

Video analīze ar *Dziļo* apmācību: BigData paveiktais

K.Sudars

8. decembris 2015

Aktivitāte: Lielapjoma pilsētvides video un citu sensoro 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

1. posma apakšuzdevumi:

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

2. posma apakšuzdevumi:

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

Grupas cilvēki: Kaspars Sudars, Roberts Kadiķis, Ričards Cacurs

2. posma darbības virzieni:

- * Skaitļošanas vides sakārtošana

Serveris ar 4 x Nvidia Tesla K20 GPU => Klāsteris ar 13 K40 GPU + Rocks

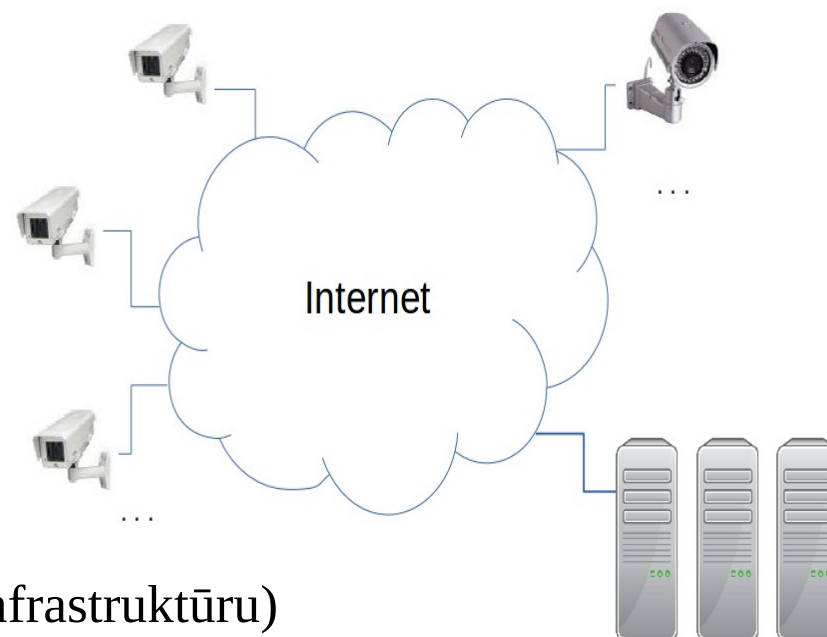
- * Dziļās apmācības rīku analīze

Pārslēgšanās darbā uz TF (iznācis 2015 nov.)

- * MNT arhitektūra uzdevuma risināšanai

“Dati → CNN → RNN → Rezultāts”

Face2vec, Autoencoder



Data gathering and analysis servers

Datu savākšanas un apstrādes koncepcija:

- * Izmantot esošo infrastruktūru

(gan eksistējošus hw/sw risinājumus, gan Edi infrastruktūru)

- * Edi **TestBed** video savākšana ar 40 videokamerām

(simulē reālos datu savākšanas apstākļus)

- * Savākto datu klāsterizācija ar *dziļo apmācību* konkrētu problēmu risināšanai (piem., objektu atpazīšana dažādās kamerās jeb sekošana, personu identificējošas informācijas aizsardzība un slēpšana video)

TensorFlow - Google's latest machine learning system, open sourced for everyone

http://googleresearch.blogspot.com/2015/11/tensorflow-googles-latest-machine_9.html



Google Research Blog

The latest news from Research at Google

Priekšrocības:

- Python API
- multi GPU
- fleksibilitāte (p., LSTM, loops)
- Google komandas atbalsts
- Autograd

Alternatīvas:

- Caffe
- Torch
- Veles no Samsung
- Digits uc.

TensorFlow - Google's latest machine learning system, open sourced for everyone

Monday, November 09, 2015

Posted by Jeff Dean, Senior Google Fellow, and Rajat Monga, Technical Lead

Deep Learning has had a huge impact on computer science, making it possible to explore new frontiers of research and to develop amazingly useful products that millions of people use every day. Our internal deep learning infrastructure [DistBelief](#), developed in 2011, has allowed Googlers to build ever larger [neural networks](#) and scale training to thousands of cores in our datacenters. We've used it to demonstrate that [concepts like "cat"](#) can be learned from unlabeled YouTube images, to improve speech recognition in [the Google app](#) by 25%, and to build image search in [Google Photos](#). DistBelief also trained the Inception model that won Imagenet's [Large Scale Visual Recognition Challenge in 2014](#), and drove our experiments in [automated image captioning](#) as well as [DeepDream](#).

While DistBelief was very successful, it had some limitations. It was narrowly targeted to neural networks, it was difficult to configure, and it was tightly coupled to Google's internal infrastructure - making it nearly impossible to share research code externally.

Search blog ...



Research at Google

google.com/+ResearchatG

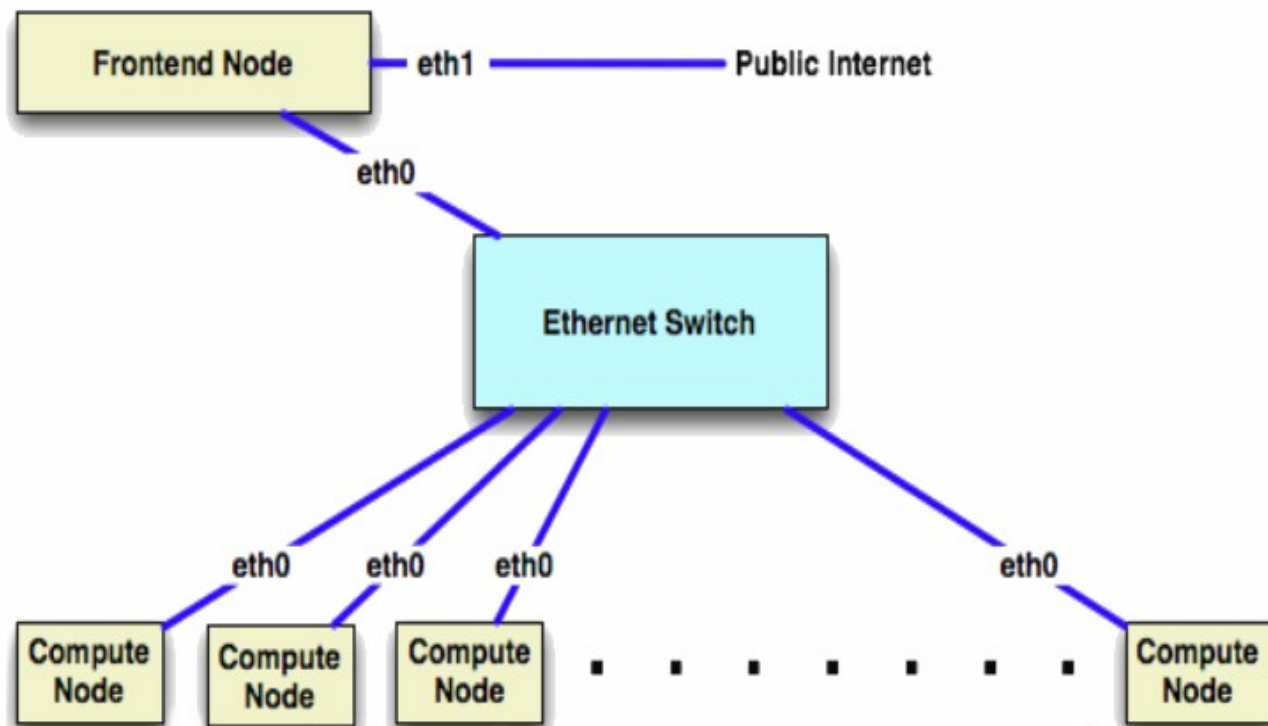
vx, CS+x



Sekot

+ 1 153 729

Labels



- 13 serveri, kur 1 centrālais serveris un 12 skaitļošanas serveri.
- Serveru hardware:
 - 2 Intel Xeon E5-2650 procesori, 2.6 GHz 8 kodoli.
 - 64 GB RAM
 - nVidia Tesla K40 grafiskās kartes (tikai skaitļošanas serveriem):
 - 1.43 Tflops double precision performance
 - 12 GB atmiņa



Uzstādīta Linux Rocks 6.2 distribūcija (Balstīta uz CentOS).



Ietver vairākas programmatūru pakas, kā piemēram:

- Ganglia - resursu monitoringa rīks
<http://10.13.137.133/ganglia>
- SGE (Oracle Grid Engine) – nodrošina izklaidētu procesu izpildi un pārraudzīšanu
- Programmēšanas un izpildīšanas rīki dažādām valodām: Java, Python, C/C++
- Trūkumi, nav programmatūru pakas video kartes draiveriem un programmatūras izstrādei, nepieciešams attiecīgo programmatūru uzstādīt un konfigurēt pašiem.
- **Nepieciešams uzstādīt neironu tīklu programmatūras platformas (TensorFlow, Caffe etc), un saprast kā tās izmantot vairāku skaitļošanas mašīnu vidē.**



cr_AlexNet
Image Classification Model

Delete Job

Solver

[solver.prototxt](#)

Network (train/val)

[train_val.prototxt](#)

Network (deploy)

[deploy.prototxt](#)

Raw caffe output

[caffe_output.log](#)

Dataset

[cats_and_rabbits](#)

Done Nov 09, 04:48:07 PM

Image Size

256x256

Image Type

COLOR

Create DB (train)

1974 images

Create DB (val)

658 images

Encoding

png

Job Status Done

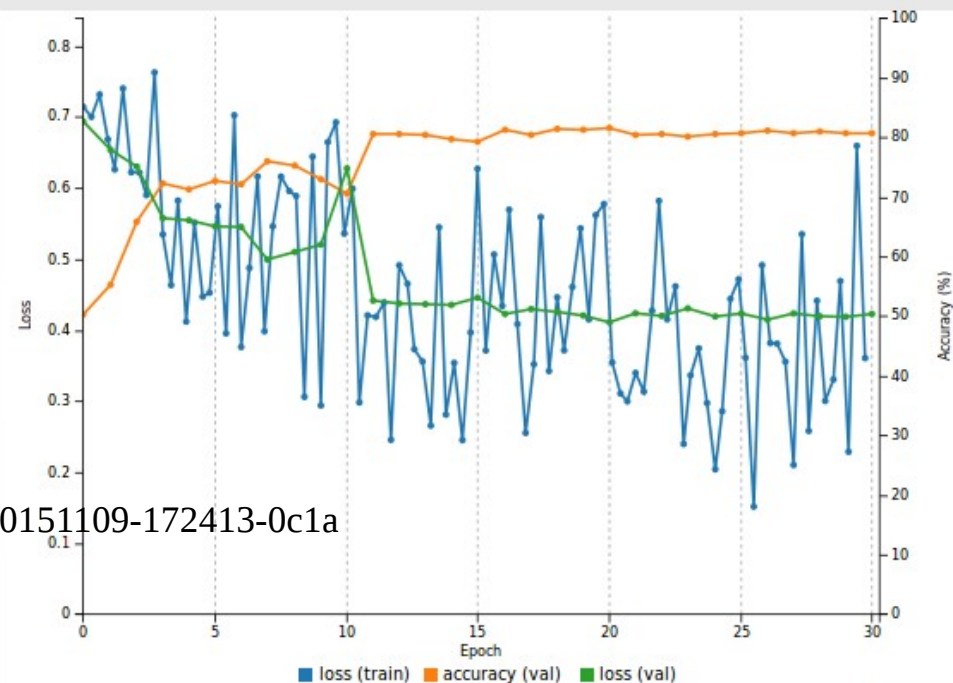
- Initialized at Nov 09, 05:24:13 PM (1 second)
- Running at Nov 09, 05:24:14 PM (4 minutes, 47 seconds)
- Done at Nov 09, 05:29:00 PM (Total - 4 minutes, 47 seconds)

Train Caffe Model Done ▾

“Kaķis vai Zaķis” detektors DIGITS-2 vidē

Klasifikācija pie
marķētiem datiem

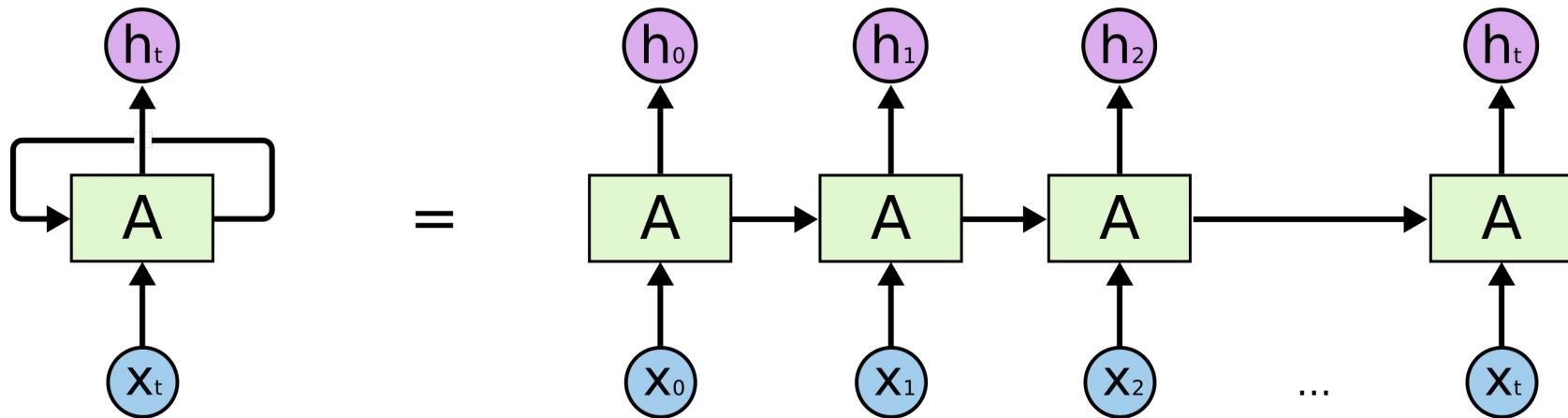
<http://10.13.137.138:5000/models/20151109-172413-0c1a>



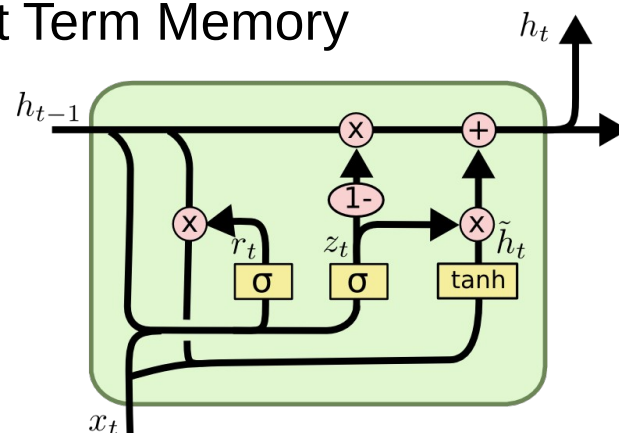
View Large

RNN

- RNN iekļauj atgriezeniskas saites, kas ļauj iemācīties laikā mainīgas pazīmes



- LSTM – Long Short Term Memory



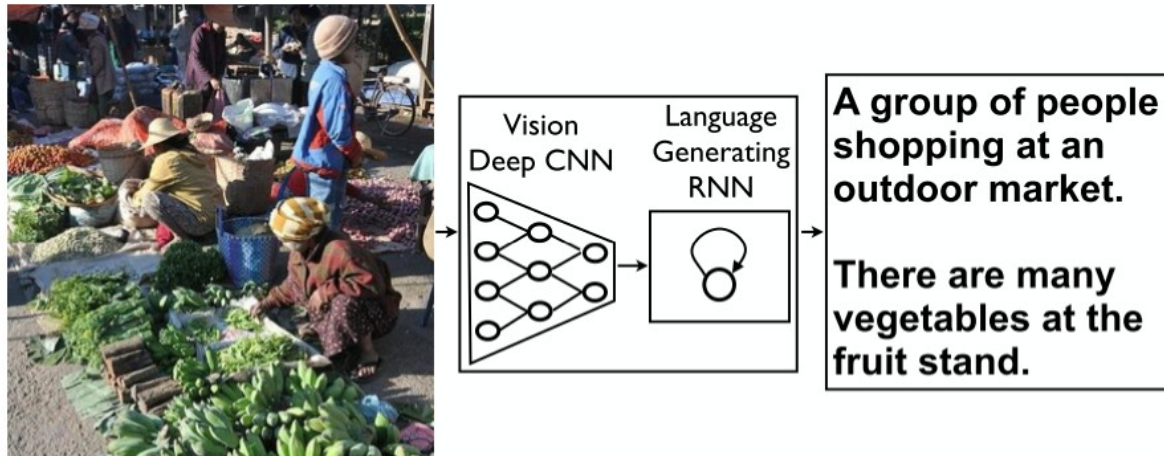
$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t])$$

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t])$$

$$\tilde{h}_t = \tanh(W \cdot [r_t * h_{t-1}, x_t])$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t$$

- HTM – Hierarchical Temporal Memory. Alternatīvs risinājums no Numenta



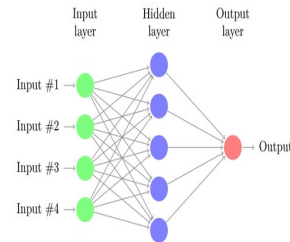
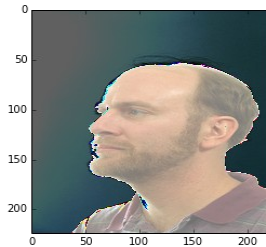
- Neural-talk: <https://github.com/karpathy/neuraltalk2>
- Neural-storyteller: <https://github.com/ryankiros/neural-storyteller>



A few boys ride their skateboards on the bridge.

- A group of people standing on the bridge.
- A couple of guys that are next to bridge
- A group of young guys riding a skateboard across a bridge.
- People walking on a bridge with much luggage

Face2vec



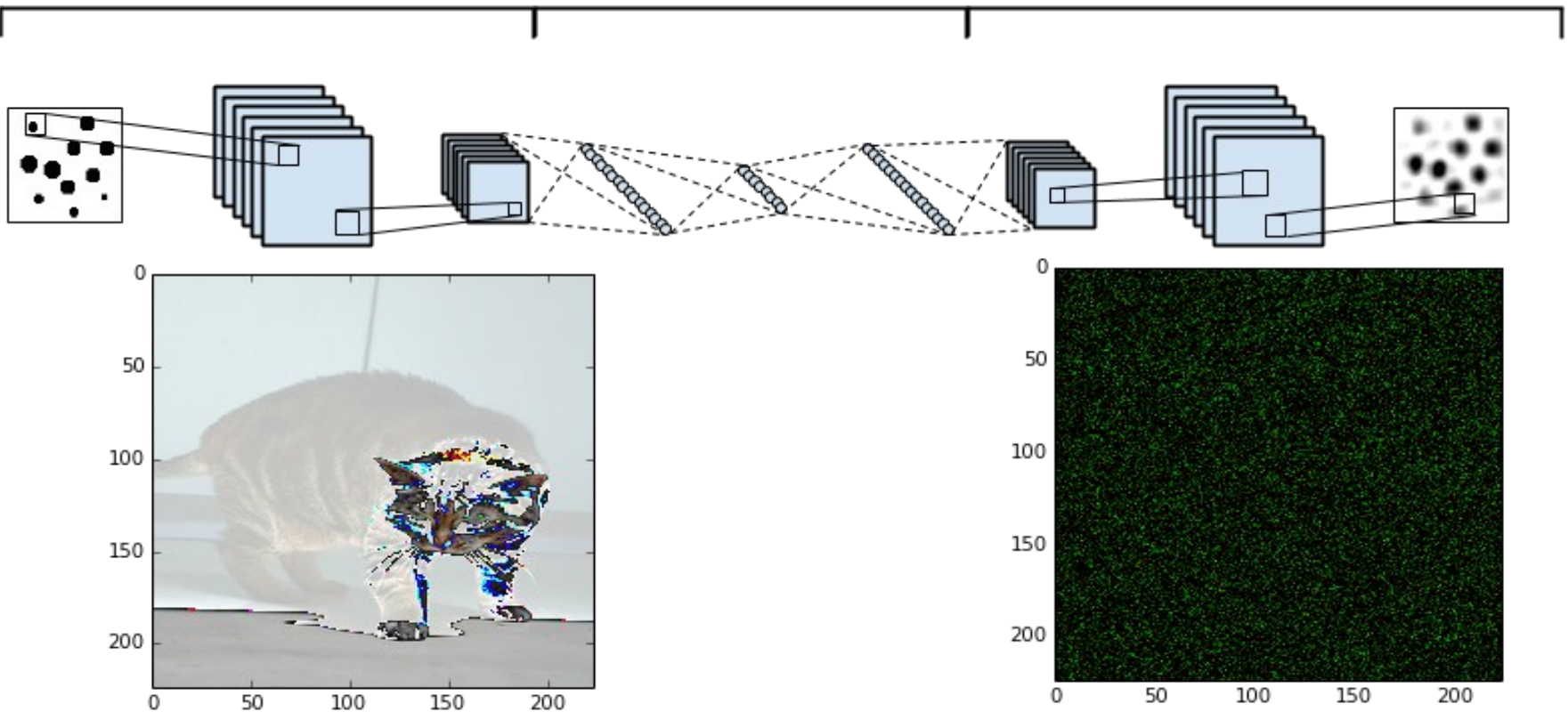
Pazīmju vektors

Sparse Auto-Encoder

Convolution step (convolution + pooling)

Fully connected encoding step

Deconvolution step (deconvolution + unpooling)



Informācijas apmaiņa DL jomā

- arXiv.org
- GitHub.com
- GitXiv.com



[About](#) [Competitions](#) [Categories ▾](#)

GitXiv [Register](#) [Sign In](#)



Sequence to Sequence – Video to Text

Generate a description of events in a video with LSTMs

DEEP LEARNING (DL) RECURRENT NEURAL NETWORKS (RNN) GENERATIVE

samim 2 points 16 hours ago 0 Comments

arXiv

Subhashini Venugopalan - Marcus Rohrbach - Jeff Donahue - Raymond Mooney - Trevor Darrell - Kate Saenko

University of Texas at Austin - University of California, Berkeley - University of Massachusetts, Lowell 4 - International Computer Science Institute, Berkeley

Real-world videos often have complex dynamics; and methods for generating open-domain video descriptions should be sensitive to temporal structure and allow both input (sequence of frames) and output (sequence of

GitHub

github-user: vsubhashini

The network is currently supported by the recurrent branch of the Caffe fork in my repository or Jeff's repository but are not yet compatible with the master branch of Caffe.

[Read more...](#)

Dependencies:

Links

- <https://vsubhashini.github.io/s2vt.html>.
- <https://www.youtube.com/watch?v=-xNI7e7YgDk>.
- <https://www.youtube.com/watch?v=pER0mjzSYaM>.
- <https://www.youtube.com/watch?v=IGaAoW8bA4c>.

<http://visual.ly/paperscape-tool-visualizing-arxiv?view=true>

Paveiktais:

- * Apgūts TF priekš MNT apmācības (neaizmirstot par pārējām vidēm)
- * MNT struktūra “Dati → CNN → RNN → Rezultāts”



Tālāk:

- * “Dati → CNN → RNN → Rezultāts” piemērs pilsētu drošībai
- * Face2vec piemērs
- * Sparse Auto-Encoder nemarķētiem datiem
- * Rocks linux klāsteris

