pool vrstvy a jedné average pool vrstvy – celkově tedy obsahuje 50 vrstev. Tento model využívá reziduálních připojení - připojení vstupu ResNet-50 modelu na výstup, kdy je spočítána suma právě tohoto vstupu (identity) a zpracovaného vstupu. Výhodou tohoto dodatečného připojení identity na výstup je, že modelu nejsou přidány žádné nové dodatečné parametry. Tento princip řeší problém tendence vyšší chybovosti u modelů s vyšším počtem vrstev. Také je vyřešený problém mizejících gradientů. Dalšími známými variantami ResNet jsou např. ResNet-101 (obsahující 101 vrstev) nebo ResNet-152 (obsahující 152 vrstev) [25].

### 3.7.3 MobileNet

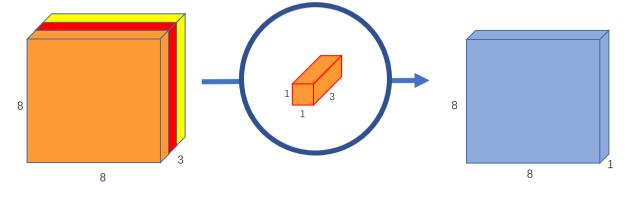
Architektura MobileNet byla poprvé představena roku 2017. Svojí myšlenkou se vyhýbá trendu tvoření komplikovanějších sítí o více vrstvách pro dosažení vyšší přesnosti. Místo toho se autoři zaměřili na výpočetně limitované platformy – na tvorbu vhodné sítě pro mobilní a vestavěné aplikace, robotiku, nebo samořídící auta. MobileNet využívá tzv. hloubkově oddělitelné konvoluce (Depthwise Separable Convolutions) [22].

Princip základních kroků hloubkově oddělitelné konvoluce je zobrazen na obrázcích 3.6 a 3.7. Při hloubkově oddělitelné konvoluci je v první části provedena operace konvoluce bez změny hloubky – každé z jader je aplikováno pouze na 1 odpovídající kanál vstupu (nikoliv na všechny jako u běžné operace konvoluce). Následně probíhá bodová konvoluce – rozměry velikosti jádra pro bodovou konvoluci jsou  $1 \times 1$  [49].

Výhodou tohoto přístupu oproti klasické operaci konvoluce je rapidní snížení počtu aplikace operace násobení – např. při velikosti batche obrázků o velikosti 256 s rozměry jednoho barevného snímku  $12 \times 12 \times 3$  pixely se jedná o snížení operací násobení z hodnoty 1228800 na 52952. Síť je pak schopná zpracovat více dat za nižší čas [49].



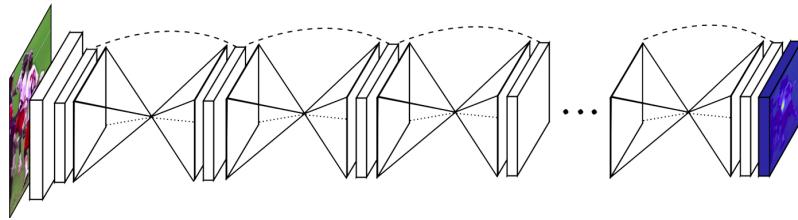
Obrázek 3.6: Hloubkově oddělitelná konvoluce – použití tří konvolučních jader [49]



Obrázek 3.7: Bodová konvoluce - transformace 3-kanálového obrazu na 1-kanálový [49]

### 3.7.4 Hourglass

Hourglass je typem konvoluční neuronové sítě představené v roce 2016. Poprvé byla využita na problematiku predikce pozice lidského těla. Hlavním úkolem této architektury je zachytit informaci přes různá rozlišení vstupního obrazu. Cílem této symetrické topologie (viz obrázek 3.8) je převést obraz na velmi nízké rozlišení pomocí pooling vrstev, následně probíhá zvětšení rozlišení. Při malém rozlišení síť dokáže efektivně pracovat s velkým kontextem, avšak se ztrácí informace o pozici. Při zvyšování rozlišení se naopak ztrácí hodně kontextu. Proto síť kombinuje rysy přes více rozlišení [32].



Obrázek 3.8: Symetrická topologie Hourglass [32]

### 3.7.5 EfficientNet

Pokud chceme dosáhnout vyšší přesnosti modelu, je běžným způsobem škálovat model za pomoci využití větší páteřní sítě (škálování do hloubky) nebo zvětšením rozlišení vstupního obrázku (škálování rozlišení). Přestože tyto metody individuálního škálování pouze na základě jednoho z faktorů zvyšují přesnost modelu, obvykle vyžadují zdlouhavé ruční ladění a u sítí obsahujících mnoho vrstev se může vyskytnout problém saturace a mizejících gradientů. Například ResNet dokáže být škálován z verze ResNet-18 na verzi ResNet-200 zvýšením počtu vrstev, řeší problém mizejících gradientů na základě funkce identity, avšak se jedná o rozsáhlou síť s mnoha parametry a dlouhým trénovacím časem. Cílem společnosti Google bylo navrhnout síť o méně parametrech, která dokáže být nasazená např. na mobilních zařízeních nebo méně výkonných strojích a je méně náročnou na výpočetní zdroje a trénovací čas. V roce 2019 byla poprvé prezentována EfficientNet se základním principem efektivního složeného škálování modelu [45, 44].

Můžeme provést následující typy škálování:

- škálování rozlišení – pokud má obraz vyšší rozlišení, můžeme z něj získat komplexní rysy a detailnější textury,
- škálování do hloubky – pokud chceme zpracovávat více komplexní rysy (obraz s vyšším rozlišením), tak síť musí obsahovat více vrstev,
- škálování do šířky – pro zpracování komplexnějších rysů a detailnějších struktur v obrazu síť potřebuje více kanálů (více získaných map rysů) [45, 44].

EfficientNet umožňuje provádět všechny tři formy škálování zároveň pomocí heuristiky, což se dle experimentů projevilo na zlepšení celkového výkonu neuronové sítě ve srovnání s individuálním škálováním vždy pouze na základě jednoho z faktorů užívaným v minulosti. Bylo zavedeno několik verzí modelu B0 – B7, kdy B0 je základním modelem, od kterého jsou všechny další modely odvozeny. Velikost vstupního obrazu pro model B0 je  $224 \times 224$  pixelů. Architektura B0 byla navrhnuta s využitím NAS (Neural Architecture Search) s využitím optimalizace přesnosti a efektivity (FLOPS) [45, 44].

NAS je přístupem, který slouží k nalezení co nejlepší architektury neuronové sítě pro daný problém. Jedná se o automatizaci procesu ručního ladění modelu a snahu o nalezení složitějších architektur na základě učení. Ve vyhledávacím prostoru je testováno velké množství možných architektur a je vybrán model maximalizující fitness funkci – model odpovídající nejlépe řešení daného problému [18].

Při škálování je pracováno s koeficienty škálování  $\alpha = 1,2$  (škálování do hloubky),  $\beta = 1,1$  (škálování do šířky) a  $\gamma = 1,15$  (škálování rozlišení). Pokud se tedy rozlišení obrazu zvýší o 15 %, pak se počet vrstev sítě má zvýšit o 20 % a počet kanálů o 10 % [45, 44].

## 3.8 Přístupy pro detekci objektů

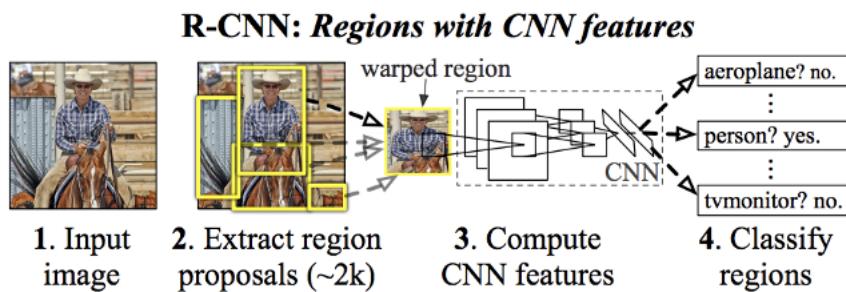
Problematika detekce objektů je jednou ze základních úloh strojového učení. Úkolem je vykreslit predikovaný bounding box okolo daného objektu. Naivním přístup k detekci by mohl být založen na pohyblivém okně přes dané pixely obrazu, kdy každé podokno je ohodnocené na základě predikce modelu skóre pro danou třídu. Podokna s vysokým skórem by pak byly částmi výsledného bounding boxu. Tento přístup je však výpočetně náročný. Proto se užívá jiných moderních přístupů.

### 3.8.1 R-CNN

R-CNN (Region Based Convolutional Neural Network) byla poprvé představena v roce 2014 [20]. Tento přístup kombinuje extrakci rysů z vybraných oblastí obrazu a konvoluční neuronovou síť. Je vybráno pouze např. 2000 oblastí na základě algoritmu selektivního výběru. Pro predikci není tedy využit velký počet oblastí, jako by tomu mohlo být u naivního algoritmu využívajícího posuvného okna, ale pouze zmenšený počet oblastí. Algoritmus selektivního výběru užívá greedy algoritmus (výběr lokálního optimu v každé fázi) pro rekursivní spojování podobných oblastí do regionů větší velikosti [17].

Princip R-CNN modelu je znázorněn na obrázku 3.9. Regiony jsou poté konvertovány na rozměry na základě modelu konvoluční neuronové sítě, která extrahuje vektory rysů o 4096 dimenzích. Extrahované vektory rysů jsou poté klasifikovány pomocí klasifikátoru SVM (Support Vector Machines) do daných tříd. Algoritmus vedle predikce dané třídy objektu také predikuje čtyři hodnoty offsetu pro získání vyšší přesnosti polohy bounding boxu [17].

Nevýhodami daného přístupu je stálá vysoká doba trénování a celková nepoužitelnost nasazení v reálném čase. Regionů je stálá vysoký počet, což vyžaduje tisíce zpětných průchodů modelem pro detekci objektů [17].

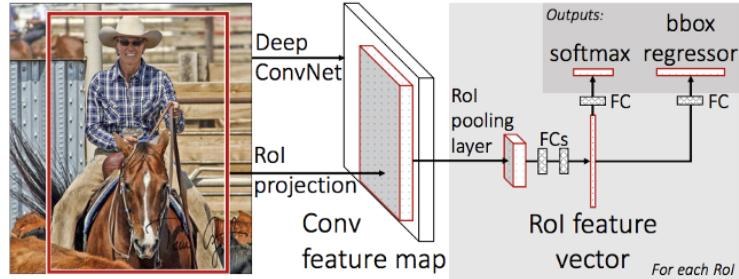


Obrázek 3.9: Princip R-CNN [17]

### 3.8.2 Fast R-CNN

Ross Girshick, jeden z předních autorů přístupu R-CNN, v roce 2015 vyřešil některé z nevýhod předešlého přístupu [19]. V R-CNN byl vysoký počet regionů postupně zpracováván konvoluční neuronovou sítí, v Fast R-CNN je na vstup modelu přiveden celý vstupní obraz. Je získána konvoluční mapa rysů, ze které identifikujeme vhodné regiony, které jsou s využitím ROI pooling vrstvy převedeny na pevnou délku. Tyto vektory mohou být klasifikované pomocí MLP sítě a softmax aktivační funkce. Vedle predikované třídy získáváme i hodnoty offsetu pro daný bounding box [17].

Fast R-CNN je znatelně rychlejší jak R-CNN, protože z daného obrazu nemusí být po každé vybíráno několik tisíc regionů, které jsou následně zpracovávány modelem. Na obrázku 3.10 je znázorněn její princip. Operace konvoluce je aplikována pouze jednou pro každý obraz, ze které je poté zpracována mapa rysů. Algoritmus selektivního výběru je však stále časově náročným procesem, který je fixně implementovaný a není schopný procesu učení a tedy i možného zlepšení přesnosti při hledání vhodných regionů v obrazu [17].

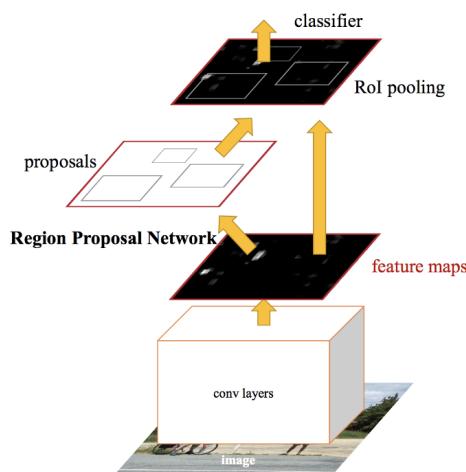


Obrázek 3.10: Princip Fast R-CNN [17]

### 3.8.3 Faster R-CNN

Faster R-CNN je metodou představenou v roce 2016, která řeší hlavně nevýhody R-CNN a Fast R-CNN z hlediska využívání algoritmu selektivního výběru [37]. Jelikož se jedná o algoritmus, který ovlivňuje výkon neuronové sítě, bylo využito myšlenky zrušení tohoto algoritmu. Princip modelu Faster R-CNN je zobrazen na obrázku 3.11. Celý obraz je opět převeden na vstup konvoluční neuronové sítě, kdy je vygenerována mapa rysů. Je ovšem využito i druhé oddělené sítě, která predikuje vhodné oblasti pro následnou detekci. Tyto oblasti jsou po změně dimenzí při průchodu ROI pooling vrstvou využity pro následnou klasifikaci obrazu v dané oblasti a je opět predikována i odchylka pro bounding boxy [17].

Faster R-CNN je díky vylepšením nedostatků přístupů R-CNN a Fast R-CNN znatelně rychlejší a je tedy vhodnou pro nasazení v reálných aplikacích [17].

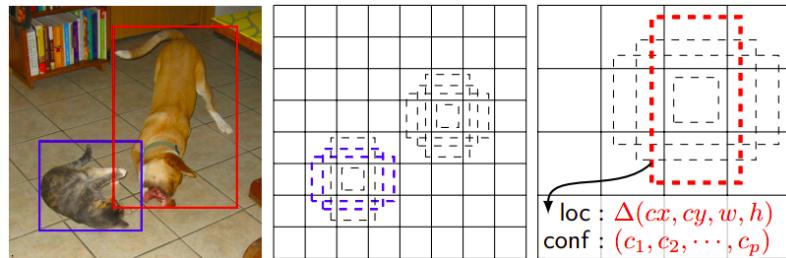


Obrázek 3.11: Princip Faster R-CNN [17]

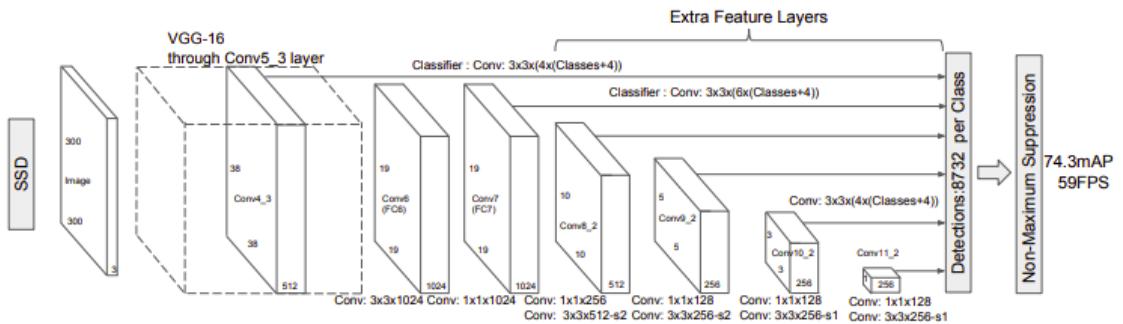
### 3.8.4 SSD

SSD (Single Shot Detector) je přístupem pro detekci objektů představeným roku 2016, který je založen na odlišném přístupu než Faster R-CNN. Principem Faster R-CNN je zpočátku hledat vhodné regiony a poté u těchto vybraných regionů provést detekci a klasifikaci. Fáze hledání vhodných regionů je u metody SSD zrušena. Místo toho se v několika konvolučních vrstvách o různém rozlišení provádí predikce bounding boxů a třídy pro každý z daných objektů (viz schéma architektury 3.13). Pro každý objekt je při trénování predikováno 8732 bounding boxů – díky tomuto vysokému množství získáme větší pokrytí dané lokace v obrazu a tak možnost nalezení bounding boxu, který má nejvyšší překryv se skutečným anotovaným bounding boxem objektu (viz obrázek 3.12). SSD následně vybírá právě na základě skóre překryvu bounding boxy se skóre vyšším jak 50 %. Následně jsou eliminovány duplicitní bounding boxy a je vybráno pouze 100 bounding boxů s nejvyšším skóre jistoty pro každý obrázek [28, 53].

Síť díky predikcím polohy bounding boxů a příslušnosti do třídy v různých měřítkách dokáže predikovat objekty rozličných velikostí. SSD je přístupem, který zvláště díky více vrstvám pro predikci v různých měřítkách objektu, dosahuje dobré přesnosti za snížení trénovacího času a může být tedy nasazeným v reálných aplikacích. Také dosahuje dobré přenosnosti i na datasetech o nižším rozlišení [28, 53].



Obrázek 3.12: Princip získání většího počtu bounding boxů pro daný objekt [28]



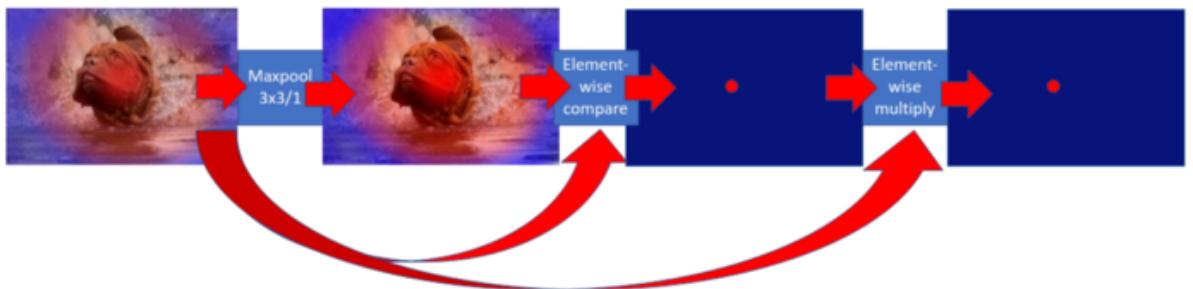
Obrázek 3.13: Architektura modelu SSD [28]

### 3.8.5 CenterNet

Mnoho přístupů k detekci (např. SSD) využívá predikování tisíců bounding boxů pro jeden objekt a vybírá bounding boxy s vysokou hodnotou překryvu. Tento přístup má následující

nevýhody – složitost je úměrná druhé mocnině počtu predikcí a model je nuten vypočítávat předpovědi pro všechny bounding boxy, což je operace náročná na čas i výpočetní prostředky. CenterNet využívá jiný způsob predikce – pro každý objekt je vytvořena tzv. *center confidence heatmap*, kdy ve středech objektu nabývá hodnota distribuce funkce svého maxima (a tedy nejvyšší jistoty). Vedle generování této mapy jsou také jako u jiných detektorů predikovány hodnoty pro rozměry bounding boxů a offsety související se středem bounding boxů predikovanými mapou jistoty. CenterNet predikuje *center confidence heatmap* pro každou třídu [2].

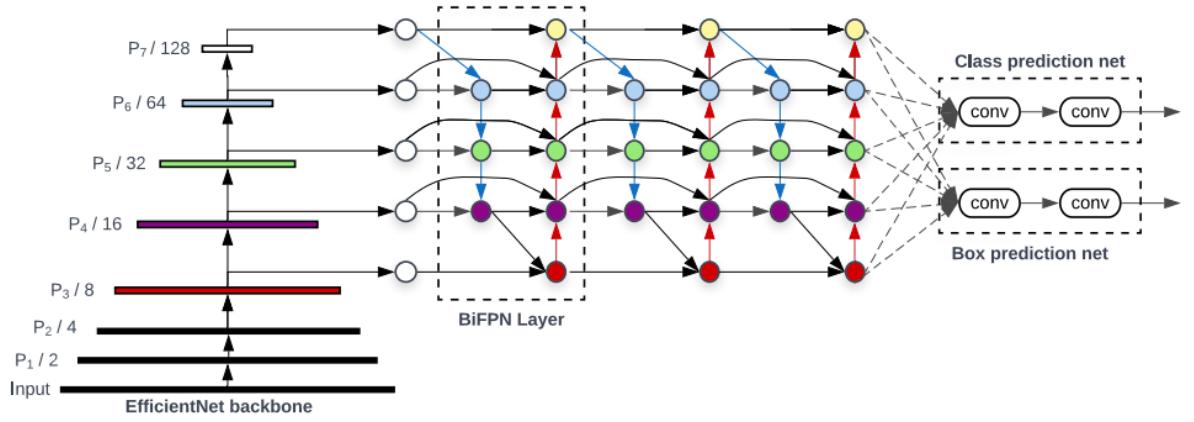
Právě užití *center confidence heatmap* nám umožňuje okamžitě odstranit irrelevantní predikce bez nutnosti predikování celého bounding boxu, princip analýzy je zobrazen na obrázku 3.14. Na mapu je aplikován max pooling a okolo samotného centra je vytvořen malý plochý region. Následně se boolovsky porovnávají jednotlivé prvky mezi vstupem a získaným výstupem po aplikaci max pooling – je vytvořena nová tepelná mapa, kdy jsou lokální maxima nastavena na hodnotu 1, zbytek je nastaven na hodnotu 0. Tato mapa se vynásobí s původním vstupem a výsledné skóre jistoty je vyprahováno. Tímto způsobem se dále pracuje jen s bounding boxy, které odpovídají lokálním maximům [2].



Obrázek 3.14: Princip analýzy *center confidence heatmap* [2]

### 3.8.6 EfficientDet

EfficientDet je modelem vyvinutým společností Google v roce 2020. Díky svému nízkému počtu parametrů se řadí k sítím s nižšími nároky na výpočetní zdroje a vyšší rychlostí trénování. Pro páteřní síť není užita např. MobileNet nebo ResNet-50 jako u předchozích metod, ale opět společností Google vyvinutá síť EfficientNet. EfficientDet neužívá FPN (Feature Pyramid Network) jako běžné algoritmy pro detekci. Při užití FPN jsou všechny rysy různých rozlišení posuzovány stejným způsobem a je jim tedy přikládána totožná důležitost. FPN je nahrazeno BiFPN (Weighted Bi-directional Feature Pyramid Network), který umožňuje analýzu a získání pouze důležitých rysů. Architektura modelu EfficientDet je znázorněna na obrázku 3.15. Každá z komponent sítě – páteřní síť EfficientNet, BiFPN a síť pro detekci bounding boxu a predikci třídy objektu – má jeden ze složených faktorů škálování a je prováděno složené škálování rozlišení, škálování do hloubky a škálování do šířky pro každou z těchto částí [46, 47].



Obrázek 3.15: Architektura EfficientDet [46]

Verze EfficientNet jsou značeny v rozsahu B0 – B7 a verze EfficientDet označením D0 – D7. Každá z verzí EfficientDet má také přiřazenou hodnotu  $\phi$ , která figuruje ve výpočtu dalších charakteristik celého modelu [46, 47].

## Kapitola 4

# Návrh řešení a implementace

Tato kapitola se zabývá návrhem a implementací řešení. V první řadě jsou popsány dvě nejdůležitější části práce – dataset otisků prstů a použité modely. Následně je kladen důraz na předzpracování vstupního otisku prstu ještě na přivedení na vstup samotné konvoluční neuronové sítě. Všechna onemocnění také musí být anotována bounding boxy s danými třídami.

### 4.1 Databáze otisků prstů

Databáze onemocnění otisků prstů získaných od pacientů je součástí stále pokračujícího výzkumu ve výzkumné skupině STRaDe na Fakultě informačních technologií v Brně. Ve své práci se zabývám v současné době pouze následujícími onemocněními – atopickým ekzémem, bradavicemi, psoriázou a dyshidrózou. Kompletní databáze však obsahuje další onemocnění, z nichž mnohá byla popsaná v kapitole 2.3. Databáze však není příliš rozsáhlou co se týče jednotlivých kategorií a proto je nutné dataset rozšířit i o syntetické otisky prstů vygenerované generátorem. Tvorba generátoru syntetických otisků prstů a onemocnění je dalším předmětem stále probíhajícího rozšiřování a výzkumu.

Onemocnění	Počet snímků v datasetu
Atopický ekzém	1107
Bradavice	96
Psoriáza	326
Dyshidróza	247

Tabulka 4.1: Počty snímků pro onemocnění reálných pacientů

Další databází živých otisků prstů je databáze americké organizace NIST – Národní institut standardů a technologie. Jedná se konkrétně o NIST Special Database 4, která obsahuje 4000 snímků získaných převážně z daktyloskopických karet s neznámými běžnými poškozeními.

## 4.2 Anotace dat

Každý snímek je anotován v XML formátu. Pro anotaci reálných otisků prstů s onemocněními byl užit program labelImg dostupný pod MIT licencí<sup>1</sup>. Níže je uveden jeden z příkladů anotace snímku s onemocněním atopického ekzému:

```
<annotation>
    <folder>atopic_png</folder>
    <filename>atopic_eczema_FP_00008.png</filename>
    <path>D:\images\atopic_eczema_png\atopic_png\atopic_eczema_FP_00008.png</path>
    <source>
        <database>Unknown</database>
    </source>
    <size>
        <width>527</width>
        <height>597</height>
        <depth>1</depth>
    </size>
    <segmented>0</segmented>
    <object>
        <name>atopic</name>
        <pose>Unspecified</pose>
        <truncated>0</truncated>
        <difficult>0</difficult>
        <bndbox>
            <xmin>13</xmin>
            <ymin>297</ymin>
            <xmax>159</xmax>
            <ymax>332</ymax>
        </bndbox>
    </object>
</annotation>
```

Snímky získané generátorem obsahují anotaci bounding boxů onemocnění v JSON formátu:  
[{"Type": "wart", "p1": [126, 232], "p2": [198, 304]}]

Pro potřeby práce byly tyto anotace v JSON formátu převedeny do XML formátu. XML formát mimo jiné obsahuje důležité informace ohledně pozice jednotlivých bounding boxů (`<xmin>`, `<ymin>`, `<xmax>` a `<ymax>`) a název dané třídy onemocnění (`<name>`).

## 4.3 Předzpracování otisků prstů

Všechny otisky prstu byly před samotným začátkem trénování konvoluční neuronové sítě předzpracovány. Šedotónové snímky jsou normalizovány (viz obrázek 4.2). Jedná se o metodu, která má za účel snížit rozptyl šedotónových hodnot podél hřebenů a údolí a usnadnit tak další zpracování [21]. Papilární linie jsou obecně po aplikaci normalizace více viditelné a lépe rozeznatelné.

---

<sup>1</sup><https://github.com/tzutalin/labelImg>

Následně je obraz prahován pomocí metody Otsu (viz obrázek 4.3). Rozdílem této metody narození od globálního prahování je princip výběru prahu. Při globálním prahování je tento práh konstantní pro všechny snímky (obvykle hodnoty 127), Otsu metoda tento práh zjišťuje speciálně pro každý snímek, konkrétně podle histogramu. Je tedy nalezen optimální práh díky principu minimalizace váženého vnitrotřídního rozptylu [33].

Obraz je segmentován s využitím morfologických operací. Ze snímku je tedy vybrán pouze otisk prstu a pozadí zůstává bílé. Mnoho snímků databáze pochází z daktyloskopických karet, které jsou naskenované – pozadí je tedy různých variant, jeho podoba není potřebná při další analýze a proto je tímto způsobem odstraněno. Morfologické operace jsou operacemi, které mohou být provedeny na binárním obrazu složeném z pixelů dvou možných barev (obvykle černobílé snímky). Každá morfologická operace využívá dva vstupy – originální obraz a jádro. Základními operacemi jsou eroze a dilatace. Eroze slouží k odstranění a ztenčení hranic objektu. Jádro je aplikováno na pixely obrazu stejným principem jako 2D konvoluce – pixel v původním obrázku (hodnoty 0 nebo 1) bude považován za 1 pouze v případě, pokud všechny pixely pod jádrem jsou hodnoty 1, jinak je nastaven na hodnotu 0. Dilatace je opakem eroze, kdy je pixel obrazu nastaven na hodnotu 1 pokud aspoň jeden pixel pod jádrem je hodnoty 1. Linie objektu v popředí se stávají silnějšími [34].



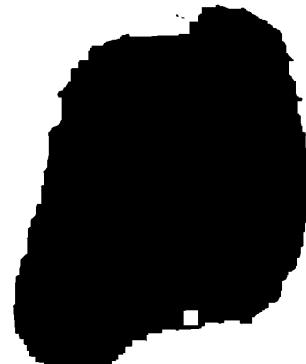
Obrázek 4.1: Snímek atopického ekzému



Obrázek 4.2: Normalizovaný obraz



Obrázek 4.3: Otsu prahování



Obrázek 4.4: Získaná maska

Pro předzpracování otisků prstů bylo využito jádro velikost  $23 \times 23$  a morfologická operace morfologického otevření – metoda eroze následována dilatací. Morfologické otevření

je vhodným přístupem pro odstranění šumu v obrazu, v našem případě se bude jednat o vyplnění světlých oblastí mezi papilárními liniemi, tak aby vznikla maska určující pozici otisku prstu (viz obrázek 4.4). Původní snímek a získaná maska jsou sečteny – výsledkem je přezpracovaný otisku prstu s bílým pozadím (viz obrázek 4.5).



Obrázek 4.5: Výsledný přezpracovaný snímek

## 4.4 Použité metody pro detekci a klasifikaci

Cílem této práce je analyzovat více různých přístupů pro detekci a klasifikaci onemocnění otisků prstů – bude experimentováno s mnoha typy detekčních modelů popsaných v kapitole 3.8. Konkrétně byly prozatím vybrány následující modely, bude dbán ohled na experimentování pro různé přístupy k detekci, tak varianty páteřních sítí:

- SSD MobileNet V2 FPNLite 320x320,
- SSD ResNet50 V1 FPN 640x640 (RetinaNet50),
- Faster R-CNN ResNet50 V1 640x640,
- Faster R-CNN ResNet101 V1 640x640,
- EfficientDet D0 512x512,
- EfficientDet D1 640x640,
- CenterNet HourGlass104 512x512,
- CenterNet Resnet101 V1 FPN 512x512.

### 4.4.1 TensorFlow a TensorFlow Model Garden

TensorFlow je open-source knihovnou zaměřující se na strojové učení. Umožňuje efektivně provádět nízko úrovňové tenzorové operace na CPU, GPU a TPU. Je schopna škálovat výpočty s využitím mnoha zařízení (např. clustery se stovkami GPU). Keras je vysokoúrovňové API knihovny TensorFlow - klade důraz na uživatelskou přívětivost a nezbytné abstrakce – umožňuje snadným způsobem stavět modely strojového učení pro rozličné úlohy [26].

TensorFlow Model Garden je knihovnou dostupnou pod licencí Apache 2.0. Obsahuje implementace state-of-the-art řešení modelů a poskytuje mimo jiné podporu pro detekci a

klasifikaci objektů – TensorFlow 2 Detection Model Zoo<sup>2</sup>[51]. Všechny modely byly před-trénovány na COCO 2017 datasetu<sup>3</sup> – jednom z nejužívanějších datasetů strojového učení. Tento dataset obsahuje více jak 200 000 anotovaných snímků řazenými do 80 objekto-vých kategorií, každý objekt je poté zařazen také do subkategorie – např. objekty kategorií *bicycle, car, motorcycle, airplane, bus, train, truck* nebo *boat* jsou zařazeny do superkategorie *vehicle* [40].

#### 4.4.2 Načítání a konfigurace modelu

Každý z detekčních modelů z knihovny TensorFlow Model Garden obsahuje konfigurační soubor s názvem `pipeline.config`. Tento soubor má mnoho parametrů vhodných pro další experimentování. Pro počáteční trénování modelu byl zvolen model SSD MobileNet V2 FPNLite 320x320, který je z knihovny jedním z nejrychlejších.

Pro vyzkoušení trénování a predikce byly zatím využity tři třídy onemocnění - bradavice, ekzém a dyshidróza. Pro trénovací a validační dataset je třeba vytvořit tzv. TF records, které obsahují obsahují vstupní obrázky reprezentované v bytech, anotované pozice bounding boxů s třídami a další dodatečné informace.

#### 4.4.3 Trénování modelu a následná predikce

Pro zkušební počáteční trénování prvního z detekčních modelů bylo využito prostředí Colab Pro s GPU T4 a P100. Při následné predikci je vybráno v každém obrazu 100 boun-ding boxů s nejvyšším skóre jistoty, uživateli je poté vykresleno 8 bounding boxů s aspoň 50 % mírou jistoty pro klasifikaci. Současně jsou do textového souboru zapsány i informace o jednotlivých nejistějších predikcích. Následně bude potřeba model vyhodnotit na testo-vacím datasetu na základě např. překryvu anotovaných a predikovaných bounding boxů a správnosti klasifikované třídy onemocnění, nebo pomocí dalších technik.

### 4.5 Výstup detekce a klasifikace onemocnění otisků prstů

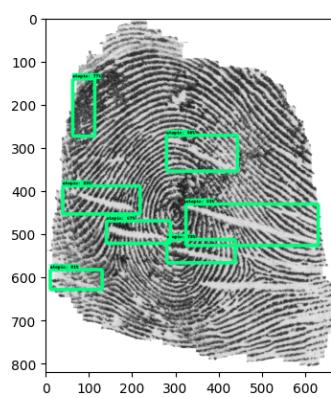
Výstupem detekce a klasifikace onemocnění otisků prstů je konzolová aplikace umožňující načtení, konfiguraci a trénování vybraného modelu ať už od počátku nebo od uvedeného checkpointu. Snímky z testovacího datasetu jsou následně predikovány natrénovaným modelem a jsou vykresleny bounding boxy s klasifikací a klasifikačním skóre jistoty pro dané onemocnění. Výsledky predikce jsou zapsány do textového souboru, následně bude na tes-tovací sadě ohodnocena i přesnost modelu. Při trénování modelu bude možnost upravovat různé parametry, vizualizovat metriky a počítat výkonnost.

Prvotní trénování modelu SSD MobileNet V2 FPNLite 320x320 probíhalo 50 000 ite-rací. Rovnoměrným dílem byly zastoupeny reálné snímky pacientů trpících atopickým ek-zémem, vygenerované syntetické otisky bez onemocnění, vygenerované syntetické otisky s onemocněním bradavic a vygenerované syntetické otisky s onemocněním dyshidrózy. Vý-sledné ukázkové predikce natrénovaného modelu jsou následující:

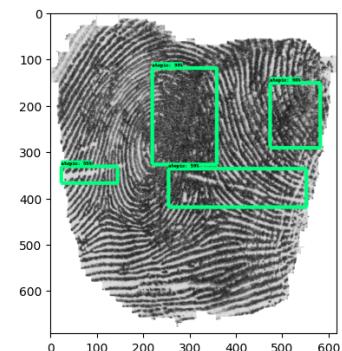
---

<sup>2</sup>[https://github.com/tensorflow/models/blob/master/research/object\\_detection/g3doc/tf2\\_detection\\_zoo.md](https://github.com/tensorflow/models/blob/master/research/object_detection/g3doc/tf2_detection_zoo.md)

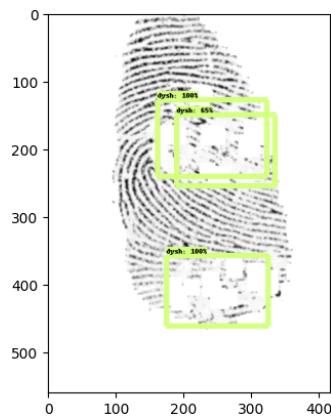
<sup>3</sup><https://cocodataset.org>



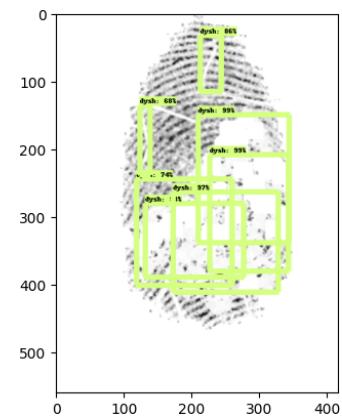
Obrázek 4.6: Predikce atopického ekzému



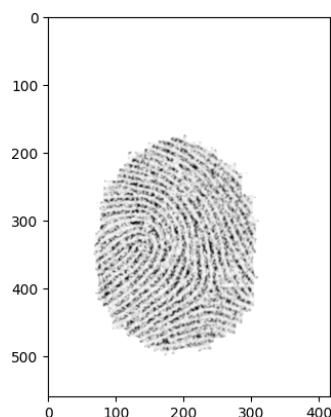
Obrázek 4.7: Predikce atopického ekzému



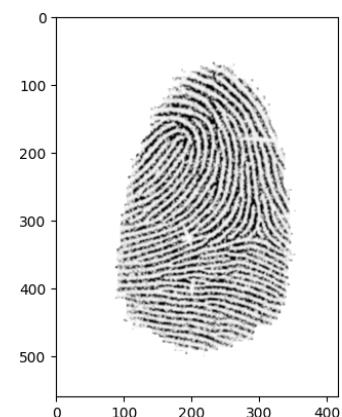
Obrázek 4.8: Predikce dyshidrózy



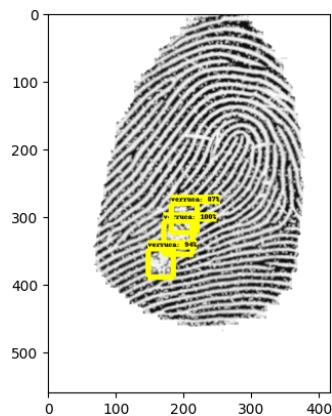
Obrázek 4.9: Predikce dyshidrózy



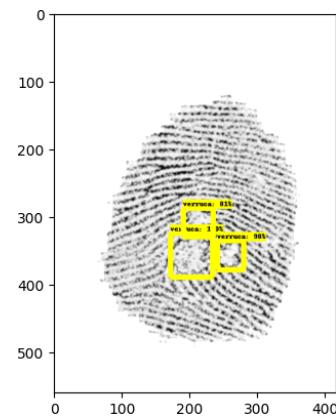
Obrázek 4.10: Model neučinil žádné predikce pro otisk prstu bez onemocnění



Obrázek 4.11: Model neučinil žádné predikce pro otisk prstu bez onemocnění



Obrázek 4.12: Predikce bradavic



Obrázek 4.13: Predikce bradavic

Při predikci se do souboru `predictions.txt` zapisují informace o predikcích se skóre jistoty aspoň 50 %. Níže je ukázka predikce pro snímek 4.6:

```
FILE NAME: atopic_eczema_FP_00183.png
DETECTIONS:
CLASS: atopic eczema
DETECTION SCORE: 98.44 %
CLASS: atopic eczema
DETECTION SCORE: 89.35 %
CLASS: atopic eczema
DETECTION SCORE: 80.77 %
CLASS: atopic eczema
DETECTION SCORE: 78.55 %
CLASS: atopic eczema
DETECTION SCORE: 76.88 %
CLASS: atopic eczema
DETECTION SCORE: 67.33 %
CLASS: atopic eczema
DETECTION SCORE: 62.71 %
```

# Kapitola 5

## Závěr

Cílem tohoto semestrálního projektu bylo nastudovat literaturu týkající se zpracování otisků prstů, biometrie a konvolučních neuronových sítí. Je kláden důraz na popis onemocnění, které mohou oblast prstu postihnout a mnohdy ovlivnit proces rozpoznávání. Při studii konvolučních neuronových sítí jsem se zaměřila na výčet jejich architektur s důrazem na ty z nejnovějších. Modelů pro detekci objektů také existuje celá řada, opět byla snaha popsat u každého z modelů jakým způsobem případně vylepšuje jiná již existující řešení a na jakém je založen principu.

Pro následnou analýzu bylo vybráno několik architektur, se kterými bude dále experimentováno a bude probíhat jejich porovnání na základě různých hledisek. Budou provedeny experimenty i s parametry jednotlivých modelů. Následně byla oannotována část datasetu otisku prstů a jejich předzpracování s využitím morfologických operací. Také byl implementován kód pro načtení a konfiguraci modelu. Prvotní zkušební trénování pak bylo provedeno s modelem SSD MobileNet V2 FPNLite 320x320 s využitím klasifikace onemocnění atopického ekzému, bradavic a dyshidrózy. Nutno podotknout, že toto trénování proběhlo jen na části reálného datasetu atopického ekzému a pouze na syntetických snímcích s dalším onemocněním. Dataset tedy byl zatím pouze částečným a bude rozšířen i o snímky pacientů trpící téměř chorobami a bude experimentováno i s onemocněním psoriázy. Dále byla implementována metoda pro predikci a vizualizaci bounding boxů pro testovací dataset. Model bude muset být následně na testovacím datasetu ohodnocen - ať už z hlediska procentuálního překryvu, porovnání predikované a anotované třídy onemocnění nebo dalších technik.

Bude také potřebné vygenerovat více syntetických snímků s onemocněními, protože dataset reálných otisků prstů s onemocněními je velmi malým pro učení konvoluční neuronové sítě. Avšak bude potřeba otestovat model na snímcích z reálného datasetu, aby byl schopen model rozpoznávat a detektovat onemocnění otisků prstů reálných pacientů.

# Literatura

- [1] AFTERACADEMY. *Mastering Activation Functions in Neural Networks* [online]. AfterAcademy, prosinec 2019 [cit. 2022-01-20]. Dostupné z: <https://afteracademy.com/blog/mastering-activation-functions-in-neural-networks>.
- [2] ALMOG, U. *CenterNet, Explained* [online]. Towards Data Science, 2021 [cit. 22. prosince 2021]. Dostupné z: <https://towardsdatascience.com/centernet-explained-a7386f368962>.
- [3] ALONSO-FERNANDEZ, F., ROLI, F., MARCIALIS, G., FIERREZ, J., ORTEGA-GARCIA, J. et al. Performance of fingerprint quality measures depending on sensor technology. *J. Electronic Imaging*. 2008, sv. 17, č. 1, s. 1–11. ISSN 1017-9909.
- [4] BAROTOVÁ Štěpánka. *Detector of Skin Diseases by Fingerprint Technology*. Brno, CZ, 2017. Bakalářská práce. Vysoké učení technické v Brně, Fakulta informačních technologií. Dostupné z: <https://www.fit.vut.cz/study/thesis/20057/>.
- [5] BROWNLEE, J. *When to Use MLP, CNN, and RNN Neural Networks* [online]. Machine Learning Mastery, červenec 2018 [cit. 2022-01-19]. Dostupné z: <https://machinelearningmastery.com/when-to-use-mlp-cnn-and-rnn-neural-networks/>.
- [6] BROWNLEE, J. *A Gentle Introduction to Pooling Layers for Convolutional Neural Networks* [online]. Towards Data Science, 2019 [cit. 23. prosince 2021]. Dostupné z: <https://machinelearningmastery.com/pooling-layers-for-convolutional-neural-networks/>.
- [7] BROWNLEE, J. *Loss and Loss Functions for Training Deep Learning Neural Networks* [online]. Machine Learning Mastery, leden 2019 [cit. 2021-12-09]. Dostupné z: <https://machinelearningmastery.com/loss-and-loss-functions-for-training-deep-learning-neural-networks/>.
- [8] BROWNLEE, J. *How to Choose an Activation Function for Deep Learning* [online]. Machine Learning Mastery, leden 2021 [cit. 2021-12-01]. Dostupné z: <https://machinelearningmastery.com/choose-an-activation-function-for-deep-learning/>.
- [9] BRZOBOHATÁ, K. *Dermatoglyfika* [online]. 2017 [cit. 2021-12-09]. Dostupné z: <https://is.muni.cz/el/sci/jaro2017/Bi5120c/Dermatoglyfika.pdf>.
- [10] DINESH. *CNN vs MLP for Image Classification* [online]. Medium, 2019 [cit. 23. prosince 2021]. Dostupné z: <https://medium.com/analytics-vidhya/cnn-convolutional-neural-network-8d0a292b4498>.

- [11] DRAHANSKY, M., DOLEZEL, M., URBANEK, J., BREZINOVA, E. a KIM, T. Influence of Skin Diseases on Fingerprint Recognition. *BioMed Research International*. 2012, sv. 2012, s. 144–157. ISSN 2578-8892.
- [12] DRAHANSKÝ, M. *Biometrické systémy - Studijní opora* [online]. 2006 [cit. 2021-12-07]. Dostupné z:  
[https://www.fit.vutbr.cz/study/courses/BIO/private/BIO\\_Studijni\\_opora.pdf](https://www.fit.vutbr.cz/study/courses/BIO/private/BIO_Studijni_opora.pdf).
- [13] DRAHANSKÝ, M., BŘEZINOVÁ, E., ORSÁG, F. a HEJTMÁNKOVÁ, D. Dermatologické faktory ovlivňující snímání otisků prstů. *Kriminalistika*. 2010, sv. 42, č. 3. ISSN 1210-9150.
- [14] DRAHANSKÝ, M. a KANICH, O. *Rozpoznávání podle otisků prstů* [online]. 2020 [cit. 2021-12-07]. Dostupné z:  
[https://www.fit.vutbr.cz/study/courses/BIO/private/04\\_BIO\\_Prednaska\\_OK\\_2020.pdf](https://www.fit.vutbr.cz/study/courses/BIO/private/04_BIO_Prednaska_OK_2020.pdf).
- [15] DYLEVSKÝ, I. a JEŽEK, P. *Kožní ústrojí - Anatomie kůže* [online]. [cit. 2021-12-10]. Dostupné z: <https://vos.palestra.cz/skripta/anatomie/12a1.htm>.
- [16] FOOTE, K. D. *A Brief History of Neural Networks* [online]. Dataversity, listopad 2021 [cit. 2021-12-25]. Dostupné z:  
<https://www.dataversity.net/a-brief-history-of-neural-networks/>.
- [17] GANDHI, R. *R-CNN, Fast R-CNN, Faster R-CNN, YOLO — Object Detection Algorithms* [online]. Towards Data Science, září 2018 [cit. 2021-12-03]. Dostupné z:  
<https://towardsdatascience.com/r-cnn-fast-r-cnn-faster-r-cnn-yolo-object-detection-algorithms-36d53571365e>.
- [18] GEY, T. *What is Neural Architecture Search? And Why Should You Care?* [online]. Towards Data Science, září 2021 [cit. 2021-12-15]. Dostupné z:  
<https://towardsdatascience.com/what-is-neural-architecture-search-and-why-should-you-care-1e22393de461>.
- [19] GIRSHICK, R. Fast R-CNN. *Computing Research Repository*. 2015, sv. 1504. ISSN 2331-8422.
- [20] GIRSHICK, R. B., DONAHUE, J., DARRELL, T. a MALIK, J. Rich feature hierarchies for accurate object detection and semantic segmentation. *Computing Research Repository*. 2013, sv. 1311. ISSN 2331-8422.
- [21] HONG, L., WAN, Y. a JAIN, A. Fingerprint Image Enhancement: Algorithm and Performance Evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1998, sv. 20, č. 8, s. 777–789. ISSN 0182-8828.
- [22] HOWARD, A. G., ZHU, M., CHEN, B., KALENICHENKO, D., WANG, W. et al. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *Computing Research Repository*. 2017, sv. 1704. ISSN 2331-8422.
- [23] INFORMEDHEALTH.ORG. *Warts: Overview* [online]. Cologne, Germany: Institute for Quality and Efficiency in Health Care (IQWiG), listopad 2019 [cit. 2021-11-24]. Dostupné z: <https://www.ncbi.nlm.nih.gov/books/NBK279586/>.

- [24] JAIN, A. a FENG, J. Latent Fingerprint Matching. *IEEE transactions on pattern analysis and machine intelligence*. 2011, sv. 33, s. 88–100. ISSN 1939-3539.
- [25] KAUSHIK, A. *Understanding ResNet50 architecture* [online]. OpenGenusIQ, březen 2020 [cit. 2021-12-14]. Dostupné z: <https://iq.opengenus.org/resnet50-architecture/>.
- [26] KERAS. *About Keras* [online]. Keras, 2021 [cit. 2021-12-25]. Dostupné z: <https://keras.io/about/>.
- [27] KOLB, L. a FERRER BRUKER., S. J. *Atopic Dermatitis* [online]. StatPearls Publishing, srpen 2021 [cit. 2021-11-24]. Dostupné z: <https://www.ncbi.nlm.nih.gov/books/NBK448071/>.
- [28] LIU, W., ANGUELOV, D., ERHAN, D., SZEGEDY, C., REED, S. et al. SSD: Single Shot MultiBox Detector. *Lecture Notes in Computer Science*. 2016, s. 21–37. ISSN 1611-3349.
- [29] MALTONI, D., MAIO, D., JAIN, A. K. a PRABHAKAR, S. *Handbook of Fingerprint Recognition*. 2. vyd. Springer, 2009. ISBN 978-1-84882-253-5.
- [30] NAIR, P. A. a BADRI, T. *Psoriasis* [online]. StatPearls Publishing, srpen 2021 [cit. 2021-11-24]. Dostupné z: <https://www.ncbi.nlm.nih.gov/books/NBK448194/>.
- [31] NATIONAL HEALTH SERVICE. *Raynaud's* [online]. NHS, říjen 2020 [cit. 2021-12-21]. Dostupné z: <https://www.nhs.uk/conditions/raynauds/>.
- [32] NEWELL, A., YANG, K. a DENG, J. Stacked Hourglass Networks for Human Pose Estimation. *Computing Research Repository*. 2016, sv. 1603. ISSN 2331-8422.
- [33] OPENCV. *Image Thresholding* [online]. OpenCV, 2021 [cit. 22. prosince 2021]. Dostupné z: [https://docs.opencv.org/3.4/d7/d4d/tutorial\\_py\\_thresholding.html](https://docs.opencv.org/3.4/d7/d4d/tutorial_py_thresholding.html).
- [34] OPENCV. *Morphological Transformations* [online]. OpenCV, 2021 [cit. 22. prosince 2021]. Dostupné z: [https://docs.opencv.org/3.4/d7/d4d/tutorial\\_py\\_thresholding.html](https://docs.opencv.org/3.4/d7/d4d/tutorial_py_thresholding.html).
- [35] PERNICKÝ, R. *Pokročilé rozpoznávání podle otisků prstů* [online]. 2020 [cit. 2021-12-21]. Dostupné z: [https://www.fit.vutbr.cz/study/courses/BIO/private/08\\_BIO\\_Prednaska\\_RP\\_2020.pdf](https://www.fit.vutbr.cz/study/courses/BIO/private/08_BIO_Prednaska_RP_2020.pdf).
- [36] PERUMANOOR, T. J. *What is VGG16? — Introduction to VGG16* [online]. Great Learning, září 2021 [cit. 2021-12-13]. Dostupné z: <https://medium.com/@mygreatlearning/what-is-vgg16-introduction-to-vgg16-f2d63849f615>.
- [37] REN, S., HE, K., GIRSHICK, R. B. a SUN, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *Computing Research Repository*. 2015, sv. 1506. ISSN 2331-8422.
- [38] RUDER, S. *An overview of gradient descent optimization algorithms*. 2017 [cit. 2021-12-20]. Dostupné z: <https://arxiv.org/abs/1609.04747>.
- [39] SHARMA, S. *Activation Functions in Neural Networks* [online]. Towards Data Science, září 2017 [cit. 2021-12-01]. Dostupné z: <https://towardsdatascience.com/activation-functions-neural-networks-1cbd9f8d91d6>.

- [40] SIMALANGO, M. F. *What Object Categories / Labels Are In COCO Dataset?* [online]. Amikelive, 2018 [cit. 22. prosince 2021]. Dostupné z: <https://tech.amikelive.com/node/718/what-object-categories-labels-are-in-coco-dataset/>.
- [41] SMITH, M. P., LY, K., THIBODEAUX, Q., BHUTANI, T., LIAO, W. et al. Acrodermatitis continua of Hallopeau: clinical perspectives. *Psoriasis: Targets and Therapy.* 2019, sv. 9, s. 65–72. ISSN 2230-326X.
- [42] SOUKUP, T. a VELETA, T. Systémová sklerodermie v roce 2017. *Vnitřní lékařství.* 2018, č. 2, s. 146–154. ISSN 1801–7592.
- [43] SVORADOVÁ, V. *Generování onemocnění kůže do syntetických otisků prstů z SFinGe.* Brno, CZ, 2019. Bakalářská práce. Vysoké učení technické v Brně, Fakulta informačních technologií. Dostupné z: <https://www.fit.vut.cz/study/thesis/21893/>.
- [44] TAN, M. a LE, Q. V. *EfficientNet: Improving Accuracy and Efficiency through AutoML and Model Scaling* [online]. Google AI Blog, květen 2019 [cit. 2021-12-14]. Dostupné z: <https://ai.googleblog.com/2019/05/efficientnet-improving-accuracy-and.html>.
- [45] TAN, M. a LE, Q. V. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. In: *Proceedings of the 36th International Conference on Machine Learning.* Long Beach, California, USA: PLMR, červenec 2019, s. 6105–6114. ISBN 978-1-5108-8698-8.
- [46] TAN, M., PANG, R. a LE, Q. V. EfficientDet: Scalable and Efficient Object Detection. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* Seattle, WA, USA: IEEE, červenec 2020, s. 10778–10787. ISBN 978-1-7281-7168-5.
- [47] TAN, M. a YU, A. *EfficientDet: Towards Scalable and Efficient Object Detection* [online]. Google AI Blog, duben 2020 [cit. 2021-12-13]. Dostupné z: <https://ai.googleblog.com/2020/04/efficientdet-towards-scalable-and.html>.
- [48] ŠALKO, M. *Detekce a klasifikace poškození otisku prstu s využitím neuronových sítí.* Brno, CZ, 2020. Bakalářská práce. Vysoké učení technické v Brně, Fakulta informačních technologií. Dostupné z: <https://www.fit.vut.cz/study/thesis/22567/>.
- [49] WANG, C.-F. *A Basic Introduction to Separable Convolutions* [online]. Towards Data Science, srpen 2018 [cit. 2021-12-25]. Dostupné z: <https://towardsdatascience.com/a-basic-introduction-to-separable-convolutions-b99ec3102728>.
- [50] WANG, C.-F. *The Vanishing Gradient Problem* [online]. Towards Data Science, leden 2019 [cit. 2021-12-14]. Dostupné z: <https://towardsdatascience.com/the-vanishing-gradient-problem-69bf08b15484>.
- [51] YU, H., CHEN, C., DU, X., LI, Y., RASHWAN, A. et al. *TensorFlow Model Garden* [online]. 2020 [cit. 2021-12-22]. Dostupné z: <https://github.com/tensorflow/models>.
- [52] ZBOŘIL, F. a ZBOŘIL, F. V. *Přednáška 9 - Rozpoznávání* [online]. 2021 [cit. 2021-12-21]. Dostupné z: [https://wis.fit.vutbr.cz/FIT/st/cfs.php.cs?file=%2Fcourse%2FIZU-IT%2Flectures%2FIZU%2F1819izu\\_9.pdf&cid=12818](https://wis.fit.vutbr.cz/FIT/st/cfs.php.cs?file=%2Fcourse%2FIZU-IT%2Flectures%2FIZU%2F1819izu_9.pdf&cid=12818).

- [53] ZEREN, M. T., AYTULUN, S. K. a KIRELLI, Y. Comparison of SSD and Faster R-CNN Algorithms to Detect the Airports with Data Set Which Obtained From Unmanned Aerial Vehicles and Satellite Images. *European Journal of Science and Technology*. 2020, č. 19, s. 643–658. ISSN 2148-2683.