

Sintētisko datu ģenerēšana pašbraucošo auto apmācībai



Maksims Ivanovs
LU DF un EDI, doktorants

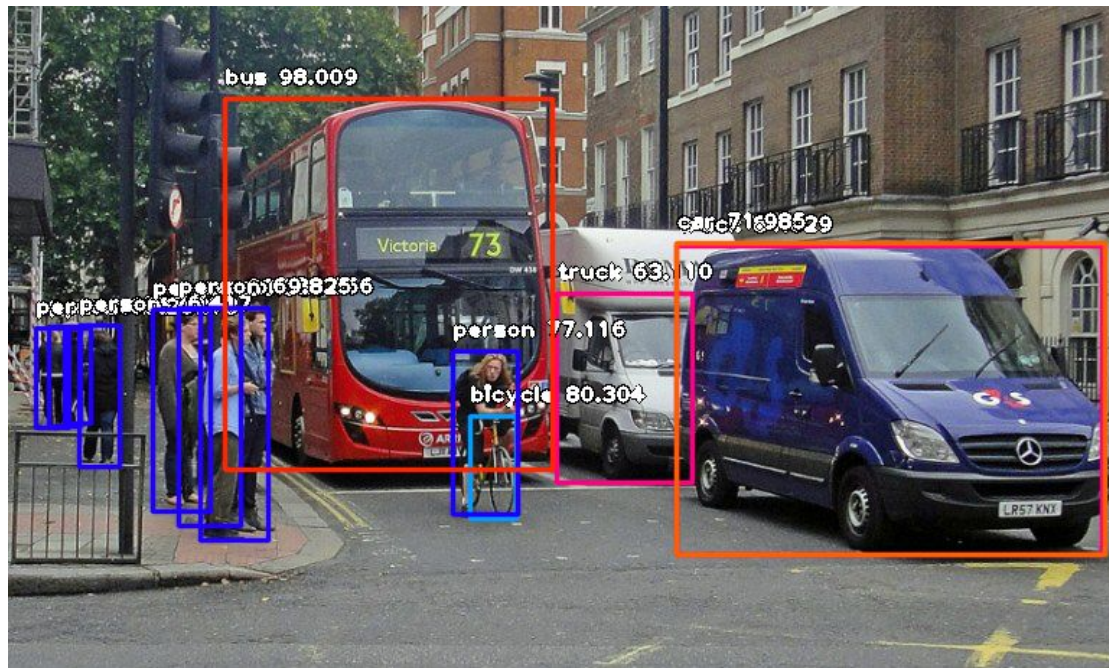
Rīga, 2020. gada 22. aprīlis

Saturs

- Pašbraucošo auto apmācība: mērķu definēšana
- Mašīnmācīšanās paradigma
- Mākslīgie neironu tīkli
- Dziļā mašīnmācīšanās
- Sintētiskie dati
- Sintētiskie dati pašbraucošo mašīnu apmācībai
- Kopsavilkums

Pašbraucošo auto apmācība: mērķi (1)

Objektu atpazīšana



Pašbraucošo auto apmācība: mērķi (2)

Semantiskā segmentācija



Pašbraucošo auto apmācība: mērķi (3)

Semantiskā segmentācija



Semantiskā segmentācija ar Xception71 mākslīgo neironu tīklu

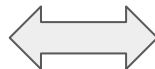
Pašbraucošo auto apmācība: mērķi (4)

Prasības - attēlu apstrādei jānotiek:

- Automātiski
- Ātri
- Precīzi

Vai to var panākt ar tradicionālām programmēšanas paradigmām!?

```
Test.py x
1 num = 10
2 if (num == 0):
3     print("Number is Zero")
4
5 elif (num > 5):
6     print("Number is greater than 5")
7
8 else:
9     print("Number is smaller than 5")
10 |
```



*Cilvēkam ir 2 kājas, ķermenis, galva, 2
rokas un...
Mašīnai ir četri riteņi un...
Kokam ir stumbrs, zari un...*

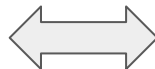
Pašbraucošo auto apmācība: mērķi (4)

Prasības - attēlu apstrādei jānotiek:

- Automātiski
- Ātri
- Precīzi

Vai to var panākt ar tradicionālām programmēšanas paradigmām!?

```
Test.py x
1 num = 10
2 if (num == 0):
3     print("Number is Zero")
4
5 elif (num > 5):
6     print("Number is greater than 5")
7
8 else:
9     print("Number is smaller than 5")
10 |
```



*Cilvēkam ir 2 kājas, ķermenis, galva, 2
rokas un...
Mašīnai ir četri riteņi un...
Kokam ir stumbrs, zari un...*

Ne gluži!

Mašīnmācīšanās paradigma

- Definīcija (Mitchell, 1997):

Definition: A computer program is said to **learn** from experience E with respect to some class of tasks T and performance measure P , if its performance at tasks in T , as measured by P , improves with experience E .

- Īpaši noder uzdevumiem, kurus grūti atrisināt algoritmiski, piemēram:
 - Balss atpazīšana
 - Rekomedācijas sistēmas
 - **Attēlu atpazīšana**
 - **Attēlu segmentācija**

Mašīnmācīšanās paradigma

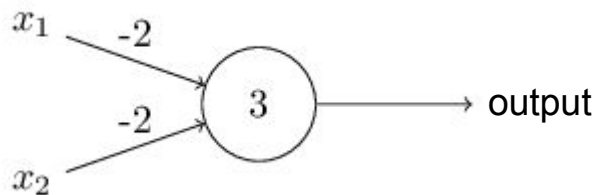
- Definīcija (Mitchell, 1997):

Definition: A computer program is said to **learn** from experience E with respect to some class of tasks T and performance measure P , if its performance at tasks in T , as measured by P , improves with experience E .

- Īpaši noder uzdevumiem, kurus grūti atrisināt algoritmiski, piemēram:
 - Balss atpazīšana
 - Rekomedācijas sistēmas
 - Attēlu atpazīšana
 - Attēlu segmentācija
- **Jautājums: kā to īstenot praksē!?**

Mākslīgie neironu tīkli (1)

Perceptrons:



$$\text{output} = \begin{cases} 0 & \text{if } w \cdot x + b \leq 0 \\ 1 & \text{if } w \cdot x + b > 0 \end{cases}$$

00 -> 1

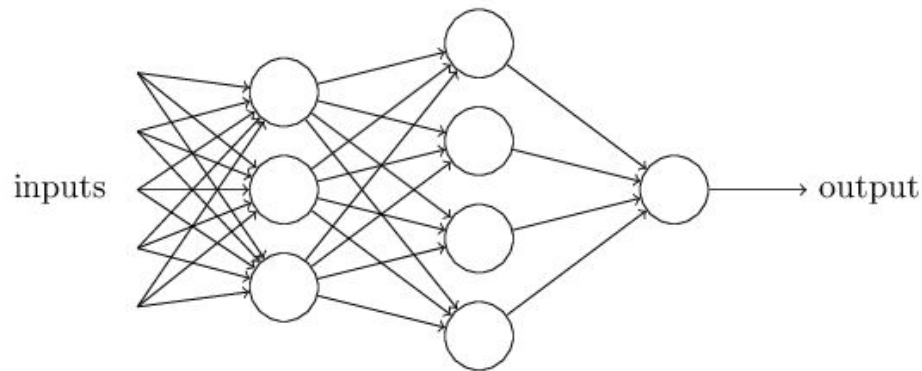
01 -> 1

10 -> 1

11 -> 0

NAND vārti

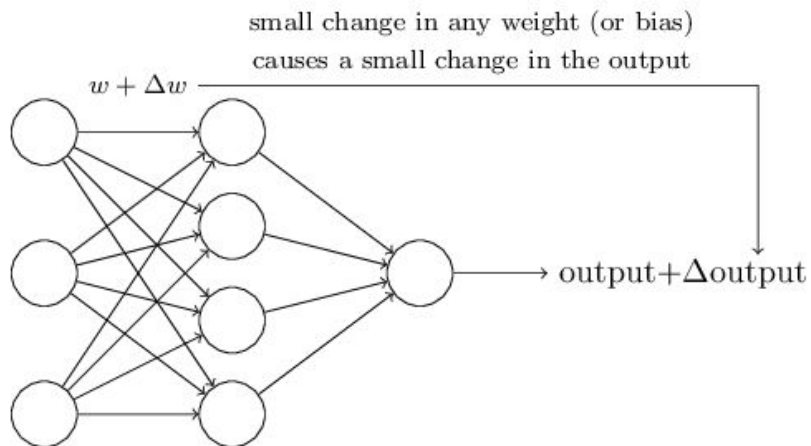
Perceptronu tīkls:



Šāds tīkls potenciāli var pieņemt sarežģītus lēmumus (piemēram, veikt attēla segmentāciju), bet to vajag apmācīt

Mākslīgie neironu tīkli (2)

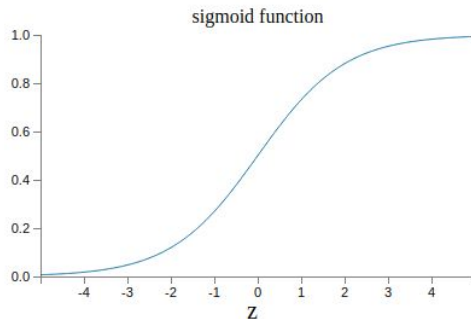
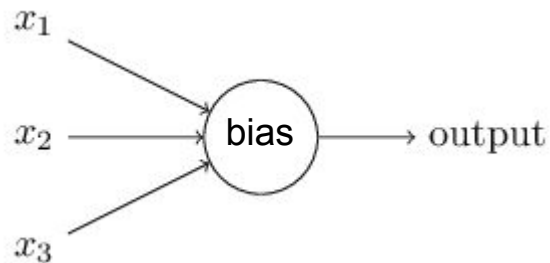
Tīkla apmācības shēma



Problēma: perceptrona izvade ir tikai 0 vai 1

Mākslīgie neironu tīkli (3)

Risinājums: sigmoīda neirons



Svarīgākā īpašība: gludums

Kāda tagad ir izvade?

~~$$\text{output} = \begin{cases} 0 & \text{if } w \cdot x + b \leq 0 \\ 1 & \text{if } w \cdot x + b > 0 \end{cases}$$~~

$$\Delta \text{output} \approx \sum_j \frac{\partial \text{output}}{\partial w_j} \Delta w_j + \frac{\partial \text{output}}{\partial b} \Delta b.$$

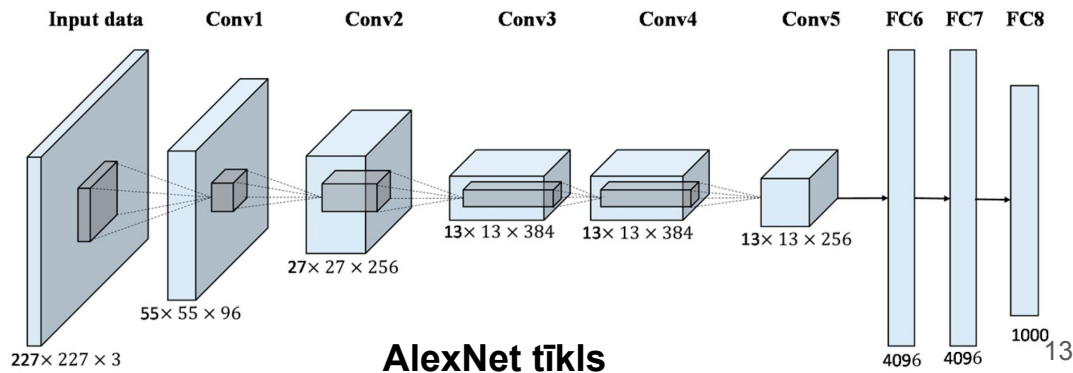
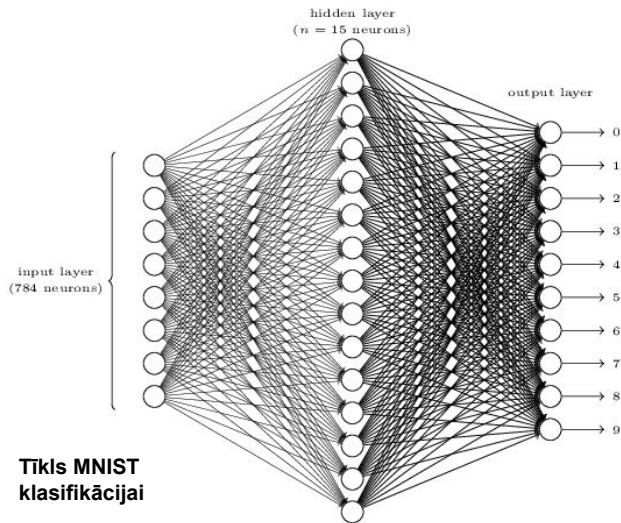
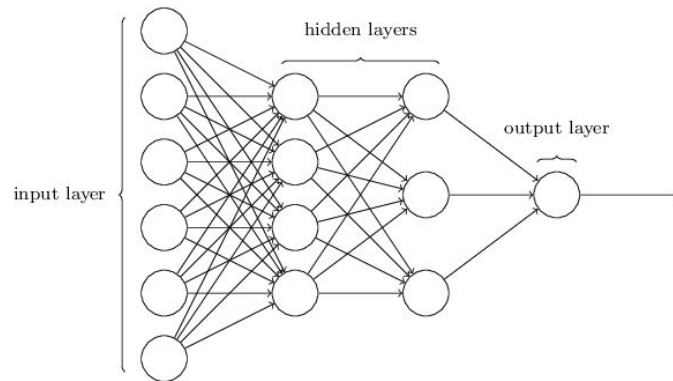
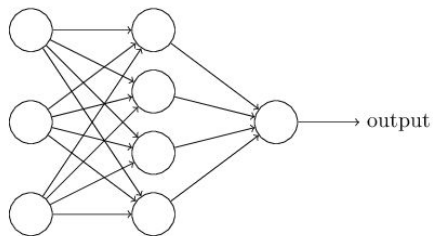
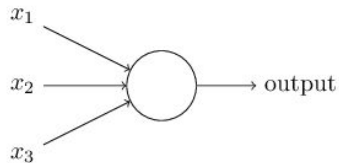
$$\sigma(w \cdot x + b),$$

σ is called the *sigmoid function*

$$\sigma(z) \equiv \frac{1}{1 + e^{-z}} \equiv \frac{1}{1 + \exp(-\sum_j w_j x_j - b)}.$$

Dziļā mašīnmācīšanās (1)

Tīkla arhitektūra

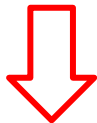


Dziļā mašīnmācīšanās (2)

Galvenā priekšrocība: labi rezultāti

Prasības:

- Liela skaitļošanas jauda (īpaši noder GPU), it īpaši tīkla apmācības posmā
- **Lielas apmācības datu kopas**



- Lielo datu pārvaldnieku (Google, Facebook utt.) dominēšana
- Apmācības datu pieejamība kļūst par vienu no svarīgākajiem faktoriem tirgū (sal. Kai-Fu Lee (2018) *AI Superpowers: China, Silicon Valley, and the New World Order*)

Risinājumi!?

Sintētiskie dati (1)

Sintētiskie dati: piemēram, ar 3D datorspēļu dzinējiem ģenerēti attēli

Kāpēc ir vajadzīgi sintētiskie dati? Jo dabiskie dati:

- Ne vienmēr ir pieejami
- Ir dārgi (ievākšana un marķēšana)
- Mēdz būt neprecīzi nomarkēti



Piemērs no GTA V datorspēles



Piemēri no MS COCO datu kopas

Sintētiskie dati (2)

Trīs pamatveidi, kā izmantot sintētiskos datus (Nikolenko, 2019):

- **Trenēt modeļus:**
 - Lai vēlāk tos izmantotu uz reāliem datiem
 - Lai vēlāk tos izmantotu, uzlabojot sintētiskos datus
- **Papildināt ar sintētiskiem datiem dabiskas datu kopas:**
 - Ģenerējot sintētiskos datus no nulles
 - Ģenerējot sintētiskos datus no esošiem dabiskiem datiem
- **Izmantot sintētiskos datus privātuma vai juridisku apsvērumu dēļ:**
 - Veselības aprūpē
 - Finansēs un soc. zinātnēs
 - Sintētiskie dati ar privātuma garantiju

Sintētiskie dati un pašbraucošās mašīnas (1)

- Pielietojums:
 - SLAM (simultaneous localization and mapping)
 - Kustības plānošana
 - Kustības kontrole
 - u.c.
- Pašbraucošo mašīnu neironu tīkli bieži vien tiek apmācīti, izmantojot stimulēto mācīšanos (*reinforcement learning*)
- Apmācība reālajā pasaulē būtu ļoti dārga un iespējams arī bīstama, tāpēc tiek izmantoti sintētiskie dati
- Svarīgs nošķīrums: statiskās sintētiskās datu kopas <-> interaktīvās 3D vides



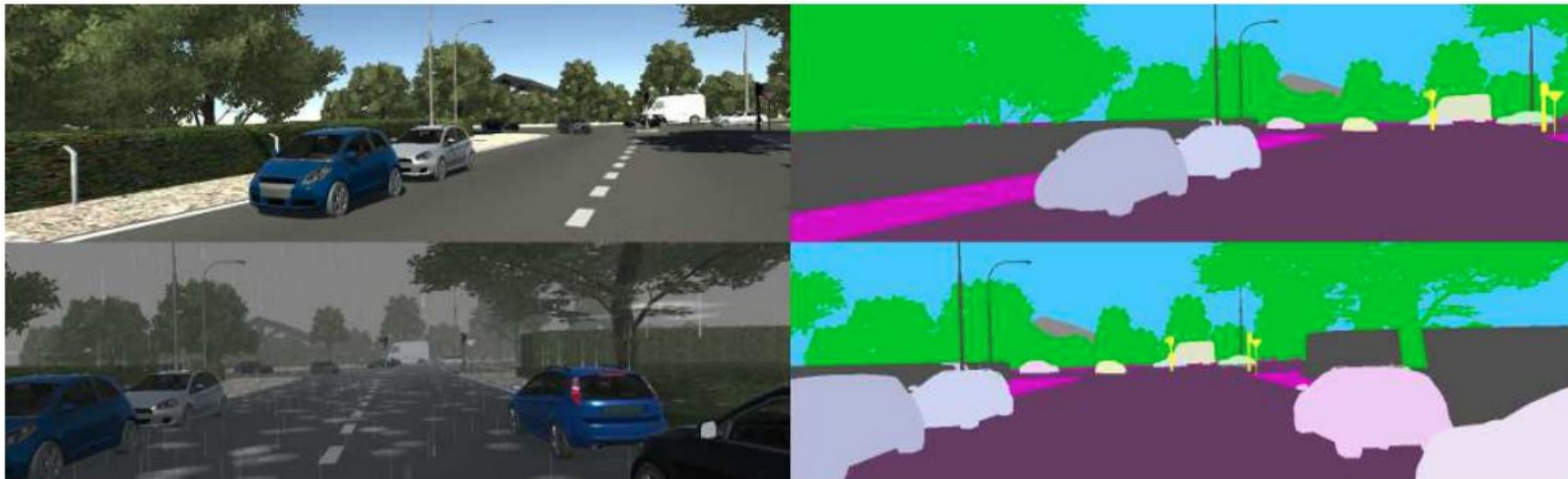
Sintētiskie dati un pašbraucošās mašīnas (2)

Virtual KITTI (*A Gaidon, Q Wang, Y Cabon, E Vig. 2016.*):

- 5 virtuālās vides un 50 sintētiskie video
- Mērķi: objektu atpazīšana, vairāku objektu izsekošana, skata un objekta līmeņa semantiskā segmentācija, optiskās plūsmas un dziļuma novērtēšana
- Eksperimentu rezultāti (objektu izsekošana):
 - Tikai neliela atšķirība starp reāliem un virtuāliem datiem
 - Būtiski: modeļi tika trenēti uz reāliem datiem un testēti uz sintētiskiem datiem, kas lielā mērā izskaidro labos rezultātus

Sintētiskie dati un pašbraucošās mašīnas (2)

Virtual KITTI (*A Gaidon, Q Wang, Y Cabon, and E Vig. 2016.*):



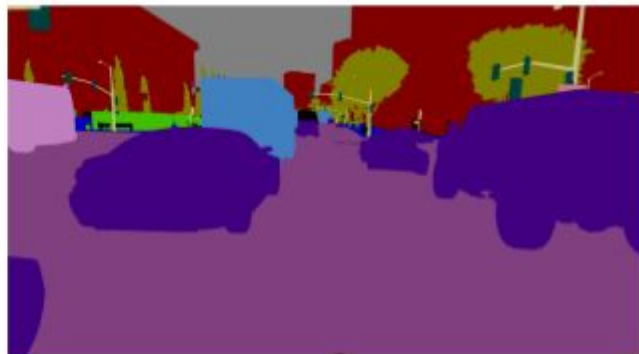
Sintētiskie dati un pašbraucošās mašīnas (3)

GTA V datu kopas (M. Johnson-Roberson, C. Barto, R. Mehta, S. N. Sridhar, K. Rosaen, and Vasudevan, 2017):

- Ceļu skati no Grand Theft Auto V videospēles
- Trīs datu kopas: 10 000, 50 000, 200 000 attēlu
- Papildus informācija: dziļuma kartes un robežlodziņi (*bounding boxes*)
- Rezultāti (objektu atpazīšana ar standarta arhitektūras Faster R-CNN tīklu):
 - Tīkls pēc treniņa uz 50 000 un 200 000 attēlu datu kopām demonstrēja augstāku precizitāti, nekā tīkls, kas tika trenēts uz dabiskas datu kopas (Cityscapes)
 - NB! Cityscapes ir 2975 attēlu

Sintētiskie dati un pašbraucošās mašīnas (3)

GTA V datu kopas (M. Johnson-Roberson, C. Barto, R. Mehta, S. N. Sridhar, K. Rosaen, and Vasudevan, 2017):



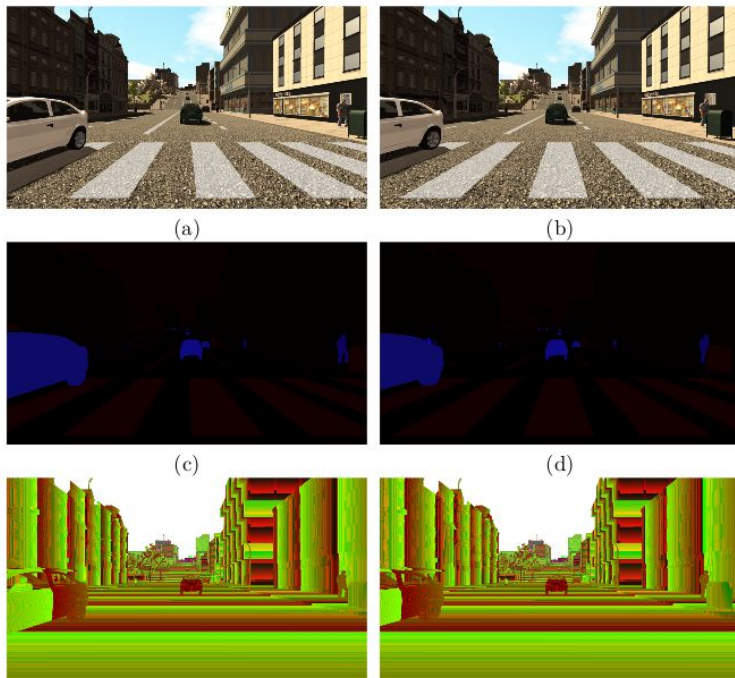
Sintētiskie dati un pašbraucošās mašīnas (4)

SYNTHIA (*G. Ros, L. Sellart, J. Materzynska, D. Vazquez, and A. M. Lopez., 2016*):

- Ņujorkas pilsētvides atveidošana ar *Unity* platformas rīkiem
- 213 000 attēlu, 13 semantiskās segmentācijas klases (gājēji, riteņbraucēji, ceļi, ēkas utt.)
- Eksperimentālie rezultāti (semantiskā segmentācija): dabiskās datu kopas papildināšana ar attēliem no SYNTHIA -> labāka precizitāte
- Vēlāk (2017) - arī SYNTHIA-SF (Sanfrancisko)

Sintētiskie dati un pašbraucošās mašīnas (4)

SYNTHIA (*G. Ros, L. Sellart, J. Materzynska, D. Vazquez, and A. M. Lopez., 2016*):

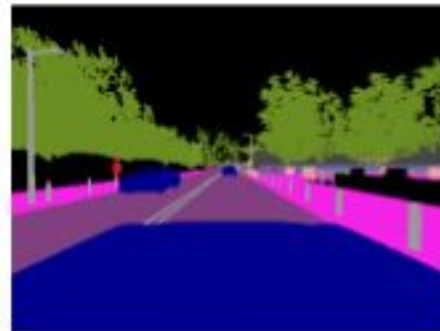


Sintētiskie dati un pašbraucošās mašīnas (5)

- **Simulācijas vide CARLA** (*CAR Learning to Act*; Dosovitskiy, German Ros, Felipe Codevilla, Antonio Lopez, Vladlen Koltun, 2017):
 - Implementācija: atvērto datu (*open-source*) paplašinājums *Unreal Engine 4*
 - Sensori:
 - RGB kameras ar maināmo izvietojumu
 - Dziļuma kartes
 - Semantiskās segmentācijas kartes (12 klases)
 - Robežlodziņi
 - Informācija par transportlīdzekli:
 - Novietojums
 - Orientācija

Sintētiskie dati un pašbraucošās mašīnas (5)

- **Simulācijas vide CARLA** (*CAR Learning to Act*; Dosovitskiy, German Ros, Felipe Codevilla, Antonio Lopez, Vladlen Koltun, 2017):



Kopsavilkums

- Ir pieejami vairāki risinājumi: gan sintētiskās datu kopas, gan 3D interaktīvās vides
- Iespējamie pētniecības virzieni:
 - Dažādu sintētisko kopu salīdzināšana
 - Atšķirību mazināšana starp sintētiskiem datiem un dabiskiem datiem
 - Dabisko un sintētisko datu sajaukšana:
 - Datu kopas ietvaros
 - Viena attēla ietvaros

Paldies par uzmanību!