



ZADÁNÍ BAKALÁ ŠKÉ PRÁCE

Název:	Detekce objekt v pta ích hnízdech pomocí neuronových sítí
Student:	Jan Havl j
Vedoucí:	Ing. Josef Pavlí ek, Ph.D.
Studijní program:	Informatika
Studijní obor:	Softwarové inženýrství
Katedra:	Katedra softwarového inženýrství
Platnost zadání:	Do konce letního semestru 2017/18

Pokyny pro vypracování

Navhnete a implementujete knihovnu umož ňující zpracovávat data z kamery umíst ěné v hnízdu (pta í budce) s cílem rozpoznat po et vajec v hnízdu . Pro rozpoznání použijte existující algoritmy um ělé inteligence využívající data získaná ze serveru PtaciOnline.cz a další sw. nástroje projektu BirdObserver (athena.pef.czu.cz).

Postupujte v t ěchto krocích:

1. Prove te detailní specifikaci požadavk .
2. Seznamte se s projektem BirdObserver a strukturou dat na serveru PtaciOnline.cz.
3. Prove te analýzu a návrh knihovny.
4. Návrh implementujte, zdokumentujte a vhodným zp sobem otestujte.
5. Knihovnu navrhnete a implementujte v jazyce Java tak, aby bylo možné ji formou závislostí (dependency) integrovat s ostatními nástroji projektu BirdObserver.

Seznam odborné literatury

Dodá vedoucí práce.

Ing. Michal Valenta, Ph.D.
vedoucí katedry

prof. Ing. Pavel Tvrdí k, CSc.
d ěkan

V Praze dne 12. ledna 2017



**FAKULTA
INFORMAČNÍCH
TECHNOLOGIÍ
ČVUT V PRAZE**

Bakalářská práce

Detekce objektů v ptačích hnízdech pomocí neuronových sítí

Jan Havlůj

Katedra softwarového inženýrství

Vedoucí práce: Ing. Josef Pavlíček, Ph.D.

7. ledna 2018

Poděkování

Chtěl bych poděkovat vedoucímu mé bakalářské práce, panu Ing. Josefu Pavlíčkovi, Ph.D., za jeho čas, ochotu a pomoc při vedení této práce. Děkuji svým rodičům a přítelkyni, kteří mi byli po celou dobu studia velkou oporou. V neposlední řadě bych chtěl poděkovat svým spolužákům, kteří mě při studiu nikdy neváhali podpořit.

Prohlášení

Prohlašuji, že jsem předloženou práci vypracoval(a) samostatně a že jsem uvedl(a) veškeré použité informační zdroje v souladu s Metodickým pokynem o etické přípravě vysokoškolských závěrečných prací.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona, ve znění pozdějších předpisů. V souladu s ust. § 46 odst. 6 tohoto zákona tímto uděluji nevýhradní oprávnění (licenci) k užití této mojí práce, a to včetně všech počítačových programů, jež jsou její součástí či přílohou, a veškeré jejich dokumentace (dále souhrnně jen „Dílo“), a to všem osobám, které si přejí Dílo užít. Tyto osoby jsou oprávněny Dílo užít jakýmkoli způsobem, který nesnižuje hodnotu Díla, a za jakýmkoli účelem (včetně užití k výdělečným účelům). Toto oprávnění je časově, teritoriálně i množstevně neomezené. Každá osoba, která využije výše uvedenou licenci, se však zavazuje udělit ke každému dílu, které vznikne (byť jen zčásti) na základě Díla, úpravou Díla, spojením Díla s jiným dílem, zařazením Díla do díla souborného či zpracováním Díla (včetně překladu), licenci alespoň ve výše uvedeném rozsahu a zároveň zpřístupnit zdrojový kód takového díla alespoň srovnatelným způsobem a ve srovnatelném rozsahu, jako je zpřístupněn zdrojový kód Díla.

V Praze dne 7. ledna 2018

.....

České vysoké učení technické v Praze
Fakulta informačních technologií

© 2018 Jan Havlůj. Všechna práva vyhrazena.

Tato práce vznikla jako školní dílo na Českém vysokém učení technickém v Praze, Fakultě informačních technologií. Práce je chráněna právními předpisy a mezinárodními úmluvami o právu autorském a právech souvisejících s právem autorským. K jejímu užití, s výjimkou bezúplatných zákonných licencí a nad rámec oprávnění uvedených v Prohlášení na předchozí straně, je nezbytný souhlas autora.

Odkaz na tuto práci

Havlůj, Jan. *Detekce objektů v ptačích hnízdech pomocí neuronových sítí*. Bakalářská práce. Praha: České vysoké učení technické v Praze, Fakulta informačních technologií, 2018. Dostupný také z WWW: (<https://github.com/havluj/EggDetector>).

Abstrakt

Tato práce se zabývá návrhem a tvorbou softwarové knihovny pro detekci objektů v obraze použitím neuronových sítí.

V první části jsme seznámeni s cílem práce a detailní specifikací požadavků. Druhá část se věnuje teorii počítačového vidění, neuronových sítí a TensorFlow. Následně je na základě získaných teoretických znalostí a specifikace požadavků vypracována analýza možného řešení. Dle analýzy jsou navržena a implementována dvě řešení. První implementace se zaměřuje na detekci objektů, druhá na rozpoznávání a klasifikaci obrazu. V poslední části práce je vybráno efektivnější řešení, které je řádně ověřeno a otestováno.

Výsledkem je softwarová knihovna, která umožňuje automaticky rozpoznávat počet vajec v daném videu. Celý program je implementován v jazyce Java a je možné ho integrovat do projektu BirdObserver.

Klíčová slova neuronové sítě, počítačové vidění, detekce objektů, strojové učení, učení s učitelem, rozpoznávání obrazu, detekce vajec, ptačí hnízda, Java, TensorFlow

Abstract

This thesis focuses on designing and creating a software library that is able to detect objects in images using neural networks.

In the first part of the thesis, goals and technical requirements are specified. The second part discusses the theory behind computer vision, neural networks, and TensorFlow. Using the technical specification and acquired theoretical knowledge, a detailed analysis of a possible solution is presented. There are two implementations of the analysis. The first implementation is developed using object detection, the second one using image recognition. In the last part of the thesis, a more effective implementation is chosen, which is then properly tested and verified.

The result of this thesis is a software library that is able to automatically detect the number of eggs in a given video sequence. The entire solution is written in Java and is easily intergratable with other parts of the BirdObserver project.

Keywords neural networks, computer vision, object detection, machine learning, supervised learning, image recognition, egg detection, bird nests, Java, TensorFlow

Obsah

Úvod	17
Struktura práce	17
1 Cíl práce	19
1.1 Nefunkční požadavky	19
1.2 Funkční požadavky	19
2 Rešerše	21
2.1 Počítačové vidění	21
2.2 Neuronové sítě	22
2.3 TensorFlow	23
3 Analýza a návrh	25
3.1 Projekt Ptáci Online	25
3.2 Představení vstupních dat	25
3.3 Současný způsob zpracování dat	26
3.4 Nefunkční požadavky	26
3.5 Funkční požadavky	27
3.6 Návrh řešení	29
3.7 Volba technologií	31
3.8 Shrnutí kapitoly	32
4 Nástroje	33
4.1 Hromadné stahování dat	33
4.2 Příprava trénovacích a testovacích dat	34
5 Implementace detekování objektů	39
5.1 Příprava vývojového prostředí	39
5.2 Příprava dat	40
5.3 Struktura projektu	41

5.4	Trénování neuronové sítě	41
5.5	Ověření funkčnosti a výsledky	46
5.6	Závěr	47
6	Implementace rozpoznávání obrazu	51
6.1	Příprava dat	52
6.2	Trénování neuronové sítě	53
6.3	Ověření funkčnosti a výsledky	54
6.4	Závěr	54
7	Implementace Java knihovny	57
7.1	Srovnání implementací neuronových sítí a volba řešení	57
7.2	Použití TensorFlow v Javě	57
7.3	Uživatelské rozhraní knihovny	58
7.4	Implementace	58
7.5	Distribuce knihovny	60
7.6	Výsledek	60
8	Ověření implementace	63
8.1	Testování funkčnosti	63
8.2	Testování distribuce	64
8.3	Měření efektivity implementace	65
8.4	Shrnutí kapitoly	66
	Závěr	67
	Literatura	69
A	Seznam použitých zkratk	73
B	Dokumentace API knihovny	75
B.1	Package org.cvut.havluja1.eggdetector	75
C	Tagger	81
C.1	Dokumentace (v AJ)	81
C.2	Zdrojový kód	81
C.3	Ukázkový výstup	81
D	Folder Trimmer	85
D.1	Dokumentace (v AJ)	85
D.2	Zdrojový kód	86
E	Konfigurace detekování objektů	89
F	Test efektivity knihovny	95

G Použité programy	99
H Obsah přiloženého média	101

Seznam obrázků

2.1	Struktura a funkce neuronu v umělé neuronové síti.	22
2.2	Struktura neuronové sítě.	23
3.1	Ukázka neupraveného snímku z ptačí budky.	26
3.2	Vejce nemusí být vždy viditelná.	27
3.3	Program určený ke klasifikaci obrazu.	29
3.4	Program určený k detekci objektů.	30
4.1	Uživatelské rozhraní nástroje LabelImg.	37
4.2	Tagger – hromadné určování počtu vajec.	38
5.1	Předtrénované modely poskytované společností Google.	42
5.2	Vizualizace průběhu trénování programem TensorBoard.	44
5.3	Úvodní obrazovka programu TensorBoard: přehled stavu trénování.	45
5.4	Použití Google Cloud Platform.	46
5.5	Model je schopný detekovat všech devět vajec i bez ohledu na to, že některá vejce jsou sotva viditelná.	47
5.6	Model je schopný detekovat i vejce v jiném prostředí.	48
5.7	Model detekuje vejce i ve složitých podmínkách.	48
5.8	Méně úspěšný model nedokáže odhalit špatně viditelné vejce.	49
7.1	Ukázka debugovacího režimu knihovny EggDetector.	61

Seznam algoritmů

2.1	Použití algoritmu „stochastic gradient descent“ poskytovaný knihovnou TensorFlow.	24
3.1	Prvotní návrh uživatelského rozhraní knihovny.	28
3.2	Snímek reprezentovaný jako vektor je vstupem pro neuronovou síť.	31
4.1	Hromadné stahování dat ze serveru athena.pef.czu.cz.	33
4.2	Výběr vhodných složek pro trénování nástrojem Tagger.	35
4.3	Ukázkový výstup programu LabelImg.	36
7.1	Získání finálního počtu vajec pro celou složku pomocí dat z jednotlivých snímků.	58
7.2	Transformace vstupního snímku bilineární interpolací.	59
7.3	Detekce objektů pomocí TensorFlow v Javě.	59
8.1	Testování softwarové knihovny EggDetector.	63
C.1	Ukázkový výstup programu Tagger.	82
D.1	Zdrojový kód nástroje FolderTrimmer.	86
E.1	Konfigurační soubor pro detekci vajec v ptačích hnízdech.	89
F.1	Test k vyhodnocení efektivity výsledné knihovny.	95

Úvod

Projekt Ptáci Online byl spuštěn Fakultou životního prostředí České zemědělské univerzity v Praze roku 2014. Hlavním cílem projektu je poskytnout vědecká data, ve formě videa z ptačích budek, široké veřejnosti [1]. Projekt se těší poměrně velké popularitě [2] a aktuálně spolupracuje s více než dvěma desítkami spolupracovníků. Mezi ně patří lidé z akademické sféry, soukromého sektoru, ale i z Ministerstva životního prostředí České republiky [1]. Vzhledem k množství pořizovaných záznamů by bylo žádoucí získávat informace z videa automaticky pomocí algoritmů umělé inteligence. Předmětem této práce je vytvořit algoritmus, který je schopen určit počet vajec v hnízdě v daném videozáznamu.

Existuje několik způsobů, jak naučit počítač „vidět“. Téměř vždy se musíme zaměřit na předzpracování obrazu, jeho vlastnosti a jeho segmentaci. Jako řešení pak můžeme zvolit různou sadu deterministických algoritmů, například pro detekci hran nebo na samotné klasifikování segmentovaného obrazu. Všechny tyto algoritmy mají však jednu společnou nevýhodu. Formálně popsat tvar nějakého objektu a vytvořit sadu pravidel, podle kterých poznáme, zda-li se jedná o hledaný objekt, je velmi těžké. Světelné podmínky nemusí být ideální, objekty se mohou překrývat, mohou být různě barevné, vzdálené nebo otočené. Všechny tyto problémy znemožňují vytvoření perfektního algoritmu manuálně. Lepším řešením je použití algoritmů umělé inteligence, jako jsou například umělé neuronové sítě. Můžeme manuálně vytvořit algoritmus, který dokáže „vytrénovat“ druhý, obecný, algoritmus k vyřešení konkrétního problému. V této práci se budeme soustředit na použití neuronových sítí pro vyřešení problémů počítačového vidění.

Struktura práce

Práce je rozdělena do čtyř částí. První část obsahuje stanovení cílů a specifikaci funkčních i nefunkčních požadavků. V další části diskutujeme teorii

počítačového vidění, neuronových sítí a knihovny TensorFlow. Třetí, analytická část, se zaměřuje na výběr vhodných technologií a postupů pro tvorbu implementace. Čtvrtá, praktická část, je rozdělena do několika kapitol. Nejprve popisuje použití a vývoj všech potřebných nástrojů. V dalších kapitolách jsou popsány dva různé způsoby implementace, jejich porovnání, otestování a validace. Na závěr je vybrána efektivnější implementace, dle které je vytvořena softwarová knihovna pro určení počtu vajec v hníždě.

Cíl práce

Cílem teoretické části práce je seznámit se s technologiemi a principy, které jsou dnes standardně používány k řešení problémů spjatých s počítačovým viděním a strojovým učením. Zaměříme se především na umělé neuronové sítě, jejich historii, současný vývoj a hlavně na principy jejich fungování. Na základě získaných teoretických znalostí je vypracována analýza řešení, která může být považována za výstup teoretické části. Cílem analýzy je jednak selekce vhodných principů a technologií pro implementaci řešení, ale i zpracování technických požadavků zároveň s popsáním současného stavu.

Cílem praktické části je implementovat řešení, které bude dostatečně rychlé a spolehlivé. Výstupem bude softwarová knihovna, která bude umožňovat rozpoznávání počtu vajec v hnízdě na jednotlivých snímcích i na sekvencích videa. Součástí knihovny bude neuronová síť, která bude řešit samotnou detekci vajec. Tento přístup umožní recyklaci výsledného programu s minimální modifikací pro řešení podobných problémů, jako například určení druhu ptáka, počtu mláďat nebo predaci hnízda.

Následující seznam funkčních a nefunkčních požadavků slouží jako přesný popis požadované formy výsledné knihovny. Tento seznam byl vytvořen ve spolupráci s vedoucím práce.

1.1 Nefunkční požadavky

- N1** Knihovna musí pracovat v reálném čase. Po zadání vstupních dat je výsledek očekáván maximálně v jednotkách sekund.

1.2 Funkční požadavky

- F1** Systém bude distribuovaný formou Java knihovny v archivu JAR¹.

¹Zkratka pro Java ARchive. „JAR je kompresní souborový formát, používaný platformou Java, založený na ZIP kompresi.“[3]

- F2** Knihovna bude distribuována ve dvou verzích:
- JAR archiv obsahující, jak přeložený zdrojový kód knihovny, tak i všechny závislosti², které jsou knihovnou vyžadovány.
 - JAR archiv obsahující pouze přeložený zdrojový kód knihovny bez externích závislostí. Do archivu bude přibalen konfigurační soubor pro systém Maven[4] s definicí závislostí.
- F3** Knihovna musí umět pracovat se strukturou dat na serveru <http://athena.pef.czu.cz/ptacionline/>.
- F4** Knihovna musí umět určit počet vajec v hnízdě pro každý jednotlivý snímek.
- F5** Knihovna musí umět určit počet vajec v hnízdě pro složku jako celek. Složka se skládá ze sekvence obrázků, která reprezentuje jednu videosekvenci.
- F6** Snímky mohou být ve formátu JPEG³ a PNG⁴.

²Zdrojový kód může používat funkce poskytované jinými softwarovými knihovnami.

³Joint Photographic Experts Group.

⁴Portable Network Graphics.

Rešerše

2.1 Počítačové vidění

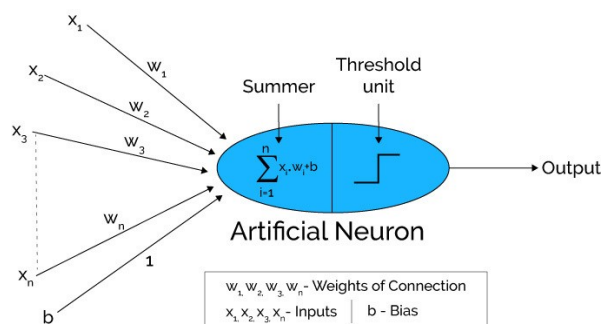
Počítačové vidění je odvětví počítačové vědy, které se zabývá získáváním informací z obrazu. Může se jednat o extrahování informací z videí, obrázků nebo například lékařských snímků [5]. Naučit počítač „vidět“ není vůbec jednoduchý úkol, i přes to, že pro náš mozek je získávání informací z obrazu přirozené a jednoduché. Obraz, jak ho vidí počítač, je pouze posloupnost systematicky uspořádaných jedniček a nul. Motivací pro hledání tvarů nebo přiřazování významu této posloupnosti dat je automatizace úkolů, které doposud vykonávali lidé. Počítačové vidět znamená číst matici pixelů⁵, které mají různou intenzitu. Veškeré operace počítačového vidění jsou založeny na analýze sousedství pixelů, jejich jasu, či barevného spektra. Mezi typické využití počítačového vidění patří například [6]:

- autonomní vozidla,
- analýza obrazu v bezpečnostní systémech,
- průmyslové roboty,
- rozpoznávání obrazu.

V této bakalářské práci se budeme zabývat použitím nástrojů k rozpoznávání obrazu. Tyto nástroje nám umožní automatizovat úkoly, které by jinak museli vykonávat lidé. Rozpoznávání obrazu můžeme rozčlenit do několika kategorií podle typu problému [6], například:

rozpoznávání obrazu: neboli klasifikace obrazu – zařazení obrazu do jedné ze známých kategorií (například na snímku je pes)

⁵Body, které tvoří obraz.



Obrázek 2.1: Struktura a funkce neuronu v umělé neuronové síti.

Zdroj: https://cdn-images-1.medium.com/max/800/0*0H1zxs0DSEBXW0iI.jpg [8].

detekce objektů: autonomní vozy se snaží detekovat a následně klasifikovat všechny objekty v jeho okolí

identifikace: rozpoznávání obličeje nebo otisku prstu

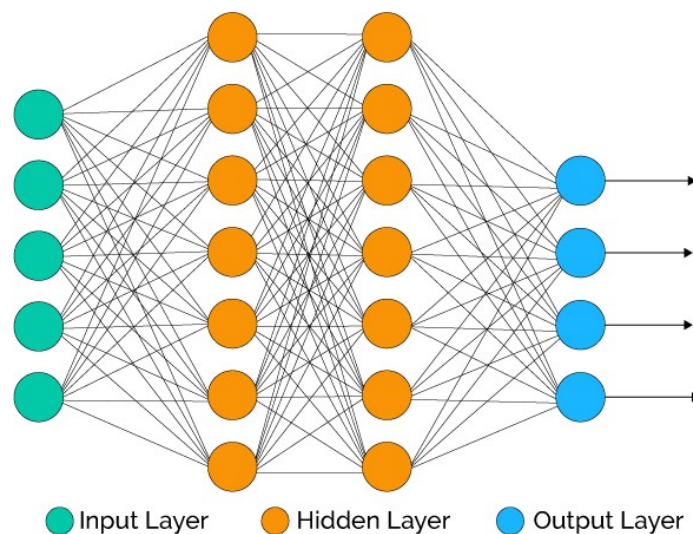
V současné době jsou algoritmy umělé inteligence [15] nejefektivnějším řešením rozpoznávání obrazu. Algoritmy založené na konvolučních neuronových sítích se dokonce přiblížily svou přesností a efektivitou lidem [7].

2.2 Neuronové sítě

Umělá neuronová síť je softwarový systém inspirovaný biologickými neuronovými sítěmi, které se nachází v lidských a zvířecích mozcích. Neuronové sítě jsou běžně používány například k rozpoznávání obrazu, filtrování spamu nebo předvídání vývoje hodnoty akcií.

Umělé neuronové sítě obsahují neurony, které jsou mezi sebou propojeny. Vstupem každého neuronu jsou výstupy neuronů v předchozích vrstvách. Tento vstup je zpracován podle interní logiky neuronu a následně je vygenerován výstup, který se stává vstupem pro neurony v následujících vrstvách. Typická interní logika neuronu pracuje s váženými vstupy a vnitřní, tzv. excitační, funkcí. Tato funkce často funguje tak, že pokud součet vážených vstupů překročí určitou hranici při průchodu touto funkcí, neuron se „excituje“ a vydá odpověď. Učení, neboli trénování neuronové sítě, představuje proces hledání optimálních vah a spojení jednotlivých neuronů. V rámci trénování můžeme měnit i topologii sítě. Neuronová síť je tedy **orientovaným váženým grafem**. Vstupem pro „první řadu“ neuronů jsou vstupní data, například jednotlivé body⁶ obrazu. Aktivace „poslední řady“ neuronů reprezentuje výsledek pro daný vstup [9]. Struktura neuronové sítě a neuronu je znázorněna na obrázku 2.2, respektive 2.1.

⁶pixels



Obrázek 2.2: Struktura neuronové sítě.

Zdroj: https://cdn-images-1.medium.com/max/1600/0*IUWJ5oJ_z6A1G7Ja.jpg [8].

Dnešní architektury neuronových sítí však mohou být mnohem komplexnější. Mohou se lišit počtem skrytých vrstev, strukturou propojení neuronů (například cyklická propojení neuronů), skryté vrstvy mohou mít paměť, atd. Mezi nejznámější typy neuronových sítí patří konvoluční neuronové sítě, modulární neuronové sítě, Boltzmannova síť, síť s dlouhodobou a krátkodobou pamětí a perceptronové sítě. Stručně představení jednotlivých typů viz [8].

V rámci této práce se zaměříme především na použití hlubokých konvolučních neuronových sítí ke klasifikaci obrazu a detekci objektů.

2.3 TensorFlow

„TensorFlow je otevřená softwarová knihovna⁷ pro numerické výpočty pomocí metody „data flow graph“ [10]. Knihovnu je možné použít například pro usnadnění matematických výpočtů, tvorbu lineárních modelů a neuronových sítí. Použitím knihovny získáme přístup k sadě připravených algoritmů, které nám značně usnadní tvorbu neuronové sítě. K dispozici jsou například algoritmy pro trénování (hledání optimálních vah k jednotlivým hranám grafu), jako je například algoritmus „stochastic gradient descent“. TensorFlow dělá použití dostupných algoritmů opravdu jednoduchým, viz algoritmus 2.1.

TensorFlow také značně usnadňuje tvorbu a konfiguraci struktury neuronové sítě, viz [11]. V neposlední řadě můžeme použít skripty pro práci s daty [12] nebo TensorFlow Object Detection API [13].

⁷Open source software.

2. REŠERŠE

Algoritmus 2.1: Použití algoritmu „stochastic gradient descent“ poskytovaný knihovnou TensorFlow.

```
1 optimizer = tf.train.GradientDescentOptimizer(learning_rate=0.001)
2 train_op = optimizer.minimize(loss=loss,
3                               global_step=tf.train.get_global_step())
```

Analýza a návrh

V této kapitole se zaměříme na strukturu projektu Ptáci Online, problémy spojené se současným způsobem sbírání dat, rozbor funkčních a nefunkčních požadavků, návrh řešení a výběr technologií. Na závěr budeme diskutovat dva různé způsoby řešení – jejich výhody a nevýhody.

3.1 Projekt Ptáci Online

„Cílem projektu je popularizovat ochranu ptáků v blízkosti lidských sídel, jejich hnízdění, včetně jeho monitoringu, s využitím speciálního technického zařízení tzv. „chytré ptačí budky“.“ [14]

Tyto „chytré ptačí budky“ slouží k monitorování hnízdícího ptactva. Každá budka obsahuje jednu nebo dvě kamery s nočním přísvitem. Ve vletovém otvoru budky je umístěna světelná brána, která spustí natáčení kamer při detekci pohybu. Dále je do budky vestavěn venkovní a vnitřní teplotní senzor, mikrofon a senzor venkovního osvětlení, který reguluje funkci přísvitu kamer. Přenos nasbíraných dat z budky probíhá přes ethernetový PoE⁸ kabel. Tento kabel zajišťuje i napájení veškeré elektroniky uvnitř budky. [1] Jak vypadá záznam z budky je vidět na obr. 3.1.

3.2 Představení vstupních dat

Ptačí budka vyprodukuje sekvenci videa pokaždé, když je v ní detekován pohyb. Kvůli energetické úspornosti kamer a množství přenášených dat mají tato videa nižší snímkovou frekvenci⁹. Z takového videozáznamu lze extrahovat množinu jednotlivých snímků, které jako celek tvoří dané video. Snímky jsou ukládány do formátu PNG a zařazeny do složek, přičemž **jedna složka reprezentuje jeden videozáznam**.

⁸Power over Ethernet.

⁹Počet snímků za sekundu.



Obrázek 3.1: Ukázka neupraveného snímku z ptačí budky.

Výsledná softwarová knihovna bude navržena pro práci s jednotlivými složkami. Načte všechny snímky v dané složce a následně je zpracuje. Uživatel bude schopen získat informace o složce jako celku i jednotlivých snímcích.

3.3 Současný způsob zpracování dat

Akademičtí pracovníci potřebují shromáždit velké množství dat, ale taková činnost je velmi časově náročná. Proto jsou na takovou práci najímáni brigádníci. Brigádníky je nutno nejprve zaškolit, aby věděli, co mají ve videích hledat. Poté sledují jedno video za druhým a získané informace zapisují do tabulek v Excelu. Tento způsob práce je drahý a časově náročný.

Jedna z informací, která nás může zajímat, je počet vajec v hníždě. Brigádník musí video otevřít, shlédnout a najít část, kde jsou vejce zřetelně viditelná (viz obr. 3.2a). Poté vejce spočítá a výsledek zapíše do tabulky. Jelikož je nutné získat co nejvíce dat, brigádníci tento proces vykonávají co nejrychleji. To vede k chybám, např. chybně spočítaný počet vajec nebo zapsání výsledku na špatný řádek tabulky.

Je snahou vytvořit systémy, které by tyto úkoly mohly provádět automaticky. Příkladem takového systému je například práce od Pavla Šumy, který se snaží automaticky zjistit počet mláďat v hníždě [16]. Dalším příkladem je práce Ing. Josefa Pavlíčka, Ph.D a spol. [17] nebo i tato práce.

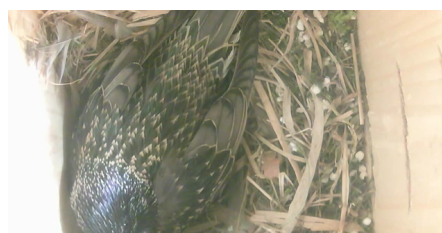
3.4 Nefunkční požadavky

3.4.1 N1

Knihovna musí pracovat v reálném čase. Po zadání vstupních dat je výsledek očekáván maximálně v jednotkách sekund.



(a) Vejce jsou zřetelně viditelná v čase 0:01.



(b) Vejce nejsou viditelná ve zbytku videa, jako je tomu např. v čase 0:14.

Obrázek 3.2: Vejce nemusí být vždy viditelná.

Z vlastností neuronových sítí vyplývá, že vyhodnocení vstupních dat bude téměř okamžité. Samotný průchod snímkem grafem, resp. průchod snímkem neuronovou sítí, není příliš časově náročný. Nutnou podmínkou však je dostupný soubor obsahující popis struktury neuronové sítě. Trénování, nebo-li hledání optimálního nastavení vah tak, aby výstupní chyba sítě byla co nejmenší, je velmi výpočetně náročná činnost, která musí být dokončena **před** použitím knihovny.

3.5 Funkční požadavky

3.5.1 F1

Systém bude distribuovaný formou Java knihovny v archivu JAR.

Neuronové sítě můžeme naprogramovat v libovolném programovacím jazyce. Tato podmínka může být tedy bez problému splněna. Výsledná softwarová knihovna pro detekci počtu vajec v hnízdě bude implementována v jazyce Java.

3.5.2 F2

Knihovna bude distribuována ve dvou verzích:

- *JAR archiv obsahující, jak přeložený zdrojový kód knihovny, tak i všechny závislosti, které jsou knihovnou vyžadovány.*
- *JAR archiv obsahující pouze přeložený zdrojový kód knihovny bez externích závislostí. Do archivu bude přibalen konfigurační soubor pro systém Maven[4] s definicí závislostí.*

Distribuci přeloženého zdrojového kódu vyřešíme pomocí nástroje pro správu, řízení a automatizaci sestavování aplikací. Maven je pro nás nejvhodnější vol-

3. ANALÝZA A NÁVRH

bou, vzhledem k druhé části podmínky – do archivu bude přibalen konfigurační soubor pro systém Maven s definicí závislostí.

3.5.3 F3

*Knihovna musí umět pracovat se strukturou dat na serveru
<http://athena.pef.czu.cz/ptacionline/>.*

Knihovna bude schopna pracovat se strukturou dat, která je popsána v kapitole 3.2. Uživatelská interakce s knihovnou bude reflektovat strukturu dat projektu:

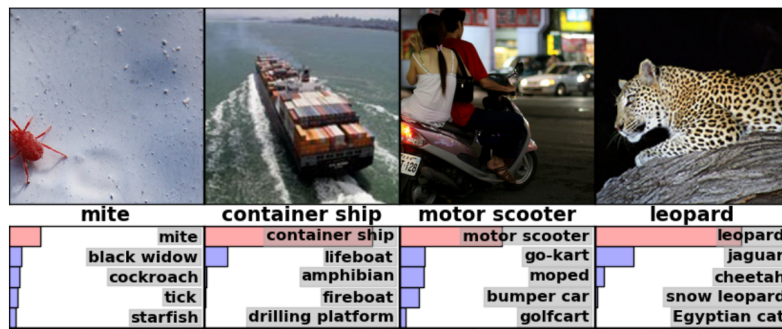
Algoritmus 3.1: Prvotní návrh uživatelského rozhraní knihovny.

```
1 // Inicializujeme EggDetector. Knihovna se připraví pro
2 // zpracovávání sekvencí.
3 EggDetector eggDetector = new EggDetector();
4
5 // EggDetectoru předáme absolutní cestu k složce, kterou
6 // chceme vyhodnotit. EggDetector nám vrátí třídu
7 // SequenceClassifier, která obsahuje veškeré informace
8 // k dané složce.
9 SequenceClassifier sequenceClassifier = eggDetector.evaluate(new
10     File("image_dir"));
11
12 // Zjistíme finální počet vajec v hnízdě pro danou složku.
13 System.out.println("final count: " +
14     sequenceClassifier.getFinalCount());
15
16 // Můžeme zjistit počet vajec v jednotlivých snímcích.
17 System.out.println("individual scores: " +
18     sequenceClassifier.getIndividualCounts());
19
20 // Ukončíme EggDetector - uvolníme informace o neuronové
21 // síti z paměti. Instance EggDetectoru se stane nepoužitelná.
22 eggDetector.closeSession();
```

3.5.4 F4

Knihovna musí umět určit počet vajec v hnízdě pro každý jednotlivý snímek.

Výsledná knihovna bude uživateli umět poskytnout informace o jednotlivých snímcích – viz algoritmus 3.1.



Obrázek 3.3: Program určený ke klasifikaci obrazu.

Zdroj: dokumentace softwarové knihovny TensorFlow [18].

3.5.5 F5

Knihovna musí umět určit počet vajec v hnízdě pro složku jako celek. Složka se skládá ze sekvence obrázků, která reprezentuje jednu videosekvenci.

Výsledná knihovna bude uživateli umět poskytnout informace o složce jako celku – viz algoritmus 3.1.

3.5.6 F6

Snímky mohou být ve formátu JPEG a PNG.

Pro zpracování snímků použijeme standartně dostupné třídy pro práci s grafikou v programovacím jazyce Java. Třídy `BufferedImage` a `Graphics2D` nám umožní zpracovat oba dva formáty – JPEG i PNG.

3.6 Návrh řešení

Jádro implementace bude tvořit neuronová síť, která bude trénována metodou „učení s učitelem“¹⁰. Tvorba implementace se bude skládat ze čtyř částí:

- Příprava trénovacích a testovacích dat.
- Trénování neuronové sítě.
- „Konzumace“ vytrénované neuronové sítě knihovnou, která je výsledkem této práce.
- Ověření funkčnosti.

¹⁰Supervised learning.

3. ANALÝZA A NÁVRH



Obrázek 3.4: Program určený k detekci objektů.

Zdroj: dokumentace softwarové knihovny TensorFlow [13].

Vytýčené cíle pro neuronové sítě mohou být různé. Prvním možným řešením je neuronová síť určená ke klasifikaci celého obrazu. Typickým vstupem je snímek, ve kterém se nachází objekt, který chceme rozpoznat. Pokročilejší typ této sítě je schopný popsat komplexnější scény, jako například „Dva lidé na pláži.“. Ukázka takového programu je vidět na obrázku 3.3. V našem případě bychom klasifikovali počet vajec v hnízdě, kde by jednotlivé výsledky reprezentovaly počet vajec na daném snímku. Tzn. výstupem takové neuronové sítě by byla jedna z předem daných kategorií, o které si s největší pravděpodobností myslíme, že popisuje, co je na daném snímku. Každá kategorie by reprezentovala jiný počet vajec v hnízdě, například kategorie 0, kategorie 1, atd.

Druhým možným řešením je komplexnější neuronová síť určená k detekci objektů. Obraz je nejprve segmentován na části, o kterých si síť myslí, že by mohly obsahovat nějaké objekty. Následně jsou tyto části klasifikovány a je rozhodnuto, zda-li a s jakou pravděpodobností se jedná o hledaný objekt. Výstupem je množina detekcí, která je tvořena umístěním objektu, typem objektu a pravděpodobnostní mírou, která reprezentuje, jak moc je si síť jistá, že se opravdu jedná o daný objekt. Typickým vstupem je libovolný snímek (viz obrázek 3.4). V našem případě bychom hledali detekce jednotlivých vajec s relativně vysokou pravděpodobností. Počet detekcí by reprezentoval počet vajec v hnízdě.

Ať už zvolíme jakýkoliv způsob implementace, pro zpracování obrazu předáme neuronové síti vektor o velikosti $300 \times 300 \times 3$ (snímek bude zmenšen na velikost 300×300 bodů a zůstane barevný – každý bod obrázku obsahuje

3 barevné složky¹¹), který reprezentuje náš snímek. Formát a tvorba tohoto vektoru viz algoritmus 3.2.

Algoritmus 3.2: Snímek reprezentovaný jako vektor je vstupem pro neuronovou síť.

```

1 int[] intValues = new int[INPUT_SIZE * INPUT_SIZE];
2 byte[] byteValues = new byte[INPUT_SIZE * INPUT_SIZE * 3];
3 intValues = ((DataBufferInt)
4     img.getRaster().getDataBuffer()).getData();
5
6 for (int i = 0; i < intValues.length; ++i) {
7     byteValues[i * 3 + 2] = (byte) (intValues[i] & 0xFF);
8     byteValues[i * 3 + 1] = (byte) ((intValues[i] >> 8) & 0xFF);
9     byteValues[i * 3 + 0] = (byte) ((intValues[i] >> 16) & 0xFF);
10 }
11 // byteValues je vstupem pro neuronovou síť

```

Knihovna nám poté vrátí požadovaný výsledek¹².

3.7 Volba technologií

V této kapitole vybereme konkrétní platformy, programovací jazyky a nástroje potřebné k implementaci praktické části této práce.

3.7.1 Platforma

Vzhledem k funkčním požadavkům je jednoznačnou volbou programovací jazyk Java. Volíme verzi Java 8 SE, která poskytuje všechny nástroje, které k doručení výsledku potřebujeme.

3.7.2 Neuronové sítě

Pro zjednodušení tvorby neuronových sítí použijeme jednu z dostupných softwarových knihoven. Mezi tři nejznámější patří TensorFlow, OpenNN a FANN¹³. Z těchto tří knihoven pouze jedna poskytuje Java API. Už jen z tohoto důvodu je TensorFlow nejlepším kandidátem.

TensorFlow má skvělou dokumentaci [19] s velkým množstvím detailních návodů [20]. Poskytuje API k detekci objektů, kde je možné využít již předem vytrénované modely [13]. Pomocné skripty určené k trénování neuronové sítě,

¹¹RGB - red, green, blue. Každý bod obrázku obsahuje informaci o koncentraci červené, zelené a modré.

¹²V případě rozpoznávání obrazu, knihovna vrátí seznam pravděpodobností pro všechny známe typy. V případě detekce objektů, knihovna vrátí seznam, pozici a typ objektů.

¹³Fast Artificial Neural Network.

uložení její struktury a ověření funkčnosti naimplementujeme v jazyce Python 3, jelikož je primárním jazykem pro interakci s knihovnou TensorFlow.

3.7.3 Práce s daty

K vytrénování neuronové sítě je potřeba velké množství trénovacích dat. Trénovacími daty jsou myšleny snímky, ke kterým manuálně dodáme všechny potřebné informace – například počet a umístění jednotlivých vajec. Abychom proces získávání a označování dat usnadnili, využijeme několik nástrojů, které jsou detailně popsány v kapitole 4.

- Pro hromadné stažení dat použijeme nástroje `bash` a `wget`. Hromadně stažená data budou obsahovat i obsah, který pro nás není užitečný. Proto použijeme nástroj, který všechna neužitečná data smaže. Více v kapitole 4.1.
- Trénovací a testovací data vytvoříme pomocí nástrojů `LabelImg` a `Tagger`. Detailní informace obsahuje kapitola 4.2.

Abychom byli schopni nově vytvořená trénovací data použít k trénování neuronové sítě, musíme je převést do standardního formátu. V našem případě musíme z binárních dat snímků a textových XML souborů vytvořit tzv. TFRecord [21]. Každý způsob implementace vyžaduje mírně odlišný formát trénovacích dat. V kapitolách 5 a 6 diskutujeme mimo jiné i přípravu trénovacích dat.

3.8 Shrnutí kapitoly

V této kapitole byl analyzován projekt Ptáci Online – byly představeny jeho cíle a popsán současný stav, jakým jsou sbírána a zpracovávána data. Byly adresovány všechny funkční i nefunkční požadavky s ohledem na strukturu projektu a dat. Byla navržena dvě možná konkurenční řešení, u kterých není předem jasné, které z nich je efektivnější a přinese lepší výsledky. Nezbyvá tedy nic jiného, než implementovat obě dvě řešení a jejich výsledky porovnat. Na závěr jsme se zaměřili na výběr vhodných principů a nástrojů potřebných k úspěšné implementaci. Byla vybrána platforma knihovny, knihovna pro práci s neuronovými sítěmi a nástroje pro přípravu dat.

V dalších kapitolách se zaměříme na tvorbu a použití všech nástrojů použitých v rámci této práce. Ale především budeme diskutovat oba dva způsoby implementace, které nakonec porovnáme.

Nástroje

Proces přípravy a zpracování dat můžeme usnadnit použitím několika nástrojů. Zaměříme se na hromadné stahování dat ze serveru `http://athena.pef.czu.cz/ptacionline/` a jejich „pročištění“. Následně představíme nástroje, ve kterých obohatíme stažené snímky o informace potřebné k trénování neuronové sítě.

4.1 Hromadné stahování dat

4.1.1 Získání dat

Abychom nemuseli stahovat snímky jeden po druhém manuálně, pomůžeme si napsáním jednoduchého skriptu, který stáhne všechny snímky za nás. K tomu nám postačí dva nástroje: `bash` a `wget`. Tento skript stáhne veškerá data, která jsou na serveru `http://athena.pef.czu.cz/ptacionline/` dostupná. Detailní dokumentace skriptu je k nalezení v příloze H).

Algoritmus 4.1: Hromadné stahování dat ze serveru `athena.pef.czu.cz`.

```
#!/bin/bash

DIRECTORY=data
URL=http://athena.pef.czu.cz/ptacionline/

wget -o log.txt -nv --show-progress -c -P "$DIRECTORY" -r -np -nH \
    --cut-dirs=1 -R index.html "$URL"
```

Stručné vysvětlení příkazu `wget`, který používáme na poslední řádce skriptu 4.1:

- Všechny složky a podsložky dostupné na serveru budou staženy do lokální složky `$DIRECTORY`.

- `-o log.txt` – Vytvoří záznam do souboru `log.txt`.
- `-nv` – Zobrazuje pouze chyby, ne varování.
- `-show-progress` – Ukáže progres stahování.
- `-c` – Pokračuj ve stahování nedokončených souborů.
- `-r` – Rekurzivně stahuj podsložky.
- `-np` – Nestahuj soubory v složkách výše, než `ptacionline`.
- `-nH` – Nestahuj do složky, která se jmenuje stejně jako doména, ale přímo do `$DIRECTORY`.
- `-cut-dirs=1` – Ve složce `$DIRECTORY` vynech první složku (`ptacionline`).
- `-R index.html` – Nestahuj soubory `.html`.

4.1.2 Čištění dat

Skript, který jsme představili v kapitole 4.1.1, stáhne veškerá data z daného serveru. Z těchto dat použijeme pouze snímky z ptačích budek, ostatní data pro nás nejsou relevantní. Abychom se v datech mohli lépe orientovat a ušetřili místo na pevném disku, bylo by vhodné nepotřebná data smazat. Pro tento účel naprogramujeme jednoduchý nástroj v programovacím jazyce Java, který všechna nepotřebná data smaže.

Zdrojový kód a dokumentace nástroje **FolderTrimmer** je k nalezení v příloze D. Výsledný program stačí spustit a složku, která má být promazána, mu předat jako argument. Například: `run.sh /home/demo/eggs/data`.

FolderTrimmer funguje ve dvou režimech. První, základní režim, smaže všechny soubory jiného typu než PNG, TXT a XML. V případě, že složka neobsahuje další složku nebo alespoň jeden soubor typu PNG, TXT nebo XML, bude smazána. Druhý režim funguje stejně jako první, ale smaže všechny složky, které neobsahují soubor `imgdata.xml`. To umožní vymazání všech dat, která nejsou relevantní pro trénování neuronové sítě. První režim spustíme příkazem `run.sh false "cesta_ke_slozce"`. Druhý režim spustíme příkazem `run.sh true "cesta_ke_slozce"`.

4.2 Příprava trénovacích a testovacích dat

Když máme k dispozici všechna potřebná data, je potřeba je připravit do takové podoby, abychom je mohli použít k trénování neuronové sítě. Příprava dat se liší podle typu implementace, který zvolíme. V případě trénování neuronové sítě k detekci objektů chceme k jednotlivým snímkům přidat informaci **kde**, **jaké velikosti** a **kolik** vajec se v nich nachází. V případě trénování

neuronové sítě pro rozpoznávání (klasifikaci) obrazu je potřeba ke snímkům přidat informaci o **celkovém počtu vajec**.

4.2.1 Tagger

Nástroj **Tagger** slouží k manuálnímu označování snímků. Pracuje na úrovni složek, kde jednotlivé složky a v nich umístěné snímky reprezentují jednu videosekvenci. Uživatel pak může program spustit a poměrně rychle označit, kolik vajec se na daných snímcích nachází. Tagger je webovou aplikací s jednoduchým uživatelským rozhraním. Uživatel je prezentován všemi snímky dané složky a má možnost u každého snímku specifikovat počet vajec. Složky jsou vybírány automaticky. Jakmile uživatel uvedl počet vajec u každého snímku, data odešle a je prezentován snímky z další složky. Tento proces se opakuje do té doby, dokud nebudou dostupné žádné složky s neoznačenými snímky. Ukázka uživatelského rozhraní je vidět na obrázku 4.2.

Výstupem programu jsou soubory `imgdata.xml`, které obsahují informace o počtu vajec na jednotlivých snímcích. Tato data jsou pak používána pro trénování neuronové sítě k rozpoznávání obrazu, kde počet vajec reprezentuje možné výstupy neuronové sítě. Takto označené snímky se dají použít i pro validaci jakéhokoliv řešení – finální softwarové knihovně předáme složku se snímky, knihovna vyhodnotí výsledky a my je poté můžeme porovnat s výsledky, které jsme manuálně nasbírali. Ukázkový soubor `imgdata.xml` je přiložen v příloze C.1.

Program je napsán v programovacím jazyce Java. Detailní popis nástroje se nachází v příloze C. Jedná se o webovou aplikaci, která je postavena na technologii Spring Boot [22]. Výpis 4.2 ukazuje algoritmus pro selekci vhodných složek.

Algoritmus 4.2: Výběr vhodných složek pro trénování nástrojem Tagger.

```
1 public static ArrayList<String> scanFolder(String location) {
2     File locFile = new File(location);
3
4     if (!locFile.exists()) {
5         return new ArrayList<>();
6     }
7
8     // example folder name: 20160430_073822_526_D
9     final Pattern pattern = Pattern.compile("\\d{8}_\\d{6}_\\d{3}_D");
10    List arr = Arrays.asList(locFile.listFiles((File file, String name) -> {
11        File workingDir = new File(file.getAbsolutePath() +
12            File.separator + name);
13        if (!workingDir.isDirectory() &&
14            !pattern.matcher(name).matches()) {
15            return false;
16        }
17    }
18    });
19    return arr;
```

```

16     File imgDataFile = new File(workingDir, "imgdata.xml");
17     if (imgDataFile.exists()) {
18         return false;
19     }
20
21     if (workingDir.list((f, n) -> {
22         File workingFile = new File(f.getAbsolutePath() +
23             File.separator + n);
24         return workingFile.isFile() &&
25             FilenameUtils.getExtension(n).equals("png");
26     }).length <= 0) {
27         return false;
28     }
29
30     return true;
31 }));
32
33 return new ArrayList<>(arr);
34 }

```

4.2.2 LabelImg [23]

Trénovací data pro detekci objektů vyžadují jiný formát než data pro rozpoznávání obrazu. Na jednotlivých snímcích je potřeba označit pozici všech vajec. Přesně za tímto účelem byl vytvořen nástroj LabelImg [23]. Uživatel si při spuštění programu zvolí složku, ve které se nachází snímky připravené k označení. Uživatel má možnost zadat typy objektů, které chce na snímcích označovat. V našem případě se jedná pouze o jeden typ objektu – vejce¹⁴. Poté je uživateli prezentován snímek, na kterém má možnost objekty manuálně ohraničit (viz obrázek 4.1).

Níže je přiložen ukázkový výstup programu LabelImg.

Algoritmus 4.3: Ukázkový výstup programu LabelImg.

```

1 <annotation>
2   <folder>eggs</folder>
3   <filename>snap0001-0000000648.png</filename>
4   <path>C:\Users\test\Desktop\eggs\snap0001-0000000648.png</path>
5   <source>
6     <database>Unknown</database>
7   </source>
8   <size>
9     <width>1280</width>
10    <height>720</height>
11    <depth>3</depth>
12  </size>

```

¹⁴Ve skutečnosti se náš objekt (label) jmenuje `egg`.

4.2. Příprava trénovacích a testovacích dat

```
13 <segmented>0</segmented>
14 <object>
15   <name>egg</name>
16   <pose>Unspecified</pose>
17   <truncated>0</truncated>
18   <difficult>0</difficult>
19   <bndbox>
20     <xmin>462</xmin>
21     <ymin>379</ymin>
22     <xmax>544</xmax>
23     <ymin>447</ymin>
24   </bndbox>
25 </object>
26 </annotation>
```








Obrázek 4.1: Uživatelské rozhraní nástroje LabelImg.

4. NÁSTROJE

[Different folder](#)

Prefill the form:

snap0001-000000668.png		<input type="text" value="0"/>
snap0002-000001419.png		<input type="text" value="2"/>
snap0003-000002341.png		<input type="text" value="2"/>
snap0004-000003160.png		<input type="text" value="8"/>
snap0005-000004082.png		<input type="text" value="8"/>

Obrázek 4.2: Tagger – hromadné určování počtu vajec.

Implementace detekování objektů

V této kapitole se zaměříme na implementaci neuronové sítě, která bude sloužit k detekci objektů. V rámci této práce nás bude zajímat pouze jeden objekt: vejce.

Abychom mohli vytrénovat nový model, který je schopný detekce vajec, je potřeba, abychom trénovací data převedli do speciálního formátu. K validaci modelu budeme používat skript, u kterého bude vstupem obrázek libovolné velikosti. Interně bude snímek reprezentován vektorem o velikosti 300 x 300 x 3 (snímek zmenšen na velikost 300 x 300 bodů a zůstane barevný – každý bod obsahuje 3 barevné složky). Formát a tvorba tohoto vektoru viz algoritmus 3.2. Jakmile provedeme detekci objektů, musíme detekované objekty klasifikovat a rozřadit do předem stanovených kategorií. Kategorie, do které budeme chtít snímky rozřadit, bude pouze **jedna**:

Kategorie „egg“: Objekt vejce.

Celé řešení – příprava dat, trénování neuronové sítě, testování funkčnosti a měření přesnosti budeme implementovat v programovacím jazyce Python 3.

5.1 Příprava vývojového prostředí

Abychom mohli používat TensorFlow Object Detection API, je potřeba následně připravit vývojové prostředí:

1. Nainstalovat Python 3 a pip3.
2. Získat knihovnu TensorFlow společně s dostupnými modely ze sdíleného úložiště:

```
git clone https://github.com/tensorflow/models
```

5. IMPLEMENTACE DETEKOVÁNÍ OBJEKTŮ

3. Nainstalovat modely:

```
cd models/research && python3 setup.py
```

4. Přidat tyto řádky do ~/.bashrc (path_to_models_directory nahradíme skutečným umístěním modelů):

```
export MODELS=path_to_models_directory
export PYTHONPATH=$MODELS:$MODELS/slim
export OBJ_DET=$MODELS/object_detection
```

5. Nainstalovat TensorFlow: sudo pip3 install tensorflow-gpu nebo sudo pip3 install tensorflow. Více informací viz [24].

6. Pro ověření instalace spustíme interaktivní Python 3:

```
1 import tensorflow as tf
2 hello = tf.constant('EggDetector!')
3 sess = tf.Session()
4 print(sess.run(hello))
```

Systém by měl odpovědět „EggDetector!“.

5.2 Příprava dat

Nástrojem LabelImg (viz kapitola 4.2.2) jsem označil **1890** vajec. Data je potřeba rozdělit na trénovací a testovací. Testovací data slouží k validaci a vyhodnocení úspěšnosti trénovaného modelu. Standardně se data dělí přibližně v poměru devět trénovacích snímků k jednomu testovacímu. Data jsem rozdělil následovně:

- **1701** výskytů vajec jako trénovací data,
- **189** výskytů vajec jako data validační.

Trénovací data přesuneme do pracovní složky. Po přesunutí je nutné opravit XML soubory vygenerované nástrojem LabelImg. V obou složkách, kde máme trénovací data a testovací data pustíme následující příkaz:

```
for fullfile in *.jpg; do
  filename=$(basename "$fullfile")
  filename="${filename%.*}"
  echo "$filename".xml
  awk -v var="$filename" \
    'NR==3{$0="\t<filename>"var".jpg</filename>"}1;' \
    "$filename".xml > temp.xml && mv temp.xml "$filename".xml
```

```

done

for fullfile in *.png; do
    filename=$(basename "$fullfile")
    filename="${filename%.*}"
    echo "$filename".xml
    awk -v var="$filename" \
        'NR==3{$0="\t<filename>"var".png</filename>"}1;' \
        "$filename".xml > temp.xml && mv temp.xml "$filename".xml
done

```

5.3 Struktura projektu

Projekt, ve kterém budeme celou implementaci tvořit, bude mít následující strukturu:

```

object-detection-training.....pracovní složka
├── training.....složka s konfigurací
│   ├── checkpoint_model ..... data modelu, ze kterého budeme vycházet
│   ├── egg_label_map.pbtxt
│   └── pipeline.config
├── test_images.....testovací obrázky k validaci
├── frozen_graph.....výsledný, exportovaný graf
├── images ..... všechna testovací a trénovací data
│   ├── test.....testovací data
│   └── train.....trénovací data
├── export_inference_graph.py
├── generate_tfrecord.py
└── xml_to_csv.py

```

Stejná struktura je i na přiloženém médiu (viz příloha H), konkrétně ve složce `src/impl/object-detection-training`. Všechny skripty, které v této kapitole používáme k trénování a validaci neuronové sítě jsou také dostupné v příloze H.

5.4 Trénování neuronové sítě

Moderní modely pro rozpoznávání obrazu mají miliony parametrů a je **extrémně** výpočetně náročné je vytrénovat. Učení „přenosem modelu“¹⁵ je technika, která ušetří náročnou práci využitím již předtrénovaného modelu a přetrénováním pouze finálních vrstev [25]. Trénování sítě „od nuly“ je standardně efektivnější, ale v rámci mého testování jsem ve výsledcích rozdíl ne-

¹⁵Transfer learning.

5. IMPLEMENTACE DETEKOVÁNÍ OBJEKTŮ

COCO-trained models {#coco-models}

Model name	Speed (ms)	COCO mAP[^1]	Outputs
ssd_mobilenet_v1_coco	30	21	Boxes
ssd_inception_v2_coco	42	24	Boxes
faster_rcnn_inception_v2_coco	58	28	Boxes
faster_rcnn_resnet50_coco	89	30	Boxes
faster_rcnn_resnet50_lowproposals_coco	64		Boxes
rfcn_resnet101_coco	92	30	Boxes
faster_rcnn_resnet101_coco	106	32	Boxes
faster_rcnn_resnet101_lowproposals_coco	82		Boxes
faster_rcnn_inception_resnet_v2_atrous_coco	620	37	Boxes
faster_rcnn_inception_resnet_v2_atrous_lowproposals_coco	241		Boxes
faster_rcnn_nas	1833	43	Boxes
faster_rcnn_nas_lowproposals_coco	540		Boxes

Obrázek 5.1: Předtrénované modely poskytované společností Google.

zaznamenal. Proto jsem se rozhodl o použití již předtrénovaného modelu a metody „transfer learning“.

5.4.1 Výchozí model

V rámci této implementace jsem zkusil použít tři různé modely, které sloužily jako startovní bod pro trénování. Nejlepší výsledky se dostavily při použití modelu `ssd_mobilenet_v1_coco`¹⁶ [26]. Modely, které nebyly tak úspěšné jsou: `faster_rcnn_nas_coco`¹⁷ a `faster_rcnn_resnet101_coco`¹⁸. Stažený model umístíme do složky `training` v pracovním adresáři. Další modely, jejich rychlost a efektivita jsou zobrazeny na obrázku 5.1.

5.4.2 Konfigurace

Nyní je čas nakonfigurovat parametry, podle kterých se bude neuronová síť řídit při učení. TensorFlow Object Detection API používá soubory `protobuf` ke konfiguraci trénovacího a evaluačního procesu. Konfigurační soubor je rozdělen do pěti částí: `model`, `train_config`, `eval_config`, `train_input_config` a `eval_input_config` [27].

Pro každý typ aplikace jsou vhodné jiné parametry. Výsledný konfigurační soubor použitý společně s modelem `ssd_mobilenet_v1_coco` je vypsán v příloze E.

¹⁶Ke stažení na http://download.tensorflow.org/models/object_detection/ssd_mobilenet_v1_coco_2017_11_17.tar.gz nebo v příloze H.

¹⁷Ke stažení na http://download.tensorflow.org/models/object_detection/faster_rcnn_resnet101_coco_2017_11_08.tar.gz.

¹⁸Ke stažení na http://download.tensorflow.org/models/object_detection/faster_rcnn_nas_coco_2017_11_08.tar.gz.

5.4.3 Trénovací data

TensorFlow Object Detection API očekává trénovací data ve formátu TFRecord [21]. Na příložené médium (viz příloha H) jsem umístil všechny skripty pro vygenerování těchto souborů z trénovacích dat.

Nejprve máme pouze jednotlivé snímky a k nim XML soubory, které obsahují dodatečné informace o daných snímcích. Cílem je dostat dva soubory: `train.record`, který obsahuje všechna potřebná data pro trénování a `test.record`, který obsahuje všechna data pro validaci a testování. Konverzní skript byl napsán podle předlohy [12]. Abychom dostali dva požadované soubory, budeme postupovat v těchto krocích:

1. Vygenerujeme dva CSV¹⁹ soubory, které budou obsahovat informace o jednotlivých detekcích v trénovacích a testovacích datech:

```
python3 xml_to_csv.py
```

2. Z kombinace CSV záznamu, vygenerovaném v předchozím kroku, a trénovacích snímků vytvoříme soubor `train.record`:

```
python3 generate_tfrecord.py \
  --csv_input=data/train_labels.csv \
  --output_path=data/train.record
```

3. Z kombinace CSV záznamu, vygenerovaném v prvním kroku, a validačních snímků vytvoříme soubor `test.record`:

```
python3 generate_tfrecord.py \
  --csv_input=data/test_labels.csv \
  --output_path=data/test.record
```

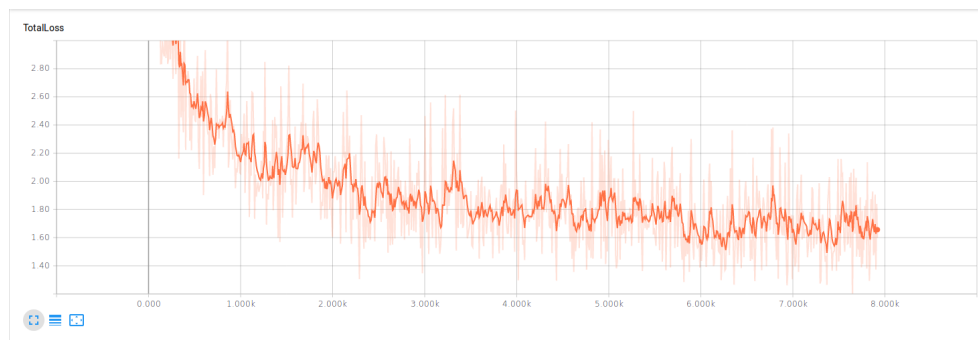
5.4.4 Trénování

V tuto chvíli máme vše připraveno a můžeme začít samotné trénování neuronové sítě. Je důležité si uvědomit, že trénování modelu je velmi výpočetně náročné a je možné, že nebudeme mít lokálně dostupný dostatečně výkonný hardware, který by tuto úlohu zvládl v rozumném čase. Máme tedy tři možnosti, kde a jak model vytrénovat. Čas potřebný k trénování je ovlivněn množstvím trénovacích dat, komplexitou modelu a hardwarovým výkonem.

První, nejjednodušší, ale nejméně výkonná varianta, je trénování na notebooku nebo stolním počítači pomocí procesoru. Tento způsob je poměrně neefektivní – například při použití metody „transfer learning“ se model trénoval

¹⁹Comma Separated Values.

5. IMPLEMENTACE DETEKOVÁNÍ OBJEKTŮ



Obrázek 5.2: Vizualizace průběhu trénování programem TensorBoard.

šest dní čistého času, než se stal použitelným. Trénování takového modelu bez použití metody „transfer learning“ by trvalo týdny.

Druhá varianta, která je komplexnější na zprovoznění, je trénování na lokálním stroji pomocí grafické karty. Tato metoda je podstatně více efektivní, než trénování na procesoru počítače. Na dvou desktopových grafických kartách *Nvidia GeForce GTX 650* byl model použitelný po 7 hodinách při použití metody „transfer learning“.

Nejpraktičtější, ale nejsložitější na zprovoznění, je trénování na vzdálených serverech. Google poskytuje na své platformě *Google Cloud Platform* produkt *ML Engine* [28], který je specializován mimo jiné i na trénování modelů – tzn. že je použit optimalizovaný hardware. Společnost Google stojí za frameworkem TensorFlow i platformou *ML Engine*, proto je jejich integrace relativně dobře vyřešená. Tento fakt dělá tzv. „cloud learning“ nejefektivnějším řešením.

5.4.4.1 Monitorování průběhu trénování

TensorFlow obsahuje nástroj, kterým můžeme monitorovat průběh trénování modelu. Program spustíme následovně:

```
tensorboard --log_dir=training --port=8080
```

Ve webovém prohlížeči otevřeme adresu `http://localhost:8080` a budeme prezentováni úvodní obrazovkou aplikace TensorBoard (viz obr. 5.3). Nástroj vizualizuje veškeré potřebné metriky – jednou z nejzajímavějších je metrika „total loss“ (viz obr. 5.2).

5.4.4.2 Trénování modelu na lokálním hardware

Trénování na jedné grafické kartě, více grafických kartách nebo na procesoru spouštíme vždy stejným způsobem:

```
python3 $OBJ_DET/train.py \  
--logtostderr \  

```

5.4. Trénování neuronové sítě



Obrázek 5.3: Úvodní obrazovka programu TensorBoard: přehled stavu trénování.

```
--train_dir=training/ \  
--pipeline_config_path=training/pipeline.config
```

Použitím parametrů můžeme například specifikovat, jak a kolik grafických karet má být využito. Pro použití dvou grafických karet jsem specifikoval parametry následovně:

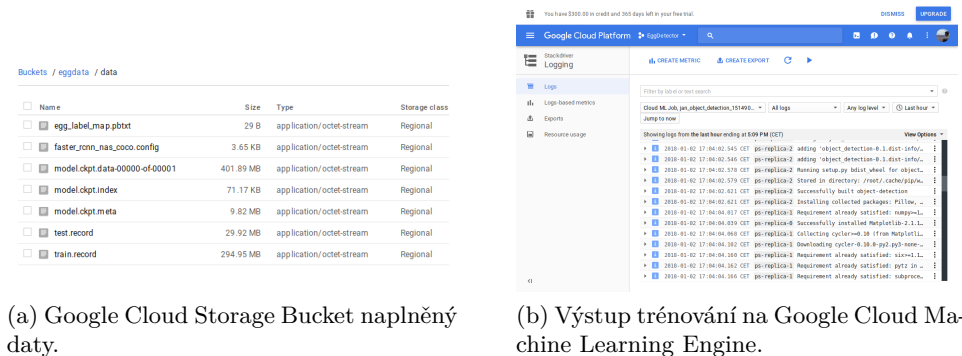
```
python3 $OBJ_DET/train.py \  
  --logtostderr \  
  --train_dir=training/ \  
  --pipeline_config_path=training/pipeline.config \  
  --num_clones=2 \  
  --ps_tasks=1
```

5.4.4.3 Trénování modelu na Google Cloud ML Engine [28]

Abychom mohli model trénovat na Google Cloud ML Engine, je potřeba několik kroků navíc. Pro detailní popis, jak zprovoznit trénování na Google platformě, viz [29].

Jakmile máme připravené naše úložiště (viz obr. 5.4a), můžeme spustit „vzdálený trénink“:

5. IMPLEMENTACE DETEKOVÁNÍ OBJEKTŮ



Obrázek 5.4: Použití Google Cloud Platform.

```
gcloud ml-engine jobs submit training \
  object_detection_${version_unique_ID} \
  --job-dir=gs://eggdata/train \
  --packages \
    dist/object_detection-0.1.tar.gz,slim/dist/slim-0.1.tar.gz \
  --module-name object_detection.train \
  --config object_detection/samples/cloud/cloud.yml \
  -- \
  --train_dir=gs://eggdata/train \
  --pipeline_config_path= \
    gs://eggdata/data/faster_rcnn_resnet101_coco.config
```

Průběh trénování je zaznamenáván do „logu“, který je vidět na obrázku 5.4b. Samozřejmě můžeme použít i nástroj TensorBoard stejně, jako kdybychom model trénovali lokálně:

```
tensorboard --logdir=gs://eggdata --port=8080
```

5.5 Ověření funkčnosti a výsledky

V případě, že se rozhodneme ukončit proces trénování, máme k dispozici několik souborů, které jsou pouhými záchytnými body²⁰ pro další iterace. Tyto soubory je vhodné uchovat, abychom měli možnost trénování obnovit tam, kde bylo naposledy ukončeno. Nicméně pro práci s modelem a detekci vajec je potřeba tyto soubory převést na jeden statický. Tomuto procesu se říká „zmražení grafu“. Graf exportujeme následovně:

```
python3 export_inference_graph.py \
  --input_type image_tensor \
```

²⁰Checkpoints.


```
--pipeline_config_path training/ssd_mobilenet_v1_coco.config \
--trained_checkpoint_prefix training/model.ckpt-VERSION \
--output_directory frozen_graph
```

Ve složce `frozen_graph` přibyl soubor typu PB²¹. Tento soubor je výsledným modelem, který můžeme použít k validaci řešení.

K validaci použijeme Python skript napsaný ve formě Jupyter Notebook [30].

```
jupyter notebook
```

Na přehledové stránce otevřeme soubor `test.ipynb` a zvolíme `Cell → Run All`. Poté jsou zobrazeny výsledky. Ukázka výstupu skriptu, ve kterém používáme nejúspěšnější model vytvořený metodou „transfer learning“ vycházejícího z modelu `ssd_mobilenet_v1_coco`, je vidět na obrázcích 5.5, 5.6 a 5.7. Méně úspěšný model je zobrazen na obrázku 5.8.



Obrázek 5.5: Model je schopný detekovat všech devět vajec i bez ohledu na to, že některá vejce jsou sotva viditelná.

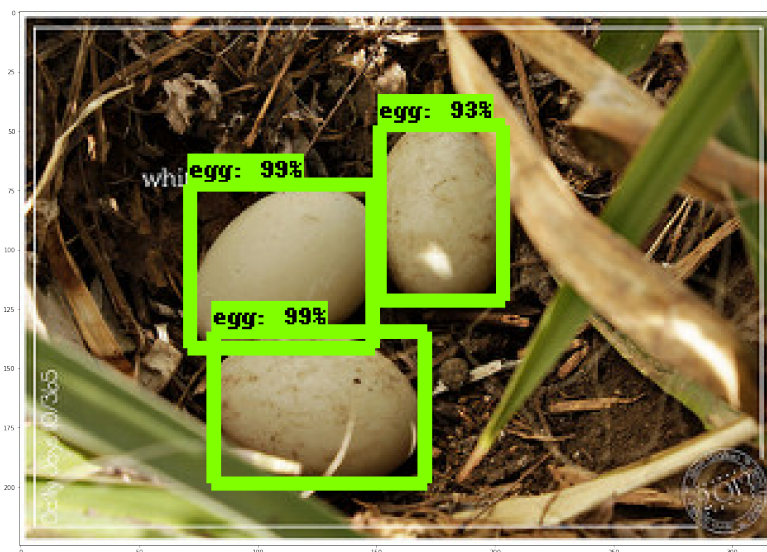
5.6 Závěr

Dokázali jsme, že řešení popsané v této kapitole funguje. Z výsledků vyplývá, že správná konfigurace (topologie) a výchozí model dokáží značně ovlivnit výsledné chování modelu i přesto, že ho vždy trénujeme stejnými daty.

Výsledky doručené modelem, který byl vytrénován metodou „transfer learning“ z modelu `ssd_mobilenet_v1_coco` a používá konfiguraci

²¹Protocol Buffer.

5. IMPLEMENTACE DETEKOVÁNÍ OBJEKTŮ



Obrázek 5.6: Model je schopný detekovat i vejce v jiném prostředí.



Obrázek 5.7: Model detekuje vejce i ve složitých podmínkách.

`training/pipeline.config`, jsou ze všech nejlepší. Proto bude tento model považován za výstup této kapitoly a na závěr bude porovnán s implementací klasifikace obrazu (viz kapitola 6).

I přes uspokojivé výsledky řešení je zde prostor pro vylepšení. Více trénovacích a testovacích dat by zajistilo ještě lepší výsledky (vzorek použitý v této kapitole obsahuje **1890 detekcí**). Předzpracování obrazu tak, aby nebyl nikdy příliš světlý nebo tmavý, by zajisté pomohlo doručit viditelné zlepšení



Obrázek 5.8: Méně úspěšný model nedokáže odhalit špatně viditelné vejce.

výsledků. Je možné, že existuje i efektivnější kombinace konfigurace a výchozího modelu.

Na závěr by bylo vhodné shrnout, že efektivita rozpoznávání vajec touto metodou je **velmi dobrá**.

Implementace rozpoznávání obrazu

V této kapitole se zaměříme na implementaci neuronové sítě, která bude mít za cíl rozpoznat, do které kategorie snímek patří. Jedná se tedy o klasifikaci do předem daných kategorií. Vstupem pro naše skripty bude obrázek libovolné velikosti. Interně bude snímek reprezentován vektorem o velikosti $300 \times 300 \times 3$ (snímek zmenšen na velikost 300×300 bodů a zůstane barevný – každý bod obsahuje 3 barevné složky). Formát a tvorba tohoto vektoru viz algoritmus 3.2. Kategorií, do kterých budeme chtít snímky rozřadit, bude jedenáct:

Kategorie „0“: Pro snímky, kde je počet vajec roven **0**.

Kategorie „1“: Pro snímky, kde je počet vajec roven **1**.

Kategorie „2“: Pro snímky, kde je počet vajec roven **2**.

Kategorie „3“: Pro snímky, kde je počet vajec roven **3**.

Kategorie „4“: Pro snímky, kde je počet vajec roven **4**.

Kategorie „5“: Pro snímky, kde je počet vajec roven **5**.

Kategorie „6“: Pro snímky, kde je počet vajec roven **6**.

Kategorie „7“: Pro snímky, kde je počet vajec roven **7**.

Kategorie „8“: Pro snímky, kde je počet vajec roven **8**.

Kategorie „9“: Pro snímky, kde je počet vajec roven **9**.

Kategorie „10“: Pro snímky, kde je počet vajec roven **10**.

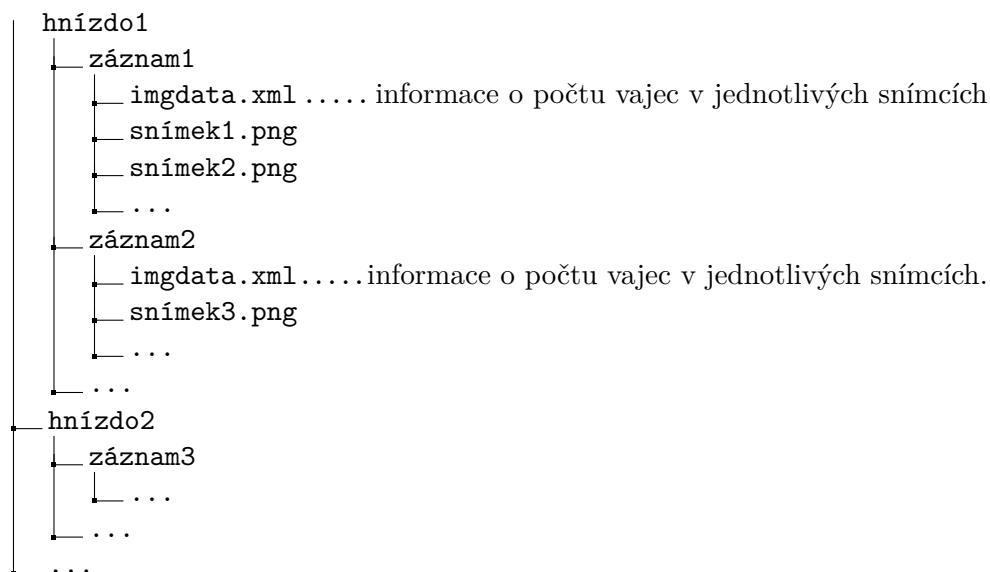
Celé řešení – příprava dat, trénování neuronové sítě, testování funkčnosti a měření přesnosti budeme implementovat v programovacím jazyce Python 3.

6.1 Příprava dat

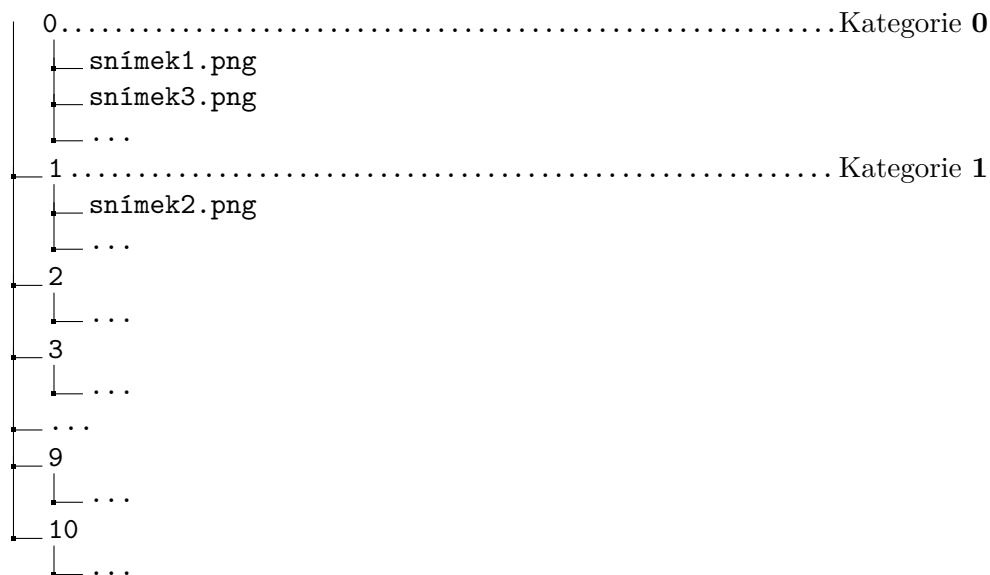
Nástrojem Tagger (viz kapitola C) jsem označil **6178 snímků**. Všechny tyto snímky budou sloužit jako trénovací nebo testovací data pro tuto kapitolu.

Abychom nemuseli vyvíjet vlastní skripty pro trénování neuronové sítě, ale mohli použít skripty standartně dostupné [31], musíme upravit strukturu trénovacích a testovacích dat.

Současná struktura dat je následující:



Nová struktura dat, které potřebujeme docílit:



Každá kategorie má vlastní složku. Do každé kategorie patří snímky s počtem vajec, který odpovídá dané kategorii. Jakmile máme trénovací data uspořádaná do požadované struktury, je vše připraveno pro trénování neuronové sítě.

6.2 Trénování neuronové sítě

Moderní modely pro rozpoznávání obrazu mají miliony parametrů a je extrémně výpočetně náročné je vytrénovat. Učení „přenosem modelu“²² je technika, která ušetří spoustu práce využitím již předtrénovaného modelu a přetrénováním pouze finálních vrstev [25]. Více informací k efektivitě tohoto řešení viz [25].

Předpokladem pro trénování neuronové sítě je nainstalovaná knihovna TensorFlow a všechny její závislosti [24]. Model, ze kterého budeme vycházet, je **Inception-v4**²³ [32], který byl vytrénován společností Google na přibližně 1,2 mil. snímků [33]. V soutěži ImageNet [7] drží model Inception-v4 nejlepšího skóre: Top-1 Accuracy²⁴ 80.2% a Top-5 Accuracy²⁵ 95.2% [34].

Googlem poskytovaný skript pro přetrénování finálních vrstev nepodporuje nejnovější model Inception-v4. Stačí však pár modifikací a můžeme nový model použít. Upravený skript se nachází v příloze H.

6.2.1 Struktura aplikace

Na disku vytvoříme složku `egg_recognition`, ve které se budeme pohybovat. Budeme potřebovat tuto strukturu:

```
egg_recognition ..... pracovní složka
├── bottlenecks
├── models
│   └── inception_v4.pb ..... předtrénovaný model
├── training_summaries
├── result ..... umístění výsledného modelu
└── training_data ..... trénovací data ve formátu specifikovaném výše
```

Poté spustíme připravený skript `retrain.py`:

```
cd egg_recognition
python3 retrain.py \
  --bottleneck_dir=bottlenecks \
```

²²Transfer learning.

²³Dostupný ke stažení na: http://download.tensorflow.org/models/inception_v4_2016_09_09.tar.gz

²⁴Odpověď modelu (ta s nejvyšší pravděpodobností) přesně odpovídala očekávanému výsledku.

²⁵Kterákoliv z 5 nejpravděpodobnějších odpovědí modelu odpovídala očekávanému výsledku.

```
--how_many_training_steps=8000 \  
--model_dir=models/ \  
--summaries_dir=training_summaries/ \  
--output_graph=result/egg_classifier_graph.pb \  
--output_labels=result/egg_classifier_labels.txt \  
--image_dir=training_data \  
--print_misclassified_test_images \  
--random_crop=16 \  
--random_scale=7 \  
--random_brightness=4
```

Hodnoty parametrů `how_many_training_steps`, `random_brightness`, `random_scale` a `random_crop` můžeme změnit. K hodnotám uvedeným výše jsem dospěl opakovaným testováním, kde tyto hodnoty produkovaly nejlepší výsledky. Skript `retrain.py` [31] poskytuje nejlepší dokumentaci všech dostupných parametrů.

6.3 Ověření funkčnosti a výsledky

Vytrénovaný a vyexportovaný model můžeme otestovat pomocí veřejně dostupného skriptu `label_image.py` [35]:

```
cd egg_recognition  
python3 ~/tensorflow/examples/image_retraining/label_image.py \  
--graph=result/egg_classifier_graph.pb \  
--labels=/tmp/output_labels.txt \  
--output_layer=final_result:0 \  
--image=test_image.jpg
```

Výsledek je prezentován v následujícím formátu (hodnoty mohou být odlišné):

```
5 (score = 0.62071)  
4 (score = 0.44595)  
6 (score = 0.43252)  
0 (score = 0.43049)  
9 (score = 0.00032)
```

6.4 Závěr

Dokázali jsme, že řešení funguje. Při hromadném testu dat zjistíme, že tato implementace **není** příliš efektivní i přesto, že náš trénovací vzorek dat je poměrně velký (**6178 snímků**). Už ze samotného výsledku při vyhodnocení pouze jednoho snímku je vidět, že si síť není jistá, do jaké kategorie daný snímek zařadit. Všechny výsledky mají buď relativně nízké skóre nebo naopak

všechny poměrně vysoké. Bohužel se nedostáváme k výsledku, ve kterém by byla pouze jedna kategorie dominantní²⁶.

Největší problém spočívá v **malých rozdílech mezi jednotlivými kategoriemi**. Vajíčka tvoří pouze malou část snímku, takže rozdíl mezi jedním nebo dvěma vajíčky je malý. **Většina plochy snímku se stává šumem**, který „mate“ tuto implementaci. Kdyby kategorie reprezentovaly dvě velmi odlišné věci, jako například auto a psa, bylo by výrazně snazší snímek mezi tyto dvě kategorie rozřadit.

Řešení by se stalo mnohem efektivnějším v případě, kdyby snímky, které síť vyhodnocuje, byly předem zpracované a upravené. Normalizace jasu a především ořezání snímku tak, aby na něm zbyla pouze vajíčka, by zvýšila efektivitu této implementace.

Tato kapitola neposkytla výsledky, ve které jsem doufal. Je zde ale prostor pro vylepšení, která by toto řešení mohla udělat efektivnějším, než detekce objektů popsána v kapitole 5.

²⁶Síť by jedné kategorii přiřadila vysoké skóre (> 80%) a zbytek kategorií by měl skóre nízké.

Implementace Java knihovny

Předmětem této kapitoly je tvorba softwarové knihovny, která je cílem této práce. Nejdříve porovnáme výsledky dvou implementací z kapitoly 5 a 6. Jako další krok budeme diskutovat, jakým způsobem integrovat vybranou implementaci do programovacího jazyka Java. Následně se zaměříme na uživatelské rozhraní knihovny a její implementaci. Na závěr vyřešíme distribuci knihovny nástrojem pro správu, řízení a automatizaci sestavování aplikací.

7.1 Srovnání implementací neuronových sítí a volba řešení

Vzhledem k výsledkům popsaných v závěrech kapitol 5 a 6 je zřejmé, že detekce objektů, popsaná v kapitole 5, je podstatně efektivnější implementací.

Nicméně, výsledná knihovna je navržena tak, že kdykoliv může být styl řešení změněn bez toho, aniž by muselo být modifikováno uživatelské rozhraní.

7.2 Použití TensorFlow v Javě

Abychom mohli použít model pro detekci vajec v naší knihovně, je nutné ho exportovat. Stačí postupovat v krocích popsaných v kapitole 5.5. Exportovaný model umístíme mezi statické zdroje knihovny.

TensorFlow poskytuje API pro jazyk Java. Abychom ho mohli používat, stačí přidat novou „závislost“ do souboru `pom.xml`:

```
1 <dependency>  
2   <groupId>org.tensorflow</groupId>  
3   <artifactId>tensorflow</artifactId>  
4   <version>1.4.0</version>  
5 </dependency>
```

Java API, poskytované knihovnou TensorFlow, je poměrně nízkoúrovňové [36]. Pro usnadnění práce použijeme wrapper²⁷ inspirovaný rozhraním, které TensorFlow poskytuje pro platformu Android.

7.3 Uživatelské rozhraní knihovny

Při tvorbě uživatelského rozhraní vycházíme z prvotního návrhu, který vznikl analýzou funkčních požadavků (viz algoritmus 3.1). Kompletní dokumentace API, které knihovna uživateli poskytuje je dostupná v příloze B. Nejlepším zdrojem ukázek použití knihovny jsou testy, které diskutujeme v kapitole 8.

7.4 Implementace

Kompletní zdrojové kódy implementace jsou dostupné v příloze H. Knihovnu zkompilejeme nástrojem Maven: `mvn clean package -DskipTests`.

Neuronová síť, která se stará o detekci a klasifikaci objektů na jednotlivých snímcích, vrací výsledek jako množinu detekcí. Každá detekce se skládá z informací o umístění objektu, typu objektu a pravděpodobnosti v procentech, že se opravdu o daný objekt jedná. Je tedy nutné stanovit hranici²⁸, která slouží jako mezník, zda-li je výsledek relevantní, či nikoliv. Jestliže nastavíme hranici například na 30%, knihovna vrátí pouze detekované objekty, o kterých si je jistá alespoň na 30%. Kdybychom nastavili hranici na 0%, knihovna by vrátila všechny možné detekce nalezené v obrázku. Celkový počet detekcí na snímku je limitován na sto. V kapitole 8.3 se zaměříme na to, jak nastavení hranice dokáže ovlivnit efektivitu knihovny.

Algoritmus 7.1: Získání finálního počtu vajec pro celou složku pomocí dat z jednotlivých snímků.

```
1 public Integer getFinalCount() {
2     TreeMap<Integer, Integer> scores = new TreeMap<>();
3
4     for (Integer val : imageScores.values()) {
5         if (scores.containsKey(val)) {
6             scores.replace(val, scores.get(val) + 1); // increment
7         } else {
8             scores.put(val, 1);
9         }
10    }
11
12    int bestGuess = 0;
13    while (!scores.isEmpty()) {
14        Map.Entry<Integer, Integer> e = scores.pollLastEntry();
```

²⁷Třída, která „obalí“ jinou třídu a přidá jí funkčnost.

²⁸Threshold.

```

15     if (e.getValue() > 1) { // threshold (how many times do we need
16         the value)
17         return e.getKey();
18     } else if (e.getValue() == 1) {
19         bestGuess = e.getValue();
20     }
21 }
22
23 return bestGuess;
24 }

```

Algoritmus 7.2: Transformace vstupního snímku bilineární interpolací.

```

1 BufferedImage thumbnail = new BufferedImage(INPUT_SIZE, INPUT_SIZE,
2     BufferedImage.TYPE_INT_RGB);
3 Graphics2D tGraphics2D = thumbnail.createGraphics(); //create a
4     graphics object to paint to
5 tGraphics2D.setBackground(Color.WHITE);
6 tGraphics2D.setPaint(Color.WHITE);
7 tGraphics2D.fillRect(0, 0, INPUT_SIZE, INPUT_SIZE);
8 tGraphics2D.setRenderingHint(RenderingHints.KEY_INTERPOLATION,
9     RenderingHints.VALUE_INTERPOLATION_BILINEAR);
10 tGraphics2D.drawImage(image, 0, 0, INPUT_SIZE, INPUT_SIZE, null);
11
12 // convert img to [INPUT_SIZE, INPUT_SIZE, 3]
13 BufferedImage convertedImg = new BufferedImage(thumbnail.getWidth(),
14     thumbnail.getHeight(), BufferedImage.TYPE_INT_RGB);
15 convertedImg.getGraphics().drawImage(thumbnail, 0, 0, null);

```

Algoritmus 7.3: Detekce objektů pomocí TensorFlow v Javě.

```

1 intValues = ((DataBufferInt)
2     convertedImg.getRaster().getDataBuffer()).getData();
3
4 for (int i = 0; i < intValues.length; ++i) {
5     byteValues[i * 3 + 2] = (byte) (intValues[i] & 0xFF);
6     byteValues[i * 3 + 1] = (byte) ((intValues[i] >> 8) & 0xFF);
7     byteValues[i * 3 + 0] = (byte) ((intValues[i] >> 16) & 0xFF);
8 }
9
10 // Copy the input data into TensorFlow.
11 inferenceInterface.feed(INPUT_NAME, byteValues, 1, INPUT_SIZE,
12     INPUT_SIZE, 3);
13
14 // Run the inference call.
15 inferenceInterface.run(outputNames, logStats);

```

```
15 // Copy the output Tensor back into the output array.
16 outputLocations = new float[MAX_RESULTS * 4];
17 outputScores = new float[MAX_RESULTS];
18 outputClasses = new float[MAX_RESULTS];
19 outputNumDetections = new float[1];
20 inferenceInterface.fetch(outputNames[0], outputLocations);
21 inferenceInterface.fetch(outputNames[1], outputScores);
22 inferenceInterface.fetch(outputNames[2], outputClasses);
23 inferenceInterface.fetch(outputNames[3], outputNumDetections);
```

7.5 Distribuce knihovny

Abychom splnili funkční požadavek **F2** (viz kapitola 1.2) o distribuci knihovny, použijeme `maven-assembly-plugin`:

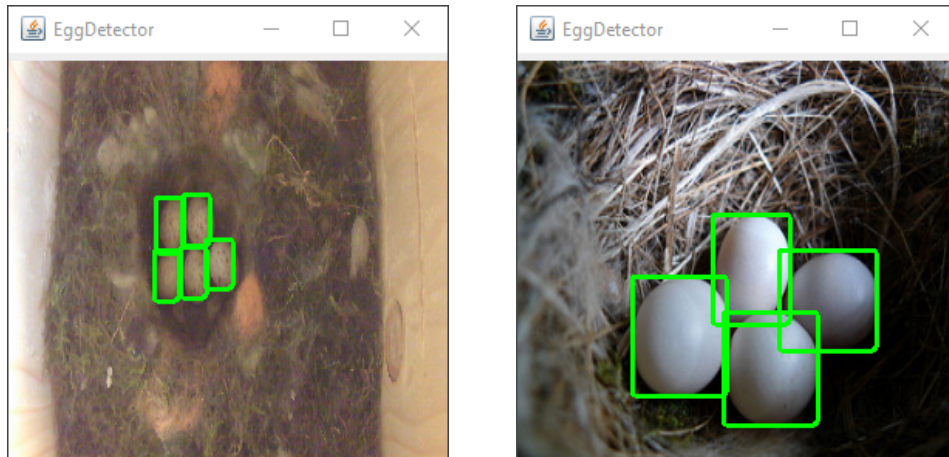
```
1 <plugin>
2   <artifactId>maven-assembly-plugin</artifactId>
3   <configuration>
4     <descriptorRefs>
5       <descriptorRef>jar-with-dependencies</descriptorRef>
6     </descriptorRefs>
7   </configuration>
8   <executions>
9     <execution>
10      <id>make-assembly</id>
11      <phase>package</phase>
12      <goals>
13        <goal>single</goal>
14      </goals>
15    </execution>
16  </executions>
17 </plugin>
```

7.6 Výsledek

Výsledkem této kapitoly je funkční softwarová knihovna v programovacím jazyce Java, která splňuje všechny cíle stanovené na začátku práce. Ke knihovně je dostupná kompletní uživatelská dokumentace.

Uživatel získá přístup k detekovanému počtu vajec v požadované složce (zjišťování celkového počtu viz alg. 7.1) i k informacím o jednotlivých snímcích. Uživatel má možnost změnit hranici, podle které knihovna určí, jaká podmnožina výsledků je relevantní. Například při nastavení hranice na 50% budou brány v potaz pouze detekce, u kterých si je knihovna jistá alespoň na 50%.

Knihovna obsahuje „debugovací“ režim, který vizualizuje, jakým způsobem jsou detekce reprezentovány interně. Ukázka „debugovacího“ režimu je vidět na obrázcích 7.1.



Obrázek 7.1: Ukázka debugovacího režimu knihovny EggDetector.

Ověření implementace

V této kapitole se zaměříme na testování, validaci a měření efektivity implementace knihovny popsané v kapitole 7. Všechny testy jsou implementovány v jazyce Java pomocí testovacího frameworku JUnit [37]. Zdrojové kódy testů jsou dostupné v příloze H.

8.1 Testování funkčnosti

Abychom dokázali, že současná implementace knihovny funguje, byly sestaveny dva „jednotkové testy“²⁹.

První test využívá pět snímků přiložených ke zdrojovým kódům testů. Nejdříve porovnává předem známý počet vajec v těchto snímcích s počtem, který vypočítá knihovna. Následně je porovnán výsledný počet vajec, který byl vypočítán z těchto pěti snímků.

Algoritmus 8.1: Testování softwarové knihovny EggDetector.

```
1 public class PackagedDataTest {
2
3     static EggDetector eggDetector;
4     static SequenceClassifier sequenceClassifier;
5
6     @BeforeClass
7     public static void setUp() {
8         eggDetector = new EggDetector();
9         sequenceClassifier = eggDetector.evaluate(new
10             File(PackagedDataTest.class.getClassLoader().
11                 getResource("sample_images").getFile()));
12     }
13
14     @AfterClass
15     public static void close() {
```

²⁹Unit testing.

```

15     eggDetector.closeSession();
16 }
17
18 @Test
19 public void testIndividualScores() {
20     Map<String, Integer> res =
21         sequenceClassifier.getIndividualCounts();
22     Assert.assertEquals((int) res.get("image1.png"), 9);
23     Assert.assertEquals((int) res.get("image2.png"), 5);
24     Assert.assertEquals((int) res.get("image3.jpg"), 3);
25     Assert.assertEquals((int) res.get("image4.jpg"), 4);
26     Assert.assertEquals((int) res.get("image5.jpg"), 3);
27 }
28
29 @Test
30 public void testFinalScore() {
31     // we expect 3, because we counted 3 eggs 2 times
32     Assert.assertEquals((int) sequenceClassifier.getFinalCount(),
33         3);
34 }

```

Druhý test porovnává výsledky knihovny s manuálně označenými daty nástrojem `Tagger` (viz kapitola C). Vstupem je pouze jedna složka s libovolným počtem snímků.

8.2 Testování distribuce

Abychom otestovali, že knihovna splňuje funkční požadavek na distribuci, vytvoříme nový projekt, ve kterém se pokusíme knihovnu použít.

Přidáme knihovnu `EggDetector` do projektu:

```

1 <dependency>
2   <groupId>org.cvut.havluja1</groupId>
3   <artifactId>eggdetector</artifactId>
4   <version>1.0</version>
5   <scope>system</scope>
6   <systemPath>
7     ${basedir}/../../../../bin/eggdetector-1.0-jar-with-dependencies.jar
8   </systemPath>
9 </dependency>

```

a použijeme ji:

```

1 EggDetector eggDetector = new EggDetector();
2 SequenceClassifier sequenceClassifier = eggDetector.evaluate(new
3   File(Main.class.getClassLoader().
4     getResource("sample_images").getFile()));

```

```

4
5 System.out.println("final count: " +
6     sequenceClassifier.getFinalCount());
7 System.out.println("individual scores: " +
8     sequenceClassifier.getIndividualCounts());
9
10 eggDetector.closeSession();

```

Lze vidět, že distribuce funguje. Zdrojové kódy externího projektu se nachází v příloze H ve složce `src/impl/libtest`.

8.3 Měření efektivity implementace

Pro určení efektivity naší implementace využijeme manuálně označených **6178 snímků** nástrojem Tagger. Porovnáme také výsledky při různých nastaveních „thresholdu“ knihovny (viz kapitola 7.4).

Zdrojový kód testu je vypsán v příloze F. Knihovnu otestujeme pětikrát na stejných datech s pěti různými nastaveními „thresholdu“. V testech se zaměříme na dvě metriky. První metrikou je poměr správně vyhodnocených složek a celkového počtu složek, vyjádřený v procentech. Tu získáme následujícím způsobem:

$$\text{úspěšnost} = \frac{\text{počet správně vyhodnocených složek}}{\text{celkový počet složek}} * 100\%$$

Druhou metrikou je rozdíl mezi očekávaným počtem vajec a výsledkem, který dostaneme použitím knihovny:

$$\text{vzdálenost} = \sum_{n=1}^{\text{počet videosekvencí}} |\text{známý počet vajec} - \text{počet vajec vypočítaný knihovnou}|$$

Druhá metrika je lepším nástrojem pro měření efektivity, protože říká, o kolik vajec se knihovna „spletla“. Z toho vyplývá, že **čím nižší číslo, tím lepší efektivita**.

8.3.1 Test efektivity s nastavením hranice na 60%

```

Found 186 directories.
EggDetector evaluated 57 directories correctly.
57/186: 30.645163% success rate.
Egg count of all folders added together: 982.
Distance (|real eggs - found eggs|): 409 eggs (smaller is better).

```

8.3.2 Test efektivity s nastavením hranice na 50%

```
Found 186 directories.  
EggDetector evaluated 59 directories correctly.  
59/186: 31.72043% success rate.  
Egg count of all folders added together: 982.  
Distance (|real eggs - found eggs|): 367 eggs (smaller is better).
```

8.3.3 Test efektivity s nastavením hranice na 40%

```
Found 186 directories.  
EggDetector evaluated 62 directories correctly.  
62/186: 33.333336% success rate.  
Egg count of all folders added together: 982.  
Distance (|real eggs - found eggs|): 336 eggs (smaller is better).
```

8.3.4 Test efektivity s nastavením hranice na 30%

```
Found 186 directories.  
EggDetector evaluated 62 directories correctly.  
62/186: 33.333336% success rate.  
Egg count of all folders added together: 982.  
Distance (|real eggs - found eggs|): 311 eggs (smaller is better).
```

8.3.5 Test efektivity s nastavením hranice na 20%

```
Found 186 directories.  
EggDetector evaluated 46 directories correctly.  
46/186: 24.731182% success rate.  
Egg count of all folders added together: 982.  
Distance (|real eggs - found eggs|): 340 eggs (smaller is better).
```

8.4 Shrnutí kapitoly

V této kapitole bylo dokázáno, že implementace knihovny funguje a je možné ji použít formou závislostí z jiného projektu. Měřením efektivity knihovny jsme došli k závěru, že nejlepším nastavením hranice pro detekci objektů je **30%**. Knihovnu tedy upravíme tak, aby toto nastavení používala jako výchozí hodnotu.

Závěr

Cílem této práce byl návrh a implementace softwarové knihovny pro určení počtu vajec v ptačích hnízdech. Byla sepsána detailní specifikace funkčních a nefunkčních požadavků, které popisují požadovanou formu knihovny. Prerekvizitou pro implementaci samotné knihovny bylo vyřešení úlohy „rozpoznávání vajec v obraze“.

Čtenáři byla prezentována veškerá potřebná teorie o strojovém učení, neuronových sítích a nástroji TensorFlow. Na základě získaných teoretických znalostí byla provedena analýza funkčních a nefunkčních požadavků; projektu Ptáci Online; současného stavu řešení a struktury vstupních dat. Následně byly navrženy dva směry implementace, kterými bychom mohli úlohu rozpoznávání vajec vyřešit. V poslední části analýzy byly zvoleny vhodné principy a technologie a bylo navrženo uživatelské rozhraní knihovny.

V další kapitole byl čtenář seznámen se všemi nástroji, které jsou potřebné k dokončení implementace. Navrhl a naimplementoval jsem dva plnohodnotné nástroje pro práci s daty a různé skripty pro jejich zpracování. Pomocí těchto nástrojů jsem manuálně připravil **8068** trénovacích dat – označení vajec na snímcích – které byly následně použity k trénování neuronových sítí.

Prvním řešením pro detekci vajec byla implementace realizována technikou „detekce objektů v obraze“. Tato implementace přinesla velmi dobré výsledky a proto jsem se ji rozhodl použít jako základ pro výslednou knihovnu (viz sekce 5.6). Druhý styl implementace nebyl tak efektivní, a proto jsem se rozhodl ho nepoužít (viz sekce 6.4). Nicméně i tento styl implementace je zajímavý a může sloužit jako užitečný základ pro čtenáře, který se bude zabývat rozpoznáváním obrazu.

Nakonec byla v kapitole 7 implementována samotná knihovna v programovacím jazyce Java, která je schopna automaticky rozpoznat počet vajec v hnízdě. Knihovnu jsem řádně otestoval a změřil její efektivitu (viz kapitola 8).

Do budoucna by bylo vhodné systém rozšířit o zpracování obrazu před samotnou detekcí vajec. Například normalizace jasu, kontrastu a expozice by

pomohla výsledky značně zlepšit. Je zde také prostor pro přípravu více trénovacích dat a vylepšení konfigurace neuronové sítě.

Cíl této práce byl v plné míře naplněn, včetně všech funkčních a nefunkčních požadavků. Největším přínosem práce je systém pro detekci objektů v obraze. Vzhledem k návrhu systému stačí, aby byl připraven dostatečně velký vzorek trénovacích dat pro jakýkoliv objekt a můžeme použít tentýž systém například pro detekci mládat v hnízdech. Tato práce také slouží jako vhodný základ pro čtenáře, kterého zajímá řešení problémů počítačového vidění pomocí neuronových sítí a nástroje TensorFlow.

Literatura

- [1] Česká zemědělská univerzita v Praze, Fakulta životního prostředí: *O projektu ptáci online*. [online]. [cit. 2017-10-27]. Dostupné z: <http://www.ptacionline.cz/o-projektu/o-projektu>
- [2] Česká zemědělská univerzita v Praze, Fakulta životního prostředí: *Ptáci online v médiích*. [online]. [cit. 2017-10-27]. Dostupné z: <http://www.ptacionline.cz/o-projektu>
- [3] Oracle: *jar-The Java Archive Tool*. [online]. [cit. 2017-11-14]. Dostupné z: <https://docs.oracle.com/javase/7/docs/technotes/tools/windows/jar.html>
- [4] The Apache Software Foundation: *Maven* [online]. 2017. Dostupné z: <https://maven.apache.org/>
- [5] Ballard, D. H.; Brown, C. M.: *Computer Vision*. Prentice Hall, 1982, ISBN 0-13-165316-4, 523 s., [cit. 2018-01-02]. Dostupné z: http://homepages.inf.ed.ac.uk/rbf/BOOKS/BANDB/Ballard__D._and_Brown__C.M.__1982__Computer_Vision.pdf
- [6] Klette, R.: *Concise Computer Vision*. Springer, 2014, ISBN 978-1-4471-6320-6, [cit. 2017-10-19].
- [7] Russakovsky, O.; Deng, J.; Su, H.; aj.: ImageNet Large Scale Visual Recognition Challenge. Technická zpráva, Stanford University, University of Michigan, Massachusetts Institute of Technology & UNC Chapel Hill, 2015, [cit. 2017-09-29]. Dostupné z: <https://arxiv.org/pdf/1409.0575.pdf>
- [8] XenonStack: *Overview of Artificial Neural Networks and its Applications* [online]. [cit. 2017-11-17]. Dostupné z: <https://medium.com/@xenonstack/overview-of-artificial-neural-networks-and-its-applications-2525c1adff7>

- [9] Zell, A.: *Simulation of Neural Networks*. Addison-Wesley, 1994, ISBN 3-89319-554-8, [cit. 2017-11-19].
- [10] Google Inc.: *About TensorFlow [online]*. [cit. 2017-12-06]. Dostupné z: <https://www.tensorflow.org/>
- [11] Google Inc.: *Building the CNN MNIST Classifier [online]*. [cit. 2017-12-23]. Dostupné z: https://www.tensorflow.org/tutorials/layers#building_the_cnn_mnist_classifier
- [12] Google Inc.: *TensorFlow: Preparing Inputs [online]*. [cit. 2017-11-22]. Dostupné z: https://github.com/tensorflow/models/blob/master/research/object_detection/g3doc/using_your_own_dataset.md
- [13] Google Inc.: *Tensorflow Object Detection API [online]*. [cit. 2017-11-17]. Dostupné z: https://github.com/tensorflow/models/tree/master/research/object_detection
- [14] Ministerstvo životního prostředí: *Ptáci online: Sledujte záběry z „chytré ptačí budky“ na budově MŽP [online]*. [cit. 2017-11-18]. Dostupné z: http://www.mzp.cz/cz/news_160608_Ptaci_online
- [15] Pavlíček, J.; Švec, V.; Tichá, I.: *Artificial intelligence in education: Can the AI Teach Them?* Czech University of Life Sciences Prague, Faculty of Economics and Management, 2015, ISBN 978-80-213-2560-9, 436 – 443 s., [cit. 2018-01-05].
- [16] Šuma, P.: *Detekce počtu mládat v ptačích hnízdech za pomoci nástrojů rozpoznání obrazu*. Bakalářská práce, České vysoké učení technické v Praze, Fakulta informačních technologií, 2017, [cit. 2017-09-05].
- [17] Hanzlík, P.; Kožíšek, F.; Pavlíček, J.: Design of intelligent decision support systems in agriculture. *INTERNATIONAL JOURNAL OF MATHEMATICS AND COMPUTERS IN SIMULATION*, ročník 9, 2015: s. 113 – 118, ISSN 1998-0159, [cit. 2018-01-05]. Dostupné z: <http://www.naun.org/main/NAUN/mcs/2015/a322002-233.pdf>
- [18] Google Inc.: *Image Recognition [online]*. [cit. 2017-11-09]. Dostupné z: https://www.tensorflow.org/tutorials/image_recognition
- [19] Google Inc.: *Tensorflow API Documentation [online]*. [cit. 2017-11-17]. Dostupné z: https://www.tensorflow.org/api_docs/
- [20] Google Inc.: *Getting Started with TensorFlow [online]*. [cit. 2017-11-17]. Dostupné z: https://www.tensorflow.org/get_started/
- [21] Google Inc.: *TensorFlow: Importing Data [online]*. [cit. 2017-11-17]. Dostupné z: https://www.tensorflow.org/programmers_guide/datasets

-
- [22] Pivotal Software, Inc.: *Spring Boot [online]*. 2018. Dostupné z: <https://projects.spring.io/spring-boot/>
- [23] Tzutalin: *LabelImg [online]*. Git code (2015). Dostupné z: <https://github.com/tzutalin/labelImg>
- [24] Google Inc.: *Installing TensorFlow [online]*. [cit. 2017-12-06]. Dostupné z: <https://www.tensorflow.org/install/>
- [25] Donahue, J.; Jia, Y.; Vinyals, O.; aj.: DeCAF: A Deep Convolutional Activation Feature for Generic Visual Recognition. Technická zpráva, UC Berkeley & ICSI, 2013, [cit. 2017-09-08]. Dostupné z: <https://arxiv.org/pdf/1310.1531v1.pdf>
- [26] Google Inc.: *Tensorflow detection model zoo [online]*. [cit. 2017-12-17]. Dostupné z: https://github.com/tensorflow/models/blob/master/research/object_detection/g3doc/detection_model_zoo.md
- [27] Google Inc.: *Configuring the Object Detection Training Pipeline [online]*. [cit. 2017-12-01]. Dostupné z: https://github.com/tensorflow/models/blob/master/research/object_detection/g3doc/configuring_jobs.md
- [28] Google Inc.: *Cloud Machine Learning Engine [online]*. [cit. 2017-10-06]. Dostupné z: <https://cloud.google.com/ml-engine/>
- [29] Govindan, S.: *Object Detection: Tensorflow and Google Cloud Platform [online]*. Pub. Nov 5, 2017 [cit. 2017-12-17]. Dostupné z: <https://medium.com/google-cloud/object-detection-tensorflow-and-google-cloud-platform-72e0a3f3bdd6>
- [30] Project Jupyter: *Jupyter Notebook Documentation [online]*. 2017. Dostupné z: <http://jupyter.org/documentation.html>
- [31] Google Inc.: *Simple transfer learning with Inception v3 or Mobilenet models. [online]*. [cit. 2017-12-03]. Dostupné z: https://github.com/tensorflow/tensorflow/blob/master/tensorflow/examples/image_retraining/retrain.py
- [32] Szegedy, C.; Ioffe, S.; Vanhoucke, V.; aj.: Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. Technická zpráva, Google Inc., 2016, [cit. 2017-10-09]. Dostupné z: <https://arxiv.org/pdf/1512.00567.pdf>
- [33] Google Inc.: *Preparing the datasets [online]*. [cit. 2017-12-12]. Dostupné z: <https://github.com/tensorflow/models/tree/master/research/slim#preparing-the-datasets>

- [34] Russakovsky, O.; Deng, J.; Su, H.; aj.: ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision (IJCV)*, ročník 115, č. 3, 2015: s. 211–252, doi:10.1007/s11263-015-0816-y.
- [35] Google Inc.: *Using the Retrained Model*. [online]. [cit. 2017-12-03]. Dostupné z: https://github.com/tensorflow/tensorflow/blob/master/tensorflow/examples/label_image/label_image.py
- [36] Google Inc.: *org.tensorflow* [online]. [cit. 2017-12-12] (TensorFlow Java API v1.4). Dostupné z: https://www.tensorflow.org/api_docs/java/reference/org/tensorflow/package-summary
- [37] Project Jupyter: *JUnit 5 User Guide* [online]. 2017. Dostupné z: <http://junit.org/junit5/docs/current/user-guide/>
- [38] Oracle: *Javadoc tool*. [online]. [cit. 2017-11-14]. Dostupné z: <http://www.oracle.com/technetwork/java/javase/documentation/index-jsp-135444.html>
- [39] *TexDoclet* [online]. 2017. Dostupné z: <http://doclet.github.io/>

Seznam použitých zkratk

- GUI** Graphical user interface
- XML** Extensible markup language
- HTML** HyperText Markup Language
- RAM** Random-access memory
- PDF** Portable Document Format
- NN** Neural Network
- CNN** Convolutional Neural Network
- RCNN** Recursive Convolutional Neural Network
- API** Application Programming Interface
- PNG** Portable Network Graphics
- JPEG** Joint Photographic Experts Group

Dokumentace API knihovny

Příloha obsahuje dokumentaci API Java knihovny, která je výsledkem této práce. Dokumentace byla vygenerována ze zdrojových kódů nástroji Javadoc [38] a TexDoclet [39]. Kód a dokumentace je psána v anglickém jazyce, proto je i text zde v angličtině.

B.1 Package org.cvut.havluja1.eggdetector

<i>Package Contents</i>	<i>Page</i>
Classes	
EggDetector.....	75
SequenceClassifier	79

B.1.1 Class EggDetector

B.1.1.1 Count the number of eggs in given images

The egg detector is a library that helps you count the number of eggs in a given folder. Egg detector works by using TeonsorFlow Object Detection API in the background. To learn more, see <https://www.tensorflow.com>.

Example usage:

```
1 EggDetector eggDetector = new EggDetector();
2 SequenceClassifier sequenceClassifier = eggDetector.evaluate(new
   File("image_dir"));
3 System.out.println("final count: " +
   sequenceClassifier.getFinalCount());
4 System.out.println("individual scores: " +
   sequenceClassifier.getIndividualCounts());
5 eggDetector.closeSession();
```

B.1.1.2 Declaration

```
1 public class EggDetector
2   extends java.lang.Object
```

B.1.1.3 Constructor summary

EggDetector()

Constructor loads the pre-trained frozen graph into memory.

B.1.1.4 Method summary

closeSession()

Closes the EggDetector session.

evaluate(File)

Runs egg detection on a given *dir*.

getMinimalConfidence()

Get the minimalConfidence setting for this instance.

isDebugMode()

Get this instance's debug mode setting.

setDebugMode(boolean)

Set this instance's debug mode setting.

setMinimalConfidence(float)

Set the minimalConfidence setting for this instance.

toString()

B.1.1.5 Constructors

- *EggDetector*

```
1 public EggDetector()
```

– **Description**

Constructor loads the pre-trained frozen graph into memory.

It also checks whether TensorFlow is supported on your platform.

B.1.1.6 Methods

- *closeSession*

```
1 public void closeSession() throws java.lang.IllegalStateException
```

– **Description**

Closes the EggDetector session. This instance of EggDetector will not be usable again.

– **Throws**

- * `java.lang.IllegalStateException` – if the session has been closed already by calling `closeSession()`

- *evaluate*

```
1 public SequenceClassifier evaluate(java.io.File dir) throws
   java.lang.IllegalArgumentException,
   java.lang.IllegalStateException
```

– **Description**

Runs egg detection on a given *dir*.

– **Parameters**

- * *dir* – a directory containing .jpg or .png files for object detection

– **Returns** –

– **Throws**

- * `java.lang.IllegalArgumentException` – if *dir* is not a directory or contains no images
- * `java.lang.IllegalStateException` – if the session has been closed already by calling `closeSession()`

- *getMinimalConfidence*

```
1 public float getMinimalConfidence()
```

– **Description**

Get the *minimalConfidence* setting for this instance.

Minimal Confidence score is used as a confidence boundary during the process of object detection. An object that has been detected

with a confidence score lower than *minimalConfidence* is ignored. An object that has been detected with a confidence score higher or equal than *minimalConfidence* is added to the final result list.

– **Returns** – This instance’s *minimalConfidence* setting.

- *isDebugMode*

```
1 public boolean isDebugMode()
```

– **Description**

Get this instance’s debug mode setting.

If debug mode is enabled (set to true), the library will open a `JFrame` for each processed image with detections graphically highlighted.

– **Returns** – debug mode setting for this instance

- *setDebugMode*

```
1 public void setDebugMode(boolean debugMode) throws  
    java.lang.IllegalStateException
```

– **Description**

Set this instance’s debug mode setting.

If debug mode is enabled (set to true), the library will open a `JFrame` for each processed image with detections graphically highlighted.

– **Parameters**

* `debugMode` – turn the debug mode on or off

– **Throws**

* `java.lang.IllegalStateException` – if the session has been closed already by calling `closeSession()`

- *setMinimalConfidence*

```
1 public void setMinimalConfidence(float minimalConfidence) throws  
    java.lang.IllegalStateException
```

– **Description**

Set the *minimalConfidence* setting for this instance.

Minimal Confidence score is used as a confidence boundary during the process of object detection. An object that has been detected

with a confidence score lower than *minimalConfidence* is ignored. An object that has been detected with a confidence score higher or equal than *minimalConfidence* is added to the final result list.

– **Parameters**

* `minimalConfidence` – *minimalConfidence* for this instance

– **Throws**

* `java.lang.IllegalStateException` – if the session has been closed already by calling `closeSession()`

• *toString*

```
1 public java.lang.String toString()
```

B.1.2 Class SequenceClassifier

B.1.2.1 A class containing object detection results for a given directory

SequenceClassifier is a data class containing the results of object detection for a given directory. When constructed, object detection is performed on all images and results are stored in memory.

Example usage:

```
1 EggDetector eggDetector = new EggDetector();
2 SequenceClassifier sequenceClassifier = eggDetector.evaluate(new
  File("image_dir"));
3 System.out.println("final count: " +
  sequenceClassifier.getFinalCount());
4 System.out.println("individual scores: " +
  sequenceClassifier.getIndividualCounts());
5 eggDetector.closeSession();
```

B.1.2.2 Declaration

```
1 public class SequenceClassifier
2 extends java.lang.Object
```

B.1.2.3 Method summary

getFinalCount()

Get the final score for the entire directory.

getIndividualCounts()

Gets the individual egg count for every image provided.

B.1.2.4 Methods

- *getFinalCount*

```
1 public java.lang.Integer getFinalCount()
```

– **Description**

Get the final score for the entire directory.

The final score is calculated as follows:

- * individual scores of images are sorted and counted
- * the highest egg count is returned as a result if we detected this egg count in at least two different images
- * if no two images contain the same egg count, the highest detected egg count is returned
- * if no eggs are detected in any of the images, 0 is returned

– **Returns** – final egg count for this instance

- *getIndividualCounts*

```
1 public java.util.Map getIndividualCounts()
```

– **Description**

Gets the individual egg count for every image provided.

– **Returns** – A map of individual scores. The key is the filename. The value is the egg count.

Tagger

Tato příloha obsahuje zdrojový kód a způsob použití nástroje **Tagger**. Kompletní dokumentace a zdrojový kód je dostupný na přiloženém médiu (viz příloha H).

C.1 Dokumentace (v AJ)

Data tagger is a simple web application used for data tagging used to train a neural network. The data strictly needs to follow the structure of `http://athena.pef.czu.cz/ptacionline/134572snaps/`

C.1.1 Usage manual

- Compile (if needed) the app with `mvn clean package`
- Run with `target/run.sh` on Linux or `target\run.bat` on Windows. Both scripts accept data location as the first and only parameter.
- For example: `target/run.sh /home/test/egg/data/nest1`

C.2 Zdrojový kód

Zdrojový kód je dostupný na přiloženém médiu (viz příloha H).

C.3 Ukázkový výstup

Po určení počtu vajec na každém snímku je do příslušné složky vygenerován soubor `imgdata.xml`, který je použit pro trénování a později testování neuronové sítě. Níže je přiložen ukázkový výpis souboru.

Algoritmus C.1: Ukázkový výstup programu Tagger.

```
1 <?xml version="1.0" encoding="UTF-8"?>
2 <java version="1.8.0_121" class="java.beans.XMLDecoder">
3   <object class="org.cvut.havluja1.tagger.model.FolderData">
4     <void property="imgData">
5       <object class="java.util.ArrayList">
6         <void method="add">
7           <object class="org.cvut.havluja1.tagger.model.ImgData">
8             <void property="eggCount">
9               <int>0</int>
10            </void>
11            <void property="name">
12              <string>snap0001-0000000668.png</string>
13            </void>
14          </object>
15        </void>
16        <void method="add">
17          <object class="org.cvut.havluja1.tagger.model.ImgData">
18            <void property="eggCount">
19              <int>2</int>
20            </void>
21            <void property="name">
22              <string>snap0002-0000001419.png</string>
23            </void>
24          </object>
25        </void>
26        <void method="add">
27          <object class="org.cvut.havluja1.tagger.model.ImgData">
28            <void property="eggCount">
29              <int>2</int>
30            </void>
31            <void property="name">
32              <string>snap0003-0000002341.png</string>
33            </void>
34          </object>
35        </void>
36        <void method="add">
37          <object class="org.cvut.havluja1.tagger.model.ImgData">
38            <void property="eggCount">
39              <int>8</int>
40            </void>
41            <void property="name">
42              <string>snap0004-0000003160.png</string>
43            </void>
44          </object>
45        </void>
46        <void method="add">
47          <object class="org.cvut.havluja1.tagger.model.ImgData">
48            <void property="eggCount">
```

```
49     <int>8</int>
50   </void>
51   <void property="name">
52     <string>snap0005-0000004082.png</string>
53   </void>
54 </object>
55 </void>
56 </object>
57 </void>
58 </object>
59 </java>
```

Folder Trimmer

Tato příloha obsahuje zdrojový kód a způsob použití nástroje `FolderTrimmer`. Kompletní dokumentace a zdrojový kód je dostupný na přiloženém médiu (viz příloha H).

D.1 Dokumentace (v AJ)

Folder trimmer is a command line tool for deleting useless data the download script downloads.

D.1.1 What does it do

- Deletes files that do not end on either `.png`, `.xml`, or `.txt`.
- If a folder does not contain another folder or at least one of the files listed above, the folder is deleted.

D.1.2 Usage manual

- Compile (if needed) the app with `mvn clean package`.
- Run with `target/run.sh` on Linux or `target\run.bat` on Windows. Both scripts accept two parameters:
 - First parameter can be `true` or `false` and it determines whether you want to delete everything except manually tagged data (containing `imgdata.xml`) in case of `true` or if you just want to delete the useless data in case of `false`.
 - Second parameter specifies the root folder of the data that should be trimmed.
 - For example to delete only useless data but keep all the folders with some image data, use:

```
target/run.sh false ~/eggdetector/data
```

- To delete everything except tagged data, run:

```
target/run.sh true ~/eggdetector/data
```

D.2 Zdrojový kód

Algoritmus D.1: Zdrojový kód nástroje FolderTrimmer.

```
1 package org.cvut.havluja1.foldertrimmer;
2
3 import java.io.File;
4 import java.io.IOException;
5
6 import org.apache.commons.io.FileUtils;
7 import org.apache.commons.io.FilenameUtils;
8
9 public class FolderTrimmer {
10     public static void main(String[] args) throws IOException {
11         File rootDir = new File(args[0]);
12
13         if (!rootDir.exists() || !rootDir.isDirectory()) {
14             throw new IllegalArgumentException("root dir does not exist");
15         }
16
17         System.out.println("finding useless data...");
18         findAndDeleteEmptyDirs(rootDir);
19         System.out.println("done");
20     }
21
22     private static void findAndDeleteEmptyDirs(File dir) {
23         final boolean[] shouldBeDeleted = {true};
24         final boolean leaveOnlyTaggedData =
25             System.getProperty("leaveonlytagged")
26                 .equalsIgnoreCase("true");
27
28         File[] toBeProcessed = dir.listFiles((file, s) -> {
29             File workingFile = new File(file, s);
30
31             // if dir -> return true and tag this folder not to be deleted
32             if (workingFile.isDirectory()) {
33                 shouldBeDeleted[0] = false;
34                 return true;
35             }
36         });
37     }
38 }
```



```
36
37     if (workingFile.isFile()) {
38         if (leaveOnlyTaggedData) { // if we want to keep only
39             tagged data
40             if (s.equals("imgdata.xml")) {
41                 shouldBeDeleted[0] = false;
42                 return false;
43             } else {
44                 if (FilenameUtils.getExtension(s).equals("png")) {
45                     return false;
46                 }
47                 return true;
48             }
49         } else { // If file and is not xml, txt or png return
50             true. If it is, tag this folder not to be deleted.
51             if (FilenameUtils.getExtension(s).equals("xml")
52                 || FilenameUtils.getExtension(s).equals("png")
53                 || FilenameUtils.getExtension(s).equals("txt")) {
54                 shouldBeDeleted[0] = false;
55                 return false;
56             } else {
57                 return true;
58             }
59         }
60     }
61     return true;
62 });
63
64 if (shouldBeDeleted[0]) {
65     try {
66         FileUtils.deleteDirectory(dir);
67         System.out.println("[D] deleting dir: " +
68             dir.getAbsolutePath());
69     } catch (IOException e) {
70         e.printStackTrace();
71     }
72 } else {
73     if (toBeProcessed.length > 0) {
74         for (File currFile : toBeProcessed) {
75             // if file -> delete
76             if (currFile.isFile()) {
77                 if (currFile.delete()) {
78                     System.out.println("[F] deleting file: " +
79                         currFile.getAbsolutePath());
80                 }
81                 continue;
82             }
83         }
84     }
85 }
```

D. FOLDER TRIMMER

```
81         // if dir -> recursive call
82         if (currFile.isDirectory()) {
83             findAndDeleteEmptyDirs(currFile);
84         }
85     }
86 }
87 }
88 }
89 }
```

Konfigurace detekování objektů

Výsledný konfigurační soubor použitý společně s modelem
ssd_mobilenet_v1_coco:

Algoritmus E.1: Konfigurační soubor pro detekci vajec v ptačích hnízdech.

```
1 model {
2   ssd {
3     num_classes: 1
4     box_coder {
5       faster_rcnn_box_coder {
6         y_scale: 10.0
7         x_scale: 10.0
8         height_scale: 5.0
9         width_scale: 5.0
10    }
11  }
12  matcher {
13    argmax_matcher {
14      matched_threshold: 0.5
15      unmatched_threshold: 0.5
16      ignore_thresholds: false
17      negatives_lower_than_unmatched: true
18      force_match_for_each_row: true
19    }
20  }
21  similarity_calculator {
22    iou_similarity {
23    }
24  }
25  anchor_generator {
26    ssd_anchor_generator {
27      num_layers: 6
28      min_scale: 0.2
29      max_scale: 0.95
```

E. KONFIGURACE DETEKOVÁNÍ OBJEKTŮ

```
30     aspect_ratios: 1.0
31     aspect_ratios: 2.0
32     aspect_ratios: 0.5
33     aspect_ratios: 3.0
34     aspect_ratios: 0.3333
35   }
36 }
37 image_resizer {
38   fixed_shape_resizer {
39     height: 300
40     width: 300
41   }
42 }
43 box_predictor {
44   convolutional_box_predictor {
45     min_depth: 0
46     max_depth: 0
47     num_layers_before_predictor: 0
48     use_dropout: false
49     dropout_keep_probability: 0.8
50     kernel_size: 1
51     box_code_size: 4
52     apply_sigmoid_to_scores: false
53     conv_hyperparams {
54       activation: RELU_6,
55       regularizer {
56         l2_regularizer {
57           weight: 0.00004
58         }
59       }
60       initializer {
61         truncated_normal_initializer {
62           stddev: 0.03
63           mean: 0.0
64         }
65       }
66       batch_norm {
67         train: true,
68         scale: true,
69         center: true,
70         decay: 0.9997,
71         epsilon: 0.001,
72       }
73     }
74   }
75 }
76 feature_extractor {
77   type: 'ssd_mobilenet_v1'
78   min_depth: 16
```

```

79     depth_multiplier: 1.0
80     conv_hyperparams {
81       activation: RELU_6,
82       regularizer {
83         l2_regularizer {
84           weight: 0.00004
85         }
86       }
87       initializer {
88         truncated_normal_initializer {
89           stddev: 0.03
90           mean: 0.0
91         }
92       }
93       batch_norm {
94         train: true,
95         scale: true,
96         center: true,
97         decay: 0.9997,
98         epsilon: 0.001,
99       }
100     }
101   }
102   loss {
103     classification_loss {
104       weighted_sigmoid {
105         anchorwise_output: true
106       }
107     }
108     localization_loss {
109       weighted_smooth_l1 {
110         anchorwise_output: true
111       }
112     }
113     hard_example_miner {
114       num_hard_examples: 3000
115       iou_threshold: 0.99
116       loss_type: CLASSIFICATION
117       max_negatives_per_positive: 3
118       min_negatives_per_image: 0
119     }
120     classification_weight: 1.0
121     localization_weight: 1.0
122   }
123   normalize_loss_by_num_matches: true
124   post_processing {
125     batch_non_max_suppression {
126       score_threshold: 1e-8
127       iou_threshold: 0.6

```

E. KONFIGURACE DETEKOVÁNÍ OBJEKTŮ

```
128     max_detections_per_class: 100
129     max_total_detections: 100
130   }
131   score_converter: SIGMOID
132 }
133 }
134 }
135
136 train_config: {
137   batch_size: 24
138   optimizer {
139     rms_prop_optimizer {
140       learning_rate {
141         exponential_decay_learning_rate {
142           initial_learning_rate: 0.004
143           decay_steps: 800720
144           decay_factor: 0.95
145         }
146       }
147       momentum_optimizer_value: 0.9
148       decay: 0.9
149       epsilon: 1.0
150     }
151   }
152   fine_tune_checkpoint:
153     "training/ssd_mobilenet_v1_coco_2017_11_17/model.ckpt"
154   from_detection_checkpoint: true
155   num_steps: 200000
156   data_augmentation_options {
157     random_horizontal_flip {
158     }
159   }
160   data_augmentation_options {
161     ssd_random_crop {
162     }
163   }
164 }
165
166 train_input_reader: {
167   tf_record_input_reader {
168     input_path: "data/train.record"
169   }
170   label_map_path: "training/egg_label_map.pbtxt"
171 }
172
173 eval_config: {
174   num_examples: 12
175   max_evals: 10
176 }
```

```
176
177 eval_input_reader: {
178   tf_record_input_reader {
179     input_path: "data/test.record"
180   }
181   label_map_path: "training/egg_label_map.pbtxt"
182   shuffle: false
183   num_readers: 1
184   num_epochs: 1
185 }
```


Test efektivity knihovny

Algoritmus F.1: Test k vyhodnocení efektivity výsledné knihovny.

```
1 package org.cvut.havluja1.eggdetector;
2
3 import java.beans.XMLDecoder;
4 import java.io.File;
5 import java.io.FileInputStream;
6 import java.io.IOException;
7 import java.util.Map;
8
9 import org.cvut.havluja1.tagger.model.FolderData;
10 import org.cvut.havluja1.tagger.model.ImgData;
11 import org.junit.AfterClass;
12 import org.junit.BeforeClass;
13 import org.junit.Test;
14
15 public class SuccessRateTest {
16
17     static final String DIR = "~/data/egg_count_images";
18     static EggDetector eggDetector;
19
20     @BeforeClass
21     public static void setUp() {
22         eggDetector = new EggDetector();
23     }
24
25     @AfterClass
26     public static void close() {
27         eggDetector.closeSession();
28     }
29
30     @Test
31     public void testHighThreshold(){
```

```
32     eggDetector.setMinimalConfidence(0.6f);
33     System.out.println(eggDetector);
34     testSuccessRate();
35 }
36
37 @Test
38 public void testMiddleHighThreshold(){
39     eggDetector.setMinimalConfidence(0.5f);
40     System.out.println(eggDetector);
41     testSuccessRate();
42 }
43
44 @Test
45 public void testMiddleThreshold(){
46     eggDetector.setMinimalConfidence(0.4f);
47     System.out.println(eggDetector);
48     testSuccessRate();
49 }
50
51 @Test
52 public void testMiddleLowThreshold(){
53     eggDetector.setMinimalConfidence(0.3f);
54     System.out.println(eggDetector);
55     testSuccessRate();
56 }
57
58 @Test
59 public void testLowThreshold(){
60     eggDetector.setMinimalConfidence(0.2f);
61     System.out.println(eggDetector);
62     testSuccessRate();
63 }
64
65 private void testSuccessRate() {
66     File workingDir = new File(DIR);
67     // get all directories
68     File[] subdirs = workingDir.listFiles((file, name) -> {
69         File tmp = new File(file, name);
70         if (tmp.isDirectory()) {
71             return true;
72         }
73
74         return false;
75     });
76
77     int totalCnt = 0;
78     int correctCnt = 0;
79
80     // distance
```

```

81     int totalEggCount = 0;
82     int lengthDifference = 0;
83
84     for (File dir : subdirs) {
85         try {
86             int foundTmp = evaluateDir(dir);
87             int expectedTmp = getExpectedCount(dir);
88
89             totalEggCount += expectedTmp;
90             totalCnt++;
91             if (foundTmp == expectedTmp) {
92                 correctCnt++;
93             } else {
94                 lengthDifference += Math.abs(foundTmp -
95                     expectedTmp);
96             }
97             System.out.println("expected: " + expectedTmp + " |
98                 found: " + foundTmp);
99         } catch (IOException | IllegalArgumentException e) {
100         }
101     }
102
103     float cntSuccessRate = ((float) correctCnt) / ((float)
104         totalCnt);
105     System.out.println("Found " + totalCnt + " directories.");
106     System.out.println("EggDetector evaluated " + correctCnt + "
107         directories correctly.");
108     System.out.println(correctCnt + "/" + totalCnt + ": " +
109         (cntSuccessRate * 100) + "% success rate.");
110     System.out.println("Egg count of all folders added together: "
111         + totalEggCount + ".");
112     System.out.println("Distance (|real eggs - found eggs|): " +
113         lengthDifference + " eggs (smaller is better).");
114 }
115
116 private int evaluateDir(File dir) throws IllegalArgumentException
117 {
118     return eggDetector.evaluate(dir).getFinalCount();
119 }
120
121 private int getExpectedCount(File dir) throws IOException {
122     FolderData decodedSettings;
123     try (FileInputStream fis = new
124         FileInputStream(dir.getAbsolutePath() + "/imgdata.xml")) {
125         XMLDecoder decoder = new XMLDecoder(fis);
126         decodedSettings = (FolderData) decoder.readObject();
127         decoder.close();
128         fis.close();
129     } catch (IOException e) {

```

F. TEST EFEKTIVITY KNIHOVNY

```
121         throw e;
122     }
123
124     int eggCount = 0;
125     for (ImgData imgData : decodedSettings.getImgData()) {
126         if (imgData.getEggCount() > eggCount) {
127             eggCount = imgData.getEggCount();
128         }
129     }
130
131     return eggCount;
132 }
133 }
```

Použité programy

TexStudio psaní bakalářské práce v \LaTeX a překlad do PDF

Notepad++ úprava textových souborů

gedit úprava textových souborů

IntelliJ IDEA Ultimate psaní implementace v jazyce Java

GIT verzování kódu

JDK8 kompilace, ladění a testování Java knihovny

Python 3 zpracování dat, trénování neuronových sítí

Tensorflow softwarová knihovna pro strojové učení

LabelImg označování trénovacích dat

javadoc generování dokumentace do HTML

TexDoclet generování dokumentace do \LaTeX

wget hromadné stahování dat

bash pomocné skripty pro zautomatizování práce

awk hromadná oprava drobných chyb v datech

Obsah přiloženého média

bin.....	adresář se spustitelnou formou implementace
doc	adresář s dokumentací implementace
text	text práce
└─ thesis.pdf	text práce ve formátu PDF
src	
└─ impl.....	zdrojové kódy implementace
└─ eggdetector.....	zdrojové kódy implementace knihovny
└─ egg_recognition	implementace rozpoznávání obrazu
└─ object-detection-training	implementace detekce objektů
└─ tools.....	zdrojové kódy pomocných nástrojů
└─ foldertrimmer	zdrojové kódy nástroje FolderTrimmer
└─ tagger.....	zdrojové kódy nástroje Tagger
└─ download.sh	skript pro hromadné stažení dat
└─ libtest	projekt pro testování distribuce knihovny
└─ thesis.....	zdrojové kód textu práce
└─ readme.txt.....	stručný popis obsahu média