

MAŠĪNMĀCĪŠANĀS CILVĒKA ATPAZĪŠANAI MACHINE LEARNING FOR HUMAN RECOGNITION

Autori: **Deņiss Surikovs, Andrejs Mežārs, Ritvars Bleive**
e-pasti: ds19074@edu.rta.lv, am18118@edu.rta.lv, rb18051@edu.rta.lv
Zinātniskā darba vadītājs: **Sergejs Kodors Dr.sc.ing., asoc. prof.**
Rēzeknes Tehnoloģiju akadēmija,
Atbrīvošanas aleja 115, Rēzekne, LV-4601

Abstract. *The aim of this work is to develop a neural network that will be able to recognize human presence. To achieve this goal, authors applied neural network architecture YOLOv5 and the open dataset COCO. The experiment was repeated three times. The study yielded a good result - the neural network was able to detect humans in images with a good precision.*

Keywords: *artificial neural network, COCO dataset, human recognition, YOLOv5.*

Ievads

Viedās monitoringa sistēmas ir piedzīvojušas strauju izaugsmi pēdējā laikā. Arvien lielāka nozīme kļūst pētījumiem par cilvēku klātbūtnes noteikšanu. Cilvēka atpazīšana un meklēšana ir būtiska, lai savlaicīgi identificētu personu, atpazītu viņas darbības un izpildītu ainas analīzi. Šim uzdevumam varbūt pielietoti konvolūcijas neironu tīkli (*convolutional neural network*, tālāk *CNN*). *CNN* ir plaši izmantoti, lai atpazītu objektus un realizētu saistītus risinājumus. [1]

Mākslīgie neironu tīkli tika veidoti, imitējot bioloģisko neironu tīklu darbību. Mākslīgais neironu tīkls ir sistēma, kas spēj apstrādāt sarežģītus ievades datus, pielietojot mašīnmācīšanās algoritmus. Mašīnmācīšanās algoritmi «iemācā» neironu tīklus atkārtot uzdevumus, pielietojot apmācības piemērus, kas atšķirībā no simboliskā mākslīgā intelekta (*symbolic AI*) nepieprasā programmēt sarežģītu un specifisku loģiku balstīto uz eksperta pieredzes. Savukārt, *CNN* ir speciāls mākslīgo neironu tīkla tips, kas ir paredzēts analizēt attēlus. [2]

YOLO ir objektu meklēšanas risinājums, kas sadala attēlus režģa sistēmā. Katra režģa šūna ir atbildīga par objektu noteikšanu sevī. *YOLO* ir viens no populārākajiem objektu meklēšanas algoritmiem, pateicoties tā ātrumam un precizitātei. *YOLOv5* ir risinājuma jaunāka versija, kas darbojas izmantojot *PyTorch* tehnoloģiju. [3]

Pētījuma mērķis ir pārbaudīt **hipotēzi**: “Vai, apmācot *YOLOv5* neironu tīklu ar 2700 attēlu lielu datu kopu, tas spēs atpazīt cilvēka klātbūtni?”. Lai iegūtu objektīvāku statistiku un uzlabotu pētījuma rezultātus, eksperiments tika atkārtots 3 reizes.

Uzdevumi: 1) sagatavot apmācības datukopu; 2) apmācīt *YOLOv5* neironu tīklus 3 reizes; 3) pārbaudīt neironu tīklu darbu.

Materiāli un metodes

Lai sasniegtu mērķi un pārbaudītu hipotēzi, neironu tīkla apmācībai tika izmantotas dažādas tehnoloģijas.

Datu kopas sagatavošana: pētījuma realizēšanai tika izvēlēta un sagatavota *COCO* datu kopa, kura, savukārt, satur 2700 attēlus. [4] Ar FiftyOne rīka palīdzību no datu kopas tika atlasīti tikai tie attēli, kuri satur cilvēkus (*skat. piemērus 1. un 2. attēlos*). [5] Datu kopa ir sadalīta trīs apakškopās: trenēšanas (70%), pārbaudes (20%) un testēšanas (10%) kopas. Šie dati tiks pielietoti *YOLOv5* neironu tīkla apmācībai, kas analizēs un apstrādās šo informāciju.



1. un 2. attēls. Attēlu piemēri no COCO datu kopas [4]

Formātu konvertēšana: lai lietotu *YOLOv5* modeli, bija nepieciešams pārveidot *COCO JSON* formātu uz *YOLOv5 PyTorch TXT* formātu. Šis uzdevums tika panākts ar *RoboFlow* rīka palīdzību. [6] Dati tika augšupielādēti *RoboFlow* vidē, kas, savukārt, atļauj ģenerēt jaunu datu kopu, pielietojot augmentācijas risinājumus, un eksportēt to *YOLOv5 PyTorch TXT* formātā.

Neironu tīkla apmācība: iegūto datu kopu bija nepieciešams importēt *Google Colaboratory (Colab)*, kur, lai trenētu *YOLOv5* modeli; tika izmantots *PyTorch - mašīnāpmācības* tehnoloģija, kas darbojas *Python* vidē (skat. 3. attēlu). Šajā posmā tika uzsākta *YOLOv5* apmācība ar konkrēto datu kopu un, respektīvi, iegūti neironu tīkla apmācības rezultāti.

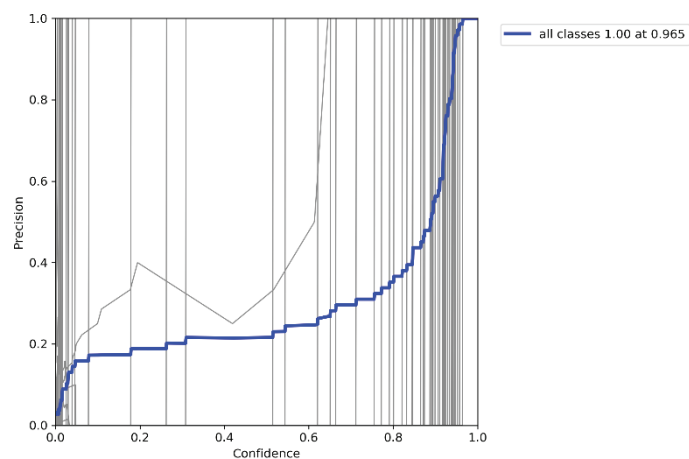
```
!git clone https://github.com/ultralytics/yolov5
%cd yolov5
%pip install -qr requirements.txt
import torch
from yolov5 import utils
display = utils.notebook_init()

torch.hub.download_url_to_file('...', 'tmp.zip')
!unzip -q tmp.zip -d ../datasets && rm tmp.zip

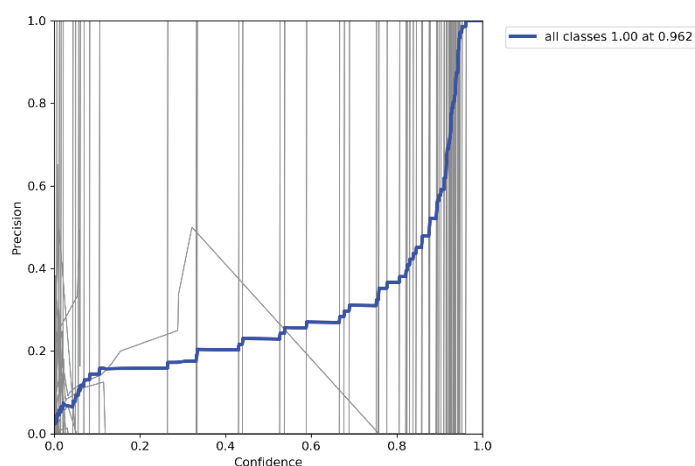
!python train.py --img 640 --batch -1 --epochs 100 --content/datasets/data.yaml --weights yolov5s.pt --cache
```

3. attēls. *Google Colab* skripts ar *YOLOv5* inicializāciju un apmācību

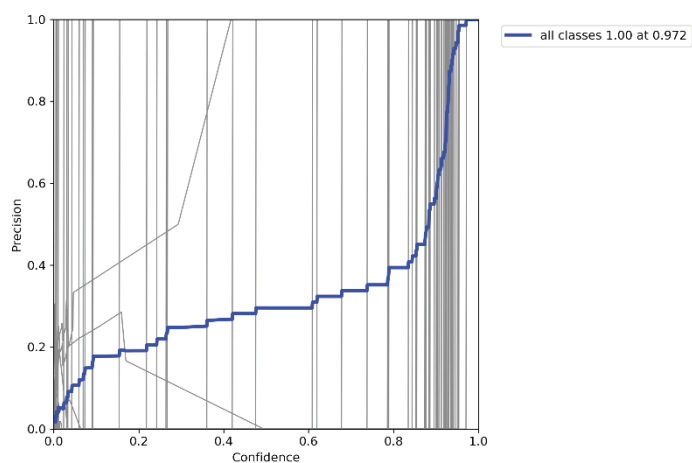
Attēlu analīze: apmācību rezultāti tika automātiski reģistrēti *TensorBoard*. Ar *TensorBoard* palīdzību tika nodrošināta rezultātu vizualizācija un kvalitātes kontrole, kas bija nepieciešami mašīnāpmācības eksperimentam (skat. 4., 5. un 6. attēlu).



4. attēls. Precizitātes līkne pēc pirmā eksperimenta



5. attēls. Precizitātes līkne pēc otrā eksperimenta



6. attēls. Precizitātes līkne pēc trešā eksperimenta

4.-6. attēlos ir redzams, ka apmācīts neironitīkls viennozīmīgi atpazīst objektus ar uzticību (*confidence*) 96% un 97%. Savukārt, grafiks attēlo eksponenciālo kritumu, kas ir saistīts ar nelīdzsvarotu datu kopu (*disbalanced dataset*). Tāpēc, lai vienkāršotu kvalitātes pārbaudi, tika nolemts vizuāli pārskatīt neironitīkla darbu.

Rezultāti un to izvērtējums

Analizējot eksperimentos iegūtos rezultātos, var novērot, ka visos eksperimenta mēģinājumos YOLOv5 algoritms ar labu precizitāti ir spējis atpazīt cilvēku esamību attēlos (*skat. 7., 8. un 9. attēlu*).



7. attēls. YOLOv5 trenēšanas rezultāti pēc pirmā eksperimenta



8. attēls. YOLOv5 trenēšanas rezultāti pēc otrā eksperimenta



9. attēls. YOLOv5 trenēšanas rezultāti pēc trešā eksperimenta

Var novērot, ka starp trīs eksperimenta mēģinājumiem būtisko atšķirību nav. Lai uzlabotu neironīklu novērtēšanas kvalitāti, būtu vērts pielietot kvantitatīvo metodi, apskatot tikai klasi ar cilvēkiem.

Secinājumi

- Hipotēze apstiprinājās, jo apmācot *YOLOv5* neironitīklu ar 2700 attēlu lielu datu kopu, tas spēj atpazīt cilvēka klātbūtni.
- Rezultātu uzlabošanai *YOLOv5* neironitīklam ir nepieciešama lielāka datu kopa, kā arī lielāka skaitļošanas veiktspēja, uz kuras notika eksperimenti.
- Objektīvāku rezultātu iegūšanai ekperments varētu būt papildināts ar citiem statistikas datiem, piemēram mainot augmentācijas procesu *RoboFlow* vidē, vai arī izmantojot *YOLOv5* analogu.
- Starp trīs eksperimentu rezultātiem nav novērojamas būtiskas atšķirības.

Acknowledgment

This research is funded by the Latvian Council of Science, project “Development of autonomous unmanned aerial vehicles based decision-making system for smart fruit growing”, project No. lzp-2021/1-0134.

Summary

The authors selected and prepared a COCO dataset, which includes 2700 images of humans. The YOLOv5 object detection algorithm was used for machine learning, as well as Roboflow and Google Colaboratory. To provide better statistics, the experiment was repeated a total of three times. The study yielded a good result - the neural network was able to detect humans in images with ~96% accuracy.

The content of the work emphasizes the possibility for the neural network to identify a person using the YOLOv5 model, summarizes the results, analyzes them, and creates the following conclusions:

- The hypothesis was confirmed because by training the YOLOv5 neural network with a data set of 2,700 images, it can detect the human presence.
- To improve performance, the YOLOv5 neural network requires a larger data set as well as the computing performance on which the experiments were performed.
- For more objective results, the experiment could be supplemented with other statistics, such as changing the augmentation process in RoboFlow, or using a YOLOv5 analog.
- There are no significant differences between the results of the three experiments.

Literatūra

1. Human detection and tracking with deep convolutional neural networks [tiešsaiste], [atsauce uz 01.05.2022]. Pieejams: <https://link.springer.com/article/10.1007/s11042-020-09579-x>
2. Konvolūcijas neironu tīkls [tiešsaiste], [atsauce uz 01.05.2022]. Pieejams: <https://www.techtarget.com/searchenterpriseai/definition/convolutional-neural-network>
3. YOLOv5: The Latest Model for Object Detection [tiešsaiste], [atsauce uz 01.05.2022]. Pieejams: <https://medium.com/axinc-ai/yolov5-the-latest-model-for-object-detection-b13320ec516b>
4. Comon Objects In Context [tiešsaiste], [atsauce uz 09.05.2022]. Pieejams: <https://cocodataset.org/#download>
5. FiftyOne [tiešsaiste], [atsauce uz 10.05.2022]. Pieejams: <https://voxel51.com/fiftyone/index.html>
6. Computer Vision Datasets [tiešsaiste], [atsauce uz 09.05.2022]. Pieejams: <https://public.roboflow.com/>