

Praktikum zo strojového učenia a umelej inteligencie na vizuálnych dátach

Andrej Lúčny

Katedra aplikovanej informatiky FMFI UK

lucny@fmph.uniba.sk

http://dai.fmph.uniba.sk/w/Andrej_Lucny

www.agentspace.org/praktikum

15

Podpora deep learning v OpenCV

- OpenCV podporuje použitie, nie trénovanie modelov deep learningu
- Nepodporuje ich beh na CUDA, len na OpenCL, t.j. využíva grafiku ale nie naplno, má efektívny beh na CPU
- modul dnn
- Podporuje formáty: caffe (.caffe .prototxt), tensorflow (.pb, .pbtxt) a pytorch

Použitie modelu dl

- Model prichádza spravidla v dvoch súboroch, textový popisuje architektúru, binárny váhy
- K modelu treba vedieť rozmery vstupného tenzora (rozlíšenie obrázku), jeho farebný model, rozsah (0..255,-128..128,0..1,-0.5..0.5, ...), priemer dát prípadne aj smerodajnú odchýlku ak sa používajú na normalizáciu vstupu

Použitie modelu dl

- K modelu treba poznať mená výstupných vrstiev siete (dajú sa zistiť z architektúry)
- K modelu treba vedieť interpretovať jeho výstup. Táto interpretácia býva pomerne komplikovaná a uhádnuť ju môže byť veľmi obtiažne
- Treba pamätať na to, že každý model má určité a pomerne malé rozlíšenie, takže relevantným častiam vstupu hrozí, že pri úprave na vstup do siete zaniknú

Kolorizácia obrazu

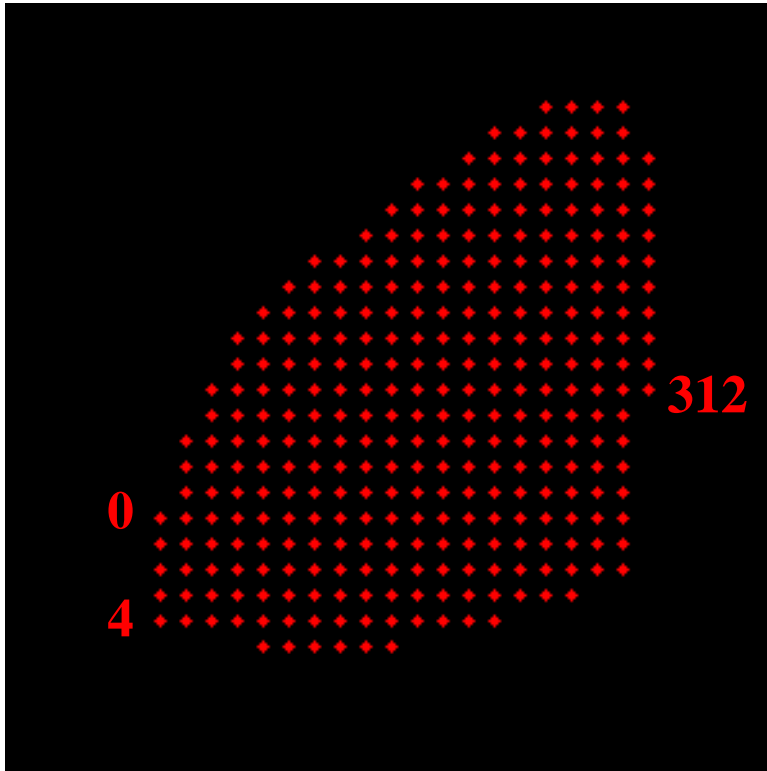


Farebný model

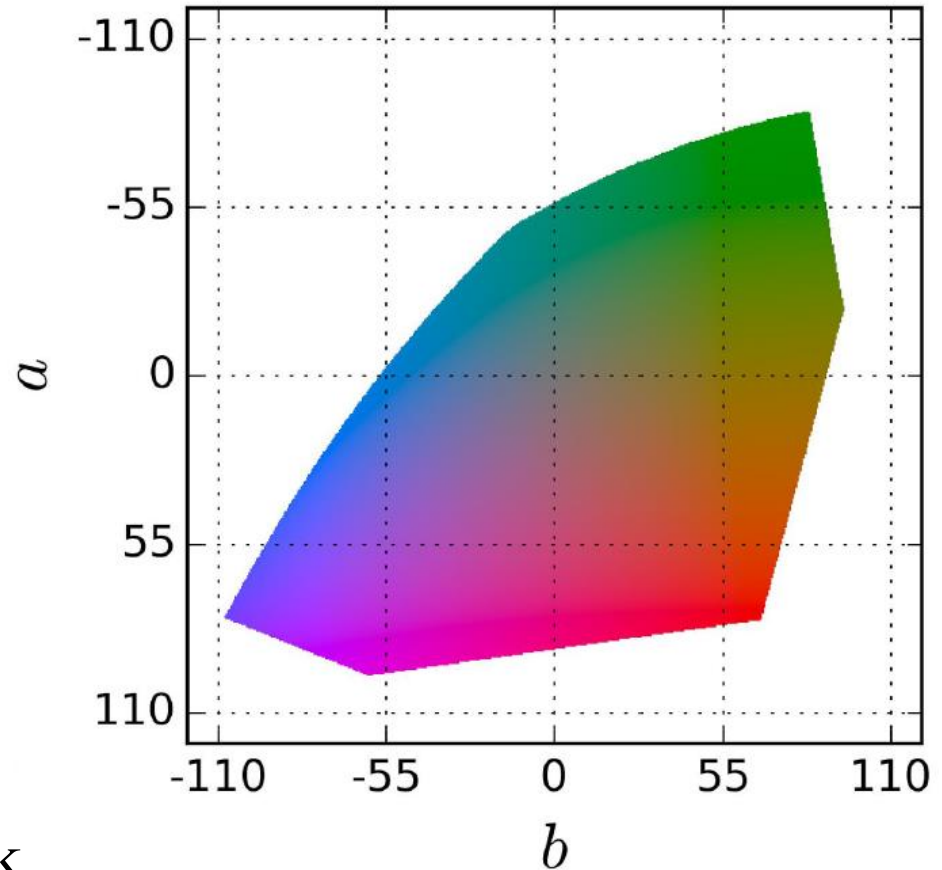
CIE Lab

L .. Intenzita (šedotónový obr.)

Colors in ab space
(continuous)



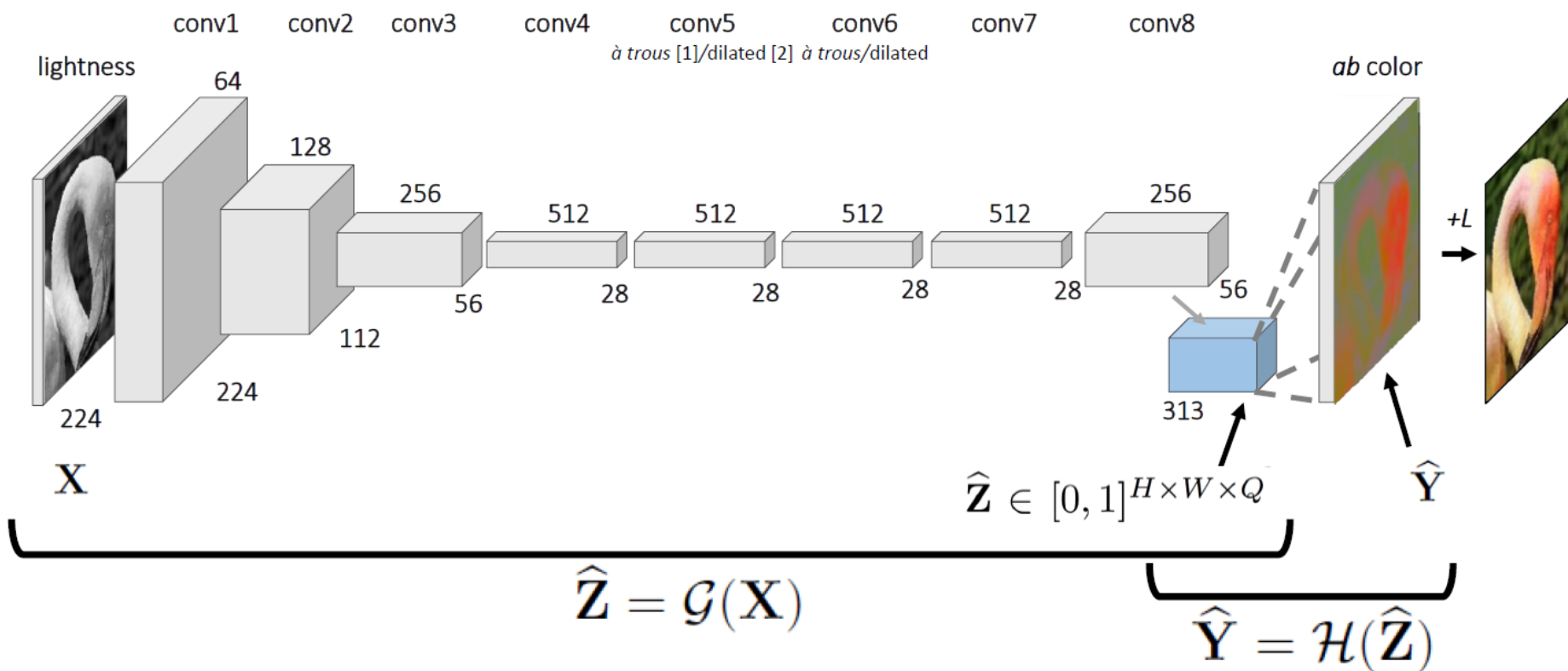
diskretizácia farebných zložiek
farba = vektor 313
pravdepodobností



a .. green-red b .. blue-yellow

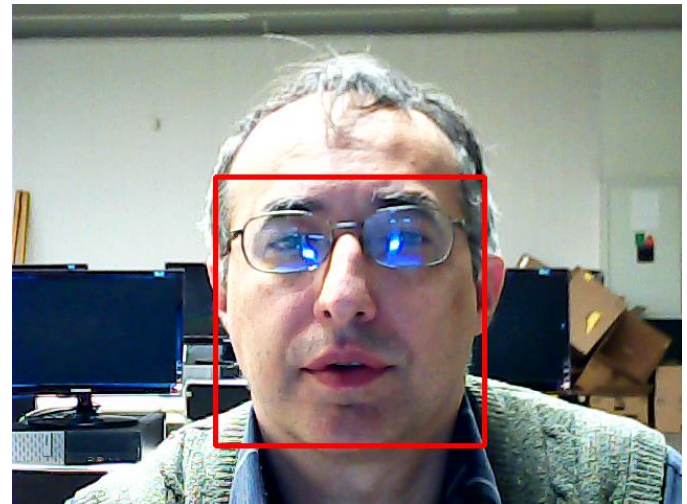
Ofarbenie šedotónových obrázkov

Network Architecture



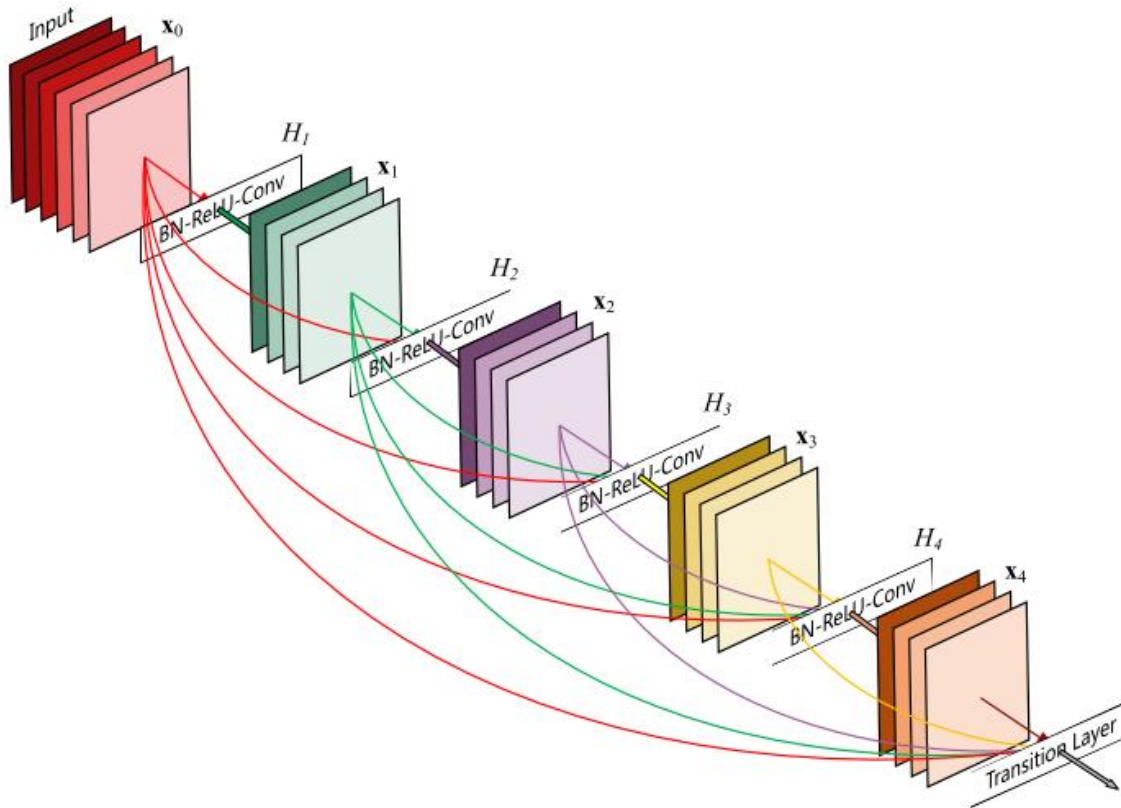
[1] Chen *et al.* In arXiv, 2016.
 [2] Yu and Koltun. In ICLR, 2016

Deep Detector tváří

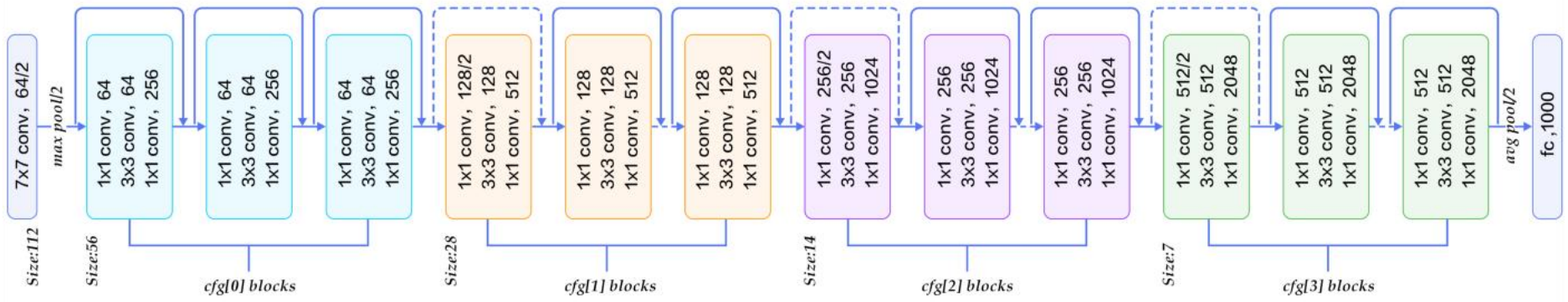


ResNet res10_300x300_ssd_iter_140000

ResNet



50 layers	$cfg=[3,4,6,3]$
101 layers	$cfg=[3,4,23,8]$
152 layers	$cfg=[3,8,36,3]$



Vektorizácia tváří na základe podobnosti



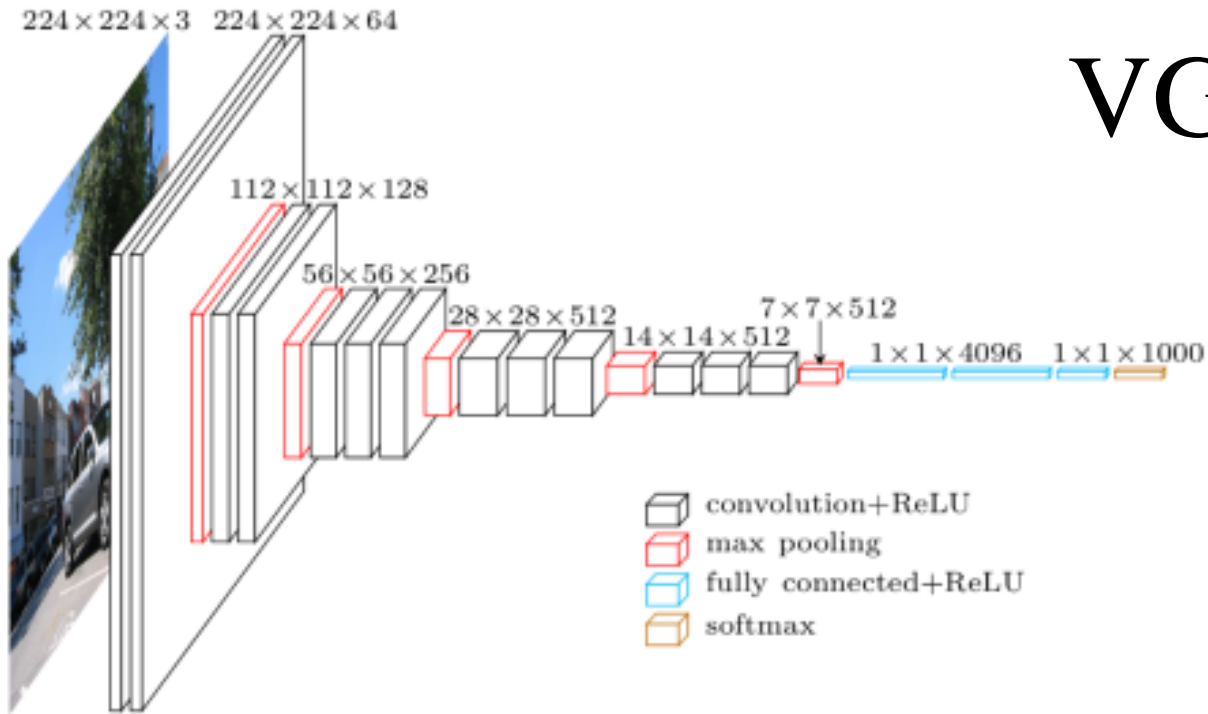
128

```
-0.08919032 0.09024902 0.040942036 -0.04157353 -0.16488977 -0.08485678  
0.015514145 0.0086573735 0.10780367 0.015699599 0.11084026 -0.033236295  
-0.28599826 0.01693825 -0.03396097 0.10043061 -0.11056056 -0.084603205  
-0.20021445 -0.13665384 0.041180193 0.054861303 0.03219283 0.053657733  
-0.14906552 -0.26555598 -0.068431035 -0.13249005 -0.012135312 -0.1647014  
0.039607167 0.047529414 -0.1974987 -0.116168454 -0.0032384843 0.006701291  
-0.031247359 -0.08913836 0.19228819 0.09850126 -0.18168111 0.040185988  
0.117058754 0.31412745 0.25148463 0.03351514 -0.01026143 -0.04951557  
0.13217147 -0.26562682 0.029647294 0.163914 0.14591782 0.08370224  
0.14086226 -0.1545196 0.12179933 0.118593425 -0.25701052 0.07275014  
0.06817722 -0.033007257 -0.081529796 -0.06928277 0.1810219 0.054374013  
-0.11748302 -0.06775741 0.1598805 -0.15249078 -0.054820616 0.09704439  
-0.12356175 -0.17020611 -0.29924074 0.0837964 0.4332566 0.16179526  
-0.21912678 -0.01701596 -0.09125681 -0.023876999 0.017025955 0.04003688  
-0.14529091 -0.07200293 -0.08375634 0.051448725 0.12148529 -8.755233E-4  
-0.022465339 0.19842789 0.07226152 -0.06475312 0.044450887 0.07831155  
-0.19028005 -0.027101355 -0.12790774 -0.003676355 0.096206255 -0.21705586  
-0.041463308 0.07904555 -0.21227759 0.12578832 0.01345009 -0.108141884  
-0.08246433 -0.010525962 -0.12190152 0.027392972 0.17223357 -0.26611778  
0.21453089 0.22108765 -0.05874655 0.09871258 0.05133348 0.08493719  
-0.045902524 0.015397225 -0.09935625 -0.14287551 0.05100835 -0.037032764  
0.026100205 0.026802793
```

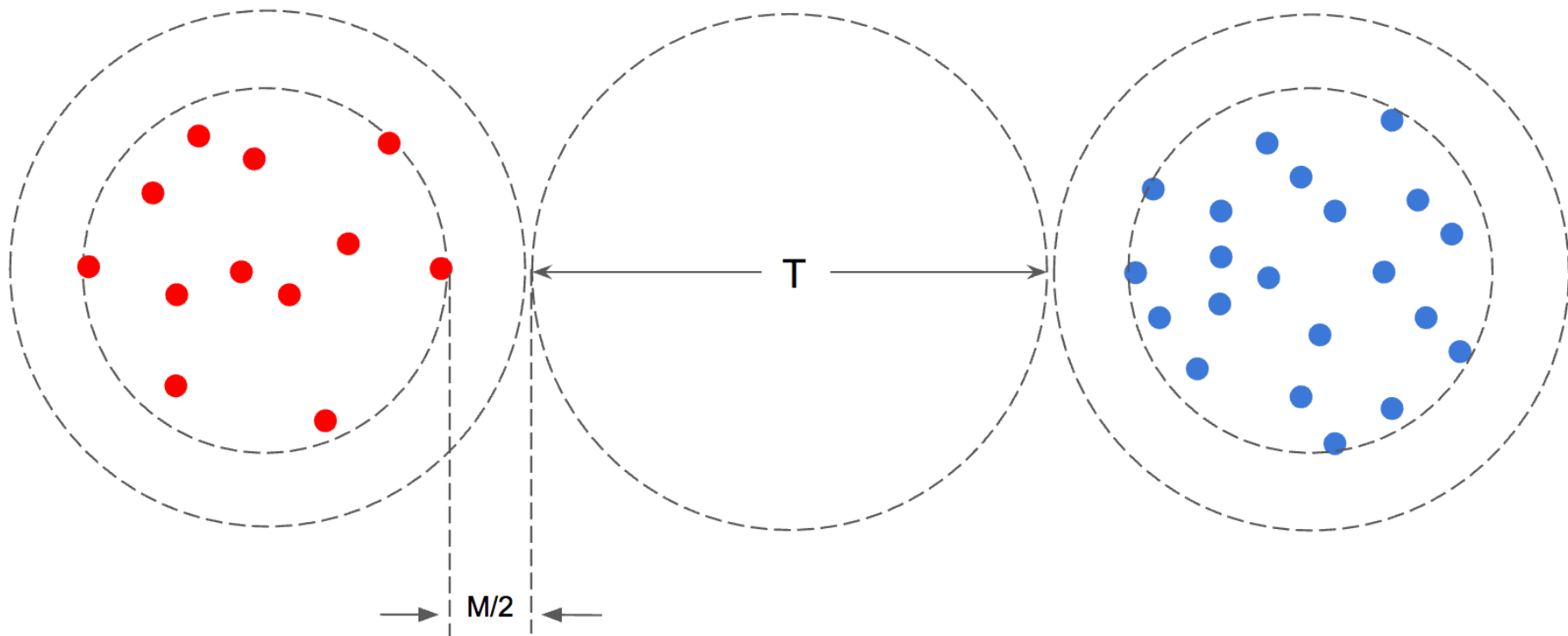
Euklidovská vzdialenosť vektorov zodpovedá nepodobnosti obrázkov

Vectorizer tváří

VGG



Metrická chybová funkcia

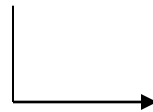


Recognizer tváří

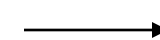


128

```
-0.08919032 0.09024902 0.040942036 -0.04157353 -0.16488977 -0.08485678  
0.015514145 0.0086573735 0.10780367 0.015699599 0.11084026 -0.033236295  
-0.28599826 0.01693825 -0.03396097 0.10043061 -0.11056056 -0.084603205  
-0.20021445 -0.13665384 0.041180193 0.054861303 0.03219283 0.053657733  
-0.14906552 -0.26555598 -0.068431035 -0.13249005 -0.012135312 -0.1647014  
0.039607167 0.047529414 -0.1974987 -0.116168454 -0.0032384843 0.006701291  
-0.031247359 -0.08913836 0.19228819 0.09850126 -0.18168111 0.040185988  
0.117058754 0.31412745 0.25148463 0.03351514 -0.01026143 -0.04951557  
0.13217147 -0.26562682 0.029647294 0.163914 0.14591782 0.08370224  
0.14086226 -0.1545196 0.12179933 0.118593425 -0.25701052 0.07275014  
0.06817722 -0.033007257 -0.081529796 -0.06928277 0.1810219 0.054374013  
-0.11748302 -0.06775741 0.1598805 -0.15249078 -0.054820616 0.09704439  
-0.12356175 -0.17020611 -0.29924074 0.0837964 0.4332566 0.16179526  
-0.21912678 -0.01701596 -0.09125681 -0.023876999 0.017025955 0.04003688  
-0.14529091 -0.07200293 -0.08375634 0.051448725 0.12148529 -8.755233E-4  
-0.022465339 0.19842789 0.07226152 -0.06475312 0.044450887 0.07831155  
-0.19028005 -0.027101355 -0.12790774 -0.003676355 0.096206255 -0.21705586  
-0.041463308 0.07904555 -0.21227759 0.12578832 0.01345009 -0.108141884  
-0.08246433 -0.010525962 -0.12190152 0.027392972 0.17223357 -0.26611778  
0.21453089 0.22108765 -0.05874655 0.09871258 0.05133348 0.08493719  
-0.045902524 0.015397225 -0.09935625 -0.14287551 0.05100835 -0.037032764  
0.026100205 0.026802793
```



Enrollment



Andy

COCO dataset + YOLO detektor



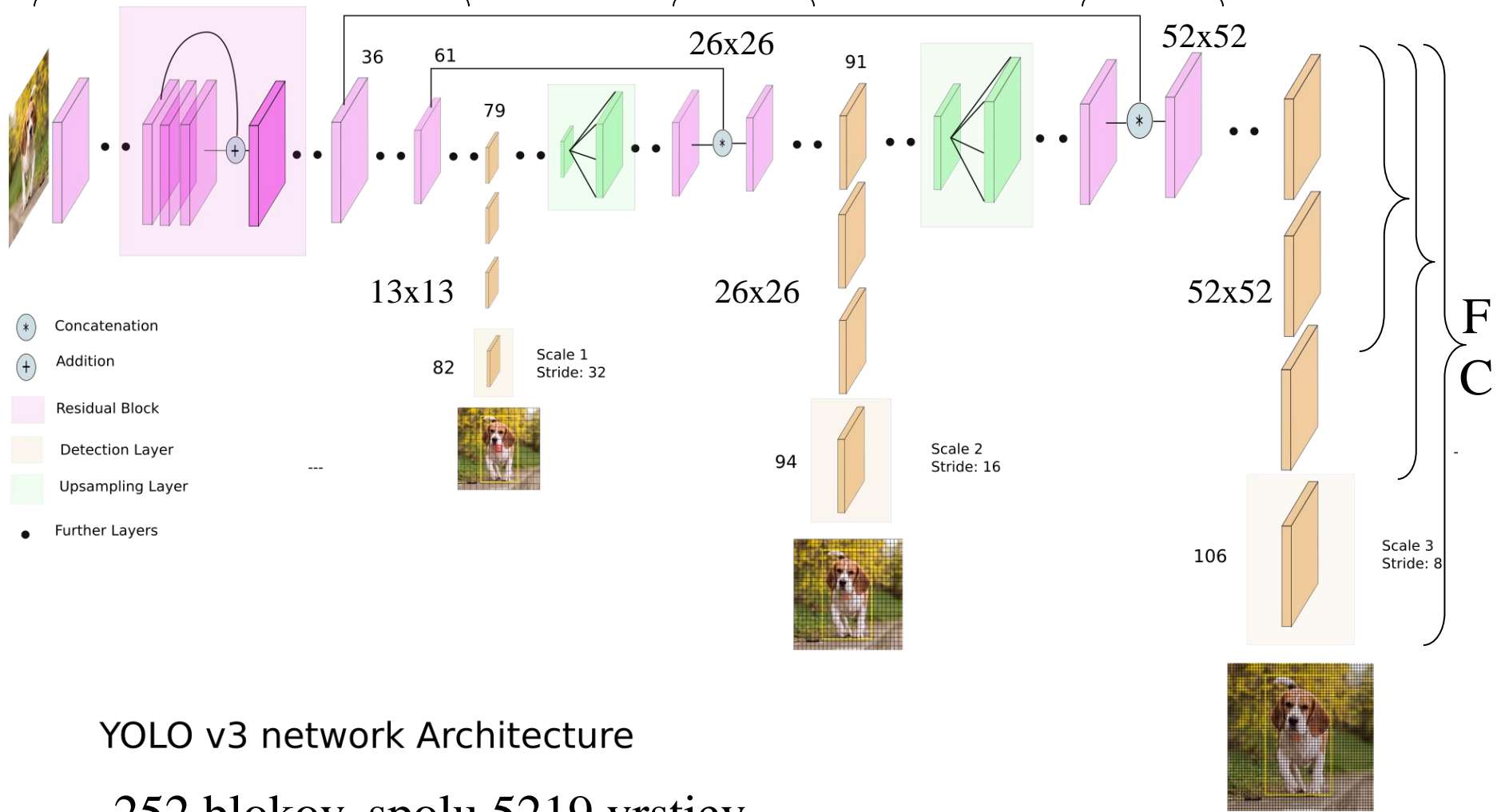
416x416
206x208
104x104
52x52
26x26
13x13

YOLO v3

Encoder

Encoder

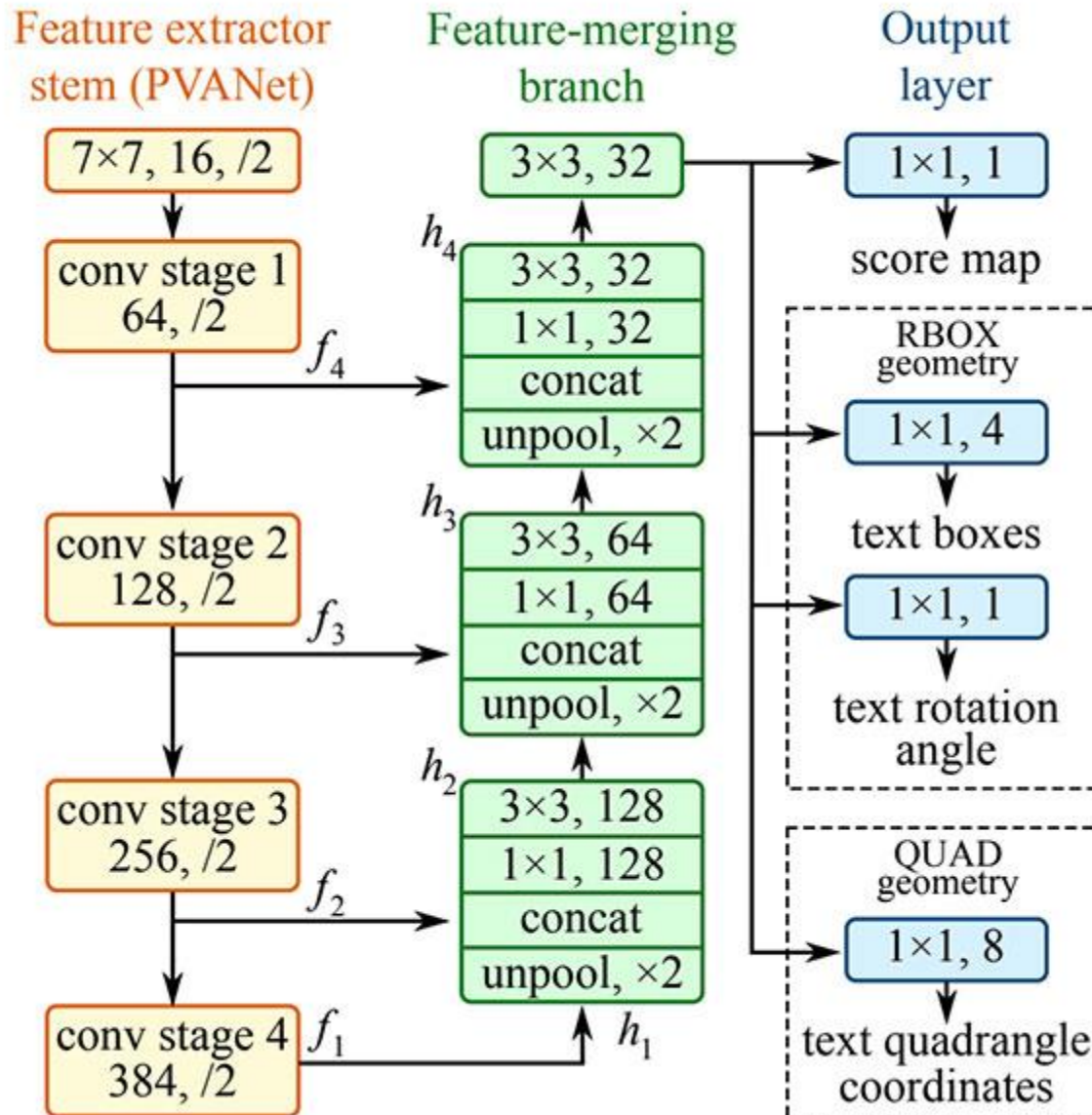
Encoder



YOLO v3 network Architecture

252 blokov, spolu 5219 vrstiev

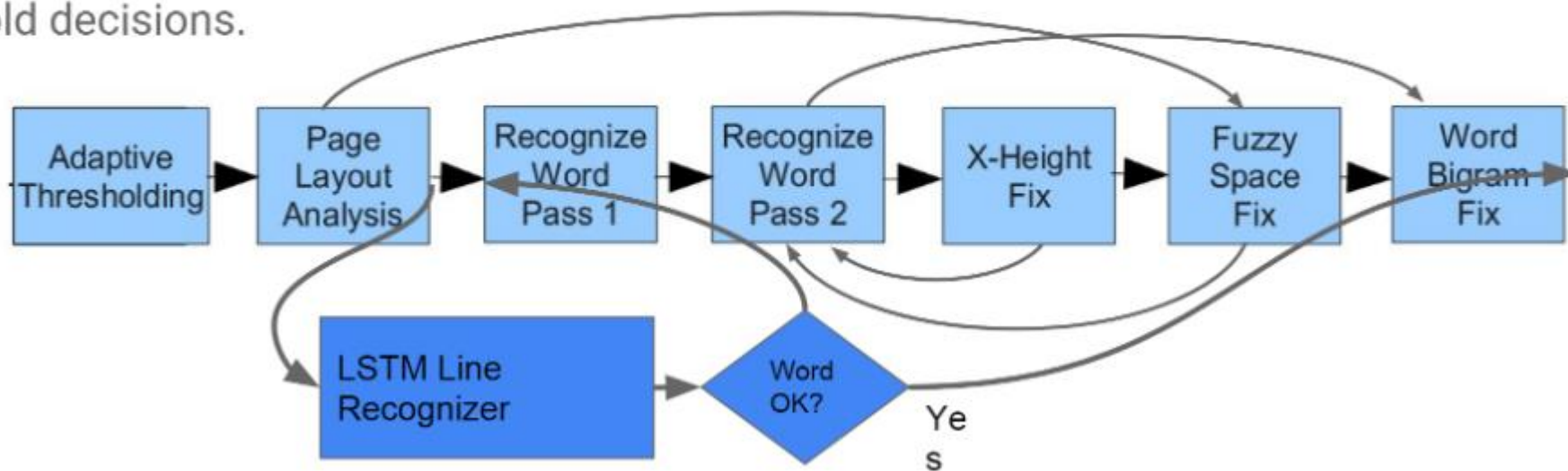
Text Detektor EAST



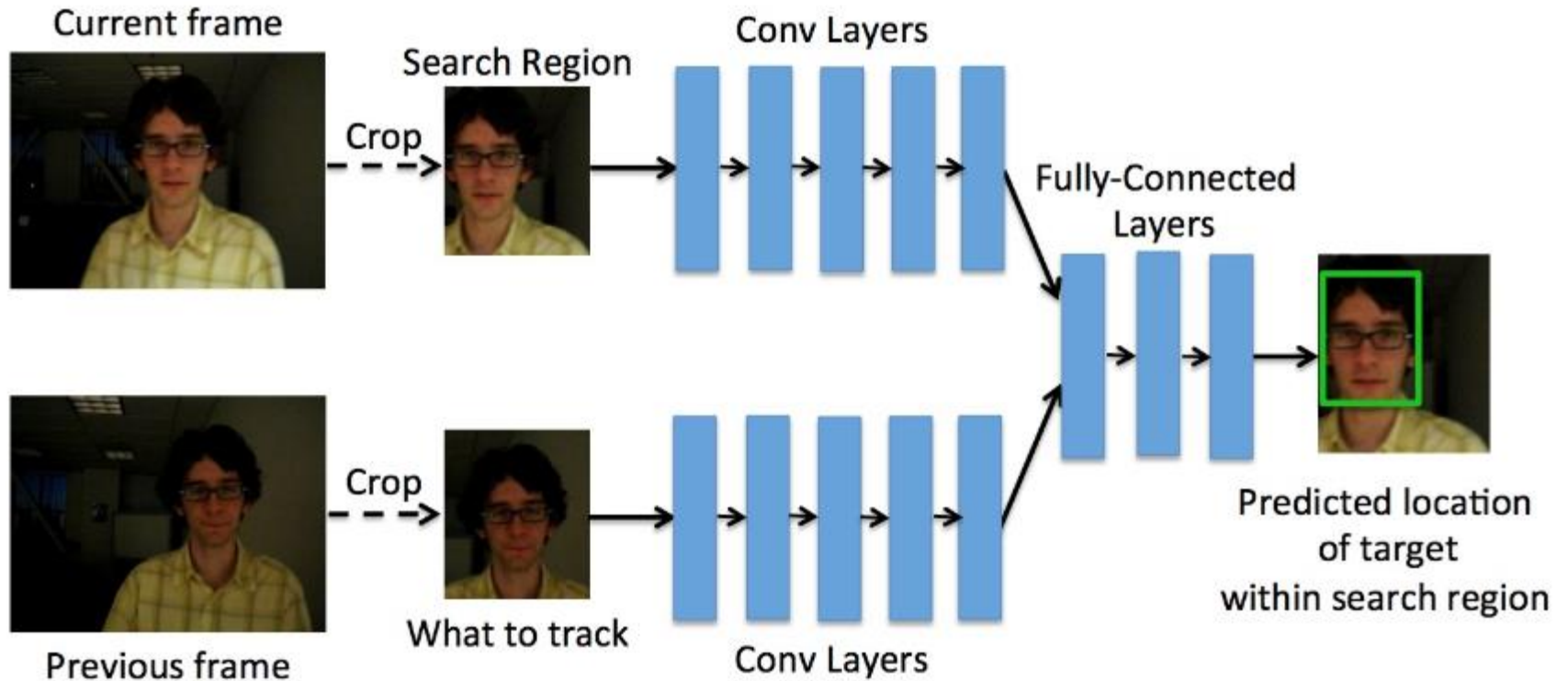
OCR Tesseract na báze LSTM

Tesseract System Architecture

Nominally a pipeline, but not really, as there is a lot of re-visiting of old decisions.



GOTURN Tracker



Model ZOO

<https://github.com/opencv/opencv/wiki/Deep-Learning-in-OpenCV>