

Česká zemědělská univerzita v Praze
Provozně ekonomická fakulta
Katedra informačního inženýrství



Diplomová práce

**Návrh a implementace autonomního stroje k vytěžování
dat ze serveru Ptacionline.cz**

Bc. Tomáš Hanuš

© 2019 ČZU v Praze

ZADÁNÍ DIPLOMOVÉ PRÁCE

Bc. Tomáš Hanuš

Informatika

Název práce

Návrh a implementace autonomního stroje k vytěžování dat ze serveru Ptacionline.cz

Název anglicky

Design and implementation of an autonomous machine for data delivery from Ptacionline.cz

Cíle práce

Cílem práce je připravit učební data pro konvoluční neuronovou síť, která umožní efektivní vyhledávání zájmových objektů (vajec a otevřených zobáčků) na snímcích ze serveru Ptacionline.cz.

Metodika

Postupujte následovně:

- 1) Ze serveru Athena.pef.czu.cz získejte požadovaná data
- 2) Obrazové soubory označte nástrojem LabelImage
- 3) Uložené výstupy z označení (ve formátu XML) uložte do složky Annotations
- 4) Připravená data předejte školiteli (ten z dat naučí vhodnou neuronovou síť)
- 5) Otestuje naučenou neuronovou síť
- 6) Definuje závěry

Doporučený rozsah práce

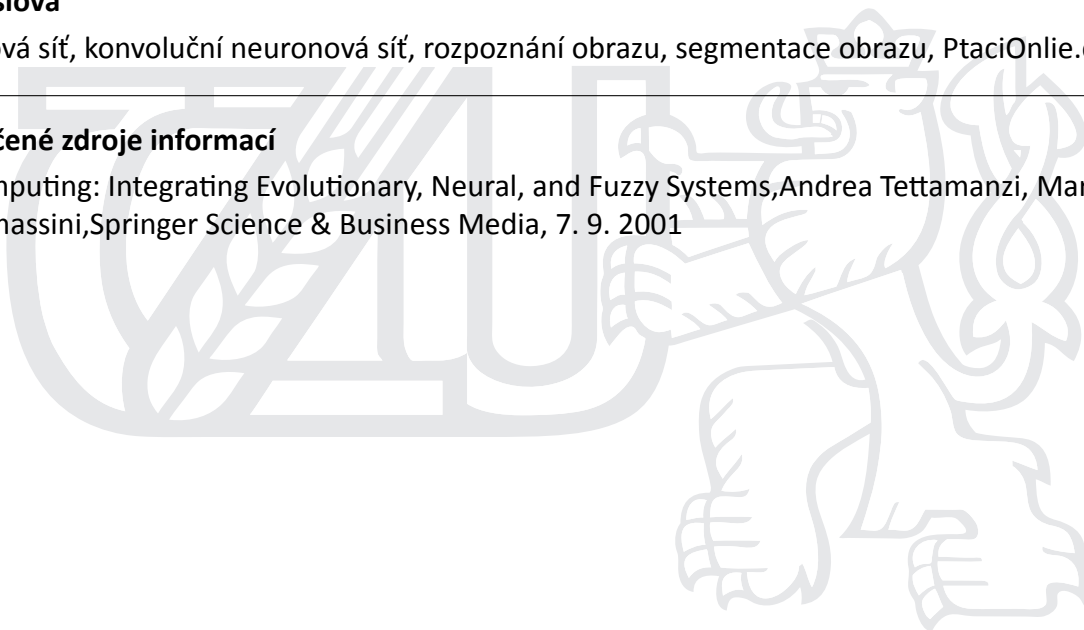
53

Klíčová slova

neuronová síť, konvoluční neuronová síť, rozpoznání obrazu, segmentace obrazu, PtaciOnlie.cz

Doporučené zdroje informací

Soft Computing: Integrating Evolutionary, Neural, and Fuzzy Systems, Andrea Tettamanzi, Marco Tomassini, Springer Science & Business Media, 7. 9. 2001



Předběžný termín obhajoby

2018/19 LS – PEF

Vedoucí práce

Ing. Josef Pavlíček, Ph.D.

Garantující pracoviště

Katedra informačního inženýrství

Elektronicky schváleno dne 21. 3. 2019

Ing. Martin Pelikán, Ph.D.

Vedoucí katedry

Elektronicky schváleno dne 21. 3. 2019

Ing. Martin Pelikán, Ph.D.

Děkan

V Praze dne 21. 03. 2019

Čestné prohlášení

Prohlašuji, že svou diplomovou práci "**Návrh a implementace autonomního stroje k vytěžování dat ze serveru Ptacionline.cz**" jsem vypracoval samostatně pod vedením vedoucího diplomové práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou citovány v práci a uvedeny v seznamu použitých zdrojů na konci práce. Jako autor uvedené diplomové práce dále prohlašuji, že jsem v souvislosti s jejím vytvořením neporušil autorská práva třetích osob.

V Praze dne 29.3.2019

Poděkování

Rád bych touto cestou poděkoval vedoucímu této diplomové práce panu Ing. Josefu Pavlíčkovi Ph.D., za ochotu a trpělivost při její tvorbě. Jsem také velmi vděčný, že jsem mohl být součástí projektu, zabývajícího se využitím umělé inteligence v praxi. Dále bych chtěl vyjádřit vděčnost svým nejbližším za jejich podporu během celého mého studia

Návrh a implementace autonomního stroje k vytěžování dat ze serveru Ptacionline.cz

Abstrakt

Diplomová práce se zabývá přípravou učební sady a jejím následným zpracováním prostřednictvím konvoluční neuronové sítě. Teoretická část práce nabízí potřebné znalosti z několika různých oblastí, které jsou s tématem autonomního rozpoznávání obrazu úzce spjata. Úvodní část teoretických poznatků je věnována základům z oblasti počítačové grafiky. Následuje kapitola pojednávající o problematice strojového vidění, zpracování a segmentace obrazu. Teoretickou část práce uzavírá kapitola o neuronových sítích. Podrobněji se zde práce věnuje konvolučním neuronovým sítím.

Praktická část spočívá v popisu procesu realizace učební sady pro automatickou detekci nakladených vajíček a otevřených zobáčků mláďat v hnízdech. Je zde popsán použitý software a také konkrétní použitý typ konvoluční neuronové sítě. Podrobně je zde popsán celý proces tvorby učební sady, od získání digitálních snímků, přes označování zájmových objektů, až po jejich finální zpracování neuronovou sítí. Následující kapitolu pak tvoří testování detekčních schopností vzniklých modelů, reflektujících kvalitu vytvořených učebních souborů. Na základě testování byly původní učební sady přepracovány a znovu aplikovány. Závěr vlastní práce tvoří porovnání výsledků původních a nově vzniklých modelů.

Klíčová slova: neuronová síť, konvoluční neuronová síť, rozpoznání obrazu, segmentace obrazu, PtaciOnline.cz

Design and implementation of an autonomous machine for data delivery from Ptacionline.cz

Abstract

The diploma thesis deals with the preparation of the learning set and its following processing by a convolutional neural network. The theoretical part of the thesis offers the necessary knowledge from several different fields, which are closely connected with the topic of autonomous image recognition. The introductory part of the theoretical knowledge is devoted to the basics of computer graphics. The following chapter is dealing with the topic of machine vision, image processing and segmentation. The theoretical part of the thesis is being closed by the chapter about neural networks. Special attention is devoted to convolutional neural networks.

The practical part consists of a description of the process of realization of a training set for automatic detection of bird eggs and open beaks in nests. It contains the informations about software necessary for realization and also the description of a specific type of convolutional neural network. In depth, there is described the whole process of creation of a learning set form the acquisition of digital images, through labeling objects of interest, to final processing by a neural network. The next chapter is where testing of detection ability of created models takes place. Obtained results are reflecting the quality of learning sets. Based on these results, the original learning sets were reworked and re-applied. The conclusion of practical part of the diploma thesis contains comparison of the original and the new models.

Keywords: neural network, convolutional neural network, image recognition, image segmentation, Ptacionline.cz

Obsah

| | |
|---|-----------|
| Úvod | 13 |
| 1.1 Cíl práce | 15 |
| 1.2 Metodika | 15 |
| 2 Teoretická východiska | 16 |
| 2.1 Počítačová grafika | 16 |
| 2.1.1 Rastrová grafika | 17 |
| 2.1.2 Vektorová grafika | 21 |
| 2.1.3 Shrnutí kapitoly | 23 |
| 2.2 Zpracování obrazu | 24 |
| 2.2.1 Image pre-processing | 24 |
| 2.2.2 Segmentace obrazu | 24 |
| Segmentace prahováním | 25 |
| Detekce hran (edge-based metody) | 26 |
| Metody hledání oblasti (region-based methods) | 27 |
| 2.2.3 Popis segmentovaných objektů | 28 |
| 2.2.4 Shrnutí kapitoly | 28 |
| 2.3 Neuronové sítě | 29 |
| 2.3.1 Biologický neuron | 29 |
| 2.3.2 Formální neuron | 31 |
| 2.3.3 Vícevrstvé neuronové sítě | 32 |
| 2.3.4 Způsoby učení neuronových sítí | 35 |
| Učení bez učitele | 35 |
| Učení s učitelem | 36 |
| 2.3.5 Konvoluční neuronové sítě | 37 |
| Vstupní vrstva | 38 |
| Konvoluční vrstva | 38 |
| Aktivační vrstva | 39 |
| Pooling vrstva | 40 |
| Plně propojená vrstva | 40 |
| Výstupní vrstva | 41 |
| Učení konvolučních sítí | 41 |
| 2.3.6 Shrnutí kapitoly | 41 |

| | | |
|----------|--|-----------|
| 3 | Vlastní práce | 43 |
| 3.1 | Princip a postup tvorby datové základny | 43 |
| 3.1.1 | Projekt Ptácionline | 44 |
| 3.1.2 | Příprava dat | 45 |
| 3.1.3 | Labeling | 46 |
| | Software LabelImg..... | 46 |
| 3.1.4 | Konvoluční neuronová síť | 50 |
| | Model YOLO | 50 |
| 3.1.5 | Proces rozpoznání objektů z obrazu..... | 51 |
| | Stream cutting | 51 |
| | Image restoration..... | 52 |
| | Image segmentation | 52 |
| | Detekce objektů..... | 53 |
| 4 | Výsledky a diskuse | 55 |
| 4.1 | Analýza výstupů | 55 |
| | První model pro rozpoznání otevřených zobáčků..... | 55 |
| | První model pro rozpoznání vajíček | 56 |
| 4.1.1 | Návrhy na vylepšení učebních sad..... | 58 |
| 4.2 | Realizace úprav | 59 |
| 4.2.1 | Analýza výstupů..... | 60 |
| | Nový model pro rozpoznání vajíček | 61 |
| | Nový model pro rozpoznání otevřených zobáčků..... | 63 |
| 4.2.2 | Porovnání výsledků obou modelů..... | 63 |
| 5 | Závěr..... | 68 |
| 6 | Seznam použitých zdrojů..... | 70 |

Seznam obrázků

| | | |
|--|--|----|
| Obrázek 1 - Pohled senzoru fotoaparátu | Obrázek 2 - Upravená finální fotografie .. | 19 |
| Obrázek 3 - Detail Bayerovi matice | | 20 |
| Obrázek 4 - Bézierova křivka | | 22 |
| Obrázek 5 - Segmentace obrazu (v pořadí zleva: vstupní obraz, detekované hrany jednotlivých objektů, segmentovaný obraz) | | 25 |
| Obrázek 6 - Biologický neuron..... | | 30 |
| Obrázek 7 – Formální neuron | | 32 |
| Obrázek 8 - Vícevrstvá dopředná neuronová síť | | 33 |
| Obrázek 9 - Aktivační funkce neuronu..... | | 34 |
| Obrázek 10 - Konvoluční síť | | 37 |
| Obrázek 11 – Konvoluce | | 38 |
| Obrázek 12 - Rectified Linear Unit | | 39 |
| Obrázek 13 – Podvzorkování..... | | 40 |
| Obrázek 14 - Úložiště zdrojových dat členěných po jednotlivých budkách..... | | 45 |
| Obrázek 15 - Spouštění LabelImg | | 47 |
| Obrázek 16 - Snímek před označením | | 47 |
| Obrázek 17 - Snímek po označení objektů | | 48 |
| Obrázek 18 - Výstup ve formátu XML..... | | 49 |
| Obrázek 19 - Určení oblasti výskytu cílových objektů..... | | 52 |
| Obrázek 20 - Příklad výstupu testování modelu na rozpoznávání otevřených zobáčků | | 56 |
| Obrázek 21 - Testování snímku obsažené v učební sadě (původní model)..... | | 57 |
| Obrázek 22 - Příklad výstupu testování modelu na rozpoznávání vajíček | | 57 |
| Obrázek 23 - Seznam nahraných modelů na portálu uspin.cz | | 60 |
| Obrázek 24 - Testování snímku obsažené v učební sadě (nový model)..... | | 61 |
| Obrázek 25 - Výsledek zpracování souboru z množiny testovacích souborů | | 62 |
| Obrázek 26 - Porovnání výsledků z obou modelů pro detekci vajíček | | 64 |
| Obrázek 27 - Porovnání přesnosti určení polohy objektů..... | | 65 |
| Obrázek 28 - Porovnání původního a nového modelu pro rozpoznávání otevřených zobáčků | | 66 |
| Obrázek 29 - Porovnání reakce modelů na obrazovou deformaci..... | | 66 |

1 Úvod

Trendem moderní doby se stává automatizace zdánlivě obyčejných lidských úkonů. Stále více se využívá metod informačních technologií k úspoře lidské práce a také finančních prostředků. V tomto směru se velmi často hovoří o využití tak zvané umělé inteligence, založené na výpočetních algoritmech, které jsou označovány jako neuronové sítě.

Projekt, díky kterému tato diplomová práce mohla vzniknout, se zabývá monitoringem ptačích hnízd. Pomocí kamerových záznamů je pravidelně a podrobně sledován celý vývoj daného druhu ptactva v několika uměle vytvořených ptačích budkách, od naklazení vajec až po opuštění jejich rodného hnízda. Tyto záznamy jsou ve formě fotografií odesílány na server, kde se ukládají a čekají na další zpracování. Fotografie jsou poté použity k vytvoření rozsáhlé učební sady pro konvoluční neuronovou síť vytvořenou za účelem autonomního rozpoznávání jednotlivých objektů zachycených na snímcích. Předmětem diplomové práce jsou konkrétně učební sady pro rozpoznávání ptačích vajíček a otevřených zobáčků.

Výstupem a smyslem celého projektu je tedy automatizovat sledování ptačích druhů s využitím umělé inteligence, aby člověk nemusel vykonávat stereotypní a zdoluhavé činnosti, které jsou časově i finančně velmi náročné. Model neuronové sítě i vytvořené učební sady by mohly být využitelné pro široké spektrum ptačích druhů, což autor práce vnímá jako významnou přidanou hodnotu celého projektu.

První část práce je zaměřena na základní druhy počítačové grafiky a jejich rozdílného praktického využití. Tato kapitola slouží jako základní stavební kámen technologií určených pro rozpoznávání objektů z digitálních obrázků.

V další fázi práce bude zpracována kapitola s názvem segmentace obrazu, kde budou sepsány postupy a metody využívané k identifikaci jednotlivých objektů na obrázku. Budou zde stručně představeny metody identifikace jednotlivých objektů, techniky jejich rozeznávání a jejich deskriptivní metody.

Teoretická východiska uzavře kapitola týkající se neuronových sítí, kde bude možné nalézt potřebná teoretická východiska nezbytná k realizaci vlastní práce. Zaměřena bude především na konvoluční neuronové sítě využívané v rámci následné vlastní práce.

V praktické části diplomové práce bude podrobně popsán postup realizace celého projektu od počátku tvorby učebních sad, přes jejich zpracování neuronovou sítí,

až po analýzu generovaných výstupů. Z konečných výsledků pak bude vyhodnocena správnost připravených podkladů a jejich případná úprava pro opětovné zpracování. V poslední části budou analyzovány získané výsledky a na jejich základě bude formulován závěr celé diplomové práce.

Konečným výsledkem praktické části by tedy měl být funkční model umělé neuronové sítě, schopný na fotografiích efektivně detekovat otevřené ptačí zobáčky a vajíčka umístěna v hnízdech.

1.1 Cíl práce

Cílem této diplomové práce je vytvoření kvalitních učebních souborů určených pro zpracování pomocí konvoluční neuronové sítě. Detekční modely vzniklé na jejich základě mají být efektivně použitelné pro autonomní rozpoznávání zájmových objektů z digitálních obrázků. Konkrétně se jedná o modely zaměřené na identifikaci ptačích vajec a otevřených zobáčků mláďat v hnízdech. Díky správně připravené datové základně bude pak pomocí umělé inteligence možné spolehlivě analyzovat rozsáhlé množství snímků ve zlomku času, který by na takovou analýzu musel vynaložit člověk. Mezi další cíle této práce se řadí získání potřebných teoretických znalostí v oblasti neuronových sítí, počítačové grafiky, segmentace obrazu nebo přípravy učební sady, nezbytných pro realizaci vlastní práce.

1.2 Metodika

Metodika teoretické části diplomové práce bude založena na studiu odborné literatury a jiných informačních zdrojů v oblasti počítačové grafiky, zpracování i segmentace obrazu, konvolučních neuronových sítí a jejich praktického využití.

Praktická část práce vychází ze získání dostatečného množství kvalitních dat (fotografií ptačích hnízd) ze školního serveru Athena.pef.czu.cz. Tato data budou dále upravena dle požadovaných kritérií a následně zpracována pomocí označovacího nástroje LabelImage. Výstupem pak bude kompletní učební sada skládající se ze zdrojových obrázků (uložených ve složce „Train“) a souborů ve formátu XML (uložených ve složce „Annotations“). Učební sada je dále předána vedoucímu diplomové práce, který provede její zpracování pomocí neuronové sítě. Výsledek a kvalita učební sady je následně autorem práce otestována a vyhodnocena pomocí serveru Uspin.cz. V poslední řadě budou definovány závěry a vyhodnocení celé práce.

2 Teoretická východiska

2.1 Počítačová grafika

Úvodní kapitola teoretických východisek této diplomové práce bude zaměřena na počítačovou grafiku neboli strojovou vizualizaci dat.

Pojem počítačová grafika sahá přibližně do poloviny dvacátého století. Od této doby vizualizace dat mnohokrát změnila svůj význam. Vývoj v oblasti informačních technologií ve všech svých aspektech byl a stále je velmi intenzivní. Neustále se vyvíjejí nové technologie výroby hardwaru, zdokonalují se logické postupy v oblasti tvorby programů nebo aplikací, posouvají se hranice rychlosti i kvality zpracování dat a neustálý posun kupředu je znát i v oblasti technologie interpretace vizualizovaných dat. Historie počítačové grafiky a vizualizace dat je velmi rozsáhlá a plná zlomových milníků, které ale v této práci nebudou podrobně dále rozebírány.

Počítačová grafika patří v současnosti k nepostradatelné součásti našeho života. Vnímáme ji neustále kolem sebe a setkáváme se s ní při běžných každodenních činnostech. Například při sledování televize, na reklamních bannerech okolo silnic nebo na etiketách potravin.

Obor počítačové grafiky je velmi komplexní a nabízí širokou škálu využití. V oblasti marketingu se grafika využívá k úpravě pořízených snímků a videí. Dále se animace obrazu uplatňuje v oblasti filmové produkce, tvorby počítačových her nebo projektování a zpracování nákresových plánů před realizací stavebních prací ve speciálních programech (autoCAD). Nejdůležitější oblastí využití počítačové grafiky pro tuto práci je její využití v rámci umělé inteligence. Konkrétně automatizace identifikování objektů z digitálních snímků.

V této kapitole se bude práce věnovat dvěma hlavním typům počítačové grafiky, které jsou v dnešní době využívány. Rastrové a vektorové. Budou definovány jejich základní principy zobrazení dat, používané formáty dat a také jednotlivé možnosti jejich uplatnění v praxi.

2.1.1 Rastrová grafika

Pojem rastrová nebo také bitmapová grafika představuje typ vizualizace dat, ve kterém je samotný obraz rozdělen do jednotlivých bodů (pixelů) uspořádaných do pravidelné mřížky, který se nazývá rastr. Každý bod v mřížce nese informaci reprezentující jeho pozici ve dvourozměrné soustavě. Má tedy souřadnice „x“ a „y“, které jasně definují jeho umístění v rastru. Každý pixel v sobě také skrývá informaci o jeho zbarvení[1].

Výsledný obraz, který se uživateli zobrazí, je tedy výsledkem rozložení jednotlivých pixelů nesoucích různé informace do dvourozměrného rastru. Celková velikost rastru a počet pixelů v něm je určena rozlišením daného obrazu. Čím více pixelů v rastru, tím čistší a kvalitnější zobrazení.

Parametry bitmapové grafiky:

Rozlišení – Označuje počet pixelů použitých k zobrazení dané informace. Ovlivňuje kvalitu zobrazení detailů. Udává se obvykle v jednotce označené jako DPI (Dot Per Inches) vyjadřující počet použitých pixelů na jednotku délky (palec)[2].

Barevná hloubka – Pojem barevná hloubka označuje v bitmapové grafice věrohodnost barevného podání jednoho pixelu. Jinými slovy určuje počet bitů sloužících k uchování informace o barevném provedení daného bodu[1].

- 8 bitů = 256 odstínů
- 16 bitů = 65 536 odstínů
- 24 bitů = 16,7 milionů odstínů

Výhody rastrové grafiky:

Hlavní předností tohoto typu počítačové grafiky je bezesporu věrohodné napodobení reality. V závislosti na rozlišení obrazu lze metodou rastrové grafiky zobrazovat reálné objekty na opravdu vysoké úrovni. Rastrová grafika spoléhá na nedokonalost lidského oka, které při dostatečně vysokém rozlišení projekce a vzdálenosti od obrazu nedokáže rozlišit

jednotlivé pixely v rastru. Výsledkem je pak pro člověka zdánlivě dokonalý obraz věrohodně reprezentující realitu.

Hodnoty jednotlivých pixelů v bitmapě je možné velmi snadno editovat za pomoci širokého spektra aplikací a nástrojů. Například mobilní aplikace některých sociálních sítí poskytují možnost nastavení „filtrů“ obrazu aplikovatelných na již pořízené digitální snímky. V důsledku použití těchto filtrů dojde k editaci hodnot pixelů v bitmapě, tedy ke změně barevného provedení daného snímku. Velmi často se využívá černobílého filtru, simulace negativu nebo filtrů zvýrazňujících barevné kontrasty fotografie.

Jak již vyplývá z odstavce výše, obraz založený na principu rastrového zobrazení lze získat pomocí digitálního fotoaparátu, mobilního telefonu nebo třeba skeneru. Pořízený obraz je možné bez problémů ořezávat. Díky tomu, že každý pixel si uchovává informace o své poloze v rastru nezávisle na svém okolí, nedochází při ořezech ke ztrátě kvality obrazu nebo jejich deformaci[1].

Nevýhody rastrové grafiky:

Mezi nedostatky využití bitmapové grafiky se řadí především omezené možnosti v oblasti úpravy velikosti obrázků. Změna celkové velikosti digitálních snímků má za následek snížení kvality výsledného zobrazení. Při přílišném zvětšení rastrového formátu dochází k roztažení mřížky rastru tak, že lidské oko je schopné vnímat ohraničení jednotlivých pixelů.

Při změně poměru stran obrazu dochází k deformaci rastrové mřížky a výsledný obraz pak působí rozmazaným dojmem. Ztrácí tím také svou schopnost věrohodně reprezentovat reálné objekty[1].

Formáty rastrové grafiky:

RAW

Jak již naznačuje název samotný, formát RAW obsahuje hrubá, téměř nijak nezpracovaná data ze senzoru. Je označován za „negativ“ digitální fotografie, a to z toho důvodu, že výsledný obraz se musí nechat vyvolat. Nejedná se o vyvolávání chemickou cestou, jak je tomu zvykem u filmových pásek, ale o výpočet obrazu pomocí počítače. Tento

formát se využívá především v profesionálním fotografování, jelikož RAW formát umožňuje později opravit určité nedokonalosti fotografie. Výsledným produktem zpracování RAW formátu bývá obvykle u fotografií níže popisovaný JPEG.

V případě standartních digitálních fotoaparátů má RAW přiřazen jen jeden barevný kanál ke každému pixelu (Bayerova matice).

Bayerova matice je základním stavebním kamenem digitálního snímání obrazu. Ve zkratce lze říci, že každý čip fotoaparátu disponuje množstvím fotodiód, které skládají digitální obraz reálného objektu. Tyto fotodiody zachycují intenzitu dopadajícího světla a rozlišují tak světlá a tmavá místa na výsledné fotografii[3].



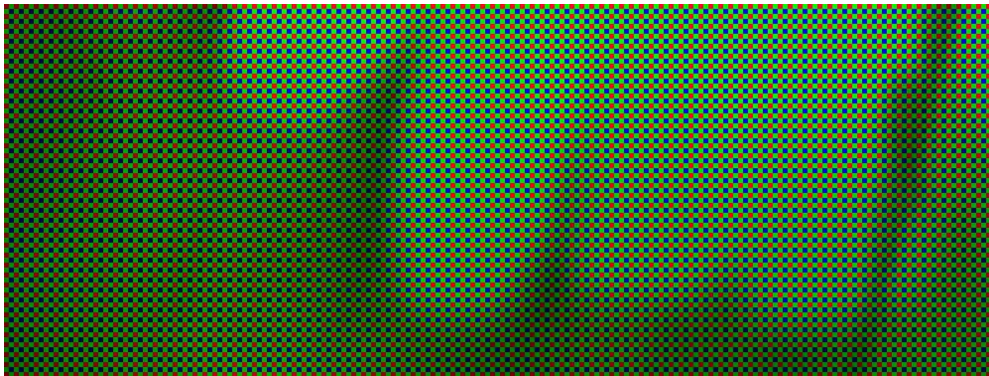
Obrázek 1 - Pohled senzoru fotoaparátu ¹

Obrázek 2 - Upravená finální fotografie ²

Pro barevné fotografování fotodiody využívají svého vlastního barevného filtru (RGB matice). Každý pixel tedy obsahuje pouze jednu barvu (červenou, zelenou nebo modrou) o různé intenzitě.

¹ Zdroj: <https://www.fotoaparar.cz/storage/a/19/1995/914g83vm-6102241.jpg>

² Zdroj: <https://www.fotoaparar.cz/storage/a/19/1995/kp069gn4-6102241-2.jpg>



Obrázek 3 - Detail Bayerovi matice ³

Pro získání výsledného snímku dochází k tak zvané interpolaci, při které dojde ke zprůměrování sousedních pixelů. V případě formátu JPEG má každý samostatný pixel již kompletní barevnou informaci, sestávající se z kombinace číselného vyjádření všech barev z RGB spektra. (viz. Kapitola JPEG).

Převod z formátu RAW do jiných běžných formátů, jako JPEG, lze provádět automaticky už v digitálních fotoaparátech nebo až později v počítačích. Rozdíl je především v kvalitě zpracování výsledného obrazu. Fotoaparáty kladou důraz především na rychlé zpracování fotografií, aby bylo možné pořizovat co nejrychleji další snímky. Výsledky tak bývají z pravidla méně kvalitní. Díky výrazně vyššímu početnímu výkonu počítačů a neomezenému množství času pro zpracování obrázků produkují počítačové sestavy ve všech směrech kvalitnější výstupy. Rozdíly můžeme vidět například v ostrosti snímku, detailech, množství šumu a dalších faktorech[3].

JPEG (The Joint Photographics Experts Group)

V současnosti se jedná o nejvíce používaný formát pro uchovávání digitálních fotografií a jinak pořízených obrázků. Do formátu JPEG se ukládají fotografie pořízené mobilními telefony nebo digitálními fotoaparáty především z důvodu úspory kapacity paměťového úložiště.

³ Zdroj: <https://www.fotoaparát.cz/storage/a/19/1995/b46wt2vz-bayer-small.jpg>

Vzniká ztrátovou kompresí dat z tzv. RAW formátů (viz. výše). Jedná se o grafický formát, který v nejběžnějším kompresním poměru 10:1 dokáže zachovat dostatečné množství informací a udržet tak potřebnou kvalitu obrazu s nižší paměťovou náročností. Využívá maximální velikost obrazu 65 535 x 65 535 pixelů, tedy až čtyři gigapixely při poměru stran obrazu 1:1[4].

JPEG využívá 24-bitovou grafiku (3x8 bitů pro červenou, zelenou a modrou barvu), která mu poskytuje využití množství 16,7 milionů různých barev, ovšem s využitím pouze 8-bitové hloubky jednotlivých barev RGB spektra. Barva jednotlivých pixelů je tedy vyjádřena trojicí čísel s hodnotou z intervalu 0-255, tedy 8 bitů pro určení odstínu každé z barev. Jejich kombinace pak vyjadřuje barevnou hodnotu jednotlivých pixelů. Formát JPEG využívá ztrátové komprese založené na metodě diskretní kosinové transformace (DCT)[5].

PNG (The portable Network Graphics)

Formát vycházející ze svého předchůdce s označením GIF. Jedná se o formát určený pro bezeztrátovou kompresi bitmapové grafiky a nabízí 24-bitovou barevnou hloubku.

Jeho hlavní předností je existence alfa kanálu (tzv. RGBA model), díky kterému je možné definovat transparentnost jednotlivých pixelů. Formát PNG je v rámci práce využíván pro přípravu učební sady pro konvoluční neuronovou síť.

2.1.2 Vektorová grafika

Vektorová grafika vychází ze základů analytické geometrie. Obraz zde není tvořen jednotlivými pixely, ale je reprezentován geometrickými útvary, tedy vektory. Vektor si lze představit jako libovolnou křivku, spojující množství jednotlivých bodů a mající definovanou svoji barevnou výplň. Takové křivky se označují jako Bézierovy křivky.

Princip vektorové grafiky je založen na poznatcích matematika Pierra Béziera a jeho metodě spojování čtyř bodů. Touto metodou lze popsat libovolnou křivku za pomoci dvou kotevních bodů (krajní body) a dvou kontrolních bodů, které určují tvar požadované křivky[7].



Obrázek 4 - Bézierova křivka⁴

Výhody vektorové grafiky

Hlavním rozdílem a také jednou z obrovských předností vektorové grafiky je libovolná manipulace s velikostí daného obrázku. To je možné právě díky tomu, že vektorová grafika je definována matematickým zápisem. Případné zmenšení nebo zvětšení obrazu se tedy provádí pomocí matematických výpočtů. Díky tomu nedochází ke ztrátě kvality obrazu, jako u grafiky rastrové.

Jednou z dalších předností je i možnost práce s jednotlivými objekty samostatně[6].

Nevýhody vektorové grafiky

Nevýhodou vektorové grafiky je složité generování obrazu. Jak je již uvedeno v předchozí kapitole, pořízení obrazu do rastrové grafiky je velmi snadné a rychlé. U vektorové grafiky je to podstatně náročnější ve všech směrech. Jelikož je vektorový obraz tvořen jednotlivými objekty, je nutné tyto objekty vygenerovat a jednoznačně definovat. Více objektů pak znamená vyšší nároky na výpočetní výkon a dobu zpracování[6]. Takto vytvořené soubory jsou pak využívány a zpracovávány speciálními grafickými programy.

Využití vektorové grafiky

⁴ Zdroj: https://upload.wikimedia.org/wikipedia/commons/thumb/8/87/Bezierova_krivka.png/220px-Bezierova_krivka.png

Vektorová grafika má svá specifická využití v oblastech, kde nelze efektivně využít grafiku rastrovou. Především se jedná o tvorbu technických výkresů, 2D animace tvořené pomocí vektorů, 3D modelování a mnoho dalších.

Moderní webové prohlížeče implementovaly podporu vektorové grafiky ve formátu SVG. Díky tomu lze v oblasti webové tvorby využívat vektorovou grafiku v kombinaci s CSS (kaskádové styly) k jednoduchým animacím objektů. Vektorové obrazy lze libovolně bez ztráty kvality zvětšovat a zmenšovat, což je velmi užitečné pro využití na responsivních webových stránkách[6].

2.1.3 Shrnutí kapitoly

V kapitole o počítačové grafice byly definovány základní pojmy z této oblasti. Byly zde představeny dva základní typy počítačové grafiky. Rastrová a vektorová. U obou těchto způsobů reprezentace digitálního obrazu byly sepsány jejich základní vlastnosti a parametry. U rastrové i vektorové grafiky byly stručně představeny jejich hlavní přednosti i nedostatky a také oblast jejich využití. Na grafických ilustracích byly pro představu uvedeny příklady, které doplňují fakta sepsaná v této kapitole. Nalézt zde můžeme i výpis nejvíce používaných formátů v oblasti ukládání digitálních obrázků. Více pozornosti bylo cíleně věnováno formátům rastrové grafiky, jako jsou formáty PNG, JPEG a RAW, které jsou velmi důležité pro rozpoznávání obrazu pomocí neuronových sítí. Obor digitalizace obrazu a jeho zpracování je velmi rozsáhlou oblastí výpočetní techniky, a proto v této kapitole byly sepsány pouze základní poznatky potřebné k realizaci praktické části diplomové práce.

2.2 Zpracování obrazu

Automatické rozpoznávání obrazu je v současnosti jednou z nejdůležitějších a nejvíce využívaných technologií pro řešení úkolů z běžného života. Strojové vidění (Machine vision), jak se toto odvětví často označuje, bylo již velmi dobře zpracováno. I přes to s sebou nese aplikace této metody v praxi nemalé množství technických problémů, které se vztahují k aktuálně vytyčenému cíli a jsou pro dané téma specifické.

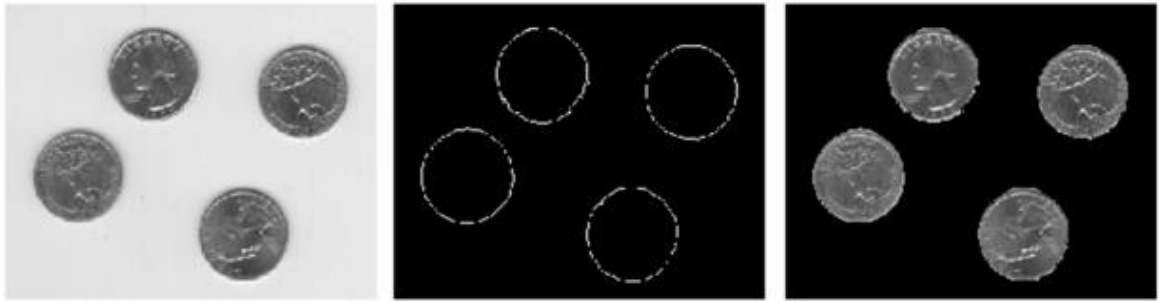
V této kapitole budou sepsány teoretické poznatky z oblasti přípravy a zpracování digitálního obrazu za účelem jejich následného využití.

2.2.1 Image pre-processing

Zpracování obrazu a rozpoznání jednotlivých objektů v nich se skládá z několika fází. V první řadě je nutné požadované snímky pořídit a zdigitalizovat. Následně dochází k tak zvanému pre-processingu digitálních obrázků. Jedná se o metodu, která má za cíl zlepšit kvalitu pořízených digitálních snímků. Soustředí se především na ostření obrazu, vyvážení stupňů šedi, odstranění silného šumu, úpravu jasu, vyvážení kontrastu a zpracování dalších faktorů obrazu díky aplikaci jednotlivých filtrů[8].

2.2.2 Segmentace obrazu

Dalším krokem k úspěšné identifikaci jednotlivých objektů v digitálním obrazu jsou metody jeho segmentace. Ty slouží k rozeznání jednotlivých objektů a jejich odlišení od celkového pozadí obrazu. Na obrázku číslo 5 je možné názorně vidět proces segmentace obrazu od vstupního souboru, až po výsledný snímek s již segmentovanými objekty.



Obrázek 5 - Segmentace obrazu (v pořadí zleva: vstupní obraz, detekované hrany jednotlivých objektů, segmentovaný obraz)⁵

V tomto směru lze efektivně využívat několik různých segmentačních technik. Některé z nich jsou stručně představeny v následujících podkapitolách.

Segmentace prahováním

Základní metodou segmentace objektů v digitalizovaném obrazu je prahování. Vychází se zde z předpokladu, že objekt a jeho pozadí mají rozdílné vlastnosti. Předem je stanovena hraniční hodnota jasu, tedy práh, který slouží jako identifikátor. Porovnávají se světelné hodnoty jednotlivých obrazových bodů zvlášť. V závislosti na tom, zda se hodnota právě zkoumaného bodu vyskytuje pod nebo nad prahovou hodnotou, dochází k vyhodnocení, jestli daný bod je součástí pozadí obrazu nebo hledaného objektu. Hodnotu prahu lze určit několika způsoby:

- Experimentálně
- Z histogramu
- Procentuálně
- Ze statistik
- Z globální znalosti

Existuje více možných způsobů prahování. Primárně se prahování dělí na prosté, s více prahy, adaptivní, částečné a prahování s hysterezí[25].

⁵ http://midas.uamt.feec.vutbr.cz/POV/Lectures/08_Popis_objektu.pdf (strana 5)

Detekce hran (edge-based metody)

Segmentace obrazu vycházející z principu detekce hran objektů tvoří rozsáhlou skupinu segmentačních algoritmů. Jde o segmentační metodu, která se principiálně skládá ze dvou částí. V první z nich dochází k identifikaci hran jednotlivých objektů a v druhé se potom z takto nalezených hran vytváří segmentace samotná. Celý proces je opět znázorněn na obrázku číslo 5.

Přístupů k realizaci detekování hran objektů v obraze je hned několik. Je zřejmé, že každá z nich díky jejich rozdílnostem bude pracovat jiným způsobem a zároveň také poskytovat odlišné výsledky finální segmentace. Nicméně výstupem každé z těchto technik detekcí hran je tak zvaná mapa hran. Tedy mapa, která je následně využita pro matematický výpočet konečné segmentace.

Bavíme-li se o detekci hrany nějakého objektu, je nutné si nejprve definovat, co hrana znamená v rámci digitálního obrazu. Za hranu lze označit určitou oblast daného obrazu, kde lze lokalizovat výraznou změnu intenzity světla. Ačkoliv to může znít jednoduše, reálně je pro autonomní stroj velmi složité jednoznačně určit polohu hrany hledaného objektu.

Velmi zásadní dopad na správné určení hrany objektů má u digitálních obrazů přítomnost šumu. Jeho vliv na výsledek lze minimalizovat v rámci pre-processingu daného snímku.

Významnou roli hraje také šířka oblasti, ve které ke změně intenzity světla dochází. Pokud je zkoumaná oblast příliš široká, lze za hranu považovat buď pouze počáteční bod změny intenzity, oblast jako celek nebo nějaký středový bod dané oblasti.

V případě, že se jedná o hranu mezi dvěma objekty, je segmentace pomocí detekce hran nerealizovatelná. V takovém případě je nutné použít některou z dalších metod, jako je srovnání se vzorem nebo statistické segmentační metody[10].

Srovnávání se vzorem (template matching)

Zde se nacházejí metody, jejichž základ tvoří určitá maska, která se defacto přikládá na právě analyzované místo. Pokud dojde ke shodě, je v daném místě detekována hrana objektu.

Výhodou takto založených metod je především jejich schopnost detekovat hrany i mezi dvěma samostatnými objekty. Další předností je také vysoká odolnost vůči šumu.

Na druhou stranu je nutné říci, že díky soustavnému propočítávání masky ke každému zkoumanému bodu disponují tyto metody obrovskými nároky na výpočetní výkon. Existují sice cesty, jak výpočetní náročnost *template matchingu* minimalizovat, stále však tyto metody zůstávají časově i procesně velmi omezené[26].

Statistické metody

Statistické metody hledání hran jsou velmi podobné výše popsaným předchůdcům. Jejich podstata spočívá v existenci jádra, jehož střed tvoří právě zkoumaný bod obrazu. Jádro má danou svoji hodnotu, která je výsledkem jisté matematické funkce zohledňující všechny body vyskytující se uvnitř vymezeného jádra. V závislosti na právě zkoumaném bodu a výsledku funkce lze potom rozhodnout, zda se v blízkosti zkoumaného bodu vyskytuje hrana[26].

Metody hledání oblasti (region-based methods)

V předchozí kapitole byly popsány segmentační techniky využívající k segmentaci obrazu detekování hran jednotlivých objektů. Jak napovídá již název této kapitoly, bude se zde jednat o techniku založenou na jiném principu. Obecně můžeme tuto skupinu charakterizovat jako metody, zaměřující se na identifikaci a segmentaci objektů jako celků. Pro správné rozpoznání jednotlivých objektů pomocí těchto algoritmů je nezbytné definovat *kritérium homogeneity oblasti*. Toto kritérium může nést různé parametry jako například: úroveň šedi, tvar textury, hodnotu barevného spektra atd.[11]. Na základě těchto kritérií pak dochází k segmentaci jednotlivých homogenních oblastí na snímku.

Výhodou těchto výpočetních algoritmů je primárně několikanásobně vyšší odolnost vůči šumu v obraze.

2.2.3 Popis segmentovaných objektů

Po dokončení celého procesu přípravy a segmentace dat přichází na řadu fáze popisu jednotlivých objektů. Popis jednotlivých objektů spočívá v jednoznačném určení jejich přesné polohy a tvaru. K definici těchto vlastností slouží příznaky (deskriptory, popisovače), které mají jednoznačně vystihovat charakteristické rysy získaných objektů. Definovat jedinečné příznaky, podle kterých by se stroj mohl sám rozhodovat, je velmi složité. Na rozdíl od neuronových sítí lidský mozek nevykonává svá rozhodnutí na základě matematického výpočtu konkrétních hodnot jednotlivých příznaků objektu. Proto by měly takové příznaky splňovat některá základní kritéria jejich užití.

V první řadě by hodnota příznaku u více stejných objektů měla být srovnatelná (Spolehlivost). Naopak, objekty z rozdílných kategorií by měly mít hodnotu příznaku jednoznačně rozdílnou (Diskriminabilita). Další důležité požadavky na deskriptory jsou:

- Invariatnost – imunita příznaku na změnu jasu, kontrastu atd.
- Efektivita – Deskriptor by měl být snadno detekovatelný
- Časová invariatnost – Stabilní hodnota příznaku při zpracování dynamických obrazů

2.2.4 Shrnutí kapitoly

Kapitola nesoucí název „Zpracování obrazu“ obsahovala soupis teoretických poznatků z oblasti možností identifikace jednotlivých objektů ze zdrojových obrázků.

V úvodní části kapitoly byly ve zkratce shrnuty možnosti předzpracování vstupních souborů neboli pre-processingu.

Podrobněji se pak kapitola zabývala tématem segmentace obrazu. Byly zde představeny základní druhy segmentačních technik. Konkrétně metoda segmentace prahováním, detekcí hran a také metoda hledání oblastí. U všech z těchto druhů byl popsán jejich základní princip fungování a jejich výhody nebo možná omezení.

Závěr kapitoly se zabývá způsobem deskripce konkrétních objektů vzniklých pomocí segmentačních algoritmů.

2.3 Neuronové sítě

Termín označuje v moderní době formu umělé inteligence, přesněji řečeno schopnost strojového učení. Jedná se prakticky o vytvoření adaptačního algoritmu, který je schopen napodobovat chování lidského mozku. Je známo, že lidský mozek zjednodušeně pracuje na principu komunikace neuronových buněk pomocí elektrických impulzů. Cílem tvorby modelů neuronových sítí je naučit stroj se samostatně rozhodovat na základě zpracování vstupních informací a dosáhnout tak požadovaných výstupů.

Existuje řada běžně využívaných konstrukčních modelů neuronových sítí s různými oblastmi jejich využití. Díky své schopnosti predikce se algoritmus neuronových sítí využívá například pro předpověď časových řad v ekonomice a finančnictví. Oblastí s využitelností neuronových sítí je díky jejich rychlému vývoji a neustálému zdokonalování algoritmů čím dál více. Sahají například i do automatizace chemických rozborů nebo dopravní signalizace. S tak zvanou umělou inteligencí se dnes člověk setkává pravidelně každý den, aniž by o tom věděl. Stávají se součástí mobilních telefonů, chytrých televizí nebo pomůcek pro handicapované lidi. Pro tuto diplomovou práci jsou podstatné modely neuronových sítí schopných rozpoznávat jednotlivé objekty na pořízených digitálních fotografiích.

V kapitole týkající se neuronových sítí jsou tedy definovány základní pojmy a principy funkce těchto algoritmů.

V úvodní části bude zjednodušeně popsána stavba biologického neuronu, který byl vzorem pro vznik umělých neuronových sítí a jejich využití v oblasti informačních technologií. Na základě získaných znalostí bude popsán princip existence tak zvaného formálního neuronu, který je základní stavební jednotkou všech umělých neuronových sítí.

Dále se bude kapitola zabývat samotným konceptem neuronové sítě. Budou představeny jejich konkrétní typy, modely a topologie. Pozornost bude věnována také možným způsobům strojového učení.

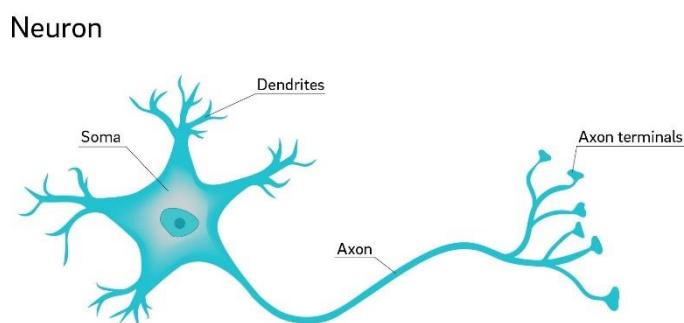
2.3.1 Biologický neuron

Lidský nervový systém je neustále zkoumanou oblastí biologie člověka. Funkce základního stavebního kamene nervové soustavy, neuronu, je již podrobně zmapována a zjednodušeně

formulované poznatky z oboru neurofyziologie jsou dostačující k definici matematického modelu neuronové sítě.

Slovem neuron se označuje nervová buňka, která je základem každé nervové soustavy. Tyto buňky jsou předurčené k uchování, zpracování a přenosu informací nezbytných pro udržení životních funkcí daného organismu. Z pohledu jejich stavby mají neurony tělo (Soma), obsahující jádro nervové buňky, své výstupní (Axony) a vstupní (dendrity) informační kanály. Z jednotlivých Axonů vystupuje množství větví označovaných jako terminály, které se díky koncové bláně (Synapse) spojují se vstupními dendrity jiných neuronových buněk a umožňují tak mezineuronovou komunikaci[12].

Názorná ukázka jednotlivých částí biologického neuronu je zobrazena na obrázku číslo 6.



Obrázek 6 - Biologický neuron

Počet neuronů v nervové soustavě člověka se pohybuje okolo množství 10^{11} až 10^{12} [13]. Během vývoje lidského jedince se počet neuronů v nervové soustavě mění. Zjednodušeně tedy lze shrnout funkci biologických neuronů následovně. Neurony zachycují vstupní informace ve formě signálů prostřednictvím obrovského množství dendritů. Signály dále proniknou do těla buňky, kde na jejich základě dochází k získání potenciálu neuronu. Dosáhne-li potenciál určité hranice (prahu), je neuron schopen vyslat signál dále pomocí axonu a předat zpracovanou informaci dalším neuronům.

Umělé neuronové sítě ve svém historickém vývoji byly silně inspirovány právě biologickou předlohou nervové soustavy živočichů. Postupem času se ale jejich vývoj osamostatnil a přímé vazby na biologii pomalu odpadávaly. Výsledné matematické modely umělých neuronových sítí jsou nyní od biologické předlohy často vzdáleny a jejich funkce byly velmi zjednodušeny.

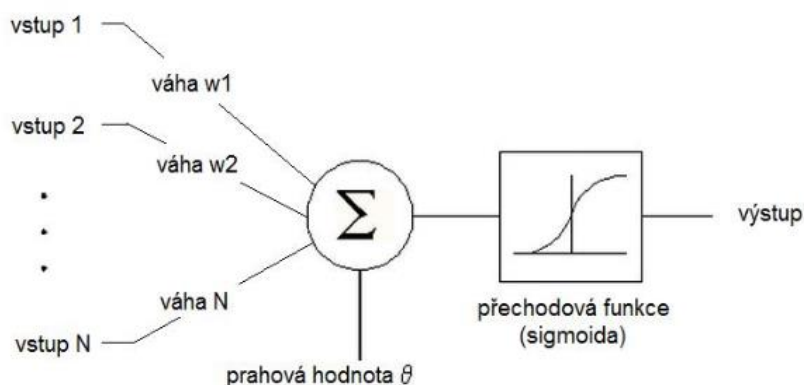
2.3.2 Formální neuron

Model formálního neuronu je postaven na zjednodušených poznacích fungování neuronu biologického. Jak již bylo napsáno v předchozí kapitole, není možné plnohodnotně okopírovat funkcionalitu biologické neuronové sítě a její princip zpracování informací. Vznikl tedy základní matematický model formálního neuronu, který dokáže ten biologický alespoň částečně napodobit. Dokáže tedy přijímat vstupní informace a na základě jejich souhrnného stavu generovat výstupy.

Matematicky je formální neuron definován vztahem:

$$Y = S\left(\sum_{i=1}^N (W_i X_i) + \theta\right)$$

- x_i – Definuje hodnotu i -tého vstupního signálu
- w_i – Definuje hodnotu i -té vstupní váhy
- $S(x)$ – Aktivační funkce neuronu
- n – Celkový počet vstupů
- θ – Práh neuronu
- Y – Výstup neuronu



Obrázek 7 – Formální neuron⁶

Dle obrázku číslo 7 je v první řadě definováno n vstupů neboli vnějších podnětů. Každému vstupu je přiřazena hodnota jeho synaptické váhy. Tělo neuronu pak má definovanou svou prahovou hodnotu neuronu. Pokud dojde k překročení prahové hodnoty neuronu, dojde k indikaci signálu na výstupu daného neuronu ve formě aktivační funkce. Aktivační funkce neuronu může mít formu binární (0/1, TRUE/FALSE) nebo bipolární (-1, 0, 1)[15].

Samostatně jsou tedy vhodné pro jednoduché rozhodovací úlohy. Například identifikace a segmentace světlých a tmavých míst z obrázku.

Formální neurony jsou samostatně schopné řešit jen jednoduché rozhodovací procesy. Až díky jejich společnému propojení informačními vazbami do různých struktur, a především do různých vrstev, vytváří komplexní funkční algoritmy schopné strojového učení.

2.3.3 Vícevrstvé neuronové sítě

V rámci neuronových sítí dochází ke kumulování neuronových buněk do jednotlivých vrstev. Tyto vrstvy se dělí na vstupní, výstupní a skryté, a společně tvoří výsledný útvar označovaný jako umělá neuronová síť.

Vstupní vrstva již dle jejího názvu zajišťuje vstupní informační tok do neuronové sítě. Neurony tvořící vstupní vrstvu mají za cíl pouze předat vstupní informace další vrstvě

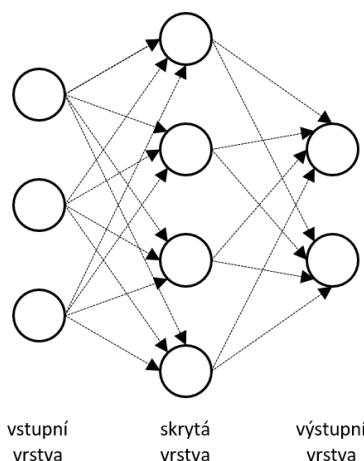
⁶ Zdroj: <https://is.mendelu.cz/eknihovna/opory/download.pl?objekt=21499>

v pořadí. Jejich aktivační funkce je lineární a informace tekoucí skrz neurony vstupní vrstvy jsou konstantní. Jejich interpretační hodnota je tedy neměnná.

Výstupní vrstva neuronové sítě produkuje výsledné hodnoty zpracování vstupních informací a připravuje je k následné interpretaci.

Na rozhraní mezi vstupní a výstupní vrstvou se u vícevrstvých neuronových sítí vyskytuje jedna nebo více vrstev skrytých. Jejich označení se odvozuje od faktu, že vstupy a výstupy těchto vnitřních vrstev nijak neovlivňují vnější okolí celé neuronové sítě.

Výslednou vícevrstvou neuronovou sítí pak lze vyjádřit orientovaným matematickým grafem, ve kterém každý neuron označuje jeden vrchol. Vrcholy jsou uspořádány do jednotlivých vrstev. Každý vrchol z první vrstvy je spojen vazbou s každým z vrcholů vrstvy následující. Vazby mezi vrcholy v rámci jednotlivých vrstev spolu propojeny nejsou (viz. Obrázek č. 8).



Obrázek 8 - Vícevrstvá dopředná neuronová síť⁷

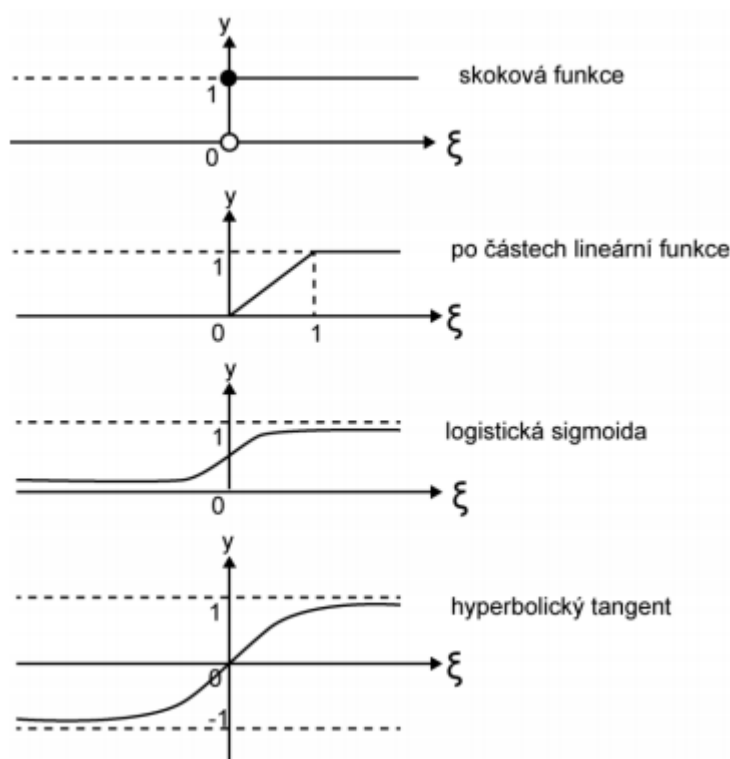
U neuronových sítí obecně tedy dochází k transformaci vstupních signálů za účelem získání určitých výstupních hodnot. Chování adaptačních algoritmů je v základu definováno třemi prvky[14]:

- Učícím pravidlem
- Aktivační funkcí

⁷ Zdroj: <https://www.napocitaci.cz/33/neuronove-site-a-princip-jejich-fungovani-uniqueidgOkE4NvrWuNY54vrLeM670eFNQh552VdDDuIZX7UDBY/>

- Množinou vazeb mezi neurony

Aktivační funkce neuronu mohou mít mnoho různých podob. V důsledku četného testování se nejčastěji využívá funkce jednotkového skoku, funkce signum, lineární nebo sigmoidální funkce. Aktivační nebo také přechodová funkce modeluje výstupní signál nabuzeného neuronu, který je dále směřován do další vrstvy neuronové sítě. Správný výběr aktivační formule má velký vliv na konvergenci výpočtu konkrétního adaptačního algoritmu. Funkce jednotlivých neuronů jsou vybírány tak, aby byla výsledná neuronová síť schopna pokrýt široký rozsah lineárních i nelineárních vztahů dle oblasti jejího využití[14].



Obrázek 9 - Aktivační funkce neuronu⁸

V rámci zkoumání průchodu informací skrz celou síť bohužel není možné analyzovat rozhodovací kroky jednotlivých neuronů uvnitř sítě. Získaným výstupem průchodu vstupních dat neuronovou sítí je tedy pouze výsledek, nikoliv však podrobný výpis kroků k jeho dosažení.

⁸ Zdroj: <http://portal.matematickabiologie.cz/res/f/neuronove-site-jednotlivy-neuron.pdf>

Topologie neuronových sítí se skládá ze dvou základních principů. Jedná se o síť s dopředným šířením signálu a zpětnovazebné síť.

Dopředné neuronové síť (angl. Feedforward neural networks) je označení pro takovou soustavu neuronů, skrze kterou pronikají všechny signály ze vstupní vrstvy až do výstupní vrstvy bez existence zpětných vazeb. Signál je tedy šířen pouze jednosměrně.

U neuronových sítí se zpětnovazebným šířením signálu (angl. Feedback Neural Network) dochází k cyklickému šíření signálu mezi jednotlivými vrstvami neuronů.

Využití neuronových sítí v praxi[24]:

- Analyzování textových znaků
- Převod řeči na text
- Analýza a segmentace obrazu
- Statistické prognózy
- Autonomní mechanické subjekty (automobily, stroje atd.)

Modelů umělých neuronových sítí existuje v současnosti poměrně velké množství. Podrobněji rozebírány v rámci této práce nebudou, ale představit se sluší alespoň tyto jejich zástupce:

- Samoorganizující se neuronové síť
- Hopfieldovy síť
- RBF síť (Radial Basis Function)

2.3.4 Způsoby učení neuronových sítí

Způsoby učení neuronových sítí se rozdělují na dvě základní oblasti. Označují se jako metody učení s učitelem a učení bez učitele.

Učení bez učitele

Při metodě učení bez učitele nemá adaptační algoritmus žádné pevně dané kritérium správnosti transformace vstupních dat. Spočívá ve shlukování sobě podobných elementů

přivedených na vstup a jejich následném rozdělení do příslušných skupin. Kritérium pro třídění vstupních dat není řízeno zásahem zvenku, ale je založeno pouze na informacích vložených na vstup neuronové sítě. Kategorizace vstupů je poté prováděna výhradně v rámci adaptačního algoritmu a jeho schopnost je možné ovlivňovat pouze úpravou synaptických vah jednotlivých neuronů.

Metoda učení bez učitele je využívána především pro dynamické sítě označované jako samoorganizující se mapy (SOM – Self Organizing Map)[16].

Učení s učitelem

Na druhé straně při učení neuronové sítě s učitelem má adaptační algoritmus definovanou konečnou množinu „ M “ dvojic „ x “ a „ y “, které představují konkrétní vstupy a spočtené výstupy zpracovávané úlohy. Tuto množinu lze označit jako příklad správného chování, tedy správné transformace vstupních dat na výstupní.

Množina „ M “ slouží nejen k samotnému učení sítě, ale také k ověření správné funkcionality. Dále se tedy dělí na množinu trénovací a testovací.

Samotné učení neuronů pak probíhá iterativně. Jednotlivé prvky trénovací množiny jsou vkládány na vstup, zjišťuje se reakce neuronu a v závislosti na odchylce výstupu dochází ke korekci jeho synaptických vah. Časový horizont potřebný k předložení všech trénovacích prvků na vstup neuronu se nazývá „*epocha učení*“[16].

Během procesu učení neuronové sítě mohou obecně nastat dva základní nežádoucí faktory. Označují se jako „*přeučení sítě*“ a „*nedoučení sítě*“.

Přeučení neuronové sítě nastává při použití nadměrného počtu neuronů v síti. Neuronová síť se v tomto případě naučí precizně rozpoznávat požadované vzory ze souboru trénovacích dat. Hodnota chyby E je velmi nízká. Avšak při zpracování dat mimo trénovací množinu nabývá chyba astronomických hodnot a neuronová síť tak ztrácí schopnost zobecnění. Řešením problému přeučení sítě může být rozšíření použité učební sady nebo zredukování počtu neuronů ve jedné ze skrytých vrstev[16].

Jako protiklad k předchozímu odstavci existuje stav označovaný jako nedoučení neuronové sítě. Ten z logiky věci nastává při využití velmi nízkého počtu neuronů.

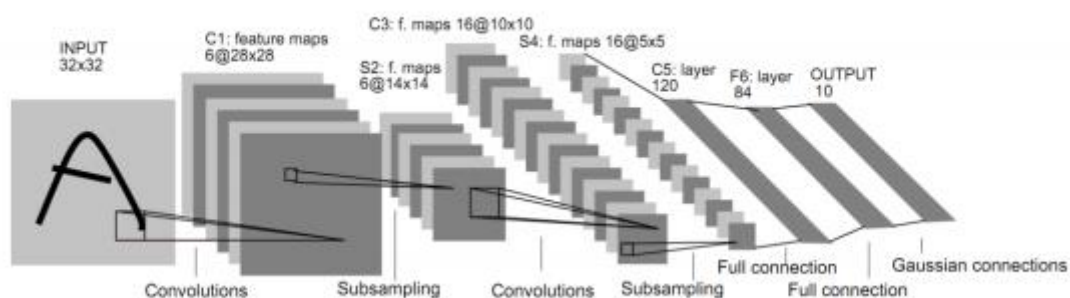
Výsledkem nedoučení sítě je opět vysoká hodnota chyby E . Přidáním neuronů do skryté vrstvy se tento problém minimalizuje. Obecně se při tvorbě neuronových sítí doporučuje začínat od menšího počtu neuronových buněk a jejich postupného přidávání. V závislosti na hodnotě chyby E se neuronová síť postupně rozšiřuje až do chvíle, kdy dosáhne svého optima. Tím lze označit stav, kdy hodnota chyby E je pro daný problém přijatelná.

2.3.5 Konvoluční neuronové síť

Konvoluční neuronové síť lze zařadit do kategorie vícevrstevných dopředných neuronových sítí používaných primárně pro automatizaci identifikace objektů v digitálním obraze za využití minimálního předzpracování. Rozpoznávání probíhá přímo z jednotlivých pixelů.

Princip konvolučních neuronových sítí vychází z modelu Neocognitronu, který byl představen v roce 1987 panem K. Fukushimau. Zjednodušeně řečeno je model Neocognitronu založen na vzájemném propojení velkého počtu neuronových vrstev s různou funkcionalitou[1].

Konvoluční neuronové síť se tedy vyznačují především svou schopností identifikace cílových objektů na základě jejich podobnosti k předloženým vzorům. Model není příliš citlivý na možné deformace, změny pozice nebo změny velikosti obrazu a díky tomu se konvoluční neuronové síť uplatňují velmi efektivně například v identifikaci ručně psaných znaků.



Obrázek 10 - Konvoluční síť

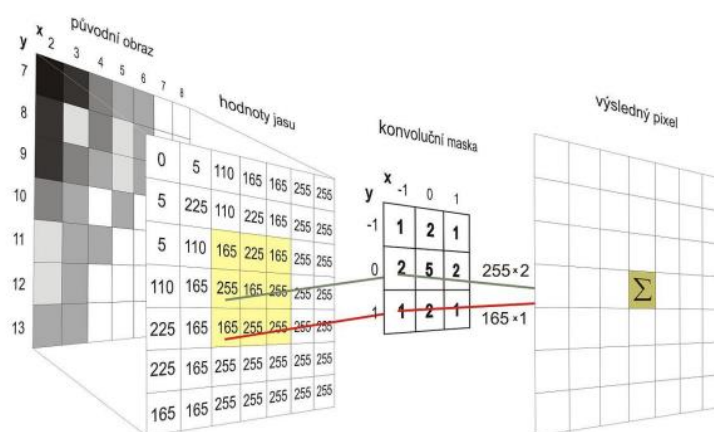
V rámci konvolučních neuronových sítí rozlišujeme několik vrstev, kde každá z nich má svoji jednoznačně definovanou funkci. Blíže jsou představeny v dalších podkapitolách.

Vstupní vrstva

Vstupní data pro konvoluční neuronovou síť tvoří matice hodnot obrazových bodů. Je tvořena hodnotami pro výšku obrázku, šířku obrázku a jeho barevnou hloubku. Pro barevné snímky je hloubka tvořena třemi kanály, tedy hodnotami RGB matice. U černobílých obrázků je kanál pouze jeden[17].

Konvoluční vrstva

K extrahování jednotlivých příznaků z obrázku slouží vrstva konvoluční. V té jsou obsaženy konvoluční filtry. Globálně se hovoří o velmi malých dvourozměrných mřížkách, často o velikosti 3x3 až 6x6 pixelů, s tří kanálovou barevnou hloubkou. V rámci dopředného kroku sítě je realizována operace konvoluce. Zde dochází k aplikaci konvolučních filtrů na každý obrazový bod digitálního snímku. Na obrázku číslo 11 je znázorněn průběh procesu konvoluce a aplikace konvolučního filtru. Výsledkem je hodnota pro jeden výsledný pixel, který je zobecněním vstupních informací.



Obrázek 11 – Konvoluce

Při zpracování digitálních obrázků se používá diskrétní konvoluční model. Matematicky je tento model pro získání výstupních hodnot definován vztahem:

$$f(x, y) * h(x, y) = \sum_{i=-k}^k \sum_{j=-k}^k f(x - i, y - j) * h(i, j)$$

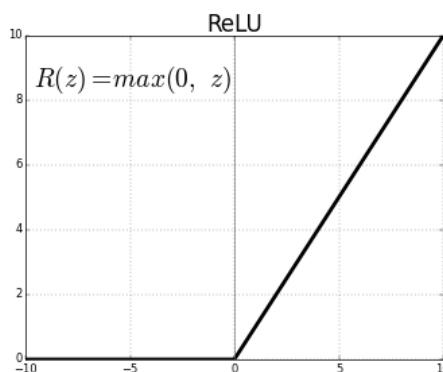
$F(x,y)$ označuje hodnotu jednoho pixelu ze vstupního obrázku f nacházející se na souřadnicích x a y . $H(i,j)$ představuje konkrétní hodnotu váhy konvolučního filtru h na souřadnicích i a j . Velikost konvolučního filtru je pak $k * k$

Neuronová síť se dokáže naučit aplikovat stejné konvoluční filtry pro vstupní snímky se stejnými příznaky. Těmi lze chápat například orientaci hran nebo oblasti snímku se specifickou barevnou hodnotou. V důsledku využití více konvolučních vrstev v rámci jedné neuronové sítě získávají jednotlivé příznaky detailnější charakteristiky[17].

Aktivační vrstva

Vrstva slouží k aplikaci aktivační neboli přechodové funkce na výstup každé vrstvy konvoluční. V některých odborných publikacích se lze setkat s klasifikací aktivační vrstvy jako součástí konvoluční vrstvy.

Konvoluční neuronové sítě využívají jako aktivační funkci Rectified Linear Unit (ReLU). Graficky znázorněna na obrázku číslo 12.



Obrázek 12 - Rectified Linear Unit⁹

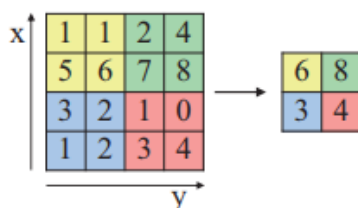
Samozřejmě je možné použít pro konvoluční sítě i jiné druhy aktivačních funkcí. Využití ReLU však dosahuje při trénování sítě několikanásobně vyšší rychlosti[17].

⁹ Zdroj: https://cdn-images-1.medium.com/max/937/1*oePAhrm74RNnNEolprmTaQ.png

Pooling vrstva

Další částí sítě bývá obvykle vrstva označovaná jako *pooling*, volně přeloženo *podvzorkovací* vrstva. Smyslem této vrstvy je redukce dimenze za účelem snížení počtu parametrů sítě. Díky tomu je celkový výpočet jednodušší, rychlejší a méně náročný na výpočetní výkon.

Nejčastěji využívanou metodou poolingů je tzv. *max-pooling*. Metoda spočívající v aplikaci filtru o rozměru $k * k$ na příznakovou mapu o rozměru $n * n$. Z přiloženého filtru je poté zachován pouze prvek s nejvyšší hodnotou. Prakticky znázorněno na obrázku číslo 13.



Obrázek 13 – Podvzorkování

Další velmi často využívanou metodou podvzorkování je využití průměru ze všech hodnot obsažených v matici filtru[22].

Plně propojená vrstva

Plně propojená vrstva (angl. Fully-connected) následuje po sérii konvolučních a podvzorkovacích vrstev. Jak již napovídá název, neurony v této vrstvě jsou plně propojeny se všemi neurony vrstvy předchozí. Využití plně propojených vrstev není nezbytné, nicméně díky nim je síť schopna se snadno naučit i nelineární kombinace získaných příznaků[22].

Výstupní vrstva

Vrstva nacházející se jako poslední v konvolučních sítích je stejně jako předchozí jmenovaná vrstva plně propojená se svým předchůdcem. Počet neuronů ve výstupní vrstvě odpovídá počtu cílových klasifikačních tříd. Při větším počtu klasifikačních tříd než dva, využívá se jako aktivační funkce výstupní vrstvy funkce zvaná *Softmax*. Jejím výstupem jsou hodnoty pravděpodobností příslušnosti k jednotlivých klasifikačním třídám. Součet těchto pravděpodobností je roven 1[22].

Učení konvolučních sítí

Z předchozích poznatků je již známo, že konvoluční sítě pracují s filtry, které se aplikují na vstupní obraz v rámci konvoluční vrstvy. V prvotní fázi učení jsou hodnoty těchto filtrů nastaveny na náhodně vygenerované hodnoty. Optimálním řešením se ukazuje být Gaussovo rozložení s hodnotou směrodatné odchylky $\sqrt{2/n}$, kdy n představuje počet vstupních neuronů v síti.

Konvoluční neuronové sítě jsou schopné si tyto filtry upravit v rámci procesu učení. Ten je založen na principu předkládání velkého množství trénovacích dat, tedy sady obrázků s příloženým popisem zájmových objektů. Postupně dochází ke zpracování celé učební sady tak, že na vstupní vrstvu sítě jsou vkládány vstupní obrázky a na výstupu dochází k vyhodnocení, jestli k obrázkům byly přidány k nim patřící deskriptory.

Dále je nutné nastavit správné váhy na vstupech neuronů (viz. Kapitola 3.3.1.). Učení konvolučních neuronových sítí je náročné na výpočetní výkon a trvá velmi dlouho. Celý proces učení se může v závislosti na rozměru neuronové sítě pohybovat v horizontu dnů, někdy až týdnů. V rámci úspory zdrojů byly vytvořeny před-učené sítě, které se po úpravě parametrů dají využít pro rozpoznávání konkrétních objektů[17].

2.3.6 Shrnutí kapitoly

V celé předchozí kapitole jsou sepsány základní pojmy a principy fungování neuronových sítí. V úvodní části byl představen biologický neuron a na jeho základě byl dále definován

neuron umělý. Zde byl definován matematický model neuronu a jeho základní funkční princip.

Ze získaných poznatků o neuronu vychází další část kapitoly, která se týká vícevrstvých neuronových sítí. Zde byl vysvětlen obecný princip fungování vícevrstvých sítí a jejich jednotlivé části. Pozornost byla také věnována způsobům učení adaptačních algoritmů. Konkrétně učení s učitelem a bez učitele. Definována byla také případná rizika pramenící z učení neuronových sítí.

V další části kapitoly se diplomová práce zabývá konkrétním typem neuronových sítí označovaných jako konvoluční. Konvoluční neuronové sítě jsou hlavním nástrojem pro automatizaci identifikace objektů v obraze. V této kapitole byly tedy rozebrány jednotlivé vrstvy konvoluční sítě a obecný princip jejich fungování.

V poslední řadě je v kapitole popsán způsob učení konvolučních neuronových sítí.

3 Vlastní práce

3.1 Princip a postup tvorby datové základny

V následujících kapitolách diplomové práce na téma „*Návrh a implementace autonomního stroje k vytěžování dat ze serveru Ptacionline.cz*“ bude podrobně zpracován postup realizace praktické části. Nejprve budou představeny základní datové zdroje, které slouží jako stavební kameny celé práce. Těmito zdroji je myšlena především organizace zaštiťující pořizování digitálních záznamů jednotlivých hnízd. Dále také zprostředkovatel datového úložiště, díky kterému je možné tato digitální data skladovat a následně zpracovávat, pomocí technik představených v teoretické části této diplomové práce.

Z množství pořízených snímků budou dále selektovány snímky vhodné pro přípravu kvalitních učebních sad pro neuronovou síť. Konkrétně se tato diplomová práce bude zabývat tvorbou učebního souboru pro rozpoznávání ptačích vajíček a otevřených ptačích zobáčků.

V první fázi bude vybráno přibližně pět set fotografií pro vytvoření každé z cílových učebních sad. Za účelem snadnějšího a časově méně náročného zpracování učebního souboru neuronovou sítí je důležité vybrané snímky zmenšit.

Předmětem další fáze práce bude segmentace vybraných obrázků a označení zájmových objektů pomocí speciálního softwaru, jehož detailnější popis bude k dispozici v jedné z následujících podkapitol. Podrobně bude zpracován i celkový postup označování jednotlivých objektů, společně s praktickými ukázkami vstupních i výstupních souborů a jejich transformace.

Následovat bude předání učebních sad vedoucímu práce Ing. Josefu Pavlíčkovi, Ph.D. za účelem jejich zpracování vytvořenou konvoluční neuronovou sítí. Po dokončení strojového učení bude probíhat testování, zda neuronová síť dokáže správně rozpoznávat objekty i na snímcích, nepocházejících z poskytnuté učební sady. Výsledky testování budou dále analyzovány, aby bylo možné rozhodnout, zda neuronová síť na rozpoznávání objektů na obrazu funguje správně. V případě, kdy chybovost rozpoznání bude příliš vysoká, bude nutné kroky tvorby učební sady opakovat a rozšířit množství použitých snímků k strojovému učení.

V okamžiku, kdy neuronová síť bude schopna efektivně rozpoznávat jednotlivé objekty z obrázků mimo učební sadu bude možné formulovat konečné závěry celého projektu, z něhož tato diplomová práce vychází. Správnost strojového učení bude reprezentováno ilustrační ukázkou.

3.1.1 Projekt Ptácionline

Projekt „Ptáci Online“ (www.birdsonline.cz) na jehož základě bylo možné vytvořit tuto diplomovou práci vznikl v roce 2014 na půdě České zemědělské univerzity v Praze v roce 2014. Od svého vzniku prošel několika vývojovými fázemi, od teoretických podkladů pro návrhy celého monitorovacího systému, přes konstrukci chytrých ptačích budek, po online zpracování a vyhodnocování zajištěných snímků.

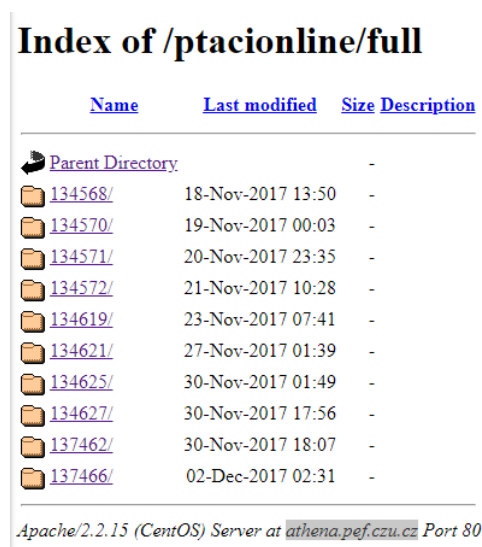
Cílem projektu je tedy realizace chytrých ptačích budek (angl. Smart Nest Box), s jejichž pomocí bude možné sledovat průběh hnízdění jednotlivých ptačích druhů v jednotlivých fázích jejich vývoje. Tedy od tvorby hnízda, přes naklazení vajec a jejich zahřívání, vylíhnutí potomků a obstarávání potravy, až po opuštění rodného hnízda. Záznamy z hnízd jsou nahrávány na speciální kamerové systémy s vlastní úložnou kapacitou a s připojením na internet, díky kterému lze získaná data sledovat v reálném čase a z pohodlí domova.










Chytrá ptačí budka je vyrobena tak, aby mohla obsahovat vestavěný počítač, pohybový senzor ve vletovém otvoru, kameru s nočním viděním, mikrofon, teplotní senzor a senzory venkovního osvětlení. Konstrukce budky je přizpůsobitelná různým ptačím druhům díky variabilitě průměru vletového otvoru. Budky jsou propojeny pomocí ethernetových kabelů do maximální vzdálenosti 80 metrů k přípojce. Napájení elektroniky v budce je realizováno pomocí technologie PoE (Power over Ethernet).

Kromě samotného přínosu a zpracování informací z oblasti ornitologie a biologie má projekt za cíl především přiblížit vědeckou činnost široké veřejnosti. Samotnou realizaci projektu se tvůrci snažili provádět za přispění vybraných studentů vysokých škol, nebo jiných zájmových organizací. Získané výstupy z kamerových záznamů pak mohou sloužit například také jako edukativní pomůcka ve školách[18].

3.1.2 Příprava dat

Snímky pořízené chytrými budkami, vybudovanými v rámci projektu Ptáci online jsou odesílány a ukládány na server *Athena*, který se nachází na půdě České zemědělské univerzity v Praze (dostupný z <http://athena.pef.czu.cz>). Zde byl vytvořen samostatný adresář pro každou z chytrých budek.



| Name | Last modified | Size | Description |
|--|-------------------------------|----------------------|-----------------------------|
|  Parent Directory | | - | |
|  134568/ | 18-Nov-2017 13:50 | - | |
|  134570/ | 19-Nov-2017 00:03 | - | |
|  134571/ | 20-Nov-2017 23:35 | - | |
|  134572/ | 21-Nov-2017 10:28 | - | |
|  134619/ | 23-Nov-2017 07:41 | - | |
|  134621/ | 27-Nov-2017 01:39 | - | |
|  134625/ | 30-Nov-2017 01:49 | - | |
|  134627/ | 30-Nov-2017 17:56 | - | |
|  137462/ | 30-Nov-2017 18:07 | - | |
|  137466/ | 02-Dec-2017 02:31 | - | |

Apache/2.2.15 (CentOS) Server at athena.pef.czu.cz Port 80

Obrázek 14 - Úložiště zdrojových dat členěných po jednotlivých budkách¹⁰

Na obrázku je možné vidět, že v rámci této diplomové práce autor pracoval se záznamy pořízenými celkem deseti různými chytrými ptačími budkami. Z každé z budky jsou pořízeny stovky různých videozáznamů, které jsou opět rozděleny do jednotlivých adresářů. Videozáznamy jsou následně rozstříhány do statických snímků ve formátu PNG o rozlišení 1280*720 pixelů. V konečném součtu jsou tedy díky projektu Ptáci online k dispozici desítky tisíc snímků, které slouží jako podkladová data pro přípravu učební sady ke zpracování konvoluční neuronovou sítí.

Prvním krokem k tvorbě kvalitní učební sady pro adaptační algoritmus je výběr vhodných podkladových dat. Za taková lze označit snímky s minimální přítomností nežádoucích

¹⁰ Zdroj: Vlastní zpracování

faktorů jako je šum nebo rozostření. Především jde o nalezení takových snímků, kde jsou zájmové objekty jasně a zřetelně identifikovatelné.

Optimální počet pro první pokus zpracování neuronovou sítí stanovil vedoucí diplomové práce Ing. Josef Pavlíček Ph.D. společně s autorem práce na 500 snímků od každého ze zájmových objektů. Tedy pro vajíčka i otevřené zobáčky.

Autor práce tedy vybral požadovaný počet obrázků vhodných k tvorbě učebních sad. Následně byly všechny snímky přejmenovány a upraveny na velikost 400*225 pixelů pro rychlejší a méně náročné zpracování konvoluční neuronovou sítí.

3.1.3 Labeling

Nezbytnou částí pro tvorbu učební sady pro neuronovou síť je tak zvaný *labeling*. Jedná se o manuální proces, při kterém dochází k označení jednotlivých cílových objektů na každém snímku zvlášť. Na základě takto definovaných pozic je každému z objektů přiřazen textový popis. Ten si lze představit jako informační mapu pro neuronovou síť určující polohu hledaných objektů. V tomto konkrétním případě byly označovány otevřené zobáčky indikující budoucí příjem potravy a následně také nakladená vajíčka v hnízdech.

Software LabelImg

Pro označování hledaných objektů v rámci této diplomové práce autor pracoval s programem LabelImg. Jedná se o software dostupný zdarma, vytvořený primárně pro Unixové operační systémy. Software byl napsán v programovacím jazyce Python.

Pro použití na operačním systému Windows je nutné instalovat produkt s názvem *Anaconda*. Jedná se o open-source distribuci pro programovací jazyky typu Python využívané pro vědecké účely v oblasti informačních technologií. Využívá se například v oblastech strojového učení, prognostických analýz nebo zpracování objemných dat.

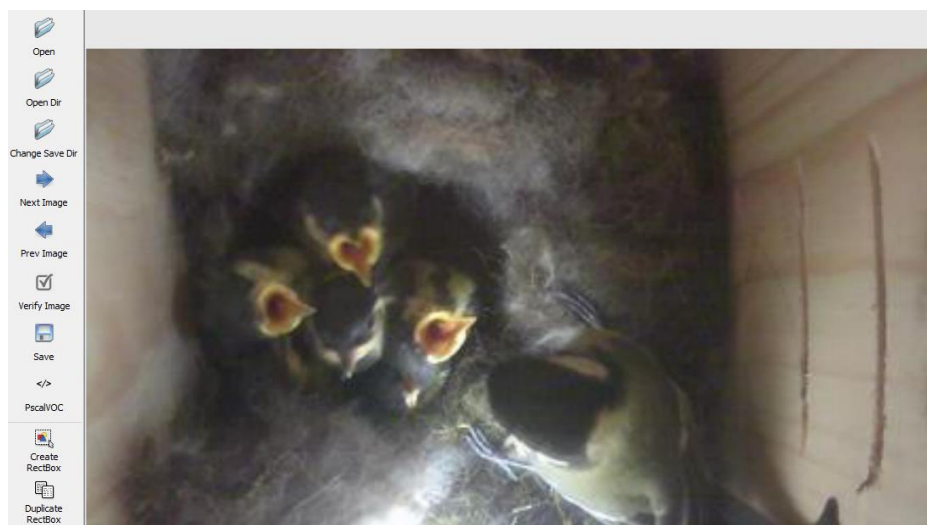
Pomocí Anaconda prompt (speciální příkazové řádky) je pak možné pomocí série jednoduchých příkazů spustit software LabelImg (viz. Obrázek číslo 15).

```
conda install pyqt=5
pyrcc5 -o resources.py resources.qrc
python labelImg.py
python labelImg.py [IMAGE_PATH] [PRE-DEFINED CLASS FILE]
```

Obrázek 15 - Spouštění LabelImg¹¹

Prostředí programu LabelImg disponuje jednoduchým a přehledným grafickým rozhraním, díky kterému je práce s ním velmi příjemná a svižná. Navíc podporuje ukládání výstupních dat ve formátu YOLO. Labeling programů existuje samozřejmě celá řada. Na doporučení vedoucího práce byl jako ideální varianta zvolen právě nástroj Labelimg.

Jak již bylo zmíněno v předchozí kapitole, snímky použité pro labeling byly upraveny do velikosti 400*225 pixelů, ve formátu PNG a se serializovanými názvy. Všechny následně byly uloženy do složky „Train“ se kterou program LabelImg dále pracoval.



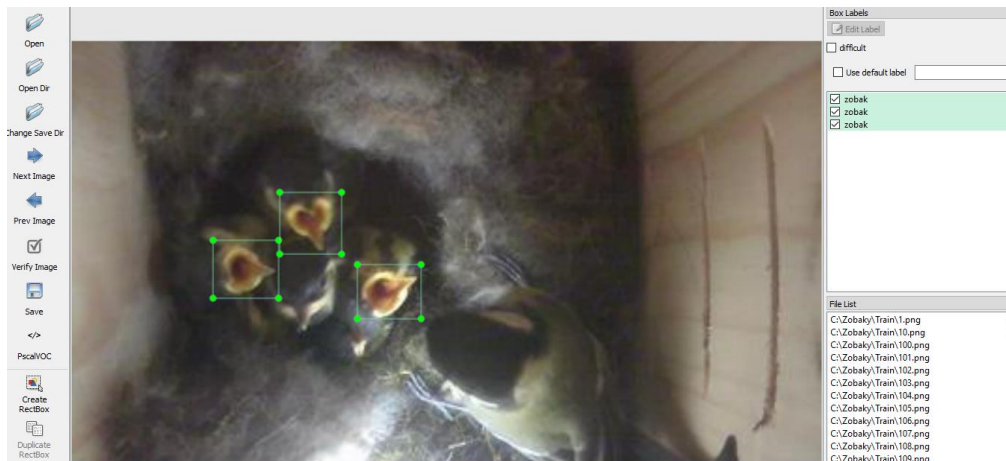
Obrázek 16 - Snímek před označením¹²

Na obrázku je zřejmé, že se v hnízdě nachází celkem šest ptáků. Jeden dospělý jedinec a pět mlád'at. Dle zadání práce jsou pro projekt podstatná pouze mlád'ata s otevřeným zobáčkem,

¹¹ Zdroj: <https://github.com/tzutalin/labelImg>

¹² Zdroj: Vlastní zpracování

reprezentující nadcházející proces krmení. Po zpracování tedy bude výsledný snímek vypadat jako na obrázku číslo 16.



Obrázek 17 - Snímek po označení objektů¹³

Vzniklé čtverce označují oblast výskytu cílových objektů, v tomto případě otevřených zobáčků. V pravé části programu je možné sledovat počet a textový popis jednotlivých objektů na snímku. U využití komplexních konvolučních neuronových sítí může být na obrázku označeno i více druhů objektů na jednou. To samozřejmě zvyšuje složitost a výpočetní náročnost celého modelu. I proto se v rámci tohoto projektu vytváří samostatný model pro každý z objektů zvlášť.

Výsledkem celého procesu označování je textový soubor ve formátu XML, který je po uložení olabelovaného snímku vygenerován a uložen pod stejným názvem do složky „Annotations“. Ten společně se zdrojovým obrázkem tvoří dvojici, díky které stroj dokáže jasně identifikovat polohu cílového objektu na snímku, a na základě toho se naučit rozpoznávat jeho vzhled.

¹³ Zdroj: Vlastní zpracování


```

<?xml version="1.0"?>
- <annotation>
  <folder>Train</folder>
  <filename>118.png</filename>
  <path>C:\Zobaky\Train\118.png</path>
  - <source>
    <database>Unknown</database>
  </source>
  - <size>
    <width>0</width>
    <height>0</height>
    <depth>3</depth>
  </size>
  <segmented>0</segmented>
  - <object>
    <name>zobak</name>
    <pose>Unspecified</pose>
    <truncated>0</truncated>
    <difficult>0</difficult>
    - <bndbox>
      <xmin>76</xmin>
      <ymin>106</ymin>
      <xmax>112</xmax>
      <ymin>139</ymin>
    </bndbox>
  </object>
+ <object>
- <object>
  <name>zobak</name>
  <pose>Unspecified</pose>
  <truncated>0</truncated>
  <difficult>0</difficult>
  - <bndbox>
    <xmin>150</xmin>
    <ymin>118</ymin>
    <xmax>187</xmax>
    <ymin>151</ymin>
  </bndbox>
</object>
</annotation>

```

Obrázek 18 - Výstup ve formátu XML¹⁴

Na obrázku výše je možné vidět výsledný popis objektů ve formátu XML. Úvodní hlavičku tvoří popis snímku, ve kterém se označování provádělo. Uveden je zde zdrojový adresář, název souboru, cesta a o pár řádek níže také barevná hloubka.

Na první pohled je zřejmé, že byly identifikovány tři objekty označené jako „zobak“. Každý z nich má přesně definované hranice pomocí souřadnic x a y . Přesněji jejich maximální a minimální hodnoty určující rozměry čtyřúhelníku použitého pro označení objektu.

¹⁴ Zdroj: Vlastní zpracování

Pro první fázi strojového učení bylo tedy autorem práce celkem připraveno a označeno 524 snímků pro otevřené zobáčky a na dalších 508 snímcích byla označena nakladená vajíčka. Tato data byla poté předána vedoucímu práce Ing. Josefu Pavlíčkovi Ph.D. pro jejich zpracování pomocí konvoluční neuronové sítě. Princip fungování takové sítě a zvolený model je popsán v následující kapitole.

3.1.4 Konvoluční neuronová síť

Pro realizaci projektu byla využita konvoluční neuronová síť. V teoretických základech práce již byl zmíněno, že se jedná o druh adaptačních algoritmů využívaných primárně pro identifikaci objektů na digitálních snímcích. Tento druh vícevrstevných neuronových sítí využívá zpravidla metody učení s učitelem. Je tedy v první řadě nutné stroji předložit učební sadu a následně pak na vzorku testovacích dat ověřit, zda učení proběhlo úspěšně a očekávanými výsledky. Úspěchem pak lze nazývat stav, při kterém neuronová síť ve své práci dosahuje přijatelné hladiny chybovosti.

K dokončení učícího procesu bylo v rámci tohoto projektu zapotřebí časového horizontu v řádu desítek hodin na každý z připravovaných modelů.

Model YOLO

Poslední variantou nejpoblárnějšího algoritmu na detekci objektů z obrázků je YOLOv3. Jedná se o adaptační algoritmus s předučenou množinou 80 jedinečných objektů, které je schopen efektivně rozeznávat. Samozřejmě je možné tvořit a implementovat další učební sady a rozšířit tuto množinu o své vlastní zájmové objekty. Algoritmu *YOLO – You Only Look Once* vyniká převážně v rychlosti zpracování vstupních dat. Navíc jeho přesnost dosahuje podobných výsledků jako u ostatních modelů tohoto typu. Jedním z hlavních konkurentů je model *SSD – Single Shot MultiBox*, který má velmi dobré výsledky, co se týče přesnosti.

Model YOLO si tedy lze představit jako nástroj schopný lokalizovat, a hlavně rozpoznat a popsat požadovaný objekt na digitálním snímku.

Princip fungování modelu YOLO spočívá v tom, že každý snímek je zpracován pomocí dopředné neuronové sítě pouze jednou, jak již napovídá celý jeho název. Model SSD pracuje na stejné bázi, ovšem YOLO dosahuje ve svém oboru výrazně vyšší rychlosti se srovnatelnou mírou chybovosti.

V prvním kroku zpracování je každý snímek rozdělen do mřížky o velikosti 13*13 buněk. Celkem tedy vznikne 169 buněk, jejichž rozměr se odvíjí od velikosti vstupního obrázku. V rámci našeho projektu jsou na vstup sítě vkládány snímky o velikosti 416*416 pixelů. Velikost jedné buňky tedy bude 32*32 obrazových bodů. Každé okénko mřížky je pak zodpovědné za predikci počtu boxů označujících jednotlivé objekty na snímku. Pro každý takto vzniklý box neuronová síť také předpovídá hodnotu sebevědomí (angl. Confidence) toho, že daný box opravdu uzavírá jeden konkrétní objekt a také pravděpodobnost, že daný objekt lze popsat konkrétní třídou. Většina takto vzniklých label boxů je později eliminována, aby jeden konkrétní objekt byl označen v ideálním případě jen jedním boxem. Odstranění probíhá v závislosti na hodnotách pravděpodobností a vah. Tato technika je označována jako **non-max suppression** (potlačení všech hodnot kromě maximálních)[23].

3.1.5 Proces rozpoznání objektů z obrazu

V této kapitole bude nastíněn princip zpracování obrazu a detekce objektů pomocí adaptačního algoritmu. K tomuto tématu se ve svém odborném článku vyjádřil i vedoucí této diplomové práce pan Ing. Josef Pavlíček, Ph.D. Jeho práce tedy slouží jako podklad pro tuto kapitolu[8]. Teoretické poznatky budou v rámci této kapitoly aplikovány konkrétně na problematiku této diplomové práce.

Stream cutting

Pojem označující proces, při kterém dochází k rozčlenění videozáznamu získaných kamerami umístěnými v chytrých ptačích budkách. Samozřejmě je možné aplikovat adaptační algoritmu i na celý videozáznam, který je vlastně tvořen posloupností jednotlivých snímků. Avšak z praktických důvodů je mnohem snazší pracovat se statickými snímky.

Image restoration

V tomto kroku dochází k očištění snímků od přirozeného šumu a také ke změně hodnoty jasu, aby byl obraz vhodný pro následný krok zpracování, kterým je segmentace.

Image segmentation

Segmentace obrazu je proces, při kterém dochází k oddělení nežádoucích objektů na obrázku od těch cílových. Podrobně jsou metody segmentace obrazu popsány v kapitole 3.2 *Zpracování obrazu*. Na základě poznatků získaných v již zmíněné publikaci byly definovány základní znalosti o vzhledu cílových objektů.

- Obrázky jsou snímány ze statické kamery ve výšce cca 50 centimetrů
- Cílové objekty se nacházejí zpravidla ve střední části obrázku
- Sledované objekty jsou otevřené zobáčky a vejčička

V našem případě tedy dojde v rámci segmentace nejprve k určení oblasti, ve které se cílové objekty mohou vyskytovat.



Obrázek 19 - Určení oblasti výskytu cílových objektů¹⁵

¹⁵ Zdroj: Vlastní zpracování

V rámci segmentace tedy dochází k oddělení nežádoucích ploch na snímku od oblasti výskytu sledovaných objektů. Z již definovaných základů víme, že vajíčka a následně i mláďata se budou nacházet přibližně uprostřed záběru. Stěny ptačí budky nás tedy v tuto chvíli nezajímají. Zelené okolí samotného hnízda také můžeme ignorovat. Pak zbývá červeně označená oblast, ve které se objekty zájmu budou zcela nepochybně nacházet. Oblast je z hlediska barevného kontrastu snadno identifikovatelná, což je jednou z hlavních podmínek pro efektivní segmentaci obrazu.

K identifikaci oblasti zájmu dochází v rámci neuronové sítě pomocí vytvořeného filtru o velikosti $k \times k$ pixelů. Ten je postupně posouván po celém digitálním snímku a pro každou jeho aplikaci je počítána průměrná barevná hodnota dané oblasti. Průměrná barevná hodnota konkrétní oblasti je poté konfrontována s intervalem hodnot, předem definovaných pro oblast výskytu cílových objektů. Velikost použitého filtru a správné nastavení barevných hodnot má zásadní vliv na přesnost určování oblasti zájmu. Pokud jsou kritéria moc přísná, může docházet k selhání celé identifikační schopnosti použitého algoritmu[8].

Detekce objektů

Posledním krokem při strojovém rozpoznávání objektů z digitálního obrazu je jejich samotná detekce. Zde už je velmi dobře patrný rozdíl mezi fungování lidského mozku a adaptačního algoritmu. Proto je v této fázi nutné brát omezení strojového učení v úvahu a definovat si základní poznatky o cílových objektech. U vajíček i zobáčků jsou základy velmi podobné.

- Místo výskytu objektů je pravděpodobně určeno vyznačenou oblastí zájmu
- Cílové objekty jsou uspořádány do malé skupinky
- Objekty mají stejné zbarvení i velikost
- Objekty se mohou překrývat

Na základě sepsaných poznatků byly nastaveny parametry adaptačního algoritmu, aby bylo dosaženo co nejlepších výsledků.

Konvoluční neuronové sítě pracují na principu učení s učitelem, jak již bylo popsáno v teoretické části práce. Pro adaptační algoritmu byly tedy vytvořeny učební sady pro oba cílové objekty, na jejichž základě by mělo být možné jednotlivé objekty strojově detekovat. Důležitou částí správného rozpoznávání objektů je nastavení hodnoty tzv. tresholdu. Nastavením jeho správné hodnoty lze eliminovat chyby detekce jako například zdvojené identifikace jednoho objektu nebo ignorování objektů s lehkou obrazovou deformací.

4 Výsledky a diskuse

4.1 Analýza výstupů

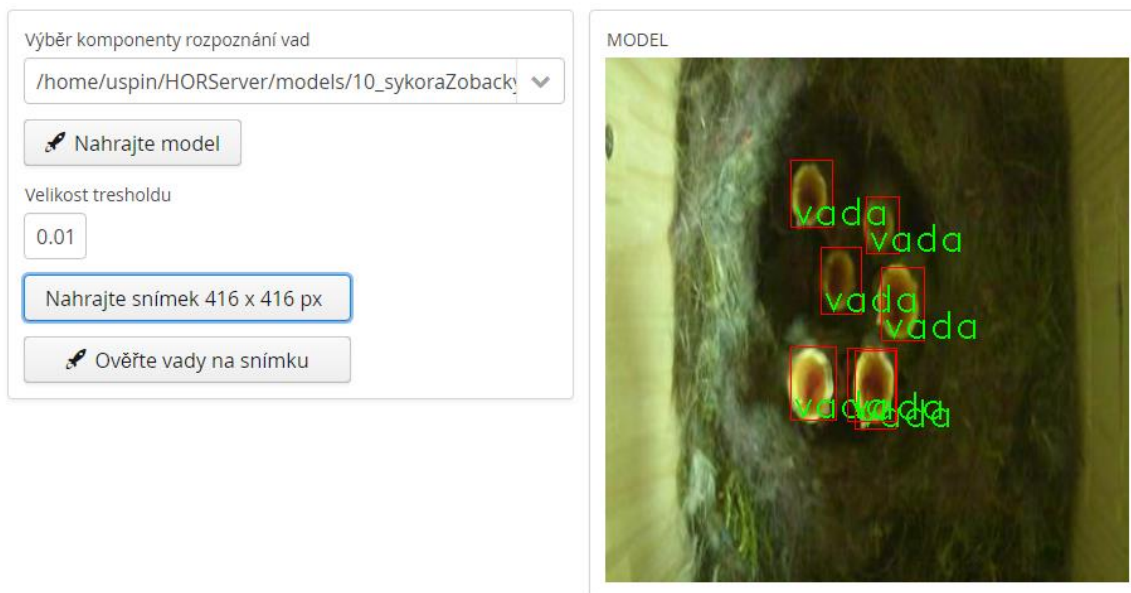
Po dokončení procesu učení je možné přistoupit k testování kvality rozpoznávání jednotlivých objektů pomocí neuronové sítě. Přes portál dostupný na www.uspin.cz je možné v online režimu ověřovat správnost fungování konvoluční neuronové sítě. Aktuálně je zde možné volit mezi čtyřmi funkčními modely. Konkrétně by síť měla být schopna rozpoznávat dospělé jedince, potravu pro mláďata, vajíčka nakladená do hnízda a otevřené zobáčky ptačích potomků. V rámci této diplomové práce jsou předmětem zájmu autora pouze vajíčka a otevřené zobáčky.

Pro ověření správného rozpoznávání objektů je neuronové síti nejprve předkládána množina vzorků, které byly součástí učební sady. Zde by se chybovost rozpoznávání jednotlivých objektů mě pohybovat na velmi nízké úrovni. Pokud neuronová síť nevykazuje známky nesprávného chování, je stroji předkládána tak zvaná testovací množina dat. Jedná se o snímky stejné velikosti a druhu, které ale nebyly součástí inicializační učební sady. Zde už může chybovost rozpoznávání kolísat v závislosti na správném nastavení konvoluční neuronové sítě nebo na rozsahu učební datové sady.

V rámci jednotlivých modelů jsou cílové objekty označeny vždy jako „vada“ obrazu.

První model pro rozpoznání otevřených zobáčků

Pro učební sadu zobáčků autor práce připravil pro první učební proces celkem 524 označených snímků. Po dokončení procesu strojového učení bylo možné poprvé vidět výsledky celé práce a vyhodnotit kvalitu zpracování učební sady. Po předložení množství vzorků obsažených v učební sadě vykazovala neuronová síť očekávanou přesnost identifikace objektů. V některých případech docházelo ke zdvojení jednoho objektu, jak je možné vidět na obrázku níže.



Obrázek 20 - Příklad výstupu testování modelu na rozpoznávání otevřených zobáčků¹⁶

Zde se dalo mluvit o velmi dobrých a přesných výsledcích. Chybovost ovšem výrazně vzrostla s množstvím snímků z množiny testovacích dat. Některé snímky neuronová síť vyhodnocovala správně, u některých ale k identifikaci objektů vůbec nedošlo nebo bylo jejich označení od objektu výrazně vychýleno.

První model pro rozpoznání vajíček

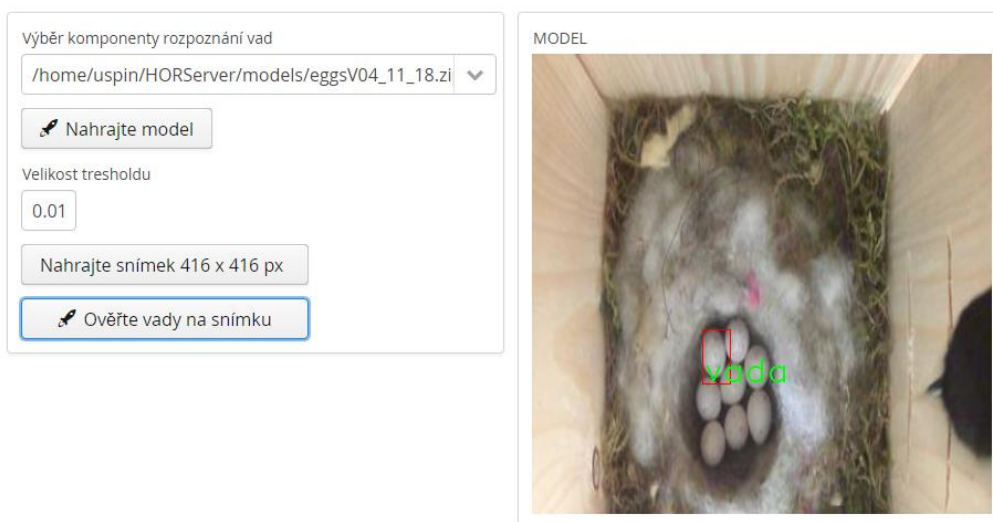
Učební sada pro identifikaci nakladených vajíček čítala 508 olabelovaných snímků. Již při prvním testování byly zjištěny velké chyby v identifikaci vajíček. Ani v rámci trénovací množiny stroj nedokázal efektivně vyhledávat a označovat jednotlivá vajíčka. Pravidelně docházelo k identifikaci pouze jednoho vajíčka, ačkoliv na obrázku se vajec vyskytovalo mnohem více.

¹⁶ Zdroj: Vlastní zpracování



Obrázek 21 - Testování snímku obsažené v učební sadě (původní model)¹⁷

Při klasifikaci výstupů z množiny testovacích souborů docházelo k úplnému selhání identifikace objektů. Vytvořený model tedy rozhodně nelze označit za použitelný pro účel projektu.



Obrázek 22 - Příklad výstupu testování modelu na rozpoznávání vajíček¹⁸

¹⁷ Zdroj: Vlastní zpracování

¹⁸ Zdroj: Vlastní zpracování

4.1.1 Návrhy na vylepšení učebních sad

Na základě zjištěných nedostatků byly vyhodnoceny následné koky, které by měly vést k výraznému zlepšení schopnosti identifikace cílových objektů pomocí neuronové sítě.

V rámci modelu zpracovávajícího ptačí vejce bylo jednoznačné, že konvoluční neuronová síť aktuálně nebyla schopná generovat přijatelné výstupy použitelné pro projekt *Ptacionline.cz*. Problém může být v rozložení jednotlivých objektů, kdy na snímcích pravidelně dochází k jejich překrývání a v důsledku toho je pro stroj obtížné identifikovat jednotlivé hrany. Kontrast mezi vajíčky a jejich okolím také nebývá příliš výrazný, což proces identifikace objektů také znepříjemňuje. Doporučením vedoucího práce Ing. Josefa Pavlíčka, Ph.D. bylo pokusit se model stabilizovat přepracováním prvotní učební sady v oblasti rozsahu a hlavně variability zpracovaných snímků. Pro učební proces tedy vznikla nová sada souborů o rozsahu 826 obrázků.

Model zajišťující identifikaci otevřených zobáčků již od počátku vykazoval výrazně lepší výsledky. Odchyly se vyskytovaly pouze v rámci souborů z testovací množiny a hladina chybovosti nebyla ani tam nijak kritická. Pro dosažení ještě lepších výsledků bylo navrženo rozšíření původní učební sady o dalších 285 vzorků. V rámci optimalizace výsledků mají být do učební sady zapojeny také méně kvalitní snímky s vyšším množstvím šumu nebo s lehkým rozostřením.

4.2 Realizace úprav

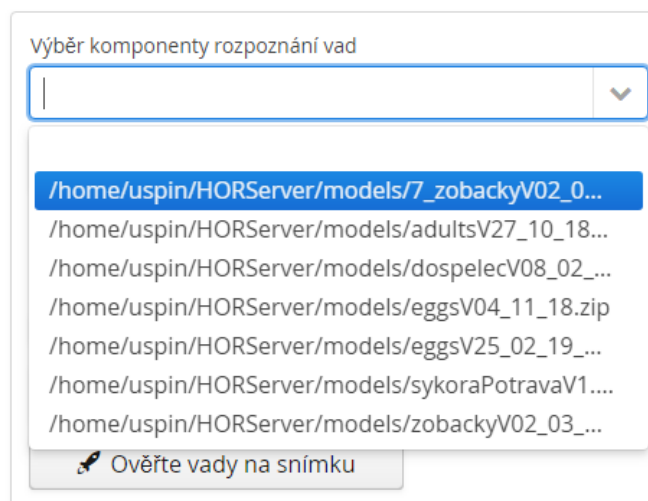
Za účelem zlepšení kvality adaptačního algoritmu a získání relevantních informací z výstupu neuronové sítě došlo k úpravě původních učebních sad. Oba trénovací soubory byly rozšířeny o přibližně tři sta dalších snímků, což by mělo mít pozitivní vliv na přesnost identifikace jednotlivých zájmových objektů.

Učební sada pro rozpoznávání zobáčků byla celá kompletně přepracována, aby lépe vyhovovala požadavkům pro úspěšné strojové učení. Důraz při tvorbě rozšířené sady byl kladen především na variabilitu použitých snímků získaných ze serveru *Athena* na akademické půdě České zemědělské univerzity v Praze.

Pro vyšší efektivitu rozpoznání otevřených zobáčků byly do učební sady zahrnuty také snímky disponující lehkou deformací obrazu.

V následující kapitole bude vyhodnoceno, zda měla tato snaha pozitivní vliv na celkové fungování konvoluční neuronové sítě a jestli výstupní data, která generuje, jsou použitelná pro statistické analýzy projektu *Ptacionline.cz*.

Na portál *uspin.cz* určený k testování kvality výsledků procesu rozpoznávání byly nahrány oba nové modely. Jejich předchůdci byly na serveru zachovány také, aby bylo možné nové výsledky porovnávat s výstupy generovanými původními modely. Všechny aktuálně existující modely nahrané na serveru jsou k vidění na obrázku.



Obrázek 23 - Seznam nahraných modelů na portálu uspin.cz¹⁹

V rámci této diplomové práce autor pracoval pouze s modely typu *eggs* a *zobacky*.

Došlo také k úpravám portálu samotného. Bylo umožněno efektivně pracovat s hodnotou tresholdu adaptačního algoritmu a díky tomu dosahovat přesnějších výsledků. Dále byla na portál uspin.cz přidána možnost automaticky měnit velikost snímku využívaného k testování. Primárním důvodem této funkce je snížení datového přenosu mezi serverem a uživatelem.

V neposlední řadě došlo k úpravě textového označení identifikovaných výsledků. V původním stavu byly objekty označovány jako „vada“, což nepůsobilo příliš dobře na potencionálního uživatele portálu. Nyní při správné identifikaci objektu dojde k označení s názvem „Detected“, což lépe odpovídá účelu, pro který je konvoluční neuronová síť využívána.

4.2.1 Analýza výstupů

V návaznosti na aplikaci nových učebních souborů a jejich zpracování konvoluční neuronovou sítí byly znovu provedeny kroky ke zjištění kvality generovaných výstupů.

Na oba modely byly nejprve aplikovány snímky obsažené v učební sadě, aby byla ověřena základní schopnost strojového rozeznávání cílových objektů. V dalším kroku byla

¹⁹ Zdroj: Vlastní zpracování

aplikována data z množiny testovacích souborů. Díky paralelnímu výskytu původních i rozšířených modelů je možné následně porovnat jejich výsledky a na jejich základě definovat závěry o vylepšení spolehlivosti nově vzniklých modelů.

Nový model pro rozpoznání vajíček

Nová učební sada pro rozeznávání vajíček v ptačích hnízdech obsahovala celkem 826 snímků. Celá sada byla přepracována pomocí programu LabelImg a nový model byl nahrán portál uspin.cz pod názvem „eggsV25_02_19_B30_E_170.zip“.

Při aplikaci souborů z množiny učební sady model vykazoval velmi nízkou chybovost, což oproti původnímu modelu bylo výrazné zlepšení.



Obrázek 24 - Testování snímku obsažené v učební sadě (nový model)²⁰

²⁰ Zdroj: Vlastní zpracování

Na obrázku výše lze vidět, že aktuální model dokáže efektivně rozlišovat jednotlivé objekty (vajíčka) z učební sady velmi přesně a spolehlivě. Tento poznatek je příslibem dramatického zlepšení schopností tohoto modelu pro další testování

Množina testovacích dat čítala 30 snímků. Při jejich výběru byl kladen důraz na jejich různorodost. Cílem testování bylo zjistit chování neuronové sítě při zpracování snímků s různou světelností, hodnotou jasu, množstvím šumu nebo výskytu jiných obrazových deformací. Testování probíhalo po jednotlivých snímcích na portále uspin.cz. Hodnota tresholdu se pohybovala v rozmezí od 0.09 do 0.3. Tato hodnota byla zvolena tak, aby bylo dosaženo objektivních výsledků se správnou vypovídající hodnotou kvality modelu.

MODEL



Obrázek 25 - Výsledek zpracování souboru z množiny testovacích souborů²¹

Na obrázku číslo 25 je vidět výrazné zlepšení modelu ve schopnosti detekce jednotlivých vajíček. Stále jsou viditelné odchylky v určení přesné polohy daného objektu, avšak celkový počet cílových objektů byl stanoven správně. Podrobné srovnání obou modelů určených pro detekci vajíček v ptačích hnízdech je předmětem kapitoly 4.2.2.

²¹ Zdroj: Vlastní zpracování

Nový model pro rozpoznání otevřených zobáčků

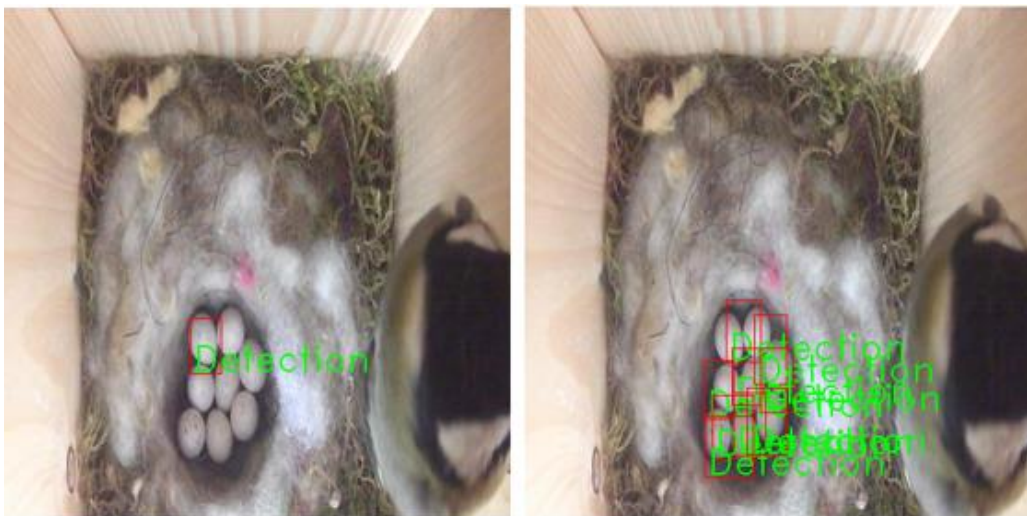
První model pro detekci otevřených zobáčků byl založen na učební sadě o rozsahu 524 označených obrázků. Z důvodu zlepšení detekčních schopností sítě v reálném čase došlo k rozšíření této učební sady o další snímky. Celkový počet digitálních obrázků použitých k tvorbě nového modelu pro rozeznávání otevřených zobáčků se rozrostl na 806.

Původní model byl schopen velmi spolehlivě rozeznávat otevřené zobáčky na vzorcích v rámci učební sady. Tato schopnost byla zachována. Zlepšení se projevilo až na souborech testovacích. Počet identifikovaných objektů z obrázku se prakticky u všech snímků rovnal nebo se alespoň výrazně přiblížil reálnému počtu, který je schopen rozeznat lidský mozek.

4.2.2 Porovnání výsledků obou modelů

Kapitola se zabývá porovnáním výsledků původních a nových modelů pro detekování otevřených zobáčků a vajíček.

Po přepracování obou učebních sad byl u modelů patrný výrazný posun v kvalitě generovaných výstupů. Marginální zlepšení nastalo v detekci ptačích vajíček. Nový model se ukázal být použitelný pro automatizaci získávání aktuálních dat z chytrých ptačích budek a jejich zpracování v rámci projektu Ptacionline.cz.



Obrázek 26 - Porovnání výsledků z obou modelů pro detekci vajíček²²

Výstupy znázorněné na obrázku byly získány pomocí portálu uspin.cz, kdy u obou použitých modelů byla nastavena identická hodnota tresholdu, aby byla zachována vypovídající hodnota tohoto srovnání. Konkrétně byla v tomto případě hodnota tresholdu nastavena na 0.15.

Na obrázku číslo 26 je zřejmé, že původní model (snímek vlevo) dokázal identifikovat na předloženém snímku pouze jeden cílový objekt. Nový model založený na přepracované a rozšířené učební sadě na stejném obrázku dokázal detekovat celkem 8 vajíček, což je reálná hodnota počtu jejich výskytu. Model se stále potýká s nedokonalostí určování přesné polohy jednotlivých vajec, nicméně pro účel analýzy celkového počtu ptačích vajíček v hnízdě není jejich přesná poloha nezbytná. Na druhé straně u některých snímků došlo i v tomto směru k přesnějším výsledkům (viz. Obrázek číslo 27).

²² Zdroj: Vlastní zpracování

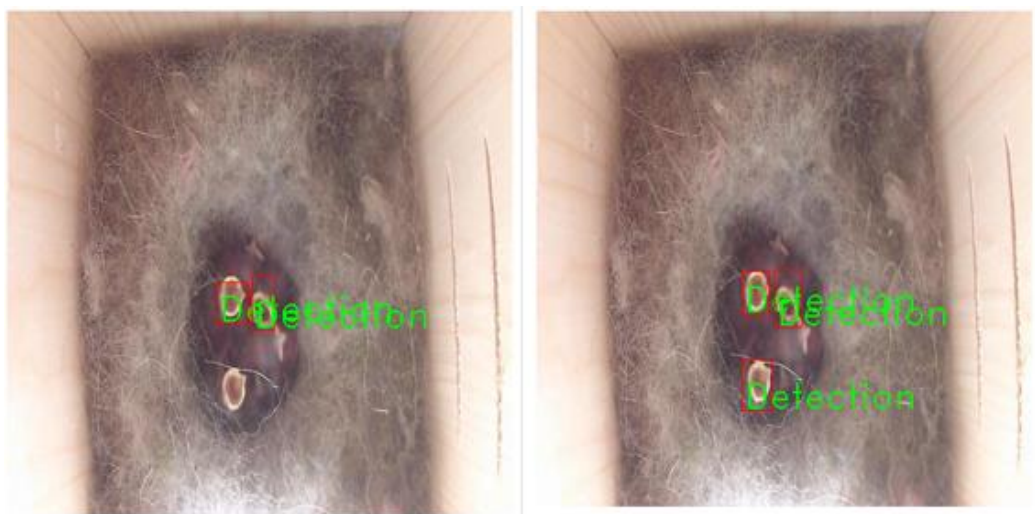


Obrázek 27 - Porovnání přesnosti určení polohy objektů²³

Výsledný model s pracovním označením „eggsV25_02_19_B30_E_170.zip“ nahraný na portálu uspin.cz lze označit jako efektivně využitelný pro monitoring počtu ptačích vajíček v chytrých budkách zkonstruovaných díky projektu Ptacionline.cz. Výstupy generované tímto modelem jsou velmi často blízké realitě a model je tedy vhodný jako podklad pro následné statistické analýzy vývoje ptačích jedinců.

Druhou částí vlastní práce bylo vytvoření funkčního modelu pro detekci otevřených ptačích zobáčků. Již původní model poskytoval obstojné výstupy, nicméně za účelem zvýšení kvality výsledného produktu byla původní množina učebních dat rozšířena. Porovnání obou vzniklých modelů ukazuje, že zvýšení počtu použitých digitálních fotografií a také zakomponování souborů s lehkou obrazovou deformací do učební sady vedlo k výraznému zlepšení chování výsledného adaptačního algoritmu.

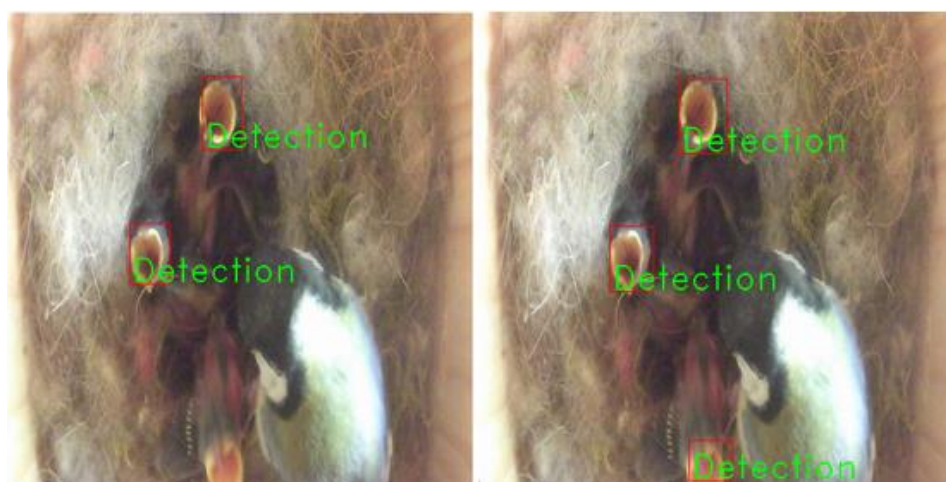
²³ Zdroj: Vlastní zpracování



Obrázek 28 - Porovnání původního a nového modelu pro rozpoznávání otevřených zobáčků²⁴

Opět byla v rámci testování nastavena identická hodnota tresholdu pro oba použité modely. Obrázek číslo 28 krásně ilustruje zlepšení detekčních schopností tohoto modelu. Původní model (na obrázku vlevo) identifikoval pouze 2 otevřené zobáčky. Nový model už byl schopen detekovat správný počet a také je zde možné vidět, že došlo i ke kalibraci polohy jednotlivých boxů označujících cílové objekty.

Reakce na zhoršené podmínky z důvodu obrazových deformací představuje následující obrázek.



Obrázek 29 - Porovnání reakce modelů na obrazovou deformaci²⁵

²⁴ Zdroj: Vlastní zpracování

²⁵ Zdroj: Vlastní zpracování

Na snímku se vyskytují celkem tři otevřené zobáčky. Dva z nich jsou jasně zřetelné a zaostřené a oba modely neměly s jejich detekcí žádné problémy. Třetí objekt byl pro původní model absolutně nerozpoznatelný kvůli mírnému rozostření. Díky zahrnutí i takto deformovaných snímků do nové učební sady vznikl model, který i hůře identifikovatelné objekty dokáže velmi spolehlivě detekovat a označit.

Nově vzniklý model pro rozpoznávání otevřených ptačích zobáčků tedy poskytuje oproti svému předchůdci ještě přesnější výsledky. Jeho nasazení do reálného provozu v rámci projektu Ptacionline.cz by mělo naplnit počáteční očekávání a sloužit tak efektivně a spolehlivě ke sledování počtu živých mláďat v hnízdech a také k detekci blížícího se procesu krmení.

5 Závěr

Diplomová práce si kladla za hlavní cíl vytvoření kvalitního učebního podkladu pro konvoluční neuronovou síť za účelem autonomního strojového rozpoznávání otevřených zobáčků a vajíček uvnitř chytrých ptačích budek pod záštitou projektu Ptacionline.cz.

Samotné praktické části diplomové práce předchází několik kapitol teoretických, které vznikly na základě studia odborných článků a publikací. V úvodu teoretické části se práce zaměřila na základní poznatky z oblasti počítačové grafiky. Konkrétně zde byly zpracovány kapitoly o rastrové a vektorové grafice.

Následují kapitoly mířící více do hloubky oblasti strojového vidění. V kapitole s názvem *zpracování obrazu* lze nalézt informace o oblasti image pre-processingu. Dále zde byly definovány základní metody segmentace obrazu, způsoby detekce objektů z digitálních snímků a také možnosti jejich popisu pro další využití.

Teoretickou část pak uzavírá rozsáhlá kapitola týkající se neuronových sítí. Na úvod je zde představen biologický neuron, který slouží jako vzor pro základní stavební jednotku umělých neuronových sítí, tzv. formální neuron. V dalších kapitolách byly definovány základní druhy topologií a způsobů učení adaptačních algoritmů. Větší pozornost autor věnoval konkrétnímu typu neuronových sítí označovaných jako konvoluční. Podrobně byly popsány jednotlivé vrstvy těchto sítí a byl také nastíněn princip jejich fungování.

Na teoretických základech bylo následně přikročeno k realizaci vlastní práce. Pro tvorbu učební sady bylo nejprve nutné vybrat vhodná podkladová data. Tím jsou myšleny digitální fotografie splňující kritéria pro použití pro tvorbu učební sady. V první fázi tvorby se jednalo o přibližně pět set snímků otevřených zobáčků a stejné množství snímků zachycujících nakladená vajíčka v hnízdech. Po úpravě velikosti snímků a jejich přejmenování došlo k jejich zpracování pomocí programu LabelImg. V prostředí tohoto softwaru došlo k označení hledaných objektů na jednotlivých snímcích. Ke každému snímku tímto došlo k vygenerování textového popisu objektů ve formátu XML. Výsledná učební sada pak obsahovala složku *Train*, která obsahovala zdrojové snímky a složku *Annotations*, kde byly uloženy objektové deskriptory.

Vzniklé učební sady byly zpracovány pomocí konvoluční neuronové sítě na bázi YOLO. Tím vznikly konkrétní modely pro detekci jednotlivých objektů. Modely byly

nahrány na portál Uspin.cz, který funguje jako uživatelské rozhraní pro testování výsledků detekčních schopností neuronové sítě.

V rámci testování vzniklých modelů byly definovány poznatky o kvalitě fungování adaptačního algoritmu. Na základě špatných prvotních výsledků bylo rozhodnuto o nutnosti celkového přepracování a rozšíření učební sady pro rozpoznávání vajíček. Za účelem zpřesnění detekčních schopností modelu pro identifikaci otevřených zobáčků byla učební sada pro tento model také rozšířena.

Nové učební sady obsahovaly více než osm set snímků pro každý ze zájmových objektů. Velký důraz byl kladen na variabilitu použitých digitálních fotografií. Do učební sady byly také zapojeny některé snímky s lehkou obrazovou deformací.

Po opětovném procesu strojového učení bylo znovu provedeno testování obou nově vzniklých modelů. Díky portálu Uspin.cz bylo možné porovnat výsledky původních i nových modelů. Nově vzniklé modely již vykazují adekvátní výsledky s přijatelnou mírou chybovosti. Učební sady vzniklé v rámci vlastní práce lze tedy efektivně použít pro učení konvoluční neuronové sítě. Adaptační algoritmus je schopný na základě těchto podkladů efektivně rozeznávat otevřené zobáčky a vajíčka uvnitř chytrých budek. Výsledná neuronová síť je vhodná pro účely portálu Ptacionline.cz, kde bude sloužit k monitoringu vývoje ptačích druhů.

Zdrojová data byla sice získávána pouze v rámci jednoho ptačího druhu, nicméně výsledné modely by měly být univerzálně použitelné napříč širokým spektrem zástupců ptačích druhů. To autor práce vnímá jako významnou přidanou hodnotu celého projektu. Nesporným přínosem vzniklých autonomních modelů je časová a finanční úspora, a dále také snížení nutnosti využití lidského faktoru při sledování ptačích druhů. Kromě statistických účelů mohou snímky z chytrých ptačích budek sloužit také jako edukativní pomůcka v oblasti ornitologie pro základní a střední školy.

Cíle stanovené pro tuto diplomovou práci se podařilo v rámci vlastní práce splnit.

6 Seznam použitých zdrojů

- [1] Bitmapová grafika, <https://wikisofia.cz>, [online], [cit. 2018-11-19], <https://wikisofia.cz/wiki/Bitmapov%C3%A1_grafika>
- [2] BECHYŇKOVÁ M., Rastrová grafika, <http://www.ivt.mzf.cz>, [online], [cit. 2018-11-19], <<http://www.ivt.mzf.cz/seminar/9-rastrova-grafika/>>
- [3] SÚKENÍK J., K čemu je formát raw, <https://www.fotoaparát.cz/>, [online], [cit. 2018-11-20], <<https://www.fotoaparát.cz/clanek/1995/proc-fotit-do-raw-raw-vs-jpeg/>>
- [4] JPEG – vše o JPEG formátu, <https://moje.tajemno.net>, [online], [cit. 2018-11-20], <<https://moje.tajemno.net/jpeg/>>
- [5] TIŠNOVSKÝ P., <https://www.root.cz>, [online], [cit. 2018-11-20], <https://www.root.cz/clanky/jpeg-kral-rastrovych-graficky-formatu/>
- [6] Vektorová grafika, <https://wikisofia.cz>, [online], [cit. 2018-11-22], https://wikisofia.cz/wiki/Vektorov%C3%A1_grafika
- [7] BECHYŇKOVÁ M., Vektorová grafika, <http://www.ivt.mzf.cz>, [online], [cit. 2018-11-22], <http://www.ivt.mzf.cz/seminar/10-vektorova-grafika/>
- [8] Pavlíček, J., Jarolímek, J., Jarolímek, J., Pavlíčková, P., Dvořák, S., Pavlík, J. and Hanzlík, P. (2018) "Automated Wildlife Recognition", *AGRIS on-line Papers in Economics and Informatics*, Vol. 10, No. 1, pp. .ISSN 1804-1930. DOI 10.7160/aol.2018.1001xx., [cit. 2019-02-15]
- [9] HORÁK K., Popis objektů, <http://midas.uamt.feec.vutbr.cz>, [online], [cit. 2018-11-27], http://midas.uamt.feec.vutbr.cz/POV/Lectures/08_Popis_objektu.pdf
- [10] Marr, D. and Hildreth, E.: Theory of edge detection Proc. Royal Soc. Lond., volume B 207, str.: 187-217, 1980., [cit. 2019-01-16]
- [11] Canny, J. F.: Finding edges and lines in images. Technical Report AI-TR-720, MIT, Artificial Intelligence Laboratory, Cambridge, MA, 1983. [cit. 2019-01-18]
- [12] Beale, R. - Jackson, T.: Neural Computing: An Introduction. J W Arrowsmith Ltd, Bristol, Great Britain 1992. [cit. 2019-02-03]
- [13] Biologické algoritmy (4) – Neuronové sítě, <https://www.root.cz>, [online], [cit. 2018-12-27], <https://www.root.cz/clanky/biologicke-algoritmy-4-neuronove-site/>

- [14] Herz, J. - Krogh, A. - Palmer, R. G.: Introduction to the Theory of Neural Computation. Addison Wesley Publishing Company, Redwood City 1991., [cit. 2019-02-20]
- [15] Neuronové sítě, <https://is.mendelu.cz>, [online], [cit. 2018-12-27],
https://is.mendelu.cz/eknihovna/opory/zobraz_cast.pl?cast=21471
- [16] IBAMU, Umělá inteligence, <http://portal.matematickabiologie.cz>, [online], [cit. 2018-12-27], <http://portal.matematickabiologie.cz/index.php?pg=analiza-a-hodnoceni-biologickyh-dat--umela-inteligence--neuronove-site-jednotlivy-neuron--adaptacni-dynamika-neuronu--uceni-bez-ucitele>
- [17] KARPATY A., Convolutional Neural Networks for Visual Recognition, <http://cs231n.github.io>, [online], [cit. 2019-01-14], <http://cs231n.github.io/convolutional-networks/>
- [18] About the project, www.birdsonline.cz, [online], [cit. 2019-02-26],
< <https://www.birdsonline.cz/en/about-the-project> >
- [19] Soft Computing: Integrating Evolutionary, Neural, and Fuzzy Systems, Andrea Tettamanzi, Marco Tomassini, Springer Science & Business Media, 7. 9. 2001 [cit. 2019-02-16]
- [20] Kvasnička, V. - Beňušková, L. - Pospíchal, J. - Farkaš, I. - Tiňo, P. - Král, A.: Úvod do teorie neuronových sítí. IRIS, Bratislava 1997., [cit. 2019-02-16]
- [21] LabelImg, <https://github.com/>, [online], [cit. 2019-01-28],
<https://github.com/tzutalin/labelImg>
- [22] UJJWAL K., An intuitive explanation of convolutional neural networks.
<https://ujjwalkarn.me>, [online], [cit. 2019-02-24], <https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/>
- [23] CHABLANI M., YOLO You only look once, real time object detection explained, <https://towardsdatascience.com>, [online], [cit. 2019-02-27],
<https://towardsdatascience.com/yolo-you-only-look-once-real-time-object-detection-explained-492dc9230006>
- [24] Tettamanzi A., Tomassini M., Soft Computing: Integrating Evolutionary, Neural, and Fuzzy Systems, Science & Business Media, 2001, ISBN 978-3-662-04335-6 [cit. 2018-06-15]
- [25] Hlaváč V., Šonka M., Počítačové vidění. Grada, Praha, 1993, ISBN 80–85424

[26] Parker J. R., Algorithms for Image Processing and Computer Vision. Wiley Publishing, Indianapolis, 2011, ISBN 978-0-470-64385-3