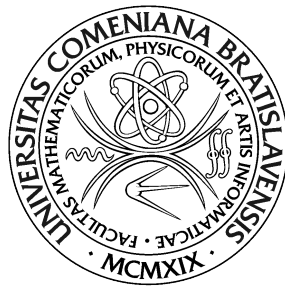


UNIVERZITA KOMENSKÉHO V BRATISLAVE  
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY



# Reidentifikácia vozidiel v snímkach z dopravných kamier

Diplomová práca

2021

Bc. Richard Dominik

UNIVERZITA KOMENSKÉHO V BRATISLAVE  
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY



# Reidentifikácia vozidiel v snímkach z dopravných kamier

Diplomová práca

Študijný program: Aplikovaná informatika  
Študijný odbor: 2511 Aplikovaná informatika  
Školiace pracovisko: Katedra aplikovanej informatiky  
Školiteľ: Ing. Viktor Kocur

Bratislava, 2021

Bc. Richard Dominik



Univerzita Komenského v Bratislave  
Fakulta matematiky, fyziky a informatiky

---

## ZADANIE ZÁVEREČNEJ PRÁCE

**Meno a priezvisko študenta:** Bc. Richard Dominik  
**Študijný program:** aplikovaná informatika (Jednoodborové štúdium, magisterský II. st., denná forma)  
**Študijný odbor:** informatika  
**Typ záverečnej práce:** diplomová  
**Jazyk záverečnej práce:** slovenský  
**Sekundárny jazyk:** anglický

**Názov:** Reidentifikácia vozidiel v snímkach z dopravných kamier  
*Re-identification of vehicles captured by traffic cameras*

**Anotácia:** Inteligentný dopravný systém (IDS) je pokročilý systém integrujúci rôzne informačné technológie s cieľom poskytnúť nástroje pre efektívnejšie, informovanejšie a bezpečnejšie využitie a návrh dopravných sietí. Dôležitou súčasťou IDS je zber dát. V kontexte cestnej dopravy je často vhodné zbierať dáta o pohybe vozidiel po rôznych cestách. Schopnosť reidentifikovať vozidlá v snímkach z rôznych dopravných kamier môže byť pri takomto zbere veľmi prospešná.

**Cieľ:** Cieľom práce je navrhnúť, implementovať a otestovať algoritmus založený na princípoch hlbokého učenia pre účely reidentifikácie vozidiel v snímkach z dopravných kamier. Súčasťou práce bude prehľad s moderných techník reidentifikácie obecné ako aj konkrétne v kontexte sledovania dopravy. Na vyhodnotenie budú využité verejne dostupné datasety a výsledky riešenia budú porovnané s existujúcimi prístupmi.

**Vedúci:** Ing. Viktor Kocur  
**Katedra:** FMFI.KAI - Katedra aplikovanej informatiky  
**Vedúci katedry:** prof. Ing. Igor Farkaš, Dr.  
**Dátum zadania:** 07.10.2020

**Dátum schválenia:** 08.10.2020

prof. RNDr. Roman Ďurikovič, PhD.  
garant študijného programu

---

študent

---

vedúci práce

Čestne prehlasujem, že túto diplomovú prácu som vypracoval samostatne len s použitím uvedenej literatúry a za pomoci konzultácií u môjho školiteľa.

Bratislava, 2021

.....

Bc. Richard Dominik

# Pod'akovanie

Pod'akovanie

# Abstrakt

Abstrakt

Klíčové slová:

# Abstract

Abstrakt EN

Keywords:

# Obsah

<b>1</b>	<b>Prehľad problematiky</b>	<b>2</b>
1.1	Reidentifikácia vozidiel . . . . .	2
1.2	Konvolučné neurónové siete . . . . .	3
1.2.1	Konvolučná vrstva . . . . .	5
1.2.2	Aktivačná vrstva . . . . .	6
1.2.3	Pooling vrstva . . . . .	6
1.2.4	Plne prepojená vrstva . . . . .	8
1.2.5	Triplet loss . . . . .	8
1.2.6	Metriky . . . . .	8
1.3	Architektúry konvulčných neurónových sietí . . . . .	9
1.3.1	VGG . . . . .	9
1.3.2	ResNet . . . . .	10
<b>2</b>	<b>Analýza datasetov</b>	<b>12</b>
2.1	VeRi-776 . . . . .	12
2.2	Stanford Cars . . . . .	13
2.3	AI City Challenge dataset . . . . .	14
2.4	Porovnanie . . . . .	15
<b>3</b>	<b>Súvisiace práce</b>	<b>16</b>



3.1	Bag of Tricks and A Strong Baseline for Deep Person Re- identification . . . . .	16
3.1.1	Triky využité při trénování . . . . .	17
3.2	VOC-ReID: Vehicle Re-identification based on Vehicle-Orientat- ion-Camera . . . . .	20

# Úvod

Úvod

# Kapitola 1

## Prehľad problematiky

V tejto kapitole sa venujeme prehľadu pojmov potrebných pre problematiku, ktorej sa táto práca venuje. Predstavíme si pojmy ako reidentifikácia vozidiel, konvolučné neurónové siete, významné architektúry a metriky pomocou ktorých sa vyhodnocujú výsledky a úspešnosť neurónových sietí pri úlohe reidentifikácie vozidiel.

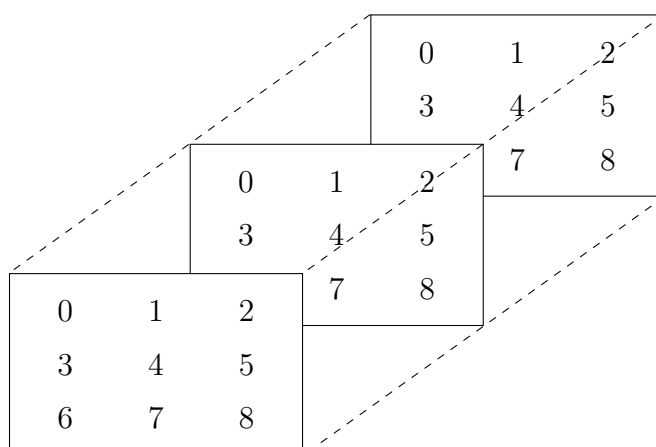
### 1.1 Reidentifikácia vozidiel

Úlohou reidentifikácie vozidiel v snímkach z dopravných kamier je nájsť zhodu rovnakého vozidla vo veľkom datasete obrázkov. Obrázky môžu byť nasmímané z rôznych kamier, orientácií, lokácií, ale aj času. Obrazové dáta taktiež môžu obsahovať rôzne oklúzie a nezaostrenosti. Na rozdiel od úlohy rozpoznávania vozidiel, ktorá sa venuje iba rozpoznaniu konkrétnej značky alebo modelu vozidla je reidentifikácia väčšou výzvou, nakoľko 2 rôzne vozidlá môžu byť vizuálne veľmi podobné. Môže ísť o rovnaký model, značku, alebo vozidlo, ktoré má rovnaký typ kolies, farbu alebo iné časti (nárazníky, svetlá a podobne). Reidentifikácia vozidiel je aktuálne veľmi populárna téma, ktorej

sa venuje oblasť počítačového videnia a jej využitie môžeme nájsť v inteligentných dopravných systémoch pre efektívnejšie navrhovanie dopravných sietí. Najaktuálnejšie prístupy sa venujú riešeniu tejto úlohy pomocou hlbokého učenia a konvolučných neurónových sietí. Mnohé prístupy profitujú z poznatkov nadobudnutých z odborných publikácii, ktoré sa venujú úlohe reidentifikácie osôb.

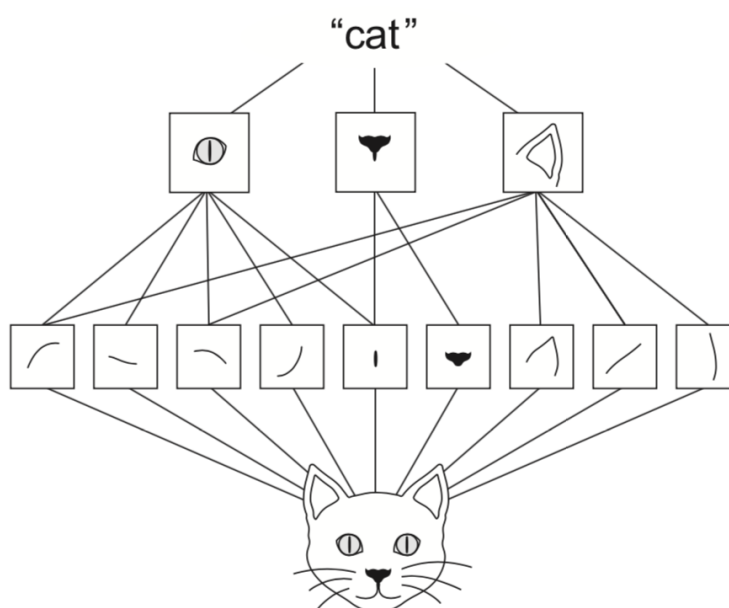
## 1.2 Konvolučné neurónové siete

Konvolučné neurónové siete (tiež známe aj ako CNNs alebo ConvNets) sú typ dopredných neurónových sietí používaných v hlbokom učení, ktoré najčastejšie pracujú s dátami, ktoré sa dajú reprezentovať pomocou mriežky. Tento typ sietí vie v dátach detekovať rôzne vzory a priradiť im význam. Ako dobrý príklad takýchto dát sú obrázky, ktoré vieme reprezentovať pomocou 3 rozmerných tenzorov s rozmermi *šírka*  $\times$  *výška*  $\times$  *hlĺbka*. V prípade obrázkov je za hlĺbku považovaný počet kanálov (v prípade RGB je počet kanálov 3 a hodnoty sú v rozmedzí 0 až 255). Na obrázku č. 1.1 môžeme vidieť grafickú reprezentáciu 3 rozmerného tenzora.



Obr. 1.1: Ukážka 3 rozmerného (3x3x3) tenzora

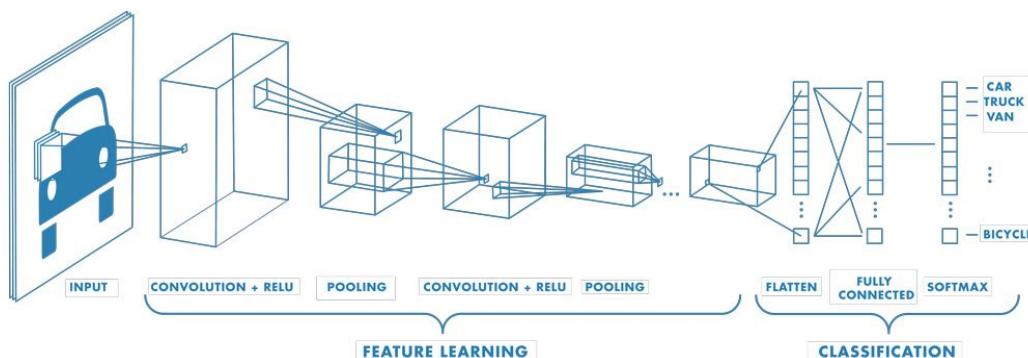
Medzi vzory, ktoré môžeme v obrazových dátach detekovať patria na vyšších vrstvách rôzne typy hrán a čím ideme vo vrstvách hlbšie, tým vieme detekovať komplexnejšie útvary vid' obrázok č. 1.3 .



Obr. 1.2: Príklad detekcie vzorov v konvolučných neurónových sieťach [Cho17]. Na vyšších vrstvách vieme detekovať rôzne typy hrán a čím ideme hlbšie, tým im vieme priradiť komplexnejší význam (oko, ucho, nos).

Jednou z hlavných výhod konvolučných neurónových sietí je ekvivariancia voči translácií vstupného obrazu. Pri vhodne zvolenom postupe tréovania sietí je možné napr. pomocou augmentácie sieť natréovať tak, aby sa výstup siete nemenil aj pri škálovaní a rotácii vstupného obrazu. Iné využitie konvolučných neurónových sietí môžeme nájsť aj v spracovaní zvuku alebo prirodzeného jazyka. Súčasťou konvolučných neurónových sietí sú skryté konvolučné vrstvy, ktoré sa najčastejšie kombinujú s vrstvami, ktoré nazývame pooling vrstvy, aktivačné vrstvy a plne prepojené vrstvy. Príklad neurónovej siete s viacerými konvolučnými vrstvami, ktoré si aj bližšie popíšeme môžeme

vidieť na obrázku č. 1.3.



Obr. 1.3: Príklad neurónovej siete s viacerými konvolučnými vrstvami [mat].

### 1.2.1 Konvolučná vrstva

Ako už názov tejto vrstvy napovedá, táto vrstva pracuje s matematickou operáciou, ktorú nazývame konvolúcia. Nakoľko konvolučné neurónové siete trénujeme na počítačoch, ktoré pracujú s diskretnými hodnotami, tak použijeme diskretnú verziu konvolúcie, ktorej definíciu môžeme vidieť na rovnici (1.1).

$$s(t) = (x * w)(t) = \sum_{a=-\infty}^{\infty} x(a)w(t - a) \quad (1.1)$$

Kde  $x$  je vstupný obraz,  $w$  je kernel a výstupom tejto operácie je mapa príznakov. Pomocou konvolúcie a zvoleného kernelu teda vieme nájsť v obraze významné miesta špecifické pre objekt, ktorý sa na danom obrázku nachádza. Ak chceme v obraze detekovať hrany, tak si za kernel zvolíme hranový detektor. Príklady takýchto detektorov sú:

1. Prewittovej hranový detektor
2. Sobelov hranový detektor

3. Robertsov hranový detektor

4. Cannyho hranový detektor

### 1.2.2 Aktivačná vrstva

Táto vrstva aplikuje na mapu príznakov z konvolučnej vrstvy nelinearitu pomocou aktivačnej funkcie. Pri konvolučných neurónových sieťach je najčastejšie ako aktivačná funkcia používaná ReLU (Rectified linear unit), ktorej definíciu môžeme vidieť na rovnici č. 1.2

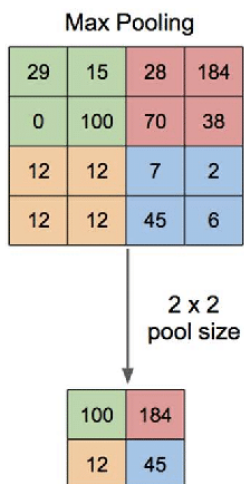
$$f(x) = \max(0, x) \quad (1.2)$$

Táto funkcia nemení veľkosť vstupu a medzi jej výhody patrí efektívnosť jej výpočtu a zníženie pravdepodobnosti problému miznúceho gradientu.

### 1.2.3 Pooling vrstva

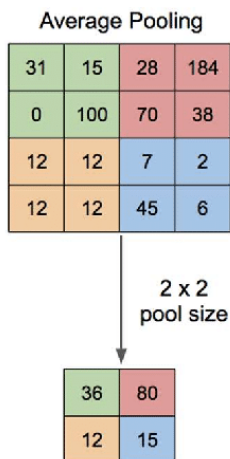
Problémom mapy príznakov z predchádzajúcich vrstiev je citlivosť na pozíciu príznakov v obraze. To znamená, že zmena veľkosti, otočenie alebo posun môžu výrazne zmeniť výslednu mapu príznakov. O riešenie tohto problému sa stará pooling vrstva a to tak, že znižuje mapu príznakov. Týmto prístupom zaručíme, že ďalšie vrstvy sa budú pozerieť už iba na najvýznamnejšie príznaky a tie menej dôležité (akými sú napríklad veľmi podrobné detaily), sa znížením rozmerov odstránia. Ďalšími výhodami tejto vrstvy sú zníženie počtu tréningových parametrov, čo znižuje počet výpočtov a teda aj času potrebného pre tréning siete a predchádza javu, ktorý nazývame overfitting. Najznámejšie funkcie využívané pre operáciu poolingú sú max pooling a average pooling.

Princípom max pooling je výber maximálnej hodnoty podľa zvoleného rozmeru filtra. Zvyčajne sa volí veľkosť filtra  $2 \times 2$ . Princíp max pooling môžeme vidieť na obrázku č. 1.4.



Obr. 1.4: Príklad operácie max pooling.

Princípom average pooling je výber priemeru hodnôt podľa zvoleného rozmeru filtra. Zvyčajne sa volí veľkosť filtra  $2 \times 2$ . Princíp average pooling môžeme vidieť na obrázku č. 1.5.



Obr. 1.5: Príklad operácie average pooling.



### 1.2.4 Plne prepojená vrstva

Poslednou fázou v konvolučných neurónových sieťach väčšinou býva plne prepojená vrstva. Ako už názov tejto vrstva napovedá, všetky neuróny tejto vrstvy sú navzájom plne prepojené ako tomu je aj v klasických neurónových sieťach. Úlohou tejto vrstvy je na základe príznakov naučených z predchádzajúcich vrstiev zaradiť objekt, ktorý bol na vstupnom obraze do správnej triedy.

### 1.2.5 Triplet loss

TODO: Triplet loss

### 1.2.6 Metriky

V tejto časti práce si popíšeme metriky, ktoré sa najčastejšie využívajú na vyhodnotenie úspešnosti systémov, ktoré sa venujú úlohe reidentifikácie.

#### **mAP**

TODO:

Je skratka z anglického Mean Average Precision

#### **CMC**

TODO:

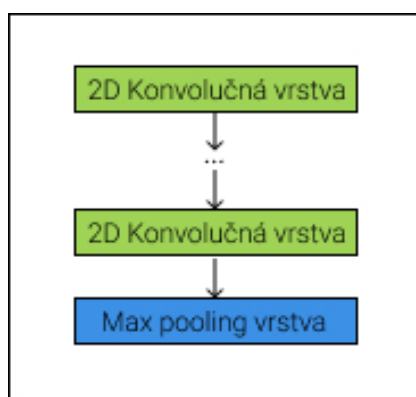
Je skratka z anglického Cumulative Matching Characteristics

## 1.3 Architektúry konvolučných neurónových sietí

V tejto podkapitole si predstavíme architektúry hlbokých neurónových sietí, ktoré v čase ich predstavenia dosahovali najlepšie výsledky a využívajú sa ako základ pri tréňovaní modelov pre úlohu reidentifikácie vozidiel.

### 1.3.1 VGG

VGG [SZ15] je skratka pre Visual Geometry Group. Táto architektúra bola predstavená v roku 2014 a skladá sa z viacerých blokov, kde každý blok obsahuje postupnosť konvolučných a max pooling vrstiev. Pri konvolučnej vrstve je veľkosť filtra  $3 \times 3$  a stride je 1. Ukážku VGG bloku môžeme vidieť na obrázku č. 1.6. Tieto bloky môžeme v rámci modelu použiť viac krát. Podľa počtu vrstiev označujeme VGG architektúry napríklad VGG16 pre 16 vrstiev a VGG19 pre 19 vrstiev. Autori tejto architektúry uvádzajú, že VGG architektúra je vhodná pre modely, kde je najväčší počet vrstiev 19 a to z dôvodu, že pri väčšom počte vrstiev sa prenášaný gradient stále znižuje, čo výrazne predlžuje čas tréňovania daného modelu.



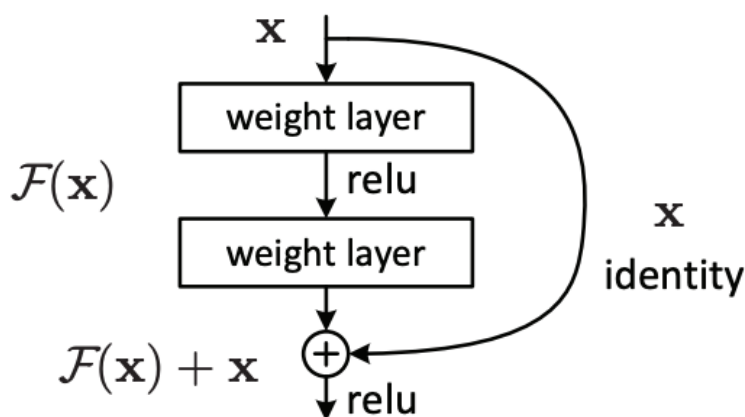
Obr. 1.6: Ukážka VGG bloku.

Vďaka tejto architektúre získal kolektív autorov v roku 2014 na ImageNet

výzve prvé miesto v klasifikácii a druhé miesto v lokalizácii. Implementáciu tejto architektúry nájdeme v každej bežne používanej knižnici pre hlboké učenie ako napríklad Keras alebo PyTorch.

### 1.3.2 ResNet

ResNet [HZRS15] je skratka pre Residual Network. Táto architektúra bola predstavená v roku 2015 a rieši problém prenášania gradientu, ktorý sme si spomenuli pri architektúre VGG. Rovnako ako VGG aj ResNet sa skladá z viacerých blokov, kde pri konvolučnej vrstve je veľkosť filtra  $3 \times 3$ . Narozdiel od VGG a iných klasických architektúr máme zavedené takzvané shortcut prepojenia. Tieto prepojenia nám umožňujú efektívnejšie prenášať gradient, čoho dôsledkom je možnosť pracovať s viacerými vrstvami ako bolo pri VGG. Najznámejšie sú verzie s počtom vrstiev 50 (ResNet 50) a 101 (ResNet 101). Ukážku ResNet bloku môžeme vidieť na obrázku č. 1.7



Obr. 1.7: Ukážka ResNet bloku [HZRS15].

Ako môžeme z ResNet bloku vidieť, shortcut prepojenie je len funkcia identity, vďaka čomu vieme dosiahnuť, že gradient bude pri spätnej propagácii

1 (lebo derivácia indetickej funkcie je stále 1). Vďaka tejto architektúre získal kolektív autorov zapojením sa v súťažiach ImageNet Large Scale Visual Recognition Challenge a Common Objects In Context (COCO) Challenge v roku 2015 prvé miesto v úlohách detekcie a lokalizácie (ImageNet) a rovnako prvé miesto v úlohách detekcie a segmentácie (COCO). Implementáciu tejto architektúry nájdeme v každej bežne používanej knižnici pre hlboké učenie ako napríklad Keras alebo PyTorch.

# Kapitola 2

## Analýza datasetov

V tejto kapitole si predstavíme sériu datasetov, ktoré sa aktuálne využívajú v počítačovom videní pri úlohe reidentifikácie vozidiel. Popíšeme si ako vybrané datasety vznikli, aké dáta obsahujú a vzájomne ich porovnáme. Niektoré z vybraných datasetov sú aj voľne prístupné, iné si však vyžadujú individuálne vyžiadanie a podpísanie súhlasu s podmienkami, ktoré si určili autori. Medzi podmienky môže patriť súhlas s využitím dát iba na nekomerčné účely a zákaz šírenia dát mimo organizácie pre ktorú je žiadosť schválená.

### 2.1 VeRi-776

VeRi-776 [LLMM18] je dataset, ktorý bol prvý krát uvedený v [LLMM16].

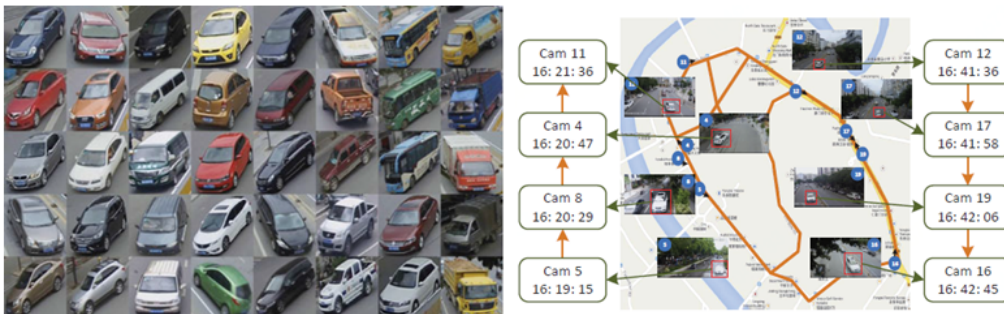
Dataset obsahuje:

1. viac ako 50 000 obrázkov
2. 776 rôznych vozidiel
3. dáta nasnímané v meste pomocou 20tich kamier

Každé vozidlo v tomto datasete je nasnímané 2 ~ 18 kamerami z rôznych uhlov, svetelných podmienok, oklúzií a v rôznych rozlíšeníach. Vozidlá disponujú aj nasledujúcimi atribútmi:

1. farba
2. typ
3. značka
4. bounding box

VeRi-776 nie je voľne dostupný. Je určený iba pre výskumné účely a pre prácu s ním je potrebné podať elektronickú žiadosť autorom datasetu. Ukážku obrázkov z tohto datasetu môžeme vidieť na obrázku 2.1



Obr. 2.1: Ukážka dát z datasetu VeRi 776.

## 2.2 Stanford Cars

Stanford Cars je voľne dostupný dataset poskytovaný Stanfordskou univerzitou. Prvý krát bol uvedený v [KSDF13].

Dataset obsahuje:

1. 16 185 obrázkov

2. 8144 trénovacích a 8041 testovacích obrázkov
3. 196 rôznych vozidiel

Dáta sú nasnímané v meste pomocou 20tich kamier. Stanford Cars je síce voľne dostupný, ale rovnako slúži iba na výskumné účely. Tento dataset ale neslúži na aplikácie pri sledovaní dopravy. Ukážku dát z tohto datasetu môžeme vidieť na obrázku č. 2.2



Obr. 2.2: Ukážka dát z datasetu Stanford Cars.

## 2.3 AI City Challenge dataset

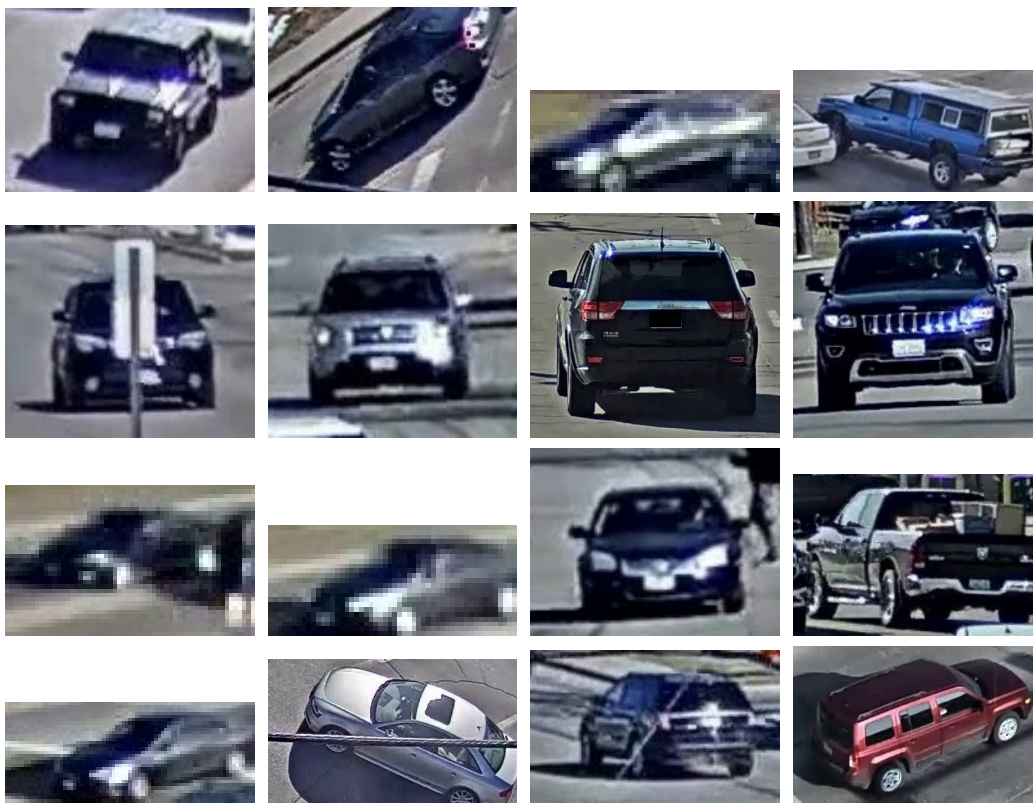
AI City Challenge dataset [NWA<sup>+</sup>21] je dataset, ktorý je súčasťou výzvy AI City Challenge . Je rozdelený na viacero stôp z ktorých je pre reidentifikáciu určená stopa 2 s názvom City-Scale Multi-Camera Vehicle Re-Identification. Tento dataset obsahuje:

1. 85 058 obrázkov
2. 52 717 trénovacích a 31 238 testovacích obrázkov
3. 440 rôznych vozidiel

4. 1103 obrázkov pre dopytovanie

5. syntetické dáta

Dáta sú nasnímane z rôznych dopravných kamier zo štátu Iowa v Spojených štátoch amerických. Obrázky boli anotované ľuďmi a poskytujú atribúty ako typ, farba vozidla a vzťahy medzi ostatnými vozidlami na obrázku. Ukážku dát z tohto datasetu môžeme vidieť na obrázku 2.3



Obr. 2.3: Ukážka dát z datasetu AI City Challenge.

## 2.4 Porovnanie

TODO: Porovnanie



# Kapitola 3

## Súvisiace práce

V tejto kapitole si predstavíme súvisiace práce, ktoré sa zaoberajú reidentifikáciou vozidiel a jednu prácu, ktorá sa zaoberá reidentifikáciou osôb. Dôvodom výberu práce zaoberajúcej sa reidentifikáciou osôb, je dôvod, ktorý sme už spomínali v predchádzajúcej kapitole, a to, že poznatky nadobudnuté z publikácií, ktoré sa venujú reidentifikácii osôb priniesli významné poznatky aj pre tému reidentifikácie vozidiel. Popíšeme rôzne prístupy autorov vybraných prác a výsledky ktoré dosiahli.

### 3.1 Bag of Tricks and A Strong Baseline for Deep Person Re-identification

V publikácii s názvom Bag of Tricks and A Strong Baseline for Deep Person Re-identification [LGL<sup>+</sup>19] sa kolektív autorov zaoberal tvorbou baseline frameworku pre úlohu reidentifikácie osôb. Venujú sa aj popísaniu a vyhodnoteniu rôznych trikov používaných v aktuálnych prístupov pre riešenie tejto problematiky. V tejto publikácii používali datasety s názvami Market1501 a DukeMTMC-reID. Ako základ pri trénovaní bol použitý ResNet 50 s predt-

rénovanými parametrami na ImageNet-e a pozmenenou dimenziou plne prepojenej vrstvy na  $N$ , kde  $N$  je počet identít v tréningových dátach. Ako ďalšie základné kroky pri tréningu takýchto modelov použili zmenu rozmerov obrázkov na rozmer  $256 \times 128$ , ktorým následne pridali nulový padding 10 pixelov a takto vytvorený obrázok ešte náhodne orezali na  $256 \times 128$  obdĺžnikový obrázok. Ďalej bol každý obrázok otočený v horizontálnom smere s pravdepodobnosťou 0.5 a použili aj mnoho iných základných prístupov, ktorým sa už nebude podrobnejšie venovať.

### 3.1.1 Triky využité pri tréningu

V tejto časti práce si popíšeme 6 trikov, ktoré podľa autorov tejto publikácie výrazne prispeli k vylepšeniu výsledkov ich baseline frameworku pre reidentifikáciu osôb.

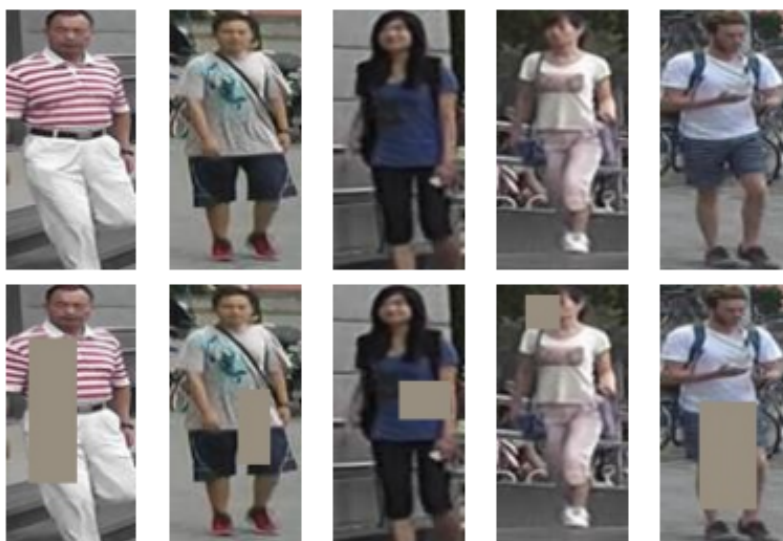
#### Warmup Learning Rate

Narozdiel od tradičných baseline prístupov, ktoré sa trénujú konštantnou rýchlosťou učenia použili autori Warmup Learning Rate, ktorý vylepšuje výkon celého re-id modelu. Hodnotu veľkosti kroku učenia môžeme vidieť na vzorci 3.1 kde  $lr(t)$  označuje learning rate počas epochy  $t$ .

$$lr(t) = \begin{cases} 3.5 \times 10^{-5} \times \frac{t}{10} & \text{ak } t < 10 \\ 3.5 \times 10^{-4} & \text{ak } 10 < t \leq 40 \\ 3.5 \times 10^{-5} & \text{ak } 40 < t \leq 70 \\ 3.5 \times 10^{-6} & \text{ak } 70 < t \leq 120 \end{cases} \quad (3.1)$$

### Random Erasing Augmentation

Random Erasing Augmentation (skrátene REA) je prístup, ktorý rieši problém rôznych oklúzií objektov. REA funguje na princípe, že si vyberie náhodný štvorcový región v obraze a vymaže jeho pixely náhodnými hodnotami. Príklad REA môžeme vidieť na obrázku 3.1



Obr. 3.1: Príklad REA, prvý riadok sú pôvodné obrázky a druhý riadok sú obrázky po aplikácii REA [LGL<sup>+</sup>19].

### Label Smoothing

TODO: Label Smoothing je trik, ktorý sa veľmi často využíva pri klasifikačných úloh aby sme predišli javu overfitting (preučenie).

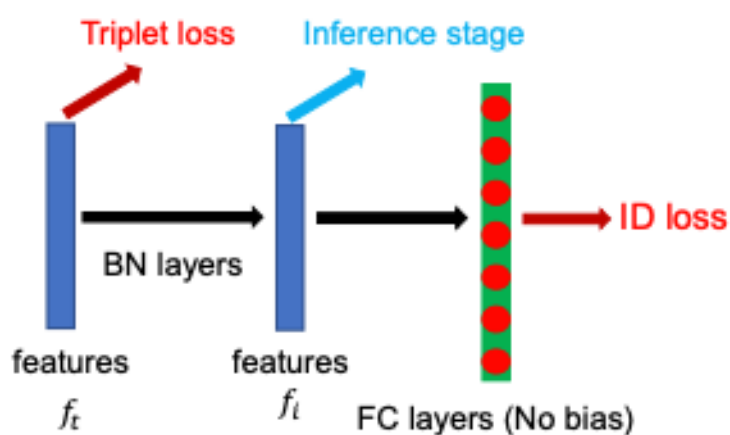
### Last Stride

Last Stride je trik, ktorý v poslednej vrstve ResNet-u, ktorý je použitý ako základ tréovania, zmenia poslednú vrstvu, (stride o veľkosti 2) na stride o veľkosti 1, čím autori získali na výstupe väčšiu mapu príznakov ( $16 \times 8$ )

čo v konečnom dôsledku zmení iba nepatrne dobu výpočtu, nezmení počet tréningových parametrov, ale zaručí výrazne vylepšenie výsledkov.

### BNNeck

TODO: BNNeck pridáva oproti štandardnému prístupu batch normalizačnú vrstvu po získaní príznakov. BNNeck môžeme vidieť na obrázku 3.2



Obr. 3.2: BNNeck [LGL<sup>+</sup>19].

### Center Loss

TODO: Definíciu Center Loss môžeme vidieť na rovnici 3.2

$$\mathcal{L}_c = \frac{1}{2} \sum_{j=0}^B \|f_{tj} - c_{y_j}\|_2^2 \quad (3.2)$$

Tréningovanie prebehlo pomocou 120 epóch a vďaka popísaným prístupom dosiahol tento baseline framework 94.5% rank-1 a 85.9% mAP na datasete Market1501 a 86.4% rank-1 a 76.4% mAP na datasete DukeMTMC-reID.

## **3.2 VOC-ReID: Vehicle Re-identification based on Vehicle-Orientation-Camera**

V publikácii VOC-ReID: Vehicle Re-identification based on Vehicle-Orientation-Camera [ZLFJ20] sa zaoberali reidentifikáciou vozidiel pomocou prístupu, ktorý nazvali VOC-ReID.

# Záver

Záver

# Literatúra

- [Cho17] François Chollet. *Deep Learning with Python*. Manning, November 2017.
- [HZRS15] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition, 2015.
- [KSDFF13] Jonathan Krause, Michael Stark, Jia Deng, and Li Fei-Fei. 3d object representations for fine-grained categorization. In *4th International IEEE Workshop on 3D Representation and Recognition (3dRR-13)*, Sydney, Australia, 2013.
- [LGL<sup>+</sup>19] Hao Luo, Youzhi Gu, Xingyu Liao, Shenqi Lai, and Wei Jiang. Bag of tricks and a strong baseline for deep person re-identification, 2019.
- [LLMM16] Xinchun Liu, Wu Liu, Tao Mei, and Huadong Ma. A deep learning-based approach to progressive vehicle re-identification for urban surveillance. In *ECCV (2)*, pages 869–884, 2016.
- [LLMM18] Xinchun Liu, Wu Liu, Tao Mei, and Huadong Ma. Provid: Progressive and multimodal vehicle reidentification for large-scale urban surveillance. *IEEE Transactions on Multimedia*, 20(3):645–658, 2018.

- [mat] Convolutional neural network. <https://www.mathworks.com/discovery/convolutional-neural-network-matlab.html>.  
Navštívené: 23. apríl 2021.
- [NWA<sup>+</sup>21] Milind Naphade, Shuo Wang, David C. Anastasiu, Zheng Tang, Ming-Ching Chang, Xiaodong Yang, Yue Yao, Liang Zheng, Pranameesh Chakraborty, Anuj Sharma, Qi Feng, Vitaly Ablavsky, and Stan Sclaroff. The 5th ai city challenge. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, June 2021.
- [SZ15] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition, 2015.
- [ZLFJ20] Xiangyu Zhu, Zhenbo Luo, Pei Fu, and Xiang Ji. Voc-reid: Vehicle re-identification based on vehicle-orientation-camera, 2020.



# Zoznam obrázkov

1.1	Ukážka 3 rozmerného (3x3x3) tenzora . . . . .	3
1.2	Príklad detekcie vzorov v konvolučných neurónových sieťach [Cho17]. Na vyšších vrstvách vieme detekovať rôzne typy hrán a čím ideme hlbšie, tým im vieme priradiť komplexnejší význam (oko, ucho, nos). . . . .	4
1.3	Príklad neurónovej siete s viacerými konvolučnými vrstvami [mat]. . . . .	5
1.4	Príklad operácie max pooling. . . . .	7
1.5	Príklad operácie average pooling. . . . .	7
1.6	Ukážka VGG bloku. . . . .	9
1.7	Ukážka ResNet bloku [HZRS15]. . . . .	10
2.1	Ukážka dát z datasetu VeRi 776. . . . .	13
2.2	Ukážka dát z datasetu Stanford Cars. . . . .	14
2.3	Ukážka dát z datasetu AI City Challenge. . . . .	15
3.1	Príklad REA, prvý riadok sú pôvodné obrázky a druhý riadok sú obrázky po aplikácii REA [LGL <sup>+</sup> 19]. . . . .	18
3.2	BNNeck [LGL <sup>+</sup> 19]. . . . .	19