Attēlu atpazīšanas un klasifikācijas AI modeļa praktisks piemērs

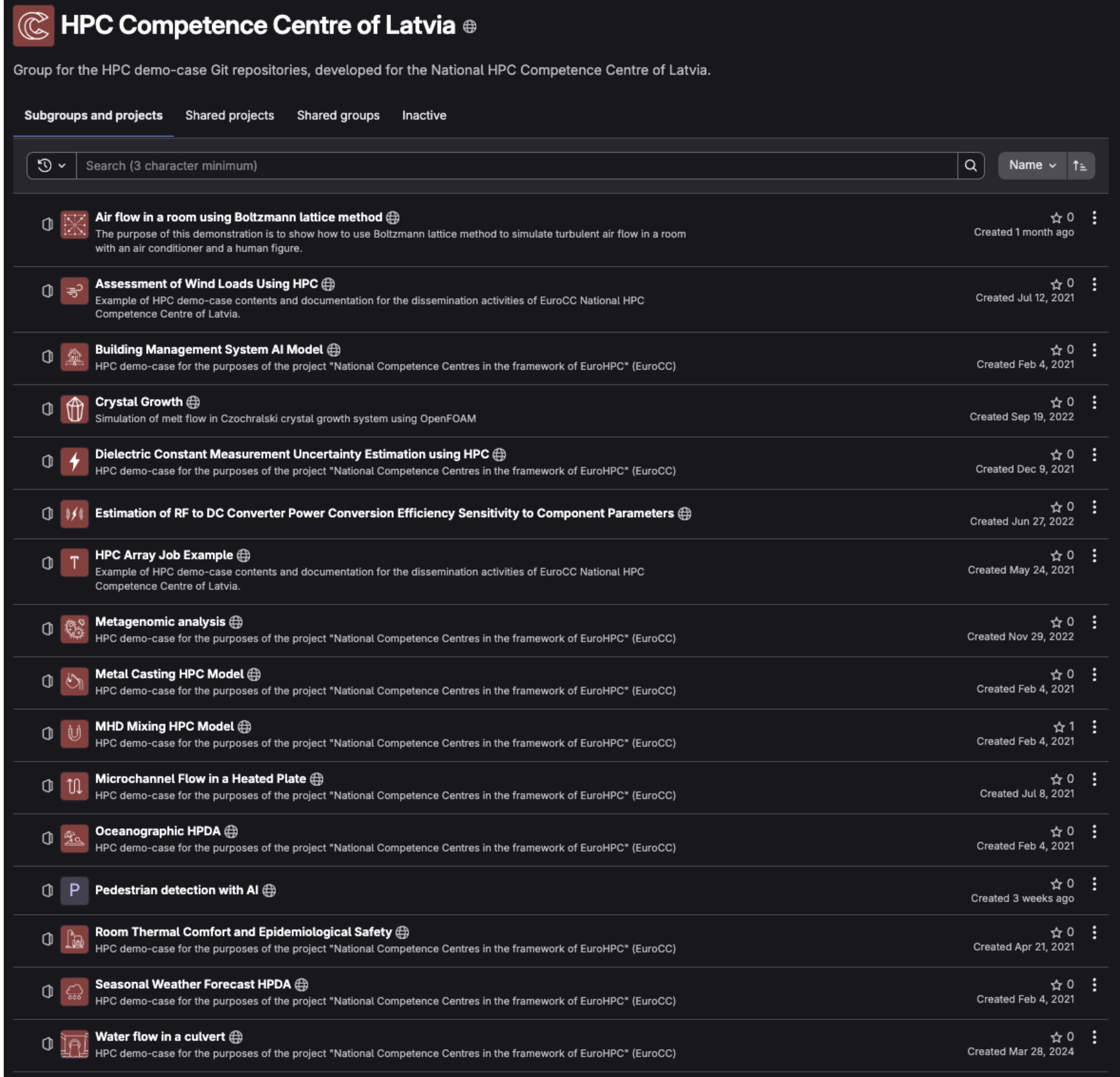
Izstrādāts EuroCC Nacionālā HPC kompetences centra ietvaros

Rūdolfs Livanovičs

28.11.2025

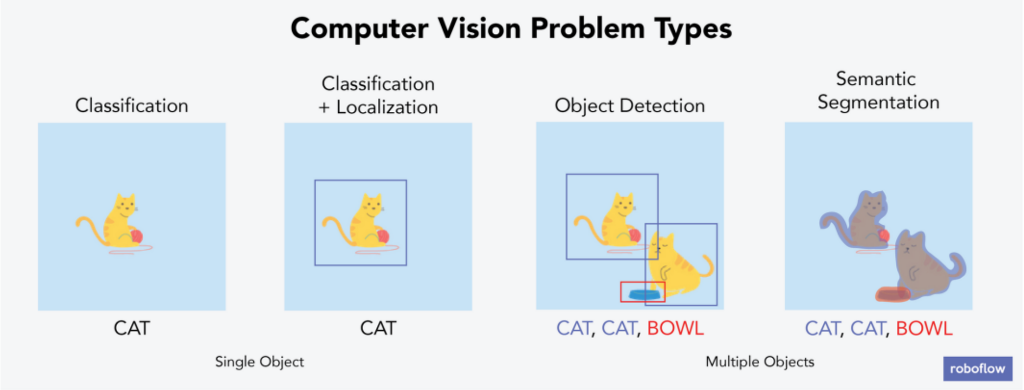
Motivācija

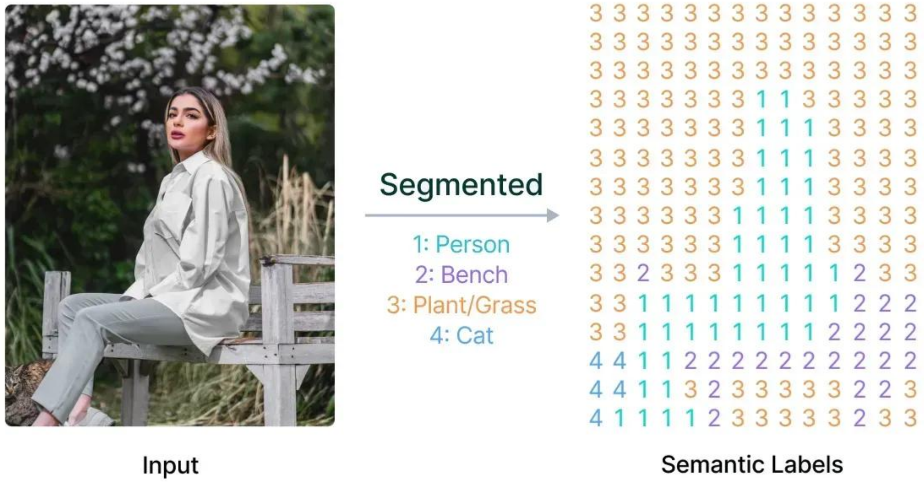
- Esošie kompetences centra demo-case piemēri jau tagad labi nosedz dažādus tradicionālos HPC aspektus
- Ar vienu izņēmumu, tie vispār nesatur mašīnmācīšanās elementus
- Praktiskos HPC popularizēšanas pasākumos nospiedošais vairākums interesentu jautā tieši par pielietojumiem mašīnmācīšanās / AI kontekstā

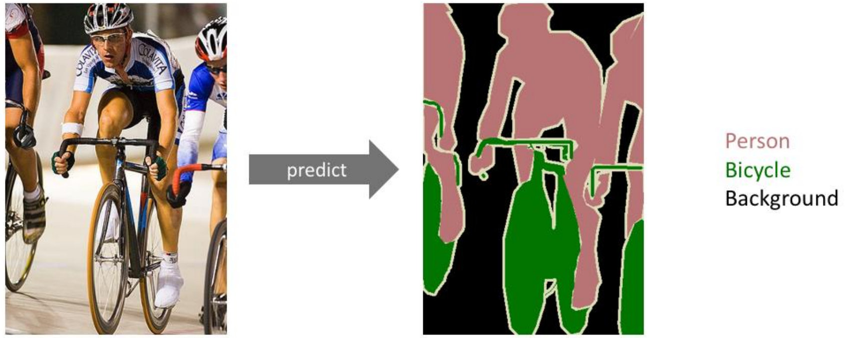


Problēmas uzstādījums

- Industriālos pielietojumos HPC-relevantā mērogā dominē divu veidu modeļi – LLM un datorredzes modeļi
- Datorredzes modeļi ir saturiski piemērotāki maza un vidēja līmeņa HPC infrastruktūrai un tiek biežāk būtiski pielāgoti vai apmācīti no jauna
- Datorredzes modeļiem ir daudzveidīgs pielietojumu spektrs ar praktisku vērtību un plašu atpazīstamību gan industrijā, gan sabiedrībā kopumā

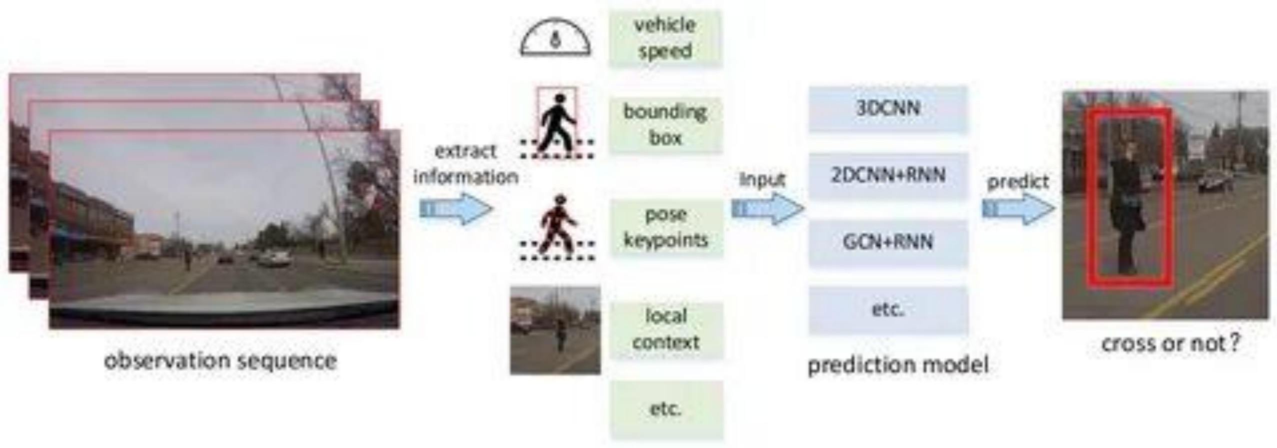






Problēmas uzstādījums

- Klasiska datorredzes problēma, kas turpina saglabāt aktualitāti: gājēju detekcija, segmentācija un uzvedības klasifikācija
- Relevanta ne tikai pašbraucošu automašīnu izveidē, bet arī ADAS (autovadītāju atbalsta) sistēmās, novērošanas kameru atbalsta sistēmās, u.c.
- Plašākā nozīmē, objektu vizuālas segmentācijas un dažādu īpašību klasifikācijas problemātika ir universāli sastopama industriālos procesos

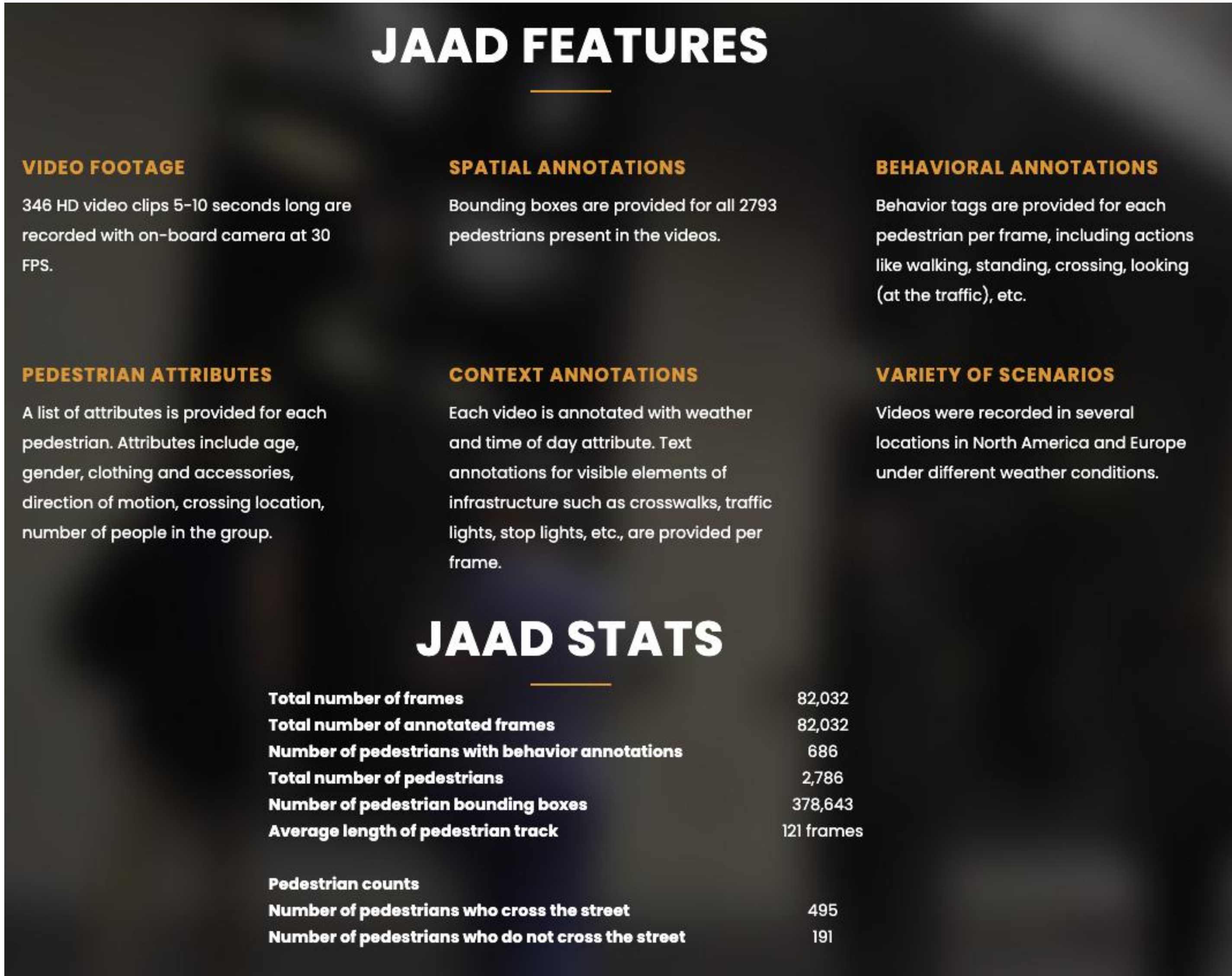


Izejas datu kopa

- Joint Attention in Autonomous Driving (JAAD) ir brīvi pieejama datu kopa, kas satur kvalitatīvi homogēnus un ekstensīvi anotētus autovadītāja perspektīvas video no dažādām Ziemeļamerikas un Eiropas pilsētām

- Šī datukopa nav pietiekami liela vai daudzveidīga state-of-the-art ADAS sistēmas izveidei, taču ir pietiekoša neliela šāda tipa sistēmas prototipa izveidei pedagoģiskiem mērķiem

- Modeļa programmatūra veidota tā, lai jaunu datukopu integrācija būtu relatīvi viegla



JAAD FEATURES

VIDEO FOOTAGE

346 HD video clips 5-10 seconds long are recorded with on-board camera at 30 FPS.

SPATIAL ANNOTATIONS

Bounding boxes are provided for all 2793 pedestrians present in the videos.

BEHAVIORAL ANNOTATIONS

Behavior tags are provided for each pedestrian per frame, including actions like walking, standing, crossing, looking (at the traffic), etc.

PEDESTRIAN ATTRIBUTES

A list of attributes is provided for each pedestrian. Attributes include age, gender, clothing and accessories, direction of motion, crossing location, number of people in the group.

CONTEXT ANNOTATIONS

Each video is annotated with weather and time of day attribute. Text annotations for visible elements of infrastructure such as crosswalks, traffic lights, stop lights, etc., are provided per frame.

VARIETY OF SCENARIOS

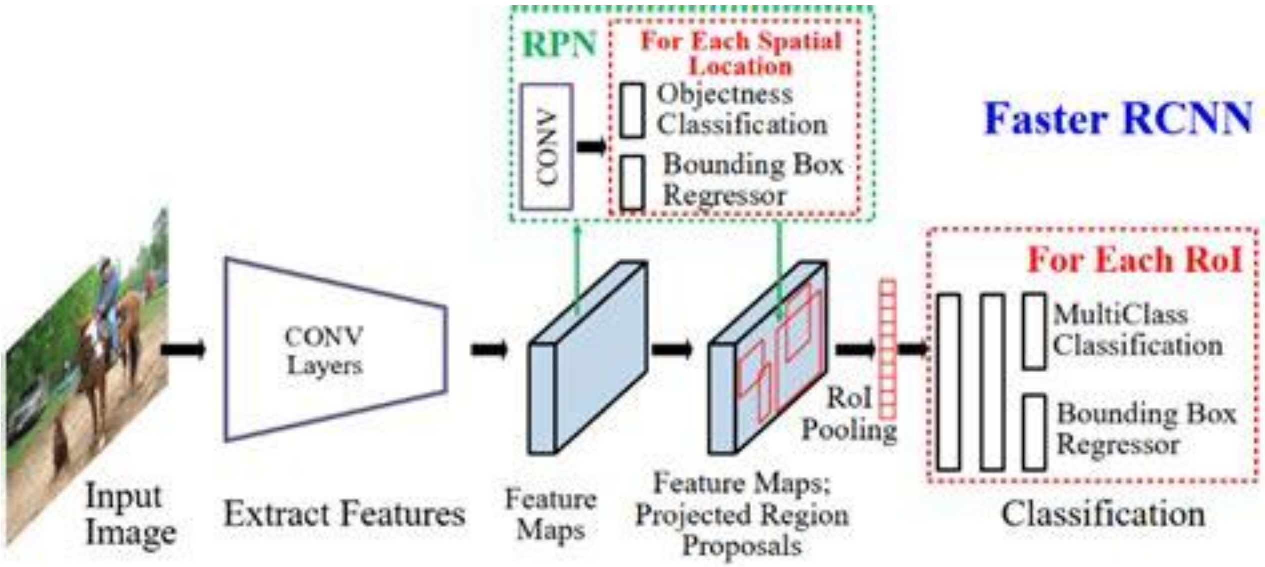
Videos were recorded in several locations in North America and Europe under different weather conditions.

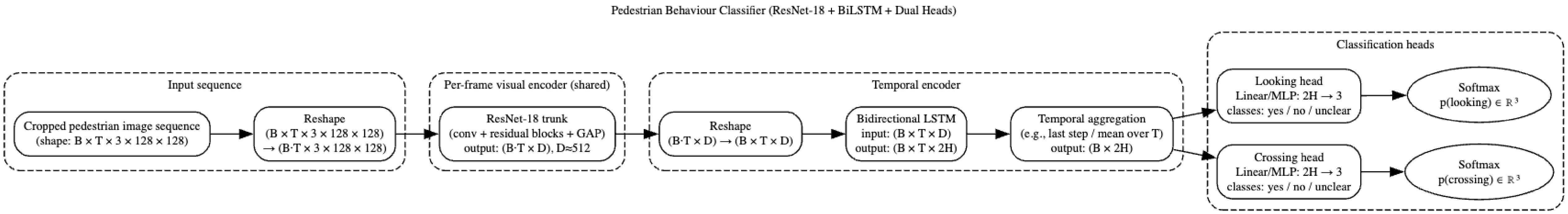
JAAD STATS

Total number of frames	82,032
Total number of annotated frames	82,032
Number of pedestrians with behavior annotations	686
Total number of pedestrians	2,786
Number of pedestrian bounding boxes	378,643
Average length of pedestrian track	121 frames
Pedestrian counts	
Number of pedestrians who cross the street	495
Number of pedestrians who do not cross the street	191

Modeļa saturs un implementācija

- **Detektors**: Faster R-CNN (ResNet-50-FPN), apmācīts vispārīgu objektu detekcijai (COCO datukopa)
- **Klasifikators**: ResNet-18 bāzes attēlu enkoderis, BiLSTM klasifikators ar divām vērtību klasēm





Modeļa saturs un implementācija

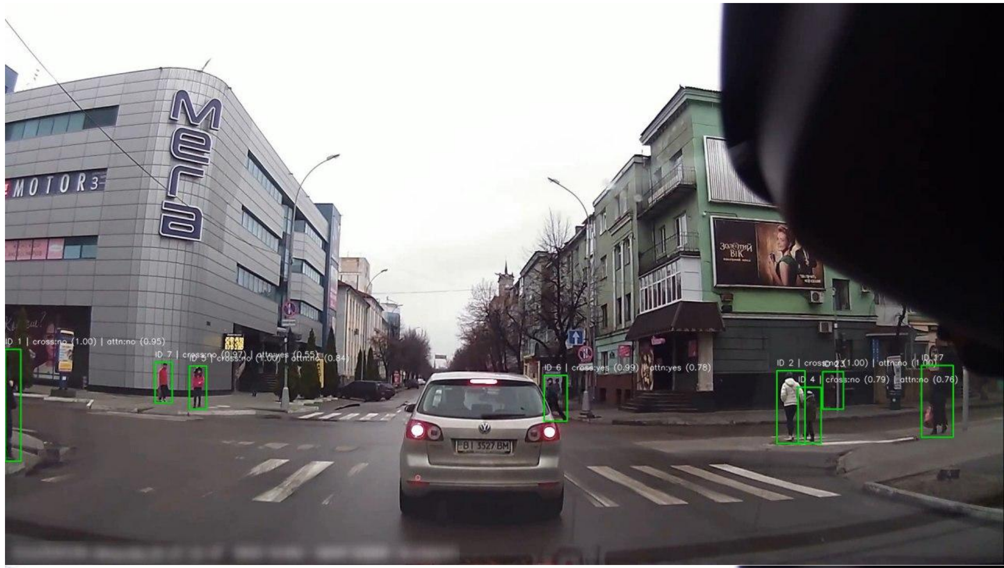
- Papildus tiešajai modeļa implementācijai, papildus infrastruktūra – automātiska datukopas sagatavošana, video apstrāde, anotāciju preprocesēšana, apmācības kontrole un validācija, inference, utt.

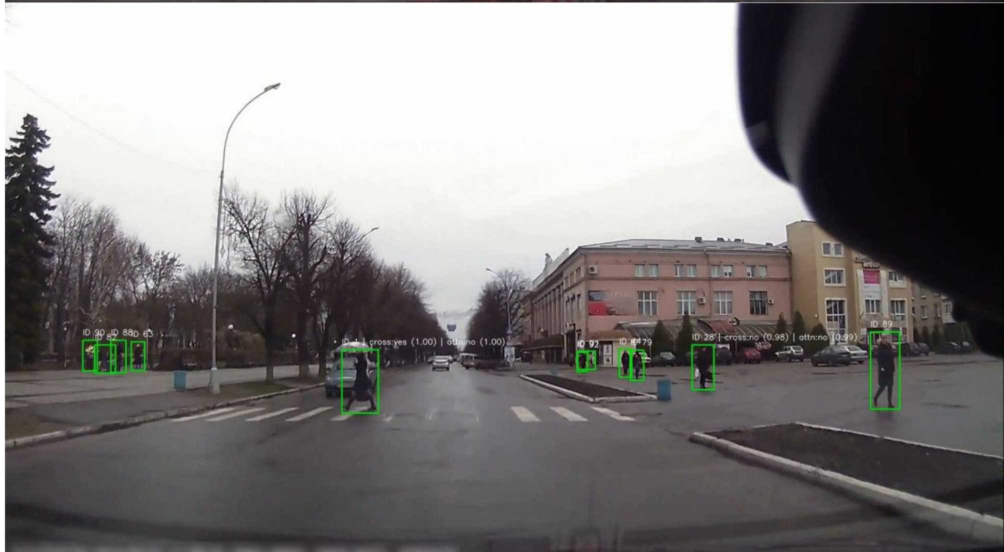
- Kopumā ~4k rindiņas Python koda, saturiski reprezentatīvs tipiskam šāda veida ML projektam industrijā – tiek izmantots uv, PyTorch, torchvision, OpenCV, pytest, utt.

- No lietotāja viedokļa unificēts interfeiss – viena komanda instalācijai, viena apmācībai, viena inferencei

- Modelis ir testēšanas fāzē un tiks vēl nedaudz papildināts, taču ir publiski pieejams:

  https://gitlab.com/eurocc-latvia/pedestrian-detection-with-ai





Demo-case izmantošana

- Modelis sagaida Linux + CUDA vidi ar vismaz vienu GPU un parametru noklusējuma vērtības pieņem LU SMI HPC klāstera GPU vidi (ar 2x RTX A6000 GPU), taču principā visu procesu iespējams testēt arī uz attiecīga klēpjdatora

- LU SMI HPC klāstera vidē jāiegūst lietošanas tiesības, jāpieslēdzas caur SSH un jāizmanto Slurm (pamācība: https://hpc.lu.lv/en/guides/slurm/).

- Modeļa lietošana ir detalizēti dokumentēta un repozitorijā ir iekļauts Slurm job faila piemērs, taču kopumā šis piemērs pieņem standarta Linux komandlīnijas lietotāja prasmes, priekšzināšanas mašīnmācīšanās jomā un Python programmēšanas prasmes.


ID 3
ID 4
ID 1
ID 2

Paldies par uzmanību.