

Detektált úthibák távolságának meghatározása sztereó kamera és LiDAR segítségével

Measure the distance of detected road deviations with stereo camera and LiDAR

SZABÓ Máté András¹, DR. URBIN Ágnes², TIZEDES László³

¹HUN-REN Számítástechnikai és Automatizálási Kutatóintézet,
Kende u. 13-17. 1111 Budapest, szabomandras@sztaki.hun-ren.hu
²Budapesti Műszaki és Gazdaságtudományi Egyetem MOGI Tanszék,
Műegyetem rkp. 9-D épület, 1111 Budapest, urbin@mogi.bme.hu
³HUN-REN Számítástechnikai és Automatizálási Kutatóintézet,
Kende u. 13-17. 1111 Budapest, tizedes.laszlo@sztaki.hun-ren.hu

Abstract

One of the most important tasks of autonomous vehicles is to sense road deviations and different obstacles. To detect these objects, I developed a convolutional neural network with supervised learning. Besides the detection, I attempted to determine their position with various sensors. I compared the accuracy of stereo vision and laser ranging and provided a valuable conclusion for designing autonomous vehicles.

Keywords: autonomous vehicle, neural network, stereo camera, LiDAR, object detection

Kivonat

Az autonóm járművek egyik legfontosabb feladata az útfelület hibáinak és a különböző akadályoknak az érzékelése. A kátyúknak, fekvőrendőröknek és egyéb objektumoknak a detektálására felügyelt tanulással konvolúciós neurális hálózatot fejlesztettem. Az érzékelésen kívül megpróbáltam ezek pozícióját is meghatározni, amire több különböző szenzort teszteltem. Összehasonlítottam a sztereó látás és a lézeres távmérés pontosságát és hasznos konklúziót szolgáltatottam autonóm járművek tervezéséhez.

Kulcsszavak: autonóm jármű, neurális hálózat, sztereó kamera, LiDAR, objektum detekció

1. BEVEZETÉS

Az elmúlt években a járműipar egyre nagyobb hangsúlyt fektetett az önvezető autók kifejlesztésére. Az egyik legfontosabb feladata ezeknek az autonóm eszközöknek az útfelület hibáinak és a különböző akadályoknak az érzékelése, amiknek az ismerete elengedhetetlen a megfelelő irányítás kialakításához. A Számítástechnikai és Automatizálási Kutatóintézetben (HUN-REN SZTAKI) fejlesztett program vizuális alapon, gépi látás segítségével képes valós időben detektálni az előre meghatározott objektumokat. A kátyúk, fekvőrendőrök és egyéb akadályok észleléséhez felügyelt tanulással neurális hálózatot hoztam létre.

Az objektumok érzékelése azonban nem elegendő, hiszen az önvezető autónak ismernie kell azok távolságát is a járműtől. Erre a feladatra a munkám során két különböző megoldást, a sztereó látást és a lézeres távmérést teszteltem. A kiértékeléshez a méréseket szinkronizálni kellett térben és időben is. A tesztek különböző eszközökkel végeztem el, majd az eredményeket számos szempont alapján értékeltem ki. Meghatároztam a sztereó látás és a lézeralapú távolságmérési módszerek egymáshoz viszonyított relatív hibáját, valamint feltérképeztem a hibák nagyságát az egyes távolságtartományokban. Kielemeztem a sztereó kamera karakterisztikáját és javaslatot tettem annak korrigálására is a pontosabb mérések elérésének érdekében. Összehasonlítottam két LiDAR szkennelési módszerét is és megvizsgáltam, hogy azok eltérései milyen hatással vannak a távolságmérés pontosságára. Végül konklúzióval szolgáltam azt illetően, hogy melyik eljárás és milyen szempontok alapján alkalmasabb önvezető autók által detektált úthibák távolságmérésére.

2. MÉRŐESZKÖZÖK

2.1. Sztereo kamera

A háromdimenziós tér feltérképezésére használt egyik leggyakoribb módszer a sztereoszkópia [1]. Ez az emberi látást imitálja, ahol a két szem helyett két, egy irányba néző kamera érzékeli a környezetet. A két egyidőben készített kép és a kamerák egymáshoz viszonyított helyzetének ismeretében háromszögeléssel kiszámítható az egyes képpontok távolsága. A sztereo kamerák nagy előnye, hogy a berendezés viszonylag olcsó, illetve a kamerák képet is készítenek, amik felhasználhatóak további alkalmazásokra, például vizuális alapú objektumdetekcióra. Azonban összehasonlítva a többi 3D érzékelő módszerrel megállapítható, hogy a távolságmérés kevésbé pontos ezeknél az eszközöknél. A sztereo kamera az egyik leggyakoribb szenzortípus önvezető autók esetén [2]. A kutatás során egy Stereolabs ZED 2 [3] sztereo kamerát használtam.

2.2. Lézer alapú távérzékelő (LiDAR)

A háromdimenziós tér feltérképezésének másik módszere az úgynevezett Time-of-Flight (ToF) metódus [4]. Ennek a lényege, hogy egy emitter lézerpulzust bocsát ki, ami az útjába kerülő tárgyról visszaverődve egy detektorba jut. A kibocsátás és az érzékelés között eltelt időből a tárgy távolsága kiszámítható. Ezt a módszert használják a lézer alapú távérzékelők (LiDAR) [5] is. Ezek az eszközök sok különböző irányba bocsátanak ki ilyen lézerpulzusokat és a távolságokat meghatározva egy pontfelhőt készítenek a szenzor környezetéről. Az ilyen eszközök egyik hátránya az árak, ami jóval magasabb, mint a korábban bemutatott sztereo kameráké. Cserébe azonban pontosabbak, illetve nagyobb terület is feltérképezhető velük. Ezek a szenzorok manapság elengedhetetlen kellékei az autonóm járműveknek [6].

A projektben két LiDAR-t, egy Ouster OS0-t [7] és egy Livox Avia-t [8] használtam. Az előbbi több lézermittor körbeforgatásával tapogatja le a környezetét, ezáltal egy 360°-os pontfelhőt előállítva. Utóbbi ennél kisebb látószöggel, viszont különleges, nyolcas alakokat leíró szkennelési metódussal rendelkezik [9]. Ezáltal a szenzor fókuszpontjában több, míg attól távolodva egyre kevesebb pont található a felhőben. A két LiDAR között jelentős árkülönbség is van, az Ouster szenzor nagyjából hatszor annyiba kerül, mint a Livox.

3. KONVOLÚCIÓS NEURÁLIS HÁLÓZAT FEJLESZTÉSE

3.1. Az adatbázis létrehozása

Az objektumok felismerésére konvolúciós neurális hálózatot kellett tanítanom [10]. Ez a hagyományos neurális hálózatok egy továbbfejlesztett verziója, amit elsősorban képfeldolgozásban és gépi látási feladatokban szoktak használni. A feladatra a YOLOv4 [11] modellt választottam, ami nagyobb FPS érték elérésére képes a többi neurális hálózattal szemben. Ez kifejezetten előnyös tulajdonság valósidejű alkalmazások számára. A detektálni kívánt objektumokat öt osztályba soroltam be. Az első kettő a fekvőrendőr, mint akadály az úton és a kátyú, mint az útfelület hibája. A maradék három a csatornalefolyó, a csatornafedő és a bicikliút jelzés. Ezekre azért volt szükség, mert az alakjuk és az útfelületen elfoglalt pozíciójuk miatt gyakran hibásan kátyúként voltak érzékelve. A használatukkal a hibás detekciók száma jelentősen csökkenthető.

Az definiált objektumokat felismerni képes neurális hálózatot felügyelt tanulással lehet létrehozni. Ennek a módszernek a lényege, hogy nagy számú annotált adatot adunk a hálózatnak, ami a tanítási folyamat végén képes lesz az új, annotálatlan adatokat is helyesen felcímkézni. A képekből álló adatbázist egyrészt nyílt hozzáférésű, interneten fellelhető forrásokból [12, 13], másrészt a saját felvételeink alapján építettem fel. Utóbbi során egy autót a megfelelő szenzorokkal felszerelve vittünk el adatokat gyűjteni. A felépített adatbázis 1609 képet tartalmaz, melyeken különböző számú, nagyságú és fajtájú objektumok szerepelnek. A reprezentativitást a kamerák, azok pozíciói, az útfelületek típusai és az időjárási viszonyok sokfélesége is biztosítja.

3.2. A neurális hálózat tanítása

A neurális hálózat tanítására a Darknet [14] nevű programgyűjteményt használtam. Ez egy olyan keretrendszer, ami képes a YOLOv4 neurális hálózat tanítására, futtatására és tesztelésére. Az adatbázist hagyományosan három részre, tanító, validációs és teszt adathalmazokra szokás felosztani. Az alkalmazásomban ezek aránya 80:10:10 százalék. A neurális hálózatot tízezer iteráción keresztül tanítottam. Az általam létrehozott modell Mean Average Precision (mAP) [15] értéke 68,7%, ami megfelelőnek mondható. Az 1. ábrán látható, hogy a modell sikeresen felismeri az előre definiált objektumokat.

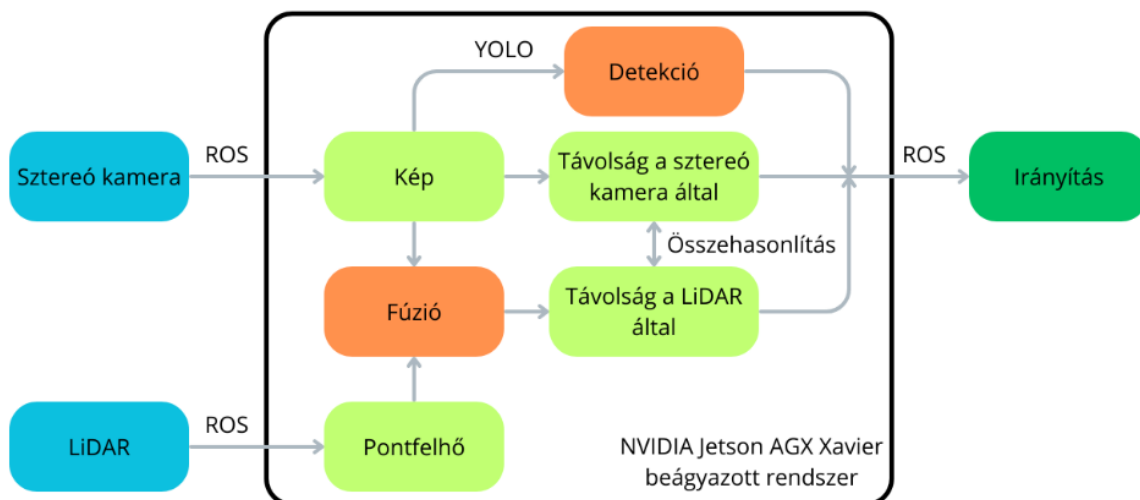


1. ábra. A neurális hálózat detekciói valós környezetben

4. A TÁVOLSÁGMÉRÉSEK SZINKRONIZÁLÁSA

4.1. A távolságméréshez használt program

Az úthibák észlelésére fejlesztett program képes az általam tanított neurális hálózat segítségével valós időben detektálni az előre definiált objektumokat és megmérni azok távolságát a járműtől. A távolságmérés történhet sztereó kamera, illetve LiDAR segítségével. A program a Robot Operating System (ROS) [16] keretrendszerben működik. A program működése a 2. ábrán látható. Ahhoz, hogy az eredményeket összehasonlítsam előbb szinkronizálnom kellett a két mérést térben és időben.



2. ábra. A program működésének folyamatábrája

4.2. A térben és időben való szinkronizálás

Egy kiterjedt objektum távolságát annak képpontjai alapján sokféle módon definiálhatjuk. Mivel a kutatásom a sztereó kamera és a LiDAR segítségével történő távolságmérés összehasonlításáról szól, így elengedhetetlen, hogy a két mérés távolság-definíciója egységes legyen.

A távolság definiálására három módszert vizsgáltam meg. Az első lehetőség egy konkrét, előre definiált pont távolságmérése, ami logikusan a detektált objektum befoglaló téglalapjának a középpontja. A második módszer a detektált objektum befoglaló téglalapjába eső képpontok távolságainak az átlaga, míg a harmadik azok mediánja.

Az egyes módszerek kiértékelését a Livox és az Ouster LiDAR mért értékeinek összehasonlításával tettem. Az volt a feltételezésem, hogy mivel mindkettő szenzor nagy precizitással bír, így a legpontosabb módszer fogja adni a legkisebb relatív hibát közöttük. A befoglaló téglalap középpontjának, az átlagos távolságnak és a medián távolságnak a relatív hibái rendre 12%, 15% és 7%. Megállapítható, hogy a legpontosabb eredményeket a medián mérési módszer szolgáltatta. A módszer előnye, hogy a kiugró értékekre, illetve az objektum nagy mértékű, akár közel 50%-os takarására sem érzékeny.

A két mérés időben való szinkronizálása azért szükséges, hogy a távolságmérések ugyanazon időpillanatban történjenek. A távolságméréseket a ROS által szolgáltatott időbélyegekkel lehet ellátni, majd ezek alapján összehasonlítani az azonos bélyeggel rendelkezőket.

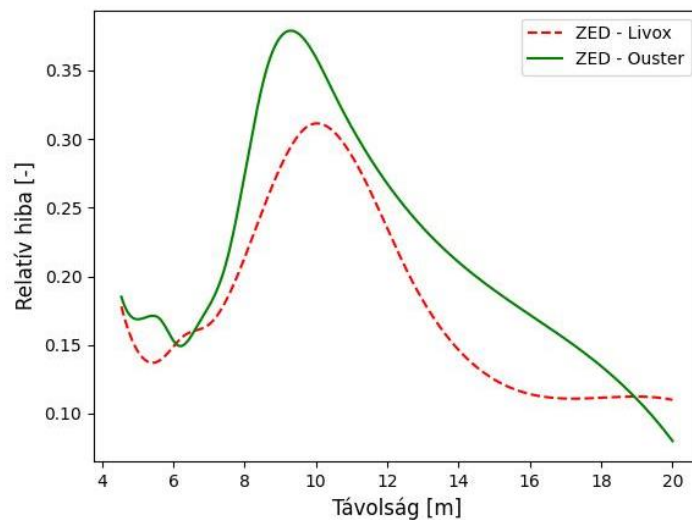
5. A TÁVOLSÁGMÉRÉSI MÓDSZEREK ÖSSZEHAJONLÍTÁSA

5.1. A sztereó kamera összehasonlítása a két LiDAR-ral

Elsődleges feladatként a sztereó kamera pontosságát elemeztem ki. Ennek érdekében összehasonlítottam a távolságértékeit mind a két LiDAR eredményeivel. Mivel a LiDAR alapú távolságmérés viszonylag pontos módszer, így ezzel értékelni lehet a sztereó kamera teljesítményét.

A ZED 2 sztereó kamera távolságmérését 648, illetve 642 detektált objektumon vettem össze az Ouster és a Livox LiDAR eredményeivel. A detektálásokra a sztereó kamera átlagos relatív hibája az Ouster-etalonnak véve 21%-os, míg a Livox-ot etalonnak véve 18%-os. A hibák jellegét elemezve megfigyelhető, hogy a detekciók nagyjából kétharmadában a sztereó kamera pozitív irányba téved, tehát távolabbit mér, mint a valóság. Ezekben az esetekben a relatív hiba is jóval nagyobb, 26%-os, illetve 24%-os a negatív tévedés 12%-ához, illetve 8%-ához képest.

A hibákat érdemes a távolság függvényében is megvizsgálni. A 3. ábrán látszik, hogy közeli objektum esetén a relatív hiba az átlagosnál kisebb, azonban egy ponton drasztikusan megnő. Ez az ugrás nagyjából 8 méteres távolság esetén lép fel. A grafikonok vége a relatív hiba mérséklődését mutatja, ez azonban a ZED 2 kamera 20 méteres maximális méréshatárának és az emiatt bekövetkező telítődésnek köszönhető. Releváns adatokat a grafikonok 10 méter alatti tartománya mutat.



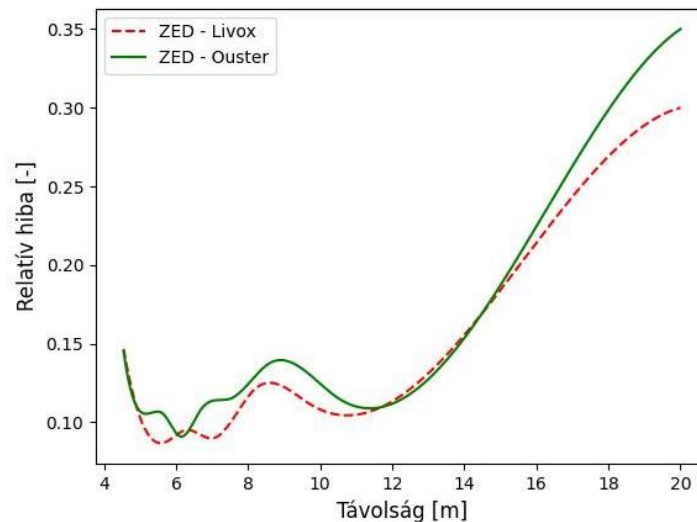
3. ábra. A sztereó kamera relatív hibája a távolság függvényében a két LiDAR-hoz képest

5.2. A karakterisztikus hibák kiküszöbölése

A sztereó kamera távolságmérése jóval több esetben téved pozitív irányban. Ebből sejthető, hogy a hibák nem csupán az eszközök különböző működéséből, hanem karakterisztikus pontatlanságokból is adódnak. Ezt okozhatja a mérés elején lévő kalibráció, valamint a kamera belső működése is. A sztereó kamera és az etalonként vett LiDAR távolságértéke között lineáris kapcsolatot feltételezve az alábbi függvényt tudjuk felírni az etalonként vett távolság (Y) és a sztereó kamera által mért távolság (X) között.

$$Y = A * X + B \quad (1)$$

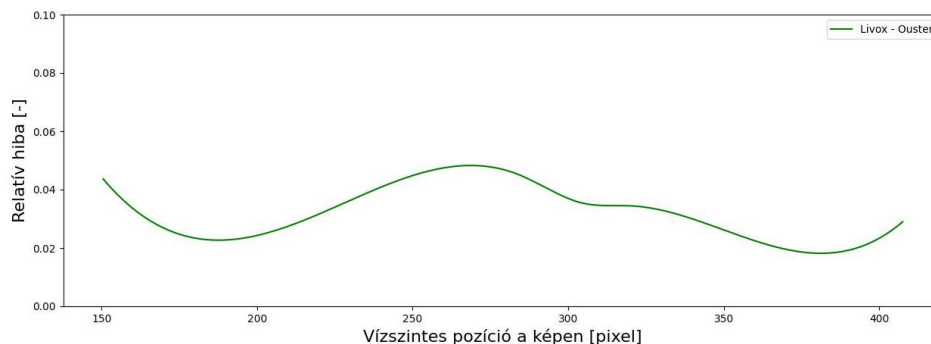
Az A és B paraméterek változtatásával az etalonhoz viszonyított relatív hiba is módosulni fog. A paraméterek helyes megválasztásával a relatív hiba minimalizálható. A relatív hiba a két paraméter függvényében széles határok között változik, minimális értéket akkor kapunk, ha az A paraméter értéke 0,62, míg a B paraméter értéke 2,1. Ebben az esetben a sztereó kamera távolságmérésének relatív hibája mindkét LiDAR-ral összehasonlítva 12%. A távolságfüggő relatív hibát a 4. ábrán lehet látni. Megfigyelhetjük, hogy a karakterisztika paramétereinek helyes beállításával a relatív hiba ugrásának helye is kitolódott 14 méterre.



4. ábra. A karakterisztika korrigálásának hatása a sztereó kamera és a két LiDAR közötti relatív hibára a távolság függvényében

5.3. Az Ouster és a Livox LiDAR összehasonlítása

A szkennelési módszerükben való eltérés miatt a Livox pontfelhője a kép vízszintes tengelye mentén a kép széléhez közelítve egyre kisebb sűrűséggel rendelkezik, míg az Ouster pontjai egyenletesen oszlanak el. Az 5. ábrán a Livox relatív hibáját láthatjuk az Ouster-hez viszonyítva a kép vízszintes tengelye mentén. Megfigyelhetjük, hogy a relatív hiba nagyjából állandó, 3-5%-os tartományban marad. Kijelenthető tehát, hogy a pontfelhő sűrűségének változása nincs hatással a relatív hiba mértékére.



5. ábra. A Livox relatív hibája az Ouster-hez képest a kép vízszintes tengelye mentén

5.4. Konklúzió

Az eredmények alapján látható, hogy a ZED 2 sztereó kamera közeli objektum távolságmérése esetén viszonylag pontos tud lenni. Ezt a pontosságát egy adott határ után jelentős mértékben elveszíti. A karakterisztika korrigálásával el lehetett érni, hogy a mérések során a relatív hiba csökkenjen és a pontatlanság ugrása kitolódjon. Amennyiben ezek az eredmények nem kielégítőek egy autonóm járműveket tervező vállalat számára, akkor a távolság mérését érdemesebb LiDAR-ral végezni.

A Livox LiDAR vizsgálatából kijelenthető, hogy a pontfelhő sűrűségének változása nincs jelentős hatással a relatív hiba változására. Ennek az oka elsősorban a távolság definiálására kiválasztott medián módszer miatt adódik, aminek a pontosságát nem befolyásolja a befoglaló téglalapon belül található pontok alacsony száma. Ezek alapján megállapítható, hogy ha egy autonóm járműveket gyártó vállalat a LiDAR-ral történő távolságmérést preferálja, akkor a Livox Avia-t is használhatja az Ouster OS0 helyett. A két eszköz közötti jelentős árkülönbség miatt pedig ez a döntés jelentősen csökkentheti az autonóm jármű gyártásának költségeit.

6. ÖSSZEFOGLALÁS

A manapság egyre inkább elterjedő autonóm járművek egyik legfontosabb feladata a környezetükben található objektumok észlelése. A kátyúk, fekvőrendőrök és egyéb akadályok detektálása kulcsfontosságú egy

önvezető autó biztonságos működése szempontjából. Ezen objektumok detektálására felügyelt tanulással neurális hálózatot hoztam létre, ami alkalmas kátyúk, fekvőrendőrök, csatornafedők, csatornalefolyók és bicikliút jelzések detektálására.

Az előre definiált objektumok távolságának mérésére két módszert, a sztereó látást és a lézeres távolságmérést teszteltem. Ehhez a két mérést térben és időben szinkronizálnom kellett. A térbeli szinkronizáció során összehasonlítottam a kiterjedt objektumok távolságának lehetséges definícióit és meghatároztam, hogy a legpontosabb eredményt a medián módszer szolgáltatja. A sztereó kamera és a két LiDAR összehasonlítása során látható volt, hogy előbbi eszköz mérsékelt pontossággal rendelkezik, ami egy bizonyos távolsági határ után jelentősen csökken. A kutatásom során kielemeztem a sztereó kamera karakterisztikáját is és annak kompenzálásával csökkenteni tudtam a mérések relatív hibáját. A két LiDAR szkennelési módszereinek elemzése során megállapítottam, hogy a medián alapú távolságmérés miatt a pontfelhő sűrűségének csökkenése nem jár a pontosság változásával.

A bemutatott programok, módszerek és mérések alkalmasak egy autonóm járműveket tervező vállalat számára is. Az általam fejlesztett neurális hálózat képes bizonyos úthibák és akadályok észlelésére, az elvégzett kutatás pedig alapot adhat egy konkrét termék szenzorainak és mérési módszereinek felülvizsgálatára is.

7. KÖSZÖNETNYILVÁNÍTÁS

A publikációban szereplő kutatást a HUN-REN SZTAKI az Európai Unió támogatásával valósította meg, az Autonóm Rendszerek Nemzeti Laboratórium keretében (RRF-2.3.1-21-2022-00002). A kutatást részben a TKP2021-NVA-01 projekt támogatta. Köszönet illeti a D3Seeron Ltd-t és az ESA Spark Funding-ot a pénzügyi támogatásért.

8. IRODALMI HIVATKOZÁSOK

- [1] S. T. Barnard, M. A. Fischler, Computational stereo, ACM Computing Surveys (CSUR), 1982, 553-572. old.
- [2] J. Kocić, N. Jović, V. Drndarević, Sensors and sensor fusion in autonomous vehicles, 2018 26th Telecommunications Forum (TELFOR), 2018, 420-425. old.
- [3] ***, Stereolabs ZED 2, Stereolabs, <https://www.stereolabs.com/products/zed-2> (Utolsó letöltés: 2024.02.19.)
- [4] L. Li, Time-of-flight camera - An introduction, Technical White Paper, 2014
- [5] J. Liadsky, Introduction to LIDAR, NPS Lidar Workshop, 2007, 1-41. old.
- [6] Y. Li, J. Ibanez-Guzman, Lidar for autonomous driving: The principles, challenges, and trends for automotive lidar and perception systems, IEEE Signal Processing Magazine, 2020, 50-60. old.
- [7] ***, Ouster OS0, Ouster, <https://ouster.com/products/hardware/os0-lidar-sensor> (Utolsó letöltés: 2024.02.19.)
- [8] ***, Livox Avia, Livox, <https://www.livoxtech.com/avia> (Utolsó letöltés: 2024.02.19.)
- [9] ***, Livox Introduces High Performance, Low Cost, Mass Market Lidar Sensors For L3/L4 Autonomous Driving Applications, Livox, <https://www.livoxtech.com/news/4> (Utolsó letöltés: 2024.02.19.)
- [10] Z. Li és tsai., A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects, IEEE transactions on neural networks and learning systems, 2021
- [11] J. Redmon és tsai., You Only Look Once: Unified, Real-Time Object Detection, Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, 779-788. old.
- [12] B. T. Passos és tsai., Cracks and potholes in road images, Mendeley Data, 2020
- [13] O. Jayasinghe és tsai., CeyMo: See More on Roads - A Novel Benchmark Dataset for Road Marking Detection, Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2022, 3104-3113. old.
- [14] A. Bochkovskiy, C. Wang, H. M. Liao, YOLOv4: Optimal Speed and Accuracy of Object Detection, arXiv, 2020
- [15] D. Shah, Mean Average Precision (mAP) Explained: Everything You Need to Know, V7 Labs, <https://www.v7labs.com/blog/mean-average-precision> (Utolsó letöltés: 2024.02.19.)
- [16] M. Quigley és tsai., ROS: an open-source Robot Operating System, ICRA workshop on open source software, 2009, 5. old.