

Univerzita Hradec Králové
Fakulta informatiky a managementu
Katedra informatiky a kvantitativních metod

Rozpoznávání objektů pomocí umělých neuronových sítí
Bakalářská práce

Autor: Lukáš Kučera
Studijní obor: Aplikovaná informatika

Vedoucí práce: Ing. Milan Košťák

Hradec Králové

Duben 2022

Prohlášení:

Prohlašuji, že jsem bakalářskou práci zpracoval samostatně a s použitím uvedené literatury.

V Hradci Králové dne 27.4.2022

Lukáš Kučera

Poděkování:

Děkuji vedoucímu bakalářské práce Ing. Milanu Košťákovi za metodické vedení práce, cenné připomínky a technické rady.

Anotace

Bakalářská práce se zabývá velice aktuálními tématy umělé inteligence a strojového vidění. V této moderní době se umělá inteligence společně se strojovým viděním používá v mnoha různých odvětvích. Bakalářská práce ve svém úvodu zodpovídá na otázku „Co je to umělá inteligence?“ a dále se zaměřuje na stručný popis historie tohoto vědního oboru. Následně je představeno několik důležitých pojmů, jako jsou umělé neuronové sítě, strojové vidění, fuzzy logika a expertní systémy, a je uvedeno několik příkladů využití umělé inteligence v různých oblastech. Také je definován pojem strojové vidění, představeny nesnáze s ním spojené a opět uvedeny příklady reálné aplikace. Cílem praktické části je vytvoření návrhu klasifikačního modelu a následná implementace. Klasifikační model pracuje s videozáznamem, přestože pro grafickou interpretaci je testování provedeno na snímcích. Dále jsou popsány použité technologie, použitá rozsáhlá datová sada (přes 50 000 snímků) s detailním popisem a sekvence účinných kroků k přípravě dat pro započítání trénování modelu. Model se projevil jako velice schopný s průměrnou jistotou klasifikace 97,4 % při trénování a 97 % při testování na 12 630 snímcích s nulovým počtem nesprávně klasifikovaných objektů.

Annotation

Title: Object recognition using artificial neural networks

The bachelor thesis deals with a very current topic – artificial intelligence and machine vision. In this modern age, artificial intelligence, along with machine vision, is used in almost every industry one can think of. The question “What is artificial intelligence?” is answered first. Then I will focus on and briefly describe the history of this scientific field. Next, I will present several important concepts, such as artificial neural networks, fuzzy logic, and expert systems. Finally, I will give examples of artificial intelligence applications in various fields. Then I will introduce the concept of machine vision, explaining the basic principle and again outlining its use. The practical part aims to create a design of the classification model and then its implementation. Although the classification model works with video recording, it is performed on images for graphical interpretation. The technologies used, the extensive data set (over 50 000 images) used with a detailed description and the steps that were taken for preparing the data to start training the model are also described. The model proved to be very capable, with an average classification accuracy of 97.4% in training and 97% in testing on 12,630 images with zero incorrectly classified objects.

Obsah

1	Úvod.....	1
2	Umělá inteligence.....	2
2.1	Historie umělé inteligence.....	2
2.2	Definice umělé inteligence.....	4
2.3	Turingův test.....	5
3	Techniky umělé inteligence.....	7
3.1	Fuzzy logika.....	7
3.1.1	Uplatnění fuzzy logiky.....	8
3.2	Expertní systémy.....	8
3.2.1	Využití expertního systému.....	8
3.3	Umělé neuronové sítě.....	9
3.3.1	Konvoluční neuronová síť.....	10
3.3.2	Uplatnění neuronových sítí.....	12
4	Využití umělé inteligence.....	14
4.1	Energetika.....	14
4.2	Marketing.....	14
4.3	Medicína.....	14
4.4	Automobilový průmysl.....	14
5	Strojové vidění.....	15
5.1	Princip strojového vidění.....	16
5.2	Komponenty systému strojového vidění.....	16
5.3	Obtíže strojového vidění.....	17
5.3.1	Iluze.....	17
5.3.2	Kvalita obrazu.....	18
5.3.3	Úhel pohledu.....	18

5.3.4	Pohyb.....	19
5.3.5	Variace objektů.....	19
5.4	Aplikace strojového vidění v reálném světě	20
5.4.1	Automobilový průmysl.....	20
5.4.2	Zdravotnictví a biomedicínský průmysl.....	20
5.4.3	Rozpoznávání obličeje	21
5.4.4	Maloobchodní průmysl.....	21
5.4.5	Výrobní průmysl	22
6	Návrh a implementace	23
6.1	Použité technologie	23
6.2	Datová sada pro CNN a klasifikační model.....	25
6.2.1	Datová sada.....	25
6.2.2	Trénování modelu	27
6.2.3	Klasifikační model	30
7	Shrnutí dosažených výsledků.....	33
7.1	Testování klasifikačního modelu	33
7.2	Výsledky.....	36
8	Závěr.....	38
9	Seznam použité literatury.....	39

Seznam obrázků

Obrázek 1 Podoba Turingova testu (Férée 2011).....	6
Obrázek 2 Aplikace filtru (MontBlanc 2020)	11
Obrázek 3 Max pooling (ComputerScienceWiki 2018).....	12
Obrázek 4 Váza nebo obličej? (Dilmen 2011)	18
Obrázek 5 Vysoká/nízká kvalita (Michelle 2012).....	18
Obrázek 6 Různé úhly (Anand 2015; Myrabella 2012)	19
Obrázek 7 Objekt v pohybu (Wilson 2012)	19
Obrázek 8 Počet snímků jednotlivých tříd (práce autora)	25
Obrázek 9 Nízká kvalita (Mykola 2018)	26
Obrázek 10 Vyšší kvalita (Mykola 2018).....	27
Obrázek 11 Příprava trénovacích dat (práce autora).....	28
Obrázek 12 Příprava testovacích dat (práce autora).....	28
Obrázek 13 Zpracování obrazu, změna rozměru (práce autora).....	29
Obrázek 14 Převod na binární matice (práce autora).....	29
Obrázek 15 Dělení datové sady 80/20 (práce autora).....	30
Obrázek 16 Posloupnost vrstev (práce autora).....	31
Obrázek 17 Dosažené výsledky trénování po dobu 20 (16) epoch s výslednou přesností 0,974 a ztrátovou funkcí 0,114 po evaluaci modelu na trénovacích datech (práce autora)	32
Obrázek 18 Výsledky predikce s uvedenou predikovanou třídou a přesností predikce (práce autora)	33
Obrázek 19 Predikce z videozáznamu (práce autora)	36

Seznam tabulek

Tabulka 1 Výsledky testování	35
------------------------------------	----

1 Úvod

Během posledních několika let se neustále klade důraz na usnadňování, automatizaci, snižování nákladů a další podobné aspekty. To je do jisté míry možné realizovat pomocí umělé inteligence, kterou lidé vypracovali do takové fáze, že jsou stroje schopné provádět intenzivní práce přesahující rámec lidských fyzických možností. Ale také mohou být společníky nebo pomocníky lidem tím, že usnadňují každodenní činnosti a jsou k dispozici skrze malá chytrá zařízení, kterými jsou například mobilní telefony. Právě počítačové vidění je jednou z oblastí umělé inteligence, která se v poslední době velice rychle rozvíjí a je optimisticky přijímána. V této oblasti jde o umožnění strojům vnímat a správně interpretovat to, co vidíme my. Právě touto rozvíjející se oblastí se teoretická část bakalářské práce zabývá a zaměřuje se zejména na klasifikaci objektů v obrazech pomocí umělých neuronových sítí, konkrétně konvolučních neuronových sítí. Zprvu je definován pojem umělá inteligence, nastíněn stručný náhled do historie, kdy tento vědní obor vznikal a ukázka principu světově známého tzv. Turingova testu. Dále budou vysvětleny základní techniky umělé inteligence, jimiž jsou expertní systémy, fuzzy logika a umělé neuronové sítě, se zaměřením na konvoluční neuronové sítě, jež budou použity v praktickém projektu. Zároveň bude uvedeno ke každé z technik několik příkladů uplatnění v reálném světě. Následně dojde k představení podoblasti zvané strojové vidění, užívané principy, nesnáze nastávající při nekvalitě, různorodosti obrazu a opět aplikace v běžném životě.

V praktické části je řešena úloha klasifikace objektu v obraze pomocí konvolučních neuronových sítí, kde je popsán postup předzpracování jednotlivých záznamů z datové sady a vložení do vhodných datových struktur, které jsou poté užity k trénování modelu. Cílem bakalářské práce je tedy seznámit čtenáře se základními principy umělé inteligence a strojového vidění a následně implementovat praktický projekt, který zahrnuje zmíněné technologie. Závěrem je provedeno testování a dosažené výsledky jsou interpretovány a shrnuty.

2 Umělá inteligence

Některé živé organismy byly obdařeny inteligencí, díky které jsou nadřazení vůči ostatním bez této schopnosti přemýšlet. Organismy mající tuto vlastnost jsou schopni reagovat na události, řešit problémy, adaptovat se a mnohé další. Lidé jsou oproti jiným organismům dokonce na vyšší úrovni z pohledu inteligence. Rozhodně nám chybí nějaké vlastnosti a schopnosti, kterými oplývají jiní živočichové, jako jsou např. velká síla, rychlost, schopnost slyšet vysoké frekvence, precizní zrak, apod. Ale naše vysoká inteligence nám umožnila vytvořit vynálezy, se kterými jsme schopni naše nedostatky vynahradit. Vynalezli jsme mikroskopy, zařízení pro snímání zvuku s příliš vysokou frekvencí, dopravní prostředky pro rychlý přesun, různé stroje pro manipulaci s masivními objekty.

2.1 Historie umělé inteligence

Na umělou inteligenci lze pohlížet jako na technický obor. Bádání v tomto oboru a tím pádem i jeho růst zajistilo mnoho aspektů, jako např. požadavky v odvětvích automatizovaného řízení, průzkum míst pro člověka nepřístupných a mnoho dalších.

Umělá inteligence jako vědní disciplína se postupně formovala v posledních 30 letech jako průsečík několika disciplín, jakými jsou např. psychologie, neurologie, kybernetika, matematická logika, teorie rozhodování, informatika, teorie her, lingvistika atd. Vývoj této disciplíny bude i nadále pokračovat a dnes už existuje sama za sebe. (Vacek 2000)

Počítačová umělá inteligence byla prvně použita ve větším měřítku v NASA k řízení a kontrole letu dálkových družic. Bez počítače by umělá inteligence nemohla existovat, a podle role, kterou počítač zastává, je dělena na slabou, kde je počítač pouze nástrojem, a silnou, kde je počítač přímo myslí a pomocí programu chápe. Je možné, že v blízké nebo daleké budoucnosti vznikne počítač, který dokáže myslet sám. (Zelinka 2003)

V roce 1941 byl v Německu poprvé použit elektronický počítač, který byl nepřiměřeně obrovský a chlazen leteckými motory, sloužící pro vojenské účely.

Na počátku roku 1950 přišel Norbert Wiener na to, že jsou všechna inteligentní rozhodnutí založena na principu zpětné vazby. Princip zpětné vazby už použil James Watt ve svém parním stroji a dost pravděpodobně netušil, že jeho objev ovlivní počátky vývoje umělé inteligence. (Wiener 1988)

V roce 1955 autoři Newell a Simon vytvořili program The Logic Theorist, považovaný za první program s umělou inteligencí. Program fungoval na principu rozkladu problému na stromový model a následného hledání větve s nejvhodnějším výsledkem. (Vacek 2000)

Roku 1957 byla testována první verze programu The General Problem Solver (GPS) od stejných tvůrců jako The Logic Theorist. Tento program byl rozšířením feedback principu a byl schopen řešit různorodé problémy. O několik let poté vznikl výzkumný tým pro umělé inteligence společností IBM. (Vacek 2000)

John McCarthy v roce 1958 přišel s novým jazykem LISP, který byl založen na zpracování seznamů a brzy ho začala používat většina vývojářů umělé inteligence. V roce 1963 dostala MIT grant v hodnotě 2,2 milionů dolarů od Ministerstva obrany USA pro zajištění technologické výhody proti Sovětskému svazu. V následujících letech vzniklo mnoho programů, např. program STUDENT, řešící algebraické problémy, nebo program SIR, rozumějící jednoduchým anglickým větám. (Searle 1980)

70. léta přinesla expertní systémy, které dokáží předpovědět pravděpodobnost řešení v určitých podmínkách. V této době přišlo na svět mnoho postupů. David Marr například uvedl teorii o strojovém vidění. Tato teorie byla o rozeznávání obrazů na základě informací o tvaru, barvě, texturách a hranách. V 70. letech také vznikl jazyk PROLOG. V 80. letech se urychlil vývoj umělé inteligence. Velice se zvýšila poptávka po expertních systémech, na které mnohé firmy spoléhaly. (Marr 2010)

Se zrychleným vývojem přišlo i praktické využití umělé inteligence. Práce Minského a Marra vedla ke vzniku kamer a počítačů pro kontrolu kvality na výrobních linkách. I armáda měla zájem o toto odvětví. (Zelinka 2003)

Přestože v historii nastaly momenty, kdy umělá inteligence přinesla několik neúspěchů a na nějaký čas jí velká pozornost věnována nebyla, postupně se stále rozvíjela. Objevily se neuronové sítě, fuzzy logika, a umělá inteligence se začala používat i v reálném životě. Dnes se s umělou inteligencí setkáme snad skoro v každém oboru. (Vacek 2000)

2.2 Definice umělé inteligence

Neexistuje pouze jedna přesná definice, ba naopak, bylo jich do dnešního dne vytvořeno velmi mnoho. Definice, vycházející z Turingova testu, profesora Marvina Minského, který tvrdí, že: „Umělá inteligence je věda o vytváření strojů nebo systémů, které budou při řešení určitého úkolu užívat takového postupu, který – kdyby ho dělal člověk – bychom považovali za projev jeho inteligence.“ (Mařík et al. 1993)

Z této definice vyplývá, že úlohy jsou tak složité, že i u člověka by vyžadovaly použití inteligence. Otázkou ale je, kdy můžeme konstatovat, že je úloha složitá? Mařík tvrdí: „Složitost lze ohodnotit počtem všech variant, které připadají při řešení v úvahu. V jednodušších úlohách lze nalézt řešení prostým přebíráním a vyhodnocováním jednotlivých variant. Ve složitějších úlohách takovýto postup není možný ani při použití superrychlých počítačů.“ Inteligentní kroky jsou takové kroky, které berou v potaz pouze smysluplné varianty a nevhodné varianty vynechají. (Mařík et al. 1993)

E. Richová píše, že: „*Umělá inteligence se zabývá tím, jak počítačově řešit úlohy, které dnes zatím zvládají lidé lépe.*“ (Rich a Knight 1991) V této definici se zrcadlí nynější stav v oblasti počítačových věd. Definice je velice výstižná, ale neobsahuje úlohy, které dosud nedokáží řešit jak počítače, tak i člověk.

Další definice, se kterými se můžeme setkat, jsou: (Zelinka 2003)

- „*Umělá inteligence je označení uměle vytvořeného jevu, který dostatečně přesvědčivě připomíná přirozený fenomén lidské inteligence.*“

- „Umělá inteligence označuje tu oblast poznávání skutečnosti, která se zaobírá hledáním hranic a možností symbolické, znakové reprezentace poznatku a procesu jejich nabývání, udržování a využívání.“
- „Umělá inteligence se zabývá problematikou postupu zpracování poznatku – osvojování a způsobem použití poznatku při řešení problému.“

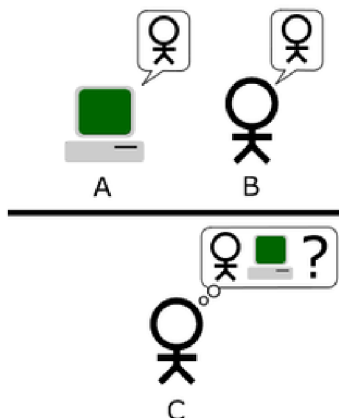
Jak již bylo zmíněno, definic existuje obrovské množství, ale jsou to definice obecné. Nesdělují nám, co to umělá inteligence opravdu je.

2.3 Turingův test

Britský matematik Alan Turing roku 1950 tvrdil, že není potřeba, aby stroj uměl myslet, postačí, když bude předstírat, že je toho schopen. Turing prezentoval tzv. Turingův test. Jedná se o test, který si klade za cíl ověřit, zda má nějaký systém umělé inteligence srovnatelnou inteligenci s člověkem. (Volná 2013)

Tato imitační hra spočívá v tom, aby člověk byl schopen rozhodnout přes písemnou interakci, zda komunikuje s člověkem nebo strojem. Subjekty, mezi něž patří testující, stroj anebo jiný člověk, jsou rozděleny do oddělených místností. Testující klade otázky v přirozené řeči a předává je druhému subjektu, kde jsou zodpovězeny počítačem, nebo jiným člověkem. Pokud alespoň 30 % respondentů vyvodí, že s ním komunikuje člověk, stroj testem projde. (Volná 2013)

Tímto způsobem například roku 2012 uspěl robot Eugen Goostman, který se vydával za 13letého ukrajinského chlapce (BBC 2014).



Obrázek 1 Podoba Turingova testu (Féréé 2011)

Obrázek 1 znázorňuje podobu Turingova testu. Princip je takový, že osoba C má jediný úkol, kterým je detekovat, jaká vystupující entita (A, B) je člověk a která pouze stroj vydávající se za člověka.

3 Techniky umělé inteligence

Kapitola postupně představuje principy fuzzy logiky, expertních systémů a umělých neuronových sítí. U neuronových sítí dojde k detailnější charakteristice konvolučních neuronových sítí, jejich zákonitostí a důležitých vrstev. Také je ke každé z technik uvedeno několik příkladů reálného využití.

3.1 Fuzzy logika

Jedním z algoritmů umělé inteligence je fuzzy logika. Zmíněná fuzzy logika pracuje odlišně od klasické logiky, která ohodnocuje výroky jako pravdivé, nebo nepravdivé (0/1) (Trenz 2009).

„Fuzzy logika je vícehodnotová logika definovaná funkcí příslušnosti prvku na intervalu $\langle 0,1 \rangle$. Jednotlivé prvky fuzzy systému integrující v sobě fuzzy logiku jsou bloky fuzzifikace (transformace reálných proměnných), fuzzy inference (pravidla, jazykové – lingvistické proměnné) a defuzzifikace (transformace jazykových proměnných).“ (Trenz 2009)

Fuzzifikací se rozumí operace, která převádí reálné proměnné do jazykových. Pro příklad si můžeme uvést reálnou proměnnou, a to cenu výrobku. Po fuzzifikaci nám vznikne jazyková proměnná s názvem cena, která může mít např. tyto atributy – malá, spíše malá, střední, spíše vyšší, vysoká. (Trenz 2009)

Fuzzy interference jsou předem definovaná transformační pravidla, která jsou modelována za pomoci podmínky: „IF podmínka THEN proved“. (Trenz 2009)

Defuzzifikace je opačnou operací k procesu fuzzifikace neboli převádí jazykové proměnné do reálných proměnných pomocí transformačních postupů. (Trenz 2009)

3.1.1 Uplatnění fuzzy logiky

Fuzzy logika se využívá pro mnoho každodenních činností spojených s používáním různých strojů, aniž bychom si to uvědomovali. Například mnozí z nás denně používají automobil, kde je právě touto logikou řešené ABS (AntiBlockiersystem – protiblokovací systém brzd). Brzdná síla je regulována na základě různých aspektů a určených interferenčních pravidel (rychlost otáčení jednotlivých kol, adheze vozovky aj.). Také pro řízení výtahu (plynulý rozjezd a brždění dle aktuální hmotnosti), řízení pračky (množství náplně), automatické ostření fotoaparátů, identifikace osob a mnoho dalších. (Trenz 2009)

3.2 Expertní systémy

Přesnou definici doposud nikdo nevyslovil. Často bývá charakterizován jako počítačový program simulující rozhodování expertů, který nevyužívá znalostí nabytých vlastní činností, ale znalostí špičkových odborníků v dané oblasti, avšak bez lidských omylů. (Celbová 1999)

Expertní systém se skládá ze dvou hlavních složek: znalostní báze obsahující znalosti, které jsou zpracovávány druhou částí, kterou je inferenční mechanismus. Inference či generování expertízy je strojové usuzování. (Celbová 1999)

Funkční expertní systém má mnoho výhod. Definitivně je levnější na provoz, nenarazíme na problém s dostupností, stabilně se rozhoduje v čase. Bohužel vytvoření takového expertního systému neboli převedení lidského rozhodování na systém fungující podle pravidel, není vůbec jednoduchá záležitost, jelikož ani samotný lidský expert není často schopen definovat postup, kterým ke správnému rozhodnutí došel. (Celbová 1999)

3.2.1 Využití expertního systému

Využití expertních systémů můžeme vyzorovat opět v mnoha odvětvích jako jsou medicína, geologie, chemie, zemědělství, matematice atd.

Typy úloh využívajících expertních systémů: (Schejbal et al. 2004)

- Monitorování – ohodnocování chování systému v reálném čase na základě porovnání s daty, které jsou pozorovány a daty, jež jsou očekávány
- Diagnostika – proces nalezení chyb na základě výsledků pozorování
- Interpretace – určení významu dat
- Plánování – nalezení sledu činností k dosažení cíle
- Konfigurace – kompletace souvisejících částí systému žadaným způsobem
- Prognózování – predikování budoucích důsledků
- Učení – inteligentní vyučování
- Ladění – tvorba předpisu pro odstranění chyb systému
- Řízení – regulace procesů

3.3 Umělé neuronové sítě

Snaha o vytvoření inteligentních strojů, schopných podobného myšlení jako u člověka, vedla ke vzniku vědního oboru – teorie neuronových sítí. Vůbec první myšlenka týkající se těchto sítí vznikla ve 40. letech 20. století. Neuronové sítě jsou struktury, které mají za úkol simulovat a implementovat některé funkce lidského mozku. Mezi hlavní vlastnosti patří adaptace a učení se. Vědci se inspirovali biologickou neuronovou sítí. (Molnár 2000)

Samotná neuronová síť je tvořena z množství neuronů a spojů. Základním stavebním prvkem je tedy neuron. Existuje i jeho zastaralý typ zvaný perceptron, se kterým se již nedá téměř setkat. Ale právě z tohoto prvně vzniklo označení vícevrstvé perceptronové sítě. (Molnár 2000)

„Umělá neuronová síť funguje tak, že vytváří spojení mezi mnoha různými procesními prvky, z nichž každý je analogický se samostatným neuronem v biologickém mozku. Tyto neurony mohou být fyzicky konstruovány nebo simulovány digitálním počítačem. Každý neuron dostává mnoho vstupních signálů. Poté, na základě vnitřního vyvažovacího systému, produkuje jednotlivý výstupní signál, který je typicky zasílán jako vstup jinému neuronu.“ (Kay 2001)

Všechny typy neuronových sítí jsou složeny ze stejných stavebních jednotek – neuronů, které mohou obsahovat různé přenosové neboli aktivační funkce, spojení mezi sebou a adaptivně-učící algoritmus. To vše pak určuje, o jakou síť se jedná. Technické neuronové sítě se dělí podle několika kritérií, které určují i jejich použití: (Zelinka 2003)

- Počtu vrstev:
 - mající jednu vrstvu
 - mající více vrstev
- Typu algoritmu učení:
 - s učitelem
 - bez učitele
- Stylu učení na síť s učením:
 - deterministické
 - stochastické

3.3.1 Konvoluční neuronová síť

Jedná se o jednu z variací tradičních neuronových sítí, navržena především pro práci s obrazovým vstupem. Tyto sítě řeší problém způsobem modelování menších částí informací, které jsou následně kombinovány pomocí hlubokých sítí. Tento proces se provádí na úrovni několika vrstev. V první vrstvě dochází k detekci hran a tvorbě šablony, respektive segmentace původního obrazu na základě detekovaných hran. Další vrstvy následně pracují s vyprodukovanými šablonami a jednoduššími objekty z obrazu a formují další sadu nových šablon s různými variacemi objektu jako je jiná pozice, osvětlení, rozměr. Poslední vrstva se snaží najít shodu se vstupním obrazem a těmito šablonami, čímž vzniká velký počet výstupů, z nichž je výsledným výstupem vážený součet. (Gollapudi a Laxmikanth 2019)

Konvoluční neuronové sítě (CNN) jsou formovány ze tří typů vrstev:

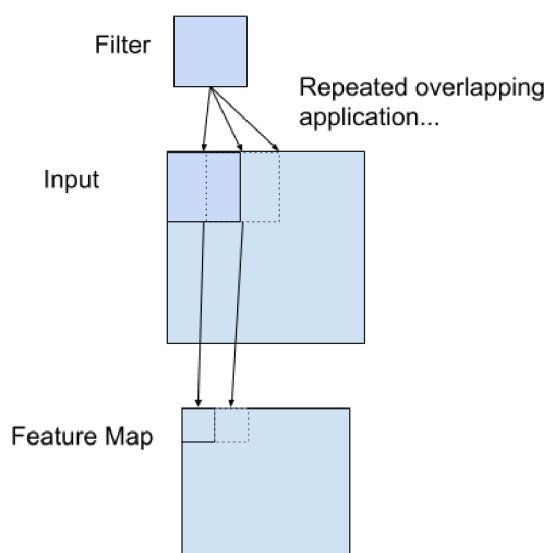
- Konvoluční vrstvy
- Pooling (sdružovací) vrstvy
- „Fully connected“ (plně propojená) vrstva

Oproti tradičním neuronovým sítím konvoluční sítě fungují trochu jiným způsobem. Důležité je zmínit, že jsou trojrozměrné (výška, šířka a hloubka). Navíc nedochází kvázání všech neuronů jedné vrstvy ke všem neuronům druhé, jde pouze o připojení k částečné sadě. Také se odlišuje výstupem, který je zde interpretován jediným vektorem skóre pravděpodobnosti. Skryté vrstvy napomáhají při rozboru prvků, což obstarávají konvoluční a pooling vrstvy. Finální klasifikaci zajišťuje fully connected (plně propojená) vrstva. (Gollapudi a Laxmikanth 2019)

Konvoluční vrstva

Konvoluční vrstva je primárním stavebním prvkem konvolučních neuronových sítí. Jejím úkolem je provádění operace s názvem konvoluce, což je lineární operace zahrnující násobení množiny vah se vstupními daty. Množina vah, respektive pole vah, je dvourozměrné a označuje se jako filtr či jádro. (MontBlanc 2020)

Filtr je záměrně volen menší, než jsou vstupní data z toho důvodu, že je možné použití stejného filtru na každou překrývající se část vstupních dat. Tento přístup zajistí například detekci určitého prvku v rozsahu celého vstupního obrazu. (MontBlanc 2020)

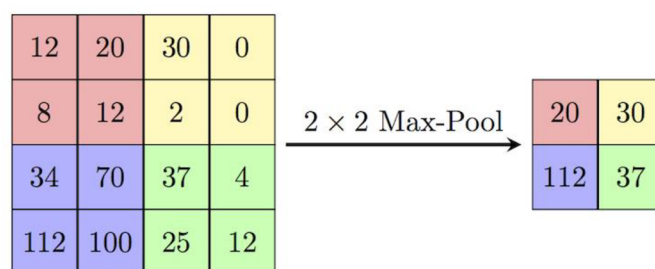


Obrázek 2 Aplikace filtru v rámci operace konvoluce (MontBlanc 2020)

Z důvodu několikanásobné aplikace filtru je výsledným výstupem dvourozměrné pole hodnot též nazývané jako mapa prvků. Tato mapa je poté předána aktivační funkci, například ReLU (Rectified Linear Activation), která bude použita v praktické části. (MontBlanc 2020)

Pooling vrstva

Sdružovací vrstva je zodpovědná za redukci parametrů pro kontrolu časování a předejití přetrénování. Standardně je na vstupu pooling vrstvy výstup z konvoluční vrstvy. Nejčastější aplikovanou technikou je max pooling, která z každé matice, jejíž velikost se nastavuje parametrem, volí maximální hodnotu (viz Obrázek 3). Tím redukuje a optimalizuje velikost mapy prvků a zároveň jsou zachovány klíčové informace o snímku. (Gollapudi a Laxmikanth 2019)



Obrázek 3 Max pooling (ComputerScienceWiki 2018)

Fully connected vrstva

Fully connected vrstva se stará o klasifikační exekuční plán finální mapy prvků. Tato vrstva přijímá data pouze jednorozměrná, tudíž je nutné data transformovat, což je možné pomocí tzv. Flatten vrstvy, která vstup transformuje do jednoho vektoru a výsledek putuje do klasické neuronové sítě, jejichž vrstvy mají za úkol predikovat třídy na základě výstupu z konvoluční vrstvy. Neurony ve fully connected vrstvě jsou plně spjaty se všemi neurony z vrstvy předešlé a fungují na stejném principu jako tradiční neuronové sítě.

3.3.2 Uplatnění neuronových sítí

V současnosti se umělé neuronové sítě používají uvnitř funkčních systémů jako pomocný nástroj, např. pro potlačení šumu či rozpoznávání znaků. Dále se s jejich aplikací také můžeme setkat při dolování dat, řízení procesů, poradenství

a rozhodování, přesné předpovídání údajů, rozpoznávání kvality výrobků, v samoobslužných mechanismech, u dopravní signalizace, v lékařství apod. (Uldrich a Jurczyk 2014)

4 Využití umělé inteligence

4.1 Energetika

V energetice se umělá inteligence používá zejména pro zlepšení údržby obnovitelných zdrojů energie, například monitorování větrných parků a distribučního vedení, vyhodnocování zvuků strojů v elektrárnách, z nichž byla schopna určit, zda stroj běží správně a případně ihned podat informace o problému. Nebo také pro sledování a vyhodnocování provozu sítí, diagnostiku a předvídání provozu nebo řízení a optimalizaci. (Řezníčková 2019)

4.2 Marketing

Pro optimální marketingové tahy je právě zde také využita umělá inteligence. Například všichni mají jistě mnoho zkušeností s bannerovými reklamami, které zobrazují právě ty produkty, o které jsme v poslední době projevili zájem, byť jen drobný (např. jednorázová prohlídka produktu na e-shopu). (Řezníčková 2019)

4.3 Medicína

V tomto velice důležitém odvětví nám je umělá inteligence také velice nápomocná. Zejména se aplikuje pro rozpoznávání obrazu (např. snímky z magnetické rezonance), poněvadž je v tomto směru přesnější než lidské oko. Také dokáže analyzovat velký objem dat s cílem stanovení nejvhodnějších léčebných postupů, ať se jedná o vzácné onemocnění nebo jemu podobné. (Řezníčková 2019)

4.4 Automobilový průmysl

Výrazný posun v zavádění umělé inteligence nyní zažívá automobilový průmysl. Zde hraje velkou roli při nahrazování lidských činností. Zejména je využita pro zvýšení bezpečnosti. Například v novějších automobilech s vyšší výbavou se můžeme setkat se systémem monitorující pozornosti řidiče, který při zaznamenání jakéhokoli problému reaguje spouštěním sekvencí varovných signálů. Také je běžné, že automobil je vybaven funkcí automatického parkování, nebo ovládání infotainmentu vozu pomocí hlasu, což sice občas nefunguje optimálně, ale časem se tato funkce jistělepší. (Rizzoli 2022)

5 Strojové vidění

Historie strojového vidění se začala psát již 70. letech 20. století. Vznikl nový vědní obor – počítačové vidění (computer vision), jehož název se uchytil a téměř každý je nyní schopen si pod tímto pojmem představit něco konkrétního či abstraktního. Tento název se stal záštitou veškerým systémům, které pracují s informacemi získanými z digitálního obrazu. Termín strojové vidění je v současnosti považován za využití počítačového vidění k řešení úloh v průmyslové automatizaci. Pod těmito úlohami si můžeme představit vizuální inspekce, počítání objektů, hlídání defektů, identifikaci barev, verifikaci kódu, kontrolu úplnosti balení, kontrolu tolerancí a nespočet jiných. Přestože je v mnoha firmách uplatňována automatizace, stále existuje velké množství operací kontrolovaných člověkem, což má vliv na kvalitu vykonání daných operací, jelikož člověk není neomylný. (Havle 2008)

Technologie strojového vidění implementuje techniky hlubokého učení. S rozsáhlým přínosem techniky hlubokého učení se staly operace jako klasifikace obrazu, detekce objektu, manipulace obrazu výrazně jednodušší a přesnější. To nám umožňuje se soustředit na výzkum značně komplexnějších aplikací této technologie, čímž se myslí např. enormně populární a rychle vyvíjející se odvětví – autonomní vozidla, humanoidi a drony. (Gollapudi a Laxmikanth 2019)

V oboru strojového vidění tím, že konstatujeme, že umožňujeme počítačům/strojům vidět, máme na mysli umožnit jim zpracovávat digitální vizuální data bez ohledu na způsob pořízení. Dá se říci, že se jedná o opak počítačové grafiky, jenž má za úkol vytvořit obrazová data z informací o existujících objektech. Strojové vidění nám umožňuje hledat konkrétní objekty v obraze/video, sledovat pohyb a jeho vzor z videa, editovat vlastnosti snímku. (Gollapudi a Laxmikanth 2019)

5.1 Princip strojového vidění

Princip strojového vidění je velice podobný jako u člověka. Místo oka, které zachytává obraz zkoumaného předmětu, je zde použita kamera, plnící stejnou funkci. Systém obraz poté vyhodnotí dle definovaného algoritmu a provede optimální akci na základě výsledku vyhodnocení. Pro zjednodušení je možné hledět na systém strojového vidění jako na obecný senzor. (Havle 2008)

Sledovaný objekt, jenž je většinou trojrozměrný, je osvětlován viditelným zářením pocházejícím z umělých zdrojů. Následně je 2D snímek, nesoucí obrazové informace, vytvářen na snímacím prvku kamery. „V kameře s polovodičovým snímacím prvem jsou měronosnými veličinami náboje či proudy poskytované prvky obrazového snímače, které jsou citlivé na světlo. Měronosné veličiny mohou nést informaci redukovanou na údaj o kombinovaném jasu barevných složek dopadajícího světla (černobílé kamery) nebo mohou nést i informaci o barvě zkoumaného předmětu (barevné kamery).“ (Havle 2008). Jelikož nám dnes pozůstává jediný a také efektivní způsob vyhodnocení obrazové informace zpracovaný počítačem, je zapotřebí analogový vstupní obraz digitalizovat za účelem možného zpracování počítačem. Počítač je schopen digitalizovaný obraz zpracovat vhodnými algoritmy pro získání informací o objektu – provádí se tzv. image processing. Zpracovaný obraz má opět digitální podobu. (Havle 2008)

5.2 Komponenty systému strojového vidění

Mezi hlavní komponenty systému strojového vidění patří objektivy a osvětlení, kamerový senzor nebo kamera, procesor a princip komunikace výsledků.

Objektivy zachycují snímek a interpretují jej senzoru ve světelné formě. Je zapotřebí spárovat kameru s vhodným objektivem pro optimalizaci systému. Aplikovatelnost světla na snímek je různorodá. Směr přicházejícího světla, jeho jas, a barva neboli vlnová délka v porovnání s barvou cíle patří mezi nejdůležitější faktory, které je potřeba promyslet při návrhu prostředí systému. Přestože je pro pořízení snímku velice důležité osvětlení, existují ještě další dva faktory, které množství tohoto příchozího světla přímo ovlivňují. Jedním z nich je komponenta,

které se říká clona. Ta se otevírá nebo zavírá v závislosti na potřebném množství světla. Při skombinování s dobou expozice je ovlivňováno množství světla dopadajícího na pole pixelů před tím, než je osvětlení aplikováno. Doba expozice, také uzávěrka, určuje dobu promítání obrazu na pole pixelů. Strojové vidění má uzávěrku řízenou elektronicky. (Control Engineering Česko 2019)

Po následném zachycení snímku se aplikují různé softwarové nástroje. V předběžném zpracování mohou být na snímek aplikovány efekty například pro zaostření okrajů, zvýšení kontrastu, výplně prostor atd. Jejich úkolem je zvýšit kvalitu snímku pro další softwarové nástroje. (Control Engineering Česko 2019)

5.3 Obtíže strojového vidění

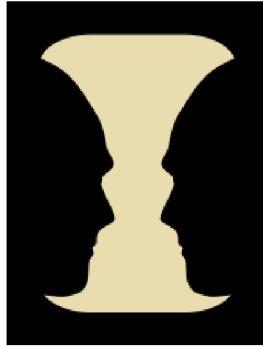
Jak již bylo zmíněno, zdrojem digitálních obrazových dat mohou být kamery, dále také webkamery, obrazové snímače a mnoho dalších. Přesnost výsledků strojového vidění velice souvisí s kvalitou interpretovaných obrazů či videí a také na zařízení používaného pro snímání vstupních obrazových dat. (Gollapudi a Laxmikanth 2019)

Snímané obrazy či obrazové záznamy mohou mít mnoho aspektů, které komplikují aplikaci strojového vidění. Některé aspekty mohou i našemu lidskému oku, respektive mozku, který zpracovává obraz z oka, způsobit potíže při klasifikaci objektu. (Gollapudi a Laxmikanth 2019)

5.3.1 Iluze

Iluze v obraze mohou být velice matoucí. Podle studie výzkumníků z Brown University v USA, kteří se snažili naučit počítače identifikovat optické iluze, je model naučený model schopen vnímat iluze stejně jako člověk. (Powell 2018)

Co například zobrazuje Obrázek 4? Vázu nebo 2 obličeje z profilu?



Obrázek 4 Váza nebo obličej? (Dilmen 2011)

5.3.2 Kvalita obrazu

Jedním z nejdůležitějších aspektů je kvalita obrazu, protože má velký vliv na přesnost výsledků zpracovaného vstupního obrazu. Nekvalitní obraz může způsobit šum a také kostičkování při přibližování a pokud nejsme schopni my rozeznat objekt, pravděpodobně toho nebude schopen ani stroj.



Obrázek 5 Vysoká/nízká kvalita (Michelle 2012)

5.3.3 Úhel pohledu

Další potíže přináší různorodost objektů při pořizování jejich snímků z různých úhlů. Každý objekt totiž nemusí vypadat ze všech úhlů totožně (viz Obrázek 6).



Obrázek 6 Různé úhly (Anand 2015; Myrabella 2012)

5.3.4 Pohyb

U pohybu tomu je podobně jako při pořizování snímků z různých úhlu. Podoba objektů (viz Obrázek 7) v pohybu se může velice lišit a pro systém jsou takové objekty často složitě identifikovatelné. Také při rychlejším pohybu může nastat motion blur efekt, což s sebou přináší neostrý a rozmazaný obraz, který opět vede ke obtížnější identifikaci.



Obrázek 7 Objekt v pohybu (Wilson 2012)

5.3.5 Variace objektů

Často také nastává problém s mnoha variacemi jednotlivých objektů. Pokud systém danou variaci, která se výrazněji odlišuje od té původní, nikdy předtím neviděl, šance na správnou klasifikaci je velice malá.

Všechny tyto zmíněné aspekty, s výjimkou iluzí, lidem problémy nedělají. Počítač bohužel není schopný například detekovat psa v poloze, kterou nikdy „neviděl“ (nebyla mu poskytnuta pro učení), což by u člověka hrozit nemělo.

5.4 Aplikace strojového vidění v reálném světě

Několik aplikací, kde je strojové vidění využíváno, již bylo zmíněno v kapitole o umělé inteligenci, jelikož pár z nich bylo částečně spjatých právě se zpracováním obrazu a spolu s umělou inteligencí tvoří robustnější a komplexnější systémy, ale nyní bude dále představeno několik specifických uplatnění využívajících čistě strojového vidění.

5.4.1 Automobilový průmysl

Automobilový průmysl se v posledních letech důrazně zaměřuje na dozrávání technologie autonomních vozů. Tím pádem je do trhu autonomních vozů investováno obrovské množství peněz. Na rok 2023 se předpokládá velikost trhu 833 miliard CZK z důvodu vysokých investic. (Isakova 2021)

Strojové vidění se v tomto průmyslu používá již řadu let, primárně pro zvýšení bezpečnosti a modernizace automobilů. Ty jsou vybaveny různými senzory a kamerami, jejichž úkolem je sbírání obrazových záznamů a dat, týkajících se dění kolem automobilu. Auta tedy dokáží detekovat značky rychlostního limitu v reálném čase během jízdy, upozornit řidiče při parkování v zóně se zákazem parkování nebo proaktivně nacházet volná parkovací místa. (Gollapudi a Laxmikanth 2019; Gupta et al. 2021)

5.4.2 Zdravotnictví a biomedicínský průmysl

Průmysl zdravotnictví a biomedicíny také zaznamenal výrazné množství užití strojového vidění. Krom pokroků ve způsobu formování a ukládání dat, souvisejících s lidskými orgány, došlo u prostředku pro interpretaci a vizualizaci zmíněných dat k drastické změně pomocí algoritmů strojového vidění. (Isakova 2021)

Nyní je možné dosáhnout preventivního opatření u rakoviny a ostatních genetických onemocnění s pomocí detekce potenciálních výskytů v poskytnutých snímcích z mikroskopu s vysokou přesností. Navíc jsou roboti schopni provést komplexní operace efektivně a důkladně. (Gollapudi a Laxmikanth 2019)

5.4.3 Rozpoznávání obličeje

Detekce a rozpoznávání obličeje jsou některé z nejvýznamnějších příkladů technologie strojového vidění. Do styku s aplikací v této formě se dostáváme s mnoha podobách. Nejčastěji se můžeme setkat s odemykáním chytrých mobilních zařízení pomocí obličeje, na letištích u samoobslužných bran, kde je porovnávána fotka z cestovního pasu s fotkou pořízenou kamerou, hledání totožnosti pachatele trestného činu zachyceného na kamerovém záznamu v databázi a mnoho dalších. (Isakova 2021)

5.4.4 Maloobchodní průmysl

Kamery instalované v maloobchodech umožňují obchodníkům sbírat ohromný objem vizuálních dat, usnadňující navrhování a zprostředkování zkušenosti zákazníků a zaměstnanců. Díky vývoji systémů strojového vidění zpracovávající tato data je digitální transformace skutečného průmyslu mnohem dosažitelnější. (Rizzoli 2022)

Například systém automatického doplnění dokáže zachytit obrazová data a provést skenování zásob tím, že sleduje položky na policích v milisekundových intervalech. Systém poté může upozornit na vyprodání zásob a tím napomáhá se správou. (Rizzoli 2022)

Dále jsou užívány také systémy pro detekci krádeží založené na zpracování obrazu z kamer v reálném čase. Lze je využít k detekci podezřelého chování, fyzických útoků nebo nehody, čímž se zlepší bezpečnost. (Rizzoli 2022)

V době pandemie onemocnění covid-19 byla nejen Česká republika tísňena opatřením, které diktovalo maximální počet osob na m² v obchodech. Zde vznikl systém automatického počítání osob vstupujících do a odcházejících z obchodu,

který zautomatizoval tuto operaci. Počítání davu je dále také užitečné pro správu personálu, front a poskytuje nám užitečná analytická data. (Rizzoli 2022)

Příkladem online obchodu využívajícího strojové vidění je obchod s brýlemi. Zákazníci mohou simulovat, jak by obroučky vypadaly na jejich obličeji atd. (Gollapudi a Laxmikanth 2019)

5.4.5 Výrobní průmysl

Výrobní průmysl již nějakou dobu obsahuje širokou škálu automatizačních řešení se strojovým viděním ve svém jádře. Pomáhá automatizovat kontrolu kvality, minimalizovat bezpečnostní rizika a zvyšovat výrobní efektivitu. (Rizzoli 2022)

Velkokapacitní výrobní závody často nedosahují 100% přesnosti v detekování vad vyráběných produktů. Kamerové systémy mohou shromažďovat data v reálném čase a využívat strojové vidění k jejich analýze a porovnávat je s předdefinovanou sadou standardů kvality. Tento systém napomáhá identifikovat defekty na makro a mikroúrovni ve výrobní lince, což s sebou přináší usnadnění výrobního procesu a snižování nákladů. (Rizzoli 2022)

Dalším přínosem je také automatizování procesů montáže a správy produktů. Například Tesla je dobře známým průkopníkem plně automatizovaných výrobních procesů. (Isakova 2021)

Tento seznam je pouhým stručným nahlédnutím pod povrch aplikací strojového vidění, jelikož škála uplatnění je nesmírně široká. S jeho příchodem a dozráním do dosavadní úrovně se enormní množství činností, prací a operací zautomatizovalo, což přineslo mnohé výhody a usnadnění.

6 Návrh a implementace

Cílem praktické části bakalářské práce je návrh a následná implementace projektu, který dokáže klasifikovat objekty z kamerového záznamu v reálném čase. Pro tento případ byl zvolen model konvoluční neuronové sítě.

Samotná implementace klasifikačního modelu je pouze nedílnou součástí celého projektu. Zprvu bylo třeba zajistit data k trénování modelu. Za vhodnou datovou sadu byla zvolena již připravená a volně dostupná sada s 43 třídami německého dopravního značení s celkovým počtem přes 50 000 obrazových záznamů. Tyto záznamy jsou rozděleny na trénovací a testovací. Data byla vhodně rozřazena do adresářů.

Zpracování všech obrazových dat k předání ve požadovaném tvaru bylo velice důležitým krokem. Postupně docházelo k normalizaci obrazu, změně tvaru a barevného kanálu, transformace vektoru tříd na binární matici a dělení datové sady na data trénovací a validační.

Poté byl model trénován a následně testován na cca 12 000 snímcích pro interpretaci výsledků. Samozřejmě došlo i na jednoduchý test modelu při klasifikaci obrazu z videozáznamu, kde předvedl téměř stejný výsledek v přesnosti, jako tomu bylo u snímků.

6.1 Použité technologie

Python

Python je interpretovaný, multiparadigmatický, dynamicky typovaný jazyk se širokým spektrem využití. Obsahuje nespočet modulů a knihoven pro strojové vidění, což byl z části důvod volby. (Python Software Foundation 2022)

OpenCV

OpenCV (Open source computer vision) je bezplatná multiplatformní knihovna strojového vidění. Rámec knihovny obsahuje nástroje, jiné knihovny a moduly, které mají vestavěnou podporu pro implementaci aplikací počítačového vidění. Je to jeden z nejrozšířenějších nástrojů s velice silnou vývojářskou komunitou

a historií sahající až do roku 1999. Je také populární svým velikým rozsahem komerčního použití v reálném světě. OpenCV obsahuje více než 500 funkcí, které napomáhají v mnohých oblastech strojového vidění. Podporuje C/C++, Python a také Javu. Vývoj OpenCV začal ve výzkumné laboratoři Intelu během naléhání na pokrok v přístupech pro vytváření aplikací náročných na CPU. Byl koncipován jako způsob, jak učinit infrastrukturu strojového vidění univerzálně dostupnou. (Gollapudi a Laxmikanth 2019; Darkcrizt 2020; Jiří 2010)

Numpy

Numpy je základní balíček určený pro vědecké výpočty pro jazyk Python. Poskytuje převážně různé variace polí (maskovaná pole, matice) a řadu operací, které se s nimi dají provádět. Hlavním důvodem jeho použití je 50krát vyšší rychlost oproti seznamům v jazyce Python díky použití opravdového pole. (NumPy Developers 2022; W3Schools 2022)

Sklearn

Sklearn (scikit-learn) je open-source knihovna určená pro strojové učení. Obsahuje desítky vestavěných algoritmů a modelů strojového učení a poskytuje nástroje pro přizpůsobení modelu, předzpracování dat, vyhodnocování. (scikit-learn developers 2022)

Keras

Keras je open-source knihovna fungující jako rozhraní pro TensorFlow a také poskytuje rozhraní jazyku Python pro umělé neuronové sítě. (Team 2022)

Mathplotlib

Mathplotlib je komplexní knihovna pro práci s vizualizacemi v jazyce Python. Také poskytuje objektově orientované API, které umožňuje vkládat grafy do aplikací pomocí obecných nástrojů GUI. (The Matplotlib Development team 2021)

TensorFlow

TensorFlow je open-source end-to-end knihovna určená pro strojové učení a umělou inteligenci. Obsahuje široké spektrum nástrojů, knihoven a různých zdrojů k usnadnění práce v odvětví strojového učení. (TensorFlow)

6.2 Datová sada pro CNN a klasifikační model

Pro trénování modelu byla zvolena již existující volně dostupná datová sada německých dopravních značek.

6.2.1 Datová sada

Datová sada obsahuje přes 50 000 digitálních obrázků kategorizovaných do 43 tříd. Rozložení elementů do jednotlivých tříd není rovnoměrné. V některých případech je rozdíl výrazný, což může ve finále vést ke komplikacím při klasifikaci tříd s nižším počtem záznamů. (Mykola 2018)

|---- data

|---- Meta # náhledové obrázky jednotlivých tříd - .png

|---- Test # obrázky pro testování - .png

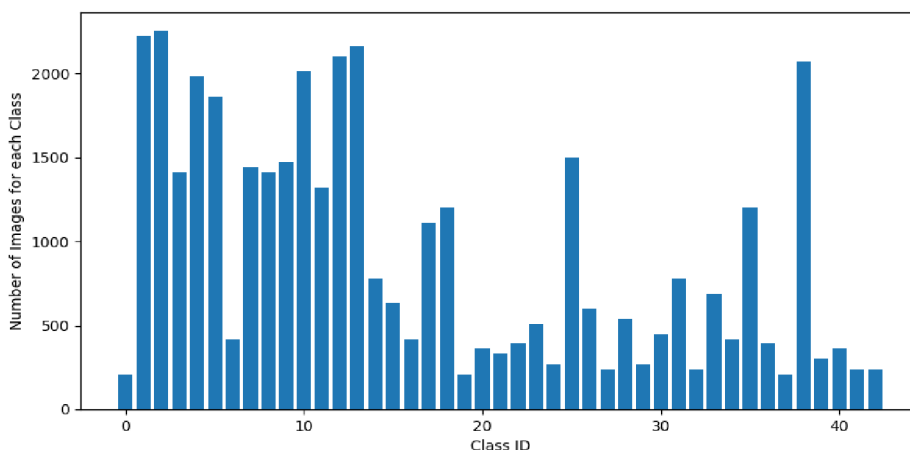
|---- Train

|---- 0 – 42 # podadresáře jednotlivých tříd s obrázky

|---- Meta.csv # informace o obrázcích v Meta adresáři - .csv

|---- Test.csv # informace o obrázcích v Test adresáři - .csv

|---- Train.csv # informace o obrázcích v Train adresáři - .csv

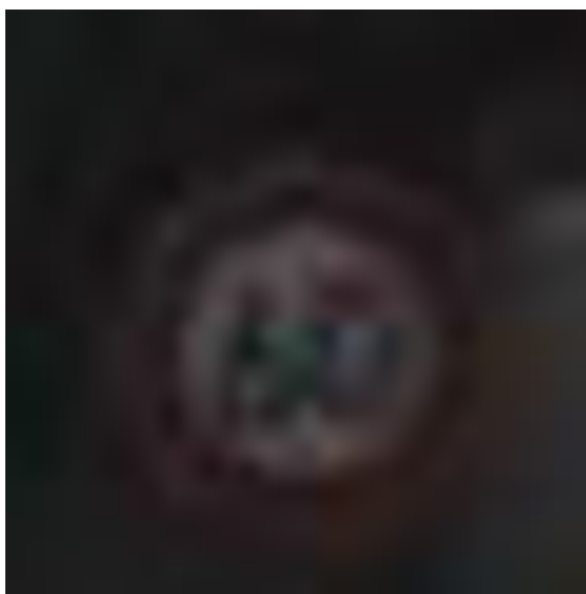


Obrázek 8 Sloupcový graf reprezentující počet snímků jednotlivých tříd (práce autora)

Součástí datové sady jsou dopravní značení zákazová (maximální povolená rychlost, zákaz vjezdu, zákaz předjíždění), výstražná (chodci, zvěř, práce na silnici, náledí), příkazová (směr jízdy) a značení upravující přednost (hlavní komunikace, dej přednost v jízdě, stop, kruhový objezd).

Každý podadresář (z celkem 43) ve složce data/Train, nesoucí název s číslem identifikační třídy, obsahuje rozdílný počet záznamů (viz Obrázek 8). V každém z nich se nachází obrázky příslušné třídy s lišící se kvalitou, úhlem, pod kterým byl objekt zachycen, velikostí, přiblížením, natočením. Některé mají tak nízkou kvalitu, že by pravděpodobně ani člověk na první pohled nebyl občas schopen rozeznat, o jaké značení se jedná. To vyloženě není na škodu, dokonce je to chtěné, jelikož ne vždy je obraz z videokamery zachycen perfektně. Tím pádem by měl být model schopen objekt klasifikovat, ačkoliv přesnost může být v tomto případě nižší, či klasifikuje objekt jako jemu podobný. Rozsah počtu snímků jednotlivých tříd se v datové sadě pohybuje v intervalu $< 210; 2250 >$.

Ve složce data/Test se nachází 12 630 obrázků, jenž budou poté použity pro účel evaluace natrénovaného modelu k zjištění jeho výsledné přesnosti.



Obrázek 9 Příklad obrázku dopravní značky s nízkou obrazovou kvalitou (Mykola 2018)



Obrázek 10 Příklad obrázku dopravní značky s vysokou obrazovou kvalitou (Mykola 2018)

6.2.2 Trénování modelu

Před zahájením samotného trénování neuronové sítě je nutné se všemi obrazovými záznamy v datové sadě provést několik operací, které přinesou lepší výsledky trénování.

Načítání datové sady

Jak již bylo zmíněno, k trénování modelu bude použita již připravená sada se strukturou, obsahující adresáře dělící obrazová data do podadresářů a také tabulkové soubory, ve kterých jsou uloženy informace o všech obrázcích jako je cesta, třída, šířka, výška atd.

Pro načítání dat z *Train* adresáře a přiřazování identifikační třídy nebyla zvolena cesta čtení rozsáhlé tabulky ze souboru *Train.csv*, ale cesta přímého přiřazování tříd jednotlivým obrázkům při jejich zpracování. Jednoduše z toho důvodu, že ze struktury je zřetelně vidět rozdělení obrazových záznamů do adresáře, jehož název nese číslo příslušné klasifikační třídy.

```

def prepare_train_data(path):
    images = []
    labels = []
    classCount = len(os.listdir(path))
    for x in range(0, classCount):
        showProgress("Loading train sources: ", x, classCount)
        imageDir = path + "/" + str(x)
        imageList = os.listdir(imageDir)

        for y in imageList:
            currentImg = cv2.imread(imageDir + "/" + y)
            currentImg = cv2.resize(currentImg, image_resolution) # 50x50
            images.append(currentImg)
            labels.append(x)
    print("")

    images = np.array(images)
    labels = np.array(labels)

    return images, labels

```

Obrázek 11 Ukázka kódu, který zajišťuje přípravu trénovacích dat (práce autora)

Následkem toho je tedy možné vždy každému obrázku přiřadit identifikační třídu dle toho, jaký adresář obrázků je zrovna čten, respektive se jedná o dvě pole s odpovídajícími hodnotami na stejných indexech. Pro konzistentnost dat jsou také při čtení upraveny jejich rozměry. Zde byla zvolena velikost 50 × 50 pixelů.

```

def prepare_test_images(path, file):
    images = []
    imageList = os.listdir(path + "Test")
    labels, paths = get_labels_and_paths(file)

    current = 0
    for imagePath in paths:
        showProgress(
            str("Loading test sources: "),
            current,
            len(imageList)
        )

        currentImg = cv2.imread(path + "/" + imagePath)
        currentImg = cv2.resize(currentImg, image_resolution) # 50x50

        images.append(currentImg)
        current += 1

    print("")
    images = np.array(images)

    return images, labels

```

Obrázek 12 Příprava testovacích dat (práce autora)

V druhém případě u přípravy testovacích dat v adresáři *Test* jsou prvně připraveny třídy i s cestou k cílovému souboru za pomoci čtení tabulkového souboru *Test.csv*, ze kterého jsou převzaty potřebné hodnoty (*classId*, *path*). Následně je analogicky, jako tomu bylo u trénovacích dat, vytvořeno pole obrázků s odpovídající třídou v poli tříd na stejném indexu a taktéž došlo ke změně velikosti.

Normalizace, změna rozměru a vektor tříd

Dalším důležitým krokem pro přípravu dat je normalizace transformováním barevného kanálu na stupně šedi, což nám z části eliminuje rozdílné osvětlení a také urychlí CNN.

```
def process_image(img):
    img = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY) # greyscale
    img = cv2.equalizeHist(img) # contrast
    img = img / 255 # converting to <0,1>
    return img

def reshape(npArray):
    return npArray.reshape(npArray.shape[0], npArray.shape[1],
                           npArray.shape[2], 1)
```

Obrázek 13 Zpracování obrazu, změna rozměru (práce autora)

Co se týče rozměrů obrázků, balíček *Keras* požaduje jeden rozměr navíc, který koresponduje s barevným kanálem. Jelikož došlo ke změně obrazu na monochromatický, je zapotřebí doplnit hodnotu 1. V případě RGB kanálu by byla doplněna hodnota 3. Následně jsou veškeré pixelové hodnoty, které se nachází v intervalu $\langle 0; 255 \rangle$, převedeny do intervalu $\langle 0; 1 \rangle$, což zajistí vyšší přesnost trénování. Při dělení celočíselných hodnot totiž dochází s nepřesností z důvodu zaokrouhlování.

```
# Converting to binary matrix
y_train = to_categorical(y_train, classCount)
y_valid = to_categorical(y_valid, classCount)
test_labels = to_categorical(test_labels, classCount)
```

Obrázek 14 Převod na binární matice (práce autora)

Další nedílnou součástí sekvence kroků je převedení kategorických dat na data číselná, respektive vektoru tříd na matici tříd. Jednou z technik používaných k této konverzi se nazývá *One-hot encoding*. Tento krok je nutný pro předání dat v požadovaném formátu modelu. (Educative)

Dělení datové sady

V OpenCV knihovně se nachází metoda *train_test_split*, které stačí poskytnout již připravená pole jak obrazových dat, tak příslušných tříd a definovat poměr dělení. Metoda vrátí rozdělená data v poměru, který jí byl určen. Byl zvolen poměr 4:1 (0,2), což znamená, že 80 % dat se použije pro trénování a 20 % pro validaci.

```
# Preparing data
train_images, train_labels = prepare_train_data(dataTrainPath)

# Splitting for train and validation
X_train, X_valid, y_train, y_valid = train_test_split(train_images,
                                                    train_labels, test_size=0.2)
```

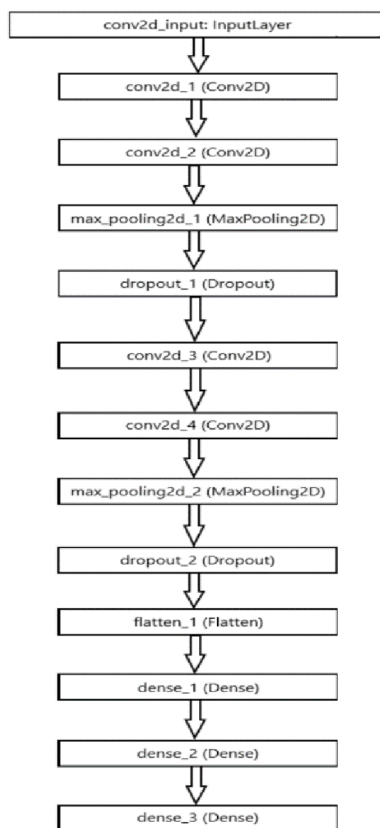
Obrázek 15 Dělení datové sady 80/20 (práce autora)

6.2.3 Klasifikační model

Model obsahuje celkově 12 vrstev. Mezi ně patří klíčové vrstvy Conv2D, dále také MaxPooling2D, Dropout, Flatten a Dense.

Prvními dvěma vrstvami jsou Conv2D, které se starají o práci se vstupními daty, jež přichází ve formátu dvourozměrné matice. Pro počet filtrů byla zvolena hodnota 32, jelikož při vyšších hodnotách by docházelo k duplicitám. Velikost kernelu pro tyto první dvě vrstvy je 5×5. Aktivační funkcí, jim přidělenou, je ReLU (Rectified Linear Activation), která se ve spojení s neuronovými sítěmi osvědčila. První vrstvě je také nutné definovat tvar vstupních dat, čímž se rozumí tvar všech vstupních obrázků. To bylo nastíněno v normalizaci dat, kde docházelo ke změně tvaru na (50, 50, 1), přičemž 1 znamená, že obrázky jsou ve stupních šedi. Model tedy bude očekávat data s tímto tvarem. Poté následuje MaxPooling vrstva, jež převzorkuje vstup. Dále se nachází Dropout vrstva napomáhající předcházení přehlcení. Stejně vrstvy se opět opakují s tím rozdílem, že zprvu nejsou použity dvě Conv2D vrstvy, nýbrž pouze jedna. Tuto část zakončuje vrstva Flatten vyrovnávající vstup.

Posledními vrstvami jsou dvě Dense vrstvy o velikosti 43 neuronů a opět již zmíněná Dropout, která přijde mezi ně.



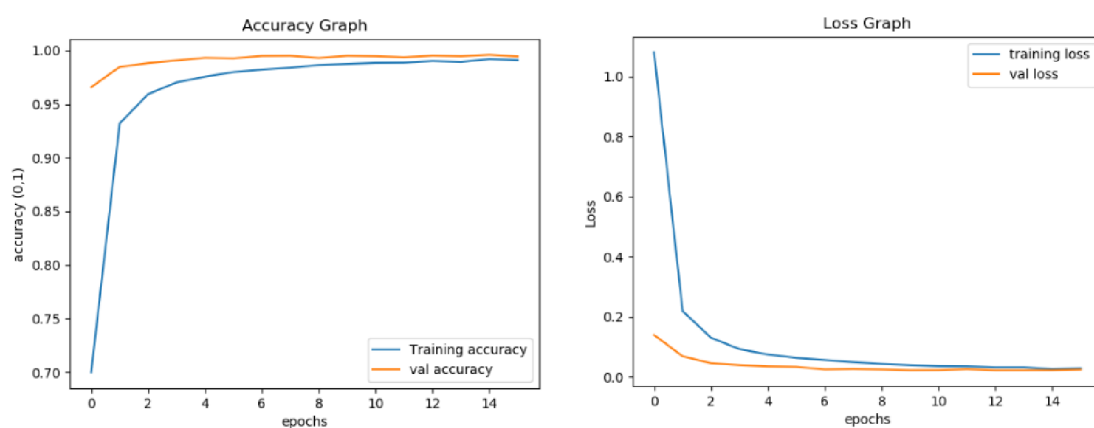
Obrázek 16 Posloupnost vrstev (práce autora)

Po vytvoření modelu bylo zahájeno trénování po dobu 20 epoch s definovanými zpětnými voláními EarlyStopping a ReduceLRonPlateau, které dokážou pozitivně ovlivnit průběh a výsledky trénování.

Algoritmus EarlyStopping má za cíl minimalizovat ztrátu tím, že při trénování modelu je tato metrika monitorována a na konci každé epochy je prováděna kontrola ztráty. Avšak je možné monitorovat i jiné metriky. Pokud se již hodnota sledované metriky nesnižuje, v závislosti na definovaných parametrech (min_delta a patience) je trénování předčasně ukončeno. Min_delta udává minimální změnu, podle které lze kvalifikovat, zda se jedná o zlepšení, či zhoršení. Patience říká, po kolika epochách bez zlepšení je trénování ukončeno.

ReduceLROnPlateau opět monitoruje udávanou metriku (zde stejně jako u EarlyStopping je hlídána ztráta) a jeho úkolem je po zjištění stagnace v učení, po daný počet epoch, snížit míru učení o definovanou hodnotu.

Zpětné volání EarlyStopping ukončilo trénování modelu již po 16. epoše, jelikož se přesnost klasifikace nijak výrazně neměnila a nemělo by smysl pokračovat. Je možné, že po dokončení všech 20 epoch by byla klasifikační přesnost dokonce menší než aktuální.

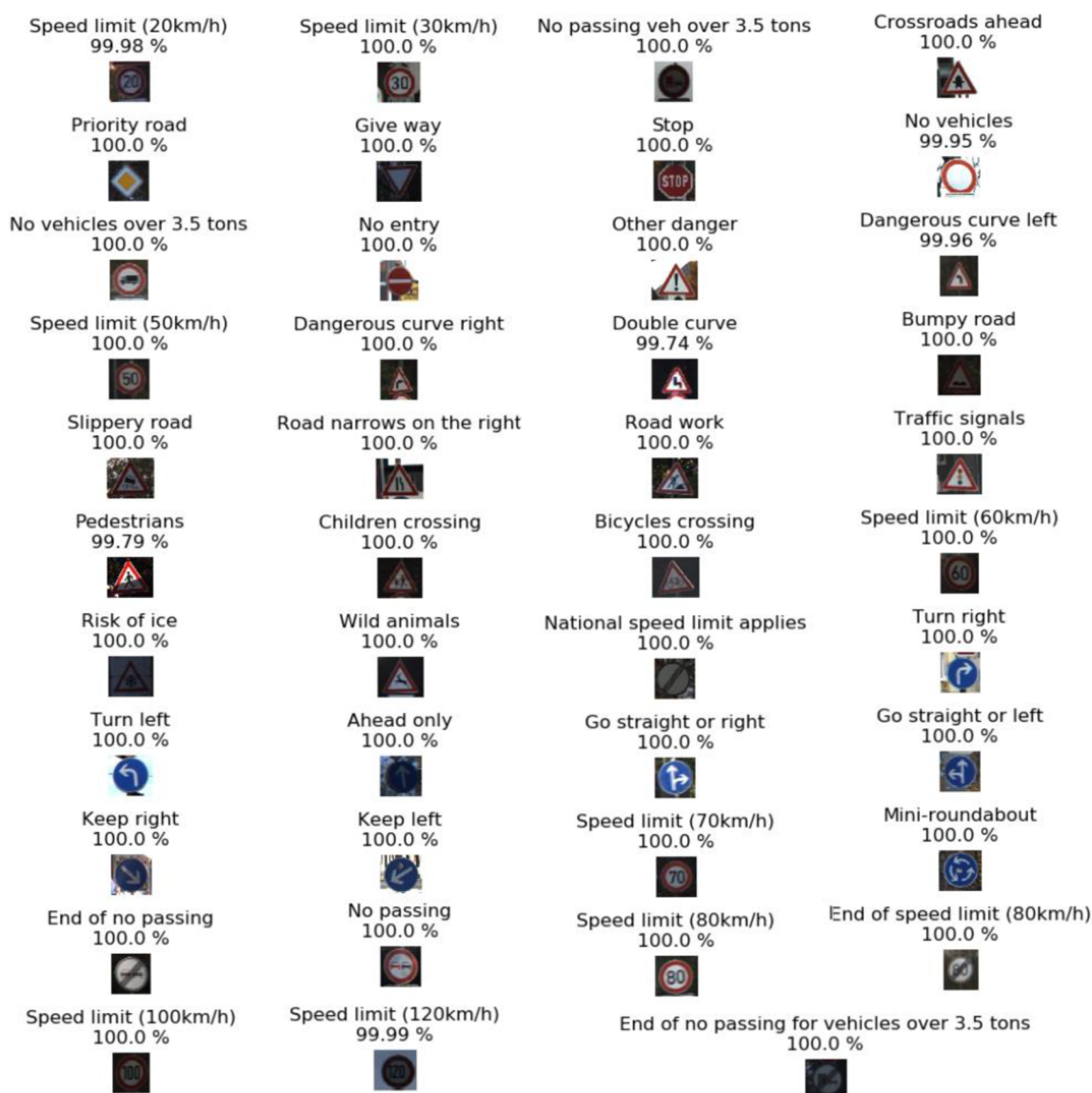


Obrázek 17 Dosažené výsledky trénování po dobu 20 (16) epoch s výslednou přesností 0,974 a ztrátovou funkcí 0,114 po evaluaci modelu na trénovacích datech (práce autora)

7 Shrnutí dosažených výsledků

7.1 Testování klasifikačního modelu

Přestože je projekt zaměřen na klasifikaci objektů z obrazu v reálném čase, pro účel interpretace výsledků v práci bylo vhodnější otestovat model pouze na vybraných snímcích. Dá se říct, že tento způsob se nijak neliší od testování na videu, jelikož práce se záznamem je podobná.



Obrázek 18 Výsledky predikce s uvedenou predikovanou třídou a přesností predikce (práce autora)

Pro grafickou interpretaci (viz Obrázek 18) byl vybrán z každé třídy jeden snímek, který prošel procesem klasifikace pomocí natrénovaného modelu. Snímky byly vybrány z datové sady ze složky s testovacími daty data/Test tak, aby byly zároveň rozpoznatelné lidským okem. Testování bylo tedy provedeno pokaždé na jednom snímku ze všech tříd.

K získání relevantnějších výsledků bylo provedeno testování na datech opět ze složky data/Train, kde se nachází 12 630 záznamů. Prvně došlo k uložení cest obrázků do dvourozměrného pole, kde index vnějšího pole určuje klasifikační třídu a ve vnitřním poli se nachází všechny cesty, které byly pro danou třídu načteny z csv souboru data/Test.csv. Následně bylo pole předáno modelu, kde se načtl obrázek s danou cestou a prošel procesem klasifikace.

Klasifikační třída	Počet záznamů pro trénování	Počet testovaných záznamů	Nesprávně klasifikováno	Průměrná přesnost
0	210	60	0	0,919
1	2220	720	0	0,988
2	2250	750	0	0,997
3	1410	450	0	0,964
4	1980	660	0	0,989
5	1860	630	0	0,981
6	420	150	0	0,915
7	1440	450	0	0,971
8	1410	450	0	0,967
9	1470	480	0	0,996
10	2010	660	0	0,983
11	1320	420	0	0,981
12	2100	690	0	0,991
13	2160	720	0	0,998
14	780	270	0	0,999
15	630	210	0	0,994

16	420	150	0	0,998
17	1110	360	0	0,998
18	1200	390	0	0,973
19	210	60	0	0,999
20	360	90	0	0,915
21	330	90	0	0,858
22	390	120	0	0,992
23	510	150	0	0,968
24	270	90	0	0,987
25	1500	480	0	0,973
26	600	180	0	0,903
27	240	60	0	0,950
28	540	150	0	0,992
29	270	90	0	0,995
30	450	150	0	0,884
31	780	270	0	0,979
32	240	60	0	0,989
33	689	210	0	0,990
34	420	120	0	0,988
35	1200	390	0	0,997
36	390	120	0	0,990
37	210	60	0	0,974
38	2070	690	0	0,995
39	300	90	0	0,985
40	360	90	0	0,888
41	240	60	0	0,957
42	240	90	0	0,956

Tabulka 1 Výsledky testování

Tabulka 1 znázorňuje, s kolika obrázky z každé třídy byl model trénován a také s kolika byl následně testován. Tento nedostatek (rozdíl v počtu) byl zmíněn při charakteristice použité datové sady. Každopádně zde byl i ten nejnižší počet

snímků pro trénování u několika tříd (210) dostačující pro správnou klasifikaci. Výkyvy v přesnosti jsou zde minimální. Směrodatná odchylka nabývá hodnoty 0,035. Model byl natolik přesný, že nedošlo k žádné chybné klasifikaci.



Obrázek 19 Predikce z videozáznamu (práce autora)

Jelikož byla práce primárně zaměřena na klasifikaci objektu z videozáznamu, byl model testován i na záznamu z videokamery, kde byl sejmout pouze snímek obrazovky pro účet interpretace. Zde stejně jako u předešlého testování byl model schopen správně klasifikovat objekt s přesností odpovídající té tabulkové.

7.2 Výsledky

Klasifikační model prokázal svou vysokou přesnost již při trénovací fázi, kde se tato hodnota blížila 97,5 %, což je průměrná přesnost při evaluaci. Aby nedošlo k přetrénování modelu, použité zpětné volání *EarlyStopping* ukončilo proces trénování již po 16 epochách. Takže se k původním 20 epochám model ani nedostal, což v tomto případě bylo chtěné, jelikož výsledná přesnost bez brzkého ukončení mohla být i nižší, než aktuálně je.

Svou přesnost předvedl i v testovací fázi, kdy mu bylo poskytnuto prvně po jednom obrázku z každé třídy a poté velké množství testovacích (12 630 obrázků) k provedení klasifikace, kde model uspěl s průměrnou přesností 0,970 (97 %) se směrodatnou odchylkou 0,035. Přestože některé obrazové záznamy nebyly dosti kvalitní (malé rozlišení, jas, šum, rozmazání), nedošlo ani k jedné chybné

klasifikaci. Dále byl stanoven předpoklad, že bude mít model drobné potíže s klasifikací objektů, které náleží třídě s nízkým počtem trénovacích záznamů. Tento předpoklad se neprokázal a je tedy možné říci, že nejnížší počet (210) obrazových záznamů několika tříd je plně dostačující pro správnou klasifikaci.

8 Závěr

Cílem této práce bylo nastínit čtenáři problematiku spojenou s umělou inteligencí, strojovým viděním a také navrhnout a implementovat klasifikační model, který dokáže v obrazovém záznamu klasifikovat vybraný objekt.

V teoretické části byl zprvu definován pojem umělá inteligence s následným nahlédnutím do historie a představen významný Turingův test. Poté práce nastínila základní techniky a principy umělé inteligence, mezi které spadá expertní systémy, fuzzy logika a umělé neuronové sítě a představila příklady jejich užití v reálném světě. Dále byla představena podoblast umělé inteligence zvaná strojové vidění, její principy, doprovázející problémy a reálné využití.

Praktická část řešila úlohu klasifikace objektu v obraze pomocí konvolučních neuronových sítí, kde byla popsána sekvence kroků pro předzpracování jednotlivých záznamů z datové sady a vložení do vhodných datových struktur, které byly poté užity k trénování modelu. Pro interpretaci výsledků byl model testován na snímcích jednotlivých tříd, kde prokázal vysokou jistotu při evaluaci po trénování s průměrnou hodnotou 97,4 % a následně při samotném testování s průměrnou hodnotou 97 %. Všechny testovací snímky byly klasifikovány správnou třídou.

Dalšími aspekty pro rozvoj práce by definitivně mohly být detekce objektu v obraze pro následnou klasifikaci, rozšíření aktuální datové sady o nové třídy a větší počet snímků u některých méně početných tříd a tvorba GUI pro modifikaci předmětu klasifikace (videozáznam v reálném čase, uložený videozáznam, snímek).

9 Seznam použité literatury

- ANAND, Nithi, 2015. *Low angle of The Eiffel Tower* [online]. photo. 2. srpen 2015. [vid. 2022-04-07]. Dostupné z: <https://www.flickr.com/photos/nithiclicks/26950990696/>
- BBC, 2014. Computer AI passes Turing test in „world first“. *BBC News* [online]. [vid. 2022-02-17]. Dostupné z: <https://www.bbc.com/news/technology-27762088>
- CELBOVÁ, Iva, 1999. *Úvod do problematiky expertních systémů* | *Ikaros* [online] [vid. 2022-02-17]. Dostupné z: <https://ikaros.cz/uvod-do-problematiky-expertnich-systemu>
- COMPUTERSCIENCEWIKI, 2018. *File:MaxpoolSample2.png - Computer Science Wiki* [online] [vid. 2022-04-07]. Dostupné z: <https://computersciencewiki.org/index.php/File:MaxpoolSample2.png>
- CONTROL ENGINEERING ČESKO, 2019. Co je to strojové vidění a jak může pomoci? *Vše o průmyslu* [online] [vid. 2022-03-01]. Dostupné z: <https://www.vseoprumsly.cz/automatizace/kontrola-procesu/co-je-to-strojove-videni-a-jak-muze-pomoci.html>
- DARKCRIZT, 2020. OpenCV knihovna pro rozpoznávání objektů v obrazech a fotoaparátech. *Desde Linux* [online]. [vid. 2022-03-01]. Dostupné z: <https://blog.desdelinux.net/cs/opencv-knihovna-pro-rozpozn%C3%A1v%C3%A1n%C3%AD-objekt%C5%AF-v-obrazech-a-fotoapar%C3%A1tech/>
- DILMEN, Ataturk svg: NevitNevit, 2011. *Graphical representation* [online]. 11. srpen 2011. [vid. 2022-04-07]. Dostupné z: https://commons.wikimedia.org/wiki/File:Face_or_vase_ata_01.svg
- EDUCATIVE, [b.r.]. One-hot encoding in Python. *Educative: Interactive Courses for Software Developers* [online] [vid. 2022-03-20]. Dostupné z: <https://www.educative.io/edpresso/one-hot-encoding-in-python>
- FÉRÉE, Hugo, 2011. *English: The „standard interpretation“ of the Turing Test, in which player C, the interrogator, is tasked with trying to determine which player - A or B - is a computer and which is a human. The interrogator is limited to only using the responses to written questions in order to make the determination.* [online].

21. říjen 2011. [vid. 2022-04-07]. Dostupné z: https://commons.wikimedia.org/wiki/File:Turing_Test_Version_3.svg

GOLLAPUDI, Sunila a V. LAXMIKANTH, 2019. *Learn computer vision using OpenCV: with deep learning CNNs and RNNs*. Berkeley, Calif.: Apress. Professional and applied computing. ISBN 978-1-4842-4260-5.

GUPTA, Abhishek, Alagan ANPALAGAN, Ling GUAN a Ahmed Shaharyar KHWAJA, 2021. Deep learning for object detection and scene perception in self-driving cars: Survey, challenges, and open issues. *Array* [online]. **10**, 100057 [vid. 2022-02-17]. ISSN 2590-0056. Dostupné z: doi:10.1016/j.array.2021.100057

HAVLE, Otto, 2008. *Strojové vidění I: Principy a charakteristiky*. 4.

ISAKOVA, Tatsiana, 2021. Top Computer Vision Examples in 2021. *InData Labs* [online] [vid. 2022-02-28]. Dostupné z: <https://indatalabs.com/blog/computer-vision-examples>

JIŘÍ, Prymus, 2010. Rozšíření knihovny pro zpracování obrazu. 81.

KAY, Alexx, 2001. *Co jsou to umělé neuronové sítě?* [online]. [vid. 2022-04-08]. Dostupné z: <https://www.scienceworld.cz/technologie/co-jsou-to-umele-neuronove-site-4077/>

MARR, David, 2010. *Vision: a computational investigation into the human representation and processing of visual information*. Cambridge, Mass: MIT Press. ISBN 978-0-262-51462-0.

MAŘÍK, V., O. ŠTĚPÁNKOVÁ a Jiří LAŽANSKÝ, 1993. *Umělá inteligence*. Praha: Academia. ISBN 978-80-200-0502-1.

MICHELLE, 2012. High Resolution Images VS. Low Resolution Images: A Short Primer for Beginners. *Contemporary Communications, Inc.* [online]. [vid. 2022-04-07]. Dostupné z: <https://ccideas.com/chatter/high-resolution-images-vs-low-resolution-images-a-short-primer-for-beginners/>

MOLNÁR, Karol, 2000. *Úvod do problematiky umělých neuronových sítí* [online] [vid. 2022-02-17]. Dostupné z: <http://www.elektrorevue.cz/clanky/00013/index.html>

MONTBLANC, 2020. *Jak fungují konvoluční vrstvy v neuronových sítích s hlubokým učením? | Mont Blanc* [online]. [vid. 2022-03-23]. Dostupné z: <https://montblanczone.com/cs/jak-funguji-konvolucni-vrstvy-v-neuronovych-sitich-s-hlubokym-ucenim/>

MYKOLA, 2018. *GTSRB - German Traffic Sign Recognition Benchmark* [online] [vid. 2022-03-20]. Dostupné z: <https://kaggle.com/datasets/meowmeowmeowmeowmeow/gtsrb-german-traffic-sign>

MYRABELLA, 2012. *Français : La tour Eiffel vue de la Cité de l'architecture et du patrimoine au palais de Chaillot (Paris, France)* [online]. 1. listopad 2012. [vid. 2022-04-07]. Dostupné z: https://commons.wikimedia.org/wiki/File:Eiffel_tower_from_Cite_Architecture_Chailot.jpg

NUMPY DEVELOPERS, 2022. *NumPy documentation — NumPy v1.22 Manual* [online] [vid. 2022-03-23]. Dostupné z: <https://numpy.org/doc/stable/>

POWELL, Selina, 2018. *Scientists teach computers to recognise optical illusions* [online] [vid. 2022-02-21]. Dostupné z: <https://www.aop.org.uk/ot/science-and-vision/research/2018/09/27/scientists-teach-computers-to-recognise-optical-illusions>

PYTHON SOFTWARE FOUNDATION, 2022. *General Python FAQ — Python 3.10.3 documentation* [online] [vid. 2022-03-23]. Dostupné z: <https://docs.python.org/3/faq/general.html#what-is-python>

RICH, Elaine a Kevin KNIGHT, 1991. *Artificial intelligence*. 2. ed. New York: McGraw-Hill. ISBN 978-0-07-100894-5.

RIZZOLI, Alberto, 2022. *27+ Most Popular Computer Vision Applications in 2022* [online] [vid. 2022-02-28]. Dostupné z: <https://www.v7labs.com/blog/computer-vision-applications>

ŘEZNÍČKOVÁ, Tereza, 2019. *Umělá inteligence v praxi* [online] [vid. 2022-02-17]. Dostupné z: <https://www.spcr.cz/aktivity/z-hospodarske-politiky/12904-umela-inteligence-v-praxi>

SCIKIT-LEARN DEVELOPERS, 2022. *Getting Started. scikit-learn* [online] [vid. 2022-03-23]. Dostupné z: https://scikit-learn/stable/getting_started.html

SEARLE, John R, 1980. Minds, brains, and programs. *THE BEHAVIORAL AND BRAIN SCIENCES*. 8.

SCHEJBAL, Ctírad, Vladimír HOMOLA a František STANĚK, 2004. *Geoinformatika - Expertní a neuronové systémy* [online] [vid. 2022-02-17]. Dostupné z: <http://geologie.vsb.cz/geoinformatika/kap07.htm>

TEAM, Keras, 2022. *Keras documentation: About Keras* [online] [vid. 2022-03-23]. Dostupné z: <https://keras.io/about/>

TENSORFLOW, [b.r.]. TensorFlow. *TensorFlow* [online] [vid. 2022-03-23]. Dostupné z: <https://www.tensorflow.org/>

THE MATPLOTLIB DEVELOPMENT TEAM, 2021. *Matplotlib — Visualization with Python* [online] [vid. 2022-03-23]. Dostupné z: <https://matplotlib.org/>

TRENZ, Oldřich, 2009. *Umělá inteligence I* [online] [vid. 2022-02-17]. Dostupné z: https://is.mendelu.cz/eknihovna/opory/zobraz_cast.pl?cast=21852

ULDRICH, Miloš a Tomáš JURCZYK, 2014. Neuronové sítě a jejich využití. 2.

VACEK, Ludvík, 2000. *Historie a současnost umělé inteligence* [online] [vid. 2022-02-28]. Dostupné z: <https://www.fi.muni.cz/usr/jkucera/pv109/2000/xvacek.htm>

VOLNÁ, Eva, 2013. UMĚLÁ INTELIGENCE. 138.

W3SCHOOLS, 2022. *Introduction to NumPy* [online] [vid. 2022-04-07]. Dostupné z: https://www.w3schools.com/python/numpy/numpy_intro.asp

WIENER, Norbert, 1988. *The human use of human beings: cybernetics and society*. New York, N.Y: Da Capo Press. The Da Capo series in science. ISBN 978-0-306-80320-8.

WILSON, Gregory, 2012. *English: Sarah, fastest cheetah* [online]. 2012. [vid. 2022-04-07]. Dostupné z: [https://commons.wikimedia.org/wiki/File:Sarah_\(cheetah\).jpg](https://commons.wikimedia.org/wiki/File:Sarah_(cheetah).jpg)

ZELINKA, Ivan, 2003. *Umělá inteligence - hrozba nebo naděje?* Praha: BEN - technická literatura. ISBN 978-80-7300-068-4.



Zadání bakalářské práce

Autor: Lukáš Kučera

Studium: I1900218

Studijní program: B1802 Aplikovaná informatika

Studijní obor: Aplikovaná informatika

Název bakalářské práce: Rozpoznávání objektů pomocí umělých neuronových sítí

Název bakalářské práce AJ: Object recognition using artificial neural networks

Cíl, metody, literatura, předpoklady:

Cíl práce: Prozkoumat problematiku strojového vidění a implementovat ukázkovou aplikaci pro zvolenou úlohu.

1. Nastínit základní pojmy spojené s umělou inteligencí
2. Prozkoumat principy strojového vidění
3. Návrh implementace aplikace pro detekci vybraných objektů
4. Implementace a testování navrhovaného řešení
5. Závěr

1) GOLLAPUDI, Sunila. Learn Computer Vision Using OpenCV: With Deep Learning CNN and RNN. APRESS L.P., 2019. ISBN 1484242602.

2) GUPTA, Abhishek, Alagan ANPALAGAN, Ling GUAN a Ahmed KHWAJA. Deep learning for object detection and scene perception in self-driving cars: Survey, challenges, and open issues. Array [online]. 2021 [cit. 2021-10-14]. Dostupné z: <https://www.sciencedirect.com/science/article/pii/S2590005621000059>

Garantující pracoviště: Katedra informatiky a kvantitativních metod,
Fakulta informatiky a managementu

Vedoucí práce: Ing. Milan Košťák

Datum zadání závěrečné práce: 12.1.2021