

Reidentifikácia vozidiel v snímkach z dopravných kamier

Autor: Bc. Richard Dominik
Vedúci práce: Ing. Viktor Kocur, PhD.
FMFI UK 10.12.2021
www.st.fmph.uniba.sk/~dominik3/



O čom bude dnešná prezentácia ?

Problém a využitie

{1}

{2}

Dataseťy

Existujúce prístupy

{3}

{4}

State of the art

Výsledky implementácie

{6}

{7}

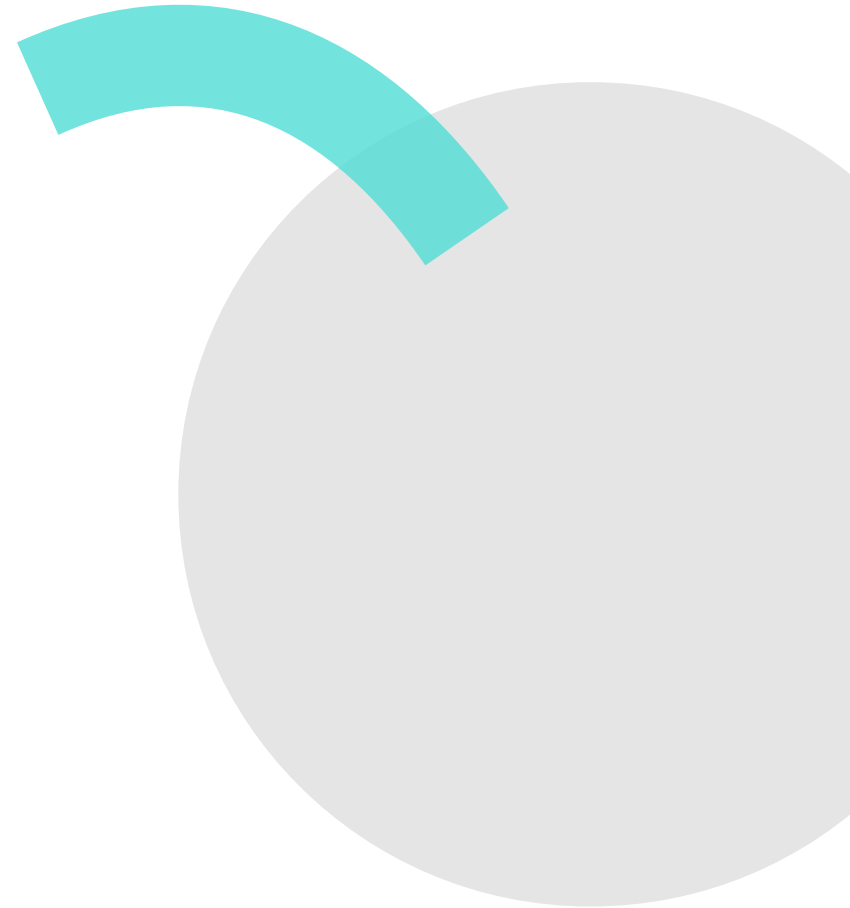
Ďalšie ciele

Problém a využitie

- zhoda rovnakého vozidla na snímkach z veľkého datasetu obrázkov (nasnímané pomocou dopravných kamier)
- rôzne kamery, orientácie, čas, lokácie, oklúzie, nezaostrenosť ...
- podobný tvar, model, farba, výrobca...
- reidentifikácia vozidiel != detekcia vozidiel

Problém a využitie

- aktuálna téma v oblasti počítačového videnia
- príbuzné k téme reidentifikácie osôb
- vieme identifikovať auto v rôznych bodoch mesta (odkiaľ kam išlo)
- využitie v inteligentných dopravných systémoch (efektívnejšie navrhovanie dopravných sietí)



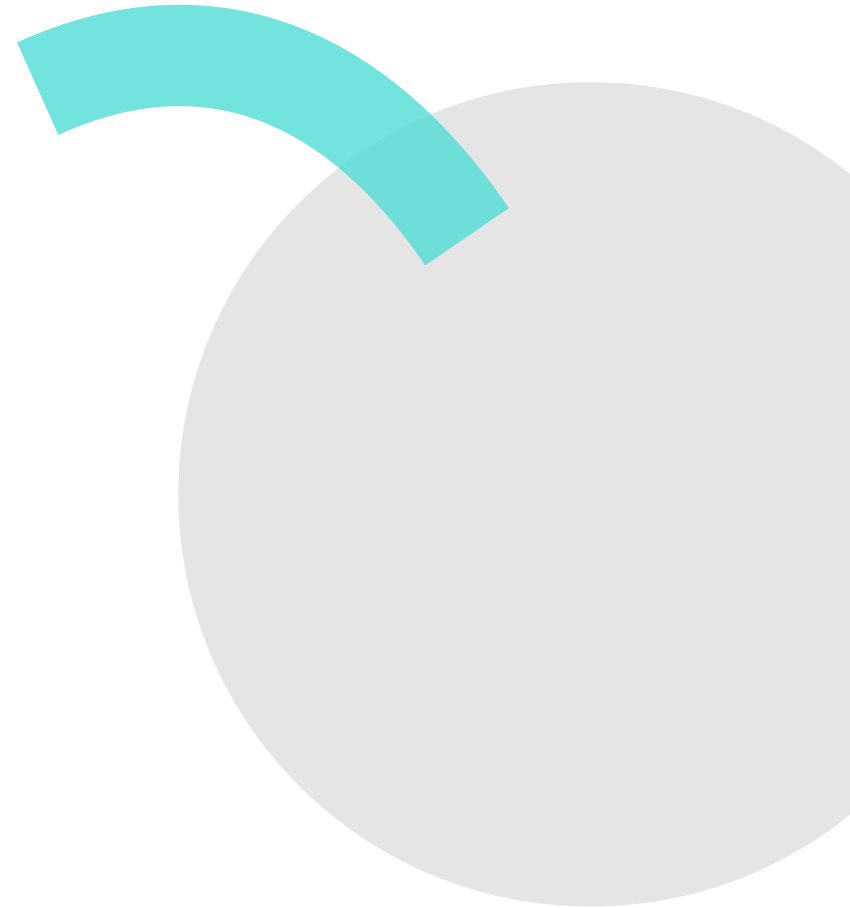
Náročnejšie prípady ?!



Obrázok: Ukážka prípadov zlyhania reidentifikácie (a, b rovnaký smer) a (c, d podobné pozadie)

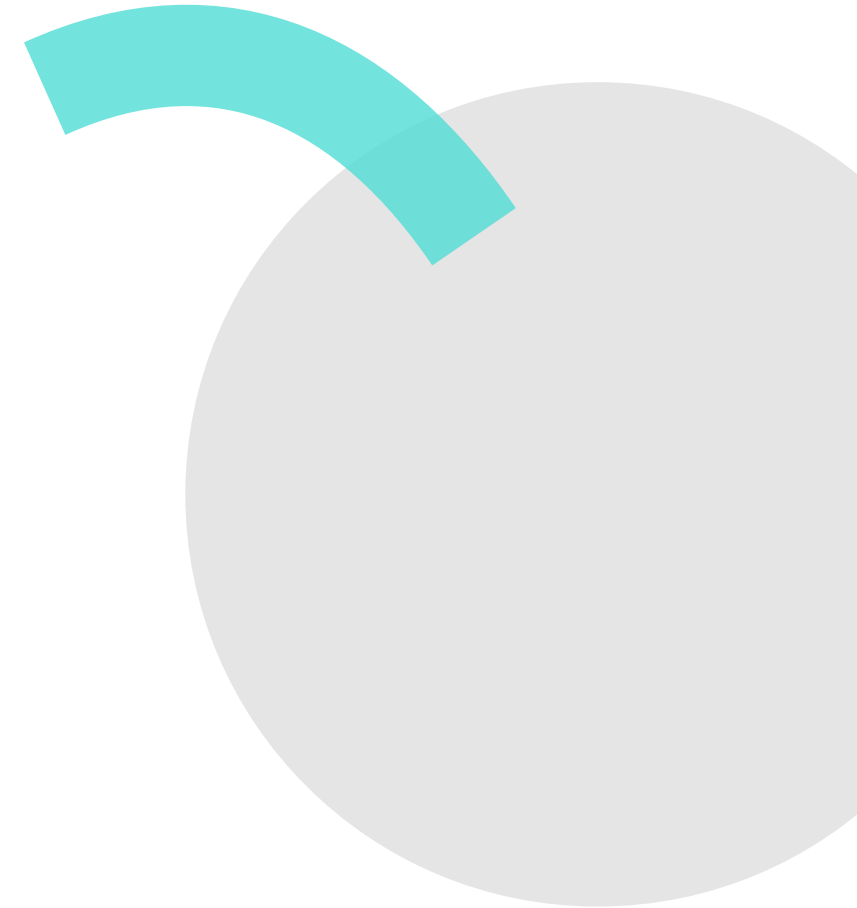
Datase~~ty~~

- **AI City Challenge dataset**
- **VeRi-776**
- Stanford Cars
- VERI – Wild
- CompCars
- VRAI
- Vehicle-1M
- VehicleX
- VehicleID
- BoxCars 116K



AI City Challenge dataset

- dáta nasnímané z dopravných kamier v USA (štát Iowa)
- 85 058 obrázkov
- 52 717 trénovacích a 31 238 testovacích obrázkov
- 440 rôznych vozidiel
- nasnímané pomocou 46 kamier
- anotované ľuďmi (farba, model, typ...)

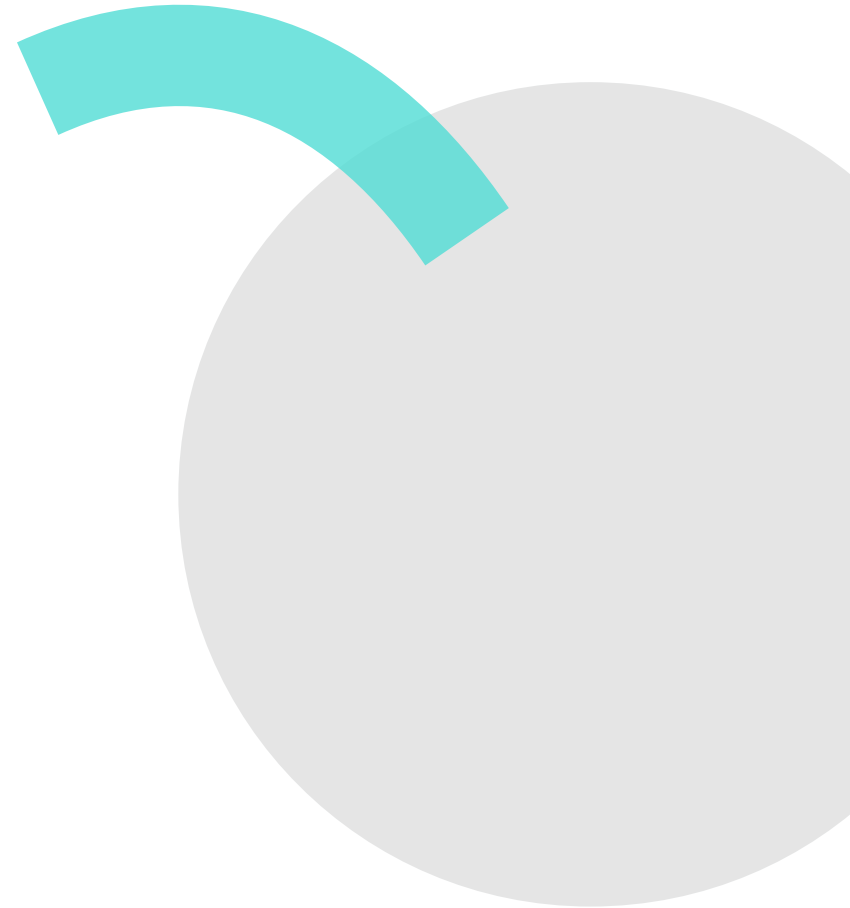


AI City Challenge dataset



VeRi-776 dataset

- 50 000+ obrázkov
- 776 rôznych vozidiel
- nasnímané pomocou 20 kamier
- rôzne pohľady, rozlíšenia, svetelné podmienky, oklúzie
- anotácie (Bbox, typ, farba, značka)



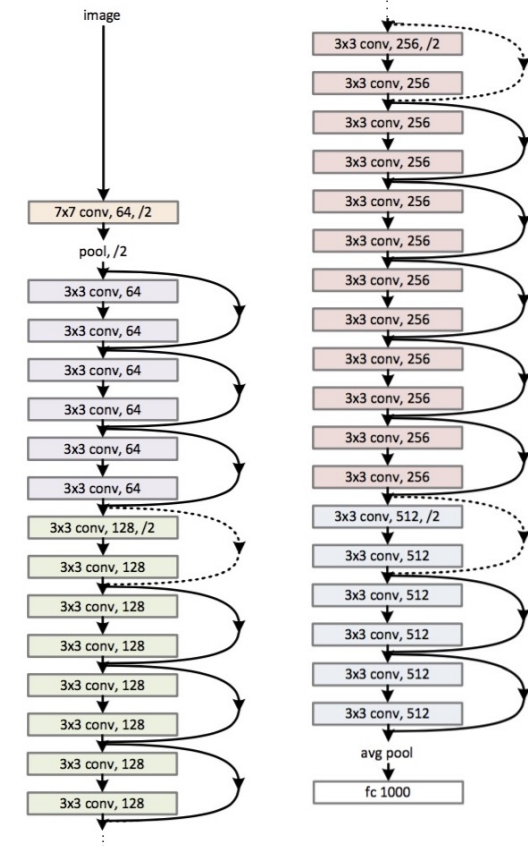
VeRI-776 dataset



Zaujímave poznatky z publikácií, ktoré využívali CNN

- využitie ResNetu ako backbone
- zmena rozmerov vstupných obrázkov a ich následná normalizácia
- trénovacie triky z publikácie zaoberajúcou sa reidentifikácie osôb (Warmup learning rate, REA)
- Variačný autoenkóder (VAE) - template vozidla bez loga výrobcu, diaľničných známkok, vzoru kolies, svetiel
- Adam optimalizátor
- iné prístupy boli napr. VOC reid (Vehicle, Orientation, Camera)

34-layer residual



Trénovacie triky

- Warmup learning rate (t je číslo epochy)

$$lr(t) = \begin{cases} 3.5 \times 10^{-5} \times \frac{t}{10} & \text{if } t \leq 10 \\ 3.5 \times 10^{-4} & \text{if } 10 < t \leq 40 \\ 3.5 \times 10^{-5} & \text{if } 40 < t \leq 70 \\ 3.5 \times 10^{-6} & \text{if } 70 < t \leq 120 \end{cases}$$

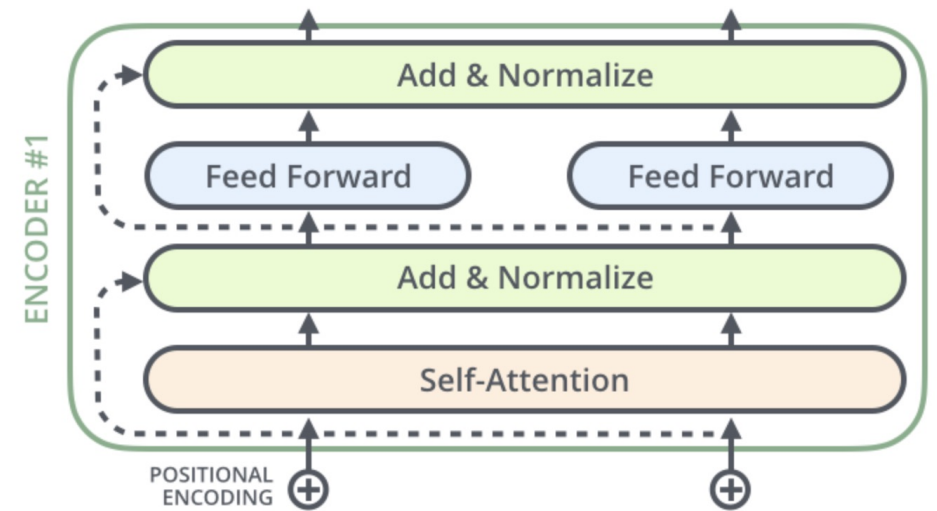
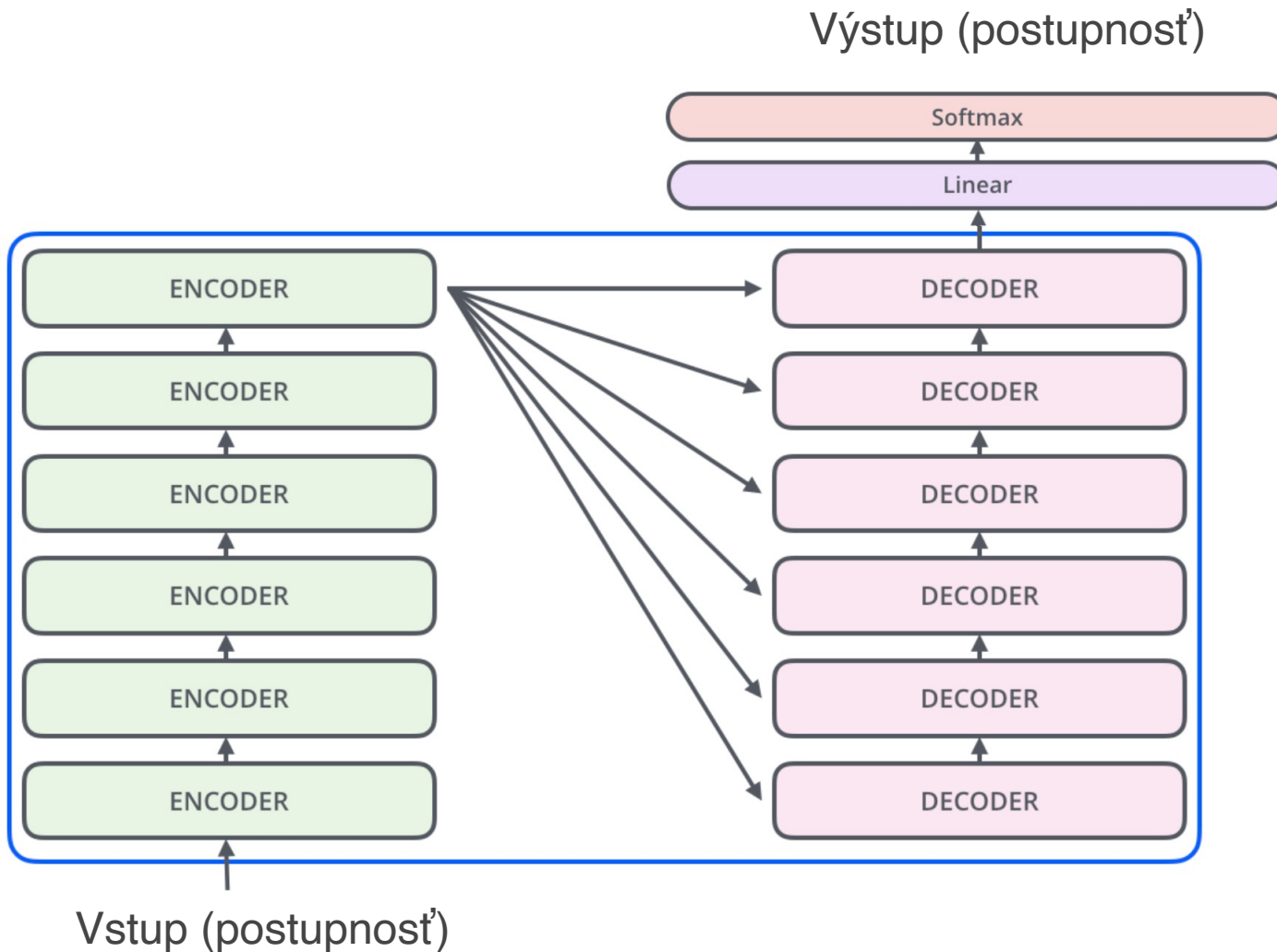
- Random Erasing Augmentation (REA)



Transformery

- začali sa objavovať publikácie, ktoré sa snažili túto úlohu riešiť pomocou Transformerov (ako náhradu za CNN)
- Transformer bol prvý krát predstavený v roku 2017 v súvislosti NLP, kde dosahoval state of the art výsledky s výrazne menším časom potrebným pre tréningovanie
- v počítačovom videní sa aktuálne využíva ViT a Swin Transformer
- ViT (október 2020, Google)
- Swin Transformer (marec 2021, Microsoft) + Swin Transformer V2 (november 2021)
- podľa paperswithcode.com state of the art výsledky pri detekcii objektov na dataseťe COCO test-dev (Swin transformer)

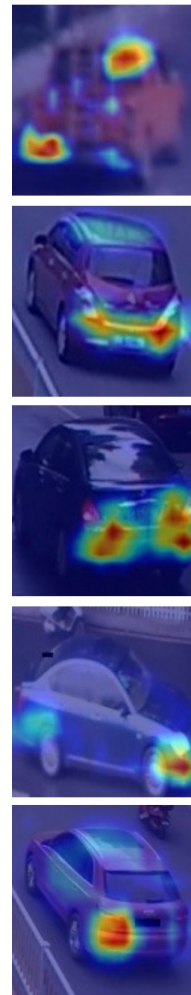
Transformery (princíp)



CNN vs Transformer



(a)



(b)



(d)

Obrázok: Grad-cam a) pôvodný obrázok, b) CNN backbone, d) Transformer (ViT) backbone

State of the reidentifikácia vozidiel Veri 776

$$\text{MAP} = \frac{\sum_{q=1}^Q \text{AveP}(q)}{Q}$$

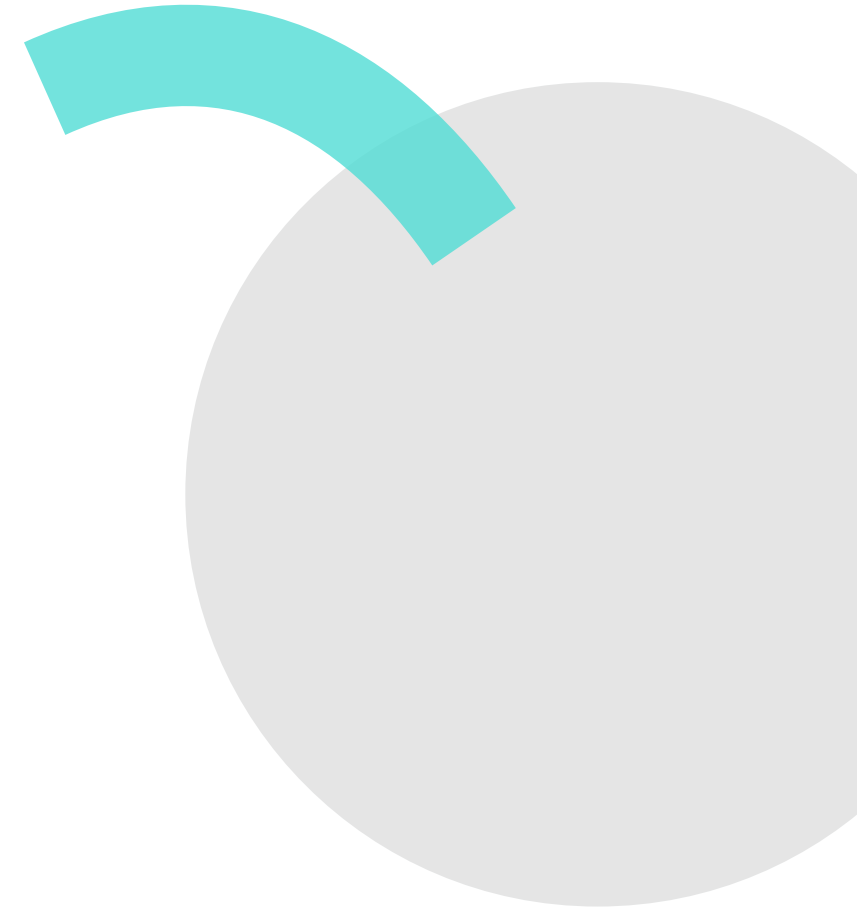
Rank	Model	mAP↑	Rank-1	Rank1	Rank5	Extra Training Data	Paper	Code	Result	Year	Tags
1	RPTM	87.4	96.2	96.2	98.1	×	Relation Preserving Triplet Mining for Stabilizing the Triplet Loss in Vehicle Re-identification			2021	
2	A Strong Baseline	87.1				×	A Strong Baseline for Vehicle Re-Identification			2021	
3	vehiclenet	83.41	96.78			✓	VehicleNet: Learning Robust Feature Representation for Vehicle Re-identification			2020	
4	TransReID	82.3	97.1			×	TransReID: Transformer-based Object Re-Identification			2021	
5	ANet	81.2	96.8	96.8	98.4	×	AttributeNet: Attribute Enhanced Vehicle Re-Identification			2021	
6	CAL	74.3	95.4		97.9	×	Counterfactual Attention Learning for Fine-Grained Visual Categorization and Re-identification			2021	
7	QD-DLF	61.83				×	Vehicle Re-identification Using Quadruple Directional Deep Learning Features			2018	

Výsledky implementácie v rámci predmetu PS2

- Natrénovanie TransReID¹ a vehiclereid_baseline¹ (CNN) (kód je implementovaný vo frameworku Pytorch)
- Google Cloud (NVIDIA Tesla T4 16GB, 7.5GB RAM)
- úprava hyperparametrov, sh skripty, v prípade vehiclereid_baseline uprava deprecated funkcií
- TransReID – 120 epôch (cca 43 hodín)
- vehiclereid_baseline – 100 epôch (cca 10 hodín)

¹linky na forky uvedených repozitárov s potrebnými úpravami sa nachádzajú na webovej stránke o tejto diplomovej práci

Výsledky implementácie v rámci predmetu PS2



- TransReID¹
 - mAP: 78.1% (v publikácii dosiahli autori 82%)
 - Rank-1 : 96.7% (v publikácii dosiahli autori 97.1%)
- vehiclereid_baseline¹
 - mAP: 22.61%
 - Rank-1 : 63.65%
 - dôvodom zlých výsledkov je skutočnosť, že tento kód neobsahuje žiadne trénovacie triky, augmentáciu...

¹linky na forky uvedených repozitárov s potrebnými úpravami sa nachádzajú na webovej stránke o tejto diplomovej práci

Ďalšie ciele v rámci implementácie

- Swin Transformer backbone (následne porovnanie s existujúcimi prácami s CNN a ViT)
- využitie tréningových trikov z publikácie o reidentifikácii osôb (warmup learning rate, REA...)
- riešenie pre dataset AI City Challenge
- Ablation Study (ktoré triky majú aký vplyv)

Zdroje

- Pirazh Khorramshahi, Neehar Peri, Jun-cheng Chen, and Rama Chellappa **The Devil is in the Details: Self-Supervised Attention for Vehicle Re-Identification**
(<https://arxiv.org/pdf/2004.06271.pdf>)
- Hao Luo , Youzhi Gu , Xingyu Liao , Shenqi Lai, Wei Jiang **Bag of Tricks and A Strong Baseline for Deep Person Re-identification**
(https://openaccess.thecvf.com/content_CVPRW_2019/papers/TRMTMCT/Luo_Bag_of_Tricks_and_a_Strong_Baseline_for_Deep_Person_CVPRW_2019_paper.pdf)
- Ze Liu, Yutong Lin, Yue Cao, Han Hu, Yixuan Wei, Zheng Zhang, Stephen Lin, Baining Guo **Swin Transformer: Hierarchical Vision Transformer using Shifted Windows**
(<https://arxiv.org/pdf/2102.04378.pdf>)
- Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, Neil Houlsby **An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale** (<https://arxiv.org/pdf/2010.11929.pdf>)

Zdroje

- Shuting He, Hao Luo, Pichao Wang, Fan Wang, Hao Li, Wei Jiang **TransReID: Transformer-based Object Re-Identification** (<https://arxiv.org/pdf/2103.14030.pdf>)
- <https://www.aicitychallenge.org/2021-data-and-evaluation/>
- <https://github.com/JDAI-CV/VeRidataset>
- <https://paperswithcode.com/sota/object-detection-on-coco>
- <https://paperswithcode.com/sota/vehicle-re-identification-on-veri-776>
- <https://editor.analyticsvidhya.com/uploads/38371XTo6Q.png>
- <https://jalammar.github.io/illustrated-transformer/>

Ďakujem za pozornosť !

Q&A