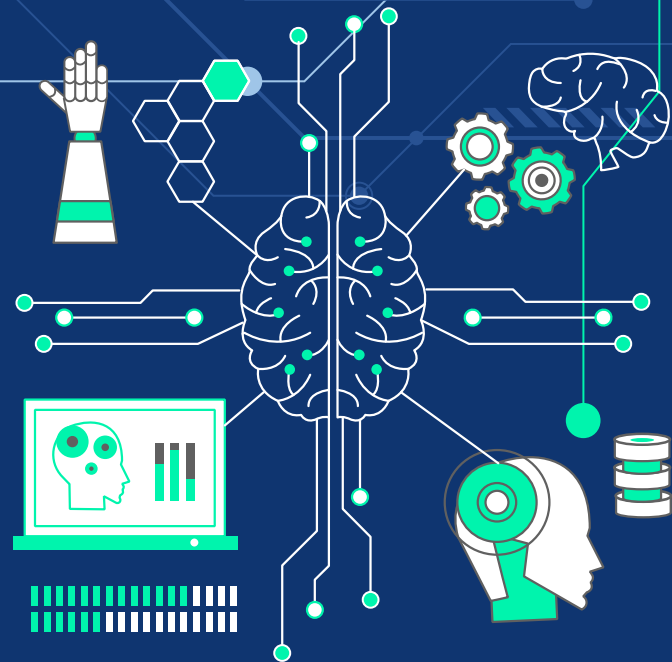




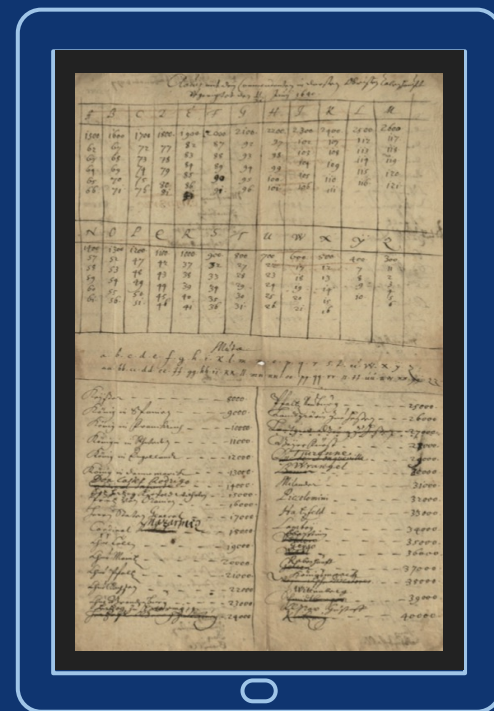
POROVNANIE METÓD UMELEJ INTELIGENCIE NA ROZPOZNÁVANIE RUKOU PÍSANÝCH ČÍSLIC

Ing. Eugen Antal, PhD. – Bc. Filip Mikuš

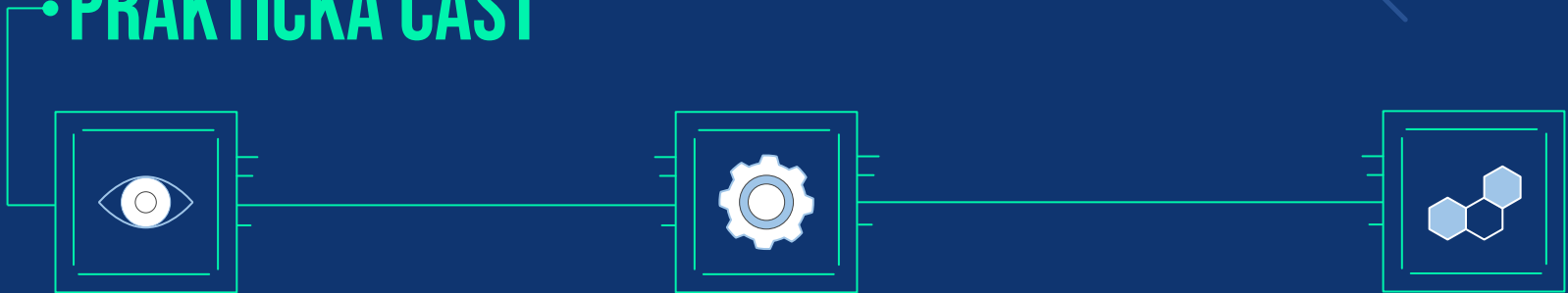


MOTIVÁCIA

- digitalizácia historických dokumentov
- uchovanie dobových dokumentov a listín - zníženie nákladov na konzerváciu, zastavenie degradácie, transkripcia údajov a dát
- interpretácia a dešifrovanie získaných údajov a dát
- práca zapojená do výskumného projektu VEGA 2/0072/20 – Moderné metódy spracovania šifrovaných archívnych dokumentov



• PRAKTICKÁ ČASŤ



ANALÝZA

- detektory a formáty datasetov
- datasety rukou písaných číslíc

IMPLEMENTÁCIA

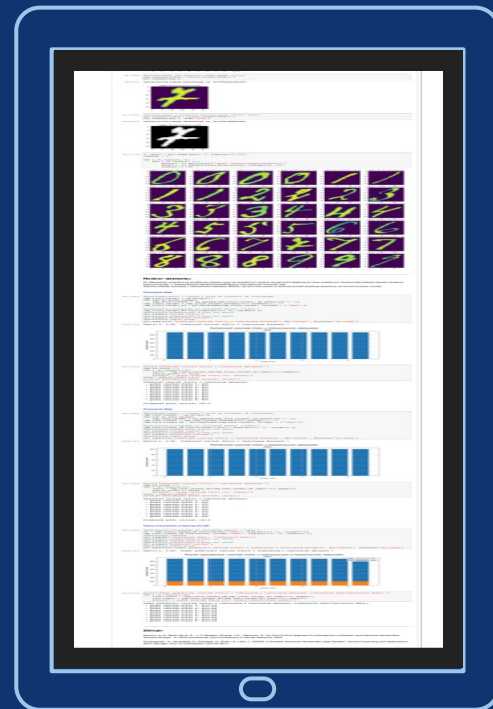
- aplikácia na predspracovanie obrazu
- konverzia formátu datasetu ARDIS
- detektor Mask R-CNN
- detektor YOLO v5

VYHODNOTENIE

- vyčíslenie úspešnosti detektorov
- porovnanie úspešnosti detektorov

ANALÝZA – DATASETY

Meno	Trénovacie vzorky	Testovacie vzorky	Rozmer vzorky	Farebný formát	Formát dát
ARDIS IV	6 600	1 000	28x28	Grayscale	.csv
Chars74K	550		1200x900	Natural	.png, .m
DIDA	252 860		Various	Natural	.jpg
EMNIST	240 000	40 000	28x28	Grayscale	.csv
MNIST	60 000	10 000	28x28	Grayscale	.csv
SEMEION	1 593		16x16	Grayscale	.csv
USPS	7 291	2 007	16x16	Grayscale	.h5

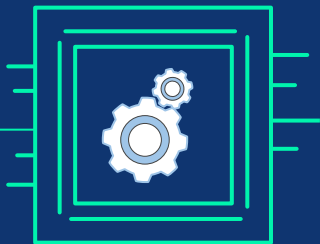


IMPLEMENTÁCIA – APLIKÁCIA NA PREDSPRACOVANIE OBRAZU

- oknové grafické užívateľské rozhranie
- metódy predspracovania: grayscale filter, odšumenie, binarizácia, Canny detekcia hrán, morfologické operácie
- využité knižnice OpenCV a Tkinter

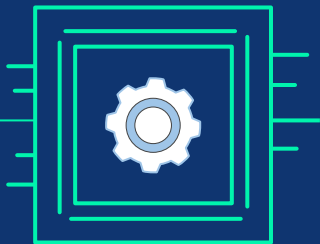


IMPLEMENTÁCIA – DETEKTORY



MASK R-CNN

- dvojúrovňový detektor (RPN, CNN)
- segmentácia detekovaných objektov
- formát datasetu COCO (.json)
- využíva knižnicu TensorFlow 2



YOLO V5

- jednoúrovňový detektor (CNN)
- nízka časová náročnosť detekcie
- formát datasetu YOLO Darknet TXT (.txt)
- využíva knižnicu PyTorch

IMPLEMENTÁCIA – VYUŽITÉ DATASETY

01

ARDIS

- referenčný klasifikačný dataset
- nutnosť konverzie (COCO, YOLO Darknet TXT)

02

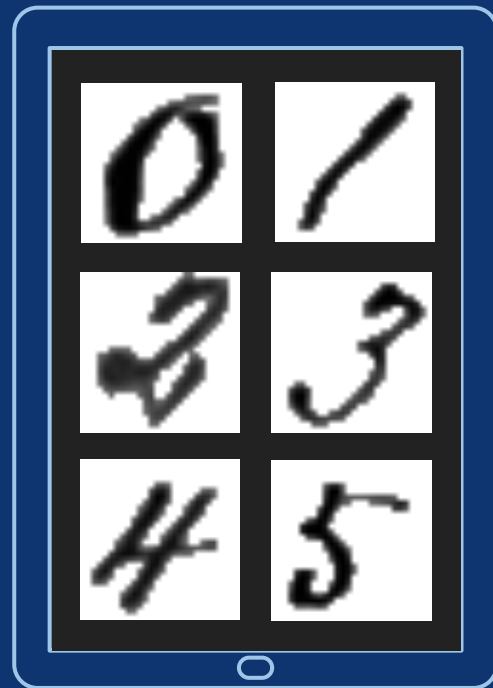
VEGA

- dataset vytvorený v rámci projektu VEGA 2/0072/20
- vzorky zaznamenávané ručne

03

VEGA-CROPPED

- dataset vytvorený rozdelením vzoriek VEGA datasetu
- veľkosť vzorky 128x128 pixelov



IMPLEMENTÁCIA – VYUŽITÉ DATASETY

01

ARDIS

- referenčný klasifikačný dataset
- nutnosť konverzie (COCO, YOLO Darknet TXT)

02

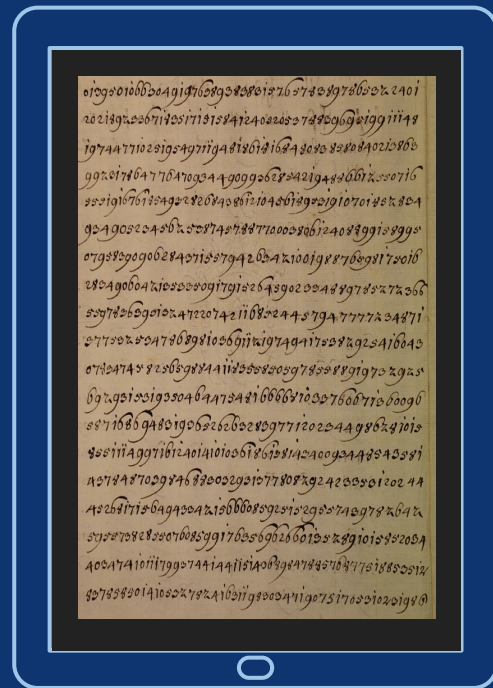
VEGA

- dataset vytvorený v rámci projektu VEGA 2/0072/20
- vzorky zaznamenávané ručne

03

VEGA-CROPPED

- dataset vytvorený rozdelením vzoriek VEGA datasetu
- veľkosť vzorky 128x128 pixelov



IMPLEMENTÁCIA – VYUŽITÉ DATASETY

01

ARDIS

- referenčný klasifikačný dataset
- nutnosť konverzie (COCO, YOLO Darknet TXT)

02

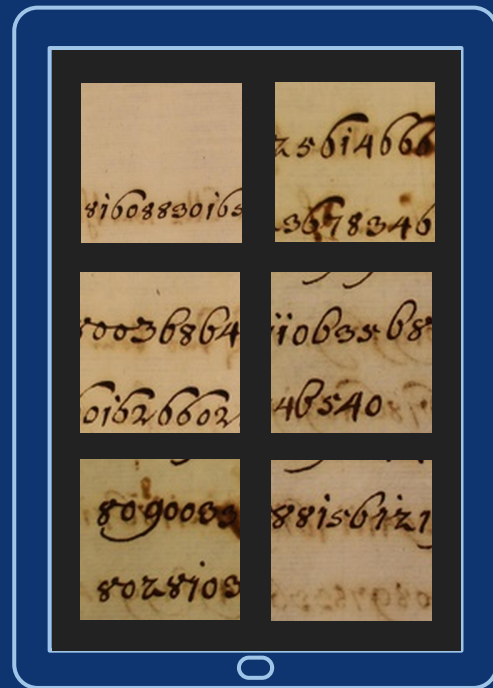
VEGA

- dataset vytvorený v rámci projektu VEGA 2/0072/20
- vzorky zaznamenávané ručne

03

VEGA-CROPPED

- dataset vytvorený rozdelením vzoriek VEGA datasetu
- veľkosť vzorky 128x128 pixelov



IMPLEMENTÁCIA – KONVERZIA DATASETU ARDIS

01

• ARDIS.CSV

- extrakcia vzoriek a metadát z .csv

02

• DÁTA – GRAYSCALE
INVERZIA

- prevod vzoriek z maticovej do invertovanej grayscale obrazovej reprezentácie

03

• DÁTA - RGB

- prevod vzoriek z grayscale do RGB obrazovej reprezentácie

IMPLEMENTÁCIA – KONVERZIA DATASETU ARDIS

04

ANOTÁCIE – POZÍCIE, OZNAČENIA

- vytvorenie anotácii pre pozíciu a označenie číslíc

05

ANOTÁCIE - MASKY

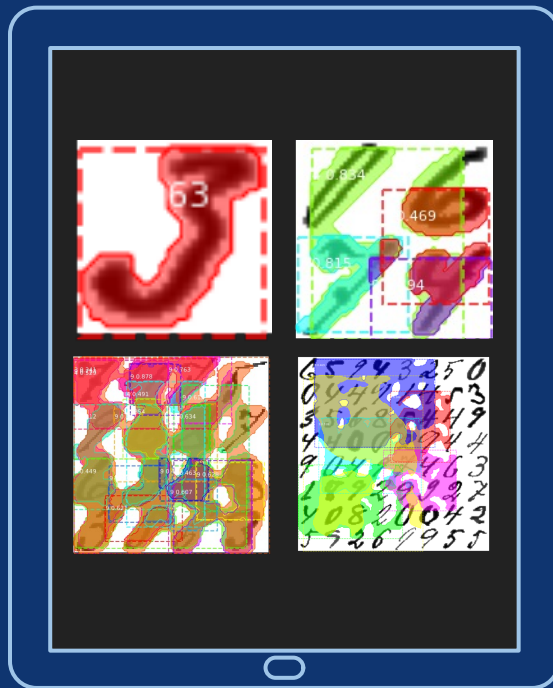
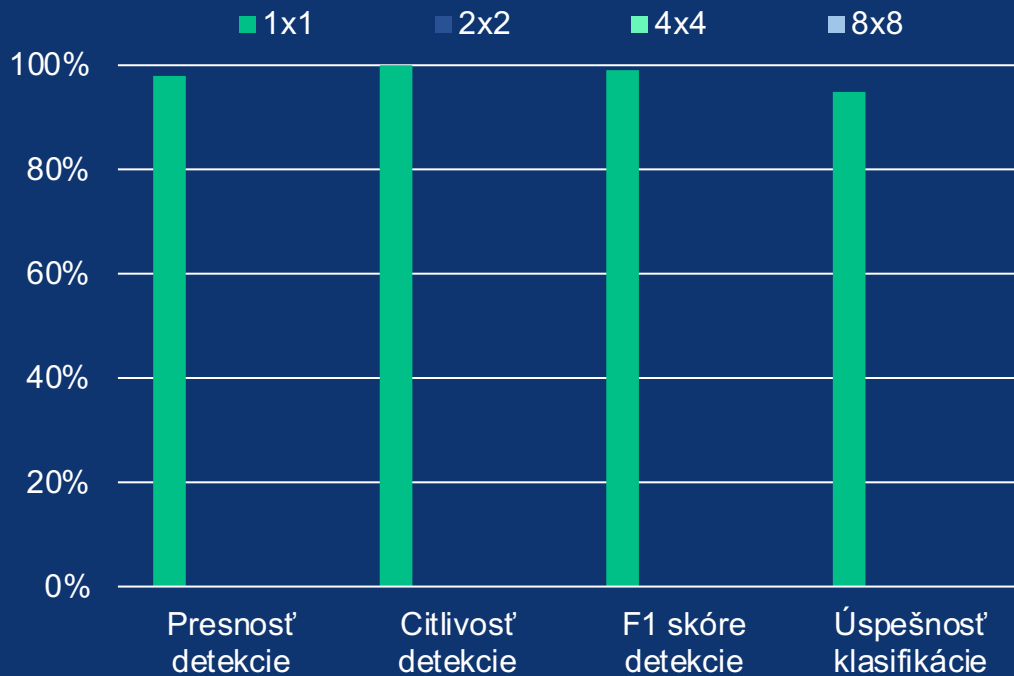
- vytvorenie anotácii pre masky číslíc pomocou detekcie kontúr

06

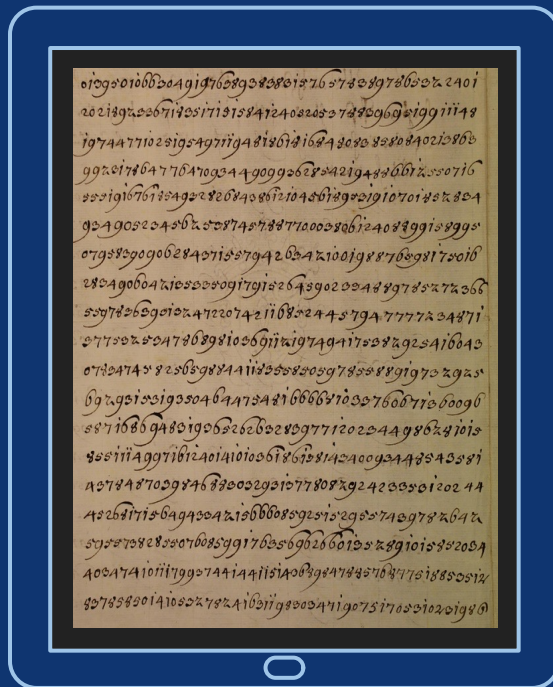
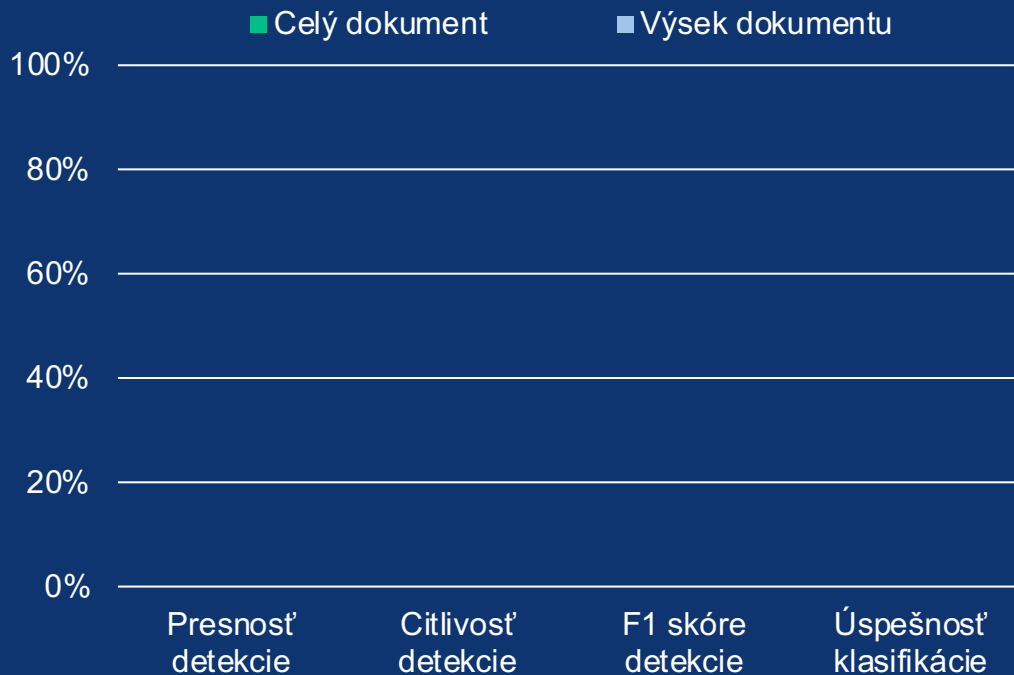
ARDIS_COCO.JSON/ ARDIS_YOLO.TXT

- vytvorenie finálneho anotačného súboru

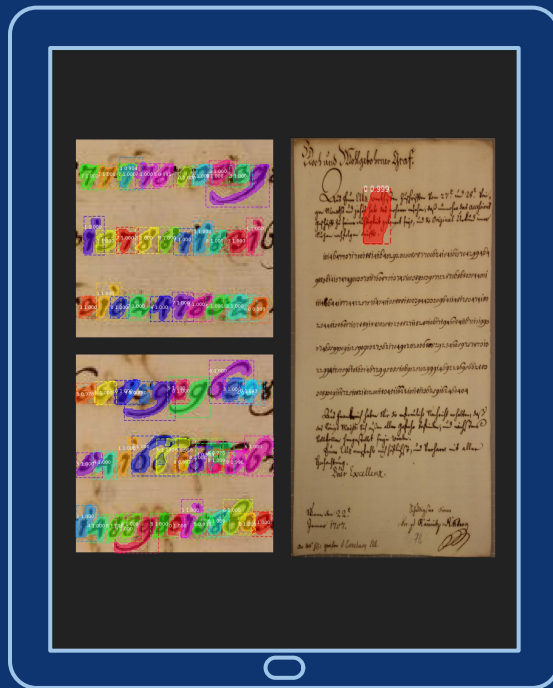
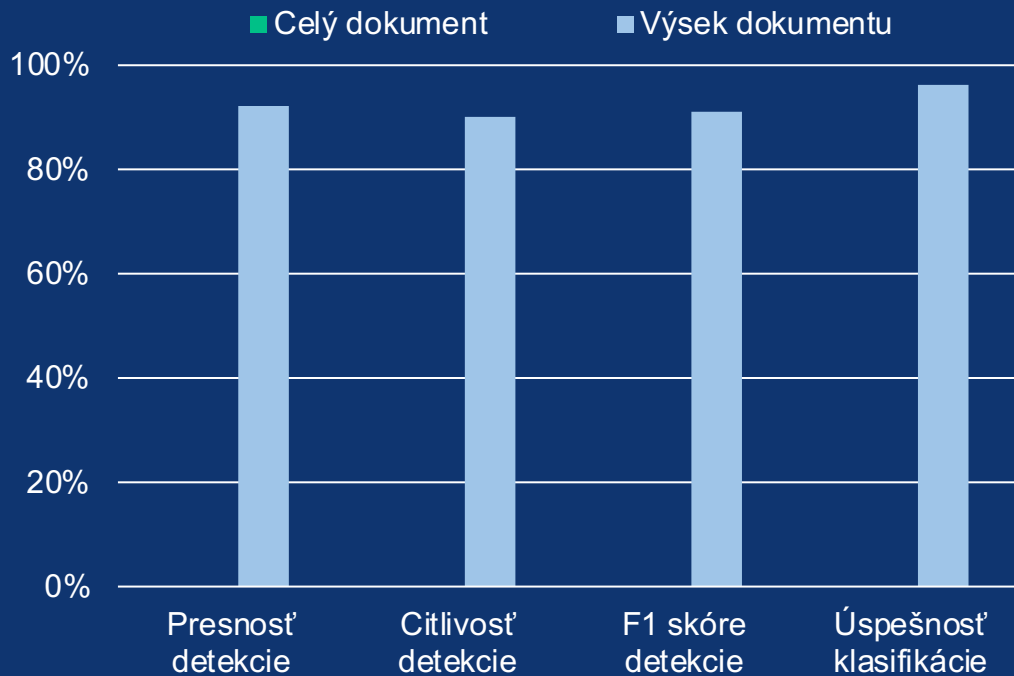
IMPLEMENTÁCIA – MASK R-CNN: ARDIS



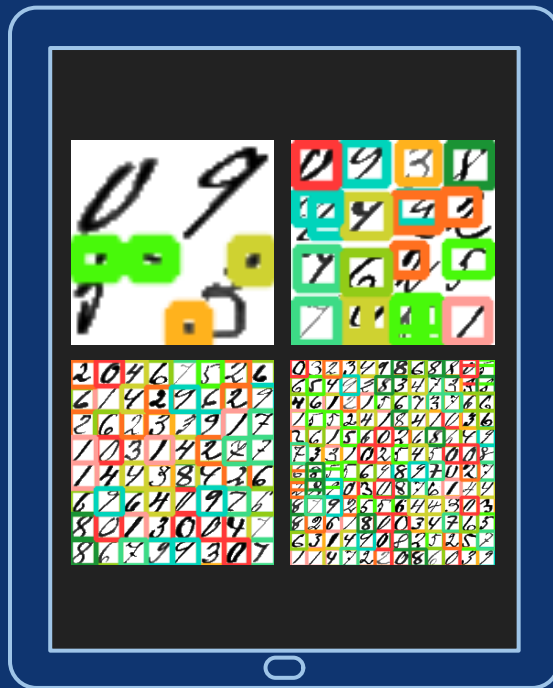
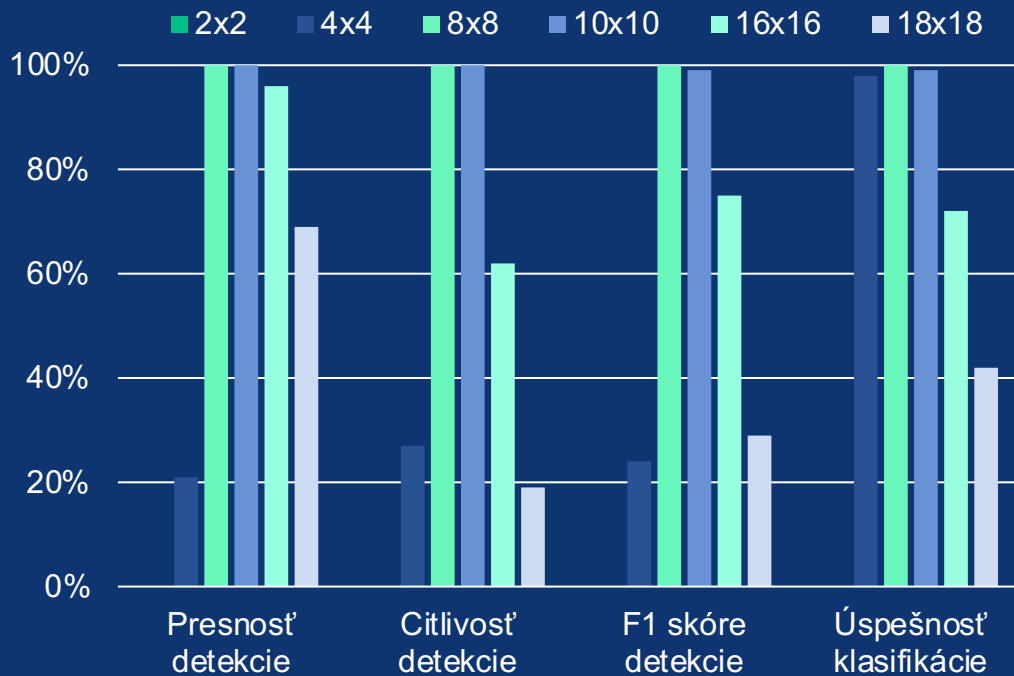
IMPLEMENTÁCIA – MASK R-CNN: VEGA



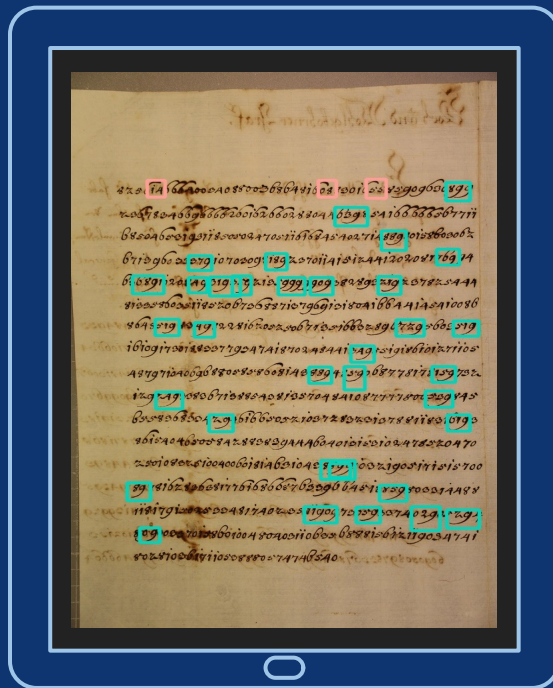
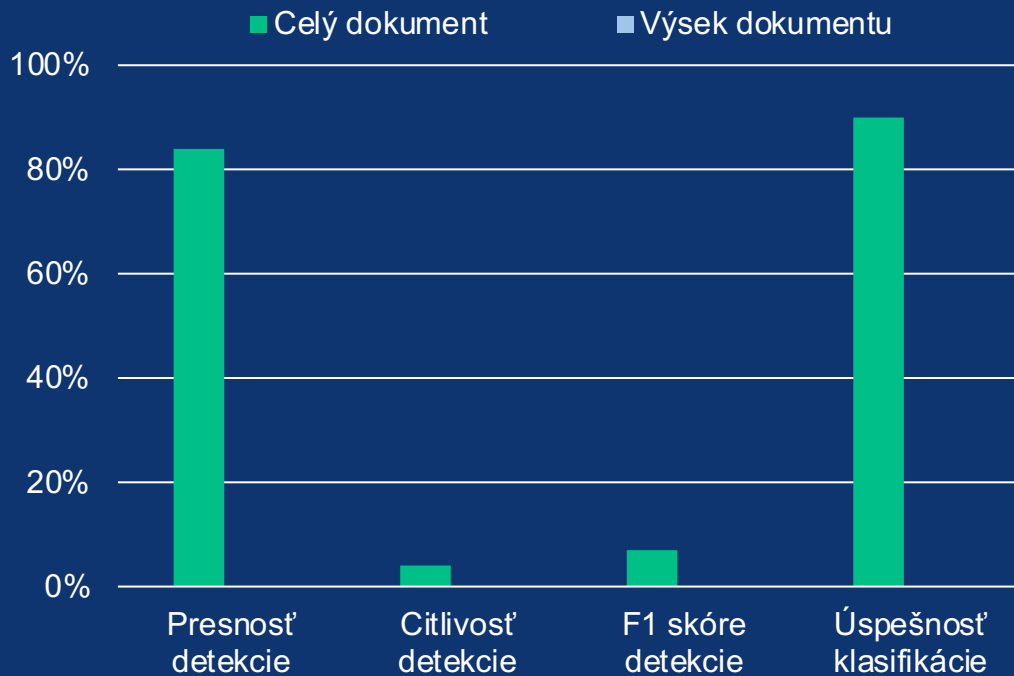
IMPLEMENTÁCIA – MASK R-CNN: VEGA-CROPPED



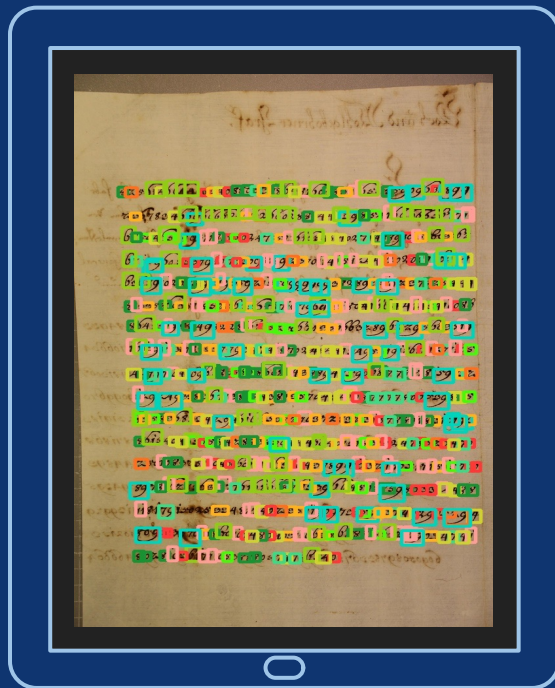
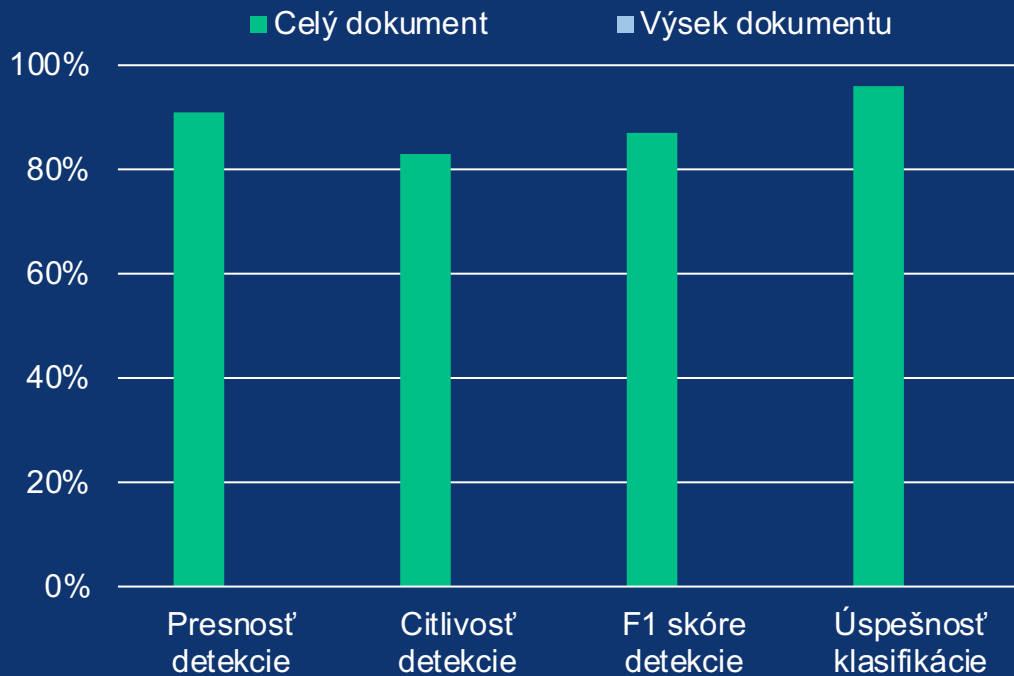
IMPLEMENTÁCIA – YOLO V5: ARDIS



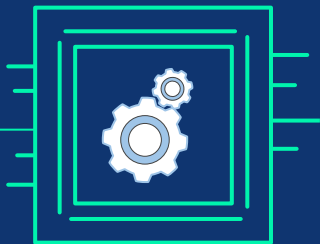
IMPLEMENTÁCIA – YOLO V5: VEGA



IMPLEMENTÁCIA – YOLO V5: VEGA-CROPPED

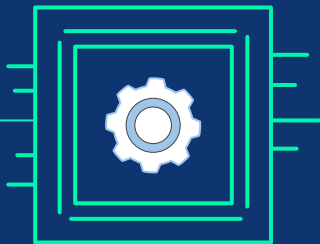


• VYHODNOTENIE – DETEKTORY



MASK R-CNN

- najlepšie výsledky: VEGA-Cropped
- problém s detekciou malých objektov
- potreba obdobnej štruktúry tréningových a reálnych vstupov
- vyššia časová náročnosť tréningu a detekcie



YOLO V5

- najlepšie výsledky: ARDIS, VEGA-Cropped
- problém s detekciou malých a veľkých objektov
- jednoduchá konfigurácia modelu
- nižšia časová náročnosť tréningu a detekcie

ZHRNUTIE

- ciele práce splnené
- poznatky a výstup využiteľné pri voľbe vhodného datasetu, predspracovaní obrazu alebo konfigurácii celého optimálne fungujúceho modelu detektora



ĎAKUJEM ZA POZORNOST



OTÁZKY

1. PREČO NEBOLA APLIKÁCIA NA PREDSPRACOVANIE VYUŽITÁ V PROCESE ROZPOZNÁVANIA ČÍSLIC?

Aplikácia mala za úlohu pokryť proces predspracovania z všeobecnej štruktúry systémov HCR. V práci sme avšak pracovali so systémami (Mask R-CNN, YOLO v5), ktoré vďaka svojej architektúre (NN) pre optimálne fungovanie nevyžadujú predspracovanie obrazového vstupu.



OTÁZKY

2. AKÁ BOLA ÚSPEŠNOSŤ AUTOMATIZOVANEJ TVORBY ANOTÁCIÍ POMOCOU KONTÚR?

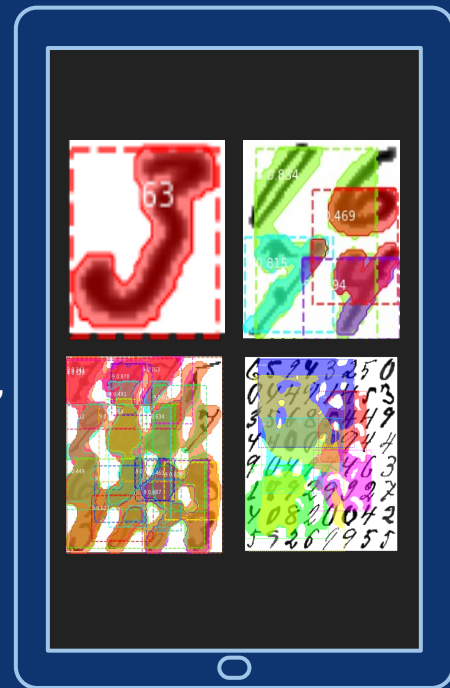
Úspešnosť automatizovanej tvorby anotácií masiek bola 100%. Tento jav má za následok forma referenčného datasetu ARDIS, ktorého vzorky zachytávajú normalizované a odšumené číslice bez nedokonalostí vo farbenom formáte Grayscale, čo uľahčuje segmentáciu číslice od bieleho pozadia.



OTÁZKY

3. OBR.42, AKO SI VIETE VYSVETLIŤ ZHORŠUJÚCU SA PRESNOSŤ DETEKcie PRI VÄČŠÍCH MRIEŽKACH? SKÚŠALI STE RÔZNE KONFIGURÁCIE RPN SIETE?

Mask R-CNN na návrh regiónov záujmu využíva samostatnú RPN sieť. Nakoľko bola daná konfigurácia trébovaná na vzorkách ARDIS s jednou číslicou zaberajúcou celú plochu vzorky, RPN sieť nebola optimálne natrébovaná. Boli vyskúšané rôzne konfigurácie hyperparametrov RPN_ANCHOR_SCALES, RPN_ANCHOR_RATIOS a RPN_ANCHOR_STRIDE, avšak bez zlepšenia pri koncovej detekcii a segmentácii.



OTÁZKY

4. KAPITOLA 5.3.3, V PRÍPADE ROZDELENIA OBRAZU DO BLOKOV 128X128 VZNIKNÚ NEROZPOZNANÉ ČÍSLICE NA ROZHraniACH BLOKOV. AKO BY STE VEDELI TENTO PROBLÉM RIEŠIŤ?

Číslice na rozhraniach blokov je možné rozpoznať pomocou vytvorenia nových výsekov 128x128, ktorých vertikálna a horizontálna os súmernosti kopíruje pôvodné hranice rozdelenia blokov.

