Inteligentná analýza obrazu a hlboké strojové učenie

Riešitelia: Bc. Ľubomír Králik, Bc. Rastislav Papšo, Bc. Anton Moysey, Bc. Terézia Mária Houbová, Bc. Dominik Kornhauser
Garant: Ing. Peter Tarábek, PhD.

ŽILINSKÁ UNIVERZITA V ŽILINE
Fakulta riadenia a informatiky

Embedding priestor

Embedding priestor predstavuje n-dimenzionálny priestor, ktorého hlavnou vlastnosťou je uloženie dat v priestore tak, aby vzdialenosť medzi dátami s rovnakými vlastnosťami bola čo najmenšia. Vďaka tomu vytvárajú dáta v embedding priestore zhluky (clustre), pričom dáta s rovnakými vlastnosťami typicky tvoria jeden zhluk.

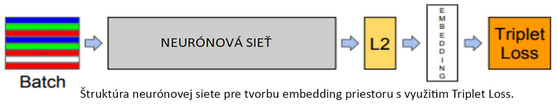
Embedding priestor môže byť následne využitý v rámci zhlukovej analýzy dát, klasifikácie dát, detekcie anomálií a poškodených dát v dátových množinách a iných metódach dátovej analýzy.

Pri experimentoch s embedding priestorom boli použité dátové množiny MNIST, FashionMNIST a CIFAR10.

Tvorba embedding priestoru

Pre tvorbu embedding priestoru sme využili neurónovú sieť s použitím špecifickej účelovej funkcie **Triplet Loss**. Vstupom do tejto funkcie sú 3 vzorky - kotva (A), pozitívna vzorka (P) a negatívna vzorka (N). Pozitívna vzorka a kotva patria do rovnakej triedy. V procese trénovania dochádza k oddeleniu pozitívnych a negatívnych párov o vzdialenosť α, nazývanú margin.
V rámci experimentov bolo snahou modifikovať TripletLoss tak, aby jednotlivé triedy mohli formovať viacero zhlukov.

TripletLoss(A,P,N,margin) = max{d(A,P) - d(A,N) + margin, 0}

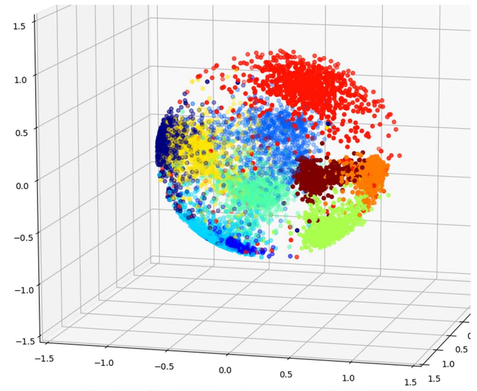


Štruktúra neurónovej siete pre tvorbu embedding priestoru s využitím Triplet Loss.
Zdroj: https://www.tensorflow.org/addons/tutorials/losses_triplet, upravené.

Experimenty

Výsledky (presnosť) pre dataset CIFAR10 na testovacej množine:

a) baseline (Softmax+CrossEntropy) - **86.67%**
b) rozdelenie trénovacích tried do podtried s využitím predtrénovanej siete ResNet50 na datasete ImageNet, metódy UMAP a algoritmu K-Means - **81.03%**
c) rovnomerné rozdelenie vzoriek do podtried so zachovaním početnosti podtried podľa b) - **85.92%**
d) rovnomerné rozdelenie vzoriek do 5 (8) podtried - **86.46%** (**86.36%**)
e) Dotrénovanie modelov b),c),d) pomocou MeanSquaredError k centroidom: priemerné zlepšenie o **0.57**%
f) vzorka P je považovaná za pozitívnu, ak spĺňa podmienku $\rho_{P,A} \geq t$, inak je označená za negatívnu - **80.33%**



Vizualizácia 3D embedding priestoru datasetu FashionMNIST

Detekcia rakoviny prsníka

Cieľom tejto časti projektu je detegovat a klasifikovať nádor rakoviny prsníka z mamografických snímkov. Pre prácu bol využitý dataset **CBIS-DDSM,** ktorý pozostáva zo samotných mamografických snímkov spolu s maskou, ktorá konkrétne lokalizuje pozíciu nádoru a metaúdajov, napr. o type nádoru - benígny, malígny. Trénovacia množina obsahuje 1248 pacientov, testovacia množina obsahuje 349 pacientov. Snímky sú rozdelené do dvoch kategórií: Calcification a Mass.

Počet snímkov v jednotlivých kategóriach	Train	Test
Calcification benign	1002	197
Calcification malignant	544	129
Mass benign	681	231
Mass malignant	637	147

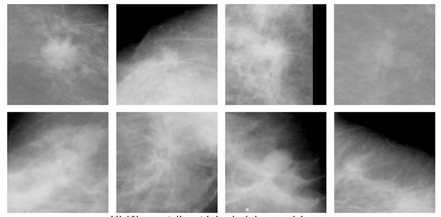
Dosiahnuté výsledky	Train	Test
Klasifikácia nádorov	96.18%	88.36%
Klasifikácia celej snímky	88.62%	71.36%

Detekcia

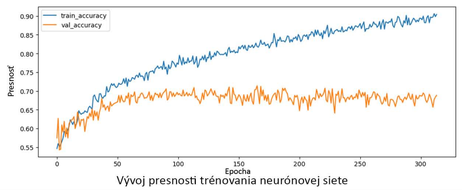
Metóda detekcie nádoru bola založená na prístupe, v ktorom bola použitá architektúra neurónovej siete z Deep Learning Frameworku s názvom Detectron. Detectron predstavuje implementáciu metód Fast R-CNN (detekcia objektov v obrazových dátach) a Mask R-CNN (sémantická segmentácia obrazu). Model Detectronu bol následne natrénovaný na uvedenom datasete spolu s informáciou o ohraničení jednotlivých nádorov.

Klasifikácia

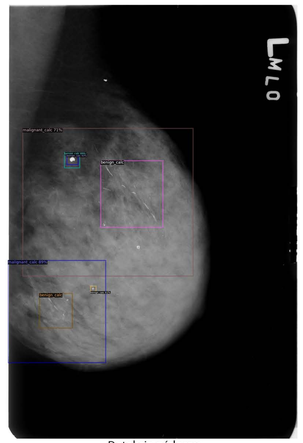
Pre klasifikáciu boli použité dva prístupy - **klasifikácia** samotných **nádorov** a **klasifikácia celej snímky**. Pri klasifikácií nádorov sa jednotlivé nádory z obrázka vystrihli, čím vznikli obrázky o rozmere 224x224 pixelov. Motiváciou k zvoleniu tohto rozmeru bolo využitie predtrénovaných sietí na datasete ImageNet, ktoré majú vstupný rozmer 224x224. Takýmto spôsobom sme predtrénovanú sieť dotrénovali na naše dáta, s jej neskorším využitím na klasifikáciu celej snímky za pomoci konceptu *plne konvolučnej neurónovej siete* (Fully Convolutional Neural Network).



Ukážka vystrihnutých obrázkov s nádorom



Vývoj presnosti trénovania neurónovej siete



Detekcia nádorov

Vizuálna kategorizácia vozidiel na základe počtu náprav

Kategorizácia vozidiel je jednou z dôležitých úloh pri analýze dopravy. Pri kategorizácii sa dá postupovať viacerými spôsobmi, napr. podľa dĺžky vozidiel alebo vizuálnych vlastností vozidiel. Projekt sa zameriava na kategorizáciu vozidiel na základe počtu náprav na vozidle.

Prvý prístup je založený na detekcii kolies a ich následnému priradeniu k detegovaným vozidlám. Používajú sa dve neurónové siete: prvá detekuje vozidlá a druhá detekuje kolesá.

Druhý prístup je zameraný na priame kategorizovanie vozidiel, kde počet náprav reprezentuje triedu vozidla. Tento prístup je robustnejší voči oklúzií vozidiel, avšak zložitejší na samotné trénovanie neurónovej siete.

V oboch prístupoch sa používa neurónová sieť ResNet-50 s využitím frameworku Detectron2.

Dataset

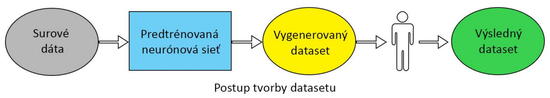
V rámci projektu boli vytvorené dva datasety: dataset pre **detekciu kolies** a dataset pre **klasifikáciu vozidiel**. Ich tvorba prebiehala v poloautomatizovanom režime s využitím predtrénovanej neurónovej siete. Jej výstup bol manuálne odkontrolovaný a v prípade chyby bol tento výstup manuálne upravený.

Dosiahnuté výsledky

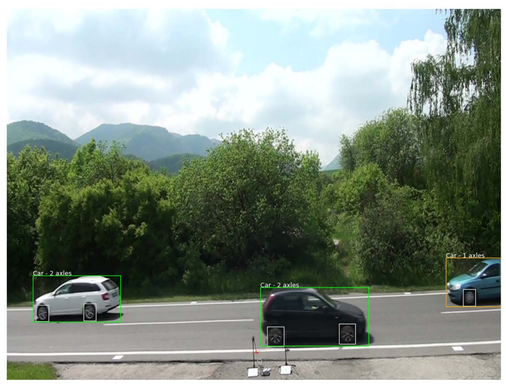
Detekcia kolies	IoU = 0.5	IoU = 0.75
Average Precision	99.90%	92.50%
Average Recall	100.00%	94.70%

Počet náprav	Počet snímkov	Presnosť klasifikácie
2	1269	100%
3	121	100%
4	21	0%
5	231	100%

Neurónová sieť pre klasifikáciu vozidiel dosiahla na testovacej sade celkovú presnosť 98,70%. Vozidlá so 4 nápravami neboli správne rozpoznávané z dôvodu ich nedostatočného zastúpenia v trénovacom datasete.



Postup tvorby datasetu



Detekcia vozidiel a kolies