

Dron ovládaný gesty

Michal Průšek

Školitel: Ing. Adam Novozámský, Ph.D.

České vysoké učení technické
Fakulta jaderná a fyzikálně inženýrská

8. září 2023

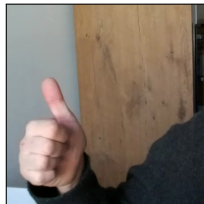
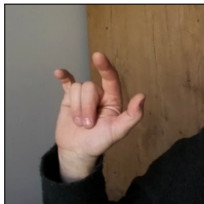
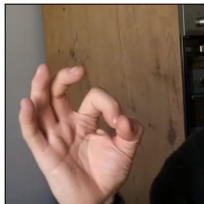
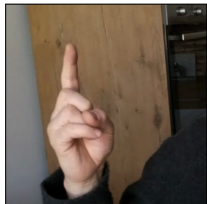
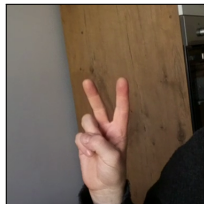
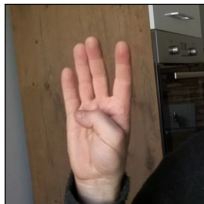
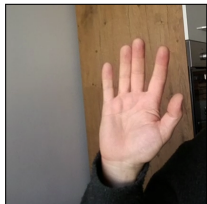


- 1 Cíle práce
- 2 Přístup založený na metodách zpracování obrazu
 - Zpracování obrazu obličeje
 - Zpětná projekce histogramu
 - Kontura ruky
 - Vyhodnocení segmentace
 - Klasifikace gesta naší metodou
 - Výsledky klasifikace
- 3 Přístup založený na metodě neuronových sítí
- 4 Videoukázka
- 5 Odpovědi na otázky oponenta



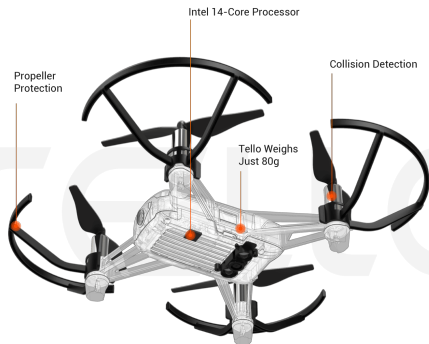
Cíle práce

- 1 vyřešit ovládání dronu a přístup k obrazovým datům
- 2 navrhnout metodu rozpoznávání gest
- 3 sestavit dataset gest
- 4 ovládat dron pomocí předvedeného gesta



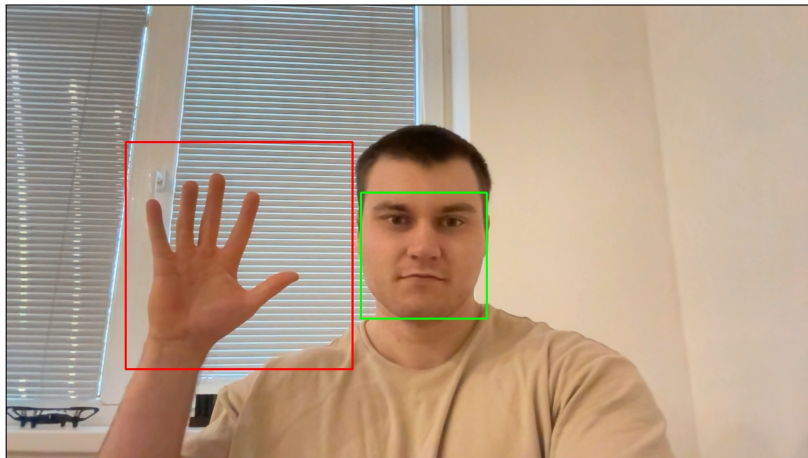
Ukázka dronu

- 720p - 30 FPS kamera
- cca 10 × 10 cm
- 81 gramů



model dronu DJI Tello

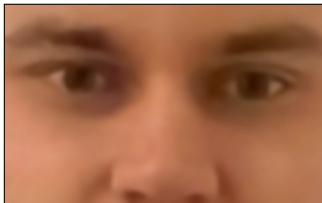




rozpoznání obličeje pomocí MediaPipe



Postup segmentace



Ořez obrazu obličeje



ořez obličeje



Otsuova metoda - maximalizace mezitřídního výběrového rozptylu:

$$k^* = \operatorname{argmax}_{k \in \{0, \dots, I-1\}} \sigma_B^2(k) \quad (1)$$



Otsu



ROI Otsu: 5 × 4



bitový součet



Morfologické transformace a zakrytí očí

Eroze - relace mezi O a S :

$$O \ominus S := \{(a, b) \in O \mid (a, b) + (c, d) \in O, \forall (c, d) \in S\} \quad (2)$$



bitový součet Otsu a
ROI Otsu



eroze bitového součtu
elementem 5×5



zakrytí očí s využitím
MediaPipe





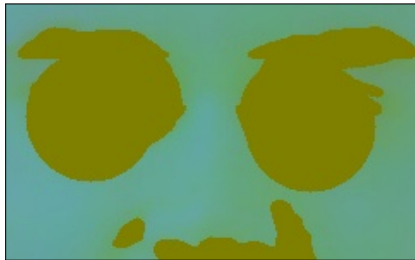
finální podoba obličeje v RGB



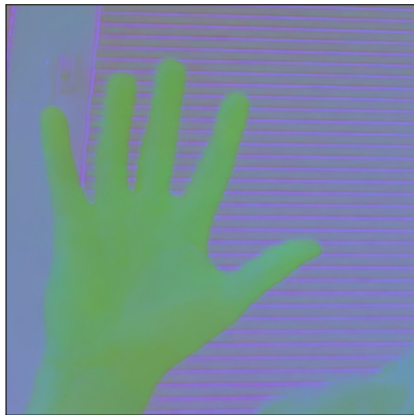
finální podoba obličeje v YCbCr



Zpětná projekce histogramu



finální podoba obličeje v YCbCr



vyříznutý obraz ruky v YCbCr



Výsledek zpětné projekce histogramu

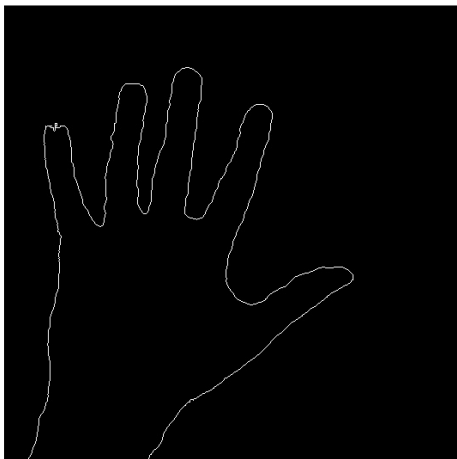


výsledek zpětné projekce
(normalizovaný 8-bitový obraz)



výsledek zpětné projekce po
prahování ($P = 7$)

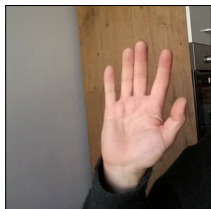




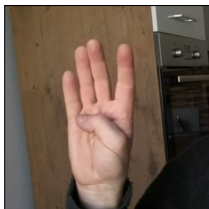
kontura binárního obrázku ruky



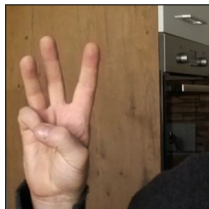
Připomenutí tříd gest



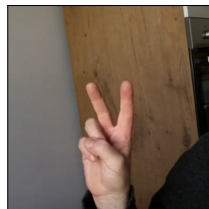
"Five"



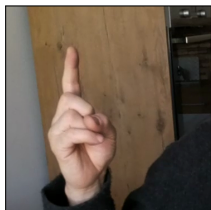
"Four"



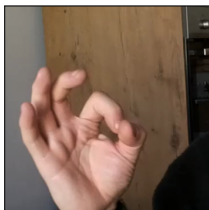
"Three"



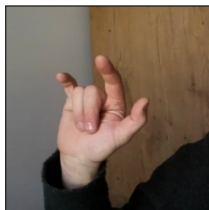
"Two"



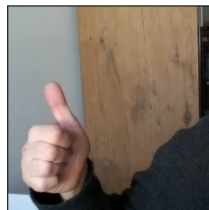
"One"



"Circle"



"Rocknroll"



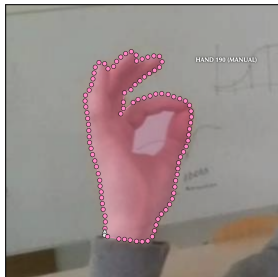
"Thumb"



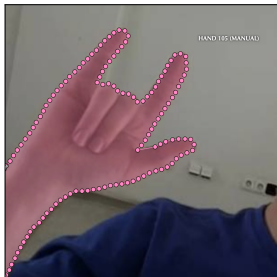
Manuální anotace datasetu ve CVAT

- CVAT = Computer Vision Annotation Tool
- 1602 manuálně anotovaných obrázků gest
- 8 tříd gest (cca 200 obrázků gest z každé třídy)
- 10 subjektů, různá pozadí, různé světelné podmínky
- dostupné na:

<https://www.kaggle.com/datasets/adamnovozmsk/hands>



Referenční polynom
gesta "Circle"



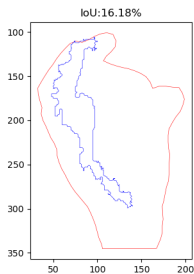
Referenční polynom
gesta "Rocknroll"



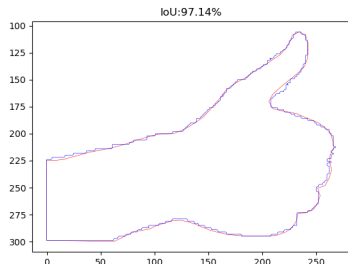
Referenční polynom
gesta "Circle"

Vyhodnocení pomocí metody IoU

- IoU = Intersection over Union
- obsah průniku polynomů / obsah sjednocení polynomů
- nejvyšší medián IoU: **88,57%** pro $P = 7$



Špatná segmentace gesta na obrázku



Úspěšná segmentace gesta na obrázku



Popis datasetu:

- 11963 obrázků gest
- 8 tříd gest (cca 1500 obrázků gest z každé třídy)
- 1 subjekt, 3 různá pozadí, různé světelné podmínky
- dostupné na:
<https://www.kaggle.com/datasets/adamnovozmsk/hands>

Zpracování datasetu:

- 1 kontura gesta
- 2 Fourierovy deskriptory - prvních 20



Klasifikace zkoumaného gesta

- 1 určení kontury
- 2 Fourierovy deskriptory - prvních 20
- 3 Eukleidovské vzdálenosti FD aktuálního gesta s FD gest v klasifikačním datasetu
- 4 klasifikace: 5 nejbližších sousedů



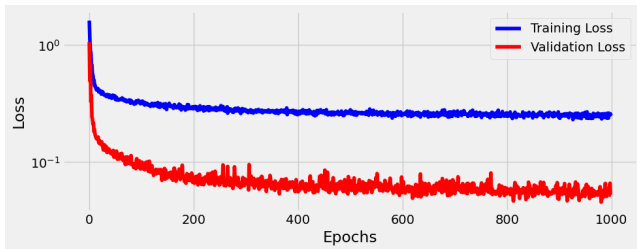
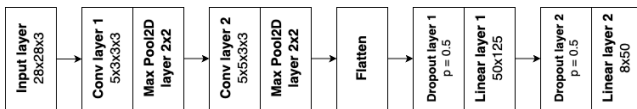
Výsledky klasifikace pomocí naší metody

- poměr rozdělení datasetu: 4:1 (referenční × testovací)
- gesta z testovací množiny klasifikujeme pomocí naší metody (porovnáváme s gesty z referenční)
- úspěšnost klasifikace: **83,67%** (5 nejbližších sousedů, 20 FD)

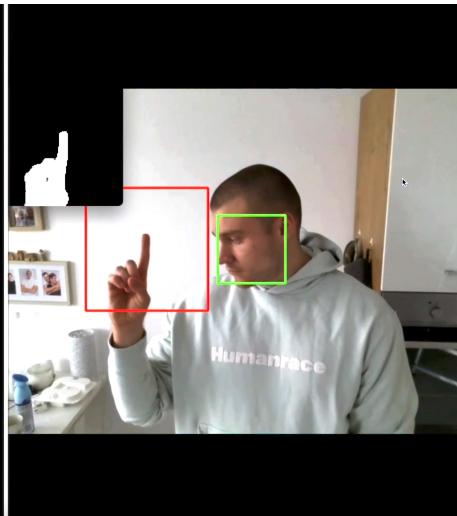


Klasifikace pomocí metody neuronových sítí

- úspěšnost klasifikace na trénovací množině obrázků gest **99,35%**
- úspěšností klasifikace na validační množině **99,03%**



Ukázka funkčnosti dronu



Děkuji za pozornost.

