



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

ÚSTAV AUTOMATIZACE A MĚŘICÍ TECHNIKY

DEPARTMENT OF CONTROL AND INSTRUMENTATION

SYSTÉM PRO SLEDOVÁNÍ A KLASIFIKACI OBJEKTŮ NA OBLOZE

SYSTEM FOR TRACKING AND CLASSIFICATION OF OBJECTS IN THE SKY

SEMESTRÁLNÍ PRÁCE

SEMESTRAL THESIS

AUTOR PRÁCE

AUTHOR

Bc. Jakub Franka

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. Ilona Janáková, Ph.D.

BRNO 2023



Semestrální práce

magisterský navazující studijní program **Kybernetika, automatizace a měření**

Ústav automatizace a měřicí techniky

Student: Bc. Jakub Franka

ID: 220976

Ročník: 2

Akademický rok: 2023/24

NÁZEV TÉMATU:

Systém pro sledování a klasifikaci objektů na obloze

POKYNY PRO VYPRACOVÁNÍ:

Úkolem studenta je navrhnout systém počítačového vidění pro detekci, trasování a klasifikaci létajících objektů na obloze. Zadání zahrnuje řešení po HW i SW stránce. Předpokládá se klasifikace do kategorií typu: letadlo, dron, pták, hmyz, případně dalších. Pro klasifikaci bude testováno i využití konvolučních neuronových sítí.

1. Seznamte se s danou problematikou. Proveďte rešerši existujících přístupů.
2. Navrhněte vhodnou kombinaci a uspořádání hardwarových komponent – kamera s optikou + PC/minipočítač.
3. Navržené zařízení realizujte.
4. Pořídeť dostatečně rozsáhlou a pestrou databázi reálných snímků/sekvencí.
5. Navrhněte algoritmy detekce a trasování objektů.
6. Navrhněte klasifikátor objektů.
7. Implementujte zvolené algoritmy do navrženého zařízení.
8. Vše otestujte. Definujte omezuječí podmínky. Zhodnotěte.

SP = 1,5,6.

DOPORUČENÁ LITERATURA:

JUREČKA, Tomáš. Detekce a klasifikace létajících objektů. Brno, 2021. Diplomová práce. Vysoké učení technické v Brně, FEKT, Ústav automatizace a měřicí techniky. Vedoucí práce Ilona Janáková.

SCHUMANN, Arne, Lars SOMMER, Johannes KLATTE, Tobias SCHUCHERT a Jurgen BEYERER. Deep cross-domain flying object classification for robust UAV detection. In: 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). IEEE, 2017, s. 1-6. ISBN 978-1-5386-2939-0. Dostupné z: doi:10.1109/AVSS.2017.8078558

Termín zadání: 18.9.2023

Termín odevzdání: 4.1.2024

Vedoucí práce: Ing. Ilona Janáková, Ph.D.

doc. Ing. Petr Fiedler, Ph.D.

předseda rady studijního programu

UPOZORNĚNÍ:

Autor semestrální práce nesmí při vytváření semestrální práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č.40/2009 Sb.

ABSTRAKT

Táto práca sa zaobrá využitím počítačového videnia v oblasti detektie, trasovania a klasifikácie lietajúcich objektov v reálnom prostredí. Cieľom je vytvoriť robustný systém, schopný efektívne pracovať v nepriaznivých podmienkach a presne identifikovať rôzne typy objektov v oblohe. Práca postupuje od teoretických základov, výberu metód, až po návrh a implementáciu systému počítačového videnia.

KĽÚČOVÉ SLOVÁ

počítačové videnie, detekcia objektov, trasovanie, klasifikácia, lietajúce objekty, mikro-počítače, dataset

ABSTRACT

Překlad abstraktu This work deals with the use of computer vision in the field of detection, tracking and classification of flying objects in a real environment. The goal is to create a robust system, capable of working effectively in adverse conditions and accurately identifying different types of objects in the sky. The work progresses from theoretical foundations, choice of methods, to the design and implementation of a computer vision system.

KEYWORDS

Překlad klíčových slov computer vision, object detection, tracking, classification, flying objects, microcomputers, dataset

FRANKA, Jakub. *Systém pro sledování a klasifikaci objektů na obloze*. Semestrálna práca. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav automatizace a měřicí techniky, 2024. Vedúci práce: Ing. Ilona Janáková, Ph.D.

Vyhľásenie autora o pôvodnosti diela

Meno a priezvisko autora: Bc. Jakub Franka

VUT ID autora: 220976

Typ práce: Semestrálna práca

Akademický rok: 2023/24

Téma záverečnej práce: Systém pro sledování a klasifikaci objektů na obloze

Vyhlasujem, že svoju záverečnú prácu som vypracoval samostatne pod vedením vedúcej/cého záverečnej práce, s využitím odbornej literatúry a ďalších informačných zdrojov, ktoré sú všetky citované v práci a uvedené v zozname literatúry na konci práce.

Ako autor uvedenej záverečnej práce ďalej vyhlasujem, že v súvislosti s vytvorením tejto záverečnej práce som neporušil autorské práva tretích osôb, najmä som nezasiahol nedovoleným spôsobom do cudzích autorských práv osobnostných a/alebo majetkových a som si plne vedomý následkov porušenia ustanovenia § 11 a nasledujúcich autorského zákona Českej republiky č. 121/2000 Sb., o práve autorskom, o právach súvisiacich s právom autorským a o zmene niektorých zákonov (autorský zákon), v znení neskorších predpisov, vrátane možných trestnoprávnych dôsledkov vyplývajúcich z ustanovenia časti druhej, hlavy VI. diel 4 Trestného zákonníka Českej republiky č. 40/2009 Sb.

Brno
.....
podpis autora*

*Autor podpisuje iba v tlačenej verzii.

POĎAKOVANIE

Obsah

| | |
|--|-----------|
| Úvod | 12 |
| 1 Detekcia objektov v obraze | 13 |
| 1.1 Prahovanie | 13 |
| 1.1.1 Globálne prahovanie | 13 |
| 1.1.2 Lokálne prahovanie | 15 |
| 1.2 Detekcia hrán | 15 |
| 1.2.1 Detekcia horizontu pomocou Houghovej transformácie | 17 |
| 1.3 Detekcia pohybu | 18 |
| 2 Popis objektov | 21 |
| 2.1 Popis tvaru | 21 |
| 2.2 Fotometrické príznaky | 22 |
| 3 Klasifikácia | 23 |
| 3.1 Klasifikátory | 23 |
| 3.2 Konvolučné neurónové siete | 23 |
| 4 Detekcia a klasifikácia v jednom kroku | 25 |
| 4.1 R-CNN | 25 |
| 4.2 YOLO | 26 |
| 5 Návrh usporiadania hardvérových komponentov | 27 |
| 5.1 Jednodoskový počítač | 27 |
| 5.2 Kamera | 27 |
| 5.3 Ovládacie prvky | 28 |
| 5.4 Uchytenie kamery | 28 |
| 6 Tvorba datasetu | 29 |
| 6.1 Získavanie dát | 29 |
| 6.2 Úprava, rozdelenie a rozšírenie datasetu | 29 |
| 6.3 Testovanie použiteľnosti datasetu | 30 |
| 7 Návrh softvéru | 32 |
| 7.1 Použitý jazyk a knižnice | 32 |
| 7.2 Štruktúra aplikácie | 32 |
| 7.3 Implementácia | 33 |
| 7.4 Návrh grafického rozhrania | 34 |

| | |
|----------------------------|----|
| Závěr | 37 |
| Literatúra | 38 |
| Zoznam symbolov a skratiek | 39 |
| Zoznam príloh | 40 |

Zoznam obrázkov

| | | |
|-----|--|----|
| 1.1 | Funkcia prahovania s prahom 0.5 pre svetlé pozadie | 13 |
| 1.2 | Jednoduché prahovanie | 14 |
| 1.3 | Automatické prahovanie (podľa známeho rozloženia) | 14 |
| 1.4 | Adaptívne prahovanie (podľa známeho rozloženia) | 15 |
| 1.5 | Cannyho hranový detektor | 17 |
| 1.6 | Detekcia horizontu pomocou Houghovej transformácie | 18 |
| 1.7 | Postup aktualizovania dynamického modelu pozadia | 19 |
| 1.8 | Odčítanie pozadia | 20 |
| 3.1 | Príklad architektúry konvolučnej neurónovej siete | 24 |
| 5.1 | návrh nastaviteľného uchytenia kamery | 28 |
| 5.2 | bloková schéma zariadenia | 28 |
| 6.1 | Metriky presnosti naučeného modelu v programe roboflow | 30 |
| 6.2 | Grafické zobrazenie priebehu učenia modelu roboflow | 31 |
| 6.3 | Exportovanie vytvoreného datasetu | 31 |
| 7.1 | Záložka na výber poskytovateľa obrazu | 34 |
| 7.2 | Záložka na výber metódy detektie objektov | 35 |
| 7.3 | Záložka na výber metódy sledovania objektov | 35 |
| 7.4 | Záložka na výber detekovaných tried | 36 |

Zoznam tabuliek

Zoznam výpisov

Úvod

V súčasnej dobe hrá počítačové videnie dôležitú rolu v mnohých aplikáciách. Táto práca sa zameriava na jeho využitie pri návrhu systému schopného detektovať, trasovať a klasifikovať objekty lietajúce na oblohe. Cieľom je vytvoriť robustný systém schopný nasadenia v reálnych nepriaznivých podmienkach so schopnosťou detekcie a klasifikácie rôznych typov živých aj umelo vytvorených lietajúcich objektov. Je požadovaná možnosť použitia rôznych prístupov spracovania vstupných dát a porovnanie ich výkonu a presnosti.

Táto práca bude postupovať od teoretických základov do konkrétnego návrhu a implementácie popísaného systému počítačového videnia.

Prvá kapitola sa venuje základným princípom detekcie objektov v obraze. Sú tu preskúmané existujúce metódy s porovnaním ich výhod a nevýhod, čo povedie k výberu vhodných metód pre navrhovaný systém.

V druhej kapitole bude popísané získavanie charakteristík popisujúcich jednotlivé detektované objekty. Zohľadňujú sa tu aspekty ako tvar, farba a ďalšie príznaky, na ktorých záleží klasifikácia v ďalšom kroku.

Na základe popisu objektov bude v ďalšej kapitole popísaná problematika klasifikácie, teda rozpoznávania a zaradenia detektovaných objektov do definovaných tried podľa ich typu (lietadlo, helikoptéra, vták, hmyz, ...).

Štvrtá kapitola predstaví moderné metódy integrujúce detekciu a klasifikáciu do jedného procesu. Cieľom je optimalizácia efektivity a jednoduchosti architektúry systému.

Následne bude navrhnutý a odôvodnený výber hardvérových komponentov systému. Budú porovnané mikropočítače vhodné na aplikáciu v tejto práci a zvolené vhodné periférie na získavanie dát a užívateľské rozhranie.

Pre správne naučenie modelov použitých v systéme je nevyhnutné vytvoriť dostačne rozsiahlu a správne navrhnutú anotovanú množinu dát, v prípade algoritmov počítačového videnia ide o databázu snímkov s označenými a triedenými objektami. Šiesta kapitola sa zameriava na tvorbu tejto databázy (datasetu) a jej validáciu.

V poslednej kapitole bude zdokumentovaný návrh softvéru pre systém detekcie, klasifikácie a trasovania objektov. V tejto kapitole bude priblížená štruktúra aplikácie, nástroje použité pri vývoji a návrh jej užívateľského rozhrania.

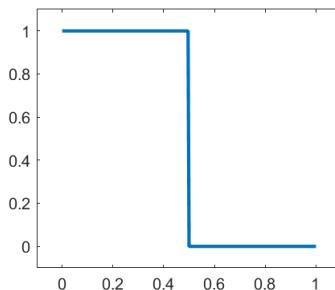
1 Detekcia objektov v obraze

Pre niektoré techniky klasifikácie objektov je nutné najprv v obraze tieto objekty nájsť. Existuje k tomu veľké množstvo metodík, pri čom každá z nich má svoje výhody a nevýhody. Je teda nutnosťou zistiť, ktorá metóda najviac vyhovuje použitiu pre túto prácu.

1.1 Prahovanie

Prahovanie je jasová transformácia, pri ktorej je šedotónový obraz rozdelený na dve a viac oblastí podľa jasovej úrovne pixelov. O tom, do ktorej z oblastí je daný pixel priradený rozhoduje prah, ktorého hodnotu je treba určiť. Výsledkom je obraz s len dvomi úrovňami jasu - binárny obraz.

Pri správnom predspracovaní obrazu prahovaním je možné v ňom jednoducho oddeliť objekty od pozadia, v prípade ak sa jas objektu dostatočne líši od jasu pozadia. Ak sa však jasové hodnoty objektu a pozadia príliš podobajú, nemusí existovať dostatočne robustné určenie prahu oddelujúceho objekty od pozadia. [3] [5]



Obr. 1.1: Funkcia prahovania s prahom 0.5 pre svetlé pozadie

1.1.1 Globálne prahovanie

Globálne prahovanie je často používaná technika, pri ktorej je použitý jediný prah na celý obraz, čo nemusí byť vhodné v prípade, ak je jasová úroveň pozadia rôzna v rôznych miestach obrazu. [3]

- **Jednoduché prahovanie** je technika, pri ktorej je globálny prah určený dopredu ručne. V prípade, že sa jasová úroveň pozadia v čase mení, je vhodnejšie použiť techniky s adaptívnym nastavením prahu.
- Prahovanie **podľa známeho rozloženia** predpokladá že poznáme relatívnu veľkosť oblasti, ktorú zaberá pozadie v obraze. Ak vieme, že objekt je svetlejší ako pozadie, môžeme určiť prah z kumulatívneho histogramu tak, aby relatívny

počet pixelov pod úrovňou prahu bol rovnaký ako oblasť, ktorú má zaberať pozadie.

- Algoritmus **K-Means** (K priemerov) pre zhlukovú analýzu rozdeľuje dátu do skupín s cieľom minimalizovať vzdialenosť bodov v zhluku a maximalizovať vzdialenosť medzi zhlukmi. Jedná sa o algoritmus učenia bez učiteľa. Funguje na princípe iterovaného posúvania k stredov zhlukov (v prípade binárneho prahovania je $k = 2$) smerom k priemeru hodnôt priradenému danému stredu v aktuálnom kroku.
- **Otsuova metóda** automatického určenia prahu je založená na maximalizovaní vzájomnej odchýlky medzi triedami. Využíva normalizovaný kumulatívny histogram, z ktorého určuje vzájomný rozptyl pre všetky možné hodnoty prahu a vyberá ten optimálny.



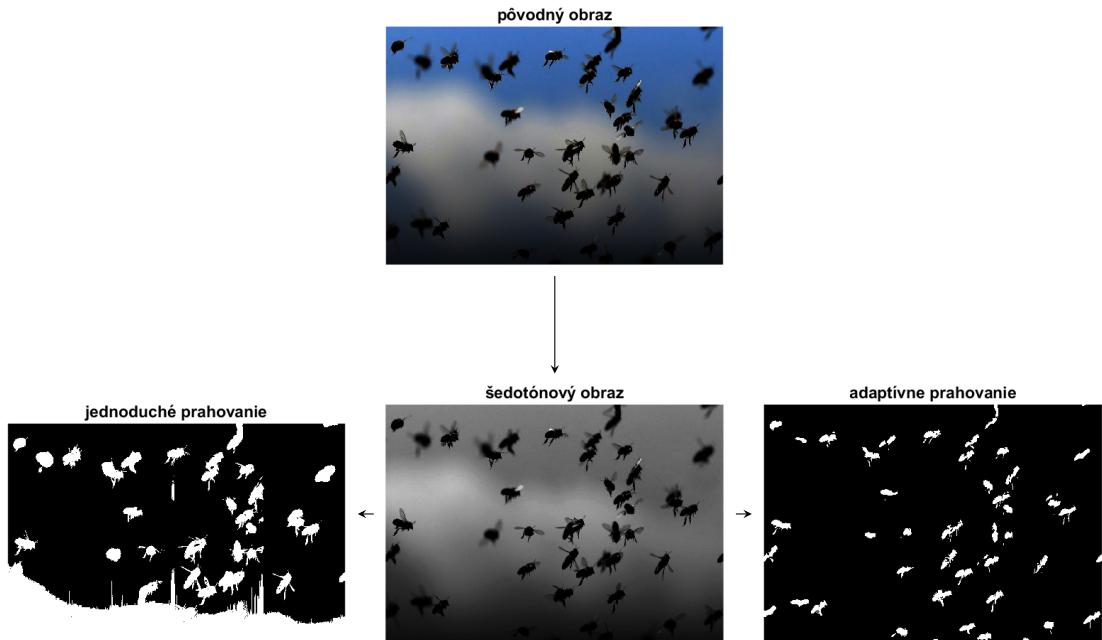
Obr. 1.2: Jednoduché prahovanie



Obr. 1.3: Automatické prahovanie (podľa známeho rozloženia)

1.1.2 Lokálne prahovanie

V prípade, že je jas pozadia rôzny v rôznych častiač obrazu, je vhodné určiť hodnotu prahu z jasovej úrovne okolia každého prahovaného pixelu. Toto takzvané adaptívne prahovanie má súčasť vyššiu výpočtovú náročnosť ako jednoduchšie globálne prahovanie, keďže je nutné určiť prah pre každý bod obrazu samostatne, je však robustnejšie voči nevhodnému pozadiu. [5]



Obr. 1.4: Adaptívne prahovanie (podľa známeho rozloženia)

1.2 Detekcia hrán

Ako jednu z metód vyhľadávania objektov v obraze je možné využiť obraz hrán a vyhľadať v ňom kontúry objektov. Opäť existuje niekoľko spôsobov ako získať obraz hrán a ako ho využiť k segmentácii obrazu, teda k detekcii objektov.

Na vytvorenie obrazu hrán z nasnímaného obrazu sa používa konvolúcia obrazu s operátorom navrhnutým tak aby aproximovalo deriváciu určitého rádu. Medzi často používané operátory patria Prewittov a Sobelov operátor, ktoré aproximujú prvú deriváciu v určitom smere (horizontálnom, vertikálnom, diagonálne). Druhú deriváciu approximuje Laplaceov operátor, používa sa tiež v kombinácii s Gaussovým filtrom (LoG operátor), čo má za následok zníženie citlivosti na šum. [6]

Komplexnejšou metódou detekcie hrán je takzvaný *Cannyho hranový detektor*. Ide o proces vo viacerých krokoch [6]. Pre jeden vstupný obraz sa postupuje nasledovne:

$$\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

Prewittov
operátor

Sobelov
operátor

Laplaceov
operátor

$$\begin{bmatrix} 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & -2 & -1 & 0 \\ -1 & -2 & 16 & -2 & -1 \\ 0 & -1 & -2 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 \end{bmatrix}$$

LoG operátor

1. Vstupný obraz býva väčšinou filtrovaný z dôvodu zníženia citlivosti na šum v obraze. Vhodné je k tomu napríklad Gaussove rozmazanie.
2. Použije sa horizontálny a vertikálny Sobelov operátor ako detektor hrán na získanie obrazu hrán v oboch smeroch I_x a I_y .

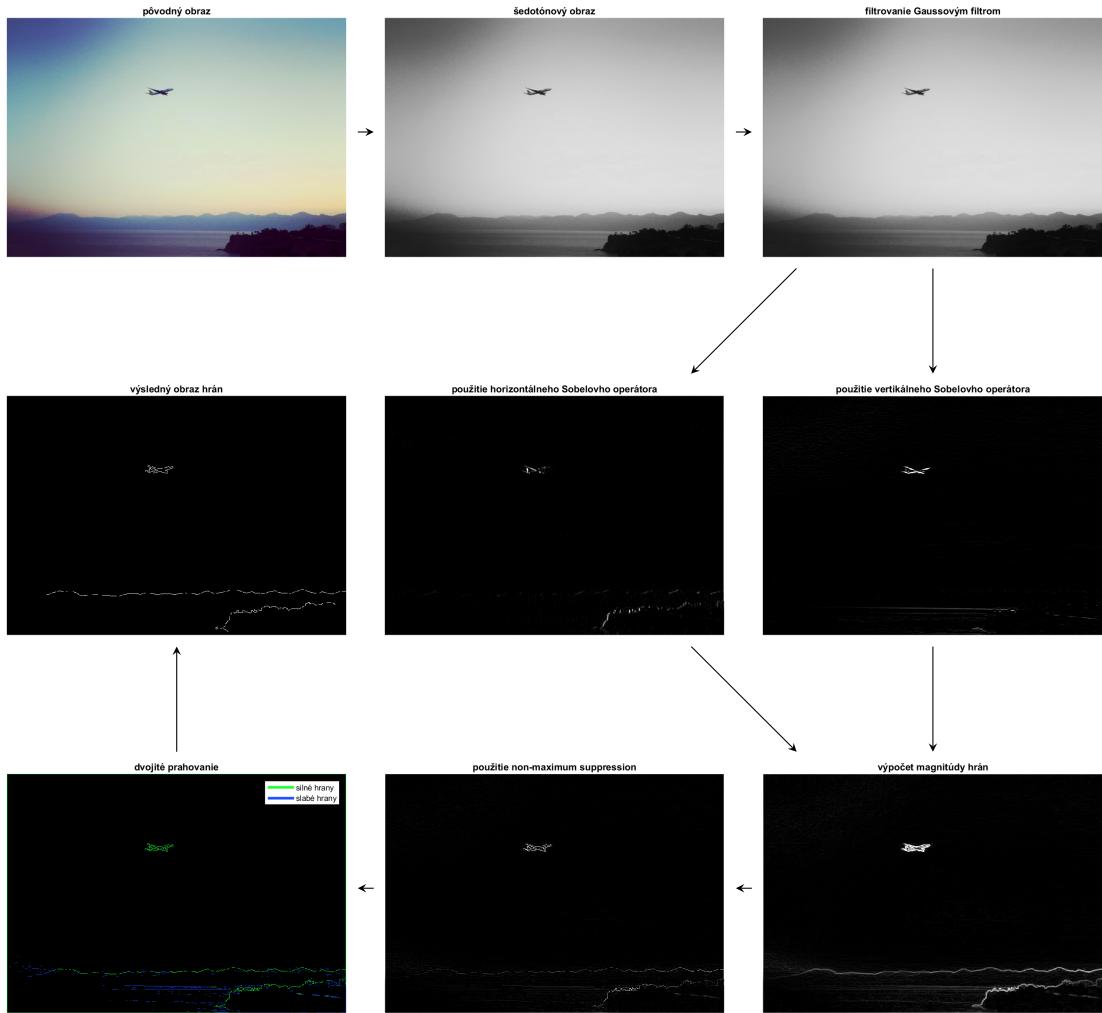
$$I_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} * I \quad I_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * I$$

3. Z toho je možné určiť magnitúdu a smer gradientu:

$$G = \sqrt{I_x^2 + I_y^2} \quad \theta = \text{atan2}(I_x, I_y)$$

4. Ku zníženiu redundancie detekovaných hrán sa použije algoritmus *non maxima suppression* (potlačenie nemaximálnej hodnoty). Použitím tohto algoritmu dôjde k zúženiu hrán, pričom sa ponechajú len tie časti hrán, ktorých magnitúda je väčšia ako v ich okolí.
5. Dvojitým prahovaním sa v obraze hrán nájdú tri úrovne:
 - **silné hrany** sú hrany vyššie ako obidva prahy. Do výsledného obrazu hrán sú určite pridané.
 - **slabé hrany** sú hrany medzi dvomi prahmi.
 - **nie hrany** sú tie hrany ktorých magnitúda je nižšia ako obidva prahy.
6. Iteratívne sú slabé hrany dotýkajúce sa silných hrán označované za silné, až kým sa žiadna slabá hrana silnej nedotýka. Vo výslednom obraze sú ponechané len silné hrany.

Obraz hrán je možné ďalej spracovať napríklad morfologickými operáciami: otvorením odstrániť krátke nevýznamné hrany, zatvorením spojiť hrany blízko pri sebe.

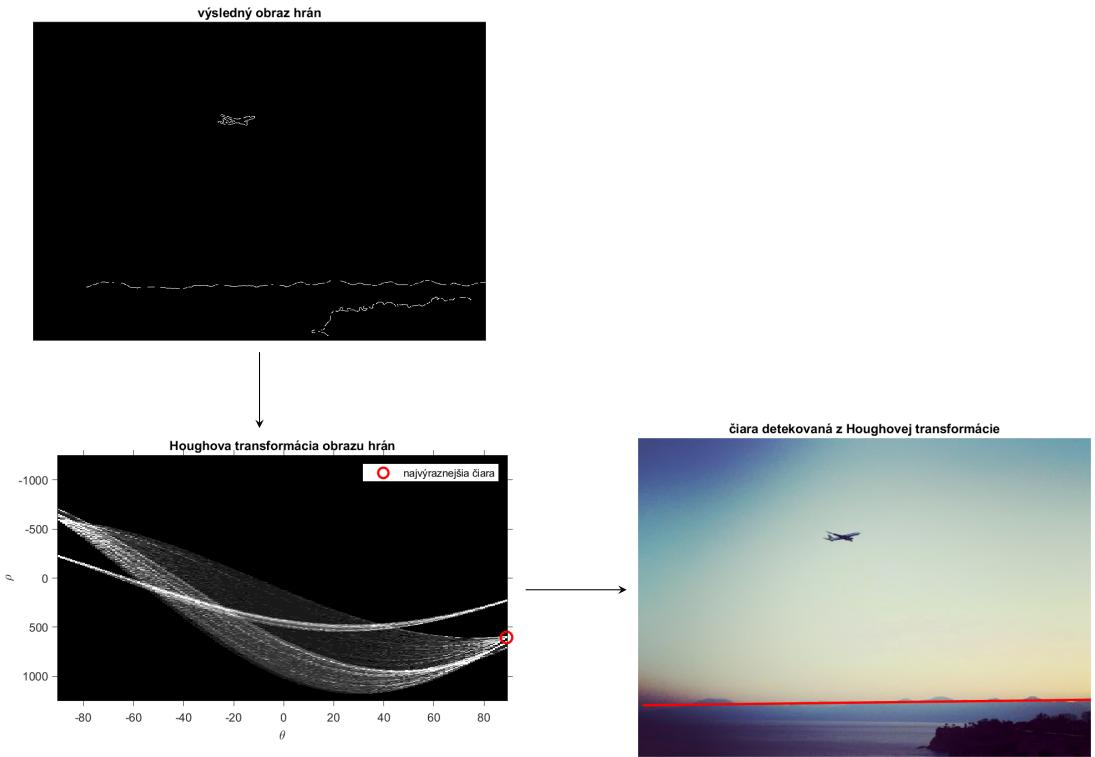


Obr. 1.5: Cannyho hranový detektor

Vyhľadávaním uzavretých kontúr v obraze hrán je možné identifikovať potenciálne objekty v obraze.

1.2.1 Detekcia horizontu pomocou Houghovej transformácie

Pokial časť obrazu obsahuje horizont, je vhodnejšie detektovať lietajúce objekty len na oblohe nad ním, aby sa predišlo falošným detekciám. Horizont je možné detektovať z obrazu hrán pomocou Houghovej transformácie, podľa práce [2]. Vo výsledku Houghovej transformácie sa horizont prejaví ako najvýraznejšia čiara.



Obr. 1.6: Detekcia horizontu pomocou Houghovej transformácie

1.3 Detekcia pohybu

V prípade, že je požadovaná detekcia pohybujúcich sa objektov, je možnosťou detektie práve detekovaním ich pohybu. Jednou z jednoduchých metód detektie pohybu sú rozdielové metódy, pri ktorých je vypočítavaný rozdielový snímok z viacerých snímkov zo sekvencie $\{I_1, I_2, \dots, I_n\}$. Rozdielový snímok môže byť:

- **jednostranný** - je najjednoduchší, nenesie informáciu o smere pohybu, vyhodnocuje sa len v miestach, kde $I_1(x, y) > I_2(x, y)$.

$$D(x, y) = \begin{cases} 0 & I_1(x, y) - I_2(x, y) < \epsilon \\ 1 & I_1(x, y) - I_2(x, y) \geq \epsilon \end{cases}$$

- **obojstranný** - dostaneme ho upravením vzťahu pre jednostranný rozdielový snímok použitím absolútnej hodnoty. Tým je dosiahnutá zameniteľnosť $I_1(x, y)$ a $I_2(x, y)$, vyhodnocuje sa teda na celom obraze. Neodstraňuje nedostatok informácie o smere pohybu, tá však nemusí byť nutnosťou.

$$D(x, y) = \begin{cases} 0 & |I_1(x, y) - I_2(x, y)| < \epsilon \\ 1 & |I_1(x, y) - I_2(x, y)| \geq \epsilon \end{cases}$$

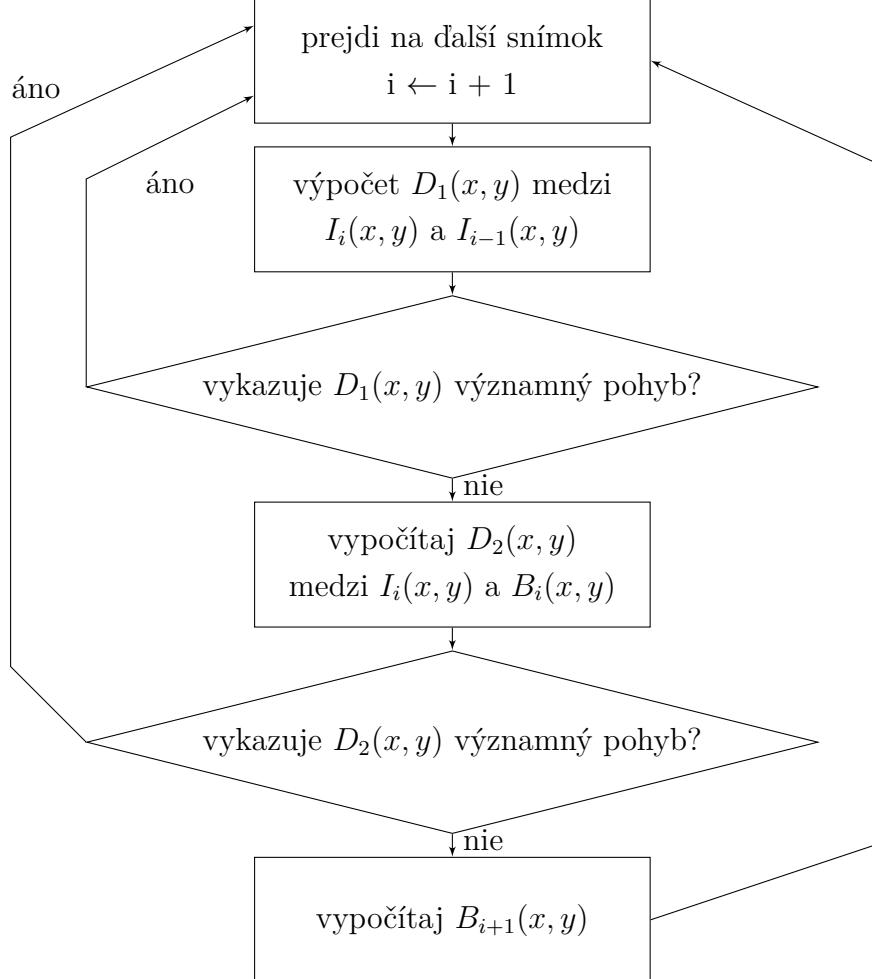
- **kumulovaný** - je vytvorený váženým súčtom jedno alebo obojstranných rozdielových snímkov. Je teda možné priemerovať pohyb za určitý čas určením každej váhy $\omega = \frac{1}{N-1}$ alebo určiť rôznym snímkom rôznu váhu a získať tak informáciu o smere pohybu.

$$D(x, y) = \sum_{i=1}^{N-1} \omega_i \cdot D_i(x, y)$$

Komplexnejšou metódou detektie pohybu je vytvorenie rozdielového snímku nie medzi nasledujúcimi snímkami, ale aktuálnym snímkom a vytvoreným modelom pozadia. Ten je možné zostaviť:

- ako jeden obraz pozadia bez objektov.
- ako priemerný snímok niekoľkých snímkov pozadia.
- ako dynamický model - iteratívnym aktualizovaním podľa aktuálneho snímku.

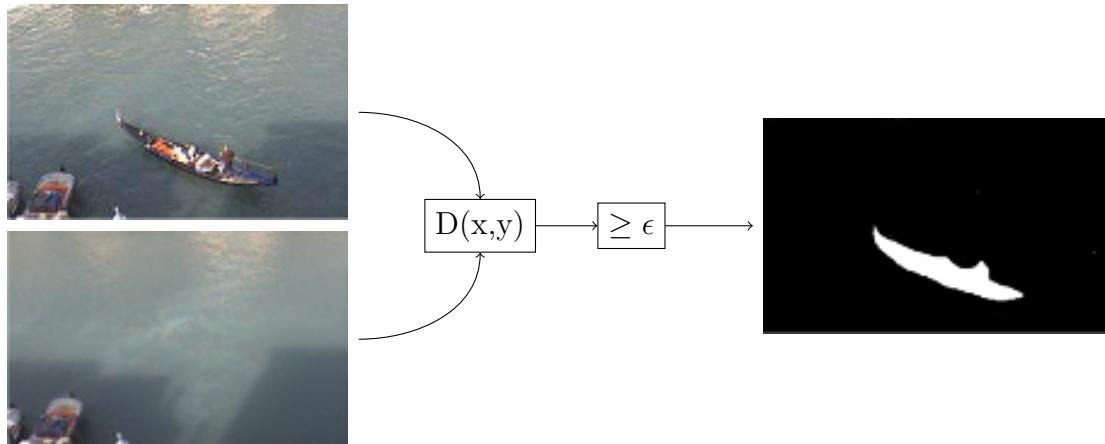
Tým sa model pozadia postupne prispôsobuje malým zmenám v prostredí.



Obr. 1.7: Postup aktualizovania dynamického modelu pozadia

Výpočet nového modelu pozadia $B_{i+1}(x, y)$ môže prebiehať napríklad pomocou lineárneho zabúdania, teda $B_{i+1}(x, y) = \alpha \cdot I_i(x, y) + (1 - \alpha) \cdot B_i(x, y)$.

Obraz obsahujúci detekovaný pohyb je získaný rozdielom a prahovaním aktuálneho snímku a modelu pozadia. [7]



Obr. 1.8: Odčítanie pozadia

2 Popis objektov

Po nájdení objektov v obraze nasleduje ich popis pomocou príznakového vektoru. Ten predstavuje matematickú reprezentáciu segmentu obrazu vektorom o ľubovoľnom počte rozmerov. Jeho prvky popisujú farbu, textúru alebo tvar objektu. Samotný objekt nemusí byť rekonštruovateľný z jeho popisu, cieľom je podľa neho rozlísiť od seba dva rôzne objekty.

Vypočítané hodnoty príznakového vektoru majú niekoľko využití:

- **Párovanie objektov** v dvoch snímkoch, teda určenie, ktoré body alebo objekty v jednom snímku korelujú s tými v druhom. Výhodou je to pri vyhľadávaní rovnakého objektu v po sebe nasledujúcich snímkoch pri sledovaní pohybu, či pri hľadaní rovnakého objektu v obraze z dvoch kamier.
- **Klasifikácia** objektov v obraze. Objekty rôznych tried majú často rôzne hodnoty určitých príznakov, je však nutné určiť ktoré príznaky sú pre klasifikáciu relevantné.

2.1 Popis tvaru

Jedny z najjednoduchších príznakov vychádzajú z tvaru objektu, používajú sa s binárnym obrazom získaným napríklad prahovaním. [4] Sú to napríklad:

- **Obsah** oblasti objektu, teda počet pixelov predstavujúcich objekt v binárnom obraze.
- **Obvod** je počet pixelov na hrane tvaru objektu. Diagonálne hrany môžu byť pri výpočte obvodu nadhodnotené v prípade, ak počítame s 4-okolím alebo podhodnotené, ak počítame s 8-okolím bodov na hrane. Je teda správnejšie počítať 1 pixel pre pravouhlé hrany a $\sqrt{2}$ pre diagonálne hrany.
- **Kompaktnosť** alebo kruhovosť je podobnosť tvaru ku kruhu. Vypočítava sa ako $\frac{Obvod^2}{Obsah}$ a je vždy vyššia alebo rovná kompaktnosti kruhu, čo je 4π .
- **Excentricita** môže mať niekoľko definícií, jednou z nich je pomer dĺžky najdlhšej úsečky v tvare objektu s dĺžkou najdlhšej úsečky na ňu kolmej. Taktiež je to excentricita elipsy s rovnakým momentom druhého rádu, teda pomer vzdialenosí ohnisiek a dĺžky hlavnej osy.
- **Pozdĺžnosť** je pomer výšky a šírky minimálneho opísaného obdĺžnika.
- **Pravouhlosť** alebo podobnosť s obdĺžnikom je pomer obsahu objektu a obsahu minimálneho opísaného obdĺžnika. Môže nadobúdať hodnoty do 1 a limitne môže klesať k 0.
- **Orientácia** alebo uhol hlavenj osi.
- **Eulerovo číslo** je topologický príznak značiaci rozdiel počtu spojených komponentov a dier v tvare objektu. Topológia popisuje vlastnosti tvaru, ktoré sa

nemenia úpravami okrem delenia a spájania častí tvaru.

- **Počet konkávností** nie je topologickým príznakom, napriek jeho podobnosti s nimi. Získa sa spočítaním spojitéh oblastí v rozdielneho obrazu objektu a jeho konvexného obalu.
- **Konvexnosť** je možné určiť z pomeru obsahu objektu ku obsahu jej konvexného obalu.

2.2 Fotometrické príznaky

Príznaky, ktorých hodnota závisí na jasových úrovniach popisovaného segmentu obrazu sa nazývajú fotometrické príznaky. Môže íst o priemernú ($B_{priemer}$), minimálnu (B_{min}), maximálnu (B_{max}) hodnotu jasu objektu, rozdiel extrémov v objekte ($B_{dif.}$), priemerov jasu objektu (Ω) a pozadia (Φ) ($B_{dif.pozadia}$), či príznaky vyplývajúce z histogramu (H) alebo geometrických momentov.

$$B_{priemer} = \frac{1}{N} \cdot \sum_{(x,y) \in \Omega} f(x,y) \quad B_{max} = \max(f(x,y))$$

$$B_{dif.pozadia} = B_{priemer(\Omega)} - B_{priemer(\Phi)} \quad B_{min} = \min(f(x,y))$$

$$H_{priemer} = \frac{1}{N} \cdot \sum_{q=1}^N q \cdot h(q) \quad H_{kontrast} = \frac{1}{N} \cdot \sum_{q=1}^N (q \cdot h(q) - H_{priemer})^2$$

$$H_{energia} = \frac{1}{N} \cdot \sum_{q=1}^N h(q)^2 \quad H_{entropia} = \frac{1}{N} \cdot \sum_{q=1}^N h(q) \cdot \log_2 h(q)$$

Z geometrických momentov, ktorých hodnoty sú závislé na afinných transformáciach (translacia, rotácia a úprava mierky) je možné úpravou získať príznaky na transformácii nezávislé. [9] Všeobecne je moment rádu (p,q) segmentu l(x,y) daný:

$$m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q l(x,y) dx dy \quad m_{pq} = \sum_X \sum_Y x^p y^q l(x,y) \\ (\text{definícia}) \quad (\text{Pre konečný diskrétny obraz})$$

Centralizované momenty μ_{pq} , vztiahnuté k tažisku objektu $(x_t, y_t) = (\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}})$, získavajú nezávislosť na transláciu:

$$\mu_{pq} = \sum_X \sum_Y (x - x_t)^p \cdot (y - y_t)^q \cdot l(x,y)$$

Z centralizovaných momentov je možné určiť momentové invarianty, základných sedem je nazývaných Huovými momentami.

3 Klasifikácia

Ďalším krokom je zaradenie detekovaných objektov do tried, teda ich klasifikácia. Klasifikácia objektov vychádza z ich popisu.

3.1 Klasifikátory

Klasifikátory sú algoritmy schopné automaticky kategorizovať dátu do dvoch a viac tried. Modely určené na klasifikáciu sú učené s učiteľom, je potrebné pomerne veľké množstvo anotovaných dát. Výsledkom je schopnosť generalizovať klasifikáciu na nových neznámych dátach.

Existuje niekoľko typov klasifikátorov, napríklad **Lineárne klasifikátory** rozdeľujú príznakový priestor nadrovinami, **Bayesov klasifikátor** vypočítava pravdepodobnosť zaradenia do triedy.

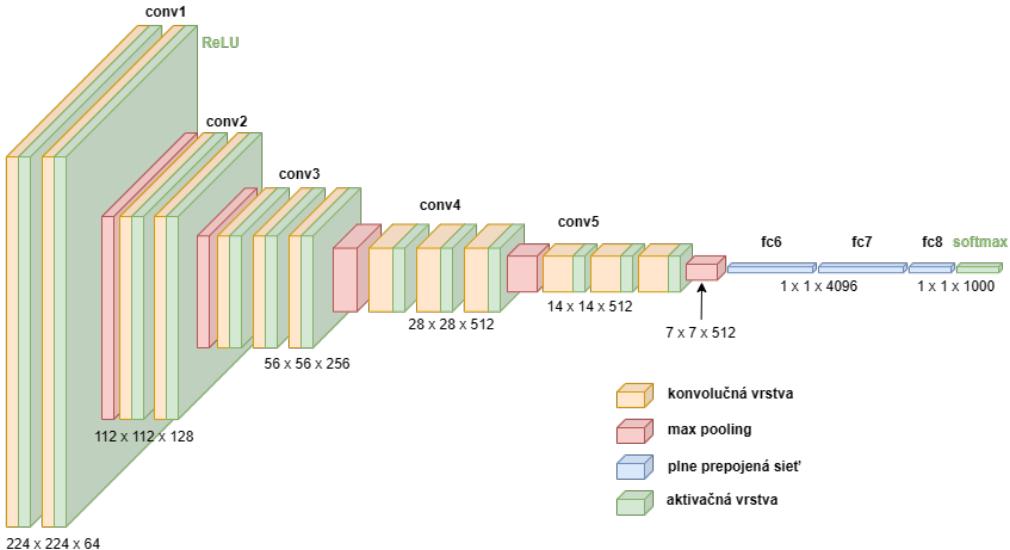
3.2 Konvolučné neurónové siete

Na rozdiel od plne prepojenej neurónovej siete, sú *Konvolučné neurónové siete* (*Convolutional neural network*) (CNN) prispôsobené na prácu s obrazom ako so vstupným signálom. Pracujú na princípe výpočtu príznakov pomocou konvolučnej časti siete a klasifikáciou plne prepojenou sieťou. Ich architektúra je usporiadaná do vrstiev, štandardne po sebe nasledujú:

1. **Konvolučná vrstva** - používa konvolučné filtre na extrakciu príznakov z obrazu.
2. **Aktivačná vrstva** - aplikuje aktivačnú funkciu, často napríklad *Usmernená lineárna jednotka* (ReLU) na každý bod. Zavádzajú do siete nelinearitu a pomáha zachytávať zložitejšie vzory.
3. **Pooling vrstva** - redukuje priestorové rozlíšenie vstupných dát. Tým sa zníži výpočtová náročnosť nasledujúcich vrstiev. Najčastejšie sa používa *max pooling*, teda výber najvyššej hodnoty v okolí.

Väčšinou je použitá kombinácia niekoľko **konvolučných** vrstiev nasledovaných **aktivačnou** vrstvou, po ktorých **pooling** vrstva pripraví redukovaný vstup pre ďalšie vrstvy. Takýchto kombinácií nasleduje niekoľko, posledná je pripojená na plne prepojenú neurónovú siet. Konvolučná časť slúži na vyhľadávanie príznakov, plne prepojená časť na klasifikáciu pomocou nich.

Rozmery vrstiev závisia na tvare vstupného signálu. V prípade šedotónového obrazu je každý bod obrazu reprezentovaný jednou skalárhou hodnotou, v prípade farebného záleží počet hodnôt na pixel na farebnom modeli. Klasifikáciou z farebného obrazu získavame väčšie množstvo informácií, ale je vyžadovaný vyšší výkon, nakoľko



Obr. 3.1: Príklad architektúry konvolučnej neurónovej siete

sa konvolučné filtre aplikujú po jednom na každú zložku obrazu (teda v prípade RGB modelu na červenú, zelenú a modrú zložku samostatne).

Výstup z poslednej vrstvy plne prepojenej časti siete je výstupom z celej CNN. Výstupy klasifikačnej neurónovej siete s úlohou klasifikovať viacero tried, mávajú formáty:

- Jeden výstup, ktorého hodnota určuje triedu, do ktorej sú vstupné dátá sieťou priradené.
- Vektor výstupov s tolkými hodnotami, na koľko tried je naučená sieť klasifikovať. Hodnoty výstupov môžu reprezentovať:
 - Ohodnotenie triedy (score-based classification). V tomto prípade čísla na výstupe nemajú samostatne význam, ale ako predpoveď je určená trieda s najvyšším skóre.
 - Pravdepodobnosť zaradenia do danej triedy. Tá je získaná použitím výstupnej aktivačnej vrstvy typu *softmax*, ktorá spôsobí že každý z výstupov má hodnotu od 0 do 1 a súčet všetkých výstupov je 1.

4 Detekcia a klasifikácia v jednom kroku

4.1 R-CNN

Konvolučné neurónové siete založené na regiónoch (*Region-based CNN*) (R-CNN) sú skupina modelov strojového učenia založená na CNN. Ich cieľom je nájdenie a klasifikovanie objektov v obraze. Ich výstupom je množina rámčekov ohraničujúcich objekty a im pridelené triedy.

R-CNN pracujú v niekoľkých etapách:

1. **Selektívne vyhľadávanie** - vstupný obraz je spracovaný a sú extrahované *Regióny záujmu* (*Regions of Interest*) (RoI), teda orámované oblasti obrazu, ktoré by mohli obsahovať objekty alebo ich časti. Počet takto navrhovaných oblastí môže dosahovať niekoľko tisíc.
2. **Extrakcia príznakov** - každý RoI je predložený ako vstup pre naučenú CNN, ktorá vyprodukuje príznakový vektor pre každý RoI.
3. **Klasifikácia** - pomocou skupiny modelov *Metódy podporných vektorov* (*Support Vector Machine*) (SVM) je podľa príznakov extrahovaných CNN, klasifikovaný každý RoI do jednej z určených tried alebo ako pozadie - teda nepatriaci do žiadnej triedy.
4. **Regresia ohraničení** - konečný krok, zvyšujúci presnosť orámovania objektov. Používa sa pri ľom naučený model nezávislý na mierke nazývaný *bounding box regressor*. Jeho výstup je štvorrozmerný, skladá sa z polohy a rozmerov ohraničujúceho obdĺžnika.

R-CNN sú efektívne, no náročné na výpočtový výkon. Z toho dôvodu boli vyvinuté rýchlejšie varianty *Fast R-CNN* a *Faster R-CNN*.

Fast R-CNN aplikuje CNN na celý vstupný obraz a extrahuje z neho mapu príznakov. Až na výslednú mapu aplikuje takzvaný RoI *pooling*, ktorý extrahuje príznaky pre každý RoI, pomocou okna pevnej veľkosti. Zvyšok modelu pracuje podobne ako R-CNN, s využitím plne prepojenej siete na klasifikáciu a generovanie orámovaní.

Faster R-CNN nadväzuje na *Fast R-CNN* nahradením selektívneho vyhľadávania modelom typu *Siet návrhu regiónov* (*Region Proposal Network*) (RPN). Vstupný obraz prechádza predučenou CNN, sú extrahované príznaky. RPN využije nájdené príznaky aby určila, kde sa nachádzajú potenciálne objekty v obraze. V konečnom kroku sú klasifikované navrhnuté ohraničené oblasti podľa príznakov extrahovaných v predošлом kroku. *Faster R-CNN* je dostatočne rýchla na použitie v reálnom čase.

4.2 YOLO

You only look once (YOLO) je populárny algoritmus detekcie, ktorý na rozdiel od R-CNN rozdeluje obraz do mriežky a postupne aplikuje klasifikátor. Postupuje nasledovne:

1. Vstuný obraz je rozdelený na mriežku $S \times S$, ktorej veľkosť závisí na verzii YOLO.
2. V každej bunke mriežky je predikovaný určitý počet ohraničených oblastí a im priradené skóre dôveryhodnosti. Tie značia úroveň istoty, že oblasť obsahuje objekt a že dané ohraničenie je správne.
3. Je použitých viacero klasifikátorov. YOLO predikuje viacero oblastí za každú bunku. Počas trénovalia je jednému z prediktorov udelená zodpovednosť za predikovanie daného objekto podľa toho, ktorý z nich dosahuje najväčší *Pomer prieniku voči zjednoteniu* (*Intersection over Union*) (IOU).
4. Každá bunka predikuje rozdelenie pravdepodobnosti všetkých tried.

Hlavnou výhodou YOLO je rýchlosť. Všeobecne pracuje rýchlejšie a s menším požadovaným výpočtovým výkonom. Na druhú stranu má nevýhody spôsobené obmedzením veľkosti orámovaných oblastí. Presnosť YOLO býva nižšia pri primalých objektoch, pre použitia vyžadujúce vyššiu presnosť sa môže javiť ako výhodnejšie použiť iné modely.

5 Návrh usporiadania hardvérových komponentov

Táto kapitola sa zameriava na návrh a implementáciu prenosného zariadenia pre detekciu, klasifikáciu a trasovanie lietajúcich objektov. Cieľom návrhu je vytvorenie zariadenia s možnosťou nasadenia v čo najrôznejších podmienkach. Jednou z podmienok je teda jeho prenosnosť, možnosť napájania z akumulátora a nezávislosť na iných zariadeniach pri jeho nastavení a používaní.

5.1 Jednodoskový počítač

Jedným z najčastejšie používaných jednodoskových počítačov je rada *Raspberry Pi*. V čase návrhu bol najnovší model *Pi 4 model B*. Tento model disponuje:

- 1.5GHz ARM Cortex-A72 procesorom so 4 jadrami,
- do 8 Gb pamäte RAM,
- rozhraním HDMI s možnosťou pripojenia dvoch monitorov,
- 4 USB, z toho 2 USB 3.0,
- MIPI CSI konektor pre pripojenie kamery.

Ďalšou možnosťou je použitie jednodoskového počítača špeciálne navrhnutého na použitie v aplikáciách počítačového videnia, či všeobecne umelej inteligencie (Nvidia Jetson Xavier, Google Coral Dev Board, Rock Pi, Nvidia Jetson Nano, a podobné). *Raspberry Pi* má v porovnaní nižší odber, je kompaktnejšie, jednoduché na použitie a má dobrú kompatibilitu s operačnými systémami. Dokumentácia a knižnice pre *Raspberry Pi* sú jednoducho dostupné a zariadenia majú garantovanú dlhodobú podporu softvéru aj hardvéru. Z týchto dôvodov bolo *Raspberry Pi* zvolené na použitie v zariadení.

5.2 Kamera

Raspberry Pi dovoľuje pripojenie kamery cez MIPI konektor. Je teda najjednoduchšie použiť kamerový modul vyrobený konkrétnie pre *Raspberry Pi*. Najaktuálnejší oficiálny *Raspberry Pi* kamerový modul v čase návrhu je 12Mpx *Raspberry Pi Camera 3* so senzorom Sony IMX708 a automatickým ostrením. Parametre kamery a objektívu sú:

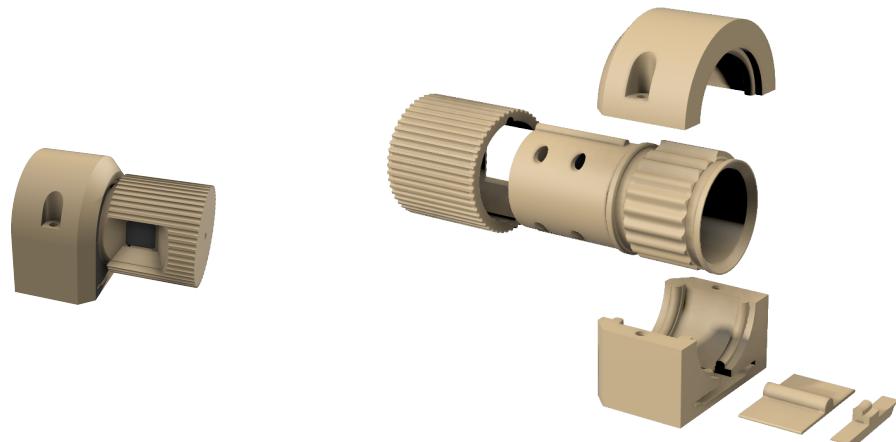
- **Ohnisková vzdialenosť:** 4,74 mm
- **Horizontálne zorné pole** 66 °
- **Vertikálne zorné pole** 41 °

5.3 Ovládacie prvky

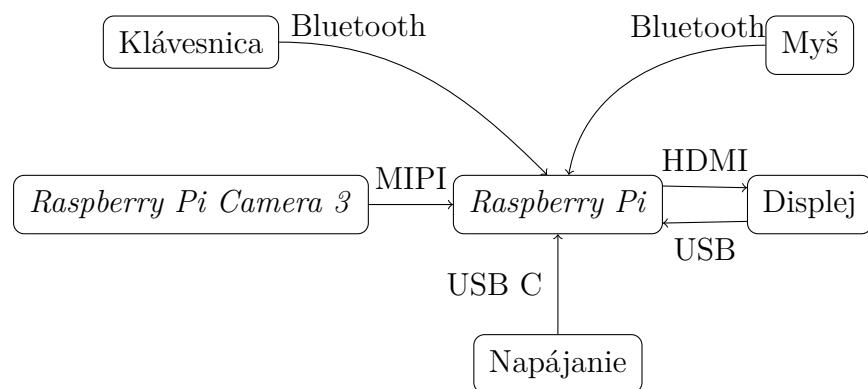
Ako najlepší ovládaci a zobrazovací prvok bol z hľadiska prenosnosti a kompaktnosti zvolený dotykový displej, konkrétnie 7 palcový IPS displej od firmy Waveshare. Táto veľkosť je postačujúca na zobrazovanie výsledkov detekcie aj ovládanie zariadenia. Rozhranie HDMI značne zjednodušuje prepojenie a inštaláciu displeja.

5.4 Uchytenie kamery

Praktickou výhodou prenosného zariadenia je možnosť jeho umiestnenia na statív. Kedže cielom je kamerou mierit smerom nahor na oblohu, bolo by užívateľské rozhranie pri nesprávnej navrhnutej orientácii kamery a displeja natočené neprakticky nadol. Namiesto pevnej konfigurácie uhla medzi displejom a kamerou bolo navrhnuté pohyblivé uchytenie kamery v kryte zariadenia, vyrobiteľné 3D tlačou.



Obr. 5.1: návrh nastaviteľného uchytenia kamery



Obr. 5.2: bloková schéma zariadenia

6 Tvorba datasetu

Správne naučenie modelov a ich schopnosť generalizovať z veľkej miery záleží na kvalite použitého datasetu (množiny dát). Tvorba datasetu je teda klúčový krok pri práci so systémami strojového videnia.

Cieľom pri tvorbe datasetu je zabezpečiť dostatočnú kvalitu a správnosť dát, dosiahnuť početne podobné zastúpenie každej triedy a správne reprezentovať rôznorodosť dát v rámci jednotlivých tried. Je vhodné, aby boli vzorky získané pri rôznych podmienkach, podľa plánovaného použitia výsledného modelu.

6.1 Získavanie dát

Kedže táto práca nadvázuje na prácu [1], tvorí dataset v nej vytvorený základ pre dataset použitý v tejto práci. Navyše bol rozšírený o niekoľko ďalších open source datasetov a vlastných anotovaných fotografií.

Na tvorbu datasetu bol použitý program *roboflow*, ktorý umožňuje načítavanie existujúcich datasetov, nahrávanie vlastných fotografií, anotáciu, spájanie dostupných datasetov, augmentáciu datasetu a generovanie anotačných súborov rôznych formátov.

Po pridaní všetkých častí datasetu a kontrole správnosti anotácie boli počty inštancií jednotlivých tried:

- **vták:** 9047
- **dron:** 1190
- **helikoptéra:** 509
- **hmyz:** 249
- **lietadlo:** 845

6.2 Úprava, rozdelenie a rozšírenie datasetu

Niektoré triedy sú reprezentované výrazne menším množstvom jednotlivých inštan- cii, čo je spôsobené nižšou dostupnosťou dát. Väčšina fotografií vtákov zo vzdialenosťi, pri ktorej by mali byť detekované, pochopiteľne obsahuje niekoľko desiatok jedincov, trieda vták je teda nadmerne zastúpená. Je možné tento problém čiastočne vyriešiť augmentáciou datasetu, čo program *roboflow* umožňuje priamo pri genero- vaní novej verzie. V budúcnosti je vhodnejším riešením ďalšie rozšírenie datasetu o snímky s objektami aktuálne nedostatočne reprezentovaných tried.

Aby bolo zabránené preučeniu modelu, musia byť dátá, na ktorých je model učený, rôzne od vylidačných a testovacích. Anotované obrazové dátá boli rozdelené do skupín s počtom:

- **trénovanie:** 2367
- **validácia:** 572
- **testovanie:** 290

Snímky boli predspracované upravením veľkosti na 640×640 px a boli aplikované augmentácie, ktorými bol počet inštancií v trénovacej množine rozšírený na 7101:

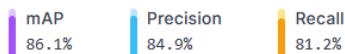
- Rotácia -30° až $+30^\circ$
- Zkosenie $\pm 15^\circ$ vertikálne aj horizontálne
- Posun odtieňu -35° až $+35^\circ$
- Zmena saturácie -25% až $+25\%$
- Zmena expozičie -20% až $+20\%$
- Rozmazanie do 0,8 px

6.3 Testovanie použiteľnosti datasetu

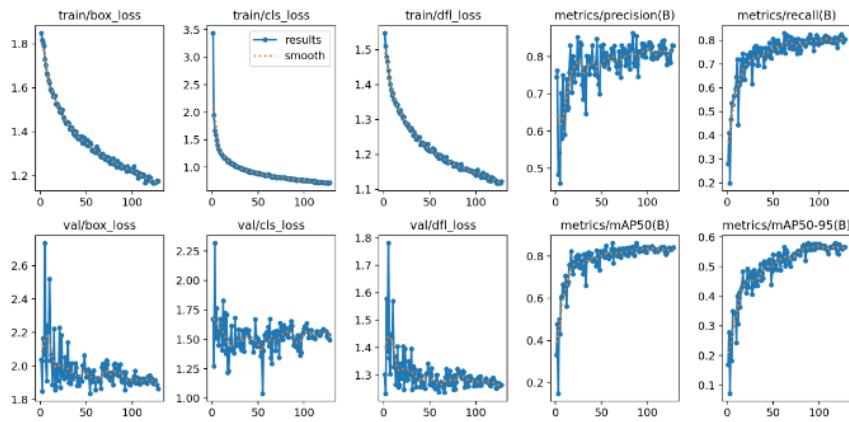
Použiteľnosť modelu bola testovaná naučením modelu typu *Roboflow 3.0 Object Detection* priamo v aplikácii *roboflow*, ktorý dosahoval úroveň presnosti 86.1 % pri zvolenom variante modelu *fast*. Pri dlhšom učení modelu určeného na reálne použitie sa dá predpokladať, že dosiahnutá presnosť bude vyššia.

Obrázok 6.1 zobrazuje metriky presnosti naučeného modelu vypočítané v programe *roboflow*, tak ako sa v ňom zobrazujú. Vypočítané metriky sú:

- **Stredná priemerná presnosť (Mean Average Precision) (mAP)**
priemer priemerných presností za všetky triedy.
- **Presnosť (Precision)** = $\frac{TP}{TP+FP}$, značí mieru správnej predikcie, v momente, keď model nejakú hodnotu predikuje.
- **Senzitivita (Recall)** = $\frac{TP}{TP+FN}$, predstavuje pomer počtu správnych predikcií ku počtu skutočne správnych hodnôt danej triedy.

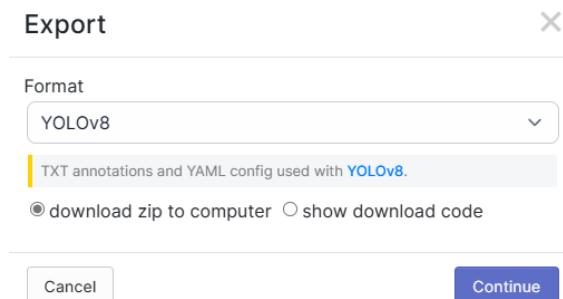


Obr. 6.1: Metriky presnosti naučeného modelu v programe roboflow



Obr. 6.2: Grafické zobrazenie priebehu učenia modelu roboflow

Výsledná vygenerovaná verzia datasetu je z programu exportovateľná v niekoľkých formátoch, napríklad pre učenie siete YOLO alebo pre použitie s knižnicou *tensorflow*. Pri exporte je stiahnutý komprimovaný priečinok obsahujúci obrazové dátá a anotáciu.



Obr. 6.3: Exportovanie vytvoreného datasetu

7 Návrh softvéru

Pri návrhu a tvorbe aplikácie systému detekcie lietajúcich objektov bol kladený dôraz na:

- spustiteľnosť a použiteľnosť na počítači,
- ovládateľnosť dotykovým panelom,
- modulárnosť a rozšíriteľnosť,
- spätná väzba a grafické zobrazenie výsledkov užívateľovi,
- možnosť používať, testovať a porovnávať rôzne metódy detekcie, klasifikácie a trasovania.

7.1 Použitý jazyk a knižnice

Použitím programovacieho jazyka *python* bola získaná kompatibilita zo zariadeniami a operačnými systémami, ktoré podporujú interpretér jazyka *python*, teda hlavne *Raspberry Pi* s nainštalovaným operačným systémom *Raspberry Pi OS* a osobné počítače, čo uľahčuje vývoj aplikácie a umožňuje jeho použitie na ľubovoľnom počítači.

Na tvorbu grafického rozhrania bola použitá knižnica *Rozhranie toolkitu Tcl/Tk (Tk interface)* (tkinter). Tá umožňuje vytvárať grafické používateľské rozhranie priamo z kódu v jazyku *python*. Obsahuje tiež objekty časovačov, vďaka ktorým je možné jednoducho a spoloahlivo ovládať kontinuálny beh aplikácie. Užívateľské rozhranie vytvorené knižnicou tkinter je kompatibilné s ovládaním klávesnicou a myšou, aj dotykovým displejom.

Knižnica *OpenCV* obsahuje implementáciu množstva algoritmov spracovania obrazu a počítačového videnia. Umožňuje základné operácie s videom aj jednotlivými snímkami ako načítanie, ukladanie a zobrazovanie, pričom podporuje rôzne formáty obrazových dát. Sofistikanejšie úlohy, ktoré knižnica rieši, zahrňujú detekciu, klasifikáciu, sledovanie pohybu. Knižnica *OpenCV* je kompatibilná s viacerými programovacími jazykmi, vrátane jazyku *python*.

7.2 Štruktúra aplikácie

Aplikácia je zložená z modulov reprezentujúcich jednotlivé implementované metódy detekcie, klasifikácie či trasovania alebo riadiacich modulov ovládajúcich tok dát medzi modulmi. Základnými typmi modulov tvoriacich aplikáciu sú:

- **Hlavná aplikácia** - existuje len jeden variant, je zodpovedný za ovládanie systému, užívateľské rozhranie a spúštanie ostatných modulov. Pre každý z ďalších jednotlivých typov modulov je hlavnému modulu priradený zoznam

existujúcich modulov daného typu, z ktorých si užívateľ môže vybrať ten aktívny, ktorý sa aktuálne použije na daný účel.

- **Poskytovateľ obrazu** - slúži na získavanie obrazových dát. Každý cyklus začína vyžiadaním nového obrazu od aktuálne zvoleného poskytovateľa obrazu hlavným modulom. Existujú dva varianty:
 - **Video** - pri spustení aplikácia načíta všetky videá uložené v priečinku `./videos/` a pridá každému jeden objekt reprezentujúci dané video ako zdroj obrazu.
 - **Kamera** - ak je v systéme nainštalovaná *python* knižnica pre ovládanie kamery, je do zonamu poskytovateľov obrazu pridaný objekt reprezentujúci kameru.
- **Detektor objektov** - po získaní aktuálneho obrazu je aplikovaný zvolený algoritmus detekcie. Detektoru je predaný obraz, jeho výstupom je zoznam rámikov oblastí v ktorých sa potenciálne nachádza objekt. V prípade zvolenia metódy detekcie a klasifikácie v jednom kroku je preskočený výber klasifikátora a výstupom je aj klasifikácia.
- **Klasifikátor** - ak je zvolený algoritmus detekcie, sú oblasti s potenciálnym obsahom objektu predané klasifikátoru. Klasifikátor vráti hlavnému modulu zaradenie objektov do tried, prípadne označenie oblasti za pozadie. V hlavnej aplikácii je následne graficky zobrazené orámovanie a klasifikácia objektu.
- **Sledovač (Tracker)** - na ušetrenie výkonu je možné aplikovať detekciu a klasifikáciu len raz za určitý počet snímkov, a už detekované a klasifikované objekty sledovať pomocou sledovačov z knižnice *OpenCV*.
- **Tester výsledkov** - v prípade, že je zvolené video ako zdroj obrazu a v priečinku `./video_annotation/` je pridaný *Súbor s čiarkou rozdelenými hodnotami (Comma Separated Values)* (csv) s rovnakým menom ako video obsahujúci správne orámovanie, je spustený testovací modul, ktorý ohodnocuje detekciu a klasifikáciu podľa určenej správnej hodnoty zo súboru.

Jednotlivé moduly sú vytvárané dedením od abstraktnej triedy predstavujúcej rozhranie modulu s aplikáciou. Pre každý z typov modulu je pripravená jedna abstraktná rodičovská trieda obsahujúca niektoré zdieľané premenné, ako je názov modulu a funkcie, ktoré musí modul implementovať. Povinnou funkciou je napríklad funkcia vyžiadania si ďalšieho snímku pre modul poskytovateľa obrazu.

7.3 Implementácia

V aplikácii je ako modul detekcie a klasifikácie aktuálne implementovaný model YOLO, pričom aplikácia automaticky pridá detekované naučené modely. Tie musia

byť uložené v priečinku `./yolo/[model]/`, kde je `[model]` nahradený názvom modelu. V tomto priečinku je uložený model samotný v súbore `model.onnx` a csv súbor so zoznamom názvov a farebných označení tried vo formáte:

`"názov", "červená", "zelená", "modrá"`

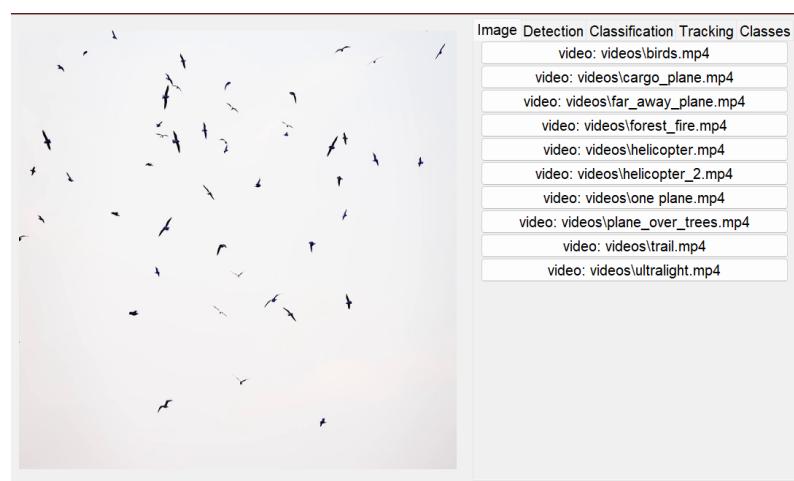
Tento súbor má názov `classes.csv`. Aplikácia načíta všetky nájdené modely spolu s popisom tried a vloží ich do zoznamu detektorov.

7.4 Návrh grafického rozhrania

Grafické užívateľské rozhranie (*Graphical User Interface*) (GUI) umožňuje intuitívne ovládanie aplikácie a zobrazovanie výsledkov detekcie, klasifikácie a sledovania. Ovládanie jednotlivých modulov aplikácie je rozdelené do záložiek. Každá záložka umožňuje zvoliť, ktorá z inštancii modulu každého z typov sa aktuálne používa.

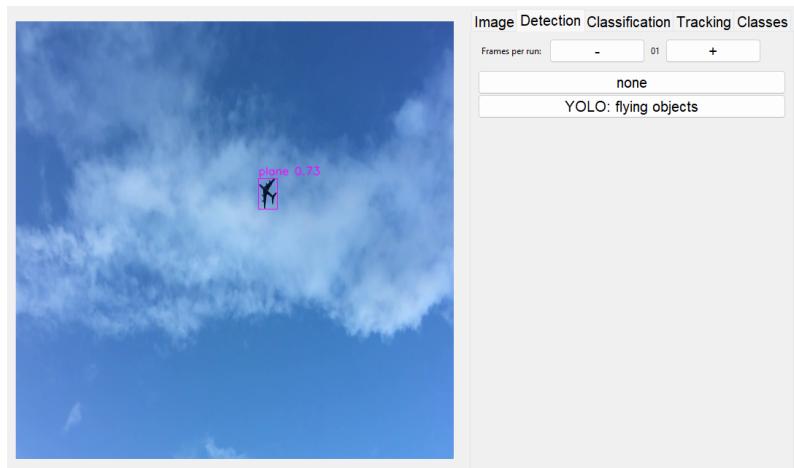
V ľavej časti obrazovky je zobrazený aktuálny snímok s výstupom detekcie a klasifikácie. Zobrazujú sa v ňom orámovaná detekované objektov, ich priradenie do tried a hodnotenie dôveryhodnosti klasifikácie. Ak existuje súbor s anotáciou správneho orámovania a klasifikácie pre aktuálne prehrávané video, sú zobrazené aj tie.

V prvej záložke je užívateľovi umožnené zvoliť zdroj obrazu. Ak je aplikácia spustená na *Raspberry Pi* a je nainštalovaná kižnica ovládania kamery, je do zoznamu zdrojov pridaná kamera. Pri spustení aplikácie sú načítané videá z priečinku `./videos` a sú zobrazené v zozname na záložke zdrojov spolu s kamerou. Kliknutím na názov zdroja je zdroj zvolený a užívateľovi sa z neho začne zobrazovať obraz.



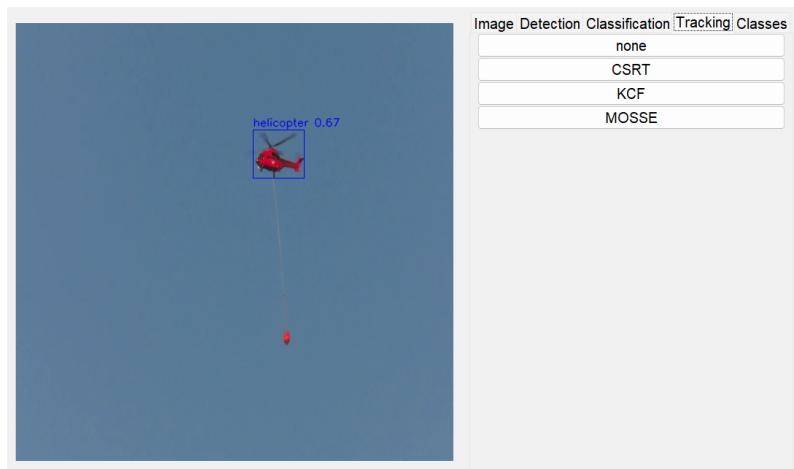
Obr. 7.1: Záložka na výber poskytovateľa obrazu

V druhej záložke výberu detektora si užívateľ vyberie metódu detekcie alebo metódu súčasnej detekcie a klasifikácie, pri ktorej nie je použiteľný výber klasifikátora na ďalšej záložke. Zároveň je na tejto záložke nastaviteľná períoda spúšťania detekcie. Medzi snímkami, na ktorých je detekcia spustená, sa buď zobrazované ohraničenie nemení alebo je akutalizované sledovačom.



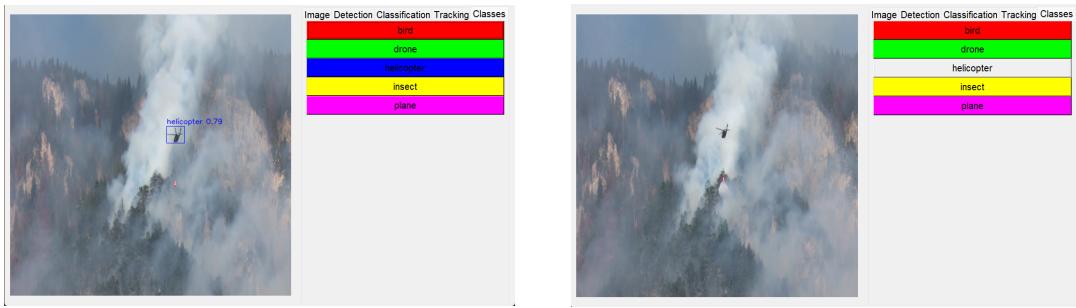
Obr. 7.2: Záložka na výber metódy detekcie objektov

V záložke sledovania objektov je umiestnený výber zo sledovačov implementovaných v knižnici *OpenCV*. Sledovač je použitý na snímkoch medzi spustením detektcie.



Obr. 7.3: Záložka na výber metódy sledovania objektov

Posledná záložka obsahuje menu s možnosťou zapnúť a vypnúť zobrazenie detekcie konkrétnych tried objektov.



Obr. 7.4: Záložka na výber detekovaných tried

Záver

V tejto práci bol popísaný návrh systému detekcie, trasovania a klasifikácie lietajúcich objektov. V kapitolách teoretickej časti práce boli vysvetlené teoretické základy postupov a metód detekcie objektov v obraze, ich popisu a následného zaradenia do tried. Dôraz bol kladený na rozdiely medzi metódami, ich výhody a nevýhody.

Nasledujúce kapitoly sa venovali samotnému návrhu systému. Bol navrhnutý a odôvodnený výber hardvérových komponentov a ich prepojenie. Pre jednoduchosť, spoľahlivosť a dostatočný výkon bola zvolená platforma *Raspberry Pi* s dotykovým panelom ako užívateľským rozhraním a kamerou od spoločnosti *Raspberry*.

Ďalej bol zdokumentovaný návrh databázy anotovaných snímkov určených na trénovanie klasifikátorov. Pri vytváraní bol použitý program *roboflow*. Výsledná databáza vznikla spojením datasetu z práce [1], niekoľkých voľne dostupných datasetov spolu s osobne získanými a anotovanými fotografiemi. Správnosť takto získanej databázy bola overená naučením modelu v systéme *roboflow* v nastavení na odskúšanie (fast). Na testovacej množine tento model dosahoval úroveň mAP 86,1 %.

Nakoniec bol popísaný softvér pre navrhovaný systém, jeho štruktúra a implementácia jeho častí. Spomenuté boli nástroje a knižnice použité pri jeho vývoji. Softvér zahrňuje moduly pre výber zdroja dát, detekciu objektov, ich klasifikáciu, sledovanie a kontrolu presnosti porovnaním s anotáciou. Systém umožňuje výber aktuálne používanej metódy a porovnanie medzi metódami. Kedže bol použitý programovací jazyk *python* a knižnice dostupné na väčšine zariadení a operačných systémoch, je možné program spustiť aj na bežnom počítači.

Ciele pokračovania tejto práce zahŕňajú ďalšie rozširovanie datasetu, implementáciu viacerých metód detekcie a klasifikácie a vyhodnotenie a porovnanie prístupov.

Literatúra

- [1] JUREČKA, Tomáš. Detekce a klasifikace létajících objektů. Brno, 2021. Diplomová práca.
- [2] JANOUŠEK, Jiří, Jan NOVOTNÝ, Petr MARCOŇ a Anna SIRUCKOVA. Algorithms for Flying Object Detection. In: 2018 Progress In Electromagnetics Research Symposium (PIERS-Toyama) [online]. Toyama, Japonsko, 2018, s. 782-783 [cit. 2024-01-03]. ISBN 978-4-8855-2316-8. ISSN 1559-9450. Dostupné z: doi:10.23919/PIERS.2018.8598196
- [3] MORSE, Bryan. Lecture 4: Thresholding [online]. Brigham, 1998 [cit. 2024-01-03]. Dostupné z: https://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/MORSE/threshold.pdf. Brigham Young University.
- [4] MORSE, Bryan. Lecture 9: Shape Description (Regions) [online]. Brigham, 1998 [cit. 2024-01-03]. Dostupné z: https://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/MORSE/region-props-and-moments.pdf. Brigham Young University.
- [5] BUHL, Nikolaj. Image Thresholding in Image Processing. Encord [online]. 2023 [cit. 2024-01-03]. Dostupné z: <https://encord.com/blog/image-thresholding-image-processing/>
- [6] SAMINA. EDUCATIVE, INC. What is Canny edge detection? EDUCATIVE, INC. Educative [online]. Educative, 2023 [cit. 2024-01-03]. Dostupné z: <https://www.educative.io/answers/what-is-canny-edge-detection>
- [7] OPEN SOURCE COMPUTER VISION. How to Use Background Subtraction Methods [online]. 2023 [cit. 2024-01-03]. Dostupné z: https://docs.opencv.org/4.x/d1/dc5/tutorial_background_subtraction.html
- [8] LEUNG, Kenneth. How to Easily Draw Neural Network Architecture Diagrams [online]. In: . [cit. 2024-01-03]. Dostupné z: <https://towardsdatascience.com/how-to-easily-draw-neural-network-architecture-diagrams-a6b6138ed875>
- [9] LI, Erbo, Yazhou HUANG, Dong XU a Hua LI. Shape DNA: Basic Generating Functions for Geometric Moment Invariants [online]. Cornell University, 2017, 3-8 [cit. 2024-01-03]. Dostupné z: doi:10.48550/arXiv.1703.02242

Zoznam symbolov a skratiek

| | |
|----------------|---|
| ReLU | Usmernená lineárna jednotka |
| CNN | Konvolučné neurónové siete (Convolutional neural network) |
| R-CNN | Konvolučné neurónové siete založené na regiónoch (Region-based CNN) |
| RoI | Regióny záujmu (Regions of Interest) |
| SVM | Metódy podporných vektorov (Support Vector Machine) |
| RPN | Siet návrhu regiónov (Region Proposal Network) |
| YOLO | You only look once |
| IOU | Pomer prieniku voči zjednoteniu (Intersection over Union) |
| tkinter | Rozhranie toolkitu Tcl/Tk (Tk interface) |
| csv | Súbor s čiarkou rozdelenými hodnotami (Comma Separated Values) |
| GUI | Grafické užívateľské rozhranie (Graphical User Interface) |
| mAP | Stredná priemerná presnosť (Mean Average Precision) |
| KNN | K najbližších susedov (K Nearest Neighbours) |

Zoznam príloh