

Detekcia vizuálneho smogu pri cestách

Michal Zrubec

O práci

➤ Školiteľ:

RNDr. Zuzana Černeková, PhD.

➤ Cieľ práce:

Naštudovať metódy spracovania obrazu, detekcie významných oblastí a neurónové siete. Preskúmať možnosť využitia neurónových sietí na detekciu reklamných plôch popri cestách. Detekovať významnosť reklám na základe saliency a následne verifikovať pomocou eyetrackera. Navrhnutú metódu otestovať a vyhodnotiť výsledky.

O čo vlastne v práci ide?

1. Real-time detekcia reklamných plôch pri cestách s využitím neurónových sietí

O čo vlastne v práci ide?

1. Real-time detekcia reklamných plôch pri cestách s využitím neurónových sietí
2. Detekcia významnosti reklamných plôch na základe saliency
 - Verifikácia pomocou eyetrackera

Dataset

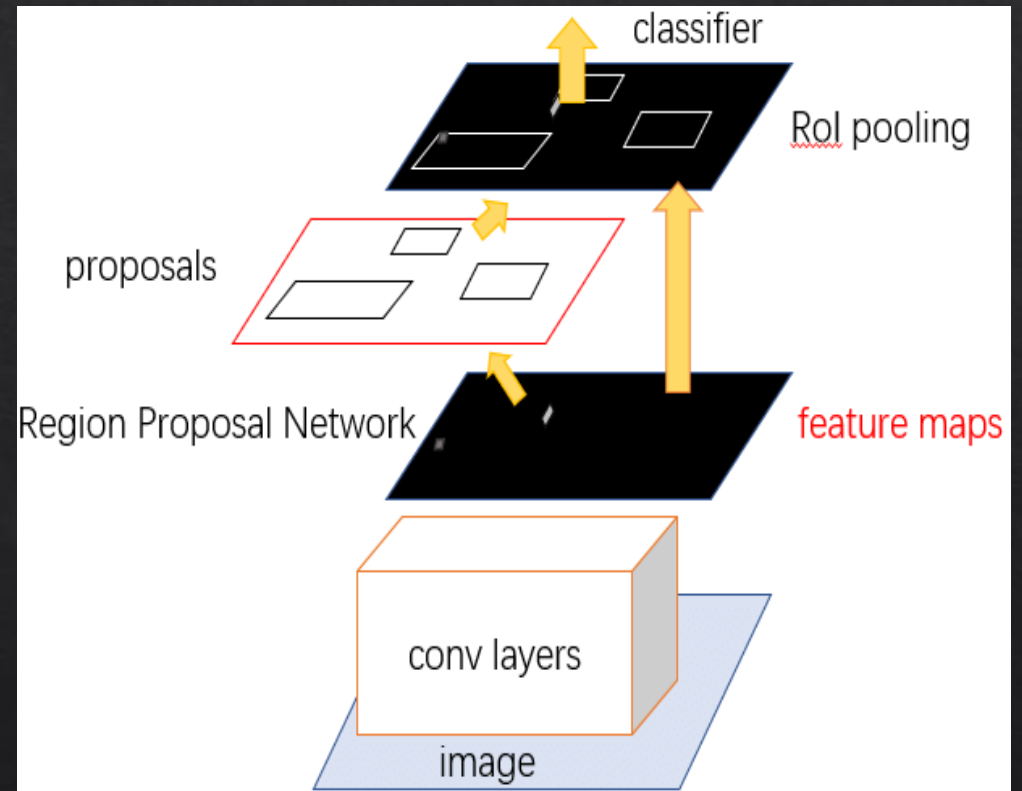
- Mapillary Vistas dataset
 - 25000 snímok rôznych veľkostí z miest a mimo miest
 - Obsahuje súbor, v ktorom sa nachádzajú anotácie objektov -> trieda „advertisement“
 - Snímky s veľkým aj menším rozlíšením
 - Rozdelený na tri množiny – **trénovacia** (18 tisíc), **validačná** (2 tisíc) a **testovacia** (5 tisíc)

Dataset

- Mapillary Vistas dataset
 - 25000 snímok rôznych veľkostí z miest a mimo miest
 - Obsahuje súbor, v ktorom sa nachádzajú anotácie objektov -> trieda „advertisement“
 - Snímky s veľkým aj menším rozlíšením
 - Rozdelený na tri množiny – **trénovacia** (18 tisíc), **validačná** (2 tisíc) a **testovacia** (5 tisíc)
- Vlastný dataset
 - Vytvorený pomocou telefónu a eyetrackera z automobilu
 - 1500 anotovaných snímok pomocou programu Roboflow
 - Rozlíšenie **1920x1080**
 - **Trénovacia, validačná a testovacia množina (70/20/10%)**

Faster R-CNN

- **Vylepšenie RCNN a Fast RCNN**
- **RPN** – výrazné zvýšenie efektivity vyhľadávania objektov



Faster R-CNN

- **Vylepšenie RCNN a Fast RCNN**
- **RPN** – výrazné zvýšenie efektivity vyhľadávania objektov
- **Detectron2**
- **fasterrcnn_resnet101_coco**

Faster R-CNN:

Name	lr sched	train time (s/iter)	inference time (s/im)	train mem (GB)	box AP	model id	download
R50-C4	1x	0.551	0.102	4.8	35.7	137257644	model metrics
R50-DC5	1x	0.380	0.068	5.0	37.3	137847829	model metrics
R50-FPN	1x	0.210	0.038	3.0	37.9	137257794	model metrics
R50-C4	3x	0.543	0.104	4.8	38.4	137849393	model metrics
R50-DC5	3x	0.378	0.070	5.0	39.0	137849425	model metrics
R50-FPN	3x	0.209	0.038	3.0	40.2	137849458	model metrics
R101-C4	3x	0.619	0.139	5.9	41.1	138204752	model metrics
R101-DC5	3x	0.452	0.086	6.1	40.6	138204841	model metrics
R101-FPN	3x	0.286	0.051	4.1	42.0	137851257	model metrics
X101-FPN	3x	0.638	0.098	6.7	43.0	139173657	model metrics

YOLO

- Velmi populární model
- Spracováva všetky objekty naraz
- Výborná rýchlosť so zachovaním presnosti



Split the image into grids

Predict bounding boxes
and confidence

Final predictions after
non-maximum suppression

YOLOv5

- **GitHub**
- Predtrénované váhy na datase COCO

Pretrained Checkpoints						
Model	size (pixels)	mAP ^{val} 0.5:0.95	mAP ^{val} 0.5	Speed CPU b1 (ms)	Speed V100 b1 (ms)	Speed V100 b32 (ms)
YOLOv5n	640	28.4	46.0	45	6.3	0.6
YOLOv5s	640	37.2	56.0	98	6.4	0.9
YOLOv5m	640	45.2	63.9	224	8.2	1.7
YOLOv5l	640	48.8	67.2	430	10.1	2.7
YOLOv5x	640	50.7	68.9	766	12.1	4.8

Mapillary Vistas Dataset

	<i>AP@0.5</i>	<i>AP@0.5:0.95</i>
YOLOv5	63,2%	48,1%
Faster R-CNN	55,8%	40,2%

Mapillary Vistas Dataset

	AP@0.5	AP@0.5:0.95
YOLOv5	63,2%	48,1%
Faster R-CNN	55,8%	40,2%

Pridanie augmentácie dát

- Náhodné otočenie
- Náhodný jas
- Šum
- Náhodné orezanie

	AP@0.5	AP@0.5:0.95
YOLOv5	58,7%	47,0%
Faster R-CNN	54,1%	42,3%

Vlastný dataset

	<i>AP@0.5</i>	<i>AP@0.5:0.95</i>
YOLOv5	96,6%	69,5%
Faster R-CNN	89,6%	58,4%

Vlastný dataset

	AP@0.5	AP@0.5:0.95
YOLOv5	96,6%	69,5%
Faster R-CNN	89,6%	58,4%

Pridanie augmentácie dát

- Náhodné otočenie
- Náhodný jas
- Šum
- Náhodné orezanie

	AP@0.5	AP@0.5:0.95
YOLOv5	93,2%	65,4%
Faster R-CNN	85,3%	58,7%

UniSal

- Dokáže modelovať významnosť na obrázkoch aj videách
- Encoder – RNN – Decoder
 - Encoder – MobileNet-V2
 - Bypass RNN – GRU (Gated Recurrent Unit) RNN (Recurrent Neural Network)
- Predtrénovaný na DHF1K datasete

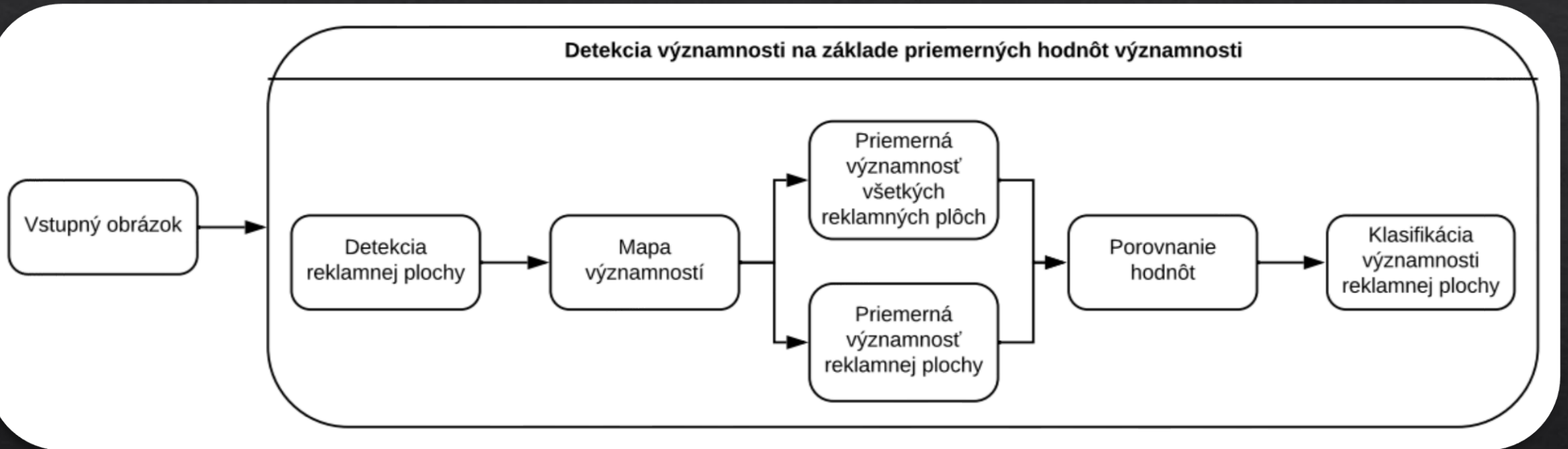


Spektrálne reziduá

- Obrázok sa pomocou Fourierovej transformácie rozdelí na fázovú a amplitúdovú zložku
- Vyhľadanie amplitúdovej zložky
- Odčítaním vyhľadanej amplitúdovej zložky od pôvodnej zostanú len najvýraznejšie oblasti na obrázku



Detekcia významnosti



Detekcia významnosti

Metóda	TP	TN	FP	FN	Presnosť (%)	Citlivosť (%)
UniSal	114	32	24	40	82,6	74,0
Spektrálne reziduá	108	24	48	30	69,2	78,3

Ukážky - Detekcia reklamných plôch



1. 6. 2022

Ukážky - Detekcia významnosti reklamnej plochy



1. 6. 2022

Detekcia významnosti reklamnej plochy



1. 6. 2022

Detekcia významnosti reklamnej plochy



1. 6. 2022

Detekcia významnosti reklamnej plochy



1. 6. 2022

Detekcia významnosti reklamnej plochy



1. 6. 2022

Použité nástroje

➤ Počítač:

- AMD Ryzen 7 2700
- 32GB RAM
- grafická karta **GeForce RTX 2070**

➤ Implementácia:

- Python
- Pytorch

➤ Verifikácia:

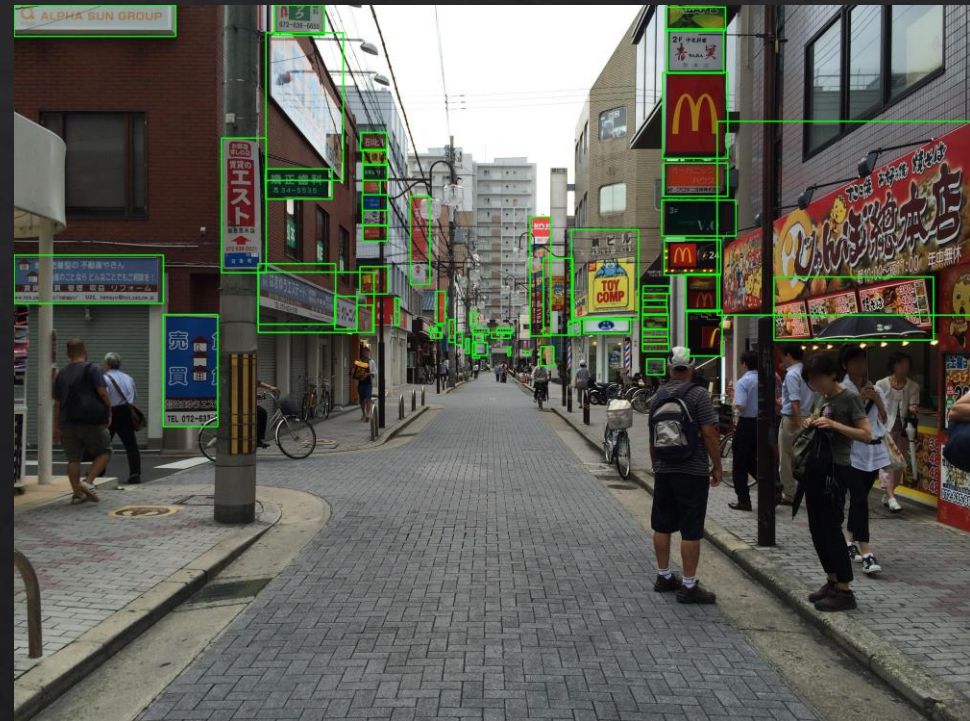
- Eyetracker Tobii Glasses 3
 - Tobii Pro Lab



Mapillary Vistas dataset vs vlastný dataset

Horšie výsledky pri Mapillary Vistas Datasete:

- Rôzne veľkosti obrázkov v datase (od 640x480 až po 5248x3936)
- V datase sú anotované aj veľmi malé reklamné plochy, ktoré pri znižovaní veľkosti obrázka pri vstupe do modelu „zaniknú“



Nesprávna detekcia



1. 6. 2022

Budúca práca

- Detekcia reklamnej plochy a jej významnosti v jednom kroku v reálnom čase.
- Ukladanie informácie o reklamných plochách – určenie najvýznamnejších typov reklamných plôch a ich najvýznamnejších vlastností

Ďakujem za pozornosť

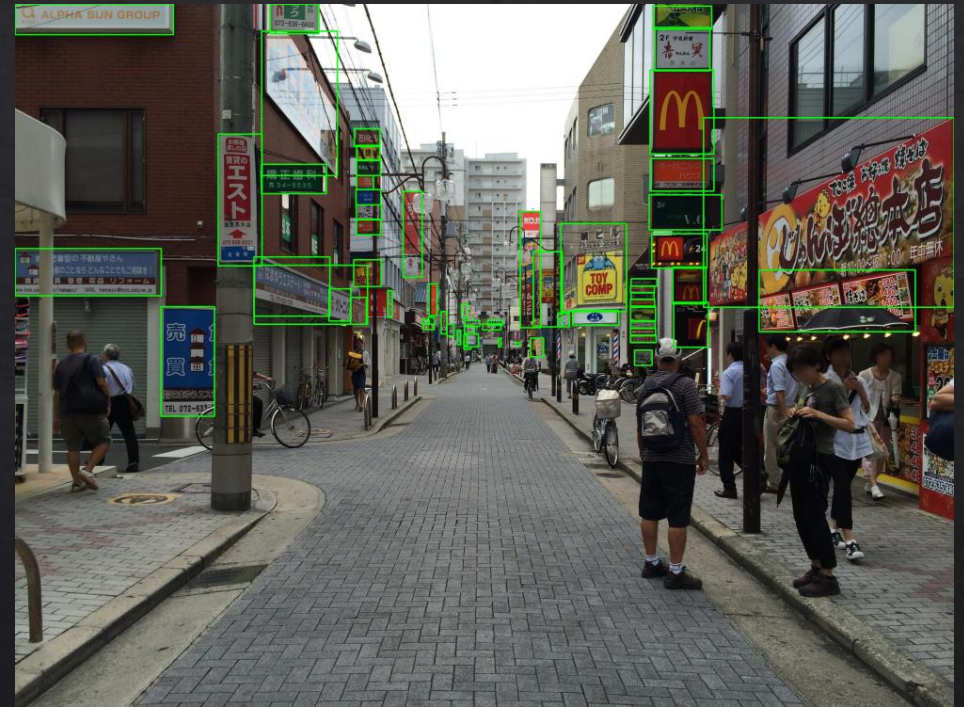
1. Skúste zhodnotiť, prečo augmentácia dát pri Mapillary Vistas datasete nepomohla zlepšeniu presnosti a aké iné druhy augmentácie by sa dali použiť (vzhľadom na povahu datasetu).

Dôvod:

- Šum mohol ešte viac zhoršiť detekciu malých reklamných plôch

Iné druhy augmentácie:

- Otočenie iba boundingboxov
- Pridanie výstrižkov z pôvodných obrázkov Image tilling



2. Navrhните ako by ste vyriešili problém s „miznúcimi“ malými billboardami pri Mapillary Vistas datasete

- Určenie podmienok pre preklamné plochy
 - Minimálna/maximálna výška, šírka
- Preatotovanie datasetu