



ZADÁNÍ BAKALÁ SKÉ PRÁCE

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

Pokyny pro vypracování

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

Postupujte v t ěchto krocích:

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

Seznam odborné literatury

Dodá vedoucí práce.

Ing. Michal Valenta, Ph.D.
vedoucí katedry

prof. Ing. Pavel Tvrdík, CSc.
d ěkan

V Praze dne 12. ledna 2017

ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE
FAKULTA INFORMAČNÍCH TECHNOLOGIÍ
KATEDRA SOFTWAREVÉHO INŽENÝRSTVÍ



Bakalářská práce

Detekce objektů v ptačích hnízdech pomocí neuronových sítí

Jan Havlůj

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

31. prosince 2017

Poděkování

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

Prohlášení

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

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

V Praze dne 31. prosince 2017

.....

České vysoké učení technické v Praze

Fakulta informačních technologií

© 2017 Jan Havlůj. Všechna práva vyhrazena.

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

0.0.1 Odkaz na tuto práci

Havlůj, Jan. *Detekce objektů v ptačích hnízdech pomocí neuronových sítí*. Bakalářská práce. Praha: České vysoké učení technické v Praze, Fakulta informačních technologií, 2017.

Abstrakt

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

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

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

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

Abstract

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

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

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

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

Obsah

Úvod	17
Struktura práce	17
1 Cíl práce	19
1.1 Nefunkční požadavky	19
1.2 Funkční požadavky	19
2 Rešerše	21
2.1 Počítačové vidění	21
2.2 Segmentace obrazu	21
2.3 Detekce objektů	21
2.4 Klasifikace obrazu	21
2.5 Neuronové sítě	21
2.6 Tensorflow	21
3 Analýza a návrh	23
3.1 Současný stav řešení problematiky	23
3.2 Technologie	24
3.3 Funkční požadavky	24
3.4 Rozpoznávání obrazu	24
3.5 Detekce objektů	24
3.6 Shrnutí kapitoly	24
4 Implementace rozpoznávání obrazu	25
4.1 Volba technologií	25
4.2 Trénovací data	25
4.3 Nástroje	25
4.4 Příprava dat	25
4.5 Trénování neuronové sítě	25
4.6 Ověření funkčnosti	25

5	Volba řešení	27
5.1	Srovnání implementací	27
5.2	Exportování grafu	27
5.3	Implementace Java knihovny	27
6	Ověření implementace	29
	Závěr	31
	Literatura	33
A	Seznam použitých zkratk	35
B	Dokumentace API knihovny	37
B.1	Package org.cvut.havluja1.eggdetector	37
C	Použité programy	43
D	Obsah přiloženého CD	45

Seznam obrázků

3.1	Vejce jsou zřetelně viditelná v čase 0:01	23
3.2	Vejce nejsou viditelná ve zbytku videa, jako je tomu např. v čase 0:14	24

Seznam algoritmů

Úvod

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

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

Struktura práce

Práce je rozdělena do 7 částí. První část obsahuje stanovení cílů a specifikaci funkčních i nefunkčních požadavků. Druhá část je teoretická. Diskutuje proble-

matiku počítačového vidění, strojového učení a neuronových sítí. Třetí, analytická část, se zaměřuje na výběr vhodných technologií a postupů pro tvorbu implementace. Čtvrtou i pátou, praktickou část, tvoří dva různé způsoby implementace, jejich porovnání, otestování a validace. Na závěr vybereme efektivnější implementaci, dle které je vytvořena softwarová knihovna v jazyce Java.

Cíl práce

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

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

Následující seznam funkčních a nefunkčních požadavků slouží k přesné definici požadavků na výsledný systém. Tento seznam byl vytvořen ve spolupráci s vedoucím práce.

1.1 Nefunkční požadavky

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

1.2 Funkční požadavky

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

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

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

²Zdrojový kód může používat funkce poskytované jinými softwarovými knihovnami.

³Joint Photographic Experts Group.

⁴Portable Network Graphics

Rešerše

- 2.1 Počítačové vidění
- 2.2 Segmentace obrazu
- 2.3 Detekce objektů
- 2.4 Klasifikace obrazu
- 2.5 Neuronové sítě
 - 2.5.1 CNN
 - 2.5.2 RCNN
- 2.6 Tensorflow

Text

Analýza a návrh

3.1 Současný stav řešení problematiky

V současném stavu neexistuje žádný program ani algoritmus, který by byl schopen videa automaticky analyzovat a získávat z nich užitečné informace. Protože čas vědců nebo akademických pracovníků, kteří se snaží nasbírat velké množství dat pro jejich výzkum, je drahý, najímají si brigádníky. Brigádníky je nutno nejprve zaškolit, aby věděli, co mají ve videích hledat. Poté sledují jedno video za druhým a získané informace zapisují do tabulek v Excelu nebo v Google Docs.

Jedna z informací, která nás může zajímat, je počet vajec v hnízdě. Brigádník musí video otevřít a najít část, kde jsou vejce zřetelně viditelná (viz. Obrázek 3.1) a kde viditelná nejsou (viz. Obrázek 3.2). Poté vejce spočítá a výsledek zapíše do tabulky. Jelikož je nutné získat co nejvíce dat, brigádníci musí svou práci vykonávat velmi rychle. To vede k chybám, např. chybně spočítaný počet vajec nebo zapsání výsledku na špatný řádek tabulky.



Obrázek 3.1: Vejce jsou zřetelně viditelná v čase 0:01



Obrázek 3.2: Vejce nejsou viditelná ve zbytku videa, jako je tomu např. v čase 0:14

3.2 Technologie

3.3 Funkční požadavky

3.4 Rozpoznávání obrazu

3.5 Detekce objektů

3.6 Shrnutí kapitoly

Text

Implementace rozpoznávání obrazu

- 4.1 Volba technologií
- 4.2 Trénovací data
- 4.3 Nástroje
- 4.4 Příprava dat
- 4.5 Trénování neuronové sítě
- 4.6 Ověření funkčnosti

Text

Volba řešení

- 5.1 Srovnání implementací**
- 5.2 Exportování grafu**
- 5.3 Implementace Java knihovny**

Text

Ověření implementace

6.0.1 Metodika vývoje

6.0.2 Testování

6.0.3 Možnosti vylepšení

Text

Závěr

Literatura

- [1] Česká zemědělská univerzita v Praze, Fakulta životního prostředí: *O projektu ptáci online*. [online]. [cit. 2017-12-27]. Dostupné z: <http://www.ptacionline.cz/o-projektu/o-projektu>
- [2] Česká zemědělská univerzita v Praze, Fakulta životního prostředí: *Ptáci online v médiích*. [online]. [cit. 2017-12-27]. Dostupné z: <http://www.ptacionline.cz/o-projektu>
- [3] Oracle: *jar-The Java Archive Tool*. [online]. [cit. 2017-11-14]. Dostupné z: <https://docs.oracle.com/javase/7/docs/technotes/tools/windows/jar.html>
- [4] The Apache Software Foundation: *Maven* [online]. 2017. Dostupné z: <https://maven.apache.org/>
- [5] Oracle: *Javadoc tool*. [online]. [cit. 2017-11-12]. Dostupné z: <http://www.oracle.com/technetwork/java/javase/documentation/index-jsp-135444.html>
- [6] *TexDoclet* [online]. 2017. Dostupné z: <http://doclet.github.io/>

Seznam použitých zkratek

GUI Graphical user interface

XML Extensible markup language

HTML HyperText Markup Language

RAM Random-access memory

PDF Portable Document Format

NN Neural Network

CNN Convolutional Neural Network

RCNN Recursive Convolutional Neural Network

API Application Programming Interface

Dokumentace API knihovny

Příloha obsahuje dokumentaci API Java knihovny, která je výsledkem této práce. Dokumentace byla vygenerována ze zdrojových kódů nástroji Javadoc[5] a TexDoclet[6]. Kód je psán v anglickém jazyce, stejně jako jeho dokumentace. Proto je text i zde v angličtině.

B.1 Package org.cvut.havluja1.eggdetector

<i>Package Contents</i>	<i>Page</i>
Classes	
EggDetector	37
SequenceClassifier	41

B.1.1 Class EggDetector

B.1.1.1 Count the number of eggs in given images

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

Example usage:

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

B.1.1.2 Declaration

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

B.1.1.3 Constructor summary

EggDetector()

Constructor loads the pre-trained frozen graph into memory.

B.1.1.4 Method summary

closeSession()

Closes the EggDetector session.

evaluate(File)

Runs egg detection on a given *dir*.

getMinimalConfidence()

Get the minimalConfidence setting for this instance.

isDebugMode()

Get this instance's debug mode setting.

setDebugMode(boolean)

Set this instance's debug mode setting.

setMinimalConfidence(float)

Set the minimalConfidence setting for this instance.

toString()

B.1.1.5 Constructors

- *EggDetector*

```
public EggDetector()
```

- **Description**

Constructor loads the pre-trained frozen graph into memory.

It also checks whether TensorFlow is supported on your platform.

B.1.1.6 Methods

- *closeSession*

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

- **Description**

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

- **Throws**

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

- *evaluate*

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

- **Description**

Runs egg detection on a given *dir*.

- **Parameters**

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

- **Returns** –

- **Throws**

- * `java.lang.IllegalArgumentException` – if *dir* is not a directory or contains no images

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

- *getMinimalConfidence*

```
public float getMinimalConfidence()
```

- **Description**

Get the *minimalConfidence* setting for this instance.

Minimal Confidence score is used as a confidence boundary during the process of object detection. An object that has been detected with a confidence score lower than *minimalConfidence* is ignored. An object that has been detected with a confidence score higher or equal than *minimalConfidence* is added to the final result list.

- **Returns** – This instance's *minimalConfidence* setting.

- *isDebugMode*

```
public boolean isDebugMode()
```

- **Description**

Get this instance's debug mode setting.

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

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

- *setDebugMode*

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

- **Description**

Set this instance's debug mode setting.

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

- **Parameters**

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

- **Throws**

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

- *setMinimalConfidence*

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

- **Description**

Set the *minimalConfidence* setting for this instance.

Minimal Confidence score is used as a confidence boundary during the process of object detection. An object that has been detected with a confidence score lower than *minimalConfidence* is ignored. An object that has been detected with a confidence score higher or equal than *minimalConfidence* is added to the final result list.

- **Parameters**
 - * `minimalConfidence` – *minimalConfidence* for this instance
- **Throws**
 - * `java.lang.IllegalStateException` – if the session has been closed already by calling `closeSession()`
- *toString*

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

B.1.2 Class SequenceClassifier

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

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

Example usage:

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

B.1.2.2 Declaration

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

B.1.2.3 Method summary

getFinalCount()

Get the final score for the entire directory.

getIndividualCounts()

Gets the individual egg count for every image provided.

B.1.2.4 Methods

- *getFinalCount*

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

- **Description**

Get the final score for the entire directory.

The final score is calculated as follows:

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

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

- *getIndividualCounts*

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

- **Description**

Gets the individual egg count for every image provided.

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

Použité programy

TexStudio psaní bakalářské práce v \LaTeX a překlad do PDF

Notepad++ úprava textových souborů

gedit úprava textových souborů

IntelliJ IDEA Ultimate psaní implementace v jazyce Java

GIT verzování kódu

JDK8 kompilace, ladění a testování Java knihovny

Python 3 zpracování dat, trénování neuronových sítí

Tensorflow softwarová knihovna pro strojové učení

LabelImg označování trénovacích dat

javadoc generování dokumentace do HTML

TexDoclet generování dokumentace do \LaTeX

wget hromadné stahování dat

bash pomocné skripty pro zautomatizování práce

Obsah přiloženého CD

readme.txt	stručný popis obsahu CD
bin	adresář se spustitelnou formou implementace
doc	adresář s dokumentací implementace
src	
impl	zdrojové kódy implementace
eggdetector	zdrojové kódy implementace knihovny
nn_training	zdrojové kódy pro trénování neuronové sítě
tools ..	zdrojové kódy pomocných nástrojů k tvorbě bakalářské práce
thesis	zdrojová forma práce ve formátu L ^A T _E X
text	text práce
thesis.pdf	text práce ve formátu PDF