

Attēlos redzamu objektu reprezentēšana ar vektoriem: Seju atpazīšanas piemērs

K. Sudars

30. nov. 2016

Aktivitāte: Lielapjoma pilsētvides video un citu sensoru signālu savākšana un analīze, izmantojot augstas veiktspējas skaitļošanas resursus

Uzdevums: Attīstīt video analīzes algoritmus un metodes iepriekš zināmu darbību vai notikumu scenāriju atpazīšanai pilsētas drošības veicināšanai

Grupas cilvēki:

Kaspars Sudars

Roberts Kadiķis

Ričards Cacurs

1. posma apakšuzdevumi (izpildīts uz 2015.g.8.decembri):

- * Apzināt labākos tehnoloģiskos risinājumus video analīzei (Rezultāts: MNT)
- * Iekārtot vidi MNT apmācībai (Rezultāts: Serveris ar 4 x Nvidia Tesla K20 GPU un Caffe, Torch apmācības vides)

2. posma apakšuzdevumi (izpildīts uz 2016.g.16.martu):

- * Skaitļošanas vides un rīku izvēle un sakārtošana (Rezultāts: *TensorFlow* izvēlēta kā galvenā MNT apmācības vide)
- * MNT arhitektūra video analīzes uzdevuma risināšanai (Rezultāts: Dati → CNN → RNN → Rezultāts)

Aktivitāte: Lielapjoma pilsētvides video un citu sensoru signālu savākšana un analīze, izmantojot augstas veiktspējas skaitļošanas resursus

Uzdevums: Attīstīt video analīzes algoritmus un metodes iepriekš zināmu darbību vai notikumu scenāriju atpazīšanai pilsētas drošības veicināšanai

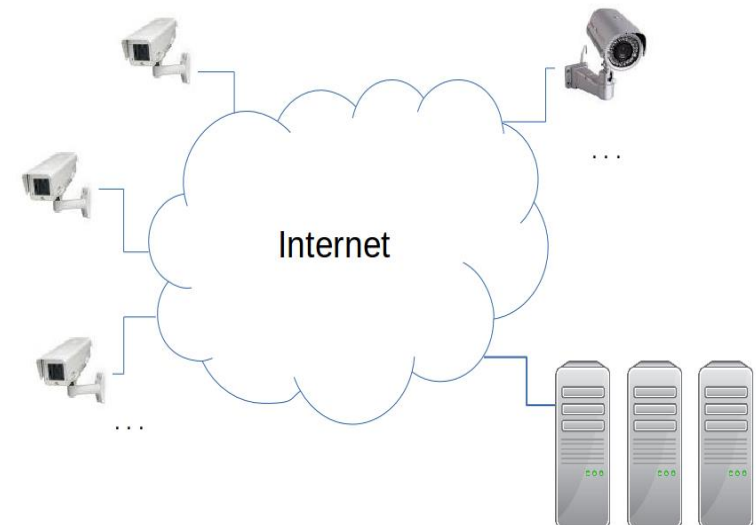
Grupas cilvēki:

Kaspars Sudars
Roberts Kadiķis
Ričards Cacurs

3. posma apakšuzdevumi:

Izstrādāt attēlu apstrādes metodes drošas pilsētvides veicināšanai

- unikālu objektu atpazīšanai video signālos
- objektu skaitīšanai ar līnijas detektoru video signālos



Data gathering and analysis servers

Unikālu objektu atpazīšana attēlos

Uzdevums:

1. Objektu, attēlotu bildē, reprezentēt ar pazīmju vektoru
2. Novērtēt *deep learning* DL pieeju deskriptoru veidošanā

State-of-the-art:

1. Florian Schroff, Dmitry Kalenichenko, James Philbin, *FaceNet: A Unified Embedding For Face Recognition and Clustering*, arXiv:1503.03832, 2015

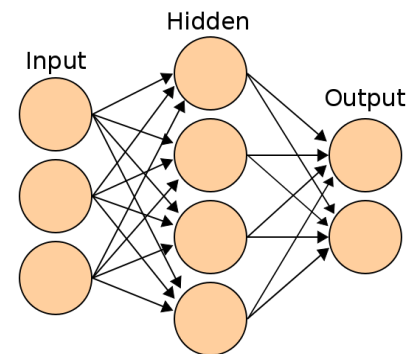
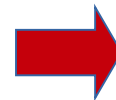
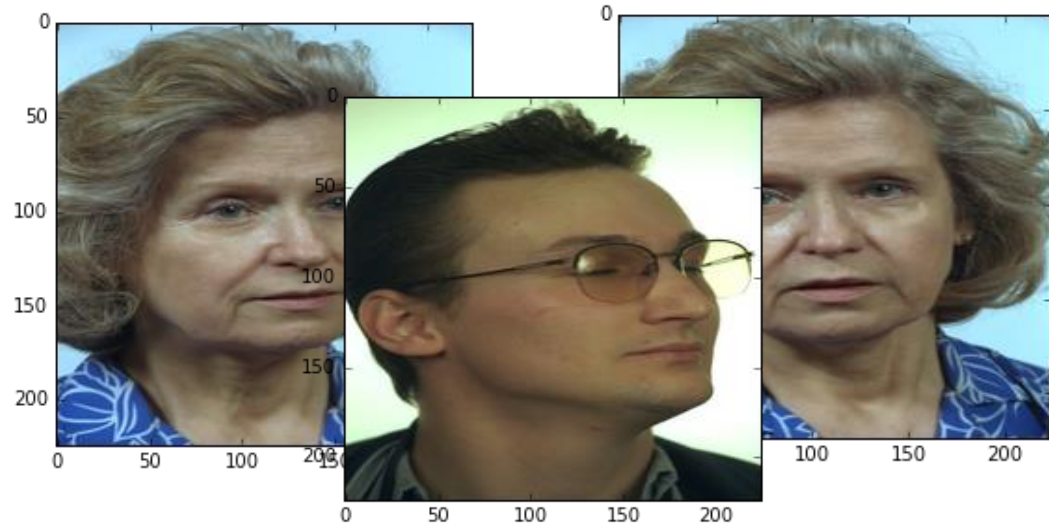
Labeled Faces in the Wild (LFW) 99.63%, On YouTube Faces DB it achieves 95.12%

2. DeepFace achieves 97.35% accuracy on LFW

COLOR-FERET

Treniņa kopa:
10148 attēli
905 klases

Pārbaudes kopa:
1190 bildes
108 klases

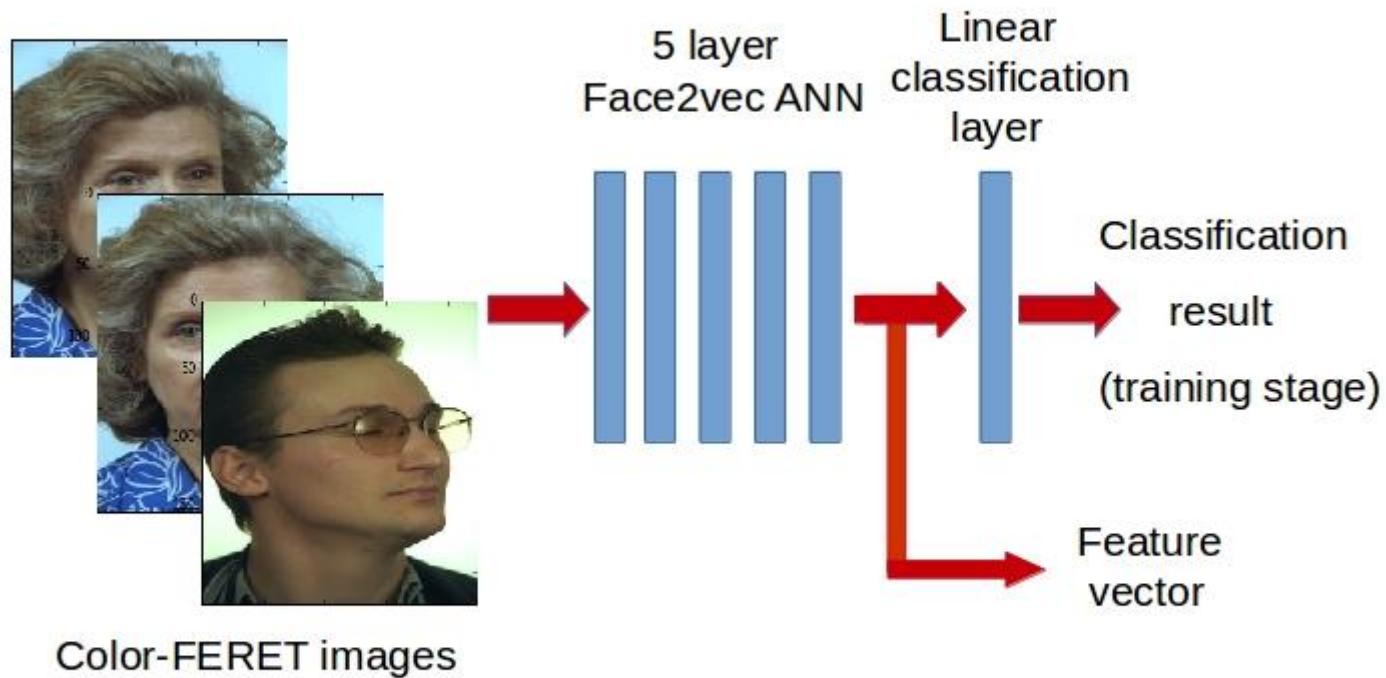


Pazīmju
vektors

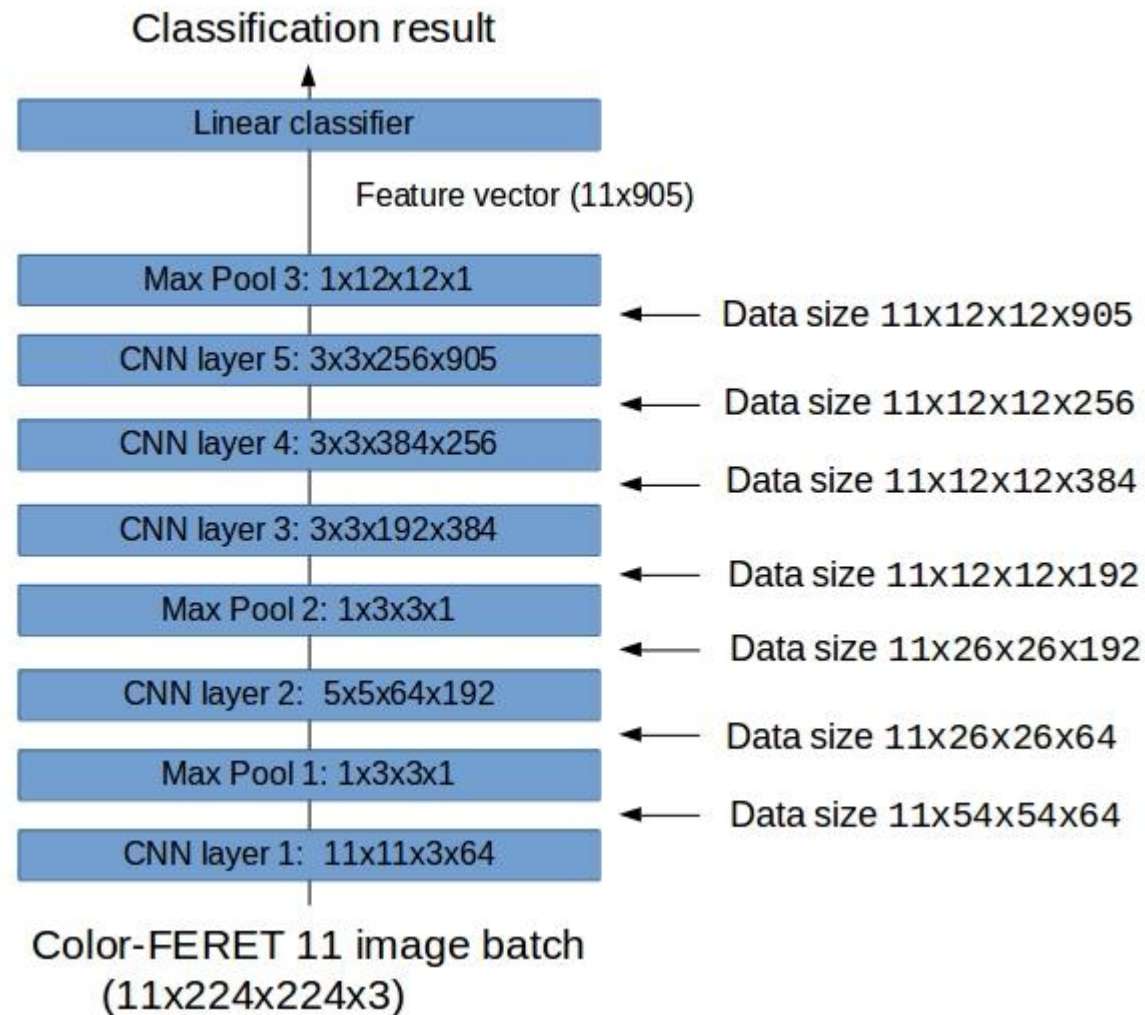
Neironu tīkls

Face2vec

Unikālu objektu atpazīšana



Neironu tīkla arhitektūra



1. Apmācība

$$\hat{\vec{v}}_i = \text{ann}(p_m, \text{Image}_i) \quad (1)$$

$$\text{Loss}(p_m) = \sum_{i=1}^I (\text{ann}(p_m, \text{Image}_i) - \vec{v}_i)^2 \quad (2)$$

$$\min_{p_m} \text{Loss}(p_m) \quad (3)$$

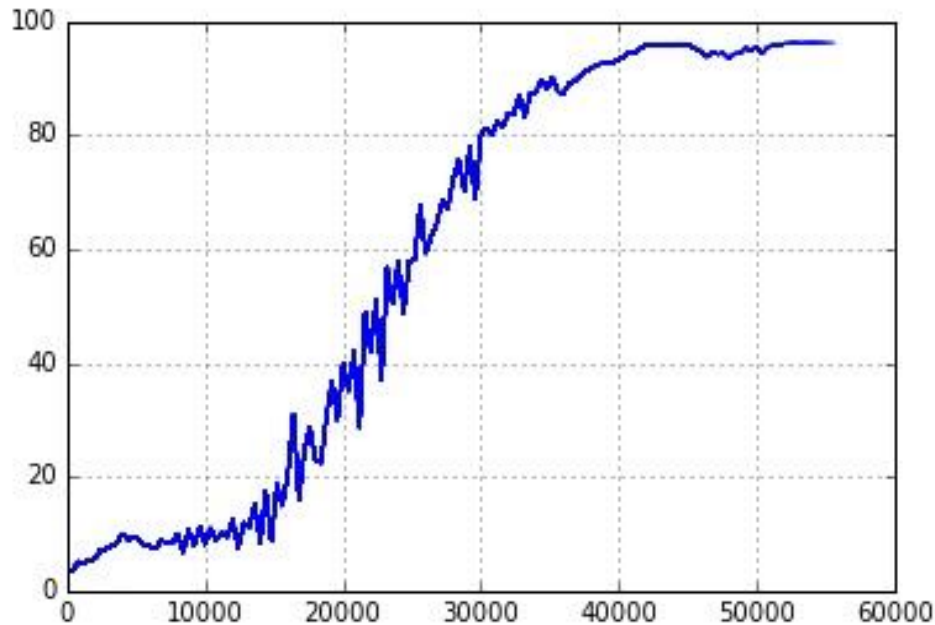
2. Atmet lineāro klasifikatoru

3. Vektoru salīdzināšana (Octave/Matlab):

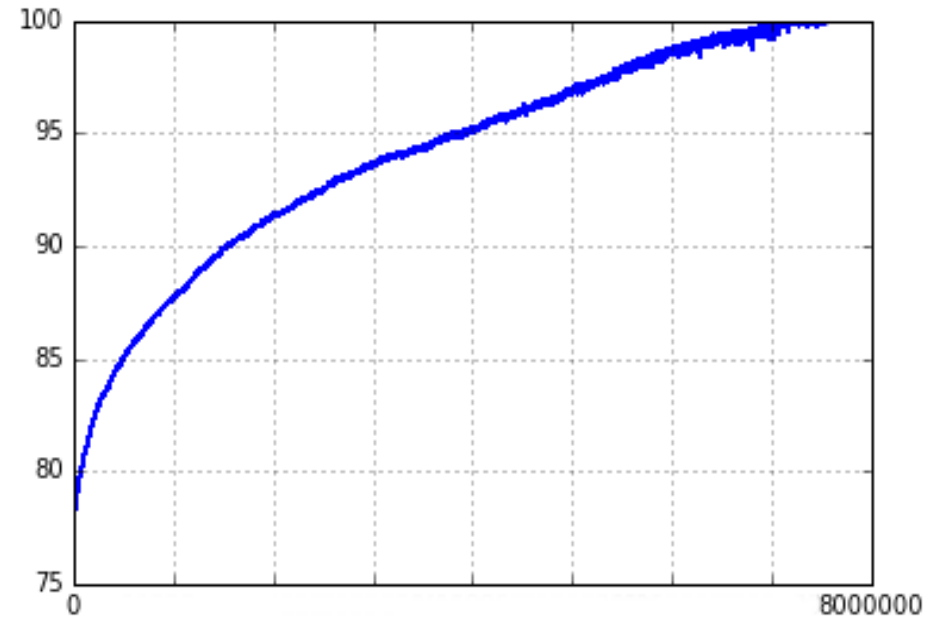
$$\mathbf{x}_i = \mathbf{x}_i - \text{mean}(\mathbf{x}_i) \quad (4)$$

$$\mathbf{x}_i = \mathbf{x}_i / \text{sum}(\text{abs}(\mathbf{x}_i)) \quad (5)$$

$$r_{ij} = \text{sum}(\mathbf{x}_i * \mathbf{x}_j) \quad (6)$$



MNT – 1, LR = 0.01



MNT – 2, LR = 0.0001

X ass: precizitāte [%], Y ass; iterāciju skaits

Apmācības sistēma:

Ubuntu 14.04 + TensorFlow HPC serveris ar Nvidia K40 GPU

1 iterācija ~0.1 sek.



	MNT-1, Color-FERET datubāze	MNT-2, Color-FERET datubāze	MNT-2, Color-FERET-fa- fb datubāze
Klase (personas vārds) atrasta uzreiz, [precizitāte, %]	52.27	87.54	93.94
Klase atrasta 2 mēģinājumos, [%]	68.51	94.87	95.96
Klase atrasta 3 mēģinājumos, [%]	77.02	96.97	97.98

Bijušie rezultāti (VPP 2. posms): Klase atrasta uzreiz 52.27 %, Divas varbūtīgākās klases 68.51 %, 3 varbūtīgākās klases 77.02 %



	MNT-3, Color-FERET-fa- fb, BatchNorm	MNT-4, Color-FERET-fa- fb, BatchNorm, TripletLoss	MNT-5, Color-FERET-fa- fb, BatchNorm, NegDB
Klase (personas vārds) atrasta uzreiz, [precizitāte, %]	87.89	71.69 (~1000 iterācijās)	94.94
Klase atrasta 2 mēģinājumos, [%]	93.94	71.69	97.98

Testējot sistēmu Omkar M. Parkhi, Andrea Vedaldi, Andrew Zisserman, Deep Face Recognition, klase atrasta ar 100 % precizitāti (apmācīta uz 2.6M bildēm no 2.6K cilvēkiem).



Pieejamās marķētās datu kopas:

MS-Celeb-1M: Challenge of Recognizing One Million Celebrities in the Real World

YouTube-8M Dataset: large-scale labeled video dataset that consists of 8 million YouTube video IDs and associated labels from a diverse vocabulary of 4800 visual entities.

IMDB-WIKI – 500k+ face images with age and gender labels

Seju atpazīšanas sistēma sasniedz 94.94% precizitāti uz COLOR-FERET datu bāzes

Pieņemts publicēšanai raksts K. Sudars, Face Recognition Face2vec Based on Deep Learning: Small Database Case, Automatic Control and Computer Science

Publicēts raksts K. Sudars, I.Bilinskis, E.Buls, Discrete Fourier Transform of the signals recovered by using high-performance Event Timers, Baltic Electronics Conference 2016, Tallinn, Estonia