

„Mesterséges intelligencia rendszerek alkalmazása a mobilitásban”

2025. tavasz
Konferencia kiadvány

MESTERSÉGES INTELLIGENCIA
RENDSZEREK ALKALMAZÁSA
A MOBILITÁSBAN
2025. tavaszi konferenciakiadvány

Lektorok:

Dr. Eisingerné Dr. Balassa Boglárka, Dr. Polák József,
Dr. Szénásy István, Dr. Prukner Péter, Dr. Kulmány István Mihály,
Dr. Busznyák Tibor, Dr. Szalay István, Dr. Enisz Krisztián,
Dr. Suta Alex, Dr. Nagy Andor, Dr. Péter Tamás,
Dr. Titrik Ádám, Dr. Speiser Ferenc Péter, Dr. Varga Zoltán

Szerkesztette:

Dr. Szauter Ferenc

Pup Dániel

Csikor Dániel

Földesi Rita

Koteczki Réka

Széchenyi István Egyetem

2025

ISBN 978-615-6443-41-0

Tartalomjegyzék

RRF-2.3.1-21-2022-00004 – MESTERSÉGES INTELLIGENCIA NEMZETI LABORATÓRIUM 6

VÁROSI KERÉKPÁROS INFRASTRUKTÚRA MODELLEZÉSE 3D-BEN: KÍSÉRLETI KERÉKPÁROS SZIMULÁCIÓS VIZSGÁLAT

GYŐR VÁROSÁBAN..... 8

BOKSÁN BENEDEK^A, AMIRA HAMMAMI^B, NAGY VIKTOR^C 8

REGIONÁLIS CAR-SHARING RENDSZER LEHETŐSÉGEI GYŐR-MOSON-SOPRON VÁRMEGYÉBEN..... 17

CSIKOR DÁNIEL^A 17

KÖZÚTI TERELŐBÓJÁK KULCSPONT-ALAPÚ DETEKTÁLÁSA 29

HOLLÓSI JÁNOS^A 29

EGYSÉGES MÓDSZERTAN JÁRMŰVEK TÁVIRÁNYÍTÁSÚ RENDSZEREKKÉ TÖRTÉNŐ ÁTALAKÍTÁSÁRA..... 47

KECSKEMÉTI ISTVÁN^A, KÖRÖS PÉTER^B 47

TECHNOLÓGIAELFOGADÁSI MINTÁZATOK AZ EGYETEMISTÁK KÖRÉBEN: KLASZTERALAPÚ MEGKÖZELÍTÉS 56

KOTECZKI RÉKA^A 56

NEURÁLIS HÁLÓZATI MODELLEK ÖSSZEHASONLÍTÓ VIZSGÁLATA JÁRMŰ-ORIENTÁCIÓ OSZTÁLYOZÁSÁRA..... 67

MAGDA DONÁT^A, HOLLÓSI JÁNOS^B 67

TRANSFORMERS AND LLMs IN COMPUTER VISION: CURRENT ADVANCES AND FUTURE DIRECTIONS 83

NORBERT MARKÓ^A, ÁRON BALLAGI^B, TAMÁS SZIRÁNYI^C 83

VIZUÁLIS ODOMETRIA ALGORITMUSOK HATÉKONYSÁGI VIZSGÁLATA..... 101

NAGY ARMAND^A, HOLLÓSI JÁNOS^B 101

A MENTÁLIS TERHELÉS KÉRDŐÍVES ÉRTÉKELÉSE SZIMULÁLT VILLAMOSVEZETÉSI KÖRNYEZETBEN..... 115

NAGY MÁRTON^A, BAGOSI ATTILA FERENC^B, NAGY VIKTOR^A 115

VÁLLALATI ÉS GAZDASÁGI SZERKEZET ÖSSZEHASONLÍTÓ ELEMZÉSE MAGYARORSZÁG ÉS CSEHORSZÁG NUTS 3

SZINTJÉN..... 131

REIDER-PESTI BEÁTA..... 131

KORSZERŰ NEURÁLIS HÁLÓZATOK ÖSSZEHASONLÍTÓ ELEMZÉSE AZ ÚTSZEGMENTÁLÁS TERÜLETÉN 150

REISINGER RICHÁRD^A, HOLLÓSI JÁNOS^B 150

SZEMLE AZ ÖNVEZETŐ JÁRMŰVEK ÉS A GÉPKOCSIVEZETŐK VEGYES KÖZLEKEDÉSÉNEK LEHETSÉGES KIHÍVÁSAIRÓL.... 163

POTENTIAL CHALLENGES FOR MIXED TRANSPORT INVOLVING SELF-DRIVING VEHICLES AND HUMAN DRIVERS..... 163

SALY GÁBOR^A 163

PARAMÉTEREZHETŐ VIRTUÁLIS KÖRNYEZET VIZUÁLIS ODOMETRIA ALGORITMUSOK FEJLESZTÉSÉHEZ 176

TÖRÖK BENCE^A, HOLLÓSI JÁNOS^B 176

MESTERSÉGES INTELLIGENCIA RENDSZEREK ALKALMAZÁSA A MOBILITÁSBAN 187

GULYÁS PÉTER, KOVÁCS ÁDÁM TAMÁS, LÉGRÁDI MÁRTON GÁBOR..... 187

RRF-2.3.1-21-2022-00004 – MESTERSÉGES INTELLIGENCIA NEMZETI LABORATÓRIUM

RRF-2.31-21-2022-00004

Konzorciumvezető: Számítástechnikai és Automatizálási Kutatóintézet (SZTAKI)

Konzorciumi partnerek:

- Széchenyi István Egyetem (SZE)
- Budapesti Műszaki és Gazdaságtudományi Egyetem (BME)
- Rényi Alfréd Matematikai Kutatóintézet (Rényi)
- Eötvös Loránd Tudományegyetem (ELTE)
- Szegedi Tudományegyetem (SZTE)
- Semmelweis Egyetem (SE)
- Kísérleti Orvostudományi Kutatóintézet (KOKI)
- Társadalomtudományi Kutatóközpont (TK)
- Nemzetbiztonsági Szakszolgálat (NBSZ)
- Kincstári Informatikai Nonprofit Felelősségű Társaság (KINCSINFO)

A projekt teljes összege: 9 439 666 600 Ft

Támogatási összeg (SZE): 470 000 000 Ft

Támogatás mértéke: 100%

A projekt időtartama: 2022.03.01. – 2025.06.30.

Projektismertető:

A Mesterséges Intelligencia Nemzeti Laboratórium projekt célkitűzése mesterséges intelligencia alkalmazások kutatása, hálózatos működés kialakítása és koordináció az alábbi kutatási területeken:

- Mesterséges intelligencia alkalmazási lehetőségeinek kutatása az IPAR 4.0 technológiai fejlesztések területén.
- Mesterséges intelligencia alkalmazási lehetőségeinek kutatása az autonóm közlekedési rendszerek technológiai fejlesztések területén.
- Mesterséges intelligencia alkalmazási lehetőségeinek kutatása az infokommunikációs technológiák fejlesztési területein:

Weboldal: <https://mi.nemzetilabor.hu/>

A projekt megvalósítását Magyarország Kormánya az Európai Unió Helyreállítási és Ellenállóképességi Eszközének finanszírozásával támogatja a Mesterséges Intelligencia Nemzeti Laboratórium program (RRF-2.3.1-21-2022-00004) keretében.

Városi kerékpáros infrastruktúra modellezése 3D-ben: kísérleti kerékpáros szimulációs vizsgálat Győr városában

Modeling urban cycling infrastructure in 3D: Experimental bicycle simulator study in the city of Győr

Boksán Benedek ^a, Amira Hammami ^b, Nagy Viktor ^c

^aSzéchenyi István Egyetem

boksanbenedek@gmail.com

^bSzéchenyi István Egyetem

hammami.amira@ga.sze.hu

^cSzéchenyi István Egyetem

nviktor@sze.hu

Absztrakt

A kerékpárszimulátor-technika az utóbbi időben egyre nagyobb figyelmet kapott a közlekedésbiztonsági kutatásokban, mivel hatékonyan szolgálja a kutatók igényeit. A módszer legfőbb kihívása azonban a valósághűség biztosítása. Egy magas fokú realizmust nyújtó kerékpárszimulátor alapfeltétele egy élethű virtuális környezet megléte. A jelen tanulmány bemutatja a magyarországi Győr városában található városi útszakasz és környezet 3D-s modelljének kidolgozását. A vizsgált környezet két infrastruktúra-verzióban is elkészült: kerékpáros nyom (piktogram) és kerékpár sáv. A különböző nyom- és sávkialakításokat egy kerékpárszimulátor segítségével végzett kísérleti vizsgálat során értékeltük, a kerékpárosok biztonságérzetének szempontjából. Teljesítmény tekintetében a virtuális környezet (3D modell) kapta a legmagasabb értékelést a valósággal való hasonlósága alapján.

Kulcsszavak: 3D modellezés, kerékpáros szimulátor, virtuális környezet, valósághűség, kerékpáros infrastruktúra

Abstract

The bicycle simulator technique has recently received increasing attention as a traffic safety research tool that meets researchers' needs. However, the main issue of this tool is the realism. A high-realistic bicycle simulator relies on a realistic virtual environment. This work introduces the development of a 3D model of an urban road environment from the city of Győr, in Hungary. Two alternative design scenarios were also created for the same environment. An experimental study was conducted, using a bicycle simulator, to evaluate cyclists' perception of safety in the three designs. In terms of performance, the virtual environment (3D model) achieved the highest rating for its similarity to the real world.

Keywords: 3D modelling, bicycle simulator, virtual environment, realism, cycling infrastructure

1. Bevezetés

A kerékpározás kiemelt jelentőséggel bír a városi közlekedésben, különösen a biztonság szempontjából [1]. Az olyan új technológiák megjelenésével, mint az autonóm járművek és az elektromos rollerek, egyre összetettebb közlekedési szituációk alakulnak ki az utakon. Mindez jelentős kihívást jelent abban, hogy minden közlekedő számára biztonságos infrastruktúrát biztosítsunk. Ennek megfelelően innovatív megközelítésekre van szükség a kerékpárosok biztonságának és viselkedésének vizsgálatához ilyen összetett közlekedési helyzetekben. Az utóbbi időben több kutatás is a kerékpárszimulátor alkalmazása felé fordult, mint hatékony kutatási eszköz felé [2,3]. A közlekedési rendszerek fejlesztése során a szimulációs környezetek – mint például a kerékpár szimulátorok – lehetőséget kínálnak a komplex rendszerkapcsolatok és az ember-gép interakciók rendszerszintű vizsgálatára és optimalizálására [4]. A szimulátorokban a 3D modellezés kulcsszerepet játszik a valósághű virtuális környezetek kialakításában. A kerékpárosok új útkialakításokra és bonyolult forgalmi helyzetekre adott észlelései és reakciói nagyban függenek attól, hogy mennyire kidolgozott az adott környezet. Ebben a folyamatban a 3D modellek részletessége meghatározó szerepet tölt be. Jelen kutatás ehhez a tudományos diskurzushoz járul hozzá azáltal, hogy bemutatja egy magyarországi, győri városi útkörnyezet 3D-s modelljének fejlesztését. A tanulmány három különböző virtuális környezeti kialakítást vizsgál egy kerékpárszimulátoros kísérlet keretében, a kerékpárosok biztonságérzetének értékelése céljából.

2. A kerékpárszimulátorok vizuális valósághűsége

A valósághű kerékpározási élmény biztosítása érdekében a kerékpárszimulátorok valósághűsége két fő összetevőre bontható: technikai és vizuális aspektusokra. Míg a technikai valósághűség a hardverelemek – például a pedálok, a fék, a kormány stb. – dinamikáját tükrözi [5], addig a vizuális valósághűség a 3D-s környezeti részletek pontossága által keltett vizuális realizmusra vonatkozik. A pontos útburkolati textúrák, a megvilágítás, az árnyékok és a geometriai arányok hiteles megjelenítése elősegíti a térbeli tájékozódást a virtuális környezetben [6]. Emellett hozzájárul a szimulált jelenet hihetőségéhez és elmerülést nyújtó élményéhez, különösen akkor, amikor új technológiákat vagy útkialakításokat vizsgálnak. A szimulátor valósághűsége hatással van a résztvevők kognitív terhelésére is, amely kapcsolatban áll a szimulátorhasználat során jelentkező mozgásbetegséggel [5]. Minél nagyobb az elmerülés

élménye (immersiveness), annál alacsonyabb a kognitív megterhelés [7]. Éppen ezért kiemelten fontos figyelmet fordítani a virtuális környezet 3D modelljének részletességére és minőségére. Egy következő tanulmány a 3D modellezésen alapuló virtuális szimulációs rendszerek tervezését és kulcstechnológiáit mutatja be, különös tekintettel az ipari, oktatási és mérnöki alkalmazásokra [8]. A rendszerek közül kiemelve a Blender egy nyílt forráskódú 3D modellezőszoftver, amely sokoldalú eszközkészletével kiválóan alkalmazható szimulációs környezetek létrehozására és vizualizálására.

3. A kerékpárszimulátorral végzett vizsgálat

A kerékpárszimulátor felépítése

A győri Széchenyi István Egyetemen kifejlesztett kerékpárszimulátor (SZEBS.1) egy eredeti kialakítás továbbfejlesztett verziója (1. ábra) [2]. A hagyományos kerékpárszimulátorokhoz hasonlóan a rendszer két alapvető egységből épül fel: szoftveres és hardveres komponensekből [3]. A kerékpárszimulátor szoftvere egy számítógépen fut, és olyan virtuális környezetet biztosít, amelyben a felhasználó interakcióba léphet a szimulációval. A szoftver a szimuláció digitális vonatkozásait foglalja magában, beleértve a szimulációs logikát (számítógépes programokat és algoritmusokat), a virtuális környezetet, valamint más szoftveres elemeket is [3]. Az SZEBS.1 szimulátor a Unity játékmotort alkalmazza.

A valóság-hű kerékpározási élmény érdekében a szimuláció valós időben reagál a felhasználó mozgásaira (pedálozás, fékezés, kormányzás). A szoftver vezérli a virtuális környezetet és a szimulációs logikát, míg a hardver biztosítja a fizikai kapcsolatot a felhasználó és a szimuláció között. A SZEBS.1 esetén a virtuális környezet modellezésére a Blender szoftver került felhasználásra, amely valós vagy képzeletbeli világ vizuális reprezentációját teszi lehetővé. A grafikai megjelenítéshez három darab 4K felbontású 55 colos OLED kijelző szolgál.



1. ábra: SZEBS.1

Esettanulmány

Az SZEBS.1 kerékpárszimulátort egy kísérlet során alkalmaztuk a SZESIM Jármű- és Közlekedésszimulációs Laboratóriumban, 2025 januárja és márciusa között. A vizsgálat célja az volt, hogy feltárja, miként befolyásolják a különböző kerékpáros infrastruktúrák az autonóm járművek (AV-k) korszakában a kerékpárosok biztonságérzetét és észleléseit. A vizsgálatban 50 önkéntes kerékpáros vett részt. A tesztet megelőzően minden résztvevő egy ismerkedő szakaszon tekert, amelynek célja az volt, hogy a kerékpárosok megismerkedjenek a szimulátor működésével és annak használatával. Ezt követően vettek részt a különböző közlekedési szituációkat modellező szimulációs teszt sorozaton.

A vizsgálat alapjául szolgáló virtuális környezet egy városi, osztott használatú útszakaszt modellez Győr városában, ahol kerékpárosok és gépjárművek osztoznak a közúti felületen. Az autonóm járművek (AV) térnyerésével azonban egyre vitatottabbá válik az osztott úthasználat koncepciója. Mivel az AV-k ismertek precíz sávkövetési algoritmusokkal rendelkeznek [9, 10], korábbi kutatások azt javasolják, hogy a járműsávok szélességét csökkentsék, és a felszabaduló területet kerékpársávok kialakítására fordítsák [11, 12]. Ezzel összhangban a vizsgálat során két útkialakítási lehetőséget teszteltünk: osztott úthasználatot (A forgatókönyv), valamint elválasztott sávos kialakítást (B, C, D és E forgatókönyvek). Az elválasztott sávos változatok esetében két különböző sáv szélességet és az útburkolati jelek hatását vizsgáltuk, az 1. táblázatban bemutatott módon.

1.táblázat: A vizsgált forgatókönyvek leírása

<i>Jellemzők</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>
Infrastruktúra leírása	Osztott úthasználat, piktogramokkal	AV-k és kerékpársávok elválasztása folytonos színes vonallal			
Útburkolati jelölés	Piktogramok 50 méterenként	Sárga folytonos vonal, kerékpáros burkolatszín nélkül		Sárga folytonos vonal, piros kerékpársáv	
Sávszélesség (m)	Összesen: 4	AV: 2,75	AV: 2,5 Kerékpár: 1,5	AV: 2,75 m Kerékpár: 1,25 m	AV: 2,5 Kerékpár: 1,5 m

4. A virtuális környezet

A vizsgálati környezet virtuális megvalósítása a Blender 3D modellezőszoftver segítségével történt. A modellezési folyamat kezdeti szakaszában a virtuális tér alapvető jellemzőinek meghatározása valósult meg, amely magában foglalta a világítási viszonyok, a hangulat és a kameranézetek beállítását. A kísérleti célok eléréséhez elengedhetetlen volt a méretpontos kialakítás és az arányos tértervezés biztosítása.

Az assetpack kidolgozása során referenciaként a győri Szövetség utca Google Street View nézete szolgált, amely alapján történt meg a környezeti elemek, a parkolók és az épületek parametrikus modellezése, valamint részleteinek kialakítása. A textúrázás során képalapú és procedurális módszerek kombinációját alkalmaztuk a valósághű megjelenés érdekében, különös figyelmet fordítva a fizikai jellemzők – például a fényvisszaverődés és a térbeli hatás – élethű leképezésére. A folyamat célja egy technikailag pontos, vizuálisan hiteles és újrafelhasználható virtuális városi környezet létrehozása volt, amely támogatja a későbbi szimulációs és interakciós vizsgálatokat.

A modellezési folyamat lezárását követően a véglegesített pálya a Unity játékmotorba került integrálásra, ahol interaktív környezetté alakítottuk. Ez a lépés magában foglalta a dinamikus elemek – különböző típusú gépjárművek, valamint gyalogosok – importálását is. A játékos megjelenítéséhez elsőként egy alapkerékpár-modell elkészítésére volt szükség, amely a szimuláció során a felhasználó által irányított jármű szerepét tölti be. Ezzel párhuzamosan egy ehhez illeszkedő dinamikus kamerarendszert is kialakítottunk, amely képes a kerékpár

mozgását valós időben követni, alkalmazkodva annak pozíciójához és irányváltásaihoz. Ez a két elem együttesen biztosítja a felhasználó számára élvezetes és koherens vizuális élményt.

A szimuláció során egy dinamikus nézőpontot biztosító, játékosnézeti kamerarendszer került alkalmazásra, amely lehetővé tette a tér interaktív és valós idejű bejárását (2. ábra).



2. ábra: Blenderben megvalósított 3D környezet megjelenítése Unity szoftverben

4. Eredmények

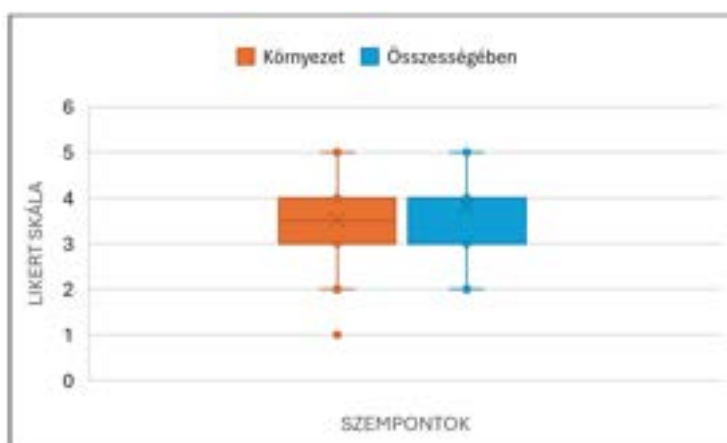
Fontos megjegyezni, hogy a jelen szakasz kizárólag a 3D modell kísérletben történő értékelésére fókuszál. Az esettanulmány eredményeit egy külön tanulmány dolgozza fel, amely önálló közleményként kerül publikálásra.

A résztvevők a kerékpárszimulátor általános teljesítményét átlagosan 3,82 pontra értékelték az ötpontos skálán. A virtuális környezet értékelése során a résztvevőknek a következő kérdésre kellett válaszolniuk egy 5-fokú Likert-skála segítségével: „Milyen mértékben tükrözi a virtuális környezet a valóságot?” A 2. táblázat a válaszok leíró statisztikai mutatóit tartalmazza. Az eredmények alapján a résztvevők úgy ítélték meg, hogy az SZEBS.1 szimulátor megfelelő vizuális valósághűséggel rendelkezik, az értékelések átlaga 3,52/5 (azaz 70%) volt.

2.táblázat A virtuális környezet értékelésének statisztikai mutatói

<i>Mutató</i>	<i>Érték</i>
Átlag	3.52
Általános Hiba	0.122
Medián	3.5
Módusz	3
Szórás	0.863
Variancia	0.744
Csúcsosság	0.456
Ferdeség	-0.263
Terjedelem	4
Minimum	1
Maximum	5

A 3. ábra szemlélteti a szimulátor teljesítményének, valamint a virtuális környezet értékelésének eloszlását. A két boxplot alakja jelentős hasonlóságot mutat, azonban néhány eltérés is megfigyelhető: a virtuális környezet értékelésének átlaga kissé alacsonyabb, mint a teljes szimulátorélményé.



3. ábra Az értékelések eloszlása a szimulátor teljesítménye és a virtuális környezet vonatkozásában

5. Összegzés

A tanulmány rámutatott arra, hogy a valósághű 3D modellezés meghatározó szerepet tölt be a kerékpárszimulátorok vizuális hűségének növelésében, ezáltal közvetlenül hozzájárulva a felhasználói élmény minőségéhez. A győri városi környezet alapján készült modell sikeresen reprodukálta a valós közlekedési helyzetet, és a résztvevők visszajelzései alapján megfelelő szintű vizuális megbízhatóságot nyújtott. A szimuláció során alkalmazott különböző forgatókönyvek – különösen a sávkialakítási megoldások és a burkolati jelzések változatai –

hasznos alapot kínálnak a jövőbeni várostervezési és közlekedésfejlesztési döntések megalapozásához. A szimulátor által biztosított élethű vizuális élmény így hozzájárulhat a kerékpárosbarát közlekedési infrastruktúra hatékonyabb tervezéséhez és kialakításához..

5. Köszönetnyilvánítás

A cikk kutatásaihoz a „Nemzeti Laboratóriumok 2020 Program – Mesterséges Intelligencia alprogram – Mesterséges Intelligencia Nemzeti Laboratórium (MILAB) létrehozása a Széchenyi István Egyetemen (NKFIH-870-21/2020)” biztosított forrást.

Irodalomjegyzék

- [1.] Christ, A. K., Costa, M., Marques, M., Roque, C., & Moura, F. (2023). Perceiving objective cycling safety: a systematic literature review. *Transportation Research Procedia*, 72, 1380–1387. <https://doi.org/10.1016/j.trpro.2023.11.601>
- [2.] Hammami, A., Borsos, A., & Sándor, Á. P. (2025). How realistic a bicycle simulator can be? - A validation study. *Multimodal Transportation*, 4(1), 100193. <https://doi.org/10.1016/J.MULTRA.2025.100193>
- [3.] Shoman, M. M., & Imine, H. (2021). Bicycle Simulator Improvement and Validation. *IEEE Access*, 9, 55063–55076. <https://doi.org/10.1109/ACCESS.2021.3071214>
- [4.] Kalincsák, I., & Polák, J. (2018). Az irányítás szerepe és helyzete a közlekedési rendszerben. In XII. IFFK 2018: Innováció és fenntartható felszíni közlekedés (pp. 274–282).
- [5.] Matviienko, A., Hoxha, H., & Mühlhäuser, M. (2023). What does it mean to cycle in Virtual Reality? Exploring Cycling Fidelity and Control of VR Bicycle Simulators. *Conference on Human Factors in Computing Systems - Proceedings*, 23. https://doi.org/10.1145/3544548.3581050/SUPPL_FILE/3544548.3581050-VIDEO-FIGURE.MP4
- [6.] Wallet, G., Sauzéon, H., Pala, P. A., Larrue, F., Zheng, X., & N’Kaoua, B. (2011). Virtual/real transfer of spatial knowledge: Benefit from visual fidelity provided in a

- virtual environment and impact of active navigation. *Cyberpsychology, Behavior, and Social Networking*, 14(7–8), 417–423. <https://doi.org/10.1089/cyber.2009.0187>
- [7.] Wintersberger, P., Matviienko, A., Schweidler, A., & Michahelles, F. (2022). Development and Evaluation of a Motion-based VR Bicycle Simulator. *Proceedings of the ACM on Human-Computer Interaction*, 6(MHCI). <https://doi.org/10.1145/3546745>
- [8.] Zhang, Y. (2023). Design and Implementation of Virtual Simulation System Based on 3D Modeling Technology. *Journal of Electronics and Information Science*, 8(6). <https://doi.org/10.23977/jeis.2023.080608>
- [9.] Jeong, Y. (2022). Interactive Lane Keeping System for Autonomous Vehicles Using LSTM-RNN Considering Driving Environments. *Sensors (Basel, Switzerland)*, 22(24), 9889. <https://doi.org/10.3390/S22249889>
- [10.] Pup D., Szakállas G. & Polák J. (2018), Research of vehicle parameter and sensor systems necessary to control autonomous vehicles. 14th IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications (MESA), Oulu, Finland, 1-8, <https://doi.org/10.1109/MESA.2018.8449146>
- [11.] Mccarville, I., Hajrasouliha, A. H., & Boswell, M. (2019). How Autonomous Vehicles will Reshape the Urban Landscape.
- [12.] Schlossberg, M., Riggs, W., Adam Millard-Ball, L. A., & Shay, E. (2018). Rethinking the street in an era of driverless cars. www.urbanismnext.com

Regionális car-sharing rendszer lehetőségei Győr-Moson-Sopron vármegyében

Possibilities of a regional car-sharing system in Győr-Moson-Sopron county

Csikor Dániel^a

^aSzéchenyi István Egyetem, Járműipari Kutatóközpont

csikor.daniel@ga.sze.hu

Absztrakt

A mobilitás a 21. századi gazdaság és társadalom egyik legjelentősebb mozgatórugója és egyben legnagyobb kihívása is. A közlekedési és logisztikai szektor az Európai Unióban 11 millió embernek ad megélhetést, ami a teljes foglalkoztatottság 5 százalékának felel meg. A vezetéssel töltött összes idő 28,6 százalékát töltik forgalmi torlódásban az EU 27 tagállamának állampolgárai átlagosan. A torlódásokhoz köthető kiadások az uniós GDP 1 százalékát teszik ki, miközben az ágazat a GDP 5 százalékát adja. A társadalom és gazdaság részéről egyre inkább fokozódik az igény a minél gyorsabb és kényelmesebb közlekedés iránt, melyhez elkerülhetetlen a vezetéstámogató rendszerek fejlesztése és elterjedése, a járművek közötti kommunikáció kialakítása, valamint az infrastruktúra korszerűsítése. A különböző autómegosztó rendszerek, mobilitási szolgáltatások globális elterjedésével jelentősen csökkenthetők lennének a forgalmi torlódások, továbbá a közúti közlekedési balesetek. A kutatás keretében elemzésre kerültek egy autómegosztó szolgáltatás kialakításához szükséges tevékenységek, feladatok a felhasználók, szolgáltatók és az autógyártók oldaláról is. Kidolgozásra került egy regionális car-sharing koncepció Győr-Moson-Sopron vármegyére vonatkozóan. A rendszer működési modelljének, előnyeinek és a lehetséges felhasználói körnek az ismertetése mellett a szolgáltatás elindításának nehézségei is megvizsgálásra kerültek. Tanulmányozásra kerültek az elérhető elektromos autó töltőállomások típusai és elhelyezkedésük a vizsgált régióban, illetve a fejlesztési, bővítési lehetőségek is. Áttekintésre kerültek a töltőállomások rendszertelen telepítésének okai, továbbá a teljesen elektromos járművek alkalmazásának kihívásai egy regionális autómegosztó szolgáltatásban.

Kulcsszavak: autómegosztás, fenntarthatóság, mobilitás, logisztika

Abstract

Mobility is one of the most important drivers of the 21st century economy and society, and one of its greatest challenges. The transport and logistics sector provides a living for 11 million people in the European Union, representing 5% of total employment. On average, 28.6% of the total driving time is spent in congested traffic by citizens of the 27 EU Member States. Congestion-related expenditure accounts for 1 percent of EU GDP, while the sector contributes 5 percent of GDP. Society and the economy are increasingly demanding faster and more comfortable transport, which will inevitably require the development and deployment of driver assistance systems, vehicle-to-vehicle communication and infrastructure modernisation. The global deployment of car-sharing systems and mobility services could significantly reduce congestion and road accidents. The research analyses the activities and tasks required to develop a car-sharing service from the perspective of users, service providers and car manufacturers. A regional car-sharing concept for the county of Győr-Moson-Sopron was developed. The operating model, the advantages of the system and the potential user base were described, and the difficulties of starting up the service were also examined. The types of electric car charging stations available and their location in the region were studied, as well as the possibilities for development and expansion. The reasons for the irregular deployment of charging stations and the challenges of deploying fully electric vehicles in a regional car-sharing service were reviewed.

Keywords: car sharing, sustainability, mobility, logistics

1. Bevezetés

A mobilitás a 21. századi gazdaság és társadalom egyik legjelentősebb mozgatórugója és egyben legnagyobb kihívása is. A közlekedési és logisztikai szektor az Európai Unióban 11 millió embernek ad megélhetést, ami a teljes foglalkoztatottság 5 százalékának felel meg. Az Európai Unióban tevékenységet folytató 915 000 közúti szállítási feladatot végző vállalkozás közvetlenül 5 millió embernek biztosít munkalehetőséget. A magyarországi munkavállalók 3,5 százalékát a közúti közlekedési szektor foglalkoztatja. Az uniós GDP 5 százalékát, megközelítőleg 500 milliárd EUR-t az áru- és személyszállítási ágazat termeli meg. Az adatokból világosan kitűnik, hogy a közlekedés mára megkerülhetetlenül részét képezi gazdaságunknak [1, 2].

Az utazás az emberiség mindennapjainak természetes velejárójává vált az idők folyamán. Az uniós állampolgárok hetente megközelítőleg 10 órát töltenek közlekedéssel és naponta hozzávetőleg 34,7 km-t tesznek meg [3]. Természetesen országonként mutatkoznak eltérések, de nincsenek igazán jelentős különbségek. Az uniós társadalom a teljes fogyasztásának 13 százalékát költi közlekedéshez kapcsolódó kiadásokra. Az ágazat óriási mértékű fellendülésének azonban nemcsak szemmel is jól látható gazdasági előnyei vannak, hanem jelentős hátrányai is, amelyek egyáltalán nem elhanyagolható mértékűek. Utazásaink során évről évre nő a forgalomban eltöltött idő aránya. A vezetéssel töltött összes idő 28,6 százalékát töltik forgalmi torlódásban az EU 27 tagállamának állampolgárai átlagosan. A közutak zsúfoltságának következtében jelentősen romlik a teherfuvarozás hatékonysága is. A torlódásokhoz köthető kiadások az uniós GDP 1 százalékát teszik ki, miközben az ágazat a GDP 5 százalékát adja. Jelentősen növelhető lenne a szektor teljesítőképessége az autópályák, autóutak és városok zsúfoltságának csökkentésével. A másik jelentős probléma a légszennyezettség folyamatos emelkedése. A közúti közlekedés, a közlekedéssel összefüggő összes szén-dioxid kibocsátás 72 százalékáért felelős [1,2,4]. A közlekedési balesetek 91 százalékát az emberi hiba, figyelmetlenség, türelmetlenség okozza. Ezeknek a mulasztásoknak az okozója legtöbbször a torlódások okozta fáradtság, mely az emberi élet természetes következménye, tartós koncentrációt követően. Olyan technológiai és szolgáltatási megoldásokat kell keresni, amelyek által teljes mértékben átalakítható a mai mobilitásról alkotott képünk [5].

A társadalom és gazdaság részéről egyre inkább fokozódik az igény a minél gyorsabb és kényelmesebb közlekedés iránt, melyhez elkerülhetetlen a vezetéstámogató rendszerek

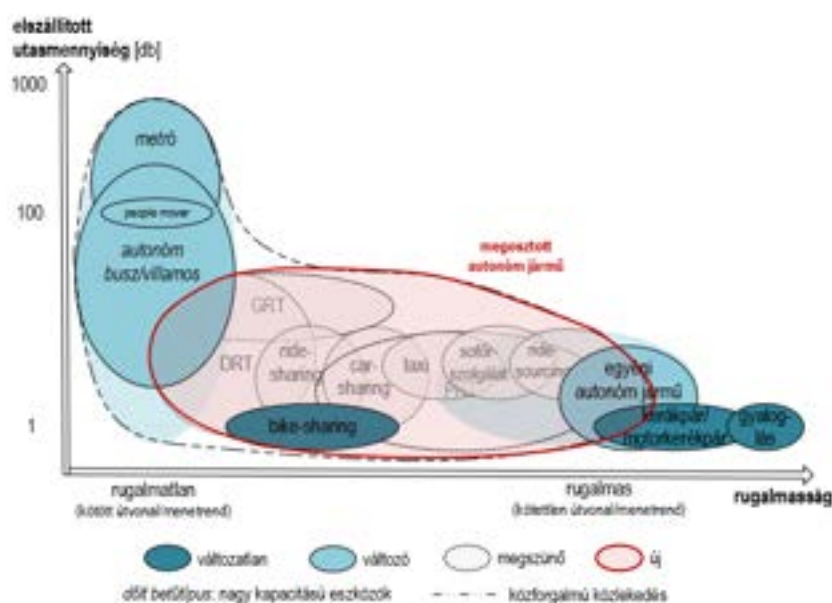
fejlesztése és elterjedése, a járművek közötti kommunikáció kialakítása, valamint az infrastruktúra korszerűsítése [2,6,7]. Nagyon ígéretes jövő áll előttünk. A különböző autómegosztó rendszerek, mobilitási szolgáltatások globális elterjedésével jelentősen csökkenthetők lennének a forgalmi torlódások, továbbá a közúti közlekedési balesetek. Jelen cikk keretében kidolgozásra került egy koncepció, hogy milyen lehetőségek vannak egy regionális rendszer kialakítására Győr-Moson-Sopron vármegyében.

2. Az autógyártók, a szolgáltatók és a jövőbeni felhasználók feladata a sikeres és hatékony car-sharing kialakításához

A különböző car-sharing rendszerek megjelenésével és ismertté válásával egyértelműen meg fog változni a jövőnk mobilitása. Számptalan alternatívát kínálnak a szolgáltatók a közlekedés megoldására a saját tulajdonú gépjármű és a tömegközlekedés helyett már napjainkban is, melyek egy része Magyarországon is elérhető és kipróbálható. A felhasználók visszajelzései alapján vannak pozitív és negatív vélemények is a rendszerekkel kapcsolatban, de ez megítélésem szerint teljesen érthető ennyire új technológiák, mobilitási koncepciók kidolgozásának a kezdetén. A szolgáltatóknak, vállalatoknak ki kell alakítaniuk az elkövetkezendő időszakban a gazdaságilag leghatékonyabb működési modellt, valamint meg kell vizsgálniuk, hogyan tudják a legnagyobb kényelmet, rugalmasságot, gyorsaságot biztosítani az ügyfeleknek a legkevesebb probléma és fennakadás mellett. A jelenlegi free-floating autómegosztó konstrukciók esetén felmerül az a probléma, hogy bizonyos területekre csoportosulnak az autók, ezáltal néhány kerületben egyáltalán nem lesz elérhető bérelhető gépkocsi, ami jelentősen csökkenti a rendszer rugalmasságát a felhasználók szempontjából. A logisztikai gondok a Budapesten működő car-sharing cégek esetén is megjelentek. Az üzemeltetőknek végig kell gondolni, hogyan lehet a legkevesebb költséggel jobb eloszlást biztosítani. Némelyik vállalkozás díjkedvezményt nyújt a régóta ki nem bérelt személygépjárművek árából, de az autók valószínűsíthetően központtól távoli elhelyezkedése miatt ez nem jelent sokaknak vonzó alternatívát. Az autonóm járművek megjelenése éppen ezért drasztikusan megváltoztathatja a car-sharing piacot. A logisztikai problémákat önállóan megoldó gépkocsik költséghatékonyabb működést eredményezhetnek a vállalatoknak. A jövőben a cégek munkatársainak nem kell áthelyeznie az autókat a jobb elérhetőség és nagyobb kihasználtság biztosítása érdekében, amit jelenleg csak sok idő és pénz ráfordításával tudnak megoldani. A bérleti percdíjakban ezek a többletköltségek természetesen megjelennek, ennek

ellenére az önvezető gépjárműveknek a bérleti díjakra gyakorolt hatását nagyon nehéz megítélni, mivel a kezdetekben várhatóan magas áron debütálnak ezek az autók. A háztól-házig szolgáltatás lehetőségének bevezetése pedig nagyfokú rugalmasságot eredményezhet az ügyfeleknek [3,4,6,8,9].

Az 1. ábra betekintést enged a különböző mobilitási lehetőségekbe az elszállított utasok számának és a közlekedési alternatívák rugalmasságának függvényében [8].



1. ábra: Mobilitási lehetőségek összehasonlítása az elszállított utasok és az alternatívák rugalmasságának függvényében¹

A cégek feladata, hogy megismertessék a polgárokkal az újfajta közlekedési konstrukciókat, amelyek megkönnyítik az emberek életét, miközben hozzájárulnak a fenntarthatóbb városok és az egészségesebb légkör kialakításához. A car-sharing szolgáltatást üzemeltető vállalkozások azonban csak akkor tudnak eredményes fejlesztéseket végrehajtani, ha megfelelő mennyiségű információval és adattal rendelkeznek a használati szokásokról és az ügyfelek elvárásairól. Az új rendszerről mindenkinek tapasztalatokat kell gyűjtenie. Mindenkinek érdeke tehát, hogy egyre többen használják ezeket a megoldásokat. A felhasználói értékelések elemzését követően lehetséges „finomhangolni”, illetve továbbfejleszteni az eddigi elképzeléseket. A kritikák és a pozitív értékelések alapján van mód az ügyfelek igényei szerinti koncepció kialakítására. Vajon melyik lesz a legjobb lehetőség a környezet védelme és a forgalmi torlódások csökkentése

¹ <https://repozitorium.omikk.bme.hu/bitstream/handle/10890/13129/ertekezes.pdf?sequence=2&isAllowed=y>

érdekében vagy az emberek szubjektivitása fog dönteni? Az elmúlt évtizedekben már sokszor bebizonyosodott, hogy egyes technológiai újítások váratlanul gyorsan hatalmas népszerűsége tettek szert, míg bizonyos fejlesztések nem tudtak átütő sikert elérni a felhasználók körében [9, 10].

Az elektromos autók térnyerése is növekszik, melyek talán mérsékelni fogják a lokális légszennyezettséget, de nem biztos, hogy a globális problémák megoldásához ez a tökéletes út.

Az önvezető gépjárművek megjelenése szintén meg fogja változtatni a jövő mobilitását. Napjainkban az autonóm gépkocsik közutakon való használatának számtalan akadálya van, melyek megoldást kívánnak. Egyre több műszaki innovációt mutatnak be az autógyártók és a szoftverfejlesztők – vezetéstámogató rendszerek a modern járművekben: adaptív távolságtartó, sávtartó, vészfékező, holtterfigyelő -, melyek szükségesek lesznek az önvezető technológiához. Számos vállalatnak vannak létező prototípusai – Nissan, Google, Tesla Motors, Bosch, Continental – melyek segítségével pontosíthatják kameráik, radarjaik, szoftvereik hatékony működését. Az autonóm autózáshoz azonban nélkülözhetetlen a hálózatba kapcsolt járművek rendszerének kidolgozása, melyhez elengedhetetlen az 5G telekommunikációs szolgáltatás kiépítése országszerte, sőt világszerte is. Infrastrukturális fejlesztésekre is szükség van, hogy a gépkocsik tudjanak kommunikálni például a parkolóházakkal, közlekedési lámpákkal és egymással is. A jogi szabályok megalkotása új feladatok elé állítja a jogászokat. Egy olyan szolgáltatásra kell szabályozást készíteni, ami még nem létezik, de jogi keretrendszer nélkül nem elképzelhető az önvezető személy- és tehergépjárműveknek a közforgalmú közlekedésben való részvétele. A legfontosabb kérdés tisztázni, hogy a hibás szoftver, technológia, infrastruktúra miatt bekövetkező balesetekért ki fogja vállalni a felelősséget, illetve hogyan lesznek használhatók az autonóm gépkocsik a hagyományos autókkal együtt városi és országúti környezetben és milyen döntést kell hoznia a járműnek, ha a baleset nem kerülhető el. Az autonóm autóktól azonban számtalan előnyt várunk: a balesetek csökkenését; egyenletesebb haladást és hatékonyabb területfelhasználást, ezáltal a torlódások mérséklését; a car-sharing elterjedését kisebb városokban, falvakban. Az autonóm autók jövőbeni megosztásának lehetősége a szolgáltatóknak a logisztika területén egyszerűbb működést, a felhasználók számára pedig kényelmesebb utazást fog biztosítani [6,7,11,12].

3. Car-sharing szolgáltatás Győr-Moson-Sopron vármegyében

Az autómegosztást tulajdonképpen magánszemélyek kezdték el, hogy rövidebb vagy hosszabb távú utazásaik költségét csökkentsék. Így alakult ki a Telekocsi rendszer, melynek előnyeit egyre többen felismerték, ezért hamar népszerű lett világszerte. Az utóbbi 10 – 20 évben azonban drasztikusan megváltozott a világ. A forgalomba helyezett gépjárművek száma nagymértékű növekedést ér el minden évben, miközben a városok levegőminősége folyamatosan romlik. Többek között sok kínai nagyvárosban folyamatosan az egészségügyi határérték felett van a levegőben lévő károsanyagok mennyisége. Az autógyártókat jogszabályok által arra kötelezik, hogy tartósan mérsékeljék a gépjárművek által kibocsátott ártalmas anyagok mennyiségét. Az utóbbi években egyre nehezebben teljesíthetők már az előírt értékek, ezért a gyártók úgy gondolják át kell alakítani a mobilitás fogalmát. A BMW és a Daimler külön-külön létrehozta a saját free-floating típusú car-sharing rendszerét, de napjainkra a két cég egyesült és ShareNow néven nyújtanak a világ nagyvárosaiban autómegosztó szolgáltatásokat. Magyarországon a Mol Magyar Olaj- és Gázipari Nyrt. reagálva a felhasználói igényekre és felismerve a fogyó kőolajkészletek lehetséges következményeit elindította Budapesten a Mol Bubi közösségi kerékpárkölcsonzó konstrukciót, majd ennek példájára a Mol Limo közösségi car-sharing-et.

A gépkocsik számának, a forgalmi torlódások és a közúti közlekedési balesetek csökkentéséhez, továbbá az országok légszennyezettségének mérsékléséhez nem elegendő azonban csak a nagyvárosokban indítani autómegosztó szolgáltatásokat, hanem ki kell terjeszteni regionális, országos szintre. Jelenleg is vannak már kezdetleges példák, amik országosan is működőképes modellek, mint például a Carpooling vagy más néven telekocsi. Jelenleg azonban a honlapon elérhető ajánlatokat átnézve főleg az országon belüli városokba, helységekre hirdetik meg a legtöbb utat, külföldre lényegesen kevesebbet. Véleményem szerint sokkal jobban kellene népszerűsíteni ezt a közlekedési formát a Győr-Moson-Sopron és Vas vármegyében napi szinten ingázó munkavállalók körében. A telekocsi szélesebb körben való elterjedése remélhetőleg a környék levegőminőségében is változást jelent majd. Úgy gondolom, hogy a határ mentén új vállalkozásokat is érdemes lenne indítani a car-sharing piacon, hiszen egy megfelelően felépített kampánnyal gazdaságilag is sikeres működést tudna elérni akár több szolgáltató is.

A már korábban ismertetett hivatalos, külső szolgáltató által nyújtott car-sharing konstrukciók közül véleményem szerint a free-floating típusú autómegosztási rendszer lenne a

legalkalmasabb a napi szinten ingázók, továbbá a falvakból a magyar kis-, közép- és nagyvállalatokhoz reggelente induló munkavállalók, valamint a városok lakóinak eredményes kiszolgálására. Az egyik legnagyobb problémát ebben az esetben is a személygépjárművek elhelyezkedése jelentené. A használók többsége reggel 6, illetve reggel 8 előtt indulna munkába többnyire falvakból vagy városi lakótelepekről. A munkaidő alatt sok autó az ipari parkok területén parkolna, ahogyan napjainkban is. Azoknak a felhasználóknak, akiknek napközben a belvárosban lenne szüksége gépjárműre nem jelentene reális alternatívát egy távoli ipari parkban parkoló bérelhető autó. Természetesen sokan dolgoznak a városon belül is, ami közlekedés szempontjából frekvenciált helyen van, de mindenképpen számolni kell a fent említett logisztikai gond megjelenésével, ha engedélyezni fogják a megosztott gépkocsik használatát az ipari parkok területén, valamint a falvakban is. A megoldást a teljesen autonóm gépjárművek elterjedése jelentené, mert, ezáltal könnyedén és folyamatosan biztosítható lenne a megosztott autóknek a városokon, régiókon, vármegyéken, tartományokon belüli optimális eloszlása, elhelyezkedése. Új lehetőségként akár háztól házig szolgáltatás is bevezethető lenne, ami növelné a rendszer népszerűségét, ezáltal vonzóbbá tenné a car-sharing rendszert. A free-floating típusú autómegosztás regionális szintű elindításával mérsékelhető lenne a forgalomban lévő gépkocsik száma, aminek a következményei egyértelműen láthatók lennének: kevesebb jármű → torlódások csökkenése → közúti közlekedési balesetek számának mérséklődése; kevesebb jármű → károsanyagkibocsátás csökkenése → jobb levegőminőség → élhetőbb városok. Az igazsághoz azonban hozzátartozik, hogy egy ilyen nagy területen akár szakaszosan – az egyes városokban, falvakban eltérő időpontban - elindított szolgáltatás számottevően nagyobb infrastrukturális beruházást igényelne a cégek részéről, továbbá a nagyvárosok belső kerületeiben üzemeltetett megoldásoknál lényegesen nagyobb gépjárműflottára lenne szükség, illetve szükséges lenne a flottában nélkülözhetetlen elektromos autók töltésére alkalmas töltőhálózat széleskörű kiépítése.

Az egyik legnagyobb kérdés egy ilyen rendszer tervének megvalósításakor, hogy a lakosság valóban használná-e. Vajon lemondanának a saját tulajdonú gépjármű birtoklásáról az emberek? Helyettesíteni tudná a szolgáltatás a magángépkocsikat? Úgy gondolom a car-sharing rendszer előnyeit sokan látják, valószínűleg a legtöbben ki is próbálnák, de vajon mennyien és mikor adnák el az autóikat egy megyék között is jól működő car-sharing rendszer esetén.

Véleményem szerint regionális szinten egy fokozatosan bővített rendszer kiépítésével lehet sikeres üzleti modellt létrehozni. A lefedettségi terület és a gépjárműflotta nem teljes körű

elindításával, nemcsak csökkenthetők a kezdeti beruházási költségek, hanem értékes tapasztalatok nyerhetők a felhasználói aktivitást, illetve a használati szokásokat illetően. Az induláskor befektetett összeg minimalizálásával gazdaságosabb működés érhető el, ezáltal az ügyfelektől kapott értékelések alapján módosíthatók az előzetes tervek és a felhasználók visszajelzései szerint fejleszthető tovább a szolgáltatás a következő években. Elképzelésem szerint Győr-Moson-Sopron vármegyében kezdetben Győrben, Mosonmagyaróváron és Sopronban lenne lehetőség egy gazdaságos car-sharing kialakítására. Napjainkban Budapesten 2 000 000 lakosra nagyjából 1 000 megosztott autó jut, de a számuk folyamatosan növekszik, mert egyre többen ismerik meg és veszik igénybe ezt a szolgáltatást. A fenti adat alapján a rendszer indulásakor Győrben 100, Mosonmagyaróváron 20, Sopronban 35 – 40 bérelhető gépjárművet kellene üzembe állítani kisebb lefedettségi területtel, majd folyamatosan növelve a városok határáig, beleértve a lakótelepeket, külvárosokat és az ipari parkokat is. Az ingázók igényei miatt Vas vármegyében is esélyes lehet sikeres autómegosztást létrehozni egy vállalkozás által. A lakosság számának figyelembevételével a jelenlegi tapasztalatok szerint Szombathelyen 40 – 50 megosztott autó gazdaságos működtetése a realitás. A szolgáltatás indulását követően a visszajelzések, minősítések kiértékelése után, illetve az anyagi szempontokat figyelembe véve lassú és óvatos ütemben növelhető a lefedettségi terület. A Sopronhoz és Szombathelyhez közeli falvaknak a rendszerbe integrálása – a „car-sharing-zóna” kiterjesztése - a következő lépés. Meglátásom szerint a free-floating autómegosztás jellegéből adódóan a „zónákat” csak városokban, falvakban, illetve nevezetességek, gyárak és ipari parkok közelében kellene és szabadna kialakítani. A továbbiakban a személygépjármű állomány növelésével lenne biztosítható az ügyfelek magas szintű kiszolgálása. A felhasználók elégedettsége az elsődleges szempont. Pozitív értékelés, valamint nyereséges üzemelés esetén érdemes lenne további magyar falvakat, helységeket is bevonni a szolgáltatásba. A percdíjak, napidíjak meghatározásánál a Budapesten működő szolgáltatók konstrukcióinak alkalmazása lehet a legoptimálisabb és leggazdaságosabb. A kihasználtság növelésével érhetőek el olcsóbb díjak. A car-sharing piacra minden bizonnyal nagy hatással lesz az önvezető személygépkocsik megjelenése és elterjedése. A fent említett modell is az autonóm járművek forgalomba állításával lesz teljes. Egy teljes régióra akkor terjeszthető ki a lefedettségi terület, ha az autók gyakran kedvezőtlen elhelyezéséből adódó logisztikai problémákra valós megoldás kerül kidolgozásra és ez minden valószínűség szerint az önvezető technológia bevezetése lesz.

4. Összegzés

A car-sharing szolgáltatások magyarországi és globális megjelenésével olyan új mobilitási alternatíva van kialakulóban, amely nagy hatással lesz az autógyártókra és a társadalomra egyaránt. A különféle lehetőségek közül nagy valószínűséggel az emberek szubjektív döntése fogja meghatározni, hogy az egyes országokban, városokban melyik autómegosztási konstrukció terjed el és válik igazán népszerűvé. A választást sok tényező befolyásolja, mint például a települések szerkezete, forgalom mértéke, az otthonok és munkahelyek közötti távolság, de hatást gyakorol a döntésre a bérelhető gépjárművek minősége, állapota, felszereltsége is. A car-sharing széleskörű elterjedésének azonban a várakozások és elemzések szerint számtalan előnye lesz. A megosztott autók használatával csökkenhet a forgalomban lévő gépkocsik száma, ezáltal mérséklődhetnek az országutakon és városokban rendszeresen kialakuló forgalmi torlódások. Az utcákon lévő kevesebb járműnek köszönhetően csökkenhet a közúti közlekedés által kibocsátott károsanyagok mennyisége, ezáltal javulhat a települések levegőminősége, így jobb és egészségesebb életfeltételek biztosíthatók az állampolgárok részére. A torlódások mérséklődésével a közlekedésbiztonság javulása várható, hiszen a statisztikai adatok szerint a közúti közlekedési balesetek 90 százaléka az emberi hibára – türelmetlenségre – vezethető vissza.

Jelen cikkben egy regionális car-sharing szolgáltatás koncepciója került kidolgozásra, melynél az igényeknek való megfelelés mellett akár gazdaságos működés is elérhető lenne. A rendszer kialakításánál azonban véleményem szerint az egyik legfontosabb szempont a lépésről-lépésre haladás elve, egy fokozatosan növekvő gépjárműflottával és lefedettségi területtel rendelkező szolgáltató tudna csak pénzügyileg stabil üzemeltetést megvalósítani. A jelenlegi tendenciák alapján a flottában elengedhetetlen lenne a teljesen elektromos meghajtású személyautók jelenléte, azonban – a töltőhálózat teljes körű kiépítéséig – csak a városokon belüli használatukat kellene engedélyezni. A villanyhajtású gépjárművek számának növekedésével szükséges lesz további gyorstöltők telepítése a városokban, illetve az autópályák és országutak mellett Győr-Moson-Sopron vármegyében.

Egy regionális car-sharing rendszer bevezetésével, ahol a flottában teljesen elektromos járművek is az ügyfelek rendelkezésére állnak, nélkülözhetetlen lesz további gyorstöltők telepítése Győr-Moson-Sopron vármegyében. A lassú töltés esetén a gépkocsik több órára is kiesnek a flottából, ami nemcsak csökkenti a felhasználók által szabadon kibérelhető járművek számát, hanem a szolgáltatóknak bevételkiesést is okoz, miközben a töltésért fizetniük kell. Egy

autómegosztási szolgáltatást nyújtó vállalat flottájában lévő autók csak a bérleti időszak alatt biztosítanak bevételt, ezért többek között a cégeknek is gondoskodniuk kell a megfelelő infrastruktúra kiépítéséről. Az autópályák mentén nagyobb igény lenne a villámtöltésre, azonban jelenleg Magyarországon a többi országhoz képest kevés a villámtöltők száma. Az optimális megoldás az lenne, ha minél több autópálya menti pihenőben alakítanának ki villámtöltőket. Győr-Moson-Sopron vármegyében az M1-es autópályán jobb a helyzet más magyarországi autópályákhoz képest, de itt is további pihenőhelyeken kellene töltőket létesíteni. Az M85-86-os gyorsforgalmi úton jelenleg nincs lehetőség töltésre.

Javaslatom szerint egy regionális car-sharing rendszer létrehozásánál kezdetben olyan korlátozást kellene bevezetni az elektromos gépjárművek használatára vonatkozóan, hogy a teljesen elektromos meghajtású gépkocsikat csak nagyobb városokban vehetnék igénybe a felhasználók. A nagy távolságokat pedig hibrid vagy benzines autók szolgálnák ki. A megoldásnak köszönhetően a városok levegőminősége javulna az elektromos autóknak és az autómegosztásnak köszönhetően, továbbá a fő közlekedési útvonalak zsúfoltsága is mérséklődne a teljes vármegyében elérhető car-sharing következményeként. A későbbiekben, a megfelelő infrastruktúra megvalósítását követően kiterjeszthető a regionális szolgáltatás a teljes flottára, azonban az elektromos autók használata a városon belül a leggazdaságosabb, hiszen a nem megújuló energiával való töltés, illetve a lítium újrahasznosításának megoldatlansága miatt a nagyméretű akkumulátorok beépítésének jelentős hátrányai vannak. Az energia előállítása és az akkumulátorok gyártása, újrahasznosítása-megsemmisítése során keletkező plusz károsanyag-kibocsátást is figyelembe véve nem elhanyagolható a nagy hatótávolságú teljesen elektromos gépjárművek környezeti terhelése globális nézőpontból.

5. Köszönetnyilvánítás

A publikációban szereplő kutatást a Széchenyi István Egyetem az Európai Unió támogatásával valósította meg, az RRF-2.3.1-21-2022-00004 azonosítójú, Mesterséges Intelligencia Nemzeti Laboratórium projekt keretében.

Irodalomjegyzék

- [1.] Európai Bizottság közlemény: A mindenkit megillető tiszta, versenyképes és összekapcsolt mobilitás felé való, társadalmilag igazságos átmenet programja, Kiadás: Brüsszel, 2017.05.31. COM (2017) 283 final
- [2.] Európai Községek Bizottsága: Cselekvési terv az intelligens közlekedési rendszerek alkalmazásának európai bevezetésére, Kiadás: Brüsszel, 2008.12.16. SEC (2008) 3084
- [3.] IPSOS Publications: Du lundi au vendredi, les Européens passent en moyenne 9h35 a se déplacer, 2017
- [4.] Európai Statisztikai Kézikönyv, EU Transport, 216
- [5.] KSH: Közlekedési balesetek
- [6.] Bännström, M., Eugensson, A., Frasher, D., Robertsson, A., Rothoff, M., Solyom, S.: Environmental, Safety, Legal and Societal Implications of Autonomous Driving Systems, 2013, Paper number: 13-0467
- [7.] Somfalvi Aliz: Mol Limo értékesítési és marketing vezetőjével interjú
- [8.] Frost and Sullivan research: megduplázódhat a közösségi autót használók száma, 2019
- [9.] Cohen, A., Jaffee, M., Shaheen, S.: Innovative mobility – carsharing outlook, 2019
- [10.] Correia, G., Jorge, D.: Carsharing systems demand estimation and defined operations – a literature review, 2013
- [11.] KSH: Helyzetkép a szállítási ágazatról, 2017
- [12.] GreenGo - interjú Michaletzky Bálinttal

Közúti Terelőbóják Kulcspont-alapú Detektálása

Keypoint-Based Detection of Road Traffic Cones

Hollósi János^a

^aSzéchenyi István Egyetem, Járműipari Kutatóközpont

hollosi.janos@sze.hu

Absztrakt

A közúti terelőbóják pontos és megbízható vizuális detektálása kulcsfontosságú mind az autonóm járműrendszerek biztonsága, mind a közúti fenntartási tevékenységek hatékonysága szempontjából. Jelen munka egy hosszú távú kutatási és fejlesztési projekt részét képezi, amelynek végső célja egy olyan autonóm robotplatform megalkotása, amely képes a terelőbóják önálló kihelyezésére és összegyűjtésére. Ebben a munkában különböző kulcspont alapú detektáló neurális hálózatok alkalmazása történik annak érdekében, hogy a bóják detektálása mellett a csúcspontjuk, mint kulcspont pontos észlelése is megvalósításra kerüljön. Ezzel egy nem megszokott felhasználási lehetőségét bemutatva a kulcspont-detektáló megoldásoknak. A kiegészítő kulcspont-információ lehetővé teszi a bóják térbeli helyzetének pontosabb becslését, ami nélkülözhetetlen a megvalósítás alatt álló robotplatform tekintetében. A kutatás során megkonstruált tanító-adathalmaz széles körű varianciát mutat környezeti és időjárási viszonyok terén, így biztosítva a hálózatok széles körű alkalmazhatóságát. Az eredmények rávilágítanak arra, hogy a kulcspont-alapú megközelítés integrációja fontos lépés az autonóm kiszolgáló robotplatform fejlesztésében, amely a jövőben jelentősen hozzájárulhat a közúti biztonság és hatékonyság növeléséhez.

Kulcsszavak: kulcspont-detektálás; terelőbója lokalizáció; neurális hálózatok; autonóm robotrendszerek; számítógépes látás a közlekedésben

Abstract

Accurate and reliable visual detection of road traffic cones is essential for both the safety of autonomous vehicle systems and the efficiency of road maintenance operations. This work forms part of a long-term research and development project aimed at creating an autonomous robotic platform capable of independently deploying and collecting traffic cones. In this study, various keypoint-based neural networks are employed not only to detect cones but also to precisely localize their apex as a keypoint. This approach presents an unconventional application of keypoint detection techniques. The additional keypoint information enables more accurate spatial localization of the cones, which is critical for the functionality of the robotic platform under development. The custom training dataset constructed for this research includes a wide range of environmental and weather conditions, ensuring the robustness and broad applicability of the models. The results highlight that integrating a keypoint-based approach is a significant step toward the development of autonomous service robots, which may considerably enhance future road safety and operational efficiency.

Keywords: keypoint detection; traffic cone localization; neural networks; autonomous robotic systems; computer vision in transportation

1. Bevezetés

A közúti terelőbóják pontos és megbízható vizuális-alapú detektálása kulcsfontosságú mind az autonóm járművek biztonságában, mind a közúti fenntartási feladatok automatizálásában. Jelen munka egy hosszú távú kutatás-fejlesztési projekt része, amelynek célja egy saját fejlesztésű és autonóm kiszolgáló robotplatform létrehozása, amely képes, többek között, közúti terelőbóják önálló kihelyezésére és begyűjtésére ipari vagy útfenntartási környezetekben. De a hosszútávú cél egy olyan általános robotplatform megalkotása, amely képes több különböző, gyakorta felmerülő feladat végrehajtására.

A gépi látás tekintetében a hagyományos objektumdetektáló módszerek elsősorban a befoglaló téglalapokra koncentrálnak. Ezzel szemben a kulcspon-alapú detektálás egyre nagyobb szerephez jut ott, ahol precíz részlet-azonosítás szükséges, például robotmanipuláció során a megfogási pontok meghatározása érdekében [1,2]. Ebben a kutatásban a terelőbója tetőpontjának (fogantyújának) kulcsponként történő detektálása lehetővé teszi a robotkar számára a stabil és pontos megfogást. Ez a nem hagyományos alkalmazási eset jól illeszkedik a fent említett kulcspon-alapú paradigma kiterjesztéséhez. Ilyen módon alátámasztva, hogy a kulcspon-alapú detektálási módszerek a legkülönbözőbb feladatok esetén is nagy hatásfokkal alkalmazhatók. Korábbi munkák a kiszolgáló robotplatformok területén főként beltéri alkalmazásokra fókuszáltak, de a kültéri, félig strukturált környezet – mint például telephelyek vagy útfelújítások – további vizuális és működési kihívásokat hoznak [3]. A bójadetektálás területén a forgalomirányításban használt markerek az autonóm járművek számára rendszerint LiDAR-alapú vagy szegmentációs módszerekkel kerülnek érzékelésre [4,5]. Jelen kutatás egy saját, különféle környezeti és időjárási körülmények között felvett adatkészletre épít, ahol beltéri és kültéri, felülnézeti képek szerepelnek, mellyel a cél annak vizsgálata is, hogy a népszerű YOLO hálózati családok milyen hatásfokkal alkalmazhatók ilyen speciális részfeladat esetén. Bár a YOLO-család modelljei széles körben bizonyítottak és a népszerű benchmark adatkészleteken – mint például a COCO [6] – rendre azt mutatják, hogy az újabb vagy nagyobb paraméterszámú változatok várhatóan jobb teljesítményt nyújtanak, ez a tendencia nem minden feladattípusnál igazolódik. Több kutatás is rávilágít arra, hogy a YOLO modellek érzékenyek az osztály-eloszlásra, a környezeti szélsőségekre és a ritkábban előforduló objektumokra. Például kimutatták, hogy class imbalance esetén a YOLOv5 esetében a mintavételezés vagy osztálysúlyozás nem javította, hanem sokszor rontotta a mAP értékeket [7]. Hasonlóképpen, extrém időjárási körülmények – mint az erős eső vagy köd – szimulált környezetben

drasztikusan csökkentették a YOLO teljesítményét, függetlenül a modell méretétől [8]. Továbbá kimutatták, hogy ha az objektumarány nagyon alacsony (pl. kis méretű és ritkán előforduló tárgyak), akkor sem a nagyobb modellek, sem a nagyobb adatmennyiség nem feltétlenül biztosítanak jelentős előrelépést a teljesítményben [9]. Egy másik vizsgálat szerint még a legújabb YOLO-változatok is nehézségekbe ütköznek finom részletek vagy részben takart objektumok felismerésében [10]. Mindez arra utal, hogy a YOLO-család fejlettsége ellenére sem garantált a folyamatos teljesítménynövekedés minden környezetben és feladattípusban, így különösen érdekes vizsgálati irány, hogy az egyes modellcsaládok miként teljesítenek olyan komplex detektálási feladatban, mint a közúti terelőbóják kulcspont-alapú felismerése. A kutatás végcélja, hogy kulcspont-alapú detektálással és a YOLO hálózat-család [11,12] különféle változataival egy robusztus, valós idejű bójalokalizáció kerüljön megvalósításra, amely megfelelő ahhoz, hogy integrálható legyen az autonóm kiszolgáló robotplatformba.

A tanulmány felépítése a következő: A 2. fejezetben bemutatásra kerül a kiszolgáló robotplatform, majd a 3. fejezetben a gépi látás alrendszerben jelenleg célként kezelt, a kutatás tárgyát képező, terelőbója felismeréshez szükséges egyedi adatkészlet tárgyalása történik. A 4. fejezetben a felhasznált neurális hálózat modellek prezentálása követ, ami után a felhasznált modellek és adatkészlet segítségével elért eredmények összegzése olvasható az 5. fejezetben, végül a 6. fejezetben az összefoglalással és kitekintéssel zárul a tanulmány.

2. Robotplatform

A hosszútávú kutatási és fejlesztési cél egy olyan autonóm mobil robotplatform megvalósítása, amely alkalmas közúti terelőbóják önálló kihelyezésére és begyűjtésére ipari, teszt pályás vagy más zárt, félig strukturált környezetben. A platform jelenlegi állapotát az 1. ábra szemlélteti, ahol felügyelet mellett, de valós munka közben látható a jármű [13]. Az alkalmazás során egyaránt követelmény a műveletek nagyfokú ismétlődése, a precíz pozicionálás, valamint a fokozott biztonsági igények kielégítése – különösen forgalomtechnikai beavatkozások vagy járműtesztelési helyszínek esetén. A robotikai rendszer egy nagy teherbírású, négykerék-meghajtású, skid-steer manőverezésű mobil platformra épül, amely képes akár 30 darab bója (kb. 150 kg össztömeg) szállítására. A navigáció GNSS RTK-alapú pozicionálással valósul meg, amely centiméteres pontosságot biztosít. A robotkar feladata a bóják lehelyezése és

begyűjtése az előre megadott koordináták alapján. Ehhez egy kollaboratív robotkar került integrálásra, melyet egy egyedi fejlesztésű elektromos megfogó (gripper) egészít ki. A platform a biztonságos működés érdekében lézerszkennerral, a környezet érzékelésére pedig többkamerás gépi látás rendszerrel van felszerelve. A rendszer működését ROS 2 alapú szoftverarchitektúra vezérli, amely két különálló fedélzeti számítógépre épül. Ezek közül az egyik a gépi látás és mesterséges intelligencia modulokat szolgálja ki, míg a másik a központi irányítást látja el, beleértve a mozgástervezést és a karvezérlést is. A vizuális észlelésért felelős neurális hálózatok futtatásához a hardver alapját egy NVIDIA® Jetson AGX Xavier™ platformra épülő Connect Tech Rudi-AGX egység képezi. A platform tervezése és megvalósítása során külön figyelem irányult a jövőbeli bővíthetőségre és más kiszolgáló jellegű feladatok végrehajtásának lehetőségére is.

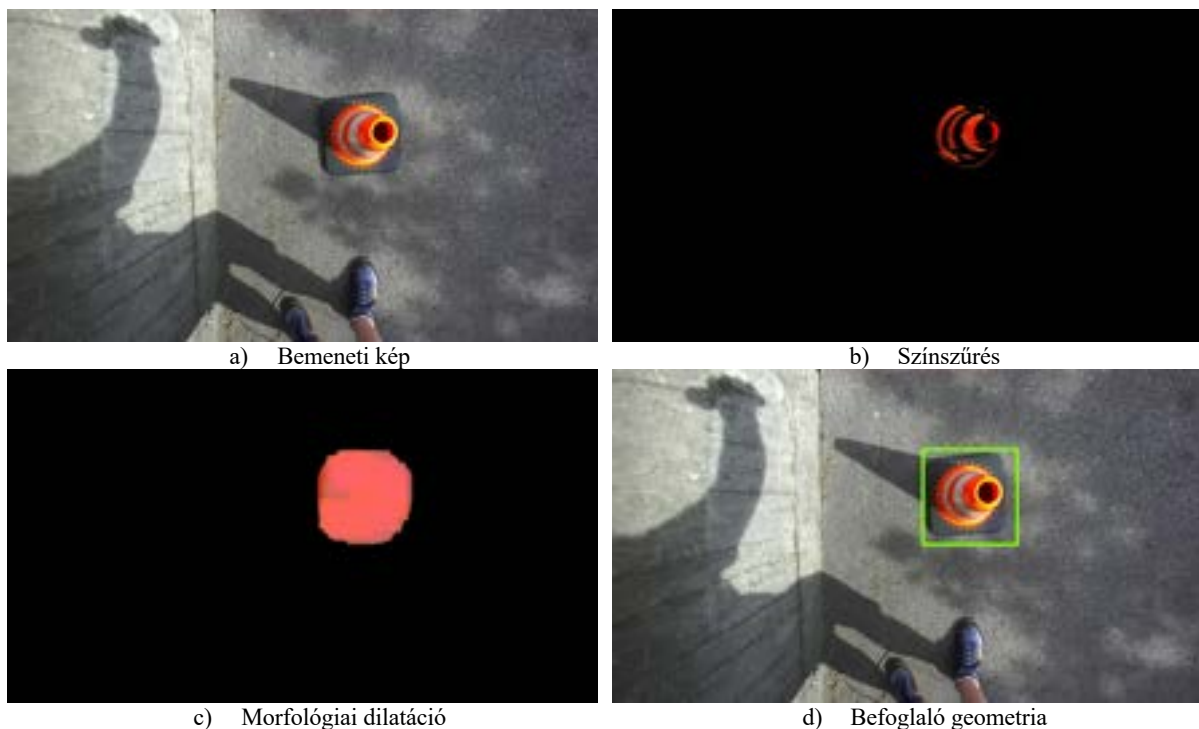


1. ábra A megvalósítás alatt álló autonóm robotplatform

3. Adatkészlet

A közúti terelőbóják automatizált észleléséhez nem állt rendelkezésre sem nyilvánosan elérhető, sem célzottan a feladathoz illeszkedő képi adatkészlet. A meglévő objektumdetektáló benchmark-adatbázisok (pl. COCO [6], KITTI [14], BDD100K [15]) nem tartalmazzak elegendő számú, felülnézetből készült, különféle környezeti viszonyokat reprezentáló bójaképet. Ezért a platform működéséhez és a neurális hálózatok tanításához elengedhetlenné vált egy saját, dedikált adatkészlet összeállítása.

Az adatrögzítés a fejlesztett robotplatform fedélzeti, Stereolabs ZED 2i típusú kamerarendszerével történt, valós beltéri és kültéri környezetekben. A felvételek készítése során cél volt a környezeti variancia maximalizálása, különböző megvilágítási körülmények (napfény, árnyék, eső, mesterséges világítás), háttérstruktúrák (aszfalt, fű, felfestések, útpadka) és kameraállások alkalmazásával. Az így létrejött adatkészlet főként felülnézeti és közel felülnézeti perspektívából ábrázolja a bójákat, amely a robotkar célzott manipulációjához szükséges. Az elkészült adatkészlet 1225 darab annotált tanítómintát és 297 darab validációs mintát tartalmaz, amelyekhez minden esetben rendelkezésre áll a bóját leíró befoglaló téglalap (bounding box), valamint a bója csúcspontja (fogantyú) mint kulcspont koordináta.

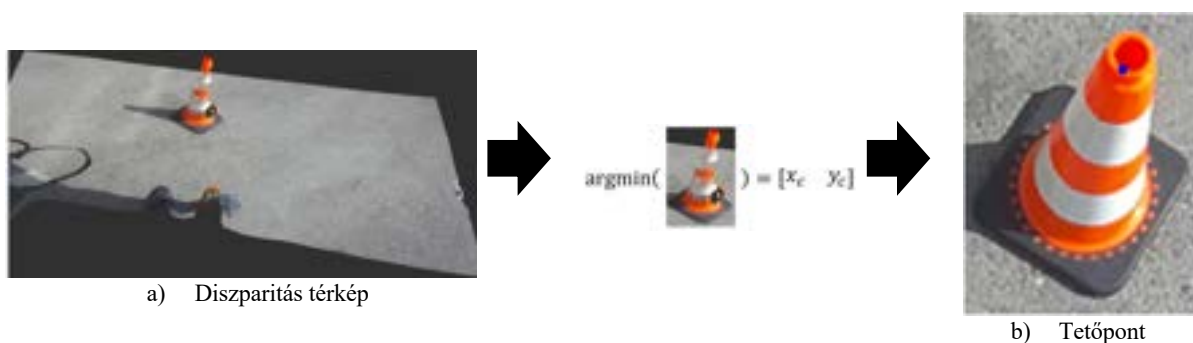


2. ábra A terelőbója képfeldolgozás-alapú szűrése

Az annotálás háromlépcsős, félig automatizált folyamat eredményeként készült el. A bóják kezdeti lokalizálása a képsíkon színszűréssel és kontúrkeresési eljárásokon [16] alapult. Minden felvételhez empirikus úton optimalizált színtartomány-szűrési hiperparaméterek kerültek alkalmazásra, figyelembe véve az adott környezeti viszonyokat (pl. fényviszonyok, háttérkontraszt). Mindezt annak érdekében, hogy a szűrési eljárás minél pontosabban meghatározza a terelőkúp színösszetétele alapján a képsík bóját tartalmazó részterületét. Az ilyen módon keletkezett bináris képen több folt formájában jelennek meg azok a területek, ahol

a bója részei találhatók. Hogy ezek egybefüggő foltot alkossanak morfológiai műveletek alkalmazása történt meg, ahol a foltok finom duzzasztásával egybefüggő terület képződik belőlük. Az ilyen módon keletkezett bináris képen a bója leírásához kontúrdetektálási eljárás [16] felhasználása történt meg, amelynek eredményeként megszületik a bóját befoglaló geometria leírása. Ezt a folyamatot foglalja össze egy példán keresztül a 2. ábra.

Miután a bója befoglaló geometriája meghatározásra került, a Stereolabs ZED 2i kamera által szolgáltatott diszparitás térkép alapján a bójához tartozó területen belül az adott legmagasabb (kamerához legközelebbi) pont került kiválasztásra. Ez jellemzően a bója peremének (karimájának) legközelebbi pontja és nem a bója tetejének tényleges középpontja. Ennek oka, hogy a bója felül lyukas, belül üreges, ezért mindenképp a karimáján fog legközelebbi pontot találni. Erre mutat egy esetet a 3. ábra.



3. ábra Diszparitás térkép alapján meghatározott tetőpont

Ahogy a 3. ábra is jól szemlélteti, ez a módszer önmagában nem képes a bójatető középpontjának meghatározására. Ezért a Hough-transzformáció [17] segítségével körkeresési módszer alapján lett meghatározva a karima, aminek már egyszerűen meghatározható a középpontja. A transzformáció segítségével a színszűrési módszerrel kijelölt területen belül a kör-szerű alakzatok detektálása történt meg. A detektált körök közül az került kiválasztásra, amelynek a legkisebb a távolsága a diszparitás térkép alapján kijelölt tetőponttal. Ennek a körnek a középpontja lett az adott bójához tartozó csúcspont, ami a kialakított adatkészlet alapját képezi. Ezt a módszert mutatja be a 4. ábra.



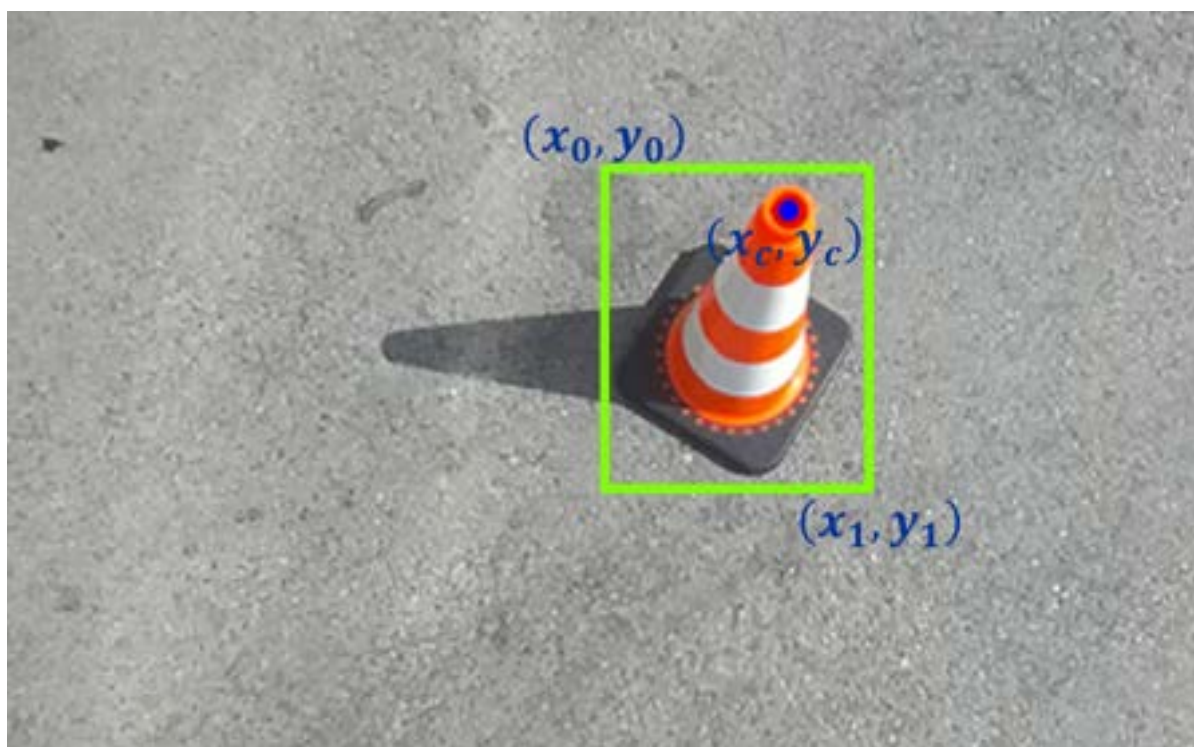
a) Detektált kör alakzatok



b) Bója tetőpontja

4. ábra Bója tetőpontjának pontosítása Hough-transzformációval

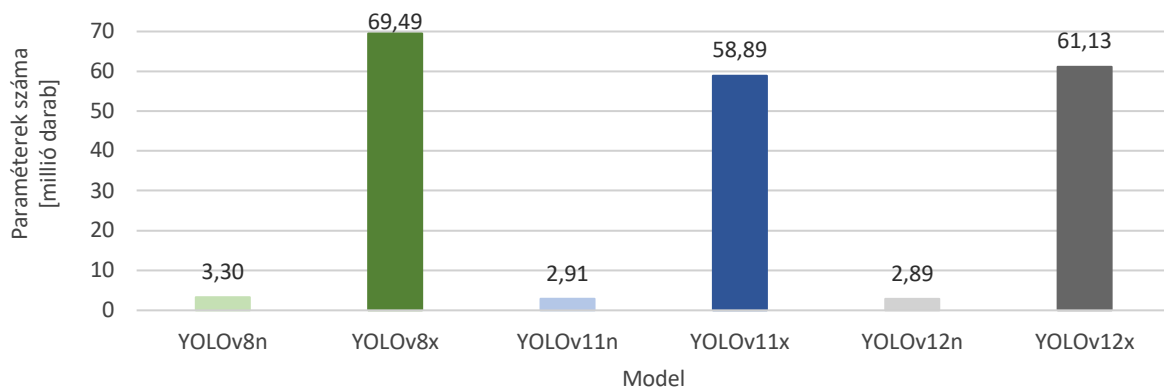
Ezzel a módszerrel létrejött egy olyan adatkészlet, ami összesen 1522 darab mintát tartalmaz. Mindegyik mintához tartozik egy színes kép, ami valamilyen kül- vagy beltéri helyszínen készült. Ezen felül egy szám-hatos, amelyet a bóját befoglaló téglalap bal felső (x_0, y_0) és jobb alsó (x_1, y_1) képsík-beli sarokpontjai, valamint a tetőpontot jelképező pont (x_c, y_c) koordinátája alkot. Ezt szemlélteti az 5. ábra is.



5. ábra Egy minta az egyedi fejlesztésű adatkészletből

4. Modellek

A bójadetektálási feladat megoldásához olyan korszerű objektumdetektáló neurális hálózatokat alkalmazása szükséges, amelyek képesek nemcsak a tárgyak befoglaló téglalapjának meghatározására, hanem azok kulcsponjtjainak regressziójára is. A modellválasztás alapját a YOLO (You Only Look Once) [11][12] hálózatsalád legújabb generációi képezték, mivel ezek az architektúrák ötvözik a valós idejű feldolgozhatóságot, a kiváló észlelési teljesítményt, valamint az egyszerűen bővíthető kimeneti struktúrát. A kutatás során a YOLOv8 [18], YOLOv11 [19] és YOLOv12 [20][21] hálózatok különböző méretváltozatai kerültek összehasonlításra, mindegyik esetben a „nano” (röviden: „n”) és az „extra-large” (röviden: „x”) konfigurációk felhasználásával. Ezeknek a komplexitását mutatja a 6. ábra.



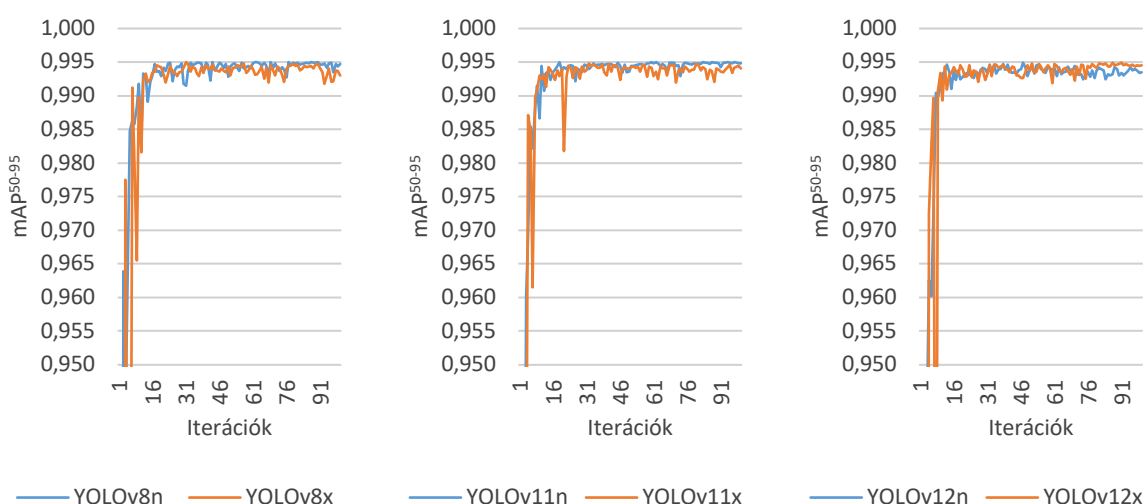
6. ábra A kutatásban felhasznált modellek komplexitása

A választott hálózatok mindegyike kiegészült egy speciális kimeneti résszel, amely a terelőbóják csúcspontját, mint kulcsponjt regressziós módon becsüli meg. A hálózatok tanítása során a befoglaló geometria (bounding box) detektálására az általánosan alkalmazott IoU-alapú (Intersection over Union) veszteségfüggvény, míg a kulcspontpozíciókhoz L1 veszteség alkalmazása történt. A YOLOv8 egyszerűsége és gyors taníthatósága miatt került bevonásra a vizsgálatba, míg a YOLOv11 a hálózaton belüli információ-fúzió továbbfejlesztett változatát képviseli. A legújabb, YOLOv12 generáció már olyan fejlesztéseket is integrál, mint az Area Attention és a FlashAttention2-alapú gyorsított figyelemmechanizmus, amelyek célja a detekciós pontosság további növelése, különösen összetett vizuális jelenetek esetén. A modellek kiválasztásánál szempont volt a robotplatform által alkalmazott hardveres környezet is, így minden modell optimalizálva lett a Jetson AGX Xavier alapú Connect Tech Rudi-AGX rendszerre történő futtatásra. Az így kialakított modellstruktúrák lehetővé teszik a különböző

erőforrásigényű alkalmazások vizsgálatát, az egyszerű beágyazott megoldásoktól egészen a nagyobb számítási kapacitást igénylő, nagy pontosságú detektálási feladatokig.

5. Eredmények

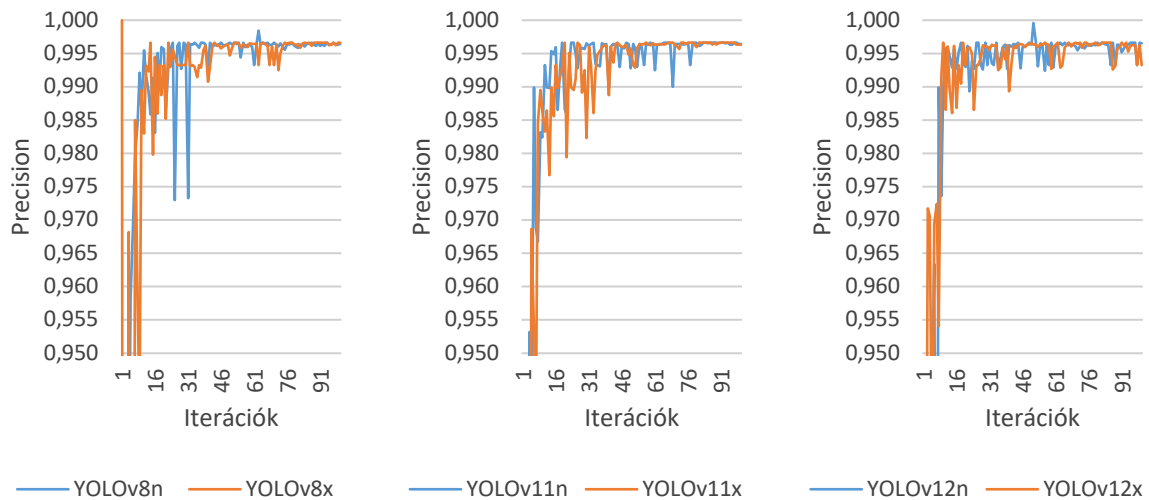
A kiválasztott modellek teljesítményének kiértékelése a saját fejlesztésű, kulcspont-információt is tartalmazó adatkészleten történt. A tanítás során minden modell 100 epoch-on keresztül került optimalizálásra, egységes tanító és validációs adathalmazon. A YOLOv8 esetében Stochastic Gradient Descent (SGD) [22][23], míg a YOLOv11 és YOLOv12 esetén AdamW [24] optimalizáló algoritmus került alkalmazásra. A hálózatok veszteségfüggvénye két komponensből állt: az objektumokat leíró befoglaló téglalapokra vonatkozó IoU-alapú veszteségből, valamint a kulcspont-regresszióhoz alkalmazott L1 veszteségből.



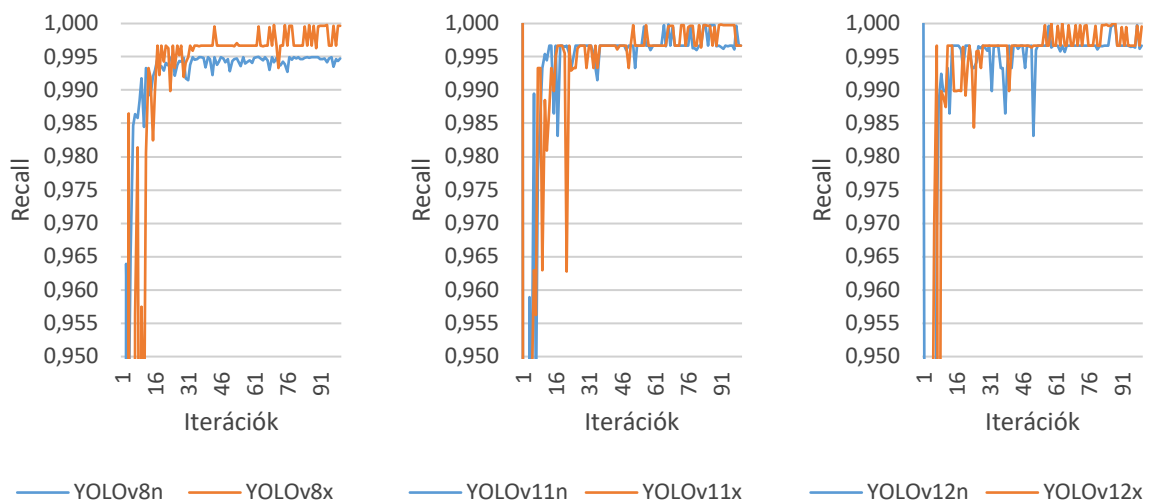
7. ábra Az egyes YOLO modellek mAP₅₀₋₉₅ tanulási görbéje

Az eredmények értékeléséhez három, az objektumdetektáló hálózatok teljesítményének mérésében széles körben alkalmazott mutató került felhasználásra. A legátfogóbb mérőszám az mAP⁵⁰⁻⁹⁵ (átlagpontosság), amely a mean Average Precision értékét jelenti több, 50 és 95% közötti IoU-küszöbön (Intersection over Union) átlagolva, így egyszerre tükrözi a lokalizáció és az osztályozás pontosságát. Ezt egészíti ki a Precision (pontosság) mutató, amely azt jelzi, hogy a detektált objektumok közül milyen arányban voltak helyes találatok, valamint a Recall (érzékenység), amely a valóban jelenlévő objektumok felismerésének arányát méri. A három metrika együttes vizsgálata lehetővé teszi, hogy a modellek teljesítményét átfogó módon

értékeljük: míg a Precision a téves riasztások minimalizálására, addig a Recall a felismerések teljességére világít rá, a mAP^{50-95} pedig a kettőt összehangoltan fejezi ki.



8. ábra Az egyes YOLO modellek Precision tanulási görbéje

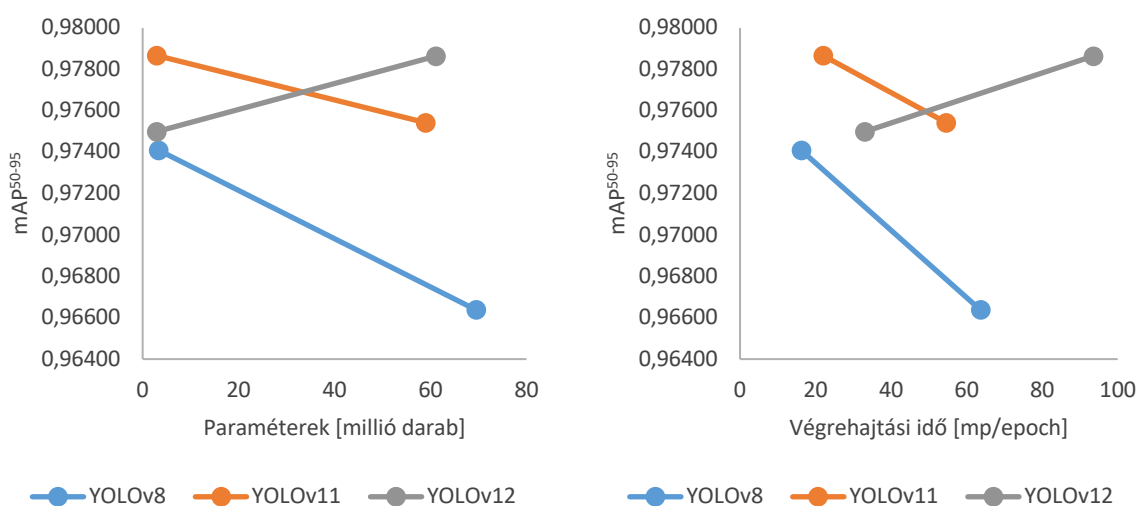


9. ábra Az egyes YOLO modellek Recall tanulási görbéje

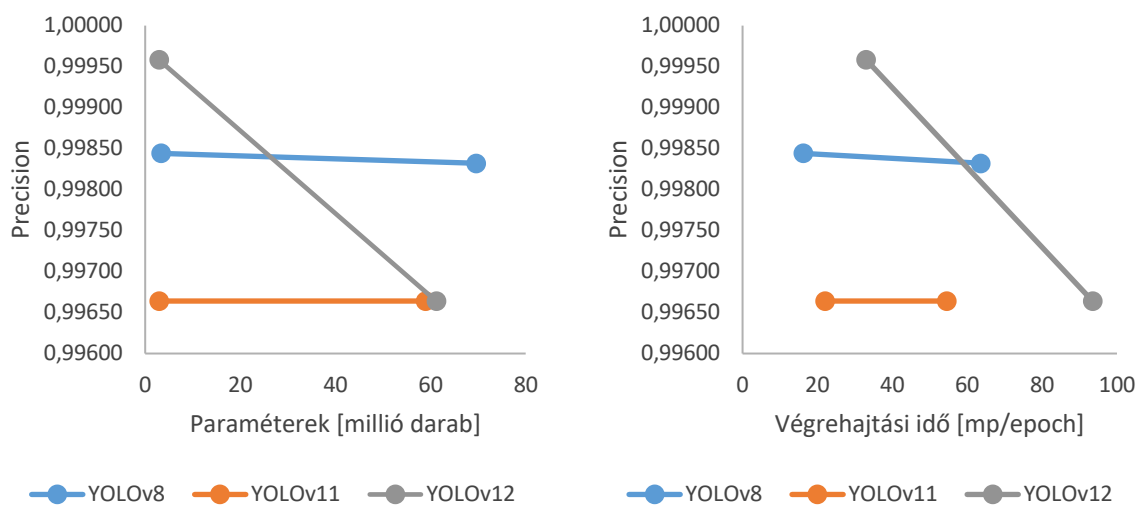
A 7-9. ábrákon az egyes modellek mAP^{50-95} , Precision és Recall metrikák szerinti értékei láthatók a tanulási folyamat során. Mindhárom modelleszám gyors konvergenciát mutatott, a kezdeti iterációk után stabilan 0,99 feletti mAP^{50-95} értékeket érve el. A YOLOv11n teljesítménye bizonyult a legmagasabbnak, ugyanakkor a YOLOv8 és YOLOv11 esetében a kisebb modellek (YOLOv8n és YOLOv11n) stabilabban teljesítettek, mint a nagyobbak (YOLOv8x és YOLOv11x), amelyeknél kisebb visszaesések figyelhetők meg. Ez jól mutatja, hogy a paraméterszám növelése nem minden esetben eredményez jobb átlagpontosságot. A

Precision mutatók szintén nagyon magas értékeket értek el (0,996 felett), ugyanakkor a különbségek kisebbek voltak, mint az mAP^{50-95} esetében. A YOLOv8 és YOLOv11 modellek kiegyensúlyozott és stabil teljesítményt mutattak. A YOLOv12 esetében a kisebb változat (YOLOv12n) kiemelkedően magas pontosságot produkált, míg a nagyobb (YOLOv12x) kissé alacsonyabb értékeket ért el, ami túlilleszkedés vagy a hálózat instabilabb konvergenciájának jele lehet. Az összes modell rendkívül magas Recall értékeket mutatott (jellemzően 0,995 felett), ami azt jelzi, hogy a valóban jelenlévő objektumok szinte kivétel nélkül felismerésre kerültek. A YOLOv12x modell teljesítménye itt is kiemelkedett, stabilan a legmagasabb értékeket hozva. Ezzel szemben a YOLOv8n teljesítménye enyhén visszaesett a Recall tekintetében, bár ez a gyakorlatban elhanyagolható mértékű.

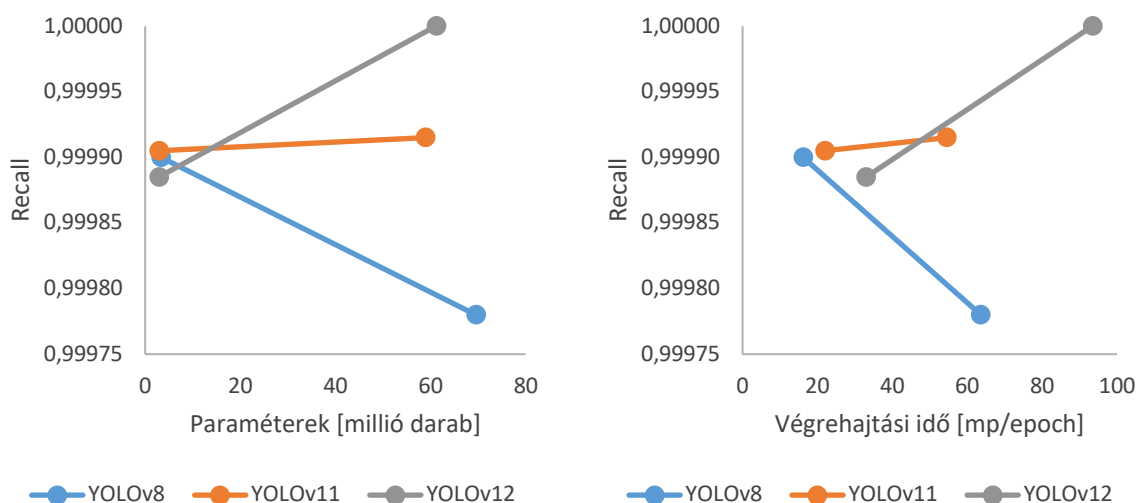
A hálózatok végső teljesítményének értékelésekor nem elegendő pusztán a pontossági mutatók vizsgálata, hanem célszerű azokat más szempontok tükrében is értékelni. Két ilyen, kiemelten fontos mérőszám a tanítható paraméterek száma és a végrehajtási idő. A tanítható paraméterek száma már bemutatásra került a 6. ábrán, ahol az egyes modellváltozatok paramétereinek nagyságrendje látható. Emellett a végrehajtási idő is meghatározó szerepet játszik, hiszen megmutatja, hogy a hálózat mennyi idő alatt képes feldolgozni a bemeneti adatokat.



10. ábra A YOLO modellek mAP^{50-95} értéke a paraméterszám és a végrehajtási idő függvényében



11. ábra A YOLO modellek Precision értéke a paraméterszám és a végrehajtási idő függvényében



12. ábra A YOLO modellek Recall értéke a paraméterszám és a végrehajtási idő függvényében

Jelen vizsgálatban nem az egyetlen képre jutó végrehajtási időt, hanem egy teljes epoch feldolgozási ideje lett figyelembe véve, amely az egész adatkészlet feldolgozását reprezentálja. Ennek jelentősége különösen abban áll, hogy a kifejlesztett rendszert valós időben, korlátozott számítási teljesítményű beágyazott hardveren kívánjuk alkalmazni, ahol a feldolgozási sebesség a pontosság mellett kritikus tényező. Az egyes hálózatok különféle hatékonysági mutatók szerinti végső eredményeit a paraméterek számának és a végrehajtási időnek a függvényében a 10-12. ábrák foglalják össze.

Az mAP^{50-95} értékek alapján a YOLOv8 és YOLOv11 esetén a nagyobb modellváltozatok teljesítménye romlott a kisebbhez képest, mégpedig hasonló mértékben, míg a YOLOv12 esetében mutatkozott csak előny a paraméterszám növelésével. A kisebb méretű modellek paraméterszáma közel azonos, ugyanakkor teljesítményük eltért: a YOLOv11n nyújtotta a legjobb eredményt, miközben a YOLOv8n és YOLOv12n kissé elmaradt ettől. Végrehajtási idő tekintetében a YOLOv8n csak minimálisan bizonyult gyorsabbnak, mint a pontosabb eredményt adó YOLOv11n. A Precision mutatók esetén szintén érdekes különbségek figyelhetők meg. A kisebb modellek közül a YOLOv12n teljesítménye volt a legjobb, ettől csak kevéssel maradt el a YOLOv8n, míg a YOLOv11n láthatóan gyengébben szerepelt a másik két megoldáshoz képest. A nagyobb méretű modellek összevetésében a YOLOv8 és YOLOv11 közötti különbségek lényegében nem változtak, azonban a YOLOv12 teljesítménye jelentősen visszaesett, egészen a YOLOv11 szintjére. Ugyanezek a tendenciák a végrehajtási idő függvényében is megfigyelhetők voltak. A Recall értékeknél a kisebb modellek teljesítménye rendkívül közel esett egymáshoz, gyakorlatilag nem mutatva érdemi különbséget. A nagyobb modellek esetén viszont eltérések jelentkeztek: a YOLOv11 teljesítménye alig változott, a YOLOv12-é javult, míg a YOLOv8-é csökkent. Ezek a különbségek a végrehajtási idő alakulásában is tükröződtek. Összességében megállapítható, hogy a hálózatok méretének növelése nem jelent automatikusan jobb eredményt, és több esetben a kisebb modellek mutattak kedvezőbb teljesítményt–idő kompromisszumot. Ez különösen releváns korlátozott erőforrású, beágyazott környezetben, ahol a valós idejű feldolgozás elsődleges követelmény.

Az egyes hálózatok tanításának különféle metrikák szerinti számszerű eredményeit az 1. táblázat foglalja össze. Mindezek alapján megállapítható, hogy a kulcspont-alapú detektálás, ebben a speciális részfeladatban, a vizsgált YOLO architektúrákkal hatékonyan megvalósítható. A YOLOv12 modelles család mutatta a legnagyobb potenciált, azonban korlátozott teljesítményű, beágyazott környezetben a YOLOv8n és YOLOv11n architektúrák jelentheti a legjobb kompromisszumot a pontosság és végrehajtási idő között.

1. táblázat A tanítások eredményei

Modell	Paraméter [M db]	Számítási idő [ms]	mAP ⁵⁰⁻⁹⁵ [%]	Precision [%]	Recall [%]
YOLOv8n	3.30	16.15	0.998440	0.999900	0.974090
YOLOv8x	69.49	63.63	0.998320	0.999780	0.966375
YOLOv11n	2.91	21.94	0.996640	0.999905	0.978660
YOLOv11x	58.89	54.55	0.996640	0.999915	0.975420
YOLOv12n	2.89	33.02	0.999580	0.999885	0.974980
YOLOv12x	61.13	93.44	0.996640	1.000000	0.978630

6. Összegzés

A kutatás célja egy olyan kulcspon-alapú vizuális bójadetektáló megoldás kialakítása volt, amely pontosan képes meghatározni a közúti terelőbóják pozícióját és csúcspontját, különösen egy autonóm kiszolgáló robotplatformba integrálva. A feladat megvalósításához saját fejlesztésű, kulcspon-annotációval ellátott képi adatkészlet összeállítása történt, amely változatos környezeti és megvilágítási viszonyok mellett ábrázolja a bójákat főként felülnézeti betekintésből. Az annotálási folyamat háromlépcsős, félig automatizált módszerrel valósult meg, amely a színszűrésen, diszparitás térképen és Hough-transzformáción alapuló pontosság-növelő eljárásokkal biztosította a kulcsponok megbízható lokalizálását.

A vizsgálathoz a YOLO hálózatsalád legújabb generációiból kerültek kiválasztásra a modellek: YOLOv8, YOLOv11 és YOLOv12 különböző méretváltozatai. Minden hálózat módosításra került, hogy képes legyen a csúcspont koordinátájának regressziójára is, majd ezek tanítása és validálása a saját adatkészleten történt.

Az eredmények alapján a YOLOv8 esetében a nagyobb modellváltozat (YOLOv8x) nem eredményezett jobb teljesítményt a kompaktabb YOLOv8n modellhez képest, sőt a tanítási stabilitás is gyengébbnek bizonyult. Hasonló teljesítménycsökkenés figyelhető meg a YOLOv11 esetében, ahol a kisebb modell ismét előnyösebb választásnak bizonyult. Ezzel szemben a YOLOv12 architektúra nagyobb méretű modellje (YOLOv12x) nemcsak pontosságban teljesít jobban, de a tanulási folyamat is stabilabb volt, és a kulcspon-detektálásban megbízhatóbb eredményt nyújtott.

Ugyanakkor minden hálózat esetében megfigyelhető, hogy a végrehajtási idő szignifikánsan nő a modell komplexitásának növekedésével. Ez különösen fontos szempont a valós idejű alkalmazhatóság szempontjából, mivel az autonóm robotplatform működése korlátozott

számítási kapacitáson zajlik. A fenti tényezők mérlegelése alapján a legkisebb, de mégis kellően pontos és gyors működést biztosító modell, a YOLOv8n, jelentheti a leginkább ajánlott megoldást a gyakorlati rendszerintegrációhoz.

A vizsgálatok igazolták, hogy a kulcspon-alapú detektálás működőképes és jól alkalmazható megközelítés a bójakezeléshez szükséges finom pozícióbecslés támogatására. A jelenlegi kutatás alapot teremt a jövőbeli kiterjesztésekhez, különösen a több nézetből történő észlelés, a valós-idejű integráció és a robotkar vezérlésének továbbfejlesztése terén.

A bemutatott kutatás nem csupán technológiai újdonság, hanem gazdasági, társadalmi és környezeti szempontból is jelentős hasznosíthatósággal bír. Gazdasági oldalról az autonóm bójakezelés csökkenti a munkaerőigényt, rövidíti az útlezárások idejét, és mérsékli a közlekedési torlódásokból adódó közvetett költségeket. Társadalmi szempontból javítja a közlekedésbiztonságot, hiszen a pontosan és gyorsan kihelyezett bóják minimalizálják az emberi hibákból fakadó veszélyeket, valamint növelik a karbantartást végző munkások biztonságát azáltal, hogy kímélik őket a forgalom közvetlen veszélyeitől. Környezeti szempontból pedig az automatizált, optimalizált bójakezelés hozzájárul a fenntartható közlekedéshez: rövidebb ideig tartó útlezárások révén kevesebb jármű vesztegel, így csökken a károsanyag-kibocsátás és a környezet terhelése. E tényezők összességében rávilágítanak arra, hogy a bemutatott módszer és a hozzá kapcsolódó autonóm robotplatform fejlesztése nemcsak technológiai újítás, hanem szélesebb értelemben is hozzájárulhat a közúti közlekedés hatékonyabbá, biztonságosabbá és fenntarthatóbbá tételéhez.

A jövőbeli tervek között szerepel a kulcspon-alapú detektálás robusztusságának növelése, különösen nehezített körülmények között, pl. erős árnyék, részleges takarás vagy extrém megvilágítás. Ennek érdekében adatkészlet-bővítés szükséges, beleértve az időjárás és szezonális varianciák, valamint változatos kamera betekintések bevonását. Az architektúrák tekintetében cél a specializált kulcspon-regressziós ágak továbbfejlesztése, pl. multi-head struktúrák, heatmap-alapú megközelítések vagy transformer-alapú lokalizáció alkalmazásával. Hosszabb távon vizsgálni szükséges az észlelés és a robotkar-mozgástervezés közvetlen összekapcsolását, valós időben adaptálódó objektum megfogási stratégiák és szenzor-fúzió bevezetésével. A kutatás végső célja egy autonóm, dinamikus környezetben is megbízhatóan működő, skálázható robotrendszer kialakítása, amely túlmutat a bójakezelésen, és más típusú tárgyak pontos és biztonságos manipulációjára is alkalmassá válik.

7. Köszönetnyilvánítás

A publikációban szereplő kutatást a Széchenyi István Egyetem az Európai Unió támogatásával valósította meg, az RRF-2.3.1-21-2022-00004 azonosítójú, Mesterséges Intelligencia Nemzeti Laboratórium projekt keretében.

Irodalomjegyzék

- [1.] H. Tian, K. Song, J. Xu, S. Ma, és Y. Yan, „Antipodal-points-aware dual-decoding network for robotic visual grasp detection oriented to multi-object clutter scenes”, Expert Systems with Applications, köt. 230, o. 120545, nov. 2023, doi: 10.1016/j.eswa.2023.120545.
- [2.] T. Li, F. Wang, C. Ru, Y. Jiang, és J. Li, „Keypoint-Based Robotic Grasp Detection Scheme in Multi-Object Scenes”, Sensors, köt. 21, sz. 6, o. 2132, márc. 2021, doi: 10.3390/s21062132.
- [3.] J. Luo, L. Zhu, L. Li, és P. Hong, „Robot Visual Servoing Grasping Based on Top-Down Keypoint Detection Network”, IEEE Transactions on Instrumentation and Measurement, köt. 73, o. 1–11, 2024, doi: 10.1109/TIM.2023.3335521.
- [4.] A. Dhall, D. Dai, és L. V. Gool, „Real-time 3D Traffic Cone Detection for Autonomous Driving”, 2019. június 5., arXiv: arXiv:1902.02394. doi: 10.48550/arXiv.1902.02394.
- [5.] I. Katsamenis, E. E. Karolou, A. Davradou, E. Protopapadakis, A. Doulamis, N. Doulamis, D. Kalogeras, „TraCon: A novel dataset for real-time traffic cones detection using deep learning”, 2022. május 24., arXiv: arXiv:2205.11830. doi: 10.48550/arXiv.2205.11830.
- [6.] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, C.L. Zitnick, „Microsoft COCO: Common Objects in Context”, in Computer Vision – ECCV 2014, köt. 8693, D. Fleet, T. Pajdla, B. Schiele, és T. Tuytelaars, Szerk., Cham: Springer International Publishing, 2014, o. 740–755. doi: 10.1007/978-3-319-10602-1_48.
- [7.] N. Crasto, „Class Imbalance in Object Detection: An Experimental Diagnosis and Study of Mitigation Strategies”, 2024, arXiv. doi: 10.48550/ARXIV.2403.07113.

- [8.] T. Kim, H. Jeon, és Y. Lim, „Challenges of YOLO Series for Object Detection in Extremely Heavy Rain: CALRA Simulator based Synthetic Evaluation Dataset”, 2023, arXiv. doi: 10.48550/ARXIV.2312.07976.
- [9.] H. Gu, H. Dong, N. Konz, és M. A. Mazurowski, „A systematic study of the foreground-background imbalance problem in deep learning for object detection”, 2023, arXiv. doi: 10.48550/ARXIV.2306.16539.
- [10.] R. Poureskandar és S. Razzagzadeh, „Improving Object Detection Performance through YOLOv8: A Comprehensive Training and Evaluation Study”, 2025, arXiv. doi: 10.48550/ARXIV.2505.11424.
- [11.] M. L. Ali és Z. Zhang, „The YOLO Framework: A Comprehensive Review of Evolution, Applications, and Benchmarks in Object Detection”, Computers, köt. 13, sz. 12, o. 336, dec. 2024, doi: 10.3390/computers13120336.
- [12.] R. Sapkota, Z. Meng, M. Churuvija, X. Du, Z. Ma, és M. Karkee, „Comprehensive Performance Evaluation of YOLOv12, YOLO11, YOLOv10, YOLOv9 and YOLOv8 on Detecting and Counting Fruitlet in Complex Orchard Environments”, 2025. február 25., arXiv: arXiv:2407.12040. doi: 10.48550/arXiv.2407.12040.
- [13.] J. Hollósi, R. Krecht, és Á. Ballagi, „Development of Advanced Intelligent Robot Platform for Industrial Applications”, ERCIM News, sz. 141, o. 41–42.
- [14.] A. Geiger, P. Lenz, C. Stiller, és R. Urtasun, „Vision meets robotics: The KITTI dataset”, The International Journal of Robotics Research, köt. 32, sz. 11, o. 1231–1237, 2013, doi: 10.1177/0278364913491297.
- [15.] F. Yu, H. Chen, X. Wang, W. Xian, Y. Chen, F. Liu, V. Madhavan, T. Darrell, „BDD100K: A Diverse Driving Dataset for Heterogeneous Multitask Learning”, in 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), jún. 2020, o. 2633–2642. doi: 10.1109/CVPR42600.2020.00271.
- [16.] S. Suzuki és K. Abe, „Topological structural analysis of digitized binary images by border following”, Computer Vision, Graphics, and Image Processing, köt. 30, sz. 1, o. 32–46, ápr. 1985, doi: 10.1016/0734-189X(85)90016-7.

- [17.] R. O. Duda és P. E. Hart, „Use of the Hough transformation to detect lines and curves in pictures”, *Commun. ACM*, köt. 15, sz. 1, o. 11–15, 1972, doi: 10.1145/361237.361242.
- [18.] G. Jocher, A. Chaurasia, és J. Qiu, Ultralytics YOLOv8. 2023. [Online]. Elérhető: <https://github.com/ultralytics/ultralytics>
- [19.] G. Jocher és J. Qiu, Ultralytics YOLO11. 2024. [Online]. Elérhető: <https://github.com/ultralytics/ultralytics>
- [20.] Y. Tian, Q. Ye, és D. Doermann, „YOLOv12: Attention-Centric Real-Time Object Detectors”, *arXiv preprint arXiv:2502.12524*, 2025.
- [21.] Y. Tian, Q. Ye, és D. Doermann, YOLOv12: Attention-Centric Real-Time Object Detectors. 2025. [Online]. Elérhető: <https://github.com/sunsmarterjie/yolov12>
- [22.] H. Robbins és S. Monro, „A Stochastic Approximation Method”, *Ann. Math. Statist.*, köt. 22, sz. 3, o. 400–407, 1951, doi: 10.1214/aoms/1177729586.
- [23.] L. Bottou, „Large-Scale Machine Learning with Stochastic Gradient Descent”, in *Proceedings of COMPSTAT’2010*, Y. Lechevallier és G. Saporta, Szerk., Heidelberg: Physica-Verlag HD, 2010, o. 177–186. doi: 10.1007/978-3-7908-2604-3_16.
- [24.] I. Loshchilov és F. Hutter, „Decoupled Weight Decay Regularization”, in *International Conference on Learning Representations*, 2017. [Online]. Elérhető: <https://api.semanticscholar.org/CorpusID:53592270>

Egységes módszertan járművek távirányítású rendszerekké történő átalakítására

Unified Methodology for Converting Vehicles to Re-mote- Controlled Systems

Kecskeméti István^a , Kőrös Péter^b

^aJárműipari Kutatóközpont

kecskemeti.istvan@ga.sze.hu

^bJárműipari Kutatóközpont

korosp@ga.sze.hu

Absztrakt

A járműautomatizálás gyorsan fejlődő területén a hagyományos járművek távvezérléssel történő utólagos felszerelése jelentős lehetőségeket kínál. Ez a tanulmány egy egységes módszertant mutat be a hagyományos járművek távvezérelhető rendszerre alakítására, amely különböző járműplatformokon adaptálható. A megközelítés a távvezérlési képességek integrálásának következetes folyamatát foglalja magában, amely az egyes járművek vezérlési architektúrájának értékelésével kezdődik, majd a szükséges hardver- és szoftverkomponensek, például érzékelők, működtetők és kommunikációs modulok tervezésével és telepítésével folytatódik. Az egyik legfontosabb jellemzője az adaptálhatóság, amely lehetővé teszi a különböző járműtípusokhoz való testreszabást a hatékonyság és a megbízhatóság veszélyeztetése nélkül. A kísérleti eredmények különböző járművek sikeres távvezérlését bizonyítják. Ez a munka rávilágít a kivitelezhetőségre és a szélesebb körű alkalmazásokra a logisztikában, a mezőgazdaságban és a vészhelyzetekre való reagálásban, a hagyományos járműtervek és a modern távvezérlési technológiák közötti szakadék áthidalására törekedve.

Kulcsszavak: távvezérlés, járműautomatizálás, távvezérelhetőség

Abstract

In the rapidly evolving field of vehicle automation, retrofitting conventional vehicles for remote control offers significant opportunities. This paper presents a unified methodology for converting traditional vehicles into remotely operable systems, adaptable across various vehicular platforms. The approach involves a systematic process for integrating remote control capabilities, starting with assessing each vehicle's control architecture and followed by designing and installing necessary hardware and software components, such as sensors, actuators, and communication modules. A key feature is its adaptability, allowing customization for different vehicle types without compromising performance or reliability. Experimental results demonstrate successful remote control of diverse vehicles. This work highlights the feasibility and broader applications in logistics, agriculture, and emergency response, aiming to bridge the gap between conventional vehicle designs and modern remote control technologies.

Keywords: remote control, vehicle automation, remote operability

1. Bevezetés

Az autonóm járművek térhódítása forradalmasította az autóiipart, azonban az autonóm rendszerek teljes körű alkalmazása még évekig várat magára.

Az önvezető járművek forradalmasíthatják a közlekedést és mindennapi életünket. Az emberi tényező kiküszöbölésével a vezetés közelebb kerül a biztonsághoz, szervezettebbé és környezetbarátabbá válik [1]. A teljes autonómia elérése azonban jelentős műszaki kutatást és pénzügyi befektetéseket igényel ahhoz, hogy a technológia rendelkezésre álljon. Ezt követheti az önvezető járművek szabályozása, amely magában foglalja a jogi felelősség meghatározását, valamint a közvélemény imázsának kialakítása, hogy a társadalom minél szélesebb rétegei elfogadják a technológiát [2,3].

Az autonóm járműveknek több feltételre van szükségük a biztonságos közlekedéshez. Először a környezetérzékelés különböző típusú érzékelők fúzióját használja a jármű körüli terep és az akadályok feltérképezésére. A beérkező adatok alapján a rendszer meghatározza, hogy milyen pályát kövessen, amelyet az alacsony szintű vezérlés referenciaértékként használ.

Az alacsonyabb szinten történik a jármű tényleges vezérlése [4], a rendszer a referenciajelek alapján állítja be a szükséges kormánysszöveget és a jármű sebességét. További fontos feladat annak biztosítása, hogy a fizikai rendszer tisztában legyen a korlátaival, például ha a referenciajelekben impulzusszerű zavarok vannak, azokat képes legyen kiszűrni.

Bár a technológia széles körű alkalmazása még távoli, a most folyó kutatómunka alapozza meg a jövőjét. Ehhez a kutatócsoportoknak már most a legújabb eszközökkel és technológiákkal kell rendelkezniük. A legtöbbször maga a járműplatform jelenti a problémát, amelyen a kutatásokat el lehet végezni. Több vállalat már most is fejleszt önvezető platformokat a meglévő járművekhez, de ezek a kész megoldások jelentős pénzügyi befektetéssel járnak.

Ebben a dolgozatban egy módszertant szeretnénk bemutatni, amellyel távvezérlési funkciókat lehet megvalósítani három teljesen különböző járművön. Ezekre az alapokra építhetők az autonóm funkciók megvalósításához szükséges érzékelők, számítógységek és szoftverkomponensek.

2. Alkalmazott járműplatformok

Ebben a cikkben három különböző járművet használunk, hogy bemutassuk, milyen lépéseket tettünk annak érdekében, hogy ezeket a rendszereket távvezérlési képességekkel szereljük fel. Mindhárom esetben ismertetjük a jármű hardverkörnyezetét, a hajtóműveket és a kommunikációs csatornákat.

A távvezérlés megvalósításához két alapvető dologra van szükség: a kormányszög és a sebesség szabályozására. Ezek lehetővé teszik a jármű irányításának átvételét, de természetesen a biztonságos működéshez ennél lényegesen többre van szükség, amit ebben a tanulmányban nem tárgyalunk részletesen.

2.1 Nissan Leaf

Az első esetben a cél egy olyan járműplatform létrehozása volt, amely az autonóm funkciók kutatására és fejlesztésére szolgál. Egy első generációs Nissan Leafet választottunk (lásd **13. ábra**), az első nagy mennyiségben értékesített tisztán elektromos járművet. A rendszerkialakítás szempontjából minden releváns érzékelő és működtető elem elérhető a CAN-hálózaton.



13. ábra Nissan Leaf, autonóm kutatási platform.

2.2. SZEmission, energiatakarékos jármű.

A Széchenyi István Egyetem egy hallgatói csoportja 2005 óta foglalkozik energiahatékony elektromos járművek fejlesztésével. A csapat minden évben részt vesz a Shell Eco-marathon energiahatékonsági versenyen, ahol a cél egy adott távolság megtétele a lehető legkevesebb energiával. 2018 óta a szervezők egy autonóm versenyt is indítottak, ahol a csapatoknak saját járműveiket önvezető funkciókkal kell felszerelniük (lásd 14. ábra), hogy teljesíteni tudják a szervezők által meghatározott feladatokat.



14. ábra SZEmission, energiahatékony versenyjármű, önvezető képességekkel.

A SZEmission versenyjármű először 2019-ben vett részt a versenyen, és azóta a diákok minden évben folyamatosan fejlesztik, mind mechanikusan, mind elektronikusan. A karosszéria méhsejtes karbon kompozitból készült, a jármű tömege pedig mindössze 95 kg. Az elektronikai rendszer teljesen egyedi, 48V-os rendszerfeszültséggel, háromfázisú PMSM motorral a hajtáshoz és az egységek közötti CAN-kommunikációval.

2.3 Gamma Komondor

A Gamma Zrt. 2010-ben kezdte meg egy olyan könnyű páncélozott jármű kifejlesztését, amely képes a nukleáris, biológiai és vegyi (ABV) szennyeződések felderítésére, elemzésére és értékelésére. Alkalmazása elsősorban a katonai és a katasztrófavédelemben van.

A Komondor könnyű páncélozott jármű (látható a 3. ábrán) méreteiben és belső kialakításában is nagyban eltér a korábban bemutatott járművektől. A meghajtáshoz nagy teljesítményű

dízelmotort használ, amely hatfokozatú automata sebességváltón keresztül juttatja el az erőt a négy kerékre.



15. ábra Gamma Komondor, könnyűpáncélos jármű.

3. Módszertan

A különböző járműplatformok távvezérléssel való felszerelése során egy egységes folyamatot követtünk. A módszertan célja, hogy a különböző hajtáslánccal, kormányművel és kommunikációs architektúrával rendelkező járműveken azonos lépések szerint építsük ki az alacsony szintű irányítást.

3.1 Általános folyamat

A távvezérlés megvalósításához két alapvető funkció vezérlése szükséges: a kormányszög és a sebesség szabályozása. A módszertan az alábbi fő lépésekből áll:

1. **A hardverkörnyezet megértése:** Minden jármű különböző architektúrával rendelkezik, ezért elsőként fel kell mérni a kommunikációs protokollokat (pl. CAN, CANOpen, J1939), a rendelkezésre álló beavatkozási lehetőségeket, valamint a biztonsági jellemzőket.
2. **A komponensek külön tesztelése:** A kormányzási és a hajtáslánc-vezérlő rendszereket először külön-külön kell vizsgálni és tesztprogramokkal ellenőrizni, hogy pontosan megértsük működésüket.

3. **Működési állapot meghatározása:** A járműveknek képesnek kell lenniük a kézi és a távvezérelt működés közötti váltásra. A rendszernek garantálnia kell, hogy a távvezérlés csak meghatározott feltételek mellett aktiválódhasson.
4. **A vezérlők hangolása:** A referenciajeleket (kormányaszög és sebesség) külső forrásból kapja a rendszer, amelyeket az alacsony szintű vezérlők hajtanak végre. A stabil működéshez szükséges a vezérlőhurkok megfelelő paraméterezése és finomhangolása.

Az egységes metodika alkalmazása lehetővé teszi, hogy különböző járműtípusokon hasonló lépések mentén valósítsuk meg a távvezérlést, miközben figyelembe vesszük az adott platform egyedi sajátosságait.

3.2 Esettanulmányok

Az általános folyamatot három különböző járműplatformon alkalmaztuk, amelyek eltérő célokra és műszaki környezetben készültek. Az alábbiakban bemutatjuk a távvezérlési képességek kialakítását a konkrét járműveken.

3.2.1 Nissan Leaf

A Nissan Leaf volt az első jármű, amelyet a rendszerrel távvezérlési funkciókkal szereltünk fel, mivel széles körben elterjedt, sorozatgyártású elektromos autóként jó alapot nyújtott a módszertan validálásához. Első lépésként részletesen feltártuk a jármű meglévő rendszereit: a kormányzást, a hajtásláncot és a fékrendszert, valamint a lehetséges beavatkozási pontokat. A kormányrendszer elektromos rásegítésű, amely a kormányoszlopra kifejtett nyomatékkal arányos rásegítő nyomatékot biztosít. A vezető által kifejtett erőt analóg érzékelő méri, amelynek jele a CAN-hálón jelenik meg. E jel felülírásával lehetőség nyílt a kormányzási szög szabályozott vezérlésére [5].

A gyorsítás a gyári elektromos gázpedál referenciajelének manipulálásával volt megvalósítható, míg a lassítás a villamos gép motorfék funkcióján, illetve szükség esetén a hidraulikus fékrendszer aktiválásán keresztül történt. A Leaf esetében a legnagyobb kihívást az jelentette, hogy a gyári rendszerek integritását megőrizve kellett beavatkozni, így a mérőrendszer szoftveres és hardveres szinten is szigorú biztonsági rétegeket kapott. Ez a platform jól szemléltette, hogyan tud a rendszer ipari szériagyártású járműveknél működni anélkül, hogy a jármű eredeti architektúráját jelentősen módosítani kellene.

3.2.2 SZEmission

A SZEmission hallgatói versenyjármű teljesen egyedi elektronikával rendelkezik, így minden vezérlőegység működése ismert és módosítható volt. Ez lehetőséget teremtett arra, hogy a mérőrendszert mélyebben integráljuk a jármű működésébe. A motorvezérlő a nyomatékreferenciát CAN-üzenet formájában kapta, a hajtás pedig fékezésre is képes volt, így mind gyorsítási, mind lassítási parancsok közvetlenül a buszon keresztül továbbíthatók voltak. A fordulatszám pontos szabályozásához inkrementális kódolót alkalmaztunk, amely nagy felbontású visszajelzést adott.

A kormányrendszer esetében a cél az energiahatékonyság maximalizálása volt. Az Ackermann-geometria betartásához szervohajtást építettek be, amely szíjhajtáson keresztül mozgatta a kerekeket. A vezérlőjelek PWM és digitális formában is adhatók voltak [6], így a mérőrendszer többféle jelkimeneti módot tudott alkalmazni. A SZEmission példája különösen jól demonstrálta a rendszer nyitottságát: itt nem meglévő gyári egységekhez kellett illeszkedni, hanem az általunk meghatározott struktúrához lehetett optimalizálni a mérési és vezérlési funkciókat. Ezáltal a rendszer alkalmassá vált versenykörnyezetben, kísérleti prototípusok vizsgálatára is.

3.2.3 Gamma Komondor

A Komondor könnyű páncélozott jármű teljesen más jellegű kihívást jelentett, mivel katonai alkalmazásra készült, és három egymástól elkülönülő CAN-hálózattal rendelkezett. Ezek közül a legtöbb a J1939 szabványt alkalmazta, amely ipari és nehézgépjárművekben elterjedt kommunikációs protokoll. A meghajtás nagy teljesítményű dízelmotorral történt, amely automata sebességváltón keresztül adta le a nyomatékot a hajtásláncre.

A kormányrendszer azonban hidraulikus volt, és nem rendelkezett elektromos rásegítéssel, így a távvezérlés itt fizikai beavatkozást igényelt. Erre a célra külső léptetőmotort alkalmaztunk, amely CANopen protokollon keresztül volt vezérelhető. A mérőrendszernek ezért nemcsak a kommunikációs illesztést, hanem a mechanikus aktuátor vezérlését is biztosítania kellett. A Komondor esete jól mutatta, hogy a rendszer a tisztán elektromos hajtású járművek mellett hagyományos, dízelüzemű és erősen mechanikus rendszerekkel rendelkező platformokra is adaptálható. Ez jelentősen kiterjesztette a mérőrendszer alkalmazási területét.

4. Eredmények

Mindhárom esetben sikeresen létrehoztunk egy alacsony szintű irányítást a korábbi módszertan szerint. A távvezérlési tesztekét először a járműben végeztük el, az alsó szintet működés közben ellenőriztük, majd a referenciajeleket vezeték nélküli kapcsolaton keresztül továbbítottuk.

A tesztek fontos feladata volt a kormányzási jellemzők meghatározása. A teljes tartományt több egyenlő részre osztottuk, és megmértük a külső és belső kerekek elfordulási szögét, majd a mért adatokból felrajzoltuk a jelleggörbét, amelyből a szoftver kiszámítja a tényleges kormányzási szöveget.

5. Értekezés

A fejlesztés során sikeresen felszereltünk három járművet távvezérlési képességekkel, amelyeket teljesen különböző alkalmazásokban fogunk használni. Minden esetben a legalacsonyabb költséggel a RIO vezérlőt építettük be. Fontos megjegyezni, hogy mindhárom esetben kísérleti járművekkel dolgoztunk, ahol a kutatás és fejlesztés a fő szempont, így a hardveres módosítások során nem vettük figyelembe az ergonómiát és a kényelmet. A projektek folytatásaként mindhárom járművet felszereltük a környezet érzékeléséhez szükséges érzékelőkkel, hogy később autonóm funkciókat tesztelhessünk rajtuk.

6. Köszönetnyilvánítás

A publikációban szereplő kutatást a Széchenyi István Egyetem az Európai Unió támogatásával valósította meg, az RRF-2.3.1-21-2022-00004 azonosítójú, Mesterséges Intelligencia Nemzeti Laboratórium projekt keretében.

Irodalomjegyzék

- [1] J. Wang, L. Zhang, Y. Huang, and J. Zhao, 'Safety of Autonomous Vehicles', *Journal of Advanced Transportation*, vol. 2020, no. 1, p. 8867757, 2020, doi: 10.1155/2020/8867757.
- [2] S. Yoo, J. Kumagai, Y. Kawabata, A. R. Keeley, and S. Managi, 'Buy and/or Pay Disparity: Evidence from Fully Autonomous Vehicles', *Applied Sciences*, vol. 14, no. 1, Art. no. 1, Jan. 2024, doi: 10.3390/app14010410.

- [3] S. Yoo and S. Managi, ‘To fully automate or not? Investigating demands and willingness to pay for autonomous vehicles based on automation levels’, *IATSS Research*, vol. 45, no. 4, pp. 459–468, Dec. 2021, doi: 10.1016/j.iatssr.2021.11.002.
- [4] J. E. Naranjo, F. Jiménez, O. Gómez, and J. G. Zato, ‘Low Level Control Layer Definition for Autonomous Vehicles Based on Fuzzy Logic’, *Intelligent Automation & Soft Computing*, vol. 18, no. 4, pp. 333–348, Jan. 2012, doi: 10.1080/10798587.2012.10643247.
- [5] G. Rödönyi *et al.*, *Identification of the nonlinear steering dynamics of an autonomous vehicle*. 2021.
- [6] Z. Pusztai and P. Kőrös, ‘Steering mechanism design for lightweight vehicle’, Jun. 2020.

Technológiaelfogadási mintázatok az egyetemisták körében: klaszteralapú megközelítés

Technology adoption patterns among university students: a cluster-based approach

Koteczki Réka^a

^aJárműipari Kutatóközpont, Széchenyi István Egyetem

koteczki.reka@ga.sze.hu

Absztrakt

A generatív mesterséges intelligencia technológiák egyre hangsúlyosabb szerepet kapnak a felsőoktatásban. A Z generáció, mint digitális bennszülöttek csoportja, kulcsszerepet játszik az ilyen technológiák elfogadásában és elterjesztésében, ugyanakkor attitűdjeik és motivációik jelentős különbségeket mutatnak. A jelen kutatás célja az volt, hogy feltárja, milyen típusú hallgatói szegmensek léteznek a mesterséges intelligenciával kapcsolatos attitűdök mentén, különös tekintettel az oktatási alkalmazásra. A kutatás során egy strukturált kérdőív került elemzésre, amely az AI elfogadásával kapcsolatos kulcstényezőket – például észlelt hasznosságot, élvezetet, attitűdöt és társadalmi befolyást – vizsgálta. A klaszterelemzés négy jól elkülöníthető csoportot azonosított: (1) AI-rajongók és innovátorok, (2) Pozitív elfogadók, (3) Mérsékeltén nyitott felhasználók és (4) Szkeptikus és óvatos felhasználók. Az egyes csoportok jelentős eltéréseket mutattak a technológia iránti nyitottság, a használati szándék, valamint az élvezeti és társadalmi tényezők mentén. Az eredmények rámutatnak arra, hogy a generatív AI elfogadása nem tekinthető egységes jelenségnek a hallgatók körében, hanem eltérő attitűdcsoportok mentén értelmezhető. A szegmentáció alapján célzott oktatáspolitikai, intézményi vagy fejlesztési stratégiák alakíthatók ki, amelyek segíthetik az AI-eszközök hatékonyabb és inkluzívabb integrációját a felsőoktatásba.

Kulcsszavak: mesterséges intelligencia, technológiaelfogadás, Z generáció, klaszterelemzés, felsőoktatás, ChatGPT

Abstract

Generative artificial intelligence technologies are playing an increasingly important role in higher education. Generation Z, as a group of digital natives, plays a key role in the adoption and dissemination of such technologies, but their attitudes and motivations vary significantly. The aim of this study was to explore the types of student segments that exist in terms of attitudes toward artificial intelligence, with a particular focus on educational applications. The research analyzed a structured questionnaire that examined key factors related to AI acceptance, such as perceived usefulness, enjoyment, attitude, and social influence. Cluster analysis identified four distinct groups: (1) AI enthusiasts and innovators, (2) positive acceptors, (3) moderately open users, and (4) skeptical and cautious users. Each group showed significant differences in

terms of openness to technology, intention to use, and enjoyment and social factors. The results indicate that the acceptance of generative AI cannot be considered a uniform phenomenon among students, but can be interpreted along different attitude groups. Based on this segmentation, targeted educational policy, institutional, or development strategies can be developed to help integrate AI tools into higher education in a more effective and inclusive way.

Keywords: artificial intelligence, technology adoption, Generation Z, cluster analysis, higher education, ChatGPT

1. Bevezetés

Az elmúlt évtizedben a mesterséges intelligencia (AI) fejlődése alapjaiban változtatta meg a mindennapi élet szinte minden területét, különösen az oktatásban, ahol az AI-alapú eszközök (pl. chatbotok, intelligens tutor rendszerek, személyre szabott tanulási platformok) fokozatosan beépültek a tanítás és tanulás folyamatába [1]. A generatív MI technológiák újfajta interaktív és adaptív tanulási lehetőségeket kínálnak, amelyek hatással vannak az oktatási szokásokra, elvárásokra és attitűdökre is [2]. A fiatalabb generációk, különösen a Z generáció – az 1997 és 2012 között született digitális bennszülöttek – technológiai szocializációja alapvetően eltér az idősebb korosztályokétól [3]. Ők már olyan környezetben nőttek fel, ahol az online jelenlét, a mobiltechnológiák és az algoritmusok által vezérelt szolgáltatások mindennaposak. Ebből következik, hogy nyitottabbak az új technológiák iránt, ugyanakkor érzékenyebbek az etikai, adatvédelmi vagy megbízhatósági kérdésekre is [4].

Klaszterelemzést lehetőséget ad arra, hogy a felhasználókat attitűdjeik, szokásaik és technológiahasználati mintázataik alapján szegmentáljuk [5]. Ez különösen fontos az oktatás területén, ahol a hallgatók heterogenitása meghatározó tényező az innovatív eszközök sikeres bevezetésében. A jelen kutatás célja tehát az volt, hogy a Z generációhoz tartozó egyetemi hallgatók körében azonosítsuk azokat a különböző attitűdcsoportokat, amelyek eltérően viszonyulnak az MI oktatási célú alkalmazásához. A kvantitatív kérdőíves felmérés során gyűjtött adatok alapján klaszterelemzést végeztünk, amely lehetővé tette, hogy a hallgatói populációt négy szegmensbe soroljuk, figyelembe véve technológiai nyitottságukat, elfogadási szintjüket, társadalmi érzékenységüket és MI-használati hajlandóságukat.

2. Szakirodalom

Az AI oktatási alkalmazásai az elmúlt évek egyik legintenzívebben kutatott területévé váltak. Az AI eszközök képesek átalakítani a tanulási környezeteket, személyre szabott tanulási lehetőségeket kínálva, növelve a tanulók produktivitását, és csökkentve az oktatók adminisztratív terheit [6]. Ezzel párhuzamosan a kutatók figyelme egyre inkább az AI elfogadásának pszichológiai, szociológiai és attitűdbeli aspektusai felé fordult.

2.1 Technológiaelfogadási modellek szerepe az AI alapú technológiák megértésében

Az oktatási célú technológiaelfogadás vizsgálatának két legismertebb modellje a Technology Acceptance Model (TAM) és az Unified Theory of Acceptance and Use of Technology (UTAUT). A TAM modell szerint a technológiahasználat szándékát leginkább az észlelt hasznosság (Perceived Usefulness – PU) és az észlelt könnyű használat (Perceived Ease of Use – PEOU) határozza meg [7]. Az elmúlt években több tanulmány is megerősítette, hogy a generatív MI eszközök, például a ChatGPT, elfogadásában ezek a tényezők kulcsszerepet játszanak [8].

A TAM modellt továbbfejlesztve a Venkatesh és munkatársai által kidolgozott UTAUT modell négy fő konstruktumot vezetett be: teljesítményelvárás (PE), erőfeszítés-elvárás (EE), társadalmi befolyás (SI) és támogató környezet (FC) [9]. A modell kiterjesztett változata, az UTAUT2 pedig új dimenziókkal – például élvezet (hedonic motivation), szokás és árérték arány – egészítette ki a fogyasztói környezetben való alkalmazás érdekében [10]. E bővítések különösen relevánsak a generatív AI-al kapcsolatos attitűdök vizsgálatakor, mivel az eszközök használata gyakran nem kötelező, hanem szabad választáson alapul.

Miközben a TAM és UTAUT modellek sikeresen magyarázzák a technológiahasználat mögött álló kognitív és affektív mechanizmusokat, e modellek gyakran átlagos válaszadókat feltételeznek, és nem veszik figyelembe a felhasználók közötti heterogenitást. Az oktatásban különösen fontos, hogy ne csak az általános trendeket értsük meg, hanem azonosítsuk az eltérő tanulói attitűdcsoportokat is – ebben a klaszterelemzés kiemelt módszertani lehetőséget kínál [11]. A klaszteranalízis segítségével meghatározhatók olyan hallgatói csoportok, amelyek hasonló attitűdökkel, motivációval vagy technológiai preferenciával rendelkeznek. Az ilyen szegmentáció lehetővé teszi, hogy az AI integrációját célzott beavatkozásokkal lehessen támogatni, például tréningek, ajánlott használati gyakorlatok vagy közösségi támogatás formájában.

2.3 Egyetemi hallgatók technológia elfogadása

Az egyetemi hallgatók technológiahasználati szokásai és attitűdjei az elmúlt években jelentős átalakuláson mentek keresztül. A digitális oktatási környezetek elterjedése, az online tanulás térnyerése, valamint az AI megjelenése új elvárásokat és kihívásokat teremtett az oktatás minden szintjén. A hallgatók MI-hez való viszonyulását számos tényező befolyásolja. Egyrészt

ott vannak a kognitív dimenziók, mint az észlelt hasznosság (PU) vagy a könnyű használhatóság (PEOU), amelyek döntően befolyásolják, hogy a hallgatók milyen értéket tulajdonítanak a technológiának a tanulási céljaik elérésében. Másrészt jelentős szerepet játszanak az affektív tényezők, például az élvezet vagy a személyes attitűd, amelyek főként a motivációs és emocionális kötődéseket tükrözik. Végül, a szociális kontextus, azaz a társak, oktatók vagy intézmények elvárásai is formálják a hallgatók döntéseit, különösen olyan újszerű és bizonytalan használati helyzetekben, mint a generatív MI alkalmazása [12].

A hallgatói sokféleség figyelembevétele különösen fontos akkor, amikor egy új technológiát önkéntes alapon lehet használni. A kötelező használattal szemben, ahol a külső elvárások dominálnak, az önkéntes alkalmazás kontextusában a belső motivációs tényezők (pl. élvezet, önhatékonyság, személyes célok) nagyobb jelentőséggel bírnak. Ennek következtében a felsőoktatási intézményeknek olyan stratégiákat kell kialakítaniuk, amelyek nemcsak a technológia hasznosságát, hanem az élményalapú használatot és a közösségi megerősítést is támogatják. Az ESG-jelentések szövegbányászati elemzését bemutató friss kutatások rámutatnak arra, hogy a strukturált, automatizált tartalomelemzés hatékony eszköze lehet az átláthatóság és az értékelhetőség javításának, ami párhuzamba állítható az MI-elfogadási mintázatok feltárásával az oktatásban [13].

3. Módszertan

A kutatás során egy kvantitatív kérdőíves felmérést alkalmaztunk, amelyet 2024 nyarán végeztünk egy magyarországi egyetem hallgatói körében. A mintát 278 fő egyetemi hallgató alkotta, akik önkéntes alapon töltötték ki az online kérdőívet. A kérdőív célja az volt, hogy feltérképezze a hallgatók AI technológiákkal kapcsolatos attitűdjeit, valamint az elfogadással és használati szándékkal összefüggő pszichológiai dimenziókat. A válaszadók kizárólag aktív hallgatói jogviszonnyal rendelkező személyek voltak. A minta nemre, szakirányra és évfolyamra nézve vegyes összetételű, ami lehetővé tette, hogy eltérő AI-hoz való hozzáállással rendelkező csoportokat azonosítsunk. A minta nem tekinthető országosan reprezentatívnak, de lefedi a különböző szakirányokat és évfolyamokat, ami megfelelő heterogenitást biztosít az attitűdök elemzéséhez. A nemek és karok aránya közel egyenletes volt, így az eredmények általános érvényessége korlátozott, de indikációt ad az egyetemi hallgatók attitűdbeli mintázataira. A résztvevők többsége nappali tagozatos alapszakos vagy mesterszakos hallgató

volt. A kérdőív több, a szakirodalomban validált skálát tartalmazott. Az AI-elfogadás vizsgálatához hat pszichológiai konstruktum alapján mértük a hallgatói attitűdöket:

- **Észlelt hasznosság** (*Perceived Usefulness – PU*)
- **Észlelt könnyű használhatóság** (*Perceived Ease of Use – PEOU*)
- **Általános hozzáállás** (*Attitűd – A*)
- **Társadalmi befolyás** (*Social Influence – SI*)
- **Élvezeti érték** (*Enjoyment – E*)
- **Használati szándék** (*Behavioral Intention to Use – BI*)

E konstruktumokat összesen 15 kijelentés alapján mértük, 5 fokozatú Likert-skálán (1 = egyáltalán nem értek egyet, 5 = teljes mértékben egyetértek). A klaszterelemzés célja az volt, hogy azonosítsuk a hallgatói populáción belül a mesterséges intelligenciával kapcsolatos attitűdök mentén elkülönülő attitűdcsoportokat. Ez lehetővé teszi, hogy ne csupán összesített értékek alapján értelmezzük az elfogadási hajlandóságot, hanem a különböző gondolkodási és viselkedési mintázatokat is megértsük. Az adatelemzést IBM SPSS 28 szoftverrel végeztük. A klaszterképzéshez előzetesen standardizáltuk a fő változókat, majd hierarchikus klaszterelemzést alkalmaztunk Ward-módszerrel és squared Euclidean távolságmértékkel a kezdeti klaszterszám becslésére. Ezt követően K-közép (k-means) klaszterezést használtunk a végleges klaszterek meghatározásához, a stabilitás és interpretálhatóság érdekében.

4. Eredmények

A klaszterelemzés célja az volt, hogy feltárja, milyen attitűdcsoportok különböztethetők meg az egyetemi hallgatók körében az AI elfogadásával kapcsolatban. Az elemzés során a válaszadók válaszaiból kiindulva négy jól elkülöníthető klasztert azonosítottunk. Ezek a klaszterek eltérő attitűdökkel, elfogadási szintekkel és technológiai nyitottsággal jellemezhetők. Az alábbiakban részletesen bemutatjuk az egyes csoportokat. Az 1. ábrán látható a klaszter csoportok százalékos megoszlása.



1. ábra: Klaszter csoportok százalékos megoszlása

(1) Mérsékeltelen nyitott technológiahasználók (n = 65)

Ez a csoport nyitott az AI használatára, de nem tekinti azt elengedhetetlennek a tanulmányai vagy a mindennapi élet szempontjából. Az észlelt hasznosság (PU) és élvezeti érték (E) dimenziók esetében közepes értékeket mutattak, míg a használati szándék (BI) és a társadalmi befolyás (SI) alacsonyabb szintű volt. E klaszter tagjai nem utasítják el a technológiát, azonban sem különösebb rajongás, sem mély integráció nem jellemzi a használatukat. Jellemző rájuk a technológiai óvatosság és az, hogy inkább eseti jelleggel alkalmazzák az MI-t, ha az kényelmi vagy hatékonysági előnyt jelent számukra.

(2) Pozitív elfogadók (n = 104)

A második klaszterbe tartozók kifejezetten pozitívan viszonyulnak a mesterséges intelligencia eszközeihez. A válaszadók magas értékeket mutattak mind az észlelt hasznosság (PU), mind az élvezeti érték (E) terén, és az általános attitűdjük (A) is kedvező. Bár nem minden válaszadó használja napi szinten az MI-t, a használati szándék (BI) kifejezetten erős bennük. Az MI-t hasznos tanulási segédeszköznek tartják, amely támogatja őket feladataik hatékonyabb elvégzésében. Különösen értékelik, ha a technológia intuitív, gyors és személyre szabott támogatást nyújt. Ugyanakkor nem feltétlenül aktív promóterei vagy „early adopterei” az új eszközöknek.

(3) Szkeptikus és óvatos felhasználók (n = 33)

A harmadik klaszter a legkisebb, ugyanakkor markánsan eltérő attitűdöket képvisel. A válaszadók alacsony szintű észlelt hasznosságot (PU), élvezeti értéket (E) és használati szándékot (BI) jeleztek, és általánosságban negatív vagy semleges attitűd jellemzi őket a mesterséges intelligencia iránt. Nem bíznak az MI-eszközök pontosságában vagy megbízhatóságában, és gyakran aggályaik vannak az etikai, adatvédelmi vagy oktatási integritás kérdéskörében is. E csoport tagjai vagy még nem használták aktívan az MI-t, vagy negatív tapasztalataik alapján döntöttek az elutasítás mellett. Számukra az MI nem nyújt hozzáadott értéket a tanulásban, sőt, egyes esetekben inkább elidegenítő tényezőként jelenik meg.

(4) AI-rajongók és innovátorok (n = 76)

Ez a csoport a legaktívabb és legpozitívabb felhasználói szegmens. A tagok rendkívül magas értékeket mutattak minden dimenzióban – különösen az észlelt hasznosság (PU), az élvezet (E) és a használati szándék (BI) terén. Az AI-rajongók gyakran saját elhatározásból kezdtek el MI-eszközöket használni, és ezek a megoldások szerves részévé váltak tanulási stratégiáiknak. Az ilyen hallgatók nem csupán elfogadók, hanem új technológiák kipróbálói, promoterei is. Számukra az MI egy olyan eszköz, amely újfajta lehetőségeket nyit meg a tanulásban, önfejlesztésben és problémamegoldásban. Emellett sokuk számára az MI használata örömteli élményt is nyújt, amely önmagában is motiváló.

A klaszterelemzés eredményei megerősítették, hogy az MI elfogadása az egyetemi hallgatók körében nem egységes jelenség, hanem komplex, többdimenziós attitűdstruktúra mentén differenciálódik. Az azonosított négy csoport különböző stratégiákat és támogatási formákat igényel, ha célunk az MI technológiák szélesebb körű és felelős integrációja az oktatásba. A hallgatók eltérő szintű nyitottsága, motivációja és tapasztalata olyan tényezők, amelyeket figyelembe kell venni az MI-eszközök fejlesztése és bevezetése során. Az azonosított négy attitűdcsoport hasonló szerkezetet mutat, mint Cabrera-Sánchez et al. [5] és Habibi et al. [4] vizsgálataiban, ahol szintén elkülöníthetők voltak a pozitív elfogadók és a szkeptikus felhasználók csoportjai. Ugyanakkor a jelen kutatásban az élvezet (enjoyment) dimenziója erősebb prediktorként jelent meg, ami a generatív MI interaktív jellegének tulajdonítható.

5. Összegzés

A kutatás célja az volt, hogy klaszterelemzéssel feltárja az egyetemi hallgatók mesterséges intelligenciával kapcsolatos attitűdjeinek és elfogadási szintjeinek sokféleségét. A 278 fős minta elemzése alapján négy világosan elkülöníthető attitűdcsoporthoz rajzolódott ki: a mérsékelt nyitott technológiahasználók, a pozitív elfogadók, a szkeptikus és óvatos felhasználók, valamint az AI-rajongók és innovátorok. Az eredmények megerősítik, hogy az AI-hez való viszonyulás nem egységes a hallgatók körében, hanem eltérő attitűdök, motivációk és tapasztalatok mentén differenciálódik. A klaszterek jellemzőinek mélyebb megértése fontos gyakorlati következményekkel jár. A pozitív elfogadók és innovátorok ösztönzése mellett a mérsékelt nyitott hallgatók célzott támogatással könnyebben bevonhatók lehetnek az MI-eszközök használatába. Ezzel szemben a szkeptikus csoport számára külön figyelmet kell fordítani a bizalomépítésre, etikai és adatbiztonsági kérdések tisztázására. Az eredmények rámutatnak arra, hogy a mesterséges intelligencia oktatási célú integrációja csak akkor lehet sikeres, ha az intézmények figyelembe veszik a hallgatók sokféleségét, és szegmentált, érzékeny bevezetési stratégiákat alkalmaznak. Az eredmények hasznosíthatók az egyetemi AI-stratégiák tervezésében, például a hallgatói szegmensekhez illeszkedő képzési programok, tréningek és bevezetési kommunikáció kialakításában. Ez hozzájárulhat a felelős és inkluzív AI-használathoz a felsőoktatásban. A jövőbeni kutatásokban érdemes tovább vizsgálni, hogy milyen tényezők segítik vagy gátolják az egyes attitűdcsoporthoz való elmozdulást a magasabb elfogadás irányába.

5. Köszönetnyilvánítás

A publikációban szereplő kutatást a Széchenyi István Egyetem az Európai Unió támogatásával valósította meg, az RRF-2.3.1-21-2022-00004 azonosítójú, Mesterséges Intelligencia Nemzeti Laboratórium projekt keretében.

Irodalomjegyzék

- [1.] Mustafa, M. Y., Tlili, A., Lampropoulos, G., Huang, R., Jandrić, P., Zhao, J., ... & Saqr, M. (2024). A systematic review of literature reviews on artificial intelligence in education (AIED): a roadmap to a future research agenda. *Smart Learning Environments*, 11(1), 1-33.

- [2.] Basch, C., Hillyer, G., Gold, B., & Yousaf, H. (2025). Understanding artificial intelligence knowledge and usage among college students: Insights from a survey on classroom, coursework, and personal applications. *Review of Artificial Intelligence in Education*, 6, e037-e037.
- [3.] Lee, C. S., Tan, L. E., & Goh, D. H. L. (2025). Examining generation Z's use of generative AI from an affordance-based approach. *Information Research an international electronic journal*, 30(iConf), 1095-1102.
- [4.] Habibi, A., Muhaimin, M., Danibao, B. K., Wibowo, Y. G., Wahyuni, S., & Octavia, A. (2023). ChatGPT in higher education learning: Acceptance and use. *Computers and Education: Artificial Intelligence*, 5, 100190.
- [5.] Cabrera-Sánchez, J. P., Villarejo-Ramos, Á. F., Liébana-Cabanillas, F., & Shaikh, A. A. (2021). Identifying relevant segments of AI applications adopters—Expanding the UTAUT2's variables. *Telematics and Informatics*, 58, 101529.
- [6.] Boussouf, Z., Amrani, H., Khal, M. Z., & Daidai, F. (2024). Artificial intelligence in education: A systematic literature review. *Data and Metadata*, 3, 288-288.
- [7.] Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, 319-340.
- [8.] Masadeh, S. A., & El-Haggag, N. (2024). Analyzing factors influencing IoT adoption in higher educational institutions in Saudi Arabia using a modified TAM model. *Education and Information Technologies*, 29(5), 6407-6441.
- [9.] Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS quarterly*, 425-478.
- [10.] Venkatesh, V., Thong, J. Y., & Xu, X. (2012). Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS quarterly*, 157-178.
- [11.] Cabrera-Sánchez, J. P., Villarejo-Ramos, Á. F., Liébana-Cabanillas, F., & Shaikh, A. A. (2021). Identifying relevant segments of AI applications adopters—Expanding the UTAUT2's variables. *Telematics and Informatics*, 58, 101529.

- [12.] Shahzad, M. F., Xu, S., & Javed, I. (2024). ChatGPT awareness, acceptance, and adoption in higher education: the role of trust as a cornerstone. *International Journal of Educational Technology in Higher Education*, 21(1), 46.
- [13.] Lukács, B., Rickards, R. C., Molnár, P., Suta, A., & Tóth, Á. (2025). ESG disclosure topics and reporting frameworks: exploratory research across automotive, construction, and energy industries. *Discover Sustainability*, 6(649).

Neurális hálózati modellek összehasonlító vizsgálata jármű-orientáció osztályozására

A Comparative Study of Neural Network Models for Vehicle Orientation Classification

Magda Donát^a, Hollósi János^b

^aSzéchenyi István Egyetem

magda.donat@hallgato.sze.hu

^bSzéchenyi István Egyetem, Járműipari Kutatóközpont

hollosi.janos@sze.hu

Absztrakt

A tanulmány célja, hogy összehasonlító vizsgálatot végezzen különböző neurális hálózati modellekkel az autók képalapú orientációs osztályozása céljából. Az imént említett neurális hálózatok az önvezető járművek döntéstámogató rendszereinek kulcsfontosságú alapját képezik. Hat különböző architektúrát, köztük előre meghatározott és saját fejlesztésű modelleket értékeltünk három különböző tanulási környezetben: képek az eredeti adatkészlet háttérével, ezek háttér nélküli és véletlenszerű háttérképekkel ellátott verziói. Az összehasonlítás elsődleges szempontjai a tanulási idő, a pontosság és a modell komplexitása voltak. Az eredmények azt mutatták, hogy a YOLO és az EfficientNet modellek kiváló teljesítményt mutattak a pontosság és a robusztusság tekintetében, míg a véletlenszerű háttérrel való tanulás elősegítette az általánosítási képességet. A kutatás hangsúlyozza, hogy a tanulási környezet és az adatmódosítási stratégiák legalább olyan fontosak, mint maga a modellarchitektúra.

Kulcsszavak: képosztályozás, neurális hálózat, önvezető járművek

Abstract

The objective of this study is to undertake a comparative investigation into various neural network models for the purpose of image-based orientation classification of cars. The mentioned neural networks are a pivotal feature in decision support systems for self-driving vehicles. Six different architectures, including predefined and self-developed models, were evaluated in three distinct learning environments: original background, without background, and with random background images. The primary criteria for comparison were learning time, accuracy, and model complexity. The findings demonstrated that the YOLO and EfficientNet models exhibited superior performance in terms of accuracy and robustness, while learning with random backgrounds promoted generalization ability. The research emphasises that the learning environment and augmentation strategies are at least as crucial as the model architecture itself.

Keywords: image classification, neural network, self-driving vehicles

1. Bevezetés

A képfelismerő algoritmusok fejlődése kulcsfontosságú az önvezető járművek biztonságának és hatékonyságának növelése szempontjából. A megbízható működéshez alapvető fontosságú a környezet objektumainak, például más járművek helyzetének és tájolásának valós idejű felismerése és értelmezése. Az önvezető technológiákban használt mesterséges intelligencia (AI) modellek autonómiát biztosítanak, és fontos kiegészítő és biztonsági funkciókat látnak el, segítve a kritikus helyzetek felismerését és elkerülését.[1,2]

E kutatás célja a különböző mesterséges intelligenciamodellek összehasonlítása az autók tájékozódásának felismerésére, különös tekintettel a tanulási idő, a pontosság és a modell összetettsége közötti összefüggésekre. Mind a végső pontosságot, mind a mögöttes okokat, például a modell architektúráját és optimalizálási viselkedését értékeljük. A kísérleteket egységes tanulási környezetben végeztük a Carvana Image Masking adathalmaz felhasználásával. Az összehasonlítás a PyTorch beépített neurális hálózatain és a YOLO architektúrán alapul. Az optimalizációs eljárásokhoz az SGD és az Adam algoritmusokat használtuk alapértelmezett paraméterezéssel, a veszteségfüggvényt pedig minden esetben a CrossEntropyLoss szolgáltatatta. A modelleket egységesen képeztük ki 32-től 64 epocháig terjedő kötegmérettel, hogy az összehasonlítás objektív és reprodukálható legyen.[3-7]

2. Adathalmazok

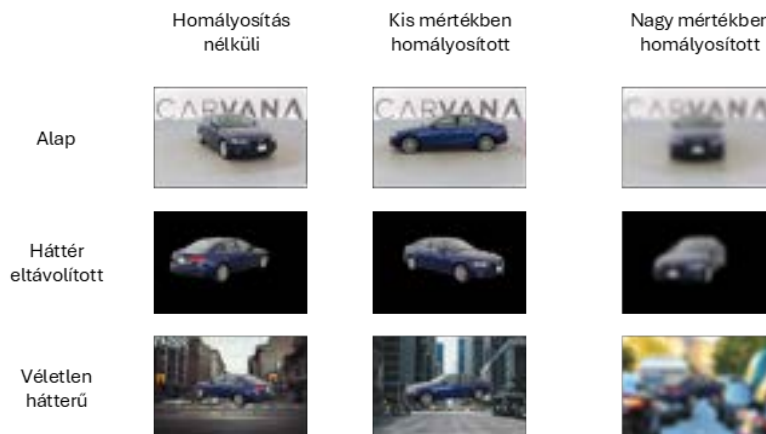
A kísérleti környezetet a Carvana Image Masking Challenge adathalmaz felhasználásával terveztük, amely összesen 5088 képet tartalmaz 318 különböző járműről. Minden járműhöz 16 képet kaptunk 22,5°-os forgási lépésekkel, így teljes 360°-os képet kaptunk. Az adatkészlet minden egyes járműhöz maszkoló képeket is tartalmazott, lehetővé téve ezzel a háttér pontos eltávolítását.

A feldolgozott képekből két további adatkészletet alakítottunk ki (lásd 1. ábra):

1. A járművekről készült képekkel összefüggésben a háttér eltávolítása során maszkoló adatokat használtak fel a képek fekete háttérrel való fedéséhez. (továbbiakban „He”, mint háttér eltávolított)
2. A kísérlet során három különböző, természetes környezetet ábrázoló hátteret fedtünk fel a háttér nélküli képekre. Az összemelési folyamat előtt a vizuális konzisztencia

biztosítása érdekében a járműveken arányos méretcsökkentést alkalmaztak.
(továbbiakban „VH”, mint véletlen háttérű)

Ennek következtében összesen három adatsor állt rendelkezésre az eredeti adatsorral együtt tanításhoz.



1. ábra: A kutatás során felhasznált adathalmazok és az azokon végzett transzformációk

A modellek robusztusságának és általánosítási képességének felmérése érdekében három szintű adatmódosítást hajtottunk végre (lásd 1. ábra).

1. A jelen vizsgálat nem tartalmaz semmilyen transzformációt (kontroll).
2. A jármű enyhe elmosódásnak van kitéve, amely a mozgás hatását hivatott szimulálni.
3. A képet nagyfokú elmosódás jellemzi, ami a nem optimális látási viszonyokat jelzi, mint például az eső vagy köd okozta látási viszonyok.

A képeket véletlenszerűen három adathalmazra osztottuk: 30% a tanításhoz, 30% a validáláshoz és 30% a teszteléshez. Ez biztosította, hogy az eredmények megbízhatóan tükrözzék a modellek általánosíthatóságát a különböző előfeldolgozási stratégiák mellett. A kutatási törekvés elsődleges célja a járművek tájolásának osztályozása volt. Ennek érdekében az egyes járművekhez tartozó 16 képet a következő 8 osztályozásba soroltuk a képsorozatban szereplő képek száma alapján (lásd 1. táblázat):

1. Közeledik (1. kép)
2. Balra közeledik (2–4. kép)
3. Balra halad (5. kép)
4. Balra távolodik (6–8. kép)

5. Távolodik (9. kép)
6. Jobbra távolodik (10–12. kép)
7. Jobbra halad (13. kép)
8. Jobbra közeledik (14–16. kép)

1. táblázat: Minta a kategóriákból

<i>ID</i>	<i>Name</i>	<i>Sample</i>
0	Közeledik	
1	Balra közeledik	
2	Balra halad	
3	Balra távolodik	
4	Távolodik	
5	Jobbra távolodik	
6	Jobbra halad	
7	Jobbra közeledik	

E kategóriák célja a járművek irányított mozgásának osztályozása, és hogy a modell a járművek helyzetéből következtetni tudjon a járművek relatív mozgásirányára. Ez különösen fontos az autonóm járművek biztonságos navigációja szempontjából, ahol a környező járművek pontos irányának felismerése kritikus fontosságú.

3. Használt modellek

A jelen tanulmányban különböző felépítésű mélytanulási modelleket vizsgáltunk az autók haladási irányának osztályozása céljából. A vizsgálatra kiválasztott modellek különböző technológiai korszakokat és architekturális megközelítéseket képviselnek, a klasszikus konvolúciós hálózatoktól kezdve a modern hatékony mélytanulási hálózatokon át a legújabb transzformátor-alapú megoldásokig. Az összehasonlítás elsődleges szempontjai a tanulási idő, az osztályozási pontosság, valamint a paraméterek száma és a modell komplexitása közötti kapcsolat voltak.

AlexNet

A modell a mély neurális hálózat architektúrájának egy sajátos típusa, amelyet számos alkalmazásban alkalmaztak, többek között számítógépes látásban és természetes nyelvi feldolgozásban. Az AlexNet-et a konvolúciós neurális hálózatok egyik legkorábbi és legismertebb képviselőjének tekintik, amely 2012-ben áttörő eredményt ért el az ImageNet versenyen. A modell nyolc rétegből áll, és körülbelül 60 millió paramétert tartalmaz. A PyTorch beépített implementációját használtuk fel, mivel az egyszerű, mégis mélyreható struktúrával rendelkezik, amely vetekszik a korszerűbb modellekkel, így optimális kiindulópontként szolgál az összehasonlításhoz. (lásd 2. ábra)[8]



2. ábra: Alexnet felépítése

ResNet-18

A ResNet-18 a Residual Neural Network (ResNet) család egyik legegyszerűbb tagja, amelyet 2015-ben mutattak be. A megközelítés elsődleges újítása a reziduális blokkszerkezet beépítése, amely megkönnyíti a kihagyott kapcsolatok tanulását, és hatékonyan kezeli a mély hálózatokra jellemző tanulási nehézségeket, többek között a gradiens eltűnésének problémáját. Ebben a tanulmányban a PyTorch könyvtár beépített implementációját használtuk, amely 18 réteget és összesen körülbelül 11,7 millió tanulható paramétert tartalmaz. [9]

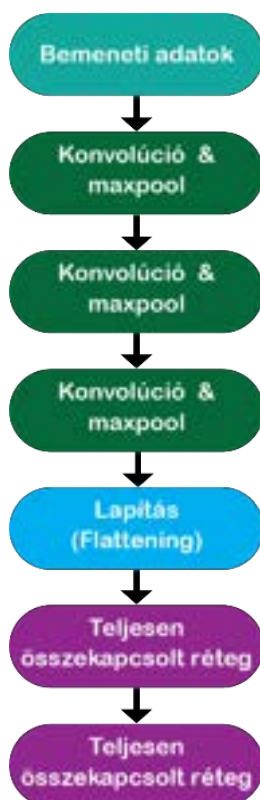
EfficientNet

Ez a modell egy gépi tanulási algoritmus, amelyet a képek jellemzőinek azonosítására és kinyerésére fejlesztettek ki. Az EfficientNet a hatékonyságot helyezi előtérbe, egyidejűleg skálázva a hálózat mélységét, szélességét és bemeneti felbontását. A tanulmányban használt

EfficientNet-B0 verzió 21,5 millió paramétert, valamint MBConv és Fused-MBConv blokkokat tartalmaz. A modell a klasszikus CNN-ekhez képest jelentősen megnövelt paraméterhatékonyságot mutat.[10]

Saját fejlesztésű CNN-modell

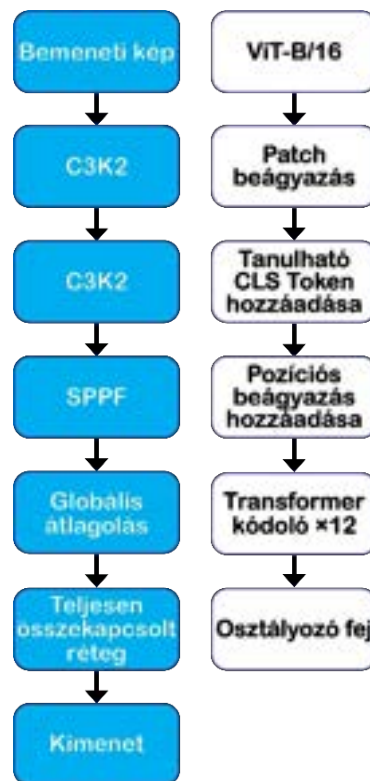
A saját fejlesztésű, mindössze 730 000 tanulható paraméterből és 6 rétegből álló kis konvolúciós hálózatot is kiértékelünk. A modellt a PyTorch keretrendszer segítségével fejlesztettük ki azzal a céllal, hogy egy egyszerű, jól érthető referenciamodellt biztosítsunk a bonyolultabb architektúrák teljesítményének értékeléséhez. (lásd 3. ábra)



3. ábra: Saját modell felépítése

YOLOv11n-cls

A YOLOv11n-cls modell az Ultralytics 2024 YOLO modellcsalád része, amelyet általában objektumfelismerésre fejlesztettek ki; ebben az esetben azonban az osztályozásra optimalizált változatot használtuk. A modell 1,6 millió paramétert tartalmaz, és a legkorszerűbb, optimalizált architektúra alacsony számítási költség mellett versenyképes teljesítményt nyújt. (lásd 4. ábra)[7]



4. ábra: Yolo11 és ViT-B/16 felépítése

ViT-B/16

A 2020-ban bevezetett Vision Transformer (ViT) architektúra jelentős eltérést jelent a hagyományos konvolúciós megközelítéstől, mivel teljesen transzformátor alapú feldolgozást alkalmaz. A használatban lévő ViT-B/16 változat 85,8 millió paraméterrel dolgozik, és a képeket 16*16 pixeles foltokban dolgozza fel egy önfigyelő mechanizmuson keresztül. A számítási igények ellenére a módszer képes pontos és általánosítható eredményeket produkálni. (lásd 4. ábra)[11]

4. Eredmények

A tanulmány célja, hogy felmérjük a modellek teljesítményét három különböző tanulási környezetben: az eredeti környezetet ábrázoló háttér, a háttér nélküli és egy véletlenszerű háttérrel rendelkező képekből álló adathalmazon. Az eredmények egyértelműen bizonyítják, hogy a tanulási környezet jelentős hatást gyakorol a modellek osztályozási pontosságára, robusztusságára és tanulási folyamatának hatékonyságára. Az ebből következő eredményeket a következő szakaszban a tanulási adatkészletek szerint csoportosítva mutatjuk be.

Alap adatkészlet

Az eredeti háttérű képeken betanított modellek magas szintű pontosságot értek el a tesztkörnyezetben. A ViT-B/16, az EfficientNet és a YOLO architektúrák magas szintű pontosságot mutattak, míg az AlexNet és a saját fejlesztésű modell szerényebb, de egyenletesebb teljesítményt mutatott. Az optimalizálók hatása ebben a környezetben modellfüggőnek bizonyult: Az Adam hozzájárulása többek között az összetettebb modellek pontosságának javítását jelentette, míg az SGD kiegyensúlyozottabb eredményeket hozott a kisebb architektúrák esetében. A háttérinformációk és a téma közötti kongruencia előnyös volt a tanulási folyamat során; ez azonban korlátozta a későbbi általánosítás képességét.

(lásd 2. táblázat)

2. táblázat: Alap adatkészleteken mért pontosságok

	<i>SGD [%]</i>	<i>ADAM [%]</i>
<i>AlexNet [8]</i>	98,28	18,67
<i>EfficientNet [10]</i>	99,53	99,35
<i>ResNet18 [9]</i>	75,87	98,99
<i>Self-Written</i>	93,54	18,67
<i>Vit_B_16 [11]</i>	79,96	22,70
<i>YOLO [7]</i>	99,77	99,6

Háttér nélküli adatkészlet

A háttér nélküli képeken betanított modellek teljesítménye összességében hasonló tendenciát mutatott; a pontosság azonban fokozott érzékenységet mutatott a tesztkörnyezet változásaira. Az AlexNet és a ResNet18 csekély mértékű pontosságcsökkenést mutatott, míg az EfficientNet és a YOLO architektúrák következetesebb eredményeket mutattak. Az optimalizálók hatása hasonló maradt, az Adam a nagyobb modellek esetében, az SGD pedig a kisebb architektúrák esetében mutatott jobb teljesítményt. A kontextuális információk hiánya azt eredményezte, hogy a modellek a tárgyak elsődleges vizuális jellemzőire koncentráltak. Ez a megközelítés azonban nem biztosította az összes modell számára a megfelelő robusztusságot a változatos környezeti kontextusokban. (lásd 3. táblázat)

3. táblázat: Háttér nélküli adatkészleteken mért pontosságok

	<i>SGD [%]</i>	<i>ADAM [%]</i>
<i>AlexNet [8]</i>	98,40	18,47
<i>EfficientNet [10]</i>	99,23	99,76
<i>ResNet18 [9]</i>	75,48	99,41
<i>Self-Written</i>	93,94	18,47
<i>Vit_B_16 [11]</i>	81,95	33,79
<i>YOLO [7]</i>	81,59	62,52

Véletlen háttérrel ellátott adatkészlet

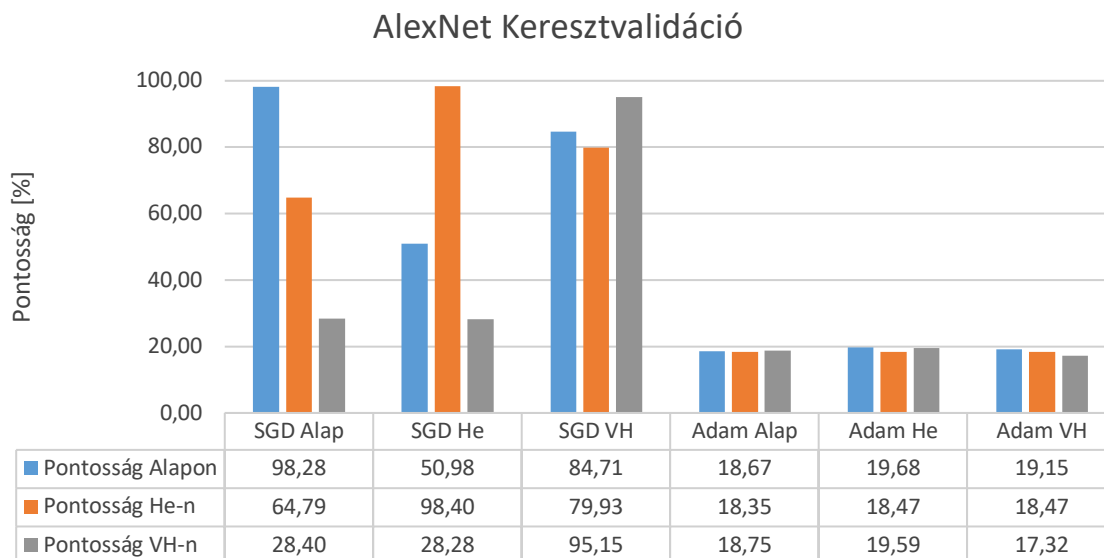
A véletlenszerű háttérű képeken tanított modellek általában gyengébb teljesítményt mutattak a saját tanulási környezetükben; a keresztellenőrzés során azonban jobb általánosítási képességet mutattak. Az AlexNet, a ResNet18 és az EfficientNet architektúrák érzékenységet mutattak a háttér változásaira, míg a ViT-B/16 teljesítménycsökkenése a globális megfigyelési mechanizmus háttérfüggőségének tulajdonítható. A YOLO architektúra bizonyult a legstabilabbnak, különösen akkor, ha az SGD optimalizálóval együtt alkalmazzák. Az Adam alkalmazása ebben a konkrét környezetben több modell teljesítményének csökkenését eredményezte. (lásd 4. táblázat)

4. táblázat: Random háttérrel ellátott adatkészleteken mért pontosságok

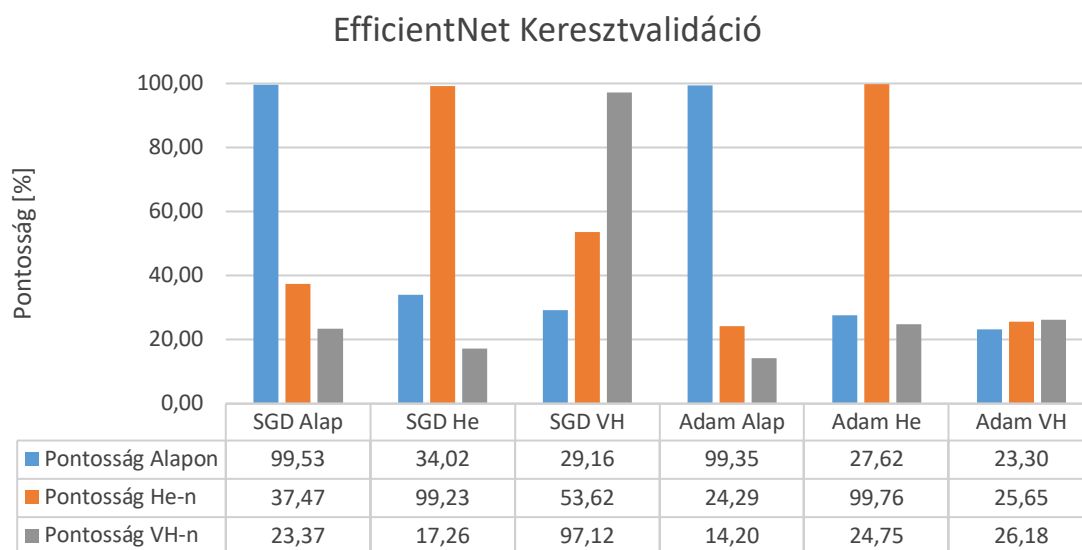
	<i>SGD [%]</i>	<i>ADAM [%]</i>
<i>AlexNet [8]</i>	95,15	17,32
<i>EfficientNet [10]</i>	97,12	26,18
<i>ResNet18 [9]</i>	70,58	93,05
<i>Self-Written</i>	87,96	17,32
<i>Vit_B_16 [11]</i>	49,97	19,05
<i>YOLO [7]</i>	99,70	26,86

A keresztellenőrzési kísérletek megerősítették, hogy a véletlenszerű háttértanulás hozzájárul a modellek robusztusságához. A háttérinformációk sokfélesége miatt a modellek kevésbé támaszkodnak a környezeti jellemzőkre, és inkább az objektumok alapvető vizuális tulajdonságait tanulják. Következésképpen a véletlenszerű háttértanulás különösen előnyös

olyan alkalmazásokban, ahol a bemeneti képek háttere ismeretlen vagy erősen változó.
(lásd 5-10. ábrák)

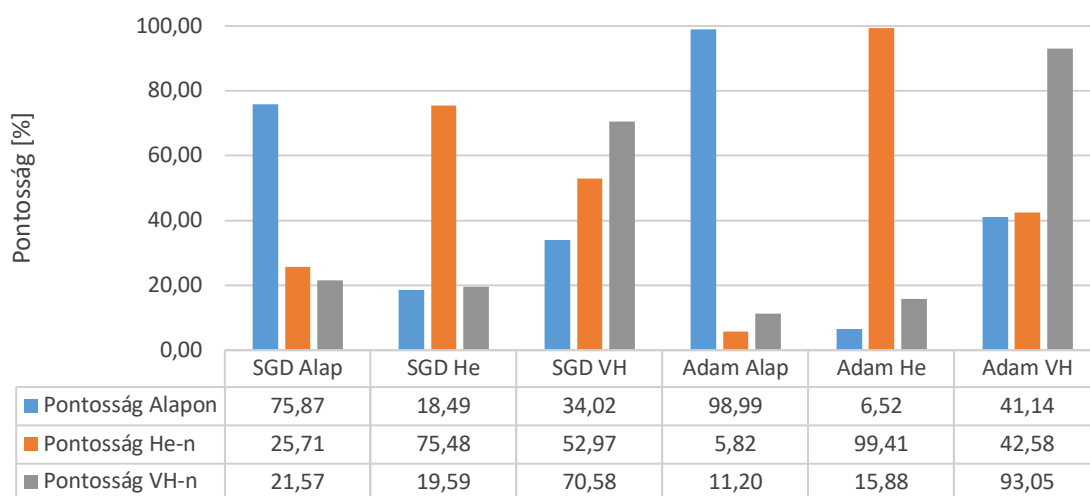


5. ábra: Alexnet keresztvalidációs mérései



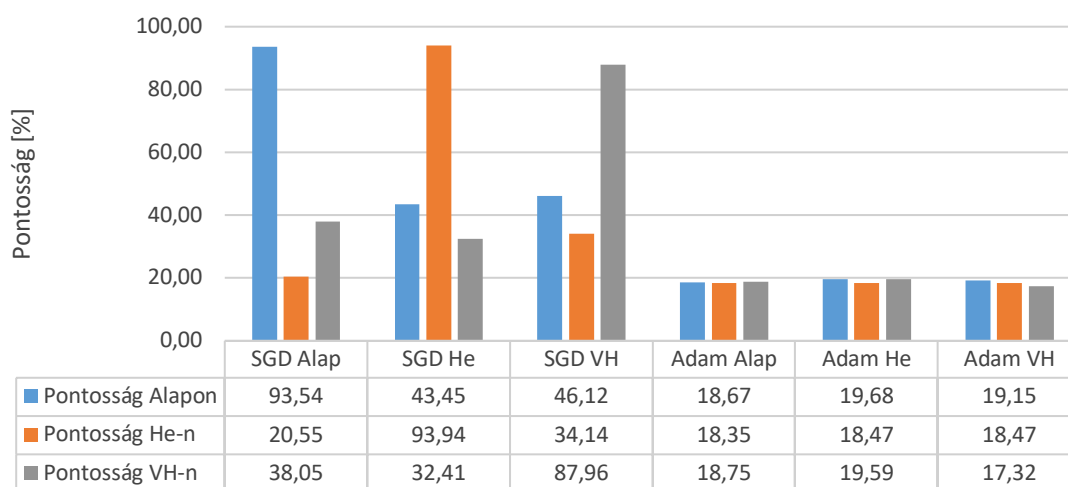
6. ábra: Efficientnet keresztvalidációs mérései

Resnet18 Keresztvalidáció



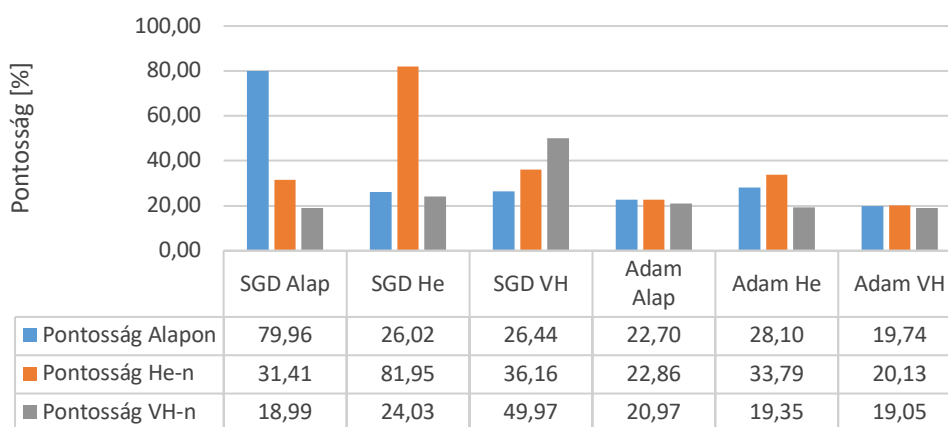
7. ábra: Resnet18 keresztvalidációs mérései

Saját modell Keresztvalidáció



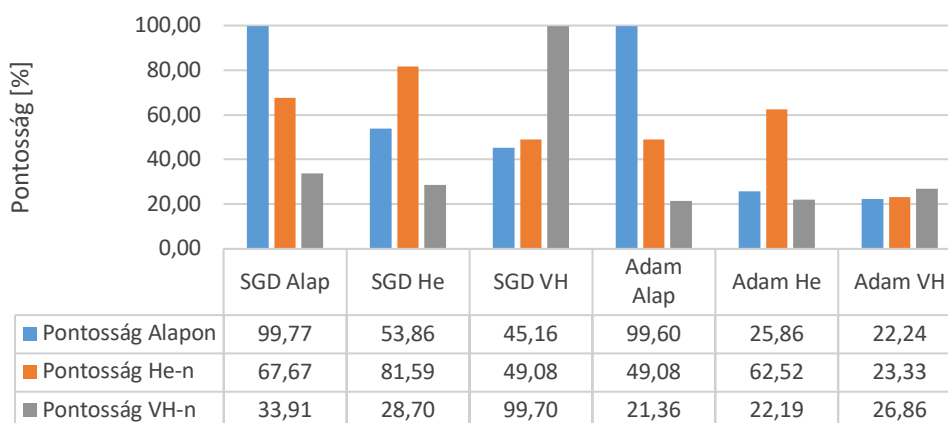
8. ábra: Saját modell keresztvalidációs mérései

Vit_b_6 Keresztvalidáció



9. ábra: Vit_b_16 keresztvalidációs mérései

YOLO Keresztvalidáció



10. ábra: YOLO keresztvalidációs mérései

Tanulási idő és modell komplexitás közötti kapcsolat

A modellek tanulási ideje és komplexitása szoros összefüggést mutatott: a kisebb számú paraméterrel rendelkező hálózatok, mint például a saját modell vagy az AlexNet, rövid tanulási időt igényeltek, míg a nagyobb architektúrák, mint például a ViT-B/16 vagy az EfficientNet, lényegesen hosszabb futási idővel rendelkeztek. A YOLO a kisebb paraméterszám ellenére stabil és gyors tanulást mutatott, míg a ViT-B/16 a globális felügyeleti mechanizmusai miatt a leghosszabb tanulási időt igényelte. (lásd 5. táblázat)

5. táblázat: Modellek paraméter száma és átlagos tanulási idejük

	<i>Paraméter szám</i>	<i>Tanulási idő (perc)</i>
<i>AlexNet [8]</i>	60 M	8
<i>EfficientNet [10]</i>	21.5 M	60
<i>ResNet18 [9]</i>	11.7 M	12
<i>Self-Written</i>	730 k	42
<i>Vit_B_16 [11]</i>	85.8M	120
<i>YOLO [7]</i>	1.6M	20

Fontos megjegyezni, hogy a megnövekedett tanulási idő és a nagyobb komplexitás nem feltétlenül garantál jobb teljesítményt, különösen a robusztusság tekintetében. Ennek illusztrálására tekintsük a véletlenszerű háttértanulás során megfigyelt általánosítási előnyt. Ez az előny független a modell méretétől, és ehelyett a tanulási környezet változékonyságától függ. Ez megerősíti azt az állítást, hogy a hatékony tanulási környezet kiválasztása túlmutat a kezdeti modellválasztáson és kulcsfontosságú szerepet tölt be a végső teljesítményeredmények és a kapott modellek gyakorlati alkalmazhatóságának meghatározásában.

5. Következtetések

A jelen tanulmány célja az volt, hogy összehasonlítsa a különböző mélytanulási modelleket a jármű tájolásának képalapú osztályozására. Az elemzés számos mérőszámot felölelt, beleértve a modellek pontosságát, a tanulási időt, az architektúra komplexitását és a különböző háttérkörnyezetekre való reagálóképességüket.

Az eredmények azt mutatják, hogy az előre elkészített, optimalizált modellek, nevezetesen az EfficientNet és a YOLO architektúrák következetes és magas teljesítményt nyújtanak, még változó környezeti feltételek mellett is. A YOLO modell különösen erősnek bizonyult a robusztusság és a tanulási sebesség tekintetében, míg az EfficientNet a pontosság tekintetében nyújtott kiváló eredményeket.

A háttér nélküli és véletlenszerű háttérrel rendelkező képeken végzett kísérletek rávilágítottak arra, hogy a tanulási környezet változatossága jelentős hatással van a modellek általánosítási képességére. A véletlenszerű háttérrel történő képzés hatékonyságát a túlillesztés

mérséklésében az bizonyította, hogy a modelleket arra ösztönözte, hogy a környezeti jellemzőkkel szemben a tárgyak alapvető vizuális jellemzőit helyezték előtérbe.

A nagyszámú paraméterrel rendelkező modellek, mint például a ViT-B/16, hosszú tanulási időt igényeltek; teljesítményük azonban nem mindig igazolta a megnövekedett számítási költségeket. Ezzel szemben a kisebb, egyszerűbb modellek - beleértve a saját hálózatot is - elfogadható pontosságot biztosítottak gyorsabb tanulási idő mellett, különösen jól strukturált környezetben.

Összességében a kutatás megerősíti, hogy a pontosság és a paraméterek száma önmagában nem elegendő kritérium a modellek kiválasztásához. Azt, hogy egy adott modell milyen hatékonysággal alkalmazható valós helyzetekben, például az önvezető járművek vizuális rendszereihez, a tanulási környezet gondos kialakítása, az adatmódosítási stratégiák és az alkalmazási cél figyelembevétele határozza meg.

6. Köszönetnyilvánítás

A publikációban szereplő kutatást a Széchenyi István Egyetem az Európai Unió támogatásával valósította meg, az RRF-2.3.1-21-2022-00004 azonosítójú, Mesterséges Intelligencia Nemzeti Laboratórium projekt keretében.

Irodalomjegyzék

- [1] J. Janai, F. Güney, A. Behl, and A. Geiger, “Computer Vision for Autonomous Vehicles: Problems, Datasets and State of the Art,” *Foundations and Trends® in Computer Graphics and Vision*, vol. 12, no. 1–3, pp. 1–308, 2020, doi: 10.1561/06000000079.
- [2] C. Badue *et al.*, “Self-driving cars: A survey,” *Expert Syst Appl*, vol. 165, p. 113816, 2021, doi: <https://doi.org/10.1016/j.eswa.2020.113816>.
- [3] A. Paszke *et al.*, “PyTorch: An Imperative Style, High-Performance Deep Learning Library,” in *Advances in Neural Information Processing Systems*, H. Wallach, H. Larochelle, A. Beygelzimer, F. d Alché-Buc, E. Fox, and R. Garnett, Eds., Curran Associates, Inc., 2019. [Online]. Available:

https://proceedings.neurips.cc/paper_files/paper/2019/file/bdbca288fee7f92f2bfa9f7012727740-Paper.pdf

- [4] L. Bottou, “Large-Scale Machine Learning with Stochastic Gradient Descent,” in *Proceedings of COMPSTAT’2010*, Y. Lechevallier and G. Saporta, Eds., Heidelberg: Physica-Verlag HD, 2010, pp. 177–186.
- [5] D. P. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization,” 2017. [Online]. Available: <https://arxiv.org/abs/1412.6980>
- [6] B. Shaler, DanGill, Maggie, M. McDonald, Patricia, and W. Cukierski, “Carvana Image Masking Challenge,” 2017.
- [7] G. Jocher and J. Qiu, “Ultralytics YOLO11,” 2024. [Online]. Available: <https://github.com/ultralytics/ultralytics>
- [8] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017, doi: 10.1145/3065386.
- [9] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778. doi: 10.1109/CVPR.2016.90.
- [10] M. Tan and Q. V Le, “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks,” 2020. [Online]. Available: <https://arxiv.org/abs/1905.11946>
- [11] A. Dosovitskiy *et al.*, “An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale,” 2021. [Online]. Available: <https://arxiv.org/abs/2010.11929>

Transformers and LLMs in Computer Vision: Current Advances and Future Directions

Norbert Markó^a, Áron Ballagi^b, Tamás Szirányi^c

^{a,b}Vehicle Industry Research Center, Széchenyi István University, Győr, Hungary

marko.norbert@ga.sze.hu, ballagi@sze.hu

^{a,c}Machine Perception Laboratory, HUN-REN SZTAKI, Budapest, Hungary

sziranyi.tamas@sztaki.hu

Abstract

Transformers have transformed computer vision by unifying perception and language through self-attention. Once dependent on cloud infrastructure, multimodal large language models (MLLMs) are now increasingly runnable locally on consumer and edge devices. This paper surveys the evolution from CNNs to transformer-based vision–language systems, outlining a taxonomy that links architectures, runtimes, and efficiency techniques to local deployment. Open models and quantization methods such as 4-bit GPTQ and LoRA/QLoRA fine-tuning enable private, low-latency inference. Together they define a decentralized, reproducible ecosystem in which multimodal intelligence becomes efficient, portable, and accessible beyond the cloud.

Keywords: Transformers, Multimodal large language models, Vision–language models, Local-first AI, Edge inference, On-device intelligence

1. Introduction

Convolutional neural networks (CNNs) defined an era of computer vision, powering classification, detection, and segmentation systems that dominated the 2010s. The introduction of transformer architectures, originally developed for natural language processing, fundamentally changed this paradigm [1]. By replacing fixed receptive fields with self-attention, transformers allowed models to capture long-range dependencies and global context, reframing vision as a sequence modeling problem. The Vision Transformer (ViT) and its hierarchical descendants extended this idea [1,3] to large-scale perception, while detection and segmentation models such as DETR treated object localization as a direct sequence prediction problem [2]. In parallel, contrastive pretraining and cross-modal alignment blurred the distinction between visual and linguistic representation. Models such as CLIP demonstrated that language could serve as a semantic scaffolding for vision [5], enabling generalization far beyond task-specific datasets. This evolution culminated in multimodal large language models (MLLMs) that integrate image, text, and increasingly video or audio inputs within a unified generative framework [6,7,9]. These systems combine perception, reasoning, and communication within a single architecture.

Yet their remarkable performance has come with a scaling cost. The largest multimodal models require hundreds of gigabytes of memory and are accessible only through proprietary APIs or cloud infrastructures. This dependence on centralized computation imposes limits on transparency, reproducibility, and accessibility. As a result, local AI development surged, which seeks to bring advanced multimodal capabilities directly to consumer and edge devices. In a local-first paradigm, the emphasis shifts from centralization to distributed capability. Models run where data are produced, enabling privacy, determinism, and real-time response without external connectivity. This transition is supported by three technological trends. First, rapid hardware progress, from NVIDIA's compact desktop GPUs to Apple's M-series systems-on-chip and Jetson-class embedded boards, has widened the feasible compute envelope for inference. Second, compression and efficiency techniques such as quantization (INT8, 5-bit, 4-bit) [10-12], pruning, and structured sparsity have dramatically reduced memory and energy requirements. Third, an expanding ecosystem of open runtimes, including llama.cpp, TensorRT-LLM, vLLM [15], MLX, and OpenVINO, now provides optimized, cross-platform execution for these models. Together, these advances are transforming multimodal intelligence from a cloud-exclusive resource into an accessible, locally deployable technology.

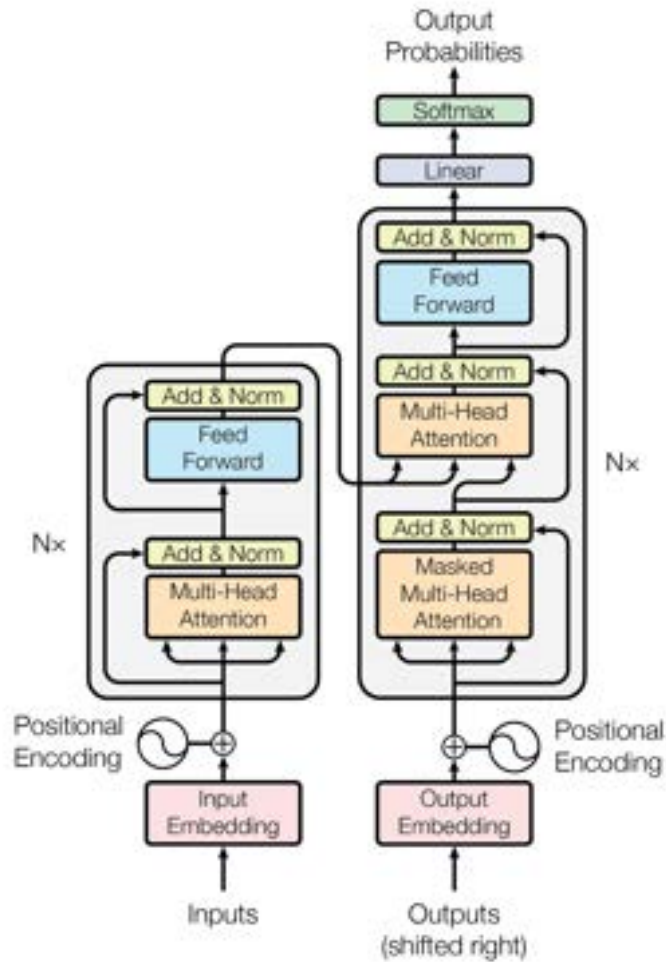


Figure 1: The Transformer Architecture, on which all modern LLMs are based on.

Despite this progress, fragmentation remains. Model lineages, quantization schemes, and runtime toolchains differ widely; licensing and distribution policies are inconsistent. Moreover, evaluation practices often ignore the realities of local deployment such as latency, VRAM limits, and energy consumption. For practitioners, the central challenge is not simply identifying the most accurate model, but navigating this heterogeneous ecosystem of trade-offs between performance, efficiency, and openness. This paper consolidates and contextualizes the emerging field of local-first multimodal AI, where large transformer-based vision–language models are optimized for on-device execution. Its key contributions are:

1. A unified framework and taxonomy that connects transformer architectures, model formats, runtimes, and accelerators to the constraints of local deployment.

2. A curated overview of open multimodal models runnable on consumer and embedded hardware, organized by scale, lineage, and quantization compatibility.
3. An integrative synthesis of enabling methods and open challenges, covering compression, fine-tuning, benchmarking, and interoperability as the foundation of sustainable, decentralized AI.

With these contributions, we aim to improve efficient and privacy-preserving local multimodal intelligence that operates independently of cloud infrastructure. The remainder of this paper is organized as follows. Section 2 and 3 provides the theoretical background on transformer architectures in vision, their extension to multimodal reasoning, and presents a taxonomy for local deployment. Section 4 surveys open multimodal models currently executable on local hardware, categorized by size, modality, and root lineage. Section 5 unifies efficiency techniques, adaptation strategies, runtimes, evaluation methods, and open challenges into a coherent framework for building and sustaining local-first multimodal intelligence. Finally, Section 6 concludes with reflections on the decentralization of AI and the future trajectory of locally deployable vision-language systems.

2. Transformers for Computer Vision

The adoption of transformers in computer vision marked a decisive shift from convolutional architectures that had dominated the field for more than a decade. The central innovation of transformers, self-attention, replaces the fixed receptive fields of convolutional filters with a mechanism that learns to weigh relationships between all tokens in an input sequence. When images are decomposed into sequences of patches, as introduced in the Vision Transformer (ViT), attention enables the model to adaptively integrate information across the entire visual field. This approach removes hand-engineered biases about locality, instead allowing context to be learned directly from data. However, the cost of global attention is high. The quadratic growth of memory and computation with respect to the number of image patches renders naive ViT models inefficient at high resolutions. Subsequent research introduced hierarchical transformers, which restrict attention to local windows and then progressively merge patches across layers. By recovering a multiscale processing structure, these models inherit some of the efficiency advantages of CNNs while retaining the contextual flexibility of attention.

Transformers also reshaped the framing of fundamental vision tasks. The DETR family demonstrated that detection and segmentation could be reformulated as sequence prediction problems, with learnable queries decoded directly into bounding boxes or masks. This eliminated the need for anchor priors, region proposals, or post-processing steps such as non-

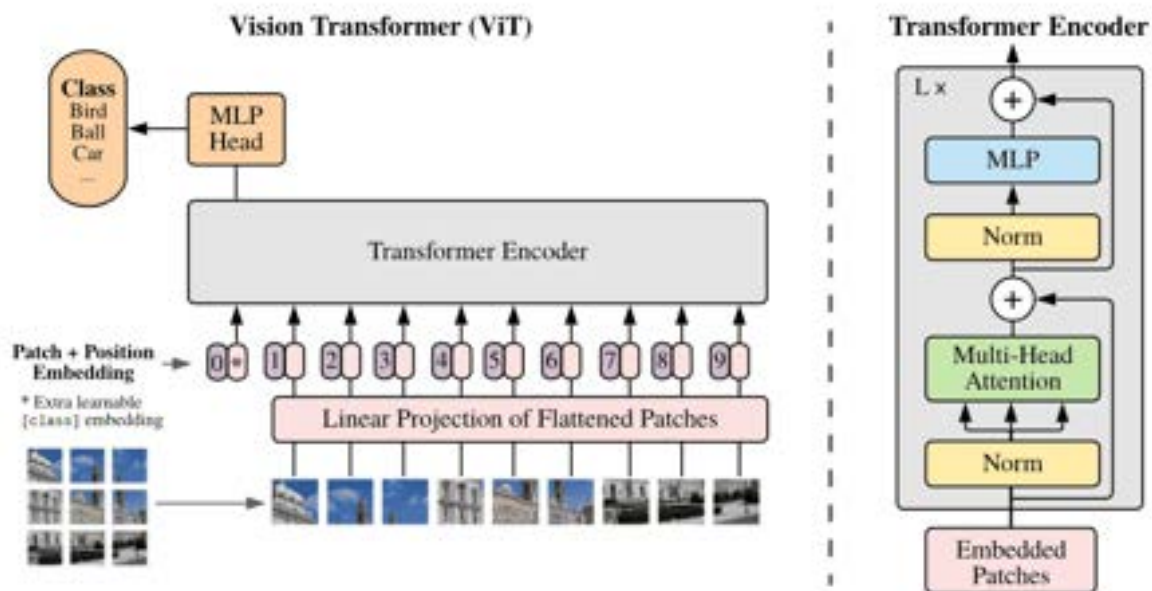


Figure 2: The Vision Transformer Architecture. This was the first work, that adapted the transformer architecture for vision purposes.

maximum suppression, producing end-to-end pipelines that are simpler and more generalizable. Pretraining strategies further amplified the utility of transformer backbones. Contrastive learning across modalities, as popularized by CLIP, showed that joint image-text supervision yields highly transferable representations, while masked autoencoding [4] produced robust encoders even from unlabelled data. These encoders now serve as standard components within larger multimodal systems. For local deployment, such pretraining has practical significance: models that have absorbed general knowledge from large datasets can be frozen or lightly adapted, reducing the need for extensive re-training on constrained hardware. Based on this, it can be seen, that transformers redefined computer vision both architecturally and conceptually. They replaced specialized pipelines with unified sequence-based frameworks, opened the door to cross-modal training, and enabled deployment on increasingly modest devices.

3. Vision-Language Models

While transformers transformed visual backbones, their integration with language models created a new family of systems: vision-language models (VLMs) and, more broadly, multimodal large language models (MLLMs) [6–9]. These architectures extend the generative and reasoning capabilities of LLMs by conditioning them on visual inputs, ranging from static images to video and, in some cases, audio. Different architectural strategies have emerged, from which the simplest is the dual-encoder design, exemplified by CLIP, where a vision encoder and a text encoder project inputs into a shared embedding space. This structure excels at retrieval and zero-shot classification but cannot, on its own, generate fluent descriptions or answers. A more expressive approach is the cross-attention fusion design, seen in Flamingo-style models, where visual tokens are passed to a frozen or lightly tuned language model via specialized cross-attention modules. This allows the LLM to remain largely intact while still grounding its generation in visual evidence. Finally, unified encoder-decoder designs interleave image and text tokens in a single sequence, allowing for tight modality coupling and more coherent multimodal reasoning, albeit at higher computational cost. Across all variants, a critical component is the projector, sometimes called a connector or Q-Former. This module maps high-dimensional vision features into the token space expected by the language model. The projector's dimensionality and the number of tokens it outputs have direct consequences for local deployment, which is the doubling the sequence length at the input inflates the memory required to store key-value caches during decoding. Thus, for constrained devices, projector efficiency is often more decisive than the parameter count of the underlying LLM.

The multimodal landscape has expanded further with extensions to temporal and auditory domains. Video-aware models introduce temporal pooling, frame sampling, or sparse attention to manage long sequences, while speech-augmented systems integrate automatic speech recognition and text-to-speech to support interactive dialogue. These extensions increase memory and latency requirements, intensifying the challenge for local deployment. From a local-first perspective, vision-language models embody both opportunity and constraint. They unlock applications such as private document analysis, edge robotics, and on-device assistive tools, but their practical use hinges on architectural frugality. Slimmed-down encoders, reduced projector outputs, and careful token budgeting are necessary to bring multimodal reasoning into the range of consumer laptops, mobile hardware, and edge accelerators.

Deploying multimodal transformers locally, whether on a workstation, laptop, or edge device, requires a coherent view of the entire inference stack. This stack extends from the model architecture itself, through the artifact and runtime layers, down to the physical accelerator. Each layer introduces distinct constraints on memory, throughput, and reproducibility, and together they define what “local-first” truly entails. At the foundation lies the model layer, which determines the computational and representational budget. The parameter count of the base language model sets the broad memory envelope, but in multimodal systems the vision encoder and projector exert equal influence. The choice of encoder, whether a compact CLIP, SigLIP, or EVA variant, affects both accuracy and the number of visual tokens fed downstream. The projector, meanwhile, translates visual embeddings into the token space of the LLM. Its dimensionality and sequence length dictate the growth of the key-value cache during generation, and thus the achievable context window under fixed VRAM. Small architectural decisions at this stage, such as capping the number of image patches or compressing feature maps, can make the difference between smooth interactivity and repeated out-of-memory crashes.

Above the model layer sits the artifact layer, which defines how the learned parameters are stored, quantized, and exchanged. The ecosystem has largely converged on a few interoperable standards. Safetensors provides a framework-native, secure serialization format; GGUF, the de facto standard for the llama.cpp runtime, enables execution of 4- and 5-bit quantized weights, and ONNX graphs favor portability across toolchains. Quantization techniques, from FP16 down to 4-bit GPTQ and AWQ variants, compress model weights by an order of magnitude while retaining acceptable fidelity. When fine-tuning is required, lightweight adapters such as LoRA [13] and QLoRA [14] permit task-specific updates without modifying the base model, a crucial feature for privacy-preserving and resource-limited contexts.

The runtime layer transforms these artifacts into executable computation. For single-device or embedded deployment, llama.cpp dominates, combining cross-platform C++ efficiency with broad quantization support and native backends for CUDA and Apple Metal. MLX, Apple’s tensor runtime, provides similar functionality within macOS environments, exploiting unified memory on M-series chips. On server-class GPUs, vLLM and Text Generation Inference (TGI) deliver high-throughput decoding through paged key–value caches and optimized scheduling. TensorRT-LLM extends this optimization to NVIDIA hardware at the kernel level, achieving sub-millisecond token latencies on Jetson and desktop GPUs. For CPU and iGPU execution,

OpenVINO and ONNX Runtime offer mixed-precision and INT8 kernels that make local inference feasible on commodity hardware. Underlying these runtimes are diverse accelerator backends. NVIDIA’s CUDA remains the reference standard for performance and tooling maturity, supporting every major quantization format and runtime. Apple’s Metal backend prioritizes efficiency per watt, a decisive advantage for laptops and mobile devices where power envelopes are tight. AMD’s ROCm stack, though newer, now provides credible support for open-source frameworks and is beginning to host quantized workloads. Intel’s CPUs and integrated GPUs, often overlooked, can reach competitive throughput with OpenVINO’s INT8 optimizations, offering a low-cost, privacy-preserving path for enterprise deployments.

Across all layers, evaluation must move beyond accuracy to include operational metrics that describe local performance. Memory consumption is the first constraint, encompassing base weights, vision encoders, projectors, and the cumulative size of key-value caches. Latency and throughput follow, measured in time-to-first-token and sustained tokens per second under realistic workloads. Energy use and thermal stability become relevant for mobile and edge platforms. Finally, licensing cannot be ignored, open weights, permissive redistribution, and commercial usage rights determine whether a model can be deployed at all in local or offline settings.

This taxonomy emphasizes that local deployment is not a single technical problem but a systems-level negotiation between architecture, representation, software infrastructure, and hardware. Each layer sets boundaries for the next, and optimization succeeds only when viewed holistically. In the remainder of this paper, this framework serves as a reference. Section 4 catalogues open multimodal models that fit within these constraints, while later sections examine the compression, fine-tuning, and runtime strategies that transform global-scale research systems into practical, self-contained local applications.

4. Local Multimodal Models

The rapid expansion of open-weight multimodal models has transformed what is possible on local hardware. Only a few years ago, image–language reasoning required cloud-scale GPUs. Today, sub-10 GB devices can host fully functional models capable of visual question answering, captioning, and grounded reasoning. This shift reflects a convergence of three trends: the modularity of transformer architectures, open-source dissemination of pretrained

weights, and aggressive efficiency gains through quantization and runtime optimization. To make sense of this emerging ecosystem, we adopt the local-first taxonomy introduced earlier. Each model can be understood along a small set of dimensions, its root language model lineage (which determines its generative capacity and context window), its vision encoder (which governs perceptual fidelity and computational cost), its modalities (image, video, audio, or text), and its deployment readiness (quantization support, runtime compatibility, and licensing).

4.1 Compact and Lightweight Models

The smallest multimodal models demonstrate how far compression and architectural pruning have advanced. Models such as DeepSeek-VL 1.3 B, MiniCPM-V 2 B, and Qwen 2.5-VL 3 B integrate visual reasoning into LLMs that can easily fit within 8 GB of memory when quantized to 4-bit. DeepSeek-VL extends the DeepSeek-LLM backbone with a lightweight visual encoder and simple cross-attention fusion, producing competitive visual understanding on minimal hardware. MiniCPM-V, developed by OpenBMB, builds on the MiniCPM-2B foundation and employs SigLIP as a vision front-end, enabling strong captioning and visual question answering in an exceptionally small footprint. Its sibling, MiniCPM-o 2.6, adds video and audio streams, showcasing the feasibility of true multimodality in models below two billion parameters.

Another representative example, Qwen 2.5-VL 3 B, from Alibaba's Qwen.AI initiative, combines a three-billion-parameter language model with a unified encoder capable of handling both images and short video clips. Microsoft's Phi-4 Multimodal 4 B extends the trend by incorporating a speech interface alongside visual and textual inputs, highlighting a shift toward conversational perception systems. All these models run natively through the llama.cpp or GGUF toolchain, achieving real-time inference on laptops or compact GPUs, and they can be fine-tuned with LoRA adapters for domain-specific applications such as industrial inspection or document parsing.

Table 1: Compact and Lightweight Multimodal Models (≤ 4 B parameters).

Model	Root LLM	Modalities	Highlights
DeepSeek-VL 1.3 B	DeepSeek-LLM 1.3 B	Image + Text	Lightweight visual encoder with cross-attention; efficient at ≤ 8 GB VRAM; supports llama.cpp / GGUF.
MiniCPM-V 2.0 (2 B)	MiniCPM-2B	Image + Text	SigLIP vision front-end; optimized for local captioning and VQA; open weights.
MiniCPM-o 2.6 (≈ 2 B)	MiniCPM-2B	Image + Video + Audio + Text	Multi-stream modality support in sub-2B model; real-time inference on laptops.
Qwen 2.5-VL 3 B	Qwen 2.5 3 B	Image + Video + Text	Unified encoder–decoder; strong visual reasoning; maintained by Alibaba (Qwen.AI).

4.2 Mid-Scale Models

Mid-scale architectures represent a balance between expressive reasoning and feasible local deployment. The LLaVA v1.6 Mistral 7 B model fuses the open Mistral-7B-Instruct backbone with a SigLIP visual encoder, yielding strong instruction-following behavior in multimodal dialogue. DeepSeek-VL 7 B follows a similar principle, integrating an optimized vision connector for more grounded reasoning. Video-capable systems such as Video-LLaVA and LLaVA-NeXT-Video extend this architecture to temporal domains, using sparse frame sampling and temporal pooling to maintain responsiveness without exploding memory requirements.

Other notable entries include Idefics 2 (8 B) from Hugging Face, which adopts a Flamingo-style cross-attention decoder for fluent image-text generation, and Fuyu 8 B from Adept AI, designed around a unified encoder–decoder framework that operates directly on raw pixels rather than discrete patches. The MiniCPM-V 2.6 (8 B) model combines the Qwen-2 backbone

with SigLIP embeddings to deliver high performance across image, video, and text tasks while retaining strong quantization support. Finally, Google DeepMind’s Gemma 3 family, spanning 4 B, 12 B, and 27 B variants, embodies the growing integration of efficient transformer training with permissive licensing and open release.

Table 2: Representative Mid-Scale Multimodal Models ($\approx 7\text{--}13$ B parameters).

Model	Root LLM	Modalities	Highlights
LLaVA-v1.6 Mistral 7 B	Mistral-7B-Instruct	Image + Text	SigLIP encoder; instruction-tuned multimodal assistant; quantized builds in GGUF.
DeepSeek-VL 7 B	DeepSeek-LLM 7 B	Image + Text	Optimized vision connector; fluent multimodal dialogue; efficient 4-bit inference.
Video-LLaVA / LLaVA-NeXT- Video 7 B	Vicuna 7B / Mistral 7B	Video + Image + Text	Temporal pooling for efficient video reasoning; open on Hugging Face.
Idetics 2 (8 B)	Flamingo-style decoder	Image + Text	Cross-attention fusion; strong captioning performance; Hugging Face release.

4.3 Large and Specialized Models

Although the frontier of open research continues to push toward ever-larger systems, only a subset remains realistic for local inference. Qwen 2.5-VL 14 B, Gemma 3 (27 B), and Kimi K2 (MoE 1 T) exemplify this scale. When quantized to 4-bit precision, such models can operate on workstations with 48-64 GB of VRAM, though they are less suited to mobile devices. Mixture-of-Experts (MoE) architectures, like Kimi K2, introduce sparsity at inference time, activating only a subset of experts per token and thus mitigating some of the computational overhead. While not yet fully portable, these architectures point toward a near-future equilibrium: very large base models distilled or sparsified into smaller, deployable sub-networks for edge execution.

Table 3: Large and Specialized Multimodal Models (> 13 B parameters).

Model	Root LLM	Modalities	Highlights
Qwen 2.5-VL 14 B	Qwen 2.5 14 B	Image + Video + Text	Large-context multimodal reasoning; runs at 4-bit on 48 GB GPUs.
Gemma 3 (27 B)	Gemma 3	Image + Text	High-capacity ViT encoder; open licensing for research; scalable quantization.
Kimi K2 (MoE 1 T)	Mistral-7B-Instruct (MoE)	Image + Text	Mixture-of-Experts design; sparse activation for efficiency; Moonshot AI release.

Across these scales, lineage coherence matters. Many current multimodal systems inherit from a small set of textual foundations, Mistral 7B, Qwen 2.x, MiniCPM-2B, and Phi-4, each bringing its own tokenizer, context handling, and quantization behavior. On the visual side, SigLIP and EVA encoders are rapidly replacing earlier CLIP variants due to improved data efficiency and robustness at lower resolutions. These pairings define not only the attainable accuracy but also runtime compatibility, since different training toolchains emit models in distinct formats.

From a deployment standpoint, several trends stand out. Quantization has evolved from a secondary optimization into an enabling technology; without 8-bit or 4-bit compression, even mid-scale models exceed consumer VRAM budgets. Adapter-based fine-tuning now provides a lightweight, repeatable workflow, allowing developers to specialize models on modest hardware. Moreover, the barrier between text-only and multimodal models is eroding, since vision connectors can often be attached to existing LLMs with limited retraining, producing effective multimodal systems with minimal additional cost.

5. Building and Sustaining Local-First Multimodal Intelligence

The emergence of locally runnable multimodal language models represents more than a technical achievement; it signals a restructuring of the AI pipeline itself. To sustain this shift, the underlying components, model efficiency, adaptation, runtime infrastructure, evaluation, and governance must evolve in concert. This section integrates the principles that, in traditional surveys, would appear as separate technical chapters. Here, they are treated as interdependent threads within a single local-first ecosystem.

Efficiency and compression form the foundation of this ecosystem. Local deployment depends on the ability to compress global-scale models into workable forms without losing their reasoning fidelity. Quantization techniques such as FP16, INT8, and 4-bit GPTQ or AWQ reduce memory usage by up to an order of magnitude while preserving fluency and perception. Structured sparsity, low-rank decomposition, and distillation extend these gains by trimming redundant parameters and transferring knowledge from large teachers to smaller students. In multimodal contexts, compression affects not only the text backbone but also vision encoders and projectors, where token reduction and adaptive pooling are critical to maintaining speed. The cumulative result is that systems once confined to data centers can now operate interactively on desktop GPUs, Apple Silicon, and even Jetson-class embedded devices.

Beyond raw efficiency lies adaptation, the process of tailoring general-purpose models to specific domains, tasks, or users. Adapter-based fine-tuning methods such as LoRA and QLoRA make this feasible on quantized bases, enabling small, domain-specific updates without retraining the entire network. Instruction-tuning with visual datasets aligns model behavior to multimodal tasks, while synthetic supervision and curated captioning help maintain alignment when real data are scarce. For local environments, this capacity is essential: it allows models to incorporate private data directly on-device, preserving confidentiality and enabling continual refinement without cloud retraining.



Figure 3: The Nvidia Jetson Orin Nano 8GB. This is currently the smallest embedded device that is able to run LLMs locally.

Complementing these model-level methods is a maturing tooling and runtime infrastructure. Portability and openness define this layer. Frameworks such as llama.cpp and GGUF enable CPU- and GPU-efficient execution; TensorRT-LLM and vLLM optimize GPU throughput. OpenVINO and MLX extend reach to Intel and Apple platforms. Together they constitute a cross-hardware runtime fabric that allows the same model artifact to run across heterogeneous devices. File formats like safetensors, ONNX, and GGUF provide stability and safety, while converter and quantizer toolchains, bitsandbytes, GPTQ-for-LLaMA, AWQ, ensure that research checkpoints can be transformed into deployable binaries. For local AI, this standardization is not auxiliary but existential, since without a reproducible toolchain, decentralization collapses into incompatibility.

To measure genuine progress, the field must rethink evaluation for local conditions. Accuracy alone is insufficient when resource budgets dominate practical feasibility. New metrics, VRAM footprint, time-to-first-token latency, sustained decoding rate, and power efficiency, capture the realities of local inference. Benchmarking must extend beyond canonical vision tasks to integrated vision-language evaluation, including captioning, OCR, VQA, and grounding. For reproducibility, standard test conditions, prompt resolution, quantization level, hardware type, should accompany reported metrics. Such grounded evaluation transforms performance from a claim into a measurable, transferable property. Efficiency and evaluation together culminate in applied use cases, where local models demonstrate their societal and industrial relevance. In

robotics, compact Vision–Language–Action loops execute perception and control onboard, minimizing latency and enabling autonomy in unconnected environments. In edge analytics, smart cameras perform defect detection or environmental monitoring without external computation. In assistive and privacy-sensitive contexts, on-device OCR and multimodal Q&A preserve user data integrity. These applications illustrate that local deployment is not a compromise, it is an expansion of capability, merging efficiency with independence.

Yet the local-first trajectory introduces new challenges. Precision loss under aggressive quantization can amplify multimodal hallucinations and degrade grounding accuracy. Data scarcity, especially for temporal and 3D modalities, limits generalization. Long-context memory remains constrained by KV-cache size and bandwidth. Legal and ethical concerns persist as open weights intersect with ambiguous licensing and the potential for misuse. Finally, reproducibility across toolchains remains difficult: a model quantized for one runtime may behave differently on another. Addressing these issues is vital if local AI is to match the transparency and accountability expected of scientific systems.

The next phase of local-first AI will center on adaptability and interoperability. Lightweight, unified Vision-Language-Action architectures promise continuous on-device learning, where models evolve with user interaction. Federated and privacy-preserving updates could synchronize local improvements without exposing raw data. Emerging sequence models with structured-state memory and streaming attention will handle long video and multimodal sequences efficiently, bridging the gap between perception and cognition. Equally transformative will be the standardization of adapter interfaces and weight formats, allowing the community to exchange fine-tuned modules across different backbones and runtimes. In parallel, the creation of self-hosted leaderboards and open evaluation suites will make local performance a measurable, evolving benchmark rather than an anecdote.

Sustaining this progress requires a culture of openness. Shared glossaries, community-maintained repositories, and continuously updated collections of model cards and quantization recipes ensure that knowledge does not stagnate. Living documentation and reproducible repositories serve as the infrastructure of scientific continuity, an especially important principle when innovation is distributed across thousands of local nodes rather than concentrated in a few cloud providers.

These intertwined dimensions, efficiency, adaptation, infrastructure, evaluation, and openness, define the modern landscape of local multimodal intelligence. They turn the once-exclusive field of large-scale vision-language modeling into a participatory, reproducible, and privacy-conscious practice. In this way, the local-first paradigm is not a retreat from scale but an evolution toward accessibility and agency, bringing the tools of multimodal reasoning into the hands of individuals and institutions everywhere.

6. Conclusion

The evolution from convolutional networks to transformer-based multimodal language models has fundamentally redefined both the scale and scope of computer vision. Where once perception and reasoning were separate domains, processed by distinct algorithms and hardware tiers, they now coexist within unified architectures capable of understanding images, video, and text through shared representations. Yet the most profound shift is not architectural but infrastructural, which is that intelligence that once required remote data centers is now migrating onto local devices.

This transition from cloud-exclusive to local-first AI is made possible by a confluence of open research and practical engineering. Transformers provided a universal architecture. The parts of these, quantization, pruning, and distillation made them lightweight, open runtimes and file formats made them portable, and adapter-based fine-tuning made them personal. These advances democratize multimodal reasoning, enabling responsive, private, and energy-efficient systems to operate directly on consumer and embedded hardware.

In essence, the transformer era has entered its second phase, one defined not by bigger models but by broader reach. The task now is to ensure that this reach remains open, efficient, and accountable. If the first wave of deep learning proved that machines can learn to see, the emerging wave of local-first multimodal AI proves that they can do so anywhere, privately, efficiently, and on our own terms.

7. Acknowledgement

The research was supported by the European Union within the framework of the National Laboratory for Artificial Intelligence. (RRF-2.3.1-21-2022-00004)

Bibliography

- [1.] A. Dosovitskiy, L. Beyer, A. Kolesnikov, et al.: *An Image is Worth 16×16 Words: Transformers for Image Recognition at Scale*, International Conference on Learning Representations (2021).
- [2.] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, S. Zagoruyko: *End-to-End Object Detection with Transformers*, European Conference on Computer Vision (2020).
- [3.] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, B. Guo: *Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows*, IEEE/CVF International Conference on Computer Vision (2021).
- [4.] K. He, X. Chen, S. Xie, Y. Li, P. Dollár, R. Girshick: *Masked Autoencoders Are Scalable Vision Learners*, IEEE/CVF Conference on Computer Vision and Pattern Recognition (2022).
- [5.] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, et al.: *Learning Transferable Visual Models from Natural Language Supervision*, Proceedings of the 38th International Conference on Machine Learning (2021).
- [6.] J.-B. Alayrac, et al.: *Flamingo: A Visual Language Model for Few-Shot Learning*, Advances in Neural Information Processing Systems (2022).
- [7.] J. Li, D. Li, S. Savarese, S. Hoi: *BLIP-2: Bootstrapping Language-Image Pre-training with Frozen Image Encoders and Large Language Models*, Proceedings of the 40th International Conference on Machine Learning (2023).
- [8.] X. Zhai, B. Mustafa, A. Kolesnikov, L. Beyer: *Sigmoid Loss for Language-Image Pre-Training (SigLIP)*, IEEE/CVF International Conference on Computer Vision (2023).
- [9.] H. Liu, C. Li, Q. Wu, Y. J. Lee: *Visual Instruction Tuning (LLaVA)*, arXiv:2304.08485 (2023).
- [10.] G. Xiao, J. Lin, M. Seznec, H. Wu, J. Demouth, S. Han: *SmoothQuant: Accurate and Efficient Post-Training Quantization for Large Language Models*, Proceedings of the 40th International Conference on Machine Learning (2023).
- [11.] E. Frantar, S. Ashkboos, T. Hoefler, D. Alistarh: *GPTQ: Accurate Post-Training Quantization for Generative Pre-trained Transformers*, arXiv:2210.17323 (2022).
- [12.] J. Lin, J. Tang, H. Tang, S. Yang, W.-M. Chen, W.-C. Wang, G. Xiao, X. Dang, C. Gan, S. Han: *AWQ: Activation-aware Weight Quantization for LLM Compression and Acceleration*, arXiv:2306.00978 (2023).
- [13.] E. J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, W. Chen: *LoRA: Low-Rank Adaptation of Large Language Models*, International Conference on Learning Representations (2022).

[14.] T. Dettmers, A. Pagnoni, A. Holtzman, L. Zettlemoyer: *QLoRA: Efficient Finetuning of Quantized LLMs*, Advances in Neural Information Processing Systems (2023).

[15.] W. Kwon, et al.: *Efficient Memory Management for Large Language Model Serving with PagedAttention (vLLM)*, arXiv:2309.06180 (2023).

Vizuális Odometria Algoritmusok Hatékonysági Vizsgálata

Efficiency Evaluation of Visual Odometry Algorithms

Nagy Armand^a, Hollósi János^b

^aSzéchenyi István Egyetem, Járműipari Kutatóközpont

nagy.armand@hallgato.sze.hu

^bSzéchenyi István Egyetem, Járműipari Kutatóközpont

hollosi.janos@sze.hu

Absztrakt

A vizuális odometria (VO) kulcsfontosságú az autonóm rendszerekben, mivel lehetővé teszi a mozgás becslését kizárólag kameraképek alapján, külső szenzorok nélkül. Jelen cikkben – egy közösen végzett kutatás keretében – különböző VO-algoritmusok kerülnek összehasonlításra a KITTI és EuRoC benchmark adatbázisok felhasználásával. A cél a legpontosabb konfigurációk azonosítása monokuláris kamerabemenet mellett. A kiértékelés során olyan teljesítménymutatókat alkalmaztunk, mint az abszolút pályahiba (ATE), a relatív orientációs hiba (RPE), valamint statisztikai hibák: MAE, MSE, MRE és FDE. A bemutatás kizárólag a legjobban teljesítő konfigurációkra terjed ki, mind kategóriánként, mind összesített rangsor alapján. Eredményeink célja, hogy iránymutatást nyújtsanak a valós idejű alkalmazásokhoz leginkább alkalmas VO-megoldások kiválasztásához, különös tekintettel a pontosság és stabilitás közötti kompromisszumokra.

Kulcsszavak: vizuális odometria, trajektória becslés, összehasonlító elemzés

Abstract

Visual odometry (VO) is a key component in autonomous systems, enabling motion estimation solely from camera images without relying on external sensors. In this paper, conducted as part of a collaborative research, various VO algorithms are compared using the KITTI and EuRoC benchmark datasets. The primary objective is to identify the most accurate configurations for monocular camera input. The evaluation is based on six standard performance metrics: Absolute Trajectory Error (ATE), Relative Pose Error (RPE), Mean Absolute Error (MAE), Mean Squared Error (MSE), Mean Relative Error (MRE), and Final Displacement Error (FDE). The analysis focuses exclusively on the top-performing configurations in both individual categories and overall rankings. The results aim to support the selection of VO solutions that are most suitable for real-time applications, with particular emphasis on the trade-offs between accuracy and stability.

Keywords: visual odometry, trajectory estimation, comparative analysis

1. Bevezetés

A vizuális odometria (VO) számítógépes látásalapú technika, amely képsorozatok elemzése révén becsli meg egy mozgó kamera pozícióját és orientációját. Különösen hasznos olyan helyeken, ahol a GPS nem elérhető vagy megbízható, például beltéren, alagutakban vagy víz alatt [1]. Bár a technológia alapjai az 1980-as évekből származnak, a valós idejű alkalmazásokat csak a 2000-es évektől tette lehetővé a hardverek fejlődése [2, 3]. A VO algoritmusok három fő kategóriába sorolhatók: klasszikus, mélytanulás-alapú és hibrid megközelítések. A klasszikus algoritmusok gyorsak és determinisztikus működésűek, ugyanakkor érzékenyek a képminőségre és a környezeti hatásokra. A tanulásalapú eljárások robusztusabbak, de erőforrás-igényesek. A hibrid modellek célja e két megközelítés előnyeinek ötvözése. Jelen tanulmány célja a VO konfigurációk teljesítményének átfogó értékelése pontosság, robusztusság, számítási sebesség és erőforrásigény szempontjából. Az azonos körülmények között végzett tesztek biztosítják az összehasonlíthatóságot és gyakorlati relevanciát [5].

2. Kapcsolódó munkák

Az utóbbi években nőtt az érdeklődés a vizuális odometria (VO) rendszerek összehasonlítása iránt, amit a SLAM technológia fejlődése és az autonóm rendszerek – például robotok, drónok, járművek – terjedése ösztönzött [1,4,6]. A korai VO-megközelítések, mint a Lucas–Kanade (LK) algoritmus [7], optikai áramláson alapultak, és máig sok alkalmazásban jelen vannak. A modern VO algoritmusok három fő csoportba sorolhatók: klasszikus, tanulásalapú és hibrid módszerek. A klasszikus megközelítések robusztus képi kulcspontdetektorokat (SIFT [8], ORB [9], BRISK [10], AKAZE [11], KAZE [12]) és leírókat (BRIEF [13], FREAK [14], RootSIFT [15], BEBLID [16]) alkalmaznak, jellemzően Brute-Force [17] vagy FLANN [18] párosítással. A tanulásalapú módszerek – például SuperPoint [19], D2-Net [20], ALIKED [21], R2D2 [22], LF-Net [23], DISK [24], DELF [25] – konvolúciós neurális hálózatokon alapulnak, és jól működnek összetett környezetekben. A legújabb fejlesztések közé tartozik a LightGlue [26], amely transformer-alapú, kontextusérzékeny párosítást végez. A hibrid megközelítések klasszikus detektorokat kombinálnak tanult leírókkal (pl. ORB2_HardNet [27], ORB2_SOSNet [28]) és fejlett párosítókkal, mint a LightGlue. Mindezek ellenére kevés egységes, összehasonlító értékelés érhető el a különböző konfigurációkról. Ez indokolja jelen

kutatásunkat, amely a PySLAM keretrendszerben vizsgálja a klasszikus és korszerű módszerek viselkedését valós körülmények között [5].

3. Alkalmazott módszerek

A vizsgálat minden elemét a PySLAM keretrendszerrel végeztük [5], amely moduláris felépítése révén rugalmasan kombinálhatóvá teszi a detektorokat, leírókat és párosítókat. Minden VO-konfiguráció ezen a platformon épült fel, az egyes futások során elmentett pályákat pedig a ground truth adatokkal vetettük össze. A teljesítménymutatók kiszámításához külön Python-alapú értékelő scriptek készültek.

3.1 Konfigurációk

Összesen 36 vizuális odometria konfigurációt értékeltünk, amelyeket alkalmazott módszertanuk, feldolgozási elveik és architektúrájuk szerint négy fő kategóriába soroltunk. A klasszikus megközelítések kézzel tervezett, széles körben alkalmazott képfeldolgozó detektorokat (SIFT [8], ORB [9], BRISK [10], AKAZE [11]) és leírókat (BRIEF [13], FREAK [14], RootSIFT [15], BEBLID [16]) alkalmaznak a jellemzők kinyerésére és reprezentációjára. A kulcspontok párosítása ezekben a konfigurációkban Brute-Force [17] vagy FLANN [18] algoritmusokkal történik, amelyek a leíróvektorok távolsága alapján végzik az illesztést. A tanulásalapú megközelítések mély neurális hálózatokon alapulnak, amelyek a detektálás és leírás együttes, adatvezérelt optimalizálását valósítják meg, például SuperPoint [19], D2-Net [20], ALIKED [21], R2D2 [22], LF-Net [23], DISK [24], DELF [25]. A hibrid módszerek klasszikus detektorokat kombinálnak tanult leírókkal, így például ORB2_HardNet [27] és ORB2_SOSNet [28] konfigurációk esetében. A LightGlue-alapú konfigurációk a transzformer-architektúrájú LightGlue [26] párosítót alkalmazzák, amely globális kontextus alapján optimalizálja a megfeleltetéseket, és klasszikus vagy tanult leírókkal egyaránt kompatibilis. Mivel működési elve alapvetően eltér a hagyományos deskriptor-párosítástól, önálló kategóriaként kezeltük. A konfigurációkat egységes névkonvenció szerint határoztuk meg, biztosítva a reprodukálhatóságot és az összehasonlíthatóságot. Az összeállított konfigurációk széles skálát fednek le mind az alkalmazott detektálási-stratégiák, mind a leírók reprezentációs képessége és a párosítási algoritmusok komplexitása tekintetében.

3.2 Adathalmazok

A vizsgálat során két széles körben használt, nyilvánosan elérhető benchmark adathalmazt alkalmaztunk, amelyek eltérő környezeti feltételek mellett teszik lehetővé a vizuális odometria algoritmusok teljesítményének értékelését és összehasonlítását.

3.2.1 KITTI

A KITTI adathalmaz kültéri közúti forgalmi eseményeket tartalmaz, amelyeket egy személygépkocsira szerelt érzékelőrendszerrel rögzítettek. A képadatokat egy sztereó kamera biztosítja, miközben a jármű különböző városi, elővárosi és autópályás szakaszokon közlekedik. A ground truth pályaadatokat nagy pontosságú GPS és inerciális mérőrendszer fúziójával állították elő, lehetővé téve a pozíció és orientáció referenciaként történő használatát a vizuális odometria algoritmusok teljesítményének objektív értékeléséhez [29, 30].



16. ábra: Képkocka a KITTI adathalmazból



17. ábra: Képkocka a KITTI adathalmazból

3.2.2 EuRoC

Az EuRoC adathalmaz beltéri környezetekben, ipari és laboratóriumi szcenáriókban készült felvételeket tartalmaz. A felvételeket egy mikronégyzetes drónra szerelt sztereó kamera és IMU érzékelők rögzítették különböző mozgásminták mellett. A ground truth referencia-pályákat

nagy pontosságú mozgáskövető rendszer (motion capture) szolgáltatta, biztosítva a milliméteres szintű pozíció- és orientációbeli pontosságot [31].



18. ábra Képkockák az EuRoC adathalmazból

3.3 Értékelési metrikák

A konfigurációk teljesítményének jellemzésére hat elterjedt metrikát használtunk [6, 10]:

1. **Absolute Trajectory Error** – A teljes pálya menti eltérés:

$$ATE = \sqrt{\frac{1}{N} \sum_{i=1}^N \|p_i^{\text{est}} - p_i^{\text{gt}}\|^2} \quad (1)$$

2. **Mean Absolute Error** – Az átlagos abszolút hiba:

$$MAE = \frac{1}{N} \sum_{i=1}^N \|p_i^{\text{est}} - p_i^{\text{gt}}\| \quad (2)$$

3. **Mean Squared Error** – A nagyobb hibákat jobban súlyozza:

$$MSE = \frac{1}{N} \sum_{i=1}^N \|p_i^{\text{est}} - p_i^{\text{gt}}\|^2 \quad (3)$$

4. **Mean Relative Error** – A hibát a megtett távolsághoz viszonyítva normalizálja:

$$MRE = \frac{1}{N} \sum_{i=1}^N \frac{\|p_i^{\text{est}} - p_i^{\text{gt}}\|}{\|p_i^{\text{gt}}\|} \quad (4)$$

5. **Relative Pose Error** – Két egymást követő pozíció közötti eltérés:

$$RPE = \frac{1}{N-1} \sum_{i=1}^{N-1} \|(p_{i+1}^{\text{est}} - p_i^{\text{est}}) - (p_{i+1}^{\text{gt}} - p_i^{\text{gt}})\| \quad (5)$$

6. **Final Displacement Error** – Az utolsó becsült és valós pozíció közötti különbség:

$$FDE = \|p_N^{\text{est}} - p_N^{\text{gt}}\| \quad (6)$$

A fenti képletekben:

- p_i^{est} : a vizuális odometria algoritmus által becsült pozíció az i -edik időpillanatban.
- p_i^{gt} : a ground truth (valós, referenciaként használt) pozíció ugyanazon i -edik időpillanatban
- N : a teljes vizsgált pozíciók száma (azaz az összehasonlított időpillanatok száma).
- p_N^{est} , p_N^{gt} : a becsült, illetve valós pozíció az utolsó (N -edik) időpillanatban.

A metrikák értékeit külső Python-szkriptek számították ki, az eredmények CSV formátumban kerültek mentésre statisztikai és grafikus feldolgozáshoz.

4. Eredmények

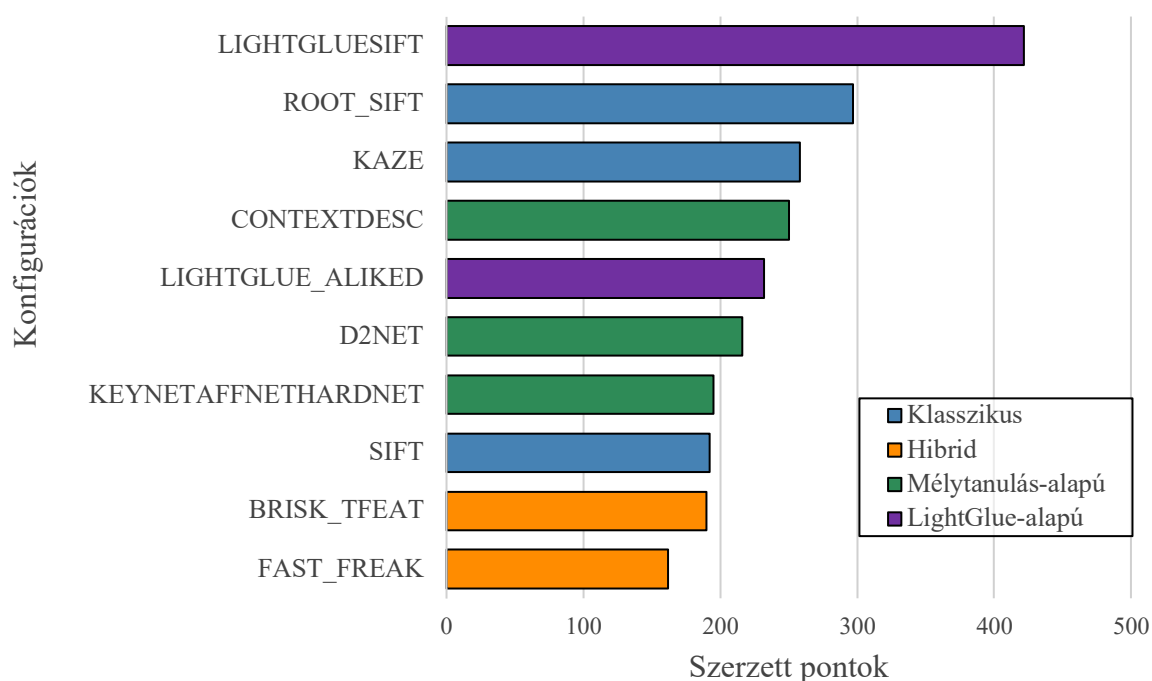
Ebben a fejezetben a különböző vizuális odometria (VO) konfigurációk kiértékelésének legfontosabb eredményeit mutatjuk be. Az értékelés során a pályabecslés pontosságát, a számítási teljesítményt és az algoritmusok stabilitását vizsgáltuk egységes tesztkörnyezetben, azonos beállítások és feldolgozási lépések alkalmazásával. A kiértékeléshez azok a több szempontot figyelembe vevő metrikák kerültek alkalmazásra, amelyek részletes képet adnak az egyes konfigurációk gyakorlati alkalmazhatóságáról. Az eredmények bemutatása a legjobb teljesítményt nyújtó konfigurációkra koncentrál.

4.1 Pontosság összehasonlítása

A vizsgálat során elsősorban azt értékeltük, hogy az egyes konfigurációk milyen pontossággal képesek rekonstruálni a kamera mozgását a ground truth pályákhoz viszonyítva. Az összehasonlításhoz hat meghatározott teljesítménymutatót alkalmaztunk: ATE, RPE, MAE, MSE, MRE és FDE [6]. Minden konfiguráció esetében a részmetrikákon elért helyezések alapján pontozás történt, ahol az első helyezett 15, a tizenötödik helyezett pedig 1 pontot kapott, majd ezek összege adta az összesített pontszámot. Az így meghatározott eredmények az 5. ábrán láthatók. Az eredmények alapján jól megfigyelhető, hogy a különböző konfigurációk között jelentős teljesítménykülönbségek alakultak ki. A legjobb eredményt a LIGHTGLUE_SIFT konfiguráció érte el, amelyben a klasszikus SIFT detektor kombinálódik a modern LightGlue [26] módszerrel. E konfiguráció stabilan teljesített mindkét

tesztadathalmazon és különösen alacsony hibát ért el az összes metrikában. A klasszikus megközelítések közül több konfiguráció is a felső rangsorba került, így a ROOT_SIFT, KAZE és SIFT, amelyek a stabil kulcspontdetektálásnak és robusztus leírásnak köszönhetően megbízható pályabecslést eredményeztek. A tanulásalapú módszerek között a CONTEXTDESC [32], a D2NET [20] és a KEYNETAFFNETHARDNET [26, 32] is kiemelkedő pontosságot értek el, bizonyítva, hogy a mélytanulás-alapú reprezentációk hatékonyan képesek kezelni a komplex képi jellemzőket. A hibrid megoldások közül a BRISK_TFEAT [10, 34] és a FAST_FREAK [10, 14] konfigurációk kiegyensúlyozott, stabil teljesítményt nyújtottak.

Összefoglalva elmondható, hogy a legjobb konfigurációkat olyan megközelítések alkotják, amelyek képesek hatékonyan egyesíteni a stabil detektálás, a megfelelő leíróreprezentáció és a pontos pozícióbecslés előnyeit.



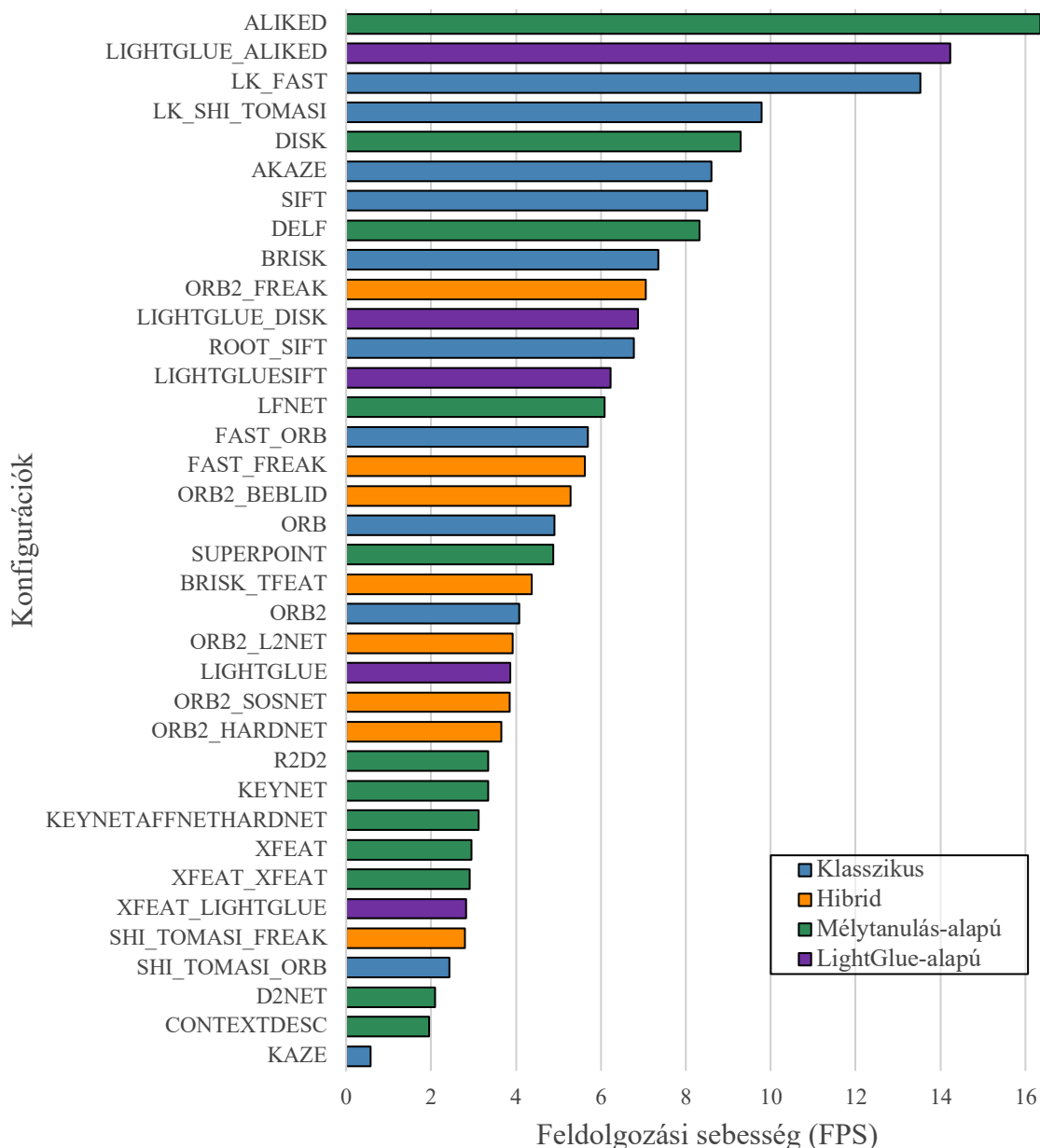
19. ábra: A 10 legpontosabb konfiguráció

4.2 Feldolgozási sebesség

A konfigurációk futásidejének értékelése során a képkockasebesség (FPS) került vizsgálatra, mivel ez kulcsfontosságú tényező a valós idejű alkalmazhatóság és a gyakorlati implementációk szempontjából. A mérések a PySLAM keretrendszerben, külön optimalizáció alkalmazása

nélkül kerültek végrehajtásra [5]. Az eredményeket a 6. ábra foglalja össze, amely a konfigurációk feldolgozási sebességét mutatja be a Brute-Force párosítót alkalmazó változatok esetén. A legmagasabb FPS-t az ALIKED érte el, ezt követték a LIGHTGLUE_ALIKED, az LK_FAST és az SHI_TOMASI_FREAK. A klasszikus konfigurációk közül az AKAZE és a SIFT is viszonylag magas feldolgozási sebességet biztosítottak.

Fontos megjegyezni, hogy az egyes konfigurációk feldolgozási idejét jelentős mértékben befolyásolja a detektorok és leírók számítási komplexitása, valamint a kulcspontok számossága is. Emellett a választott párosító algoritmusok és implementációs optimalizációk szintén érdemben módosíthatják a futásidőt. Mivel a PySLAM nem futásidő-optimalizált környezetben működik, az itt bemutatott FPS értékek elsősorban a konfigurációk egymáshoz viszonyított relatív sebességét tükrözik, de nem szolgálnak közvetlen becslésként valós idejű rendszerekre vonatkozóan. A teljesítményértékelés ennek megfelelően inkább az algoritmusok számítási igényének összehasonlító vizsgálatát szolgálja.

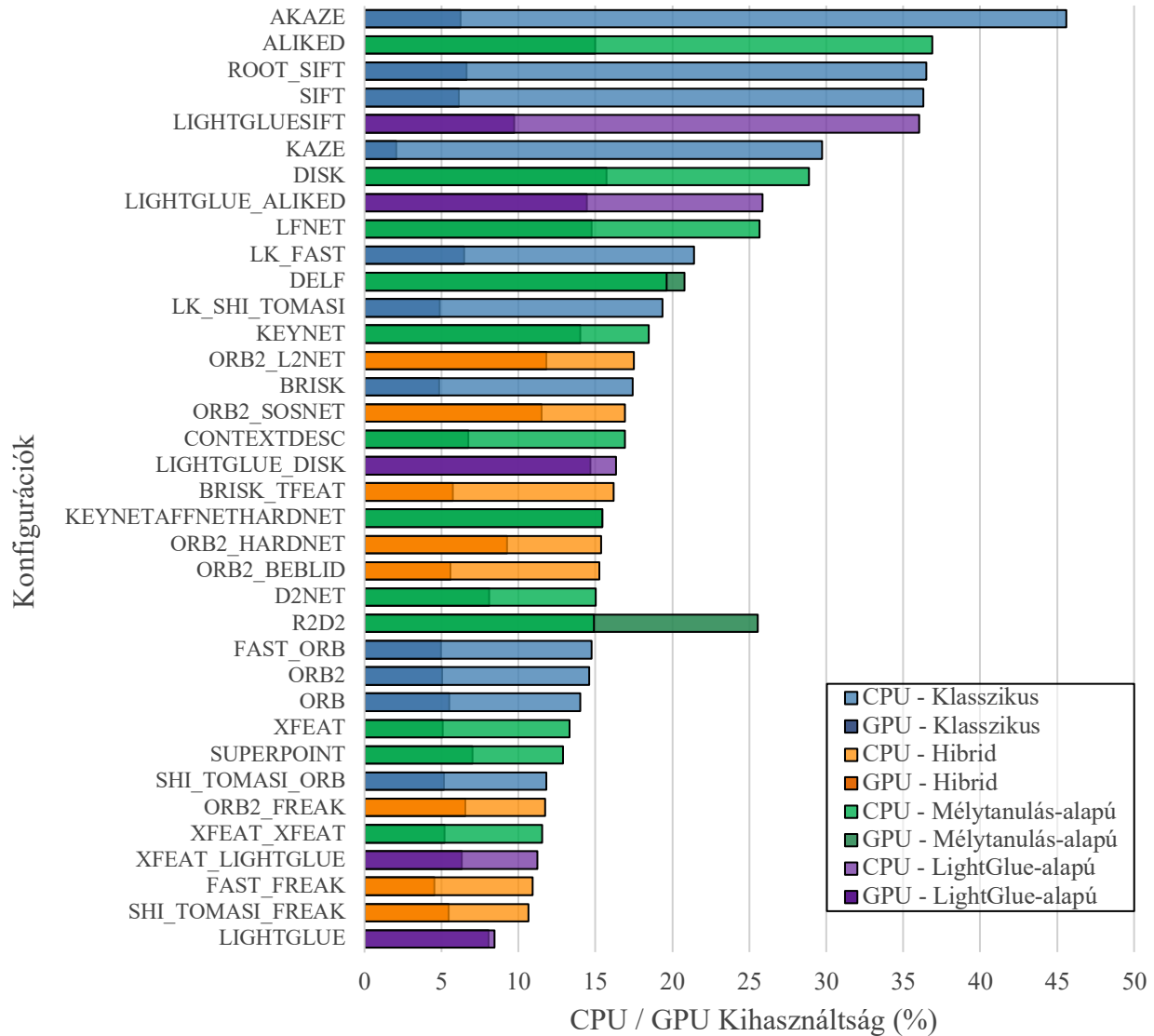


20. ábra: A konfigurációk feldolgozási sebessége

4.3 Erőforráshasználat

A 7. ábra az előzőleg bemutatott konfigurációk átlagos CPU- és GPU-terhelését mutatja, CPU kihasználtság szerint rendezve. A különböző konfigurációk erőforrásigénye jelentős eltérést mutatott. A klasszikus megközelítések (kék) főként CPU-n futottak, minimális GPU-igénnyel. A tanulásalapú modellek (zöld) elsősorban a GPU-t terheltek, míg a hibrid konfigurációk (narancs) jellemzően kiegyensúlyozottabb terhelést produkáltak. A PySLAM jelenlegi

implementációja nem teszi lehetővé a hardver teljes kihasználását, ezért további optimalizáció szükséges.



21. ábra: A konfigurációk erőforráshasználata

5. Következtetések és jövőbeli munka

A vizsgálat eredményei azt mutatják, hogy a vizuális odometria (VO) konfigurációk teljesítménye több, gyakran egymással versengő szempont szerint értékelendő. A pontosság, a futási sebesség és az erőforrásigény között rendszerint kompromisszum áll fenn, ezért az optimális beállítás kiválasztása nagymértékben függ az adott alkalmazási környezet követelményeitől. A pontosság szempontjából a **LIGHTGLUE_SIFT** konfiguráció bizonyult

a legeredményesebbnek, megelőzve a klasszikus, tanulásalapú és hibrid megközelítéseket. Általánosságban a **SIFT-alapú konfigurációk** domináltak a rangsor élén, amely rávilágít arra, hogy a klasszikus detektorok és leírók továbbra is versenyképes teljesítményt nyújthatnak a modern algoritmusokkal szemben is.

Jövőbeli fejlesztési lehetőségek:

- komplexebb adathalmazok bevonása (pl. Oxford RobotCar [35]),
- valós idejű tesztelés élő videóstreamekkel,
- futásidő optimalizálása párhuzamosítással és GPU-gyorsítással,
- a sztereó VO teljes körű támogatásának beépítése és dedikált algoritmusok integrálása,
- a párosító algoritmusok részletes teljesítményvizsgálata.

A cikk stabil, összehasonlító alapot biztosít, amelyre bővíthető és skálázható kiértékelési rendszer épülhet. Új algoritmusok, metrikák és alkalmazási forgatókönyvek integrálásával ez a keretrendszer hosszú távon is hatékonyan támogathatja a VO-kutatást és fejlesztést.

6. Köszönetnyilvánítás

A publikációban szereplő kutatást a Széchenyi István Egyetem az Európai Unió támogatásával valósította meg, az RRF-2.3.1-21-2022-00004 azonosítójú, Mesterséges Intelligencia Nemzeti Laboratórium projekt keretében.

Irodalomjegyzék

- [1.] D. Scaramuzza és F. Fraundorfer: Visual Odometry [Tutorial], *IEEE Robotics and Automation Magazine*, 18(4), 2011, 80–92.
- [2.] H. Moravec: *Obstacle Avoidance and Navigation in the Real World by a Seeing Robot Rover*, PhD-értekezés, Stanford University, 1980.
- [3.] M. Maimone, Y. Cheng és L. Matthies: Two Years of Visual Odometry on the Mars Exploration Rovers, *Journal of Field Robotics*, 24(3), 2007, 169–186.

- [4.] C. Cadena, L. Carlone, H. Carrillo, Y. Latif, D. Scaramuzza, J. Neira, I. Reid és J.J. Leonard: Past, Present, and Future of SLAM, *IEEE Transactions on Robotics*, 32(6), 2016, 1309–1332.
- [5.] L. Freda: PySLAM: A Modular Python Framework for Visual Odometry and SLAM, *arXiv preprint*, 2024, arXiv:2502.11955.
- [6.] J. Tang, Y. Ma, H. Zhu és Y. Wang: A Survey on Visual-Inertial SLAM and VO Based on Deep Learning, *HAL Open Science*, 2023.
- [7.] B.D. Lucas és T. Kanade: An Iterative Image Registration Technique, *Proc. IJCAI*, 1981, 674–679.
- [8.] D.G. Lowe: Distinctive Image Features from Scale-Invariant Keypoints, *International Journal of Computer Vision*, 60(2), 2004, 91–110.
- [9.] E. Rublee és munkatársai: ORB: An Efficient Alternative to SIFT or SURF, *Proc. ICCV*, 2011, 2564–2571.
- [10.] S. Leutenegger és munkatársai: BRISK: Binary Robust Invariant Scalable Keypoints, *Proc. ICCV*, 2011, 2548–2555.
- [11.] P.F. Alcantarilla és munkatársai: Fast Explicit Diffusion for Accelerated Features, *Proc. BMVC*, 2013.
- [12.] P.F. Alcantarilla és munkatársai: KAZE Features, *Proc. ECCV*, 2012.
- [13.] M. Calonder és munkatársai: BRIEF: Binary Robust Independent Elementary Features, *Proc. ECCV*, 2010, 778–792.
- [14.] A. Alahi és munkatársai: FREAK: Fast Retina Keypoint, *Proc. CVPR*, 2012.
- [15.] R. Arandjelović és A. Zisserman: Three Things to Improve Object Retrieval, *Proc. CVPR*, 2012, 2911–2918.
- [16.] E. Suárez és munkatársai: BEBLID: A Binary Descriptor for Fast Matching, *Pattern Recognition Letters*, 2020.
- [17.] OpenCV Documentation: cv::BFMatcher Class Reference.

- [18.] M. Muja és D.G. Lowe: Scalable Nearest Neighbor Algorithms, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014.
- [19.] D. DeTone, T. Malisiewicz és A. Rabinovich: SuperPoint: Self-Supervised Feature Detection, *arXiv preprint*, 2018, arXiv:1712.07629.
- [20.] M. Dusmanu és munkatársai: D2-Net: CNN for Detection and Description, *arXiv preprint*, 2019, arXiv:1905.03561.
- [21.] X. Zhao és munkatársai: ALIKED: Lightweight Keypoint Extraction, *arXiv preprint*, 2023, arXiv:2304.03608.
- [22.] J. Revaud és J. Glaunès: R2D2: Repeatable Detector and Descriptor, *arXiv preprint*, 2019, arXiv:1906.06195.
- [23.] Y. Ono és munkatársai: LF-Net: Learning Local Features, *arXiv preprint*, 2018, arXiv:1805.09662.
- [24.] M. Tyszkiewicz és munkatársai: DISK: Learning Features with Policy Gradient, *arXiv preprint*, 2020, arXiv:2006.13566.
- [25.] H. Noh és munkatársai: Image Retrieval with Attentive Deep Features, *arXiv preprint*, 2016, arXiv:1612.06321.
- [26.] D. Lindenberger és munkatársai: LightGlue: Feature Matching at Light Speed, *arXiv preprint*, 2023, arXiv:2204.09103.
- [27.] A. Mishchuk és munkatársai: Local Descriptor Learning Loss, *Proc. NeurIPS*, 2017.
- [28.] Y. Tian és munkatársai: SOSNet: Second Order Similarity Regularization, *arXiv preprint*, 2019, arXiv:1904.05019.
- [29.] A. Geiger, P. Lenz és R. Urtasun: KITTI Vision Benchmark Suite, *Proc. CVPR*, 2012.
- [30.] A. Geiger, P. Lenz, C. Stiller és R. Urtasun: The KITTI Dataset, *International Journal of Robotics Research*, 32(11), 2013, 1231–1237.
- [31.] J. Burri és munkatársai: EuRoC MAV Datasets, *International Journal of Robotics Research*, 35(10), 2016, 1157–1163.

- [32.] Z. Luo és munkatársai: ContextDesc: Descriptor Augmentation, *arXiv preprint*, 2019, arXiv:1904.04084.
- [33.] A. Barroso-Laguna és munkatársai: Key.Net: Keypoint Detection via CNN Filters, *arXiv preprint*, 2019, arXiv:1904.00889.
- [34.] V. Balntas és munkatársai: Learning Descriptors with Shallow CNNs, *Proc. BMVC*, 2016.
- [35.] W. Maddern, G. Pascoe, C. Linegar és P. Newman: 1 Year, 1000 km: The Oxford RobotCar Dataset, *International Journal of Robotics Research*, 36(1), 2017, 3–15.

A mentális terhelés kérdőíves értékelése szimulált villamosvezetési környezetben

Questionnaire-based assessment of mental workload in a simulated tram driving environment

Nagy Márton^a , Bagosi Attila Ferenc^b , Nagy Viktor^a

^aSzéchenyi István Egyetem

nmarton11@gmail.com; nviktor@sze.hu

^bBudapesti Közlekedési Zártkörűen Működő Részvénytársaság

bagosia@bkv.hu

Absztrakt

A városi közlekedés komplexitása és a növekvő forgalmi terhelés fokozott mentális igénybevételt jelent a villamosvezetők számára. Kutatásunk során a kognitív terhelés vizsgáltuk szimulációs környezetben, kérdőíves módszerrel kiegészítve. A vizsgálatunkban 13 villamosvezető vett részt, amely során a mentális terhelést mérjük. A vizsgálat során egy villamosszimulátorban vezettek, ahol előre beállított forgalmi szituációkkal találkoztak. Eredményeink alapján a legnagyobb kognitív terhelést a komplex, többingeres kereszteződések okozták, különösen a fonódó pályaszakasz, valamint a Halász utcai kereszteződés, amit a résztvevők is visszaigazoltak. A kérdőív alkalmazása hasznos kiegészítője volt az objektív méréseknek, lehetővé téve új, korábban nem azonosított problémás szakaszok feltárását. A kutatás hozzájárulhat a közlekedésbiztonság növeléséhez azáltal, hogy irányt mutat a kognitív terhelést csökkentő intézkedések kidolgozásához.

Kulcsszavak: kérdőív, villamosvezető, kognitív terhelés, baleset, közlekedésbiztonság

Abstract

Tram drivers face significant cognitive load due to the complexity of urban traffic. This study assessed mental workload in a tram-driving simulator using objective data and a structured questionnaire. Thirteen experienced drivers completed routes with realistic, unexpected traffic events. The questionnaire gathered demographic information, driving experience, and subjective evaluations of key scenarios. Results identified two particularly demanding locations: the shared track segment beneath Margit Bridge and the Halász Street junction. These areas were rated as cognitively taxing, aligning with prior objective findings. The questionnaire complemented the simulations by revealing additional mental strain factors not captured by measurements alone. This combined method supports identifying high-load situations and informs future efforts to reduce driver workload through infrastructure and operational changes, ultimately enhancing traffic safety.

Keywords: questionnaire, tram driver, cognitive load, accident, traffic safety

1. Bevezetés

A városi közlekedési hálózat összetett és kihívásokkal teli környezetet jelent minden közlekedő számára. A motorizáció folyamatos növekedésével a forgalom sűrűsége évről évre emelkedik, miközben a közlekedés résztvevői - a gyalogosoktól a nagytömegű járművekig - rendkívül változatosak. A városokban különösen jellemző, hogy különböző típusú és menetdinamikájú járművek osztoznak ugyanazon a pályán, ami számos konfliktushelyzethez vezethet. A közlekedők ezen felül eltérő képesítéssel és tapasztalattal vesznek részt a forgalomban: míg a hivatásos járművezetők szigorú követelményeknek kell, hogy megfeleljenek, addig a gyalogosok képzettség nélkül is részt vehetnek a forgalomban.

A nagyvárosi közösségi közlekedésben kiemelt szerepet tölt be a villamos, amely bár a KRESZ szerint közúti járműnek minősül, számos tulajdonságában eltér a hagyományos járművektől. Kötőtpályás és fém kerekeken gurul, melynek következtében a súrlódási tényező alacsony és ezáltal a járművek fék távolsága is jelentősen hosszabb. Azonban gyakran a többi járművel közös pályán halad, ezért kialakításának igazodnia kell a városi közlekedés sajátos követelményeihez. A villamosvezetők munkájuk során számos potenciális veszéllyel és konfliktushelyzettel találkoznak, amelyek elkerülése gyakran a vezetéstechnikai készségeikre és mentális terhelhetőségükre alapul. Továbbá, a balesetek elkerülése mellett figyelemmel kell lenniük az utasok biztonságára és kényelmére is, ezért kerülendők a hirtelen fékezések és gyorsítások. Mivel az agy információfeldolgozó kapacitása korlátozott, a túlzott mennyiségű inger feldolgozása figyelmi hiányosságokat és hibás reakciókat eredményezhet – különösen egy komplex közlekedési környezetben. Ezáltal tehát célunk az, hogy a villamosvezetők kognitív terheltségét különböző, innovatív módszerekkel a lehető legalacsonyabbra csökkentsük, ezzel is növelve a közlekedésbiztonságot. Ahhoz, hogy a kognitív terhelést csökkenteni lehessen, először is egy objektív mérési rendszerre van szükség, amely alkalmas a mentális terhelés mérésére, valamint okainak beazonosítására. Ez a mérési rendszer több komponensből áll: egy szemmozgáskövető szemüvegből, amely olyan adatokat rögzít, melyekből közvetlenül lehet következtetni az adott személy mentális terhelésére, fáradtságára, stb., egy járművezetői cselekvést rögzítő kamerából, valamint egy kérdőívből. A kérdőívek kulcsszerepet játszanak a kutatásokban, mivel lehetővé teszik a résztvevők szubjektív élményeinek és észleléseinek feltárását, amelyek objektív eszközökkel nem mérhetők. Segítségükkel nem várt, de releváns tényezők is azonosíthatók. Például egy figyelmi feladathoz kapcsolt kérdőív rávilágíthat a megterhelés, zavaró ingerek vagy stressz észlelt hatásaira.

2. A vizsgálat célja és rövid leírása

Az általunk elvégzett és a jövőben elvégzendő mérések elsődleges célja a villamosvezetők kognitív terhelésének vizsgálata új, innovatív és komplex vizsgálati módszerekkel. A vizsgálat megértéséhez kiemelkedő fontosságú kiemelni, hogy agyunk kapacitása véges. Az emberi kognitív rendszer – különösen az agy információfeldolgozó képessége – biológiai korlátok között működik, ezért egyszerre csak egy bizonyos mennyiségű inger feldolgozására alkalmas. Ennek következményei több jelenségben is megnyilvánulnak, például a figyelmetlenségi vakságban és a pszichológiai refrakter periódus során.[1] A figyelmetlenségi vakság azt írja le, amikor egy váratlan, jól látható vizuális inger észrevétlen marad, mivel az egyén figyelme másra irányul, vagy leköti egy párhuzamosan végzett feladat. Ez a korlátozott tudatos feldolgozási kapacitásból fakad: az emberi agy információfeldolgozó erőforrásai végesek, így, ha ezek egy adott tevékenységre összpontosulnak, más ingerek nem jutnak el a tudatos észlelés szintjére.[2] A kognitív terhelés mérésére azért van szükség, ugyanis ez a tényező nagymértékben hozzájárul a közlekedésbiztonság fokozásához. A budapesti villamoshálózaton, a balesetek bekövetkezéséért csak elenyésző mértékben felelősek a villamosvezetők, az alábbi táblázat tartalmazza a 2019-2023 között bekövetkezett balesetek számát, ami alapján a villamosvezetők az összes baleset 16%-ért voltak felelősek.

1. táblázat: A BKV Zrt. adatai alapján a villamoshálózaton bekövetkezett balesetek száma

Esemény hibássága	Saját hibás	Idegen hibás
Összesen	264 db	1647 db
Összes baleset	1911 db	
Saját hibás balesetek aránya	16%	

Azonban a villamosvezetőknek mégis nagy szerepe van a balesetmegelőzésben, megfelelő figyelemmel, előrelátással és helyes vezetéstechnika alkalmazásával a balesetek és a balesetveszélyes szituációk nagy részét el tudják kerülni. Ezt azonban csak akkor tudják megtenni, hogyha a mentális kapacitásuk nincsen túlságosan leterhelve. Az ilyen kognitív terheltséget okozó tényezőket két csoportba soroltuk: az állandó és változó tényezők csoportjai közé. Az állandó tényezők olyan tényezők, amelyek minden körülmények között a vezetés és a járművel való közlekedés részét képezik, valamint hatásukat csak kis mértékben vagy egyáltalán nem lehetséges csökkenteni. Ilyen tényezők többet között az olyan összetett kereszteződések, ahol több közlekedési ágazat találkozik egymással, valamint a forgalmat forgalomirányító fényjelző készülék nem szabályozza, csak táblás biztosítás van. Az ilyen helyszíneket

megvizsgálva és a problémák beazonosításával lehetőség van olyan változtatásokat eszközölni, amellyel csökkenthető a járművezetőkre ható kognitív terhelés. Ehhez az eszközök között az elsőbbségi viszonyok racionalizálása, valamint az infrastruktúrán történő kismértékű átalakítások elvégzése. A módosítások után újabb méréseket szükséges végezni, amely segítségével megállapítható, hogy az alkalmazott változtatások pozitív vagy negatív irányba befolyásolták a járművezetők kognitív terhelését. További megemlíthető szempontok, melyek az állandó tényezők csoportjába tartoznak:

- az utasok testi épségének és komfortjának védelme
- az aktuális időjárási viszonyokhoz való alkalmazkodás
- a jármű kötött menetdinamikai tulajdonságaihoz való alkalmazkodás
- a menetidőhöz való kötöttség
- az útvonal vonalvezetése és sajátosságai

A mentális terhelés vizsgálatakor minden esetben számolni kell bizonyos, elkerülhetetlen hatótényezőkkel. Ugyanakkor számos olyan befolyásoló tényező is létezik, melyek hatása különböző mértékben csökkenthető, sőt egyes esetekben akár teljes mértékben ki is küszöbölhető. Ezeket változó vagy változtatható tényezőknek nevezzük. Kiemelt jelentőséggel bír közülük a végállomáson rendelkezésre álló pihenőidő. Ennek növelése minden esetben javasolt, hiszen, ha a járművezetők számára elegendő idő áll rendelkezésre a személyes szükségleteik kielégítésére és a rövid pihenésre, nyugodtabban végzik munkájukat. Ezáltal csökken a mentális terhelésük, és nagyobb figyelmi kapacitás marad a forgalmi szituációk idejében történő felismerésére és az adekvát reakciók meghozatalára. További fontos változó tényező a zavaró ingerek számának minimalizálása. Ilyen zavaró tényező lehet például az utasok figyelemelterelő vagy nem megfelelő viselkedése, illetve bizonyos eszközök kezelése a forgalom lebonyolítása során. Ilyen eszköz lehet például a fedélzeti számítógép, amelynek használata menet közben bár nem tiltott mégis jelentős figyelemelvonó hatással járhat. Ezen túlmenően a jármű kezelőszerveinek elhelyezése és kialakítása is befolyásolja a vezetés közbeni terhelést. A megfelelően ergonomikus és átgondolt kezelőfelületek hozzájárulhatnak a kognitív terhelés csökkentéséhez, mivel mérséklik a felesleges figyelmi megoszlást a jármű irányítása során. Mindezen hatások és tényezők objektív mérése egy megfelelő mérési struktúra kidolgozására volt szükséges. A mérési terv kidolgozásánál fontos szempont volt a vizsgálatok megfelelő, objektív módon történő megismételhetősége, mindig ugyanazon körülmények

között és a résztvevőkkel ugyanazokat a szituációkat tudjuk megismételni. Ez azért szükséges, ugyanis a vizsgálatok során a lehető legtöbb külső behatást szükséges semlegesíteni, hogy a járművezetők tényleges kognitív terhelése jól beazonosítható és mérhető legyen. Ezen feladathoz egy villamosszimulátort használtunk fel, mely valós villamosok felépítését követi, a valóságban megtalálható útvonalakon, a valóságnak megfelelő vonalvezetéssel. A kognitív terhelés hatásának beazonosítására két módszert tudunk alkalmazni, az egyik, hogy egy-egy bizonyos helyszínt kiválasztva több résztvevő kognitív terhelését összegezve vagy átlagolva megmondható, hogy ott kiugróan magas mértékű kognitív terheltség hat a járművezetőkre. Ezen kívül egy további módszer, hogy az adott személy egy előre meghatározott útvonalon vezet, előre beállított, sok esetben váratlan szituációkra reagálva, a mérés végén pedig a saját mentális terheltségi adataihoz képest viszonyítva azonosíthatóak be a különösen nagy eltérést mutató helyszínek, valamint forgalmi szituációk.

A tanulmányunkban végzett vizsgálatához baleseti statisztikák elemzésével választottunk ki olyan csomópontokat, amelyek potenciálisan problémát jelenthetnek a járművezetők számára. Emellett figyelembe vettük a járművezetők tapasztalatait, valamint saját megfigyeléseinket is. A mérések nem a teljes hálózatra terjedtek ki, hanem több, minőségi és gyakorlati szempontok alapján kiválasztott kereszteződésre, mindkét forgalmi irányból. A kutatás középpontjában a járművezetőket érő kognitív terhelés vizsgálata áll, különböző forgalmi szituációkban – többek között keresztezésekben és egy olyan helyszínen, ahol a villamosok forgalmát úgynevezett jelzők szabályozzák. A jelzők olyan helyhez kötött berendezések, amelyek fényekkel valamilyen egyéb módon jelzéseket adnak a járművezetők számára, továbbá feladatuk a forgalom szabályozása, a biztonság megteremtése is.[3] Minden kiválasztott helyszínt két irányból vizsgáltunk, mivel a zavaró tényezők és a beláthatóság irányonként eltérhet. A helyszínek kiválasztásánál arra is törekedtünk, hogy különböző vezetői magatartásformákat igényeljenek. Fontos szempont volt, hogy a mérés ténye, azaz a megfigyelés és utólagos kiértékelés, pszichológiai nyomást helyezhet a járművezetőkre, ami torzíthatja az eredményeket. Emellett a mérőeszközök viselése is okozhat enyhe kényelmetlenséget, ami szintén hatással lehet a mentális terhelésre, azonban ez a hatás minimalizálható korszerű, szinte észrevehetetlen eszközök alkalmazásával.[4] Az adatok értékeléséhez és a mentális terhelés mértékének meghatározásához a vizsgálatban résztvevők egyéni eredményeit, valamint ezek átlagát vettük alapul.

A vizsgálat konkrét lefolyása a szimulátorban úgy történt, hogy a résztvevők két előre meghatározott vonalon vezettek végig, mindkét irányban. A mérés során kifejezett hangsúlyt fordítottunk arra, hogy a vezetés során minden érvényes szabályt és szabályozást betartsanak, valamint úgy járjanak el, mint egy valós forgalmi szituáció során. Az útvonalon előre beállított váratlan helyzetekkel találkoztak, melyre a saját belátásuk szerint reagáltak. Ezeket a szituációkat három csoportba lehet sorolni:

- egy személygépkocsi a villamossal párhuzamosan halad és váratlanul balra a villamos elé fordul
- egy gyalogos szabálytalanul a villamos elé lép
- egy úrszelvényhez közel található tárgy akadályozza a villamos forgalmát (az úrszelvény tulajdonképpen az a térrész, amit a villamos konkrétan elfoglal és minden körülmények között szabadon tartandó)

3. A kérdőívek általános leírása

Általánosan kijelenthető, hogy a pszichológiai és mérnöki kísérletekben a kérdőívek kiemelt fontossággal bírnak, mert lehetővé teszik a vizsgálatot végző személyek számára, hogy a résztvevők olyan belső tapasztalatait és nézeteit is feltárják, amelyeket pusztán viselkedéses vagy teljesítményadatokkal nem lehetne megragadni. A kérdőívek strukturált adatgyűjtést nyújtanak az emberi percepciókról és önbeszámolókról, ezzel segítve a hipotézisek tesztelését és a kísérleti eredmények értelmezését. Fontos azonban, hogy a kérdések valóban a vizsgált konstruktumot mérjék, különben félrevezető eredmények születhetnek. Nizonyos kérdőívek eredményei szorosan korrelálnak a viselkedéses mutatókkal, míg más esetekben gyengébb az összefüggés. A kérdőívek alkalmazásának hatása is jelentős: a kérdések megfogalmazása, a felépítés és a felhasználási cél kommunikációja egyaránt befolyásolja a válaszadást.[5] Korábbi kutatások kimutatták, hogy a kitöltők mérlegelik, mire használják fel a válaszaikat, és ha nem értenek egyet a célokkal vagy nem bíznak a felhasználásban, torzíthatják vagy meg is tagadhatják a válaszadást. Ez a jelenség rávilágít a bizalom és anonimitás fontosságára a hiteles adatok érdekében. Számos területen váltak a kérdőívek alapvető eszközzé: például a humán tényezőkkel foglalkozó mérnöki kutatásokban a NASA Task Load Index (NASA-TLX) kérdőívet rutinszerűen használják a mentális terhelés mérésére – eredetileg a repülésben fejlesztették ki, de ma a közlekedésben, egészségügyben és ember–gép interakcióban is

nélkülözhetetlen.[6] Az ilyen szabványosított kérdőívek lehetővé teszik a szubjektív élmények (például a terhelés érzete vagy a felhasználói elégedettség) számszerűsítését és összevetését az objektív mutatókkal, így a kutatók teljesebb képet kapnak a vizsgált jelenségről. **Összességében a kérdőívek célja és szerepe** az, hogy az emberi tényezőkre vonatkozó adatokat rendszerezett formában gyűjtsék össze és az objektív méréseket kiegészítve átfogóbb értékelést tegyenek lehetővé a kísérletekben. Ugyanakkor a megbízható következtetések érdekében elengedhetetlen a kérdőívek gondos tervezése, validálása és a válaszadói torzítások minimalizálása.

3.1. A kutatáshoz használt kérdőív bemutatása

A kutatás részét képezte egy összetett, strukturált kérdőív is, melynek elsődleges célja a vizsgálatban részt vevő villamosvezetők szubjektív tapasztalatainak, háttérinformációinak, valamint az előre meghatározott forgalmi szituációk megítélésének feltárása volt. A kérdőív több logikusan elkülönülő kérdéscsoportból állt, melyek a válaszadók személyes jellemzőire, korábbi vezetési tapasztalataira, szakmai múltjára, valamint a szimulációs vizsgálattal kapcsolatos észleléseire irányultak. A kérdőív alkalmazásával a következő kérdésekre kívántunk választ kapni:

- A járművezető tapasztalatának van-e szerepe a mentális terhelés hatásában, azaz, amennyiben egy adott személy tapasztalata egyre nagyobb egy megadott útvonalon vagy járművet vezetve, számára egy-egy lappangó, vagy esetenként veszélyes forgalmi szituáció felismerése és a szituációra való reagálás kisebb kognitív terheléssel jár-e?
- Amennyiben a vizsgált személy más járművet is vezet, az ott megszerzett tapasztalatával a villamosvezetés során is csökken-e a mentális terhelése adott szituációkban?
- Van-e szerepe a mentális terhelésnél az életkornak, illetve van-e összefüggés az életkor és a kognitív terhelés nagysága között bizonyos forgalmi szituációkban?
- Amennyiben a vizsgált személyeknek volt saját hibás balesetük, számukra okoz-e nagyobb nehézséget egy-egy veszélyhelyzet felismerése és az arra adott megfelelő reakció, valamint nagyobb mentális terheléssel jár-e ez számukra, mint olyan személyek esetében, akiknek nem volt balesetük?
- Számít-e a balesetek kivédésénél az, hogy a járművezető ismeri-e az adott útvonalszakaszt, valamint járműtípust?

A kérdőív elején demográfiai adatok kerültek rögzítésre, beleértve a vizsgálati személy nemét, életkorát, azt, hogy jelenleg villamosvezető tanuló-e, valamint, hogy visel-e szemüveget vagy kontaktlencsét. Ezt követően egy átfogó kérdéscsoport tért ki a közúti járművezetői tapasztalatokra. A válaszadóknak meg kellett jelölniük, hogy rendelkeznek-e érvényes vezetői engedéllyel különböző közúti járműkategóriákra (személyautó, autóbusz, tehergépkocsi, trolibusz stb.), milyen rendszerességgel vezetnek, továbbá történt-e velük baleset, és ha igen, az saját hibából következett-e be. Egy külön kérdésblokk az országos és egyéb vasúti járművekre vonatkozó képzettségeket vizsgálta (pl. metró, fogaskerekű, kisvasút). Itt szintén a jogosultság, a gyakorlat, valamint az esetleges balesetek előfordulása és jellege került rögzítésre.

A kérdőív legrészletesebb része a villamosvezetési tapasztalatokra koncentrált. A válaszadóknak meg kellett adniuk, hogy mikor szerezték meg villamosvezetői engedélyüket, mely típusokra rendelkeznek vizsgával, illetve milyen típusú járműveken, valamint mely vonalakon szereztek gyakorlatot. Ugyanezen a ponton tért ki a kérdőív az esetleges baleseti tapasztalatokra, valamint arra is, hogy a válaszadó saját megítélése szerint felelősnek tartja-e magát egy-egy baleset bekövetkezéséért.

Ezt követően a kérdőív a szimulátoros vizsgálat alatti élményekre kérdezett rá. A résztvevőknek nyilatkozniuk kellett arról, hogy korábban vezettek-e már szimulátorban, illetve hogy a jelen vizsgálat során tapasztaltak-e bármilyen rosszullétet vagy kellemetlenséget.

A kérdőív záró részében került sor az előre beállított forgalmi szituációk értékelésére. A vizsgálat során hat-hat szituációt modelleztünk mindkét forgalmi irányban, amelyeket a résztvevők szimuláció közben éltek át. A szituációkról készült képeket a kérdőívben is bemutattuk, és az ezekre vonatkozó kérdések négy fő szempontot vizsgáltak: mennyire találta a résztvevő az adott helyzetet veszélyesnek, mennyire volt stresszes számára, időben észlelte-e azt, valamint mennyire tartja megfelelőnek saját reakcióját. Az első három kérdésre 1–5-ig terjedő Likert-skálán kellett választ adni, míg az utóbbi nyitott szöveges válaszként szerepelt, lehetőséget adva a résztvevők önreflexiójára.

A kérdőív ilyen formában nemcsak a résztvevők általános tapasztalatait, hanem az egyes szituációkra adott szubjektív reakciókat is képes volt feltérképezni, így hasznos kiegészítése a szimuláció közben rögzített objektív adatoknak. A válaszok hozzájárultak a mentális terhelés

mélyebb megértéséhez, különös tekintettel a forgalmi szituációk komplexitására és azok észlelésére, valamint a megküzdési stratégiák feltárására.

Bajnóci utcai kereszteződés

Egy autó a villamos elé fordul (Budaörsi út felé)



A fenti szituációt (Bajnóci utca, autó a villamos elé fordul a Budaörsi út felé) mennyire találta kritikusnak a biztonság szempontjából?

1 2 3 4 5

Egyáltalán nem ☐ ☐ ☐ ☐ ☐ Nagyon kritikusnak találom

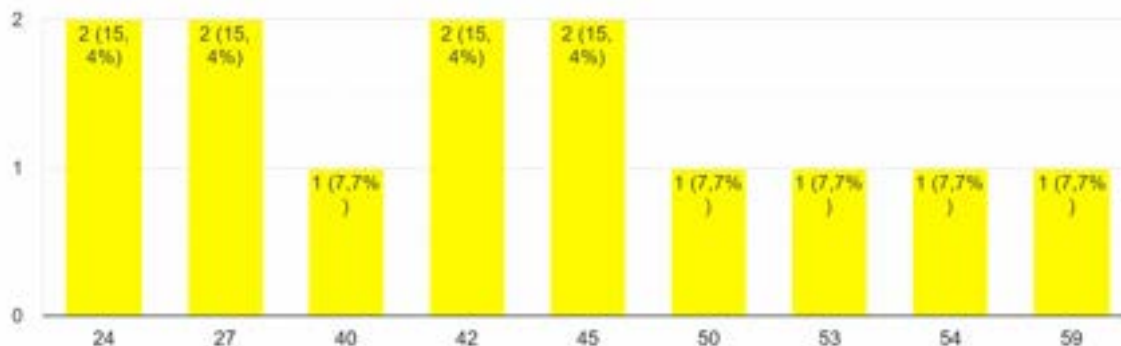
1. ábra: A kérdőív fotóval illusztrált kérdéssorára egy példa (kivonat)

3.2. A kérdőív értékelése, reflexió

A tanulmányunk során, egy pilot mérést követően, a kérdőívek végső elemzésébe 13 budapesti villamosvezetőt vontunk be. Az eredmények így nem reprezentatívak, azonban már ezen minta alapján is lehet előzetes következtetéseket levonni, az adatok jelzésértékű tendenciákat mutatnak. A kérdőív tényleges kitöltésére a mérés helyszínén került sor, közvetlenül azután, hogy a vizsgálatban résztvevő személy a villamos szimulátorban az adott szakaszt levezette és a megadott feladatokat elvégezte. A kitöltők jelentős százaléka, 69,2%-a 30 év feletti járművezető volt, valamint a kitöltők közül 84,6% férfi.

Kérjük adja meg az életkorát (csak a számot írja be, például: 25)

13 válasz



2. ábra: Az életkori adatokat ábrázoló grafikon

Szintén 9 résztvevőnek, azaz 69,2%-nak van B kategóriás jogosítványa, mellyel évente legalább 20 000 kilométert vezetnek, míg 3 résztvevő (23,1%) egyáltalán nem rendelkezik más jogosítvánnyal, mint ami a villamosvezetéshez szükséges. A villamoson kívüli infrastruktúrák tekintetében 1 személy rendelkezik HÉV vezetői képesítéssel, további 1 személy pedig a budapesti fogaskerekűre érvényes képesítéssel is. A villamosvezetés során a személygépjárművön szerzett tapasztalat is rendkívül értékes, ugyanis sok esetben a járművezetők a többi közúti közlekedő viselkedését már kitapasztalták, valamint a KRESZ szabályait is elsajátították, melyek együttesen lehetővé teszik, hogy bizonyos forgalmi szituációkra gyorsabban és magabiztosabban reagáljanak, azokat jobban előre lássák.

A baleseti előélet körében, mely itt konkrétan a közúti járművekre vonatkozik, az adatokat és eredményeket az alábbi táblázat tartalmazza. A kitöltők közül a villamos infrastruktúrán senkinek sem volt saját hibás balesete, idegen hibás három személynek volt.

2. táblázat: A balesetek számának és hibásságának alakulása a kérdőív alapján

Baleset hibássága	0 eset	1-2 eset	2-5 eset	Legalább 1 eset
Saját hibás	5	3	-	3
Idegen hibás	4	5	1	6

Azon résztvevők közül, akiknek volt saját hibás közúti balesete, saját bevallásuk alapján 75% válaszolta azt, hogy a baleset bekövetkezéséért valóban hibásnak érezte magát. Ezen eredmények alapján arra következtethetünk, hogy a résztvevők körében magas fokú a felelősségvállalás, a résztvevők nem mentetik magukat és ha úgy érzik, hibáztak, akkor azt

be is ismerik. Egy további érdekes eredmény, hogy egyik válaszadó sem választotta a „részben hibásnak érzem magam” opciót, ami arra utalhat, hogy a megkérdezett járművezetők döntően „fekete-fehér módon” ítélik meg saját felelősségüket. Ez egyaránt jelezhet magas szintű szakmai önreflexiót, de akár túlzott önhibáztatásra vagy a munkahelyi normák internalizálására is utalhat. [7] A villamosvezetők számára utasításban előírt kötelesség, hogy a baleseteket és eseményeket jelentsék a megfelelő személyek számára, a szankciók, amennyiben szükségesek, a hibák mértékének megfelelően kerülnek kiszabásra. A mentális csökkentés érdekében ezen rendszer felülvizsgálata is javasolt, egy olyan vállalati környezet kialakításával, amely kisebb mértékben az egyének szankcionálására, hanem inkább a bizonyos problémák feltárására és kollektív kiküszöbölésére fókuszál. Azonban fontos kiemelni, hogy egy ilyen súlyú munkakör betöltésénél elengedhetetlen a szankciók alkalmazása a súlyosabb esetekben.

A villamosvezetéssel kapcsolatos kérdéskör alapján a résztvevők mindegyike legalább 500 kilométert vezetett forgalomban, önállóan villamost. Az önállóság itt arra utal, hogy nem oktató, munkamódszer átadó vagy egyéb felügyelet mellett vezetett. A résztvevők jelentős része, azaz 46,2%-a több mint 50000, de kevesebb mint 250000 közötti kilométert vezetett forgalomban villamost. A résztvevők közül mindenki vezetett már korábban a szimulátorban. A mérés során a résztvevőknek két vonalszakaszon kellett végig vezetniük, az első vonalszakaszt az összes résztvevő közül 2 személy, míg a második vonalszakaszt 1 személy nem ismerte. Az útvonalszakasz nem ismerete ebben az esetben azt jelenti, hogy ott még nem vezetett önállóan, vagy nem rendelkezik a szükséges vizsgákkal (megfelelő vonalismeret) ahhoz, hogy ott önállóan villamos vezethessen. Azaz ebben az esetben nem számított, ha tanulójáráttal, próbajáráttal vagy más egyéb járáttal vezetett arra a forgalomban.

A szimulátoros mérés, illetve a szimulátorban való vezetést egy 1-től 5-ig terjedő skálán kértük, hogy értékeljék a résztvevők. Kíváncsiak voltunk arra is, hogy bármiféle rosszullét fellépett-e náluk a vizsgálat során. A járműszimulátorokban, kialakításuk vagy egyéb okok miatt előfordulhat rosszullét, mint például szédülés, fejfájás, hányinger, koncentrációs zavarok, stb. Ez azért kritikus információ a mérés tekintetében, ugyanis a tünetek a mentális terhelést negatív irányba befolyásolhatják, a figyelmet elterelhetik a tényleges vezetésről, valamint nem megfelelő körülményekhez és odafigyeléshez vezethetnek. Természetesen amennyiben a

tünetek olyan mértékűek voltak, hogy az adott résztvevő nem tudná megfelelően folytatni a szimulátoros vezetést, a mérést azonnal megszakítottuk.



3. ábra: A rosszulléttel kapcsolatos tünetekre adott válaszokat ábrázoló grafikon

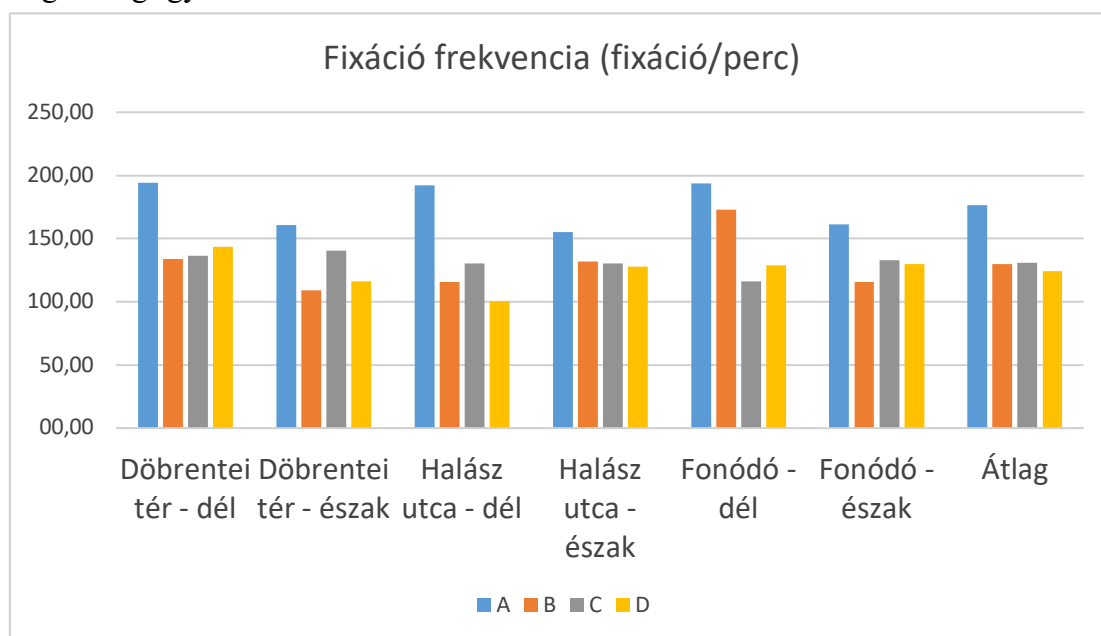
Az adatok alapján elmondható, hogy a résztvevők nagyobb százaléka tapasztalt valamiféle tünetet. Ez összefüggésben állhat a résztvevők átlagosan magasabb életkorával, továbbá a szimulátor kialakítással, azonban a konkrét eredményhez további méréseket szükséges még elvégezni.

A kérdőív egy kiemelten fontos kérdésköre, egy olyan szöveges válaszlehetőség, ahol a résztvevők a mérés után leírhatják a tapasztalataikat, valamint azt, hogy saját belátásuk szerint ők hol és milyen helyzeteket találtak kifejezetten balesetveszélyesnek, nehezen végrehajthatónak vagy túlzottan megterhelőnek, komplexnek. Ebben a szakaszban a résztvevők még nem találkoztak azzal, hogy mi számukra konkrétan milyen szituációkat állítottunk be számukra, viszont elengedhetetlen fontosságú visszajelzést adhatnak számunkra. Ebben a szakaszban a válaszok alapján feltárhatóak olyan kritikus helyszínek is, amelyekre a korábbiakban nem gondoltunk, esetlegesen, amelyek korábban nem kerültek vizsgálatra. Ezen kívül, mint az jelen kutatásunkban is látható, beigazolást nyerhet, hogy a kiválasztott forgalmi szituációk és helyszínek valóban összetettek és nehézséget okoznak a résztvevők számára. A kérdésre kapott válasz alapján a legtöbb résztvevő a villamos úrszelvényéhez közel elhelyezkedő tárgyakat tartotta kifejezetten komplexnek és megterhelőnek. Kiemelendő, hogy továbbá olyan helyzetek kerültek kiemelésre, amelyeknél több összetett feladat ellátása is szükséges a megoldáshoz, például, hogy egy villamoshoz közel elhelyezkedő tárgy található meg a pálya mellett, egy kereszteződés előtt, így figyelni kell a tárgyra, a forgalomirányító fényjelző készülék jelzésére valamint esetlegesen arra, hogyha szükséges megállni, mert a tárgy mellett a villamos biztonságosan nem fér el, a kereszteződésben a járművek és a gyalogosok

forgalmát a járművel ne akadályozza. Az ilyen összetett szituációk sokszor a személy szubjektív belátása szerint is rendkívül nagy kognitív terheléssel járnak.

A kérdőív legvégső szakaszában, mint az már a korábbiakban említésre került, az előre beállított, váratlan szituációkkal kapcsolatos kérdéseket tettünk fel a résztvevők számára. A szituációk jobb beazonosíthatósága érdekében egy-egy fotóval illusztrálásra kerültek azok, továbbá a helyszín rövid leírását is elhelyeztük a kép fölött.

Ugyanezen vizsgálat egy korábbi mérései alapján a járművezetők számára a legnagyobb mentális terhelést elsősorban a Halász utcai kereszteződés, másodrészt pedig a Margit híd alatt található fonódó pályaszakasz okozta. Ennek oka, hogy mindkét helyszín komplex, a járművezetőktől magas fokú figyelemmegosztást igényel, egyszerre a forgalom több elemét is szükséges megfigyelniük.



4. ábra: A korábbi kutatás fixáció frekvenciáját ábrázoló grafikon

A fenti grafikonon a résztvevők fixáció frekvenciája látható, azaz mennyi fixációjuk volt percenként. A fixáció az a jelenség, amikor egy személy tekintete egy bizonyos pontra hosszabb ideig fókuszál. Minél több a fixáció egy adott időszámban, a kognitív terhelés annál nagyobb mértékű.[8] A korábbi kutatásból 4 résztvevő adatait vettük figyelembe és ez alapján következtettünk arra, hogy mely helyszínek voltak a legmegterhelőbbek. A mostani tanulmányunkban a kérdőívre kapott válaszok alapján a résztvevők 61,6%-a úgy nyilatkozott, hogy legalább közepes szinten (azaz 3-mas vagy afölötti értékkel) volt számára megterhelő a fonódó szakasz déli irányú teljesítése. A Halász utcai kereszteződés szempontjából, a

résztvevők nagyobb része, 61,6%-a nyilatkozott úgy, hogy kevesebb mint közepes szinten találta megterhelőnek a déli irányú teljesítést, azonban közülük 77% nyilatkozta, hogy számára legalább közepes szinten megterhelő volt az északi irányú teljesítés.

3. táblázat: A kérdőív Likert-skálájára kapott válaszok megoszlása

Skála értékei	Fonódó – déli irány	Halász utca – déli irány	Halász utca – északi irány
1-2	5 (38,5%)	10 (77%)	3 (23,1%)
3	-	2 (15,3%)	4 (30,8%)
4-5	8 (61,5%)	1 (7,7 %)	6 (46,1%)

A korábbi mérések és a jelenlegi kutatásban kapott adatok és válaszok alapján előzetesen arra a következtetésre jutottunk, hogy valóban ezek a helyszínek nagyfokú mentális terhelést és nehézséget jelentenek a járművezetők számára. A kérdőív alkalmazásával további adatokkal egészíthető ki a kognitív terheléses vizsgálat, jobban beazonosítható az, hogy mely irányok, illetve milyen pontok azok, amik a járművezetők számára kifejezetten megterhelőek voltak.

Fontos megemlíteni, hogy a kis adathalmaz miatt ezen adatok csak tájékoztatást adnak, illetve csak előzetes következtetéseket tudunk levonni belőlük. Jelen kutatásunk elsősorban csak a hipotézisünk igazolására, valamint előzetes vizsgálati módszerek kidolgozására irányul, melyek segítségével látható, hogy valóban hasznos-e egy ilyen kérdőív kidolgozása, illetve az esetleges hibákat, valamint gyengeségeket egy átfogóbb mérés során már előre kiküszöböltnek tekinthessük.

4. Összegzés

A nagyvárosok közlekedési hálózata egy komplex rendszert alkot, a járművek és járművezetők eloszlása rendkívül sokrétű. A motorizáció növekedésével egyre több jármű van az utakon, a balesetek száma is vagy stagnáló vagy növekvő tendenciát mutat. A balesetek bekövetkezéséhez nagy mértékben szerepe van a járművezetők figyelmének és kognitív terhelésének. A közösségi közlekedés szempontjából egy kiemelt fontosságú eszköz a villamos. A villamosvezetők, habár a balesetek elenyésző százalékáért felelősek, mégis nagy szerepük van a balesetek bekövetkezésének elkerülésében. Elmondható, hogy az emberek agyi kapacitása biológiai korlátok miatt véges, egyidejűleg csak egy bizonyos számú információ fogadására és feldolgozására képes. Mivel a mentális terhelésnek nagy szerepe van a balesetek

kivédésében, elsődleges cél, hogy ezt az értéket csökkenteni tudjuk bizonyos módszerek alkalmazásával, mint például az infrastruktúra átalakításával, szabályok racionalizálásával és a villamosvezetők számára kiadott pihenőidők növelésével, számukra jobb körülmények biztosításával. Ahhoz, hogy a kognitív terhelést objektíven mérni tudjuk és ezáltal bizonyítható legyen egy-egy változtatás eszközölésének a hatása, egy átfogó mérési rendszer szükséges. Ennek a mérési rendszernek több eleme van, melynek egy része a kognitív terhelés mérésére és objektív számadatok gyűjtésére fókuszál, ilyen például a szemmozgáskövető szemüveg alkalmazása, egy további fontos eleme a szubjektív vélemények elemzése. Ehhez egy kérdőív használata a legcélszerűbb. Kérdőíveket azért érdemes használni, mert ez által megtudhatóak a résztvevők demográfiai adatai, továbbá személyes tényezők is, mint például, hogy mennyi tapasztalatuk van a villamosvezetés során, mennyi balesetük volt, továbbá, hogy hogyan értékelték a vizsgálatot. Lehetőség van olyan adatok megszerzésére is, amikre esetlegesen a vizsgálati terv kidolgozásánál nem vettünk figyelembe, mint például egy olyan helyszín is megterhelő a járművezetők számára, amiről korábban ilyen adat nem állt rendelkezésünkre. A kérdőívet egy kognitív terhelés mérésére irányuló vizsgálattal közösen alkalmazva alátámaszthatóak az ott megszerzett számadatok. Ezen kívül lehetőség van az eredmények okainak felkutatására is, tehát nem csak az tudható meg, hogy egy adott szituációban magas volt a kognitív terhelés, hanem annak konkrét okai is. Jelen kutatásunk során igazolható volt, hogy a kiválasztott helyszínek valóban nagy megterhelést jelentenek a járművezetők számára, ezen kívül fény derült egy-egy válasz okára is. Azonban fontos kiemelni, hogy jelen kutatásunk sok esetben nem egyértelmű, konkrét adattal szolgál az alacsony adathalmaz miatt. A későbbiekben további kutatások elvégzése szükséges a tényleges bizonyíthatósághoz, mely kutatások a cikk írásakor már folyamatban voltak.

5. Köszönetnyilvánítás

A cikk kutatásaihoz a „Nemzeti Laboratóriumok 2020 Program – Mesterséges Intelligencia alprogram – Mesterséges Intelligencia Nemzeti Laboratórium (MILAB) létrehozása a Széchenyi István Egyetemen (NKFIH-870-21/2020)” biztosított forrást.

Irodalomjegyzék

- [1.] R. Marois – J. Ivanoff: Capacity limits of information processing in the brain, Trends in Cognitive Sciences, 2005/9 296-305
- [2.] M. F. Tang – L. Ford – E. Arabzadeh – J. T. Enns – T. A. W. Visser – J. B. Mattingley: Neural dynamics of the attentional blink revealed by encoding orientation selectivity during rapid visual presentation, Nature communications, 2020/11.1, 434
- [3.] Villamos F.1.-F.2. számú Jelzési és forgalmi utasítás, 2017
- [4.] Juhász Anikó (2022): Mérnöki–pszichológiai szempontok: A járművezető–jármű interakció és a járművezetési környezet kialakítása. Doktori értekezés tervezet, Nemzeti Közszolgálati Egyetem, Rendészettudományi Kar. Elérhető: https://rtk.uni-nke.hu/document/rtk-uni-nke-hu/Juhasz_Aniko_ertekezeset_Tervezet_2022.pdf
- [5.] Clark IA, Maguire EA. Do questionnaires reflect their purported cognitive functions? Cognition. 2020 Feb;195:104114. doi: 10.1016/j.cognition.2019.104114. Epub 2019 Dec 20. PMID: 31869709; PMCID: PMC6963768.
- [6.] Wolf MG, Denison AJ. Survey Uses May Influence Survey Responses. Assessment. 2024 Oct;31(7):1378-1397. doi: 10.1177/10731911231213849. Epub 2023 Dec 31. PMID: 38160401.
- [7.] Högye-Nagy Ágnes – Münnich Ákos (2015): Közlekedési balesetek és szabálysértések szerepe és mérési lehetőségei a közlekedési viselkedésben: A biztonságos vezetési index kialakítása. In: Balázs Katalin – Kovács Judit (szerk.): Alkalmazott pszichológiai tanulmányok a Szociál- és Munkapszichológiai Tanszék fennállásának 25. évfordulójára. Debreceni Egyetemi Kiadó, Debrecen, 233–256.
- [8.] Nagy V. – Kovács G.: Moduláris buszvezető monitoring rendszer tervezése, Mobilitás és környezet - Jövőformáló járműipari kutatások konferencia, Győr 2022.

Vállalati és gazdasági szerkezet összehasonlító elemzése Magyarország és Csehország NUTS 3 szintjén

Comparative Analysis of Corporate and Economic Structures at the NUTS 3 Level in Hungary and the Czech Republic

Reider-Pesti Beáta

Széchenyi István Egyetem Regionális- és Gazdaságtudományi Doktori Iskola

reider-pesti.beata@ga.sze.hu

Absztrakt

A kutatás célja Magyarország és Csehország NUTS 3 szintű vállalati és gazdasági szerkezeti összehasonlító vizsgálata térinformatikai eszközök felhasználásával. A két ország összevetése több szempontból, mint például a hasonló földrajzi elhelyezkedés, a közös történelmi múlt, szocialista időszak és azt követő rendszerváltás alapján is megalapozott és indokolt. Bár a kiindulópontok hasonlóak voltak, az elmúlt két évtized során a két ország regionális fejlettségi mintázata eltérő irányokat vett. Ennek megértéséhez elengedhetetlen a gazdasági struktúrák, és válságkezelési gyakorlatok összehasonlítása, ami a vállalati adatok regionális értelmezéséhez is megfelelő keretet ad. Az elemzés alapját a két országra az EMIS adatbázisból lekért NUTS 3 szintre aggregált vállalati mutatók - cégszám, árbevétel, kötelezettségállomány - képezik. A térségek gazdasági pozíciójának átfogó értékeléséhez ezen adatok kerültek összehasonlításra a Eurostatból elérhető GDP, és bruttó hozzáadott érték (GVA) adatokkal. A területi eltérések vizionálása QGIS szoftver segítségével történt. Célom, hogy a bemutatott adatok alapján következtetéseket vonjak le a különböző mutatószámok közötti összefüggésekre, valamint a regionális vállalati jelenlétre vonatkozóan. Az eredmények jól tükrözik a két ország közötti strukturális különbségeket, és rávilágítanak a vállalati szerkezet földrajzi sajátosságaira.

Kulcsszavak: regionális gazdaságtan, vállalati kötelezettségállomány, térinformatika

Abstract

The aim of this research is to conduct a comparative analysis of the corporate and economic structures of Hungary and the Czech Republic at the NUTS 3 level using geoinformatics tools. The comparison of the two countries is well-founded and justified from multiple perspectives, including their similar geographical location, shared historical background, socialist past, and the subsequent regime change. Although their starting points were similar, over the past two decades the regional development patterns of the two countries have diverged significantly. Understanding these differences requires a comparative examination of their economic structures and crisis management practices, which also provides an appropriate framework for interpreting regional corporate data. The analysis is based on NUTS 3-level aggregated corporate indicators—number of companies, revenues, and liabilities—retrieved from the EMIS database for both countries. To comprehensively evaluate the economic position of the regions, these data were compared with GDP and gross value added (GVA) figures available from Eurostat. The spatial visualization of territorial disparities was carried out using QGIS software. The goal is to draw conclusions regarding the correlations between various indicators and the presence of corporate activity at the regional level. The results clearly reflect the structural differences between the two countries and highlight the geographical characteristics of corporate distribution.

Keywords: regional economics, corporate liabilities, geoinformatics

1. Bevezetés

A közép-kelet-európai átmeneti gazdaságok – így Magyarország és Csehország – gazdasági fejlődése az elmúlt két évtizedben számos hasonlóságot mutatott, azonban gazdasági struktúrájuk, válságokra adott válaszaik és fejlesztéspolitikai eszközeik jelentős eltéréseket tükröznek. Hárskúti [17] tanulmánya szerint a 2008-as válság Magyarországot sérülékenyebb pozícióban érte, melyet követően válaszként inkább pénzügyi stabilizációra és rövid távú politikai célokra koncentrált. Fejlesztéspolitikája inkább adminisztratív vezéreltségű volt, kevésbé kapcsolódott a piaci igényekhez. Ezzel szemben Csehországban a stabilabb pénzügyi szektornak és mérsékeltebb államadósságnak köszönhetően kevésbé volt súlyos a gazdasági visszaesés. Megoldásként hosszabb távú, strukturáltabb, exportvezérelt válaszokat alkalmazott, amelyek mögött stabilabb makrogazdasági alapok álltak. Fejlesztéspolitikája hosszú távú célokat követve jobban kapcsolódott az innovációhoz, K+F tevékenységekhez. Végh [26] elemzése megerősíti, hogy Magyarország erősen kitett volt a devizaalapú hitelezésnek, és fiskális helyzete is sebezhetőbbé tette a válságidőszakban. Csehország ezzel szemben mérsékeltebb államháztartási hiánnyal, alacsonyabb külső adóssággal és ellenállóbb bankrendszerrel rendelkezett. A két ország válságra adott válaszai közötti különbségek is jól tükrözik a tagállami válságkezelés diverzitását: míg Magyarország erőteljes állami szerepvállalásra és nem konvencionális lépésekre épített, Csehország visszafogottabb, piacbarátabb utat választott. Schweiger és Visvizi [23] politikai dimenzióban vizsgálta a két ország eltéréseit. Magyarország esetében az illiberális fordulat következtében a nemzeti szuverenitás kerül középpontban. Ezzel szemben Csehország inkább óvatos álláspontot képvisel, nem törekszik radikális különállásra. Lee [19] tanulmánya empirikusan igazolja, hogy Magyarországon, Csehországban és Lengyelországban az externális finanszírozáshoz való hozzáférés kulcsszerepet játszik a vállalatok növekedésében – különösen ott, ahol a belső források nem elegendők. Ezzel rámutat, hogy a hitelkínálat növelése ösztönözheti a gazdasági növekedést. A szerzők szerint az országoként eltérő pénzügyi és intézményi környezet eltérő mértékben befolyásolja a vállalatok teljesítményét. Körner [18] a cseh vállalati szféra adósságszerkezetének vizsgálata alapján arra jutott, hogy Csehországban a vállalatok pénzügyi struktúrája kiegyensúlyozott, jellemzően banki finanszírozást alkalmaznak, a hosszú lejáratú hitelek aránya egyre növekvő és megfigyelhető a lejáratillesség elve, ami a tőkeszerkezet válságállóság szempontjából is kedvezőbb. A gazdasági struktúrák közötti különbségeket az Európai Bizottság 2023 évi országrportjai [10, 11] megerősítik. A cseh vállalatok legfőbb

kihívásai között az energiaárak, munkaerőköltségek, globális értéklánchoz való erős kapcsolódás szerepelnek, míg Magyarországon a vállalati környezet kiszámíthatatlansága, a torzító hatású állami beavatkozások, és gyenge innovációs teljesítmény. Mindkét országban jelentős a regionális különbségek figyelhetők meg, a fővárosok a legfejlettebbek, míg a periférikus térségekben alacsonyabb a termelékenység, és a humán tőke minősége. Az EBRD [9] országjelentése szerint a két ország gazdasági helyzete között jelentős eltérések vannak. Közös pont, hogy a fellendülés expotvezérelt iparágaknak (főként autóipar) köszönhető. A jövőbeni reformok fókuszja mindkét országon a zöld átállás, digitalizáció, K+F, infrastruktúra. Magyarország esetében a Helyreállítási és Ellenállóképességi tervet még nem hagyták jóvá, így a kormány a beruházásokat előfinanszírozással hatja végre.

Jelen tanulmány aktualitását erősíti, hogy az Európai Bizottság 2023-ban több aspektusból is vizsgálta a vizsgálat a 2021–2027-es kohéziós időszak célkitűzéseit, kihívásait, és várható hatásait. A programozási jelentés [13] szerint az egyik fő kihívás az volt, hogyan lehet a COVID-19 utáni helyreállítást és a zöld, valamint digitális átmenetet integrálni a kohéziós politikába. A „tematikus koncentráció” és az éghajlatvédelemre szánt források arányának növelése is kiemelt célként jelent meg a programozás során. A Bizottság értékelési dokumentuma [12] a tervezési és értékelési rendszer jelentős egyszerűsítéseit emeli ki, például a teljesítménytartalék eltörlését és a haladási jelentések megszüntetését. A Bizottság 2024 végéig középtávú értékelést, 2029-ig hatásvizsgálatot, 2031-ig pedig visszatekintő értékelést tervez. A RHOMLO modell alapján készült ex-ante hatásvizsgálat [14] szerint a kohéziós programok pozitív gazdasági hatást gyakorolnak a beruházásokra és a foglalkoztatásra, különösen a kevésbé fejlett régiókban, ahol a relatív GDP-növekedés lehet a legjelentősebb. Ugyanakkor a kohéziós források multiplikátorhatását befolyásolja a nemzeti intézményi kapacitás, a társfinanszírozás mértéke, valamint az adminisztratív és szabályozási környezet hatékonysága.

A kohéziós politika kiemelt feladata a kevésbé fejlett térségek felzárkóztatása és a regionális egyenlőtlenségek mérséklése, amelyekhez a vállalati szektor strukturális jellemzőinek feltárása nélkülözhetetlen. A vállalati szintű vizsgálatok így jól kiegészítik a makrogazdasági és intézményi összehasonlításokat, lehetővé téve a kötelezettségállomány szerkezetének területi alapú elemzését is. A jelen kutatás célja, hogy a magyar és cseh vállalati kötelezettségállomány alakulását vizsgálja meg regionális bontásban, különös figyelemmel a kötelezettségek

mértékére és a vállalatok térbeli koncentrációjára. A vizsgálat eredményei hozzájárulhatnak a területi politikák célzottabb kialakításához és a kohéziós források hatékonyabb elosztásához.

2. Szakirodalmi háttér

A vállalati forrásszerkezet, hitelezés nem csupán a vállalatok működésének és növekedésének egyik alapvető feltétele, hanem jelentős makrogazdasági hatással is bír, különösen a gazdasági növekedés és a válságokra adott válaszok tekintetében. A szakirodalom szerint a hitelezés elősegíti a gazdasági fejlődést, valamint hozzájárul a jövedelmi egyenlőtlenségek mérsékléséhez [3]. Hosszú távú elemzések stabil kapcsolatot mutatnak a reálgazdasági hitelek és a nominális GDP alakulása között [22]. Az euróövezetben a KKV-k számára nyújtott hitelek kulcsszerepet játszanak a gazdasági növekedésben [24], és a nem pénzügyi vállalatoknak nyújtott hitelek növelése ösztönzi a vállalati növekedést [25].

A tanulmány módszertani kerete kvantitatív, vállalatszintű pénzügyi mutatókra épül, amelyek megalapozottságát számos nemzetközi, lektorált empirikus kutatás alátámasztja. Duchin és munkatársai [6] kimutatták, hogy a külső finanszírozás költségei a pénzügyi válság során jelentősen befolyásolták az amerikai vállalatok beruházási döntéseit, különös tekintettel a tőkeszerkezetre. Cowling és munkatársai [4] az Egyesült Királyságban a KKV-k körében vizsgálták, hogyan befolyásolták az állami beavatkozások és támogatások a válság utáni finanszírozást. Hasonlóan Bliss és munkatársai [3] összehasonlító elemzést végeztek a válság utáni KKV-hitelezésről, rámutatva a különböző országok közti eltérésekre.

Akbar és munkatársai [1] kvantitatív panelregressziós módszerrel elemezték, hogyan módosult a brit vállalatok tőkeszerkezete a 2008-as válság hatására. Dewally és Shao [5] a hitelpiaci szűkülés vállalati viselkedésre gyakorolt hatását vizsgálták. Fernández és munkatársai [16] elemzése szerint a válság utáni pénzügyi alkalmazkodás területileg is differenciált volt. Ozkan [20] pedig az adósság lejárat szerkezetét vizsgálta egyéni vállalatok szintjén, és megállapította, hogy a gazdasági környezet és az iparági sajátosságok jelentősen befolyásolják a pénzügyi döntéseket.

E kutatások együttesen igazolják, hogy a pénzügyi mutatók – különösen a kötelezettség/saját tőke és kötelezettség/mérlegfőösszeg arány – megfelelő indikátorok a vállalati pénzügyi stabilitás, a válságérzékenység és a területi különbségek vizsgálatára. A jelen tanulmány ezekre

az empirikusan igazolt megközelítésekre épít, területi szinten vizsgálva a vállalati adósság jellemzőit és földrajzi eltéréseit, így hozzájárulva a regionális fejlesztéspolitika és pénzügyi stabilitás közötti kapcsolat jobb megértéséhez.

3. Módszertan és eredmények

3.1 Adatbázisok és adatgyűjtési szempontok

A vállalati szintű adatok Magyarországra és Csehországra az EMIS [7] adatbázisból kerültek kinyerésre teljes körűen azon cégek vonatkozásában melyek 1 Eurónál nagyobb árbevétellel rendelkeznek. Az 1 eurós árbevétel korlát alkalmazására azért került sor, mert több mint 10 000 ilyen vállalkozás szerepelt az adatbázisban, ami meghaladta az exportálhatósági korlátot. Mivel a vizsgálat célja a működő vállalkozások elemzése, ez a szűrés érdemi torzító hatást nem jelentett. A vállalati kötelezettségállomány minél realisabb meghatározása érdekében a vizsgálat csak a működő nem pénzügyi szektorban tevékenykedő cégekre korlátozódott, vagyis a banki, biztosítási, pénzügyi tevékenységet folytató cégek kizárásra kerültek. Az első táblázat tartalmazza a kutatáshoz lekérdezett vállalatok számát arányát a két országban (lásd 1. táblázat) Annak érdekében, hogy minél teljesebb céglistával lehessen dolgozni, az EMIS adatbázisban a beszámoló évére nem került szűrés. Az adatbázisban minden cég így csak egyszer szerepel, és összességében a cégek jelentős része 2023 beszámoló adatokkal rendelkezik.

1. táblázat: A lekérdezett vállalatok száma és aránya Magyarországon és Csehországban

Ország	Eurostat 2023 vállalatok száma	Emis lekérdezett vállalatok száma	Emis Eurostat arány	Működő, nem pénzügyi szektor	2023 beszámoló aránya
Magyarország	901 020	77 676	8,62%	70 952	79,70%
Csehország	1 310 320	60 444	4,61%	56 664	72,70%

Forrás: Saját szerkesztés EMIS (2024) és Eurostat (2024) adatok alapján.

A vállalatok NUTS 3 szintű területi azonosítása az EMIS adatbázisban elérhető irányítószám, székhely, mezők alapján történt. Azokban az esetekben, ahol ezek hiányosak vagy nem egyértelműek voltak, manuálisan került sor a területi egység azonosítására a cég elnevezése és

címe alapján. A térségek gazdasági súlyának elemzéséhez a NUTS 3 szintű adatok az Eurostat [15] adatbázisból kerültek lekérésre. Az eredmények bemutatása térképi és táblázatos formában történt, amely lehetővé teszi a regionális eltérések vizuálisan is jól értelmezhető ábrázolását. A térképes megjelenítés a földrajzi különbségek azonosítását támogatja, míg a táblázatos forma a részletes összehasonlítását teszi lehetővé. E kettős megközelítés segíti a régiók közötti vállalati kötelezettség szintek strukturált értelmezését.

A tanulmányban alkalmazott pénzügyi mutatók – kötelezettség/saját tőke, kötelezettség/mérlegfőösszeg, kötelezettség/GDP, valamint egy cégre jutó kötelezettség – indokoltan tekinthetők a vállalati pénzügyi stabilitás és területi kockázatok megbízható indikátorainak. A kötelezettség/saját tőke arány a vállalat tőkeszerkezetének egészségi állapotára utal, jelezve, hogy a vállalat mennyire támaszkodik idegen forrásokra. A kötelezettség/mérlegfőösszeg arány a teljes eszközállomány finanszírozási szerkezetét mutatja meg, a túlzott eladósodottság kockázatára hívva fel a figyelmet. A kötelezettség/GDP mutató a vállalati eladósodottság makrogazdasági súlyát tükrözi, lehetővé téve a vállalati szektor területi jelentőségének összevetését a gazdasági teljesítménnyel. Az egy cégre jutó kötelezettség pedig a vállalati szintű pénzügyi teher átlagos mértékét érzékelteti, ami különösen fontos az eltérő vállalatsűrűségű régiók összehasonlításához. E mutatók alkalmazása széles körben elterjedt a nemzetközi szakirodalomban [8, 20, 21], a pénzügyi stabilitás makroszintű nyomon követésében.

3.2 Területi gazdasági teljesítmény GDP – GVA és foglalkoztatottsági adatok összehasonlítása

2023-ban a teljes magyar GDP érték 170 096 millió EUR érték volt, ezzel szemben a cseh teljes GDP 291 433 millió EUR. Ezen adatok jó tükrözik, hogy a cseh gazdaság teljesítménye jelentősen meghaladja Magyarorszáét. A második táblázat a legfejlettebb, és leelmaradottabb régió GDP, GVA és foglalkoztatottsági adatait tartalmazza (lásd 2. táblázat). NUTS 3 szintű GDP, GVA értékek alapján összeállított rangsor szerint, a két országot figyelembe véve a TOP 5-ben mindössze egy magyar térség, Budapest szerepel. Ezzel szemben az utolsó 5 helyezett között csak magyar régió található. (lásd 2. táblázat) Ez arra utal, hogy a cseh gazdasági teljesítmény nemcsak összességében nagyobb, hanem területileg is kiegyensúlyozottabban oszlik meg, míg Magyarországon Budapest dominanciája rendkívül erős, és a gazdasági

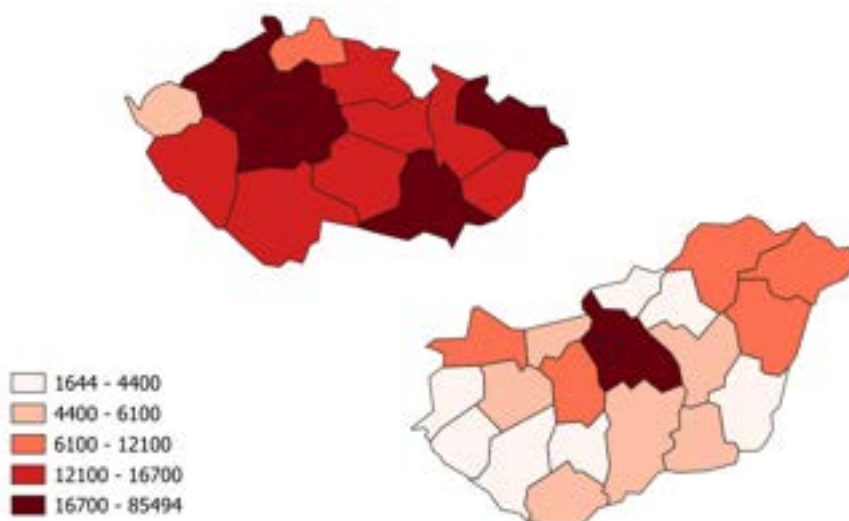
aktivitás jelentős része erre a központi régióra koncentrálódik. A területi GDP különbségeket a 1. ábra vizuálisan is jól szemlélteti. (lásd 1. ábra)

A gazdasági teljesítmény társadalmi háttérének érzékeltetése érdekében a régiók foglalkoztatottsági szintjeit is figyelembevételre kerültek. A TOP régiók foglalkoztatottsága jellemzően magas: Budapest 1,65 millió, Prága közel 1 millió foglalkoztatottal rendelkezik. Ezzel szemben az utolsó öt helyen szereplő magyar régiók – például Nógrád (49 ezer fő) vagy Tolna (82 ezer fő) – alacsony foglalkoztatási szintet mutatnak. Ez megerősíti, hogy a GDP és GVA mutatók mögött jelentős humánerőforrásbeli különbségek is meghúzódnak: a gazdaságilag fejlettebb régiókban a foglalkoztatottság is magasabb, míg a fejletlenebb térségekben alacsonyabb szintű munkaerőpiaci részvétel figyelhető meg.

2. táblázat: 2023-as GDP és GVA adatok alapján a legfejlettebb és legelmaradottabb térségek

TOP 5 régió	GDP (m EUR)	GVA (m EUR)	Foglalkoztatottság (e fő)
Hlavní město Praha	85 494	78 503	985
Budapest	75 373	65 199	1 653
Středočeský kraj	37 757	34 670	607
Jihomoravský kraj	34 587	31 759	606
Moravskoslezský kraj	27 292	25 060	555
BOTTOM 5 régió	GDP (m EUR)	GVA (m EUR)	Foglalkoztatottság (e fő)
Zala	3 669	3 174	102
Somogy	3 647	3 155	99
Békés	3 592	3 107	107
Tolna	2 931	2 536	82
Nógrád	1 644	1 422	49

Forrás: Saját szerkesztés Eurostat (2024) adatok alapján.



1. ábra: NUTS 3 szintű területi egységek GDP értékei Magyarországon és Csehországban 2023 (m EUR)

Forrás: Saját számítás Eurostat adatok alapján, térképkészítés QGIS 3.40.4 szoftverrel.

3.3 Kötelezettségállomány terheltség regionális elemzése

A kötelezettségállomány-terheltség területi elemzése során bemutatásra kerülnek a vállalatok kötelezettségeinek főbb mutatói, valamint azok területi különbségei a magyar és cseh régiók szintjén.

Összes kötelezettség területi eloszlása

A vállalatok összesített kötelezettségállománya alapján megfigyelhető, hogy mindkét országban a fővárosok – Budapest és Prága – emelkednek ki a mezőnyből. Budapest 168 366 millió euróval, míg Prága 163 947 millió euróval rendelkezik, amely országos szinten is domináns pozíciót tükröz. Emellett jelentős értéket képvisel a Pest megyei (34 821 millió EUR) és Středočeský (26 540 millió EUR) régió is, utóbbi Prága körüli gyűrűként működik. A cseh Moravskoslezský és Ústecký régiók is kiemelkednek a több mint 20 milliárd eurós kötelezettségállománnyal, jelezve a régiók erőteljes ipari és vállalati koncentrációját. Magyarországon figyelemre méltó a Győr-Moson-Sopron megyei kötelezettségállomány is, ahol az erős ipari koncentráció – különösen az autóiipari jelenlét – indokolja a magas értéket. A következő legnagyobb kötelezettségállomány Hajdú-Bihar megyében található, amely Debrecen térségének gazdasági fejlődésével, és az elmúlt évek intenzív beruházásaival magyarázható.

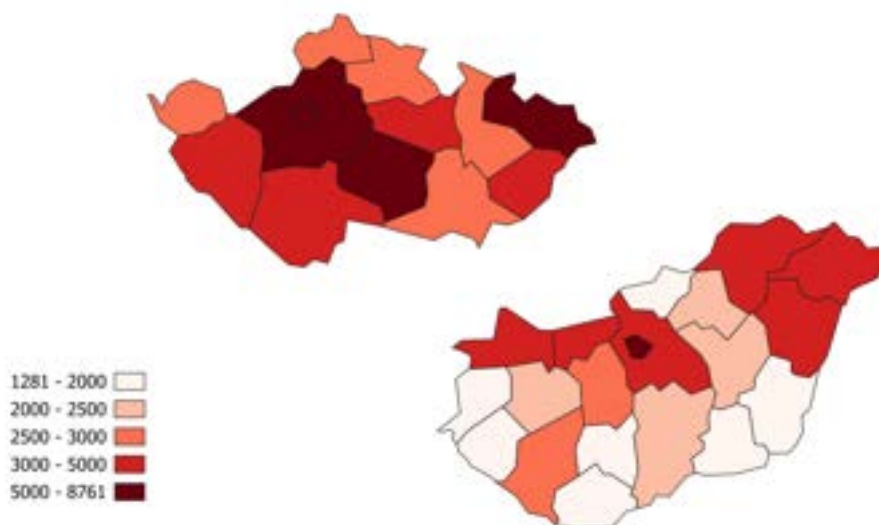
Egy cégre jutó kötelezettségállomány

A régiós cégek átlagos kötelezettségállománya alapján Csehországban Ústecký (8 761 e EUR/cég) és Prága (8 402 e EUR/cég) emelkedik ki, míg Magyarországon Budapest (6 631 e EUR/cég) vezet, ami a központi vállalati koncentrációra utal. A Moravskoslezský és Středočeský cseh térségekben az érték szintén meghaladja az 5 000 ezer EUR-t, míg a magyar vidéki térségek jellemzően 2000–4000 ezer EUR közé esnek. Az adatok azt mutatják, hogy Csehországban jellemzően nagyobb cégek, koncentráltabb kötelezettséggel működnek. A magasabb átlagos értékek nagyobb pénzügyi kitettséget jeleznek, ami a vállalatok pénzügyi stabilitásának és növekedési potenciáljának eltérő mintázatára világít rá a két ország régiói között. A 3. táblázat a kötelezettségállomány alapján számított öt legmagasabb egy cégre jutó értékkel rendelkező térséget mutatja be, míg a 2. ábra teljeskörű térképes áttekintést nyújt Magyarország és Csehország összes NUTS 3 szintű területi egységéről, lehetővé téve a területi mintázatok átfogó vizsgálatát (lásd 3. táblázat, 2. ábra)

3. táblázat: TOP 5 egy cégre jutó átlagos kötelezettségállomány térségeinkét

Térség	NUTS3 kód	Egy cégre jutó kötelezettség (e EUR)	Cégek száma
Ústecký kraj	CZ042	8 761	2 623
Hlavní město Praha	CZ010	8 402	19 512
Budapest	HU101	6 631	25 390
Moravskoslezský kraj	CZ080	5 185	4 976
Středočeský kraj	CZ020	5 177	5 127

Forrás: Saját szerkesztés EMIS (2024) adatok alapján



2. ábra: Egy cégre jutó vállalati kötelezettségállomány területi eloszlása (ezer EUR/vállalat)

Forrás: Saját számítás EMIS vállalati adatok alapján, térképkészítés QGIS 3.40.4 szoftverrel.

A vállalati kötelezettségek területileg is különbségeket tükröznek. Az egy cégre jutó kötelezettség statisztikai mutatóit a 4. táblázat foglalja össze (lásd 4. táblázat). A vizsgált régiók közül a legalacsonyabb egy cégre jutó kötelezettség Tolna megyében mutatkozott (1 281 ezer euró), míg a legmagasabb érték Prágában, a cseh fővárosban jelent meg (8 761 ezer euró). Az átlagos kötelezettség 3 354 ezer euró, de a szórás magas (1 787 ezer euró), ami a területi szóródás erősségére utal. A medián (2 855 ezer euró) az átlag alatti elhelyezkedése szintén arra utal, hogy a régiók többségében az egy vállalkozásra jutó kötelezettség az átlagérték alatt marad, és néhány kiugróan magas érték felfelé torzítja az átlagot.

4. táblázat: Egy cégre jutó kötelezettség statisztikai mutatói (e EUR)

Mutatók	Egy cégre jutó kötelezettség (e EUR)	Régió
Minimum	1 281	Tolna
Maximum	8 761	Hlavní město Praha
Átlag	3 354	
Szórás	1 787	
Medián	2 855	

Forrás: Saját szerkesztés EMIS (2024) adatok alapján

Vállalati kötelezettségállomány saját tőkéhez viszonyított aránya

Az 5. táblázat adatai bemutatják a vállalati kötelezettségállomány saját tőkéhez és GDP-hez viszonyított arányát, valamint a foglalkoztatottsággal való összefüggéseket is (lásd 5. táblázat).

A kötelezettségállomány saját tőkéhez viszonyított aránya jól kifejezi, hogy egyes régiókban milyen mértékű a külső forrásbevonás saját forráshoz képest. Prága (181,7%), Budapest (158,4%) és Győr-Moson-Sopron (174,9%) esetében az arány több mint másfélszerese a saját tőkének, amely fokozott pénzügyi kockázatot jelez. Ezekben a régiókban a GDP-érték is kiemelkedően magas, és jelentős ipari koncentráció figyelhető meg, ami részben indokolja a magasabb külső forrásigényt. Szintén magas értéket mutat Nógrád (174,7%) és Szabolcs-Szatmár-Bereg (142,1%) megyékben, alacsonyabb tőkeerőre utalhat. A legalacsonyabb arányokat a cseh Královéhradecký (58,9%) és Zlínský (60,2%) régiókban találjuk.

Vállalati kötelezettségállomány regionális GDP-hez viszonyított aránya

A mutató meghatározza, hogy mekkora pénzügyi teher nehezedik az adott térség gazdasági teljesítményére. Budapest (223,4%), Pest (149,63%) és Győr-Moson-Sopron (126,3%) kimagasló értékei azt jelzik, hogy e térségek vállalati kötelezettségállománya jelentősen meghaladja a régió bruttó hazai termékét. A cseh régiók közül Prága (191,8%), Hlavní město Praha (191,76%) és Ústecký (131,6%) hasonlóan magas arányokat mutatnak. Ezen adatok egyrészt fejlett vállalati jelenlétre utalnak, másrészt a gazdaság terheltségét is jelzik. (lásd 5. táblázat)

Vállalati kötelezettségállomány, gazdasági teljesítmény és foglalkoztatottság összefüggései

A kötelezettségállomány és a régiók munkaerő-piaci kapacitása közötti kapcsolat vizsgálata lehetővé teszi a gazdasági terhelés társadalmi dimenziójának értelmezését. A legnagyobb kötelezettségállománnyal rendelkező régiók – mint Budapest (168 milliárd EUR, 1,65 millió foglalkoztatott) és Prága (163 milliárd EUR, 985 ezer fő) – esetében a magas pénzügyi volumen kiterjedt munkaerőpiaci háttérrel párosul, amely képes lehet ellensúlyozni a pénzügyi kockázatokat. Ezzel szemben több olyan régió – például Szabolcs-Szatmár-Bereg (7,2 milliárd EUR kötelezettség, 175 ezer fő) vagy Somogy (3,7 milliárd EUR kötelezettség, 99 ezer fő) – is szerepel a listán, ahol az alacsony foglalkoztatottság mellett a kötelezettségállomány a GDP-hez viszonyítva mégis jelentős. Ez arra utal, hogy az adott térségek vállalatai a gazdasági teljesítményhez és humán erőforráshoz képest aránytalanul nagy külső forrásokat használnak,

ami fokozott sérülékenységet jelezhet. A foglalkoztatottsági adatok így fontos kontextust biztosítanak a regionális pénzügyi kockázatok értékeléséhez.

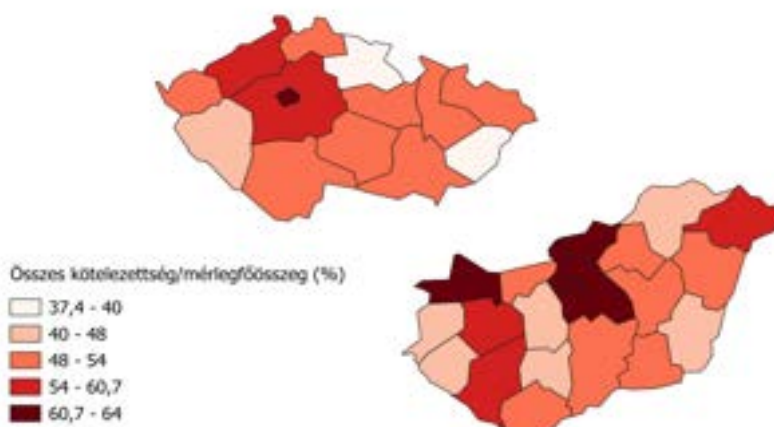
5. táblázat: A térségek vállalati kötelezettségállománya és kapcsolódó mutatói a tíz legnagyobb érték alapján

Régió	Ország	Összes kötelezettség (m EUR)	Foglalkoztatottság (e fő)	Összes kötelezettség/saját tőke	Összes kötelezettség / GDP
Budapest	Magyarország	168 366	1 653	158,44%	223,38%
Hlavní město Praha	Csehország	163 947	985	181,73%	191,76%
Pest	Magyarország	34 821	529	166,90%	149,63%
Ústecký kraj	Csehország	22 981	356	119,32%	131,64%
Győr-Moson-Sopron	Magyarország	12 621	232	174,87%	126,33%
Hajdú-Bihar	Magyarország	9 355	206	100,16%	120,70%
Komárom-Esztergom	Magyarország	7 176	146	96,97%	118,61%
Szabolcs-Szatmár-Bereg	Magyarország	7 225	175	142,14%	115,95%
Bács-Kiskun	Magyarország	6 920	196	114,86%	114,32%
Somogy	Magyarország	3 715	99	139,53%	101,88%

Forrás: Saját szerkesztés EMIS (2024) adatok alapján

Vállalati kötelezettségállomány mérlegfőösszeghez viszonyítva

A 3. ábra szemlélteti, hogy az adott régió vállalatai a mérlegfőösszegük mekkora arányát finanszírozzák kötelezettségekből. (lásd 3. ábra) A két ország összehasonlítása alapján elmondható, hogy Csehországban összességében mérsékeltebb a mérlegfőösszeghez viszonyított eladósodottság. A legmagasabb arányok Prágában (64%) és Budapesten (61,2%) tapasztalhatók, amelyek a pénzügyi és gazdasági központ szerepből adódnak. A magyarországi régiók közül Pest (62,5%), Győr-Moson-Sopron (63,6%) és Szabolcs-Szatmár-Bereg (58,7%) emelkedik ki, míg Csehországban Středočeský (55,9%) és Ústecký (54,4%) vezet. A magas értékek erőteljesebb eladósodottságot jeleznek, ugyanakkor összefüggésbe hozhatók a régiók gazdasági aktivitásával is.



6. ábra: A térségek vállalati kötelezettségállománya és kapcsolódó mutatói a tíz legnagyobb érték alapján

Forrás: Saját szerkesztés EMIS (2024) adatok alapján

4. Összegzés

Az Európai Unió kohéziós politikájának alapvető célja a regionális egyenlőtlenségek csökkentése és a fenntartható fejlődés támogatása. A szakirodalmi áttekintés alapján, a pénzügyi és gazdasági válságok új kihívásokat állítottak e politika elé, különösen a közép-kelet-európai tagállamok esetében. A válságkezelési tapasztalatok, a területi sajátosságok, valamint a fejlesztéspolitikai és intézményi válaszok mind hozzájárulnak a regionális különbségekhez.

Magyarország és Csehország példáján keresztül jól vizsgálhatók ezek a folyamatok, mivel hasonló történelmi és gazdasági pályát jártak be, ugyanakkor a kohéziós források felhasználásában, valamint a gazdasági válaszlépésekben eltérő stratégiákat alkalmaztak. A vállalati szintű elemzés ezen eltérések jobb megértéséhez járul hozzá.

A kutatás célja a két ország vállalati kötelezettségstruktúrájának területi összehasonlítása, különös tekintettel a NUTS 3 szintű térségek gazdasági teljesítményére. Az adatgyűjtés az EMIS és Eurostat adatbázisokra épül, és több mint 130 000 működő, nem pénzügyi vállalatra terjed ki. A vizsgálat során olyan mutatók kerültek meghatározásra, mint a kötelezettség / mérlegfőösszeg arány, a kötelezettség / saját tőke arány, az egy cégre jutó kötelezettség, valamint a kötelezettségek GDP-hez viszonyított aránya.

Az alkalmazott módszertan lehetővé teszi a vállalati pénzügyi helyzetek területi különbségeinek bemutatását és a gazdasági koncentrációk azonosítását. A vizsgálat térképi és

táblázatos elemzései rávilágítanak arra, hogy a két ország térségei között jelentős különbségek mutatkoznak a vállalati eladósodottság szintjében. A legnagyobb kötelezettségállomány a fővárosokban – Budapesten és Prágában – koncentrálódik, magasabb értékek még a jelentősebb gazdasági súly képviselő térségekben figyelhetők meg.

A kötelezettség és mérlegfőösszeg, valamint a saját tőke arányai alapján kimutatható, hogy a legnagyobb eladósodottsági szintek nemcsak a fejlett, központi, hanem egyes hátrányos térségekre is jellemzőek. A GDP-hez viszonyított kötelezettségarány azt mutatja, hogy több térség a vállalati szektor pénzügyi terhelése meghaladja a térség gazdasági teljesítményét, különösen Budapesten és Prágában. Az eredmények alapján megállapítható, hogy a területi pénzügyi egyenlőtlenségek mindkét országban fennállnak, és ezek figyelembevétele elengedhetetlen a hatékony fejlesztéspolitika és pénzügyi beavatkozás szempontjából.

5. Hasznosíthatóság és jövőbeni kutatási irányok

Gazdaságpolitikai döntéshozatal szempontjából az eredmények felhasználhatók a régiós támogatási programok újratervezéséhez, például térségek célzott fejlesztési forrásokhoz juttatása (pl. kamattámogatott hitelek, garanciaprogramok).

A vállalkozások számára a vizsgálat iránymutatást nyújthat pénzügyi pozíciójuk területi viszonylatú értékeléséhez. Egy adott cég a bemutatott mutatók alapján elhelyezheti magát a regionális átlaghoz képest, ami elősegíti saját finanszírozási stratégiájának kialakítását, illetve benchmarking céllal hasznosítható.

Válsághelyzet utáni újraértékeléshez a tanulmány által alkalmazott indikátorstruktúra alkalmas monitoring rendszerként történő alkalmazásra. A COVID-19, az energiaár-válság vagy jövőbeli gazdasági sokkok után ugyanezen mutatók révén jól követhetők a vállalati pénzügyi reakciók és azok területi eltérései.

Döntéstámogató rendszerekhez is integrálható a vizsgálat eredménystruktúrája, különösen a pénzintézetek, önkormányzatok és regionális gazdaságfejlesztési szereplők számára.

A társadalmi hasznosíthatóság vonatkozásában a vizsgálat segíthet az egyenlőtlen térségi fejlesztési lehetőségek igazságosabb tervezésében. A kötelezettségi szint magas értékei figyelmeztető jelzést adhatnak a munkahelyek fenntarthatóságára, a beruházási aktivitás csökkenésére. Ugyanakkor példaértékűnek tekinthetők azok a térségek (pl. Brno környéke,

vagy Győr-Moson-Sopron megye), ahol a magas kötelezettségszint ellenére sikerült versenyképes, fenntartható gazdasági struktúrát kialakítani, melyet munkahelybővítés és magasabb innovációs teljesítmény is kísért.

Jövőbeni kutatási irányként a tanulmányban alkalmazott módszertan kiterjeszthető több évre és további országokra, különösen a V4-ek (Magyarország, Csehország, Lengyelország, Szlovákia) összehasonlító elemzésére. Egy 3–5 éves időszakot felölelő trendvizsgálat lehetővé tenné a válság utáni alkalmazkodási minták feltárását, valamint annak vizsgálatát, hogy a kohéziós politika eszközei és a nemzeti fejlesztéspolitikák milyen módon befolyásolták a vállalati pénzügyi szerkezetek alakulását területi bontásban. Az időbeli kiterjesztés lehetőséget nyújtana a strukturális különbségek tartósságának vizsgálatára, valamint a régiók pénzügyi alkalmazkodóképességének elemzésére is. A kutatás továbbá kiegészíthető érzékenységvizsgálattal, amely lehetővé teszi a legnagyobb hatást gyakorló mutatók azonosítását, valamint különböző gazdasági forgatókönyvek modellezését. További pénzügyi mutatók – például likviditási mutatók – bevonása, valamint a vállalati méretkategóriák szerinti bontás új összefüggések feltárását segítené elő a régiók eltérő gazdasági struktúráinak tükrében. A pénzügyi eladósodottság és a környezeti fenntarthatóság kapcsolata szintén fontos irány. Az eladósodottabb térségekben működő vállalatok valószínűsíthetően kisebb arányban hajtanak végre környezetvédelmi beruházásokat, ami befolyásolhatja a zöld átállás sikerességét. A vállalati kötelezettségstruktúra és az ökológiai lábnyom összefüggésének vizsgálata – különösen a Green Deal célkitűzései mentén – hozzájárulhat a fenntartható és kiegyensúlyozott regionális fejlődés feltételeinek meghatározásához.

6. Köszönetnyilvánítás

A publikációban szereplő kutatást a Széchenyi István Egyetem az Európai Unió támogatásával valósította meg, az RRF-2.3.1-21-2022-00004 azonosítójú, Mesterséges Intelligencia Nemzeti Laboratórium projekt keretében.

Irodalomjegyzék

- [1.] Akbar, Shoaib – Rehman, Saeed – Ormrod, Philip: The Impact of Recent Financial Shocks on the Financing and Investment Policies of UK Private Firms. *International Review of Financial Analysis*, 26, 2013, 59–70. <https://doi.org/10.1016/j.irfa.2012.05.004>
- [2.] Beck, Thorsten – Degryse, Hans – De Haas, Ralph – van Horen, Neeltje: When Arm's Length Is Too Far: Relationship Banking over the Credit Cycle. *Journal of Financial Economics*, 127(1), 2018, 174–196. <https://doi.org/10.1016/j.jfineco.2017.11.007>
- [3.] Bliss, Barbara A. – Cheng, Ying – Denis, David J.: Corporate payout, cash retention, and the supply of credit: Evidence from the 2008–2009 credit crisis. *Journal of Financial Economics*, 115(3), 2015. 521–540. <https://doi.org/10.1016/j.jfineco.2014.10.013>
- [4.] Cowling, Marc – Liu, Weixi – Zhang, Na: Access to Bank Finance for UK SMEs in the Wake of the Recent Financial Crisis. *International Journal of Entrepreneurial Behaviour and Research*, 22(6), 2016, 903–932. <https://doi.org/10.1108/IJEBR-06-2015-0126>
- [5.] Dewally, Michaël – Shao, Yingyng: Liquidity Crisis, Relationship Lending and Corporate Finance. *Journal of Banking and Finance*, 39(1), 2014, 223–239. <https://doi.org/10.1016/j.jbankfin.2013.11.002>
- [6.] Duchin, Ran – Ozbas, Oguzhan – Sensoy, Berk A.: Costly External Finance, Corporate Investment, and the Subprime Mortgage Credit Crisis. *Journal of Financial Economics*, 97(3), 2010, 418–435. <https://doi.org/10.1016/j.jfineco.2009.12.008>
- [7.] EMIS Database: Vállalati pénzügyi adatok – Magyarország és Csehország, 2024.
- [8.] ESRB – European Systemic Risk Board (2023): Corporate credit risk and leverage in the EU: systemic implications and policy options. European Systemic Risk Board, Luxembourg: Publications Office of the European Union. <https://data.europa.eu/doi/10.2849/232900>
- [9.] European Bank for Reconstruction and Development (EBRD): Transition Report 2021–22: System Upgrade: Delivering the Digital Dividend – Central Europe and the Baltic States. London, 2021.
- [10.] European Commission: Country Report – Czechia (2023).

- [11.] European Commission: Country Report – Hungary (2023).
- [12.] European Commission: Evaluation of Cohesion Policy in 2021–2027. DG REGIO, Evaluation and European Semester, Brussels, 2023. Elérhető: https://ec.europa.eu/regional_policy/en/policy/evaluations/guidance/#1
- [13.] European Commission: Outcome of the 2021–2027 Cohesion Policy Programming: First Assessment. Publications Office of the European Union, Luxembourg, 2023.
- [14.] European Commission: The RHOMOLO Model: Ex-ante Impact Assessment of Cohesion Policy 2021–2027. Publications Office of the European Union, Luxembourg, 2023. Elérhető: <https://europa.eu/!GvwdF4>
- [15.] Eurostat: Regionális GDP, GVA, foglalkoztatottság és cégszámadatok NUTS 3 szinten, 2024.
- [16.] Fernández, Ana I. – González, Francisco – Suárez, Natalia: Bank Supply Shocks and the Substitution Between Bank and Nonbank Debt. *Journal of Corporate Finance*, 48, 2018, 122–147. <https://doi.org/10.1016/j.jcorpfin.2017.10.010>
- [17.] Hárskúti János: Magyarország válság utáni pénzügyi helyzete és a fejlesztéspolitika irányváltása. In: *Multidiszciplináris kihívások – Sokszínű válaszok*. BGF KVIK, 2012.
- [18.] Körner, Petr: The determinants of corporate debt maturity structure: evidence from Czech firms. IES Working Paper 27/2006, Charles University in Prague.
- [19.] Lee, Youngjin: The Role of External Financing in Firm Growth: Evidence from Central Europe. *International Journal of Economics and Business*, 2023.
- [20.] Ozkan, Aydın: An Empirical Analysis of Corporate Debt Maturity Structure. *European Financial Management*, 6(2), 2000, 197–212. <https://doi.org/10.1111/1468-036X.00120>
- [21.] Palliam, Ralph – Sbeiti, Wafaa – Ghosh, Dilip K.: Corporate Debt and Equity: Another Look at Their Determinants. *Frontiers in Finance and Economics*, Vol. 10, No. 2, 2013, 31–62.
- [22.] Ryan-Collins, Josh – Werner, Richard A. – Castle, Joseph: A Half-Century Diversion of Monetary Policy? An Empirical Horse-Race to Identify the UK Variable Most Likely to Deliver the Desired Nominal GDP Growth Rate. *Journal of International Financial Markets, Institutions and Money*, 43, 2016, 158–176.

- [23.] Schweiger, Christian – Visvizi, Anna (szerk.): Central and Eastern Europe in the EU: Challenges and Perspectives Under Crisis Conditions. Routledge Studies in the European Economy, London–New York, 2018.
- [24.] Tondl, Gabriele: Interest Rates, Corporate Lending and Growth in the Euro Area. Előadás, 2016. (Ha van pontos megjelenési hely vagy publikáció, kérlek egészítsd ki.)
- [25.] Vasilyeva, Regina – Turygin, Oleg – Ie, Olga – Kozlova, Maria: The Impact of Lending to Non-Financial Corporations on Economic Growth. Proceedings of CBU in Economics and Business, 2, 2021, 111–115.
- [26.] Végh Márton Zsolt: Válságkezelés innovatív módszerekkel – Közép-Európa EU-tagállamai példáin keresztül megközelítve. In: Bajmócy Zoltán, Elekes Zsuzsanna (szerk.): Innováció: a vállalati stratégiától a társadalmi stratégiáig. JATEPress, Szeged, 2013. 45–60.

Korszerű Neurális Hálózatok Összehasonlító Elemzése az Útszegmentálás Területén

Comparative Analysis of Modern Neural Networks for Road Segmentation

Reisinger Richárd^a, Hollósi János^b

^aSzéchenyi István Egyetem

reisinger.richard@hallgato.sze.hu

^bSzéchenyi István Egyetem, Járműipari Kutatóközpont

hollosi.janos@sze.hu

Absztrakt

Munkánk célja, hogy összehasonlítsuk a korszerű neurális hálózati architektúrákat egy útszegmentálási feladatban, és felmérjük teljesítményüket az adathalmaz segítségével. A kísérletek során a „Segmentation Models” Python könyvtár segítségével különböző komplexitású hálózati architektúrákat tanítunk és értékelünk, majd az eredmények alapján rangsort állítunk fel. A modellek kiválasztásánál és összevetésénél kiemelt szempont a pontosság, a feldolgozási sebesség és az erőforrás-igény. Továbbá figyelmet fordítunk a tanítási idő és a memóriahasználat alakulására is, hiszen a gyakorlati bevezetésnél ezek szintén alapvető korlátok lehetnek. Az így megszerzett tapasztalatok hozzájárulhatnak a hatékonyabb szegmentáló hálózatok kifejlesztéséhez, amelyek révén javul a feldolgozás minősége, és végső soron fokozódik az autonóm járművek biztonsága és megbízhatósága. A kísérlet célja ezen felül, hogy felmérjük a könyvtár gyakorlati alkalmazhatóságának a lehetőségét a téma fókuszában. Eredményeink rávilágítanak arra, milyen irányokba érdemes továbbfejleszteni a jelenlegi modelleket, hogy még hatékonyabb megoldásokat kapjunk.

Kulcsszavak: útszegmentálás, mély neurális hálózatok, modell összehasonlítás

Abstract

The aim of this study is to compare state-of-the-art neural network architectures in the context of road surface segmentation and evaluate their performance using the KITTI dataset. During the experiments, various network architectures of differing complexity are trained and assessed using the „Segmentation Models” Python library. Based on the results, a ranking of the models is established. Key evaluation criteria include segmentation accuracy, processing speed, and resource demands. Special attention is also given to training time and memory usage, as these pose significant practical constraints in deployment scenarios. The insights gained from this analysis can contribute to the development of more efficient segmentation networks, ultimately enhancing processing quality and improving the safety and reliability of autonomous vehicles. Additionally, this study aims to assess the practical applicability of the library within the domain of our topic. Our findings highlight promising directions for further development of current architectures toward more effective solutions.

Keywords: road segmentation, deep neural networks, model benchmarking

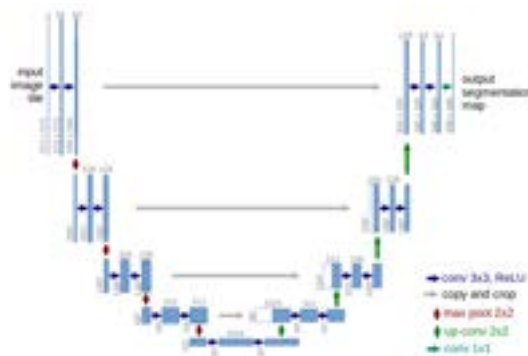
1. Bevezetés

A pontos úttestsegmentálás elengedhetetlen az autonóm járművek és vezetéstámogató rendszerek biztonságos és megbízható működéséhez. Az elmúlt tíz évben elkészített neurális hálózatok – különösen az encoder-decoder architektúrák – lényegesen javították a segmentálási pontosságot, még korlátozott mennyiségű tanítóadat esetén is. Jelen kutatás, öt, gyakran alkalmazott hálózatot mutat be, illetve hasonlít össze a KITTI adatcsomag egy részhalmazán tanítva. A modelleket a „Segmentation Models” Python-könyvtár segítségével valósítjuk meg és értékeljük ki. A modellek ugyanazokkal a paraméterekkel és ugyanazzal a forráskóddal lettek futtatva, ezt figyelembe véve egyik sem teljesen optimalizált formában kerül összehasonlításra, hiszen elsődleges cél az egyenlő feltételek alkalmazása volt. Az összehasonlítás során a segmentálási pixelpontosságot, a futási időt és a memóriaigényt vizsgáljuk, tekintettel a valós idejű alkalmazhatóság követelményeire. Továbbá bemutatásra kerülnek az architektúrák elméleti alapjai, a használt adatcsomag, illetve a használt szubkönyvtár. Különböző vizualizációkkal fogjuk szemléltetni a kutatás eredményeit és ezen eredmények közötti különbségeket, kiemelve részleteket. Összegzésünkben kitérünk a kutatás továbbfejlesztésére, illetve a jövőbeli terveket vázoljuk fel a témával kapcsolatban.

2. A használt neurális hálózatok

U-net

A U-Net nevét az „U” betűt formázó architektúrájáról kapta. (1. ábra). Klasszikus encoder-decoder felépítésű konvolúciós neurális hálózat, amit elsősorban orvosi képek segmentálására fejlesztettek ki, de esetünkben alkalmazható útszegmentálásra is.

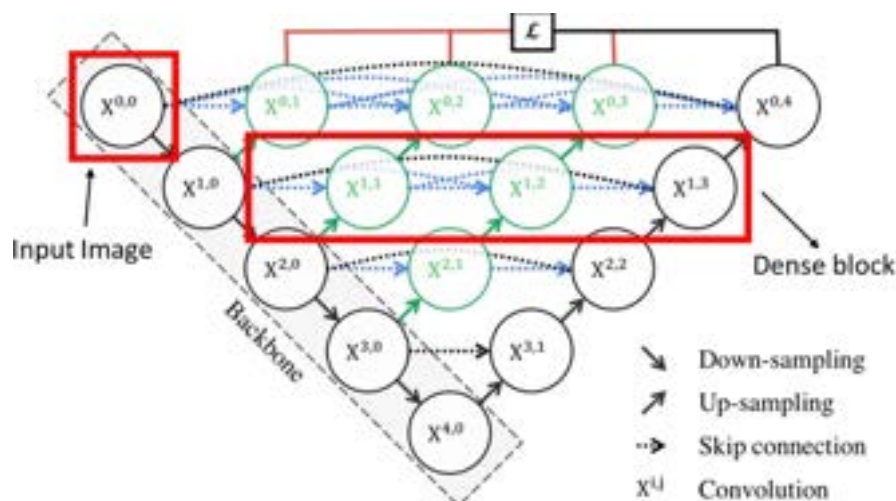


1. ábra: U-Net felépítése

Bal oldali ágon található az encoder, mely többszintű konvolúciós és pooling rétegek segítségével egyre absztraktabb jellemzőket emel ki, míg a jobb oldali decoder rétegek fokozatosan visszaállítják a térbeli felbontást (1. Ábra). A két ágot úgynevezett skip-kapcsolatok kapcsolják össze: az encoder egyező mélységű kimeneteit, összekötik a decoder azon rétegével, ahol egyezik a felbontás, ezzel közvetlen átküldve a finom részleteket, amik esetleg eltűnhetnek a jellemzők összesűritése során. Ezen tulajdonságainak köszönhetően a U-Net különösen jól teljesít, korlátozott tanítóadatmennyiség esetén. [1]

U-Net++

A U-Net++ tulajdonképpen a klasszikus U-Net egy továbbfejlesztése, mely több beágyazott skip-kapcsolatot, illetve kollektív finomítást alkalmaz. Ahelyett, hogy az encoder dekóder ágot közvetlenül összekapcsolná, hierarchikusan rendezi a skip-kapcsolatokat több szinten (2. ábra), így a jellemzőtérképek a korábbi konvolúciós blokk kimenetére is visszacsatolódnak, mielőtt a decoder rétegbe kerülnek. Ez a struktúra csökkenti az ún. „semantic gap-et” az encoder és decoder között, ezzel részletgazdagabb és finomabb maszkokat eredményezve. Memóriaigényessége révén a tanítási ideje jelentősen hosszabb, de tekintve a struktúra mélységét és bonyolultságát ez indokolt lehet. [2]



2. ábra: U-Net++ struktúrája beágyazott skip-kapcsolatokkal

MANet

A MANet (avagy Multi-scale Attention Network) a szegmentálási hatékonyságot ún. Attention (figyelem) mechanizmusok bevezetésével emeli. Architektúrát tekintve szintén egy encoder-decoder modellről van szó, ám a hagyományos skip-kapcsolatok helyett speciális attention

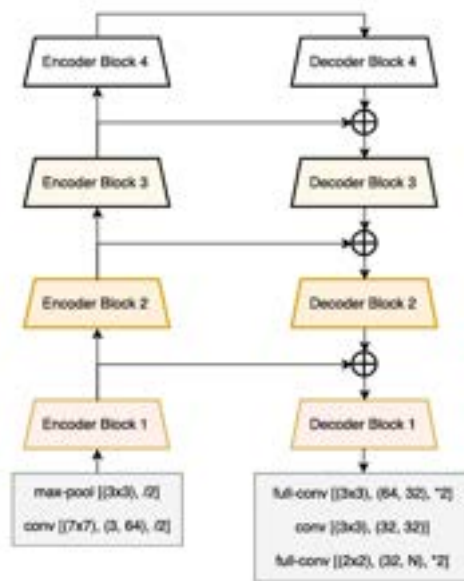
modulok finomhangolják az egyes jellemzőtérképeket. Ezek az attention blokkok több skálán képesek észlelni a kritikus jellemzőket (3. ábra). Figyelemértékeket rendelnek a különböző felbontású jellemzőkhöz, így a modell képes a releváns részleteket a háttérrel. Előnyét tekintve, a többfelbontású figyelem révén jobb az általánosító képessége, viszont ezzel szemben ezen modulok integrálása némileg megnöveli a számítási igényt, a hagyományos skip-kapcsolatos modellekhez képest. [3]



3. ábra: MANet hálózati architektúrája

LinkNet

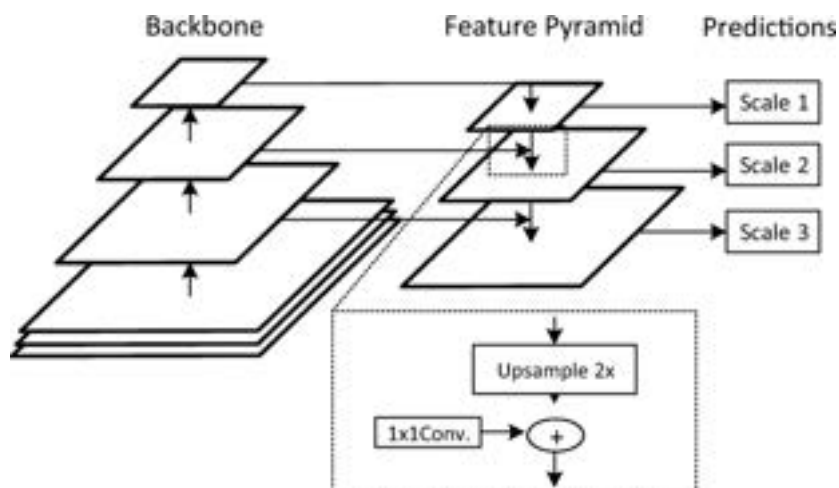
A LinkNet egy kifejezetten könnyű encoder-decoder hálózat, amelyet kifejezetten valós idejű alkalmazásokhoz fejlesztettek. A szerkezet lényege, hasonlóan a skip-kapcsolatokhoz, hogy az encoder és decoder rétegek között egyszerű „link” blokkokkal továbbítja a jellemzőket. (4. ábra) Minden encoder-blokk kimenete közvetlenül összeköttetésbe kerül a megfelelő mélységű decoder-blokkal. Ez az egyszerűbb adatútvonal lehetővé teszi az alacsonyabb paraméterszámot és memóriaterhelést, miközben elfogadható pontosságot biztosít. A LinkNet különösen alkalmas beágyazott rendszerekbe történő integrációra vagy olyan GPU alapú alkalmazásokhoz, ahol kritikus a kis késleltetés és a korlátozott hardvererőforrás. [4]



4. ábra: A LinkNet hálózati architektúrája

FPN

Az FPN (avagy Feature Pyramid Network) elsősorban objektumfelismerésben és általános képanalízisben terjedt el, de tökéletesen adaptálható útszegmentálási feladatokra is. Felépítését tekintve egy piramisstruktúrát követ: az erősebb, alacsony felbontású jellemzőkből és a gyengébb, magas felbontásúakból egyesíti a semantikus és a finom részletező információkat (5. Ábra). A megközelítés különösen előnyös komplex környezetekben. [5]



5. ábra: Az FPN hálózati architektúrája

3. A felhasznált könyvtár, adatcsomag

Segmentation Models Pytorch

A Segmentation Models Pytorch egy nyílt forráskódú Python könyvtár, amely rendkívül könnyen használható és hatékony megoldásokat kínál képfeldolgozási feladatokra, ezek között is különösen a szegmentáció területén. A könyvtár a PyTorch deep learning keretrendszerére épül és számos korszerű neurális hálózat-architektúrát tartalmaz, köztük a korábban felsorolt, vizsgált modelleket is.[6] Ezek a modellek jól konfigurálhatóak és könnyen integrálhatóak különféle alkalmazásokba, az orvosi képek feldolgozásától az általunk kutatott önvezető járművek környezetkezeléséig. Támogatja az olyan népszerű háttérhálózatokat, mint pl. a ResNet [7] vagy az EfficientNet [8] és lehetőséget biztosít előre betanított súlyok használatára is, ezzel gyorsítva a fejlesztést és javítva a teljesítményt.

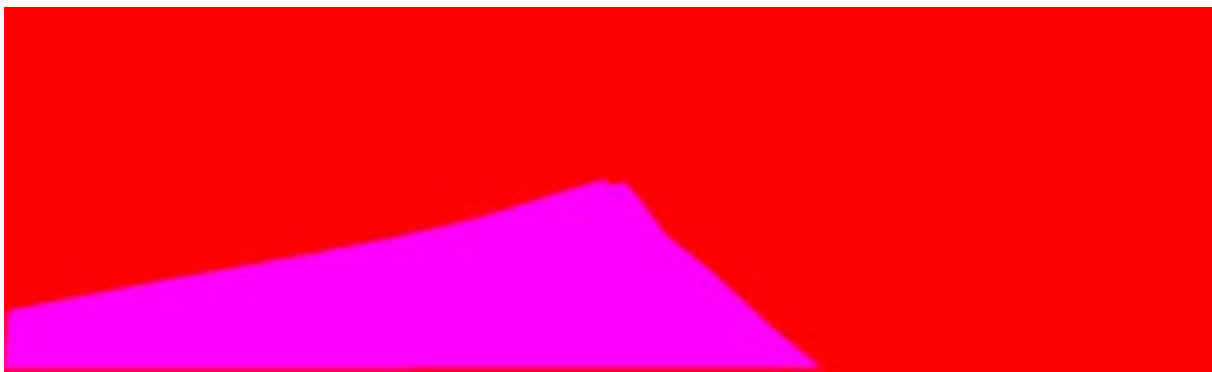
KITTI adatcsomag

A KITTI adatkészlet a Karlsruhe Institute of Technology és a Toyota Technological Institute közös projektjeként létrejött, egyik legelterjedtebb valós vezetési környezetben rögzített benchmark adatgyűjtemény és fejlett vezetéstámogató rendszerek kutatására készítették. Olyan technológiákat használtak a felvételek rögzítéséhez, illetve az adatgyűjtéshez, mint például LiDAR, GPS és IMU. A teljes adatcsomag több mint 93 000 képkockát tartalmaz, körülbelül 22 kilométernyi rögzített vezetési szekvencián, melyet a német Karlsruhe városában és annak környékén rögzítettek. A felvételi frekvencia 10 Hz. Nemcsak szegmentálásra, hanem objektumdetekcióra, mélységbecslésre és vizuális odometriára is alkalmas. Révén, hogy valós vezetési környezetből származnak az adatok, a rajta tanított modellek robusztusabbá és megbízhatóbbá válhatnak a valós környezetben történő alkalmazás során. [9]

Az általunk használt verzió egy Kaggle-n elérhető kifejezetten útszegmentáláshoz használható szelete a nagy adatkészletnek, ami 191 képet tartalmaz a hozzátartozó maszkkal. Ezek közül sem használtuk mindet a tanítás során, mert egyes maszkok kifejezetten sávszegmentáláshoz lettek készítve, ezek nekünk nem voltak relevánsak ezért csak azokat, melyeken két osztályba tudtuk sorolni az úthoz tartozó maszkot és a háttérét. (6. és 7. ábra)



6. ábra: KITTI adatsomag egy bemeneti képe



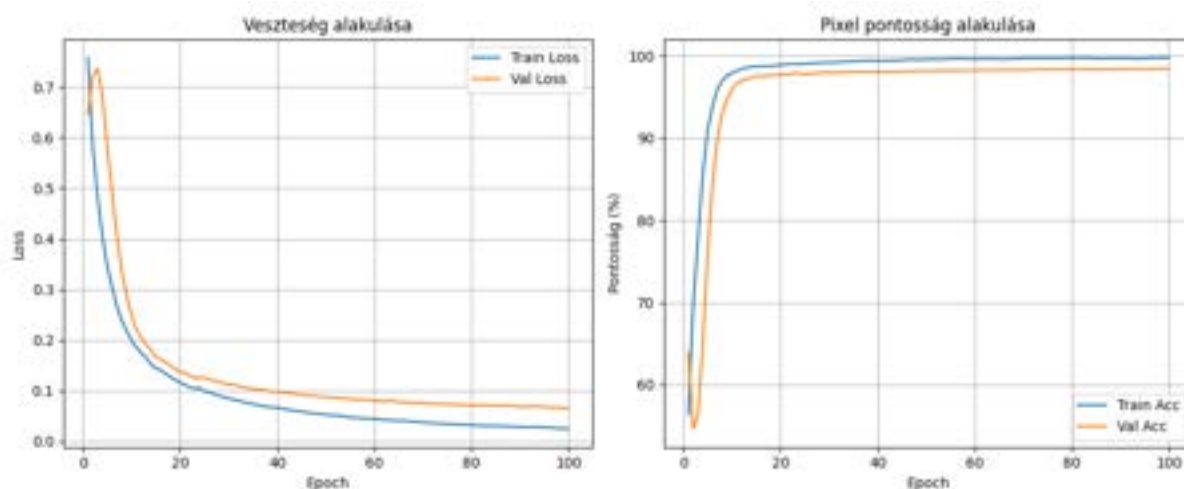
7. ábra: 6. ábrához tartozó ground truth maszk

4. Kutatás eredményei

A tanítás körülményeit tekintve majdnem minden hálózatot ugyanazokkal a hiperparaméterekkel tanítottunk egy NVIDIA GeForce 3060 GPU-n. 32-es kötegméretet használtunk minden hálózatonál a U-Net++ kivételével, ennél a magas memóriaigény miatt csak 16 volt a kötegméret. Veszteségfüggvényként a CrossEntropyLoss függvényt használtuk [10], míg optimalizálónak az Adam-et 0,0001-es tanulási rátával [11]. Minden modellt 100 epochon keresztül tanítottunk, majd a veszteség, illetve pixelpontosság változó adatait a “matplotlib” könyvtár felhasználásával vizualizáltuk. A teljes 100 epochos tanítási folyamat idejét a “time” könyvtár segítségével dokumentáltuk és ez alapján vizsgáltuk a modellek tanítási gyorsaságát. Az alábbiakban bemutatásra kerülnek a tesztelések eredményei. Az eredményeket külön szempontok szerint rendeztük sorba, majd a végén egy táblázat összegzi és értékeli.

Veszteség és pixel pontosság alakulása:

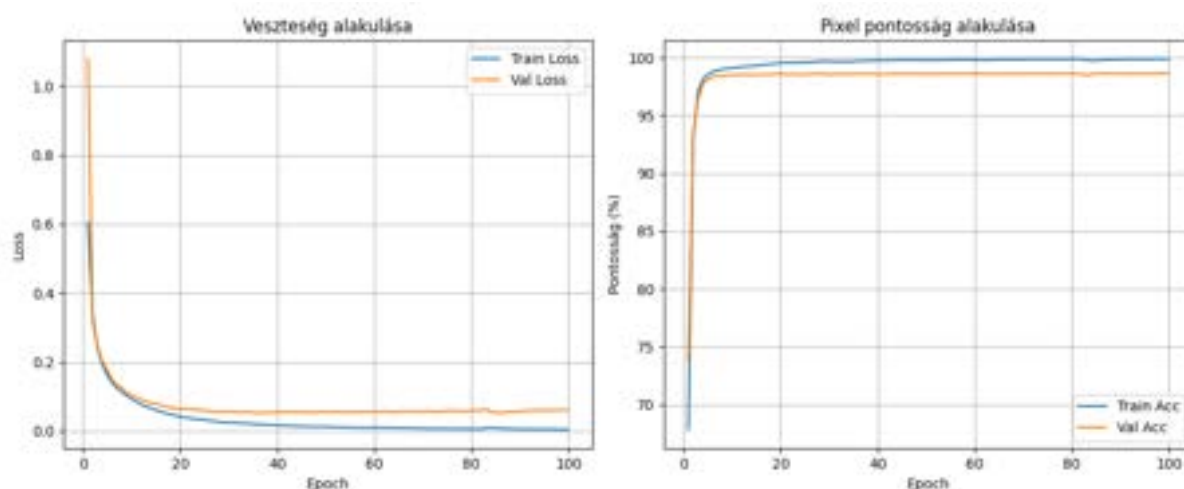
U-Net



8. ábra: A U-Net veszteség és pixel pontosság diagramjai

A legalacsonyabb mért validációs veszteség értéke 0,0571, míg a legnagyobb elért pixelpontosság 98,60% volt. A modell tanítása 593 másodpercet vett igénybe. A veszteség- és pontosság-görbék alapján jól látható, hogy a modell gyorsan konvergált, a tanulási folyamat stabil volt. (8. ábra) A U-Net egyszerűbb felépítése ellenére versenyképes teljesítményt mutatott a vizsgált adathalmazon.

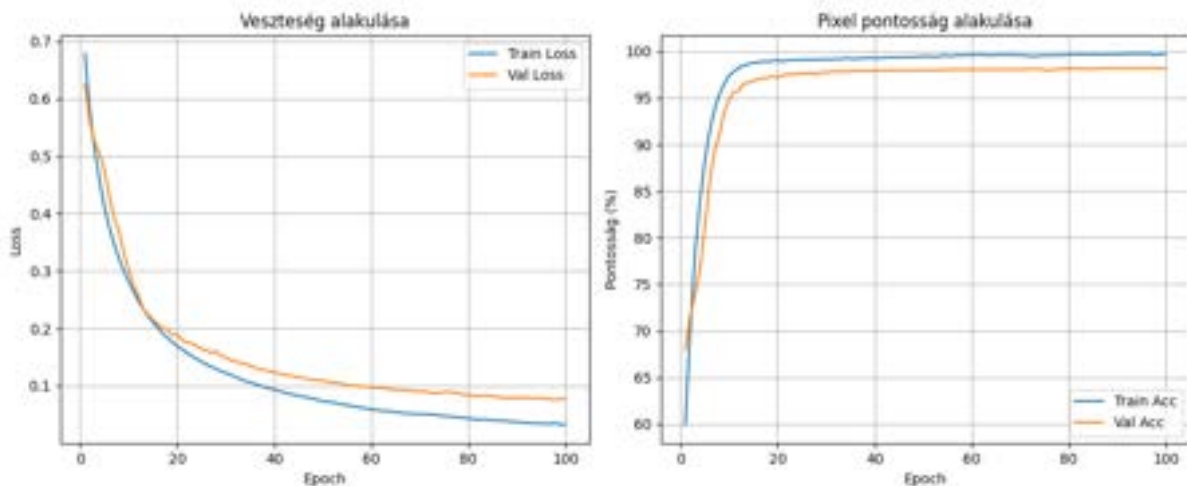
U-Net++



9. ábra: A U-Net++ veszteség és pixel pontosság diagramjai

A U-Net továbbfejlesztett változata, a U-Net++, 0,0485-ös validációs veszteséggel és 98,83% pixelpontossággal érte el a legjobb eredményt, viszont a tanulási idő is jelentősen hosszabb volt: 936 másodperc. A modell mélysége és bonyolultabb skip-kapcsolatai javították az általánosítási képességét, viszont ez magasabb számítási költséggel járt. (9. ábra)

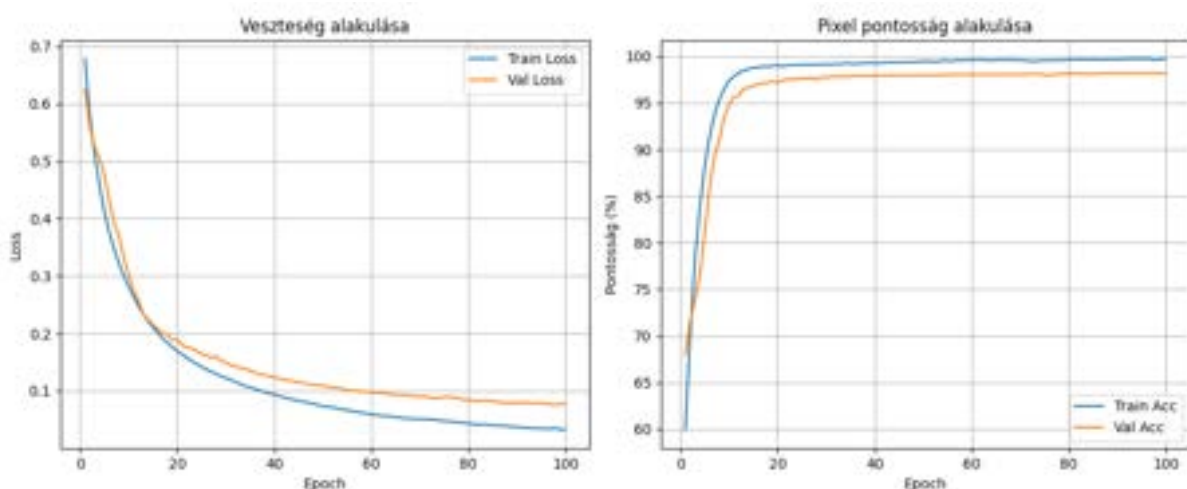
MANet



10. ábra: A MANet veszteség és pixel pontosság diagramjai

A MANet modell validációs vesztesége 0,0675 volt, a legnagyobb mért pixelpontosság pedig 98,53%. A betanítás 625 másodpercig tartott. Az attention-mechanizmusok bevezetése révén a hálózat érzékenyebbé vált a fontos jellemzőkre, de az eredmények alapján ez nem hozott jelentős előnyt más, egyszerűbb architektúrákkal szemben. (10. ábra)

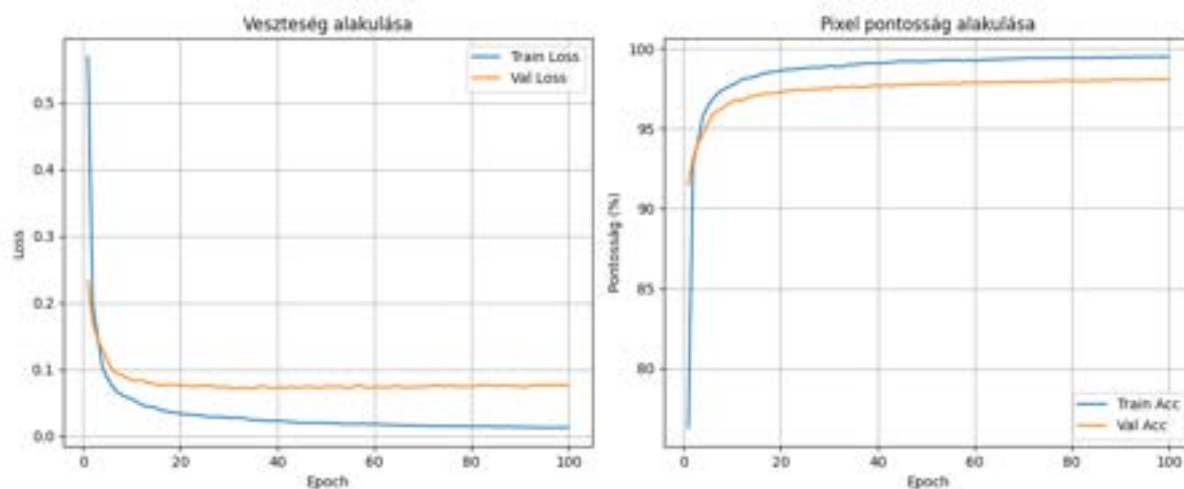
LinkNet



11. ábra: A LinkNet veszteség és pixel pontosság diagramjai

A LinkNet modell legalacsonyabb validációs vesztesége 0,0674, a legjobb pixelpontosság értéke pedig 98,50% volt. A modell tanítása 634 másodpercet vett igénybe. A modell viszonylag egyszerű architektúrája ellenére jól teljesített, és különösen gyors válaszideje miatt ideális lehet beágyazott rendszerekhez. (11. ábra)

FPN



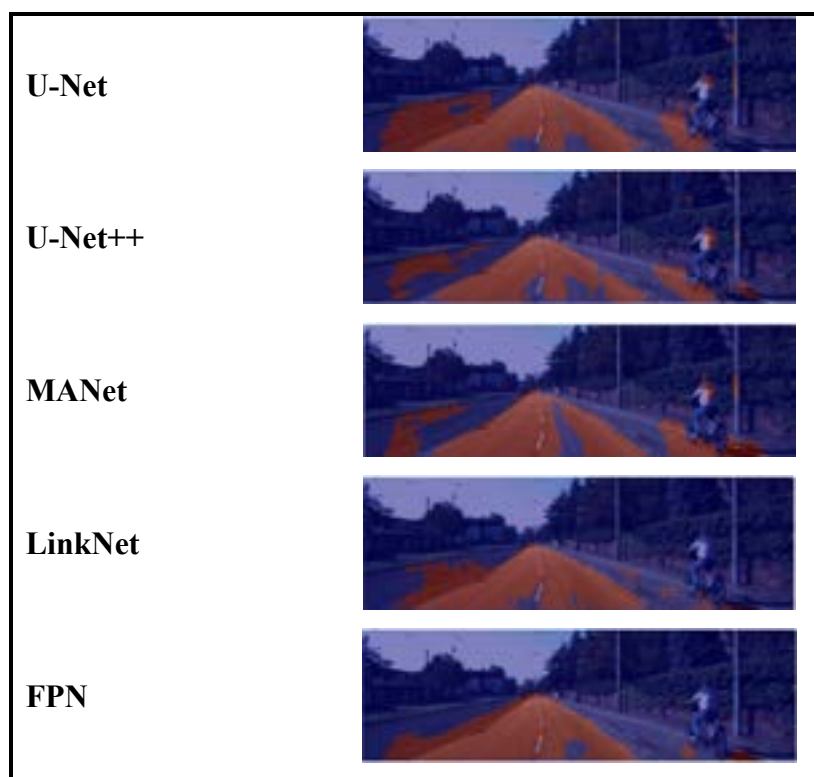
12. ábra: Az FPN veszteség és pixel pontosság diagramjai

Az FPN modell esetén a legkisebb validációs veszteség 0,0651, a legnagyobb pixelpontosság 98,52% volt. A tanulás a legrövidebb idő alatt, 526 másodperc alatt zajlott le. A teljesítmény kiegyensúlyozott (12. ábra): a veszteség és a pontosság is stabilan alakult, ráadásul ez a modell futott le a leggyorsabban, ami kiemelten fontosá válik időkritikus alkalmazásoknál, például valós idejű feldolgozás esetén.

1. Táblázat: A tanítási eredmények összegzése

Architektúra	Tanítási idő	Veszteség	Pixelpontosság	Rangsor
U-Net	593 másodperc	0,0571	98,60%	3
U-Net++	936 másodperc	0,0485	98,83%	4
MANet	625 másodperc	0,0675	98,53%	5
LinkNet	634 másodperc	0,0674	98,50%	2
FPN	526 másodperc	0,0651	98,52%	1

Az 1. táblázatban látható tehát az összesített eredmény és az általunk kiválasztott legideálisabb modellek az egyes kategóriák szerint, alatta pedig (13. ábra) az egyes hálózatokhoz tartozó egy-egy eredmény vizuális szemléltetése. A legjobb összteljesítményt az FPN érte el, amely a legrövidebb futási idő mellett versenyképes pontosságot és alacsony veszteséget mutatott, szerkezetét tekintve is kifejezetten ajánlott különböző útszegmentáló programokhoz ezért a jövőbeni kutatások során, további teszteknek és vizsgálatoknak fogjuk alá vetni az alkalmazhatósága érdekében. Bár a U-Net++ produkálta a legkisebb veszteséget és a legnagyobb pontosságot, ami nem meglepő a struktúra komplexitását figyelembe véve, a több mint 15 perces lefutási idő és a túl nagy hardver igény miatt a gyakorlati alkalmazhatósága esetünkben nem megfelelő. A LinkNet és a U-net közepes teljesítményt nyújtottak, szubjektív módon lett kiválasztva a LinkNet második helyre, csupán a tisztább vizuális eredmény miatt. A Manet magasabb veszteséggel zárt és a vizuális eredményei is zavarosak és nehezen értelmezhetőek, ezért ezt találtuk a legalkalmatlanabbnak a feladatra. A táblázat alapján jól látható, hogy a modellválasztás során nemcsak a pontosság, hanem a számítási idő és a tanítási stabilitás is kulcsfontosságú tényezők.



13. ábra: A hálózatokhoz tartozó egy-egy vizuális szemléltetés

5. Összegzés

A kutatás jövőjét tekintve az elsődleges cél a továbbfejlesztés és az optimalizálás és egy nagyobb volumenű vizsgálat lefolytatása a témakörben. A Segmentation Models könyvtár alkalmasnak bizonyult a feladat elvégzésére és egy könnyen átlátható és szerkeszthető forrást és alapot biztosít a jövőbeni kutatások során. A KITTI adatkészlet egy jóval nagyobb és részletesebb készletet biztosít, mint amit mi felhasználtunk, ezért cél, hogy más egyéb adatkészletek mellett ennek is kiaknázzuk a mélységét. Az előre elkészített benchmark adatkészletek mellett tervben van egy szimulátorban mesterségesen elkészített csomag is, ahol különböző komplex közlekedési szituációk előállításával, tudjuk finomítani és pontosítani a modellünket. Hosszútávú terv pedig egy piacképes modell felállítása, amely a kutatási eredmények alapján célfeladatra lesz optimalizálva.

6. Köszönetnyilvánítás

A publikációban szereplő kutatást a Széchenyi István Egyetem az Európai Unió támogatásával valósította meg, az RRF-2.3.1-21-2022-00004 azonosítójú, Mesterséges Intelligencia Nemzeti Laboratórium projekt keretében.

Irodalomjegyzék

- [1] Olaf Ronneberger, Philipp Fischer, Thomas Brox: *U-Net: Convolutional Networks for Biomedical Image Segmentation*, Medical Image Computing and Computer-Assisted Intervention (MICCAI) (2015), 234–241.
- [2] Zongwei Zhou et al.: *UNet++: A Nested U-Net Architecture for Medical Image Segmentation*, Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, Springer (2018), 3–11.
- [3] Jun Ma et al.: *Multi-Scale Attention Network for Semantic Segmentation*, arXiv preprint arXiv:2009.08127 (2020).
- [4] Alejandro Romera et al.: *Efficient ConvNet for Real-time Semantic Segmentation*, IEEE Intelligent Vehicles Symposium (IV) (2017), 1789–1794.

- [5] Adasri R. et al.: FPN: Feature Pyramid Networks for Object Detection, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2017), 936–944.
- [6] Ilya Berman et al.: Segmentation Models Pytorch – Python library with neural networks for image segmentation based on PyTorch, GitHub repository (2021)
- [7] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun: Deep Residual Learning for Image Recognition, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2016), 770–778.
- [8] Mingxing Tan, Quoc V. Le: EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks, Proceedings of the 36th International Conference on Machine Learning (ICML) (2019), 6105–6114.
- [9] Andreas Geiger, Philip Lenz, Raquel Urtasun: Are we ready for Autonomous Driving? The KITTI Vision Benchmark Suite, Conference on Computer Vision and Pattern Recognition (CVPR) (2012).
- [10] Diederik P. Kingma, Jimmy Ba: Adam: A Method for Stochastic Optimization, arXiv preprint arXiv:1412.6980 (2014)
- [11] Christopher M. Bishop: Pattern Recognition and Machine Learning, Springer (2006).

Szemle az önvezető járművek és a gépkocsivezetők vegyes közlekedésének lehetséges kihívásairól

Potential challenges for mixed transport involving self-driving vehicles and human drivers

Saly Gábor^a

^a Széchenyi István Egyetem, Közúti és Vasúti Járművek Tanszék – Győr

Széchenyi István Egyetem, Járműipari Kutatóközpont – Győr

saly.gabor@ga.sze.hu

Absztrakt

A jelenlegi várakozások szerint az önvezető járművek elterjedése során várhatóan lesz egy olyan időszak, amikor a hagyományos és a különböző szintű automatizáltsággal rendelkező, illetve teljesen autonóm járművek együtt fognak közlekedni az úthálózaton. Ezen járművek együttes közlekedése új kihívásokat hordoz magában, mivel a forgalmi dinamika átalakulhat, a baleseti kockázatok növekedhetnek, valamint új bizalmi problémák is fontos szerepet kaphatnak. A viselkedéstudományi vizsgálatok kimutatták, hogy az emberek gyakran másként reagálnak az önvezető járművekre, mint az emberi vezetőkre, hajlamosabbak például az agresszívebb, kockázatvállalóbb manőverekre, amikor tudják, hogy a másik közlekedő járművet egy szabálykövető algoritmus irányítja. További problémát jelent, hogy önvezető járművek szabálykövető viselkedése megváltoztathatja a forgalom dinamikáját, ami torlódásokat, váratlan lassulásokat eredményezhet. A problémák megelőzésére a fő irányvonalak az önvezető járművek közlekedésére kialakított külön sávok, illetve rugalmas szabályozások alkalmazása, amelyek gyors beavatkozást tesznek lehetővé forgalmi zavarok esetén. A közlekedésbiztonság növeléséhez nemcsak a járműveket, hanem az embereket és az infrastruktúrát is fel kell készíteni az önvezető technológiák bevezetése miatt bekövetkező változásokra.

Kulcsszavak: autonóm jármű, vegyes közlekedés, közlekedési kihívások

Abstract

According to current expectations, as self-driving vehicles become more widespread, there will likely be a period when traditional vehicles, vehicles with varying degrees of automation, and fully autonomous vehicles will all be sharing the road network. The coexistence of these vehicles poses new challenges, as traffic dynamics may change, accident risks may increase, and new trust issues may also play an important role. Behavioral studies have shown that people often react differently to self-driving vehicles than to human drivers, for example, they are more prone to aggressive, risk-taking maneuvers when they know that the other vehicle is controlled by a rule-following algorithm. Another problem is that the rule-following behavior of self-driving vehicles can change the dynamics of traffic, resulting in congestion and unexpected slowdowns. The main approaches to preventing these problems are the creation of separate lanes for self-driving vehicles and the use of flexible regulations that allow for rapid

intervention in the event of traffic disruptions. To improve road safety, not only vehicles but also people and infrastructure must be prepared for the changes that will result from the introduction of self-driving technologies.

Keywords: *autonomous vehicle, mixed traffic, transportation challenges*

1. Bevezetés

Az önvezető járművek fejlesztésének egyik legfontosabb oka a közúti biztonság növelése, az Európai Unió Vision Zero elnevezésű programjában pedig célul tűzte ki, hogy 2050-re a területén ne történhessen egyetlen közlekedési baleset miatt bekövetkező haláleset sem [1]. A fejlesztés alatt álló önvezető járművek jövőbeni elterjedésének a várakozások szerint lesz egy olyan időszaka, amikor a hagyományos, emberek által vezetett gépkocsik, illetve a számítógépek által irányított autonóm járművek közösen fognak közlekedni ugyanazon útszakaszokon [2]. Ez a helyzet újszerű kihívás lesz az infrastruktúra, a forgalomirányítás és a közlekedésben részt vevők számára egyaránt. A kutatások során meg kell vizsgálni, hogy a vegyes forgalom esetén milyen emberi viselkedés lenne megfigyelhető a közlekedési szabályokat minden esetben betartó önvezető járművekkel szemben, illetve milyen előzetes modellezési eljárások léteznek a forgalmi változásokról a vegyes forgalom során.

A szemle célja, hogy ismertesse a vegyes közlekedésben felmerülő legfontosabb kérdéseket. Ennek keretében bemutatásra kerülnek a járművezetők várható viselkedése az önvezető autókkal szemben, a vegyes forgalmat, és az ilyen közlekedési helyzetek biztonsági kockázatait vizsgáló nemzetközi kutatások, valamint a lehetséges műszaki megoldások és szabályozási javaslatok is.

2. A járművezetők viselkedése a vegyes közlekedés során

A két járműtípus különböző működése miatt előre láthatólag a vegyes közlekedés során kialakulhatnak új típusú balesetek, melyek a tisztán autonóm vagy tisztán emberi vezetők által irányított járművek forgalmában nem következtek volna be [3]. Ez jelentős kihívásokat támaszt a közlekedésbiztonsággal foglalkozó kutatók irányába.

Az emberek robotokkal szembeni viselkedését tanulmányozva kutatók megállapították, hogy az emberek három fő szempont miatt válhattak agresszívvá,

- akkor, ha a robotot irányító számítógép számítási kapacitása elmarad az emberi agyétól, és emiatt kellemetlenséget okoz az embernek,
- akkor, ha a robot teste nem rendelkezik az emberre emlékeztető részletekkel,
- akkor, ha a robot nem képes érzelmeinek megjelenítésére [4].

Ilyen összefüggést az önvezető járműveknél nem tudtak kimutatni, pontosan ellentétes reakciót válthat ki a járművezetőkből, mivel az emberi érzelmek imitálása ellentétes az önvezető jármű biztonságra törekvő, gyakran az emberi logikával ellentétes cselekedeteihez képest [4].

Már a korábbi tesztforgalomba küldött autonóm járművekkel is megtörtént, hogy a hagyományos gépkocsikat vezető emberek ismerve, hogy az autonóm jármű be fogja tartani a közlekedési szabályokat, tudatosan megnehezítették annak közlekedését, nem engedték el olyan helyzetekben, amikor elsőbbsége lett volna, vagy hirtelen gyorsításokat és lassításokat tettek az önvezető jármű megzavarására. Kínában és Dél-Koreában végzett kérdőíves felmérések alapján megállapították, hogy a járművezetők közül első sorban a fiatalabb férfiak azok, akik nagyobb valószínűséggel viselkednének agresszívebben egy önvezető járművel szemben, mint egy másik ember által irányított járművel [5].

A Szegedi Tudományegyetem kutatói egy online kérdőív segítségével vizsgálták azt, hogy az emberek személyisége hogyan befolyásolja az önvezető autókkal szembeni ellenérzéseiket és elvárásaikat. A résztvevők személyiségét öt fő szempont (nyitottság, lelkiismeretesség, barátságosság, extravertió és neuroticizmus) alapján csoportosították, valamint a technológia és az önvezető járművek iránti bizalmukat is elemezték. Az eredmények azt mutatták, hogy a személyiségnek kulcsszerepe van az autonóm járművek elfogadásában. A szorongásra hajlamos személyek általában több félelemmel tekintenek az önvezető autókra, míg a nyitottabb, barátságosabb személyiségek könnyebben bíznak meg a technológiában. A tanulmány eredményei alapján kijelenthető, hogy az önvezető járművek elfogadása nemcsak műszaki kérdés, erősen függ attól, hogy milyen típusú embereknek kell alkalmazkodniuk hozzá. Ezért a jövőben személyiséghez igazított kommunikációval és tájékoztatással célszerű növelni az emberek bizalmát az új technológia irányába [6]. Raats és kutatótársai szintén megállapították, hogy az autonóm járművek iránti bizalom alakulását nem kizárólag technikai tényezők befolyásolják. A korábbi egyéni tapasztalatok, a kulturális háttér és a terület közlekedési helyzete is fontos kérdés. A vizsgálatuk kimutatta, hogy az önvezető járművek közlekedésben történő megjelenését interdiszciplináris kutatások során szükséges elemezni [7].

Xia és kutatócsoportja, egészen más oldalról közelítette meg a problémát. Feltevésük szerint az emberek a vezetést nem egyszerűen járműmozgások sorozataként, hanem nyelvekhez hasonló módon értelmezik, ahol a vezetési műveletek, mint a lassítás vagy a sávváltás, jelentéssel bírnak, és a többi járművezető ezek alapján hozza meg a döntéseit. Ennek vizsgálatára olyan

méréseket végeztek, amelyek során a résztvevők számára forgalmi helyzeteket bemutató videófilmeket vetítettek le, miközben agyi aktivitásukat rögzítették. Az eredmények kimutatták, hogy a feltevésük igaz volt. Ezek alapján a kutatók kifejlesztettek egy rendszert, amely a járműmozgásokat kisebb egységekre bontja, majd ezekből időbeli és logikai kapcsolatok mentén összefüggő viselkedési mintákat épít fel. Ez képes a jármű pillanatnyi mozgásának követése mellett, előre jelezni a vezető szándékait is. A módszerük pontosabb előrejelzéseket adott, mint a tisztán adatalapú mélytanuló modellek. További előnye, hogy különböző forgalmi helyzeteket is képes előre jelezni. A megközelítés segítségével az önvezető járművek és a járművezetők döntéseit jobban értelmezve a vegyes közlekedés biztonságosabb módon történhet [8].

3. A hagyományos és az önvezető járművek vegyes közlekedését vizsgáló kutatások

A Chen és kutatótársai 2022-es tudományos kutatásuk során a járművezetők viselkedését vizsgálták különböző forgalmi helyzetekben a CACC, vagyis a kooperatív alkalmazkodó sebességtartó automatikával kapcsolatban. Kutatásuk során kérdőíveket töltettek ki járművezetőkkel és a válaszaik alapján a CACC-re pozitívan, negatívan és semlegesén reagáló csoportokra osztották a vizsgálatban részt vevő személyeket. A válaszok alapján szimuláció keretében vizsgálták a forgalmat, és megállapították, hogy a CACC vegyes forgalomban akkor lehet eredményes, ha a járművek legalább 85%-a használja a technológiát, 80% felhasználói arány alatt pedig a CACC rosszabb forgalomáteresztő képességet adhat, mint a tisztán emberek vezette járművekből álló forgalom [9].

Zhao és társai azt vizsgálták, hogy van-e összefüggés az önvezető járművekbe vetett bizalom és a mögötte haladó járművezető viselkedése között. Egy tesztpályán 10 járművezető személy és két jármű segítségével valós körülmények között végeztek méréseket. Az egyik jármű önvezető funkciókkal volt ellátva, de képes volt ember általi irányítással is működni, a másik jármű hagyományos gépkocsi volt. A mérések során az autonóm funkciókkal rendelkező jármű haladt elől, és jelezte, hogy önvezető vagy hagyományos üzemmódban közlekedik. Ezt a járművet követte a kísérletbe bevont járművezetők által irányított hagyományos gépkocsi. A tesztek előtt a részt vevő járművezetőkől kérdőívek segítségével megállapították, hogy mennyire bíznak meg az önvezető járművek működésében. A mérések eredményei azt

mutatták, hogy erős a korreláció, az önvezető járművet követő sofőr előfeltételezései és viselkedése között. Az új típusú járműben nem megbízó személyek nagyobb követési távolságot tartanak az önvezető jármű mögött haladva, mint egy ember vezette jármű esetén, ezzel szemben az önvezető technológiában megbízó személyek kisebb követési távolságot hagynak, mint egy ember vezette jármű esetén [10]. Mahdinia és munkatársai kísérletükben azt vizsgálták, hogyan hat az önvezető járművek jelenléte a mögöttük haladó gépkocsivezetők viselkedésére. A Texas A&M egyetem tesztpályáján kilenc sofőr követte a kísérleti járművet kétféle helyzetben, amikor az emberi vezető sebességprofilját utánozta, illetve amikor saját, automatizált sebességprofilja alapján közlekedett. Az eredmények szerint az önvezető funkcióval közlekedő gépkocsi követése a sebesség- és gyorsulás-ingadozásokat 15–23%-kal csökkentette, és a ráfutásos ütközések kockázata csökkent. A teszt résztvevői közelebb haladtak az önvezető járműhöz, de ez a kiszámíthatóbb sebességek miatt nem járt fokozott baleseti veszéllyel. Az önvezető jármű mögötti üzemanyag-fogyasztás átlagosan 10,6%-kal, a károsanyag-kibocsátás közelítőleg 4%-kal csökkent [11].

A Delfti Egyetem kutatása azt vizsgálta, hogyan alkalmazkodnak a gépkocsivezetők az önvezető járművekkel való találkozás során a követési távolságok és az előzési manőverek esetében. Egy hollandiai terepkísérletben 18 résztvevő saját személygépkocsijával közlekedett együtt kísérleti járművel, amelyet hagyományos vagy önvezető üzemmódban mutattak be számukra. A fő eredmények szerint a járművezetők hajlamosabbak voltak kisebb távolságok esetén is besorolni, amikor a másik gépkocsit önvezetőként azonosították. Hasonló módon kisebb követési távolságot tartottak az önvezető jármű mögött, mint egy ember által vezetett gépkocsi mögött. Az önvezetés biztonságosságáról szóló információk hatására a résztvevők még kisebb távolságokat hagytak a manőverekhez, és jobban bíztak az önvezető járműben. Ezzel szemben a negatív információ nem okozott jelentős viselkedésváltozást. Ez arra utal, hogy a sofőrök opportunista módon kihasználják az önvezető jármű biztonságos és kiszámítható közlekedési viselkedését [12].

Li és kutatótársai 2023-as tudományos munkájukban a helyszíni kísérletek és vezetési viselkedési kutatások tapasztalatai alapján szimulálták az emberek vezette és az önvezető járművek vegyes forgalmát. A gépkocsivezetők önvezető technológiához való hozzáállását a pszichológia viselkedési törvényeivel modellezték. Az önvezető jármű modelljének viselkedését annak megfelelően programozták, hogy egy másik önvezető jármű vagy sztochasztikus viselkedéssel rendelkező ember vezette jármű mögött halad. A kutatók a

szimulációkat az eltérő irányítású járművek különböző számbeli arányai mellett végezték el, hogy képet kapjanak arról, hogy az önvezető járművek elterjedésének különböző szakaszai hogyan befolyásolják a közlekedést [13].

Kis számú önvezető jármű esetén sok olyan forgalmi helyzet adódik, ahol különböző irányítású járművek találkoznak, ilyenkor a gépkocsivezetők és az önvezető járművek is nagyobb követési távolságokat tartanak, így a forgalom áteresztő képessége eleinte nem javul. Azonban minél több önvezető jármű vesz részt a közlekedésben, annál több esetben tudnak kisebb követési távolsággal közlekedni a közöttük megvalósuló információcsere segítségével. A torlódások száma is jelentősen csökkenthető az önvezető járművek elterjedésével. A csak önvezető járművekből álló forgalomban a torlódások száma akár a harmadára csökkenhet a kizárólag hagyományos gépkocsik közlekedéséhez képest. Az önvezető járművek alkotta forgalomban a forgalom áteresztő képessége és sebessége is nő, a forgalom egyenletes, a torlódások száma pedig csökken, mivel az önvezető járművek rövid idő alatt képesek egy vagy több szakaszba rendeződni, és állandósult vezetési állapotot fenntartani [13].

A sávváltási gyakoriság elsősorban a forgalomsűrűségtől függ, amennyiben a forgalmi sűrűség kicsi, több jármű tartja a sávját, hiszen azokban közelítőleg azonosak a szabad sebességek. A hagyományos járművek sávváltási gyakorisága az önvezető járművek növekvő arányával egyre csökken, mivel az egymást követő önvezető járművek kis követési távolságot tartanak maguk között, és az emberek nem érzik biztonságosnak a távolságokat a sávváltás megkezdéséhez. Az önvezető járművek sávváltásainak gyakorisága elterjedésük növekedésével csökken, ritkább forgalomban viszont előszeretettel váltanak sávot a sebességük fenntartásának érdekében [13].

4. A hagyományos és az önvezető járművek vegyes közlekedésének kockázatai

A Driving Safety Risk Analysis and Assessment in a Mixed Traffic Flow with Human-Driven and Autonomous Vehicles című tanulmány eredményei szerint a hagyományos és az önvezető járművek közös közlekedésének átmeneti időszaka még feltáratlan biztonsági kockázatokat rejt magában. Az önvezető járművek elterjedésének kezdeti szakaszában a két járműirányítás eltérő működési logikája és reakciói kockázatot jelentenek a közlekedésbiztonságra nézve, főként kritikus forgalmi helyzetekben, mint a hirtelen fékezés, a sávváltás vagy az előzési és elsőbbségi helyzetek. Emiatt ezen helyzetekben megnövekedhet az ütközések valószínűsége. A kutatás kimutatta, hogy a baleseti kockázat rövid távon nagyobb lehet, mint kizárólag

hagyományos vagy kizárólag önvezető forgalom esetén, ezért a vegyes közlekedés időszakát a technológiai átmenet mellett olyan kritikus időszakként is kell kezelni, amely fokozott közlekedésbiztonsági, forgalomszervezési és szabályozói figyelmet igényel. A szerzők javaslatai szerint az átmeneti kockázatok mérsékléséhez szükséges a közlekedési szabályozások felülvizsgálata, a gépkocsivezetők képzése a megváltozott forgalmi helyzetekkel kapcsolatban, valamint a forgalomirányítás intelligens eszközeinek bevezetése [2].

5. A hagyományos és az önvezető járművek vegyes közlekedését támogató megoldások

A felmerülő problémák megelőzésére az egyik lehetséges megoldás az önvezető járművek számára kijelölt sávok bevezetése lehet az utakon. Yu és társai kutatásukban a belső sávot jelölték ki erre a feladatra. A sáv kialakításának egyik célja a forgalom áteresztő-képességnek növelése, a másik a biztonság növelése azáltal, hogy kevesebb interakció történik a különböző irányítású járművek között. Az önvezető járművek 20% alatti elterjedtsége előtt azonban nem célszerű bevezetni ezt az intézkedést, mivel a sáv kihasználtsága kicsi lenne, és a hagyományos járművek közlekedését is akadályozná az, hogy eggyel kevesebb sáv áll rendelkezésükre. Nagyobb elterjedtség esetén viszont a sáv bevezetése növelhetné a hagyományos és önvezető járművek átlagos sebességét és a közlekedés biztonságát is. A biztonság tekintetében a vegyes forgalomban az önvezető sáv kijelölése jelentős előnyökkel rendelkezik, különösen az autópályák összefutó és letérő szakaszain. A megoldásnak akkor a legnagyobb a pozitív hatása, ha az önvezető járművek aránya 40% alatt van [14].

Mohajerpoor és Ramezani tanulmánya azt vizsgálta, hogyan változnak a közlekedési áramlások vegyes közlekedés esetén. A szerzők modelleket dolgoztak ki a követési távolságok várható értékére és szórására, különböző járműelrendezések esetén. A modellekben az önvezető járművek arányát valószínűségi változóként kezelték, mivel az folyamatosan változik a forgalomban. A szerzők a késéseket két sávós, jelzőlámpás utaknál is vizsgálták, különféle sáv kiosztási stratégiák esetén. Az elemzések szerint sáv kiosztási stratégiát az önvezető gépkocsik arányának kell meghatároznia. A vegyes forgalomban a sávkezelés rugalmas alkalmazása kulcsfontosságú a késések csökkentésének és a kapacitás növelésének érdekében [15].

Khondaker és Kattan kutatók a különleges forgalmi helyzetekben, mint amilyenek például a balesetek, a forgalom egyenletes haladásának megőrzése érdekében változtatható sebességkorlátozást vezetnének be az érintett útszakaszokra, amelyet betartva kisebb valószínűséggel alakulnának ki torlódások. A vegyes forgalom esetén egy ilyen irányítási rendszernek a biztonság és a fenntarthatóság szempontjait is figyelembe kell vennie a közlekedők megfelelő biztonságának megvalósításához [16].

Van Geelen és Redant kutatók a Belga Közúti Kutatási Központ (BRRC) számára készített tanulmányukban vizsgálták az önvezető járművek hatását az épített infrastruktúrára, annak érdekében, hogy az ország közlekedési hatósága el tudja kezdeni a felkészülést az önvezető járművek megjelenése miatt várható, infrastruktúrát érintő változásokra. Megállapították, hogy a fizikai infrastruktúra mellett szükséges létrehozni ennek digitális iker változatát is, ezzel segítve a navigációt, ennek megvalósításához ki kell építeni a járművek és infrastruktúra közötti kommunikációhoz szükséges műszaki hátteret. Fontosnak tartják a váltakozó irányú sávok és a dinamikus sáv kiosztás megvalósítását, ennek érdekében a sávokat elválasztó fizikai akadályok jelenlegi kialakításait felül kell majd vizsgálni [17].

1.táblázat: A vegyes közlekedés hatásait vizsgáló tanulmányok legfontosabb megállapításai

Szerzők (év)	Módszertan	Vizsgálat területe	Fő megállapítások	Következtetések és javaslatok
Chen et al. (2022)	Szimuláció (CACC)	Sebességtartó automatika hatása	CACC csak 85% feletti elterjedtség javítja a forgalmat.	Fokozatos bevezetés javasolt.
Zhao et al. (2020)	Tesztpályás kísérlet	Követési távolság és bizalom	A bizalom befolyásolja a követési távolságot.	A bizalom növelése kulcsfontosságú.
Mahdinia et al. (2021)	Valós forgalmi teszt	Emberi vezetők viselkedése	Csökkent sebességingadozás és ütközéskockázat.	Egyenletesebb, biztonságosabb forgalom.
Soni et al. (2022)	Terepkísérlet	Előzési és követési manőverek	A vezetők bátrabbak önvezető járművek mellett.	Kiszámíthatóság opportunistá viselkedést okoz.

Li et al. (2023)	Szimuláció, modellezés	Forgalom áteresztő képessége	Növekvő önvezető arány javítja az áteresztést.	Átmenetben nagyobb baleseti kockázat.
Yu et al. (2019)	Forgalmi modellezés	Autonóm sávok hatása	Sávkijelölés 20% és 40% közötti önvezető jármű arányig előnyös.	Dinamikus sávkezelés javasolt.
Mohajerpoor & Ramezani (2019)	Matematikai modellezés	Sávkiosztási stratégiák	Optimális sávkezelés az önvezető aránytól függ.	Rugalmas sávkiosztás növeli kapacitást.
Van Geelen & Redant (2023)	Infrastrukturális elemzés	Infrastruktúra hatásai	Digitális iker és V2I-kommunikáció szükséges.	Infrastruktúra fejlesztése elengedhetetlen.

A vizsgált kutatások eredményei arra utalnak, hogy az önvezető járművek és a gépkocsivezetők vegyes közlekedésének biztonságát és hatékonyságát három fő tényező határozza meg, az önvezető járművek aránya, a járművezetők viselkedése, valamint az alkalmazott forgalomszervezési megoldások. A szimulációs vizsgálatok (Chen, Li, Yu) szerint az önvezető technológia előnyei csak magas elterjedtségi arány mellett érvényesülnek, míg alacsony arány esetén a forgalom instabilabbá válhat az eltérő irányítással rendelkező járművek miatt. A valós körülmények között végzett tesztek eredményei (Zhao, Mahdinia, Soni) azt mutatják, hogy a bizalom szintje alapvetően befolyásolja a vezetői döntéseket és a követési távolságokat. A forgalmi modellek és az infrastruktúra elemzése (Mohajerpoor, Van Geelen) kimutatták, hogy a rugalmas sávkezelés és az önvezető járművek számára kijelölt sávok csökkentik a torlódásokat és növelik az útszakaszok forgalomáteresztő képességét, viszont ehhez a közlekedési infrastruktúra fejlesztése szükséges. Mindezek alapján az önvezető technológiák bevezetése a közlekedésbe összetett, többtényezős folyamat, amely technológiai, pszichológiai és infrastrukturális elemek összehangolt fejlesztését igényli a biztonságos átmenet megvalósításához.

6. Összegzés

Az önvezető járművek várható elterjedésével egyre fontosabbá válik a vegyes forgalomban rejlő kihívások azonosítása és a veszélyhelyzetek megelőzése. A szemlézett kutatások alapján megállapítható, hogy a járművezetők viselkedése jelentősen eltérhet attól függően, hogy ember vezette vagy önvezető járművel találkoznak, ami új típusú baleseti kockázatokat idézhet elő. Szimulációk és tesztpályás vizsgálatok bizonyítják, hogy az önvezető technológia megfelelő arányú elterjedése hosszú távon javíthatja a forgalom áteresztőképességét és biztonságát, ugyanakkor az átmeneti időszakban fokozott odafigyelést igényel. A közlekedésbiztonság és a forgalom áteresztő képességének javítása érdekében szükség van új megoldásokra, például önvezető járműveknek fenntartott sávok kialakítására, rugalmas sebességszabályozásra vagy az infrastruktúra digitális ikerpárjának a megvalósítására. A tudományterületeken átívelő kutatások bizonyítják, hogy az önvezető járművek bevezetése nem csupán technológiai, hanem társadalmi és szabályozási kérdés is, amely komplex, interdiszciplináris megközelítést igényel.

7. Köszönetnyilvánítás

A publikációban szereplő kutatást a Széchenyi István Egyetem az Európai Unió támogatásával valósította meg, az RRF-2.3.1-21-2022-00004 azonosítójú, Mesterséges Intelligencia Nemzeti Laboratórium projekt keretében.

Irodalomjegyzék

- [1] European Climate, Infrastructure and Environment Executive Agency. (2022). EU Road Safety: Towards “Vision Zero”. European Union. doi: 10.2840/701809
- [2] Z. Cheng, J. Zhu, Z. Feng, M. Yang, W. Zhang, and J. Chen. (2025). Driving Safety Risk Analysis and Assessment in a Mixed Driving Environment of Connected and Non-Connected Vehicles: A Systematic Survey. IEEE Transactions on Intelligent Transportation Systems, 26(5), 5747–5781. doi: 10.1109/TITS.2025.3526820
- [3] I. Barabás, A. Todorut, N. Cordos, and A. Molea. (2017). Current Challenges in Autonomous Driving. IOP Conference Series: Materials Science and Engineering, 252(1), 012096. doi: 10.1088/1757-899X/252/1/012096

- [4] Y. Lee, A. Momen, and J. LaFreniere. (2021). Attributions of Social Interactions: Driving among Self-Driving vs. Conventional Vehicles. *Technology in Society*, 66, 101631. doi: 10.1016/j.techsoc.2021.101631
- [5] P. Liu, Y. Du, L. Wang, and D. Y. Ju. (2020). Ready to Bully Automated Vehicles on Public Roads? *Accident Analysis & Prevention*, 137, 105457. doi: 10.1016/j.aap.2020.105457
- [6] M. Volosin, M. Kálnay, Á. Bánffi, N. Nyeső, G. V. Molnár, Z. Palatinus, and T. Martos. (2024). The Leading Role of Personality in Concerns about Autonomous Vehicles. *PLoS One*, 19(6), e0301895. doi: 10.1371/journal.pone.0301895
- [7] K. Raats, V. Fors, and S. Pink. (2020). Trusting Autonomous Vehicles: An Interdisciplinary Approach. *Transportation Research Interdisciplinary Perspectives*, 7, 100201. doi: 10.1016/j.trip.2020.100201
- [8] Y. Xia, M. Geng, Y. Chen, S. Sun, C. Liao, Z. Zhu, Z. Li, W. Y. Ochieng, P. Angeloudis, M. Elhajj, L. Zhang, Z. Zeng, B. Zhang, Z. Gao, and X. M. Chen. (2023). Understanding Common Human Driving Semantics for Autonomous Vehicles. *Patterns (N Y)*, 4(7), 100730. doi: 10.1016/j.patter.2023.100730
- [9] Y. Chen, D. Kong, L. Sun, T. Zhang, and Y. Song. (2022). Fundamental Diagram and Stability Analysis for Heterogeneous Traffic Flow Considering Human-Driven Vehicle Driver's Acceptance of Cooperative Adaptive Cruise Control Vehicles. *Physica A: Statistical Mechanics and Its Applications*, 589, 126647. doi: 10.1016/j.physa.2021.126647
- [10] X. Zhao, Z. Wang, Z. Xu, Y. Wang, X. Li, and X. Qu. (2020). Field Experiments on Longitudinal Characteristics of Human Driver Behavior Following an Autonomous Vehicle. *Transportation Research Part C: Emerging Technologies*, 114, 205–224. doi: 10.1016/j.trc.2020.02.018
- [11] I. Mahdinia, A. Mohammadnazar, R. Arvin, and A. J. Khattak. (2021). Integration of Automated Vehicles in Mixed Traffic: Evaluating Changes in Performance of Following Human-Driven Vehicles. *Accident Analysis & Prevention*, 152, 106006. doi: 10.1016/j.aap.2021.106006

- [12] S. Soni, N. Reddy, A. Tsapi, B. van Arem, and H. Farah. (2022). Behavioral Adaptations of Human Drivers Interacting with Automated Vehicles. *Transportation Research Part F: Traffic Psychology and Behaviour*, 86, 48–64. doi: 10.1016/j.trf.2022.02.002
- [13] X. Li, Y. Xiao, X. Zhao, X. Ma, and X. Wang. (2023). Modeling Mixed Traffic Flows of Human-Driving Vehicles and Connected and Autonomous Vehicles Considering Human Drivers' Cognitive Characteristics and Driving Behavior Interaction. *Physica A: Statistical Mechanics and Its Applications*, 609, 128368. doi: 10.1016/j.physa.2022.128368
- [14] H. Yu, S. Tak, M. Park, and H. Yeo. (2019). Impact of Autonomous-Vehicle-Only Lanes in Mixed Traffic Conditions. *Transportation Research Record: Journal of the Transportation Research Board*, 2673(6), 036119811984747. doi: 10.1177/0361198119847475
- [15] R. Mohajerpoor and M. Ramezani. (2019). Mixed Flow of Autonomous and Human-Driven Vehicles: Analytical Headway Modeling and Optimal Lane Management. *Transportation Research Part C: Emerging Technologies*, 109, 194–210. doi: 10.1016/j.trc.2019.10.009
- [16] B. Khondaker and L. Kattan. (2015). Variable Speed Limit: A Microscopic Analysis in a Connected Vehicle Environment. *Transportation Research Part C: Emerging Technologies*, 58(Part A), 146–158. doi: 10.1016/j.trc.2015.07.014
- [17] H. van Geelen and K. Redant. (2023). Connected & Autonomous Vehicles and Road Infrastructure – State of Play and Outlook. *Transportation Research Procedia*, 72, 1311–1317. doi: 10.1016/j.trpro.2023.11.592

Paraméterezhető Virtuális Környezet Vizuális Odometria Algoritmusok Fejlesztéséhez

A Parametrizable Virtual Environment for the Development of Visual Odometry Algorithms

Török Bence^a, Hollósi János^b

^aSzéchenyi István Egyetem, Járműipari Kutatóközpont

torok.bence@hallgato.sze.hu

^bSzéchenyi István Egyetem, Járműipari Kutatóközpont

hollosi.janos@sze.hu

Absztrakt

A kutatás során aiSim szimulációs környezet segítségével egy olyan integrált rendszer fejlesztése a cél, amely a szimuláció futása közben, valós időben képes szenzor adatok publikálására a ROS 2 keretrendszeren keresztül. Az aiSim egy szenzorszimulációs platform, amely lehetővé teszi, hogy a szimulációhoz egyszerre több, akár különböző szenzort vagy kamerát helyezzünk az autóra tetszőleges pozíciókban. Ezen kívül lehetővé teszi különböző környezeti és forgalmi viszonyok valósághű modellezését. Ennek köszönhetően hatékonyabban lehet vizuális odometria algoritmusokat tesztelni mivel könnyedén lehet a szimuláció paramétereit változtatni. A rendszer alapot biztosít ezen algoritmusok alapos kipróbálásához. A vizuális odometria lényege, hogy a kamerák által rögzített képek alapján határozza meg a mozgást és a környezetben bekövetkező eltolódásokat, mely kulcsfontosságú az autonóm járművek navigációjában. Ezzel a megközelítéssel nem szükséges kizárólag valódi autókra vagy az időjárásra támaszkodni, így a tesztelési folyamat sokkal rugalmasabbá válik. A rendszer célja, hogy megkönnyítse a vizuális odometria algoritmusok validálását.

Kulcsszavak: Szimuláció-alapú tesztelés, vizuális odometria, ROS 2 integráció

Abstract

The goal of this research is to develop an integrated system using the aiSim simulation environment that is capable of publishing sensor data in real time via the ROS 2 framework during the simulation. aiSim is a sensor simulation platform that allows multiple, even different, sensors or cameras to be mounted on a vehicle in arbitrary positions. It also enables realistic modeling of various environmental and traffic conditions. This allows for more efficient testing of visual odometry algorithms, as simulation parameters can be easily adjusted. The system provides a solid basis for thorough testing of these algorithms. Visual odometry determines motion and environmental shifts based on camera images, which is crucial for autonomous vehicle navigation. This approach removes reliance on real vehicles or weather conditions, making the testing process much more flexible. The system aims to facilitate the validation of visual odometry algorithms.

Keywords: simulation-based testing, visual odometry, ROS 2 integration

1. Bevezetés

A kutatás célkitűzése egy olyan rendszer megteremtése, amely lehetővé teszi azt, hogy autonóm járművek fejlesztése során valós járművekkel végzett tesztelést megelőzően egy, a valósághoz nagyfokú hasonlóságot mutató, virtuális tesztkörnyezetben is elvégezhesük a szükséges vizsgálatokat. Ez különösen fontos olyan esetekben, amikor szenzoradatokra vagy kameraképekre lenne szükség egy adott fejlesztési vagy kutatási feladathoz, de a valós körülmények között történő adatgyűjtés technikailag bonyolult vagy költséges lenne.

Vizuális odometria algoritmusok validálásánál fontos lehet specifikus zavaró tényezőkre adott rendszerreakciók vizsgálata [1]. Például fényviszony-változások, időjárási körülmények vagy forgalmi szituációk. Egy ilyen szimulációs környezet lehetőséget teremt arra, hogy reprodukálhatóan lehessen generálni releváns teszteseteket, és ehhez használható szintetikus adatokat állítsanak elő. Ezzel a megközelítéssel olyan szituációk is vizsgálhatók, amelyek a valóságban csak ritkán, vagy nehezen reprodukálható módon fordulnak elő. Például havas utak, sűrű forgalom, vagy éppen teljesen üres útszakaszok esetén.

2. Kapcsolódó munkák

A vizuális odometria (VO) területén számos kutatás foglalkozik az autonóm járművek lokalizációs és navigációs képességeinek fejlesztésével. Scaramuzza és Fraundorfer [2] átfogó áttekintést nyújtanak a vizuális odometria alapelveiről, történeti fejlődéséről és alkalmazási területeiről, különös tekintettel a robotikában és az autonóm rendszerekben betöltött szerepére. Az általuk bemutatott módszerek alapját képezik számos későbbi kutatásnak.

Howard és munkatársai [3] valós idejű sztereó vizuális odometria algoritmust fejlesztettek ki autonóm földi járművek számára, amely képes valós idejű kameramozgás becslésre egymást követő sztereó képpárokból. Ez az algoritmus jelentős előrelépést jelentett a terepi robotika területén, különösen a NASA Mars Exploration Rover programjában való alkalmazása révén.

Li és kollégái [4] a szimuláció alapú validáció fontosságát hangsúlyozzák az autonóm vezetési rendszerek fejlesztésében. Munkájukban egy formálisan definiált tesztkörnyezetet integrálnak egy ipari szimulátorral, lehetővé téve a rendszer viselkedésének szisztematikus vizsgálatát különböző közlekedési szituációkban.

Xiang és munkatársai [5] egy többkamerás vizuális odometria rendszert mutatnak be, amely virtuális LiDAR adatokat generál a környezet szegmentálása révén. Ez a megközelítés javítja a pozícióbecslés pontosságát és robusztusságát, különösen összetett környezeti feltételek mellett.

3. Rendszer bemutatása

Az aiSim egy moduláris felépítésű, nagy megbízhatóságú virtuális szimulációs környezet, amely a világon elsőként nyerte el az ISO 26262 ASIL-D minősítést az autóiipari szimulátorok között. A rendszert az aiMotive fejlesztette ki, kifejezetten járműipari funkcionális biztonsági követelményeknek megfelelő ellenőrzési és validációs célokra.

A szimulációs platform alapjául az aiSim AIR szolgál, amely mesterséges intelligencia alapú renderelési technológiát alkalmaz, lehetővé téve nagy pontosságú, több érzékelőt is lefedő szimulációk végrehajtását. Az aiSim elérhető önálló alkalmazásként, valamint moduláris SDK formájában, így rugalmasan integrálható különböző fejlesztési környezetekbe és folyamatokba. Emellett lehetővé teszi egyedi aiSim-kliensalkalmazások és szenzorok (pluginek) fejlesztését és implementálását, amelyek közvetlenül integrálhatók az aiSim szimulációs környezetbe [6,7]

Az aiSim kliens-szerver architektúrán alapul. A szerveroldali rendszer plugin-alapú struktúrával rendelkezik, amely lehetővé teszi egyedi szenzorok, járműdinamikai modellek vagy forgatókönyvek integrálását a szimulációs környezetbe. A kliensoldal fő funkciója a szimuláció konfigurálása és vezérlése. Ez magában foglalja a környezeti paraméterek és scenáriók beállítását, a szimuláció elindítását és felügyeletét, az érzékelők adatainak fogadását és feldolgozását, valamint a járművek irányításához szükséges parancsok kiadását.

A szimuláció futtatása a szerveroldalon történik, amely felelős a szükséges modulok betöltéséért, a szimulált világ időbeli léptetéséért, a szenzorfelvételek és vizualizációk ütemezéséért, továbbá a járművek fizikai viselkedésének számításáért. Emellett a szerver végzi a szimulációs forgatókönyvek feltételeinek kiértékelését is, biztosítva ezzel a realisztikus és megbízható működést [7].

A szimuláció során keletkező adatok a ROS 2 (Robot Operating System 2) keretrendszeren keresztül kerülnek publikálásra. A ROS 2 keretrendszer egyik alapvető működési elve a publikálás–feliratkozás (publish–subscribe) kommunikációs modell. Ebben az architektúrában az egyes szoftverkomponensek, úgynevezett node-ok, információt közölhetnek (publikálnak)

egy adott témakörben (topic), míg más node-ok képesek ezekre az adatokra feliratkozni és feldolgozni azokat. Az adatok publikálása azt jelenti, hogy egy szenzor vagy feldolgozó modul egy előre meghatározott topicra rendszeresen továbbít adatcsomagokat, például kameraképeket, pozícióinformációt vagy gyorsulási értékeket. A feliratkozó node ezekhez az adatokhoz valós időben hozzáfér anélkül, hogy közvetlenül ismernie kellene az adat forrását. Az aiSim és a ROS 2 közötti integrációhoz, az aiSim SDK csomagban elérhető, kifejezetten erre a célra fejlesztett toolchain-t használtam fel. Ez kiváló alapot biztosított a kutatáshoz. Ez a példa toolchain tartalmaz kifejezetten ROS 2 publikálásra programozott szenzorokat és egy "ros_bridge" nevű plugint, amely a node létrehozásáért és ezáltal az adatok publikálásáért felelős. Emellett lehetőség van olyan kliensalkalmazások létrehozására is, amelyek a ROS érzékelőkkel kompatibilisek és alkalmasak a szimulációs konfigurációk kezelésére.

A ROS 2 egy nyílt forráskódú keretrendszer, amely a ROS első generációjára épül, és fejlesztői eszközök, middleware-ek, valamint komponens-alapú architektúra révén támogatja robotikai és autonóm rendszerek létrehozását. Előnyei közé tartozik a valós idejű kommunikáció, a DDS-alapú (Data Distribution Service) adatkezelés, valamint a moduláris, többszálú és több nyelven is használható API-k. A ROS 2-t széles körben alkalmazzák autonóm járművek, ipari robotok, drónok és kutatási célú rendszerek fejlesztésében, mivel node-alapú struktúrája és gazdag szimulációs támogatottsága révén egy rugalmas és nagy teljesítményű platformot kínál. [8]

4. Szimulációs adatok publikálása

A kutatás jelenlegi szakaszában megvalósításra került egy kiterjedt scenárió, amelyben a jármű egy előre definiált útvonal mentén halad végig egy üres Budapest-modellben. A szimuláció futása közben valós időben történik a szenzorok által mért adatok továbbítása ROS 2 keretrendszeren keresztül. A példaként felhasznált toolchain modifikálása szükségessé vált ahhoz, hogy egyszerre két kameraszenzort is támogasson. Az eredeti megvalósításban amikor több kamera volt csatlakoztatva az autóhoz, csak az elsőként inicializált kamera képe jelent meg folyamatosan, és időnként villanásszerűen váltott a második kamera képkockájára. A szükséges módosításokat követően azonban a toolchain már megfelelően kezelte mindkét kamerát párhuzamosan, stabil és szinkronizált képhasználatot biztosítva. (1. ábra).



A kamera szenzorok által továbbított kép.



A szimulációban látható kép.

1. ábra: A publikált kamerák képe és a szimulátorban látható kép.

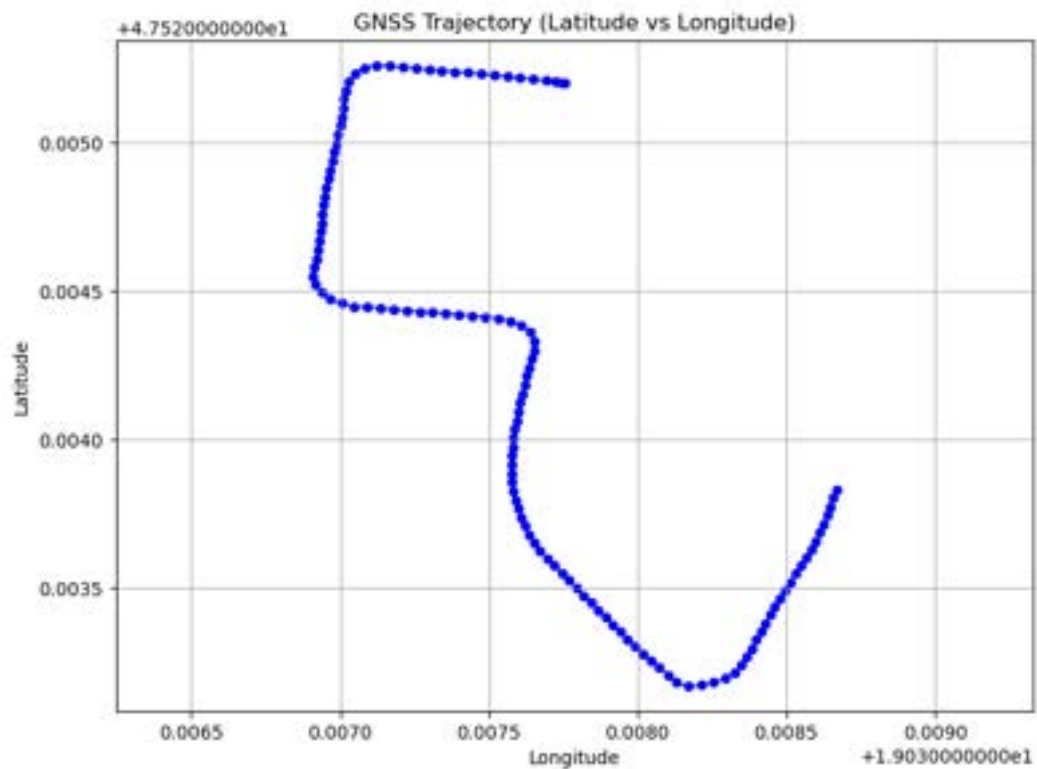
Az aiSim GUI szoftverben elérhető egy beépített Sensor Editor, amely lehetővé teszi több szenzor egyidejű elhelyezését a járművön tetszőleges pozíciókban. A szenzorok két fő kategóriába sorolhatók: real-life-based és ground truth érzékelők. (2. ábra)



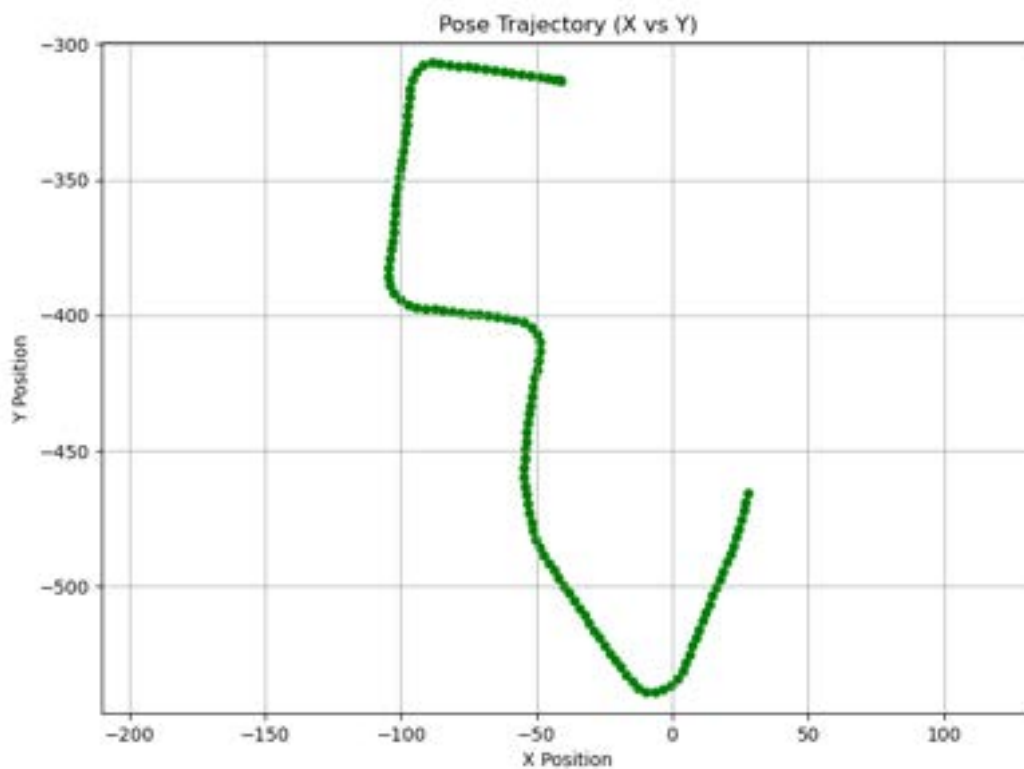
2. ábra: Sensor editor.

A real-life-based szenzorok célja a valós világban használt eszközök viselkedésének pontos szimulációja. Ezzel szemben a ground truth szenzorok nem törekednek realizmusra; ehelyett ideális, hibamentes adatokat szolgáltatnak a jármű pozíciójáról, sebességéről, gyorsulásáról vagy orientációjáról. A két kategória kombinálása lehetőséget ad komplex szimulációs környezetek létrehozására, ahol a valós érzékelők szimulált viselkedése összevethető a referenciaadatokkal.

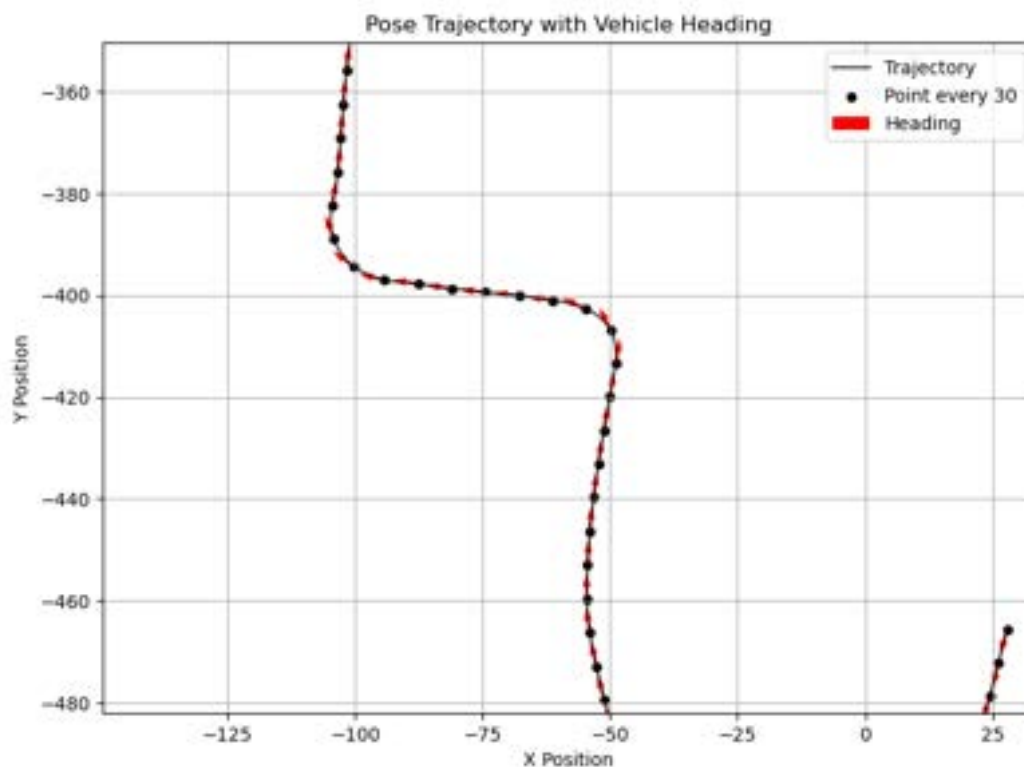
A szimulációban számos pozíció- és mozgásérzékelő is elérhető, mint például a GPS-szenzor, az IMU-szenzor, valamint a Vehicle-szenzor. Ez utóbbi egy ground truth érzékelő, amely a jármű helyzetét, irányát, sebességét és gyorsulását szolgáltatja a szimulált világ globális koordinátarendszerében. Az ezen szenzorok által generált adatok valós időben kerülnek publikálásra, lehetővé téve azok azonnali elemzését vagy vizualizációját. A jelenlegi implementációban a beérkező adatok egy egyszerű Python program segítségével kerülnek begyűjtésre majd szenzoronként külön CSV fájlokba kerülnek, majd a matplotlib könyvtár segítségével diagramokon kerülnek megjelenítésre. A 3. ábrán látható diagram a GPS szenzor által mért adatok alapján rajzolja le az autó által megtett útvonalat. A 4. ábrán a Vehicle szenzor által biztosított koordináta adatok szerepelnek. Az 5. ábra ugyanazt az útvonalat mutatja, mint a 4. ábra, de itt piros nyilakkal jelezve van az autó iránya is bizonyos koordinátáknál.



3. ábra: GPS szenzor által mért koordináták.



4. ábra: Vehicle szenzor által mért koordináták.



5. ábra: Vehicle szenzor által mért orientáció és koordináták.

Az aiSim további előnyei közé tartozik a szimulációs forgatókönyvek széles körű testreszabhatósága. A rendszer lehetőséget nyújt a napszak, időjárási körülmények, útminőség és forgalmi viszonyok részletes beállítására. (6. ábra) A kamerák esetében két gyakran előforduló, zavaró környezeti hatás sárfelverődés és kondenzáció is szimulálható. (7. ábra) (8. ábra) Ezek az effektusok jelentős hatással lehetnek a képi érzékelés minőségére, így különösen hasznosak a vizuális odometria algoritmusainak stresszteszteléséhez.



Forgalom nélkül, havas úton.



Forgalommal, nappal.



Forgalom nélkül, éjszaka.

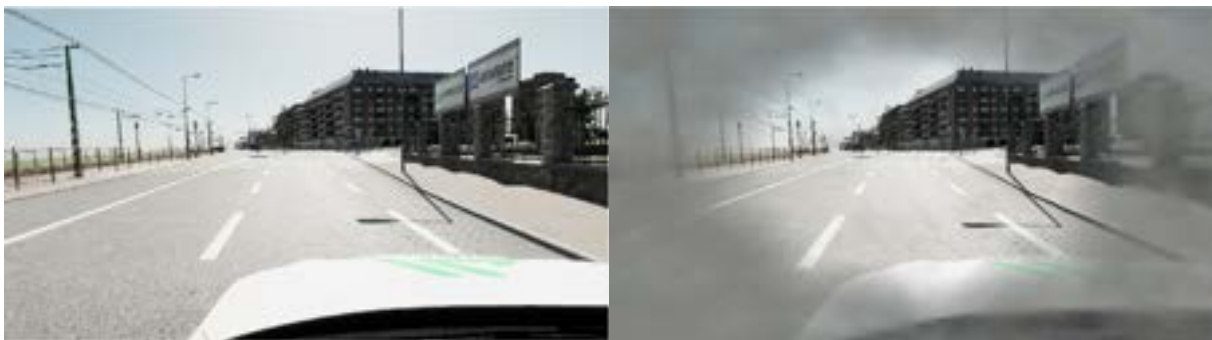
Forgalommal, napsütéses időben, rossz útminőséggel.

6. ábra: Néhány példa különböző szimulációs forgatókönyv paraméter kombinációkra.



7. ábra: Sárfelverődés szimulálása

(Balra: szennyeződés mentes felvétel, jobbra: szennyezett felvétel).



8. ábra: Kondenzáció szimulálása

(Balra: kondenzáció mentes kamerakép, jobbra: kamerakép kondenzációval).

5. Összegzés

A kutatás során olyan szimulációs rendszer került kialakításra, amely lehetővé teszi vizuális odometria algoritmusok valós idejű tesztelését, valósághű körülmények között. Az aiSim és a ROS 2 integrációja révén a rendszer képes különféle szenzoradatok szinkronizált publikálására, valamint valós környezeti feltételek – például időjárás, forgalom, vagy látási zavarok – élethű

modellezésére. A rendszer segítségével lehetőség nyílik különböző szimulációs scenáriók egyszerű konfigurálására és reprodukálására.

A jövőbeni munkák célja saját járműmodell és szenzorkonfiguráció integrálása, valamint valós környezeteken alapuló tesztpályák beillesztése a szimulációba. Mint például az egyetem közvetlen környezete. Majd ezekre a pályákra készíteni több különböző szimulációs forgatókönyvet, amikben változatos forgalom és időjárás opciók lesznek elérhetőek.

6. Köszönetnyilvánítás

A publikációban szereplő kutatást a Széchenyi István Egyetem az Európai Unió támogatásával valósította meg, az RRF-2.3.1-21-2022-00004 azonosítójú, Mesterséges Intelligencia Nemzeti Laboratórium projekt keretében.

Irodalomjegyzék

- [1] Y. X. Tan, M. B. Prasetyo, M. A. Daffa, D. S. Nitin, és M. Meghjani, “Evaluating visual odometry methods for autonomous driving in rain,” arXiv preprint arXiv:2309.05249, 2023. [Online]. Elérhető: <https://arxiv.org/abs/2309.05249>
- [2] D. Scaramuzza és F. Fraundorfer, “Visual odometry [Tutorial],” IEEE Robotics and Automation Magazine, vol. 18, no. 4, pp. 80–92, Dec. 2011.
- [3] A. Howard, H. Seraji, és E. Tunstel, “Real-time stereo visual odometry for autonomous ground vehicles,” in Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS), 2008, pp. 3946–3952.
- [4] C. Li, J. Zhang, K. Yuan, Y. Fang, és X. Li, “Simulation-based validation for autonomous driving systems,” arXiv preprint arXiv:2301.03941, 2023. [Online]. Elérhető: <https://arxiv.org/abs/2301.03941>
- [5] Z. Xiang, W. Li, X. Chen, és Y. Hu, “ViLiVO: Virtual LiDAR-Visual Odometry for an autonomous vehicle with a multi-camera system,” arXiv preprint arXiv:1909.12947, 2019. [Online]. Elérhető: <https://arxiv.org/abs/1909.12947>

[6] aiMotive, “aiSim simulation platform,” 2024. [Online]. Elérhető: <https://aimotive.com/aisim>

[7] aiMotive, “aiSim Customer Documentation,” belső dokumentáció, bizalmas, licenzmegállapodással hozzáférhető, 2025.

[8] S. Macenski, T. Foote, B. Gerkey, C. Lalancette és W. Woodall, “Robot operating system 2: Design, architecture, and uses in the wild,” Science Robotics, vol. 7, no. 66, eabm6074, 2022. máj. [Online]. Elérhető: <https://www.science.org/doi/abs/10.1126/scirobotics.abm6074>

Mesterséges intelligencia rendszerek alkalmazása a mobilitásban

The Use of Artificial Intelligence Systems in the Field of Mobility

Gulyás Péter, Kovács Ádám Tamás, Légrádi Márton Gábor

Széchenyi István Egyetem, Járműipari Kutatóközpont

gulyas.peter@ga.sze.hu, kovacs.adam@ga.sze.hu, legradi.marton.gabor@ga.sze.hu

Absztrakt

A mesterséges intelligencia (MI) rendszerek alkalmazása a modern mobilitási megoldásokban egyre nagyobb jelentőséggel bír, különösen az elektromos hajtások és a fenntartható közlekedési eszközök fejlesztésében. Jelen projekt célja egy intelligens, többcélú elektromos jármű hajtáslánc és akkumulátor menedzsment rendszer fejlesztése volt, amelyet kifejezetten egy gyermekek számára tervezett gokart platformhoz valósítottunk meg a QUATIC Technologies Kft. számára. A fejlesztés középpontjában a motorsport világán belüli tehetséggondozás támogatása, az elektromos hajtások népszerűsítése, valamint a környezeti fenntarthatóság és a hosszú távú költségcsökkentés állt. A projekt során egy meglévő gokartváz részletes műszaki felmérését követően terveztünk egy teljes elektromos hajtásláncot, amely tartalmazza az akkumulátor csomagot, a lánchajtást, valamint a motor mechanikai és elektronikai illesztését. A hajtáslánc teljesítményét úgy méreteztük, hogy az elérje a belső égésű motoros gokartok szintjét, ugyanakkor a teljesítmény szabályozhatósága révén lehetőséget biztosítson különböző képességszintű versenyzők számára történő alkalmazásra. A gépészeti és elektromos tervezés során kulcsfontosságú szempont volt a modularitás, az energiahatékonyság és a megbízhatóság. A fejlesztési folyamat során a mesterséges intelligenciát adatok gyűjtésére és automatizált adatelemzésre alkalmaztuk. A rendszer képes különböző érzékelők adatainak feldolgozására (pl. motorhőmérséklet, áramfelvétel, gyorsulás, akkumulátor-töltöttség), amelyek alapján lehetőség nyílik a hajtáslánc valós idejű optimalizálására, a terhelési ciklusok elemzésére, valamint a járműhasználat hosszú távú kiértékelésére. Az MI-alapú elemzés elősegíti a karbantartási igények előrejelzését és támogatja a felhasználói viselkedés elemzését is, amely különösen fontos szempont lehet a fiatal versenyzők fejlődésének nyomon követésében. Összességében a projekt eredménye egy intelligens, fenntartható és költséghatékony hajtáslánc megoldás lett, amely hozzájárul az elektromos gokartozás technológiai fejlődéséhez, valamint a motorsport jövő generációinak támogatásához. A mesterséges intelligencia integrálása nem csupán a technológiai színvonalat emeli, hanem új lehetőségeket is teremt az oktatásban, a versenysportban és a járműipari kutatás-fejlesztés területén.

Kulcsszavak: elektromos gokart, Mesterséges Intelligencia, tehetséggondozás

Abstract

The use of Artificial Intelligence (AI) systems in modern mobility solutions is becoming increasingly important, especially in the development of electric drives and sustainable transport. The objective of this project was to develop an intelligent, multi-purpose electric vehicle powertrain and battery management system, implemented specifically for a go-kart platform designed for children, for QUATIC Technologies Ltd. The focus of the development was to support talent development within the world of motorsports, promote electric drive systems, as well as environmental sustainability and long-term cost reduction. Following a detailed technical assessment of an existing go-kart chassis, the project involved the design of a complete electric powertrain, including the battery pack, chain drive and mechanical and electronic integration of the engine. The power output of the drivetrain was scaled to match that of an internal combustion engine go-kart, while at the same time providing the possibility of adapting the power to drivers of different skill levels through its controllability. Modularity, energy efficiency and reliability were key considerations in the mechanical and electrical design. During the development process, artificial intelligence was used for data collection and automated data analysis. The system is able to process data from various sensors (e.g. engine temperature, current consumption, acceleration, battery charge), which can be used to optimise the powertrain in real time, analyse load cycles and evaluate vehicle usage in the long term. AI-based analysis will also help predict maintenance needs and support the analysis of user behaviour, which can be a particularly important aspect in monitoring the development of young drivers. Overall, the project has resulted in an intelligent, sustainable and cost-effective powertrain solution that will contribute to the technological development of electric karting and support future generations of motorsport. The integration of artificial intelligence not only raises the technological level, but also creates new opportunities in education, racing and automotive R&D.

Keywords: *electric go-kart, Artificial Intelligence, talent management*

1. Bevezetés

Az elektromos gokartokat az elmúlt években nemcsak hobbicélokra, hanem oktatási platformként is egyre gyakrabban alkalmazzák. Például a Portói Egyetem projektje egy háromfázisú aszinkron motorral működő, alacsony költségű elektromos gokartot mutat be, amely az indukciós motorvezérlés és teljesítményelektronika oktatására készült. A rendszer felépítése egyszerű: inverter, mikrokontroller-alapú vezérlő és 24V-os DC akkumulátor biztosítja a hajtást. [2, 8, 11]

A didaktikai cél az volt, hogy a hallgatók mérési, értelmezési és vezérlési készségeket sajátítsanak el egy valós járművön, miközben a költségek alacsonyan maradnak. Az indukciós motor alkalmazása – a szervomotorokkal szemben – ipari relevanciája miatt különösen indokolt volt. [4, 8, 10]

Egy másik kutatás részletesen bemutatja az elektromos gokart mechanikai tervezését. A jármű alváza AISI 1018 acélból készült, amely jó kompromisszumot nyújt az ár, szilárdság és alakíthatóság terén. A végeselemes szimulációk során mért maximális deformációk és feszültségek alapján az alváz megfelelőnek bizonyult frontális, oldalsó és hátsó ütközési helyzetekben. [1, 5]

A jármű hajtása két 2000 W-os BLDC motorral történt, amelyek 6,16 Nm nyomatékot adtak le. A sebesség elérhette a 60 km/h-t. A rendszer nem használt differenciált és felfüggesztést sem, így a hátsó tengely merev, a kormányzás Ackermann-geometriára épült. A járművet CAN-alapú kommunikációval, LCD kijelzőkkel és akár Arduino-s vezérléssel is felszerelték. [1, 6]

A Budapesti Műszaki és Gazdaságtudományi Egyetem egy ROS-alapú (Robot Operating System) elektromos gokart-platformot fejlesztett ki autonóm járműirányítás oktatására. A platformba különféle szenzorokat, aktuátorokat és vezérlőegységeket integráltak, a kommunikáció CAN-hálózaton zajlik. A jármű elsősorban kutatási és oktatási célokat szolgál, például vezetéstámogató rendszerek és AI-alapú vezérlés tesztelésére. Ez a rendszer rávilágít arra, hogy az elektromos gokartok már nemcsak hajtástechnikai oktatási célokat szolgálnak, hanem az autonóm rendszerek kísérleti platformjaivá is váltak. [2, 7, 9]

A Tecnológico de Monterrey egy komplett tantárgyi blokkot épített ki elektromobilitás témakörében, amely során a hallgatók egy teljes elektromos járművet terveznek és integrálnak. A projekt során BLDC motorokat, hidraulikus tárcsafékrendszert, digitális műszerfalat, CAN

kommunikációs rendszert és vezérlőelektronikát kellett összeállítaniuk. A program különlegessége az ipari partner (Bosch) bevonása volt, amely releváns kihívásokat és mentori támogatást biztosított a hallgatóknak. A cél az volt, hogy valós mérnöki problémákon keresztül fejlesszék a jövő elektromosjármű-fejlesztőinek készségeit. [3, 8, 9]

2. Módszertan

A fejlesztés elsődleges célja egy olyan elektromos hajtáslánc kialakítása volt, amely teljesítmény szempontjából felveszi a versenyt a belsőégésű motorral szerelt gyerekgokartokkal. A rendszernek emellett ki kellett használnia az elektromos hajtás előnyeit: alacsonyabb karbantartási igény, csendesebb működés, szabályozható teljesítmény és fenntartható üzemelés. A konstrukció alapjául egy szabványos gokartváz szolgált, amelyet geometriáját, súlyeloszlását és szerkezeti integritását figyelembe véve alakítottunk ki. Az új hajtáslánc robusztus láncajtással valósult meg, amely jól illeszkedik az elektromos motor karakterisztikájához. A meghajtáshoz a Golden Motor BLDC-108 típusú állandómágneses, szénkefe nélküli egyenáramú motort választottuk, amelyet Trampa VESC MK6 motorvezérlővel szabályozunk. A rendszert kizárólag motoros üzemmódban használjuk, rekuperáció nélkül. A hagyományos trapézgerjesztés helyett szinuszos kommutációt alkalmazunk, amely:

- javítja a hatásfokot, ezáltal csökkenti az energiaveszteségeket,
- halkabb működést eredményez,
- egyenletesebb nyomatékot biztosít, így csökkentve a mechanikai alkatrészek terhelését és fáradását.



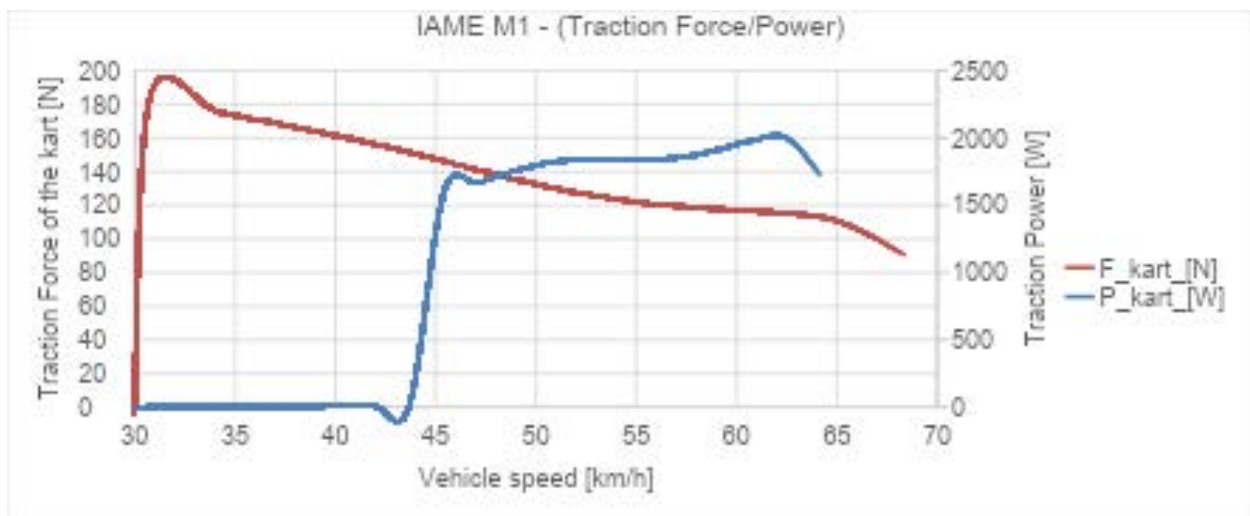
1. ábra: A választott elektromos gép: Golden Motor gyártó BLDC-108 motorja (saját kép)

A motor homlokfelületre rögzíthető, a hajtó lánckerekek köztes agyon keresztül kapcsolódnak. A hajtó lánckerekek (25, 27, 29 fogszám) variálhatók, így az áttétel és ezzel a végsebesség, illetve a gyorsulás is hangolható. A tengelyirányú és radiális rögzítést reteszkötés és hernyócsavaros biztosítás garantálja. A motor tengelye nem igényel központosító kialakítást, az agy és a felfogatás ezt mechanikailag kompenzálja. A motor rögzítése négy darab M6-os csavarral történik, ahol fontos a megfelelő előfeszítés és menetrögzítő paszta használata – különösen mivel az alumínium ház sérülékeny túlzott meghúzás esetén. A hőelvezetést a motort szabadon körbeáramló levegő biztosítja, amit szükség esetén légterelő elemmel fokozhatunk. A lánc feszessége a motortartó bak hosszirányú eltolásával állítható, így nincs szükség külön láncfeszítőre. A vibrációk és dinamikus terhelések csökkentése érdekében rezgéscsillapító alátéteket alkalmazunk a rögzítési pontokon, például az elektronikát tartó baknál is.

1. táblázat: Golden Motor BLDC-108 motor adatlapja (saját szerkesztés)

BLDC-108 motor műszaki adatai		
Névleges teljesítmény:	1500 [W]	
Névleges feszültség:	48 [V]	
Névleges áram:	40 [A]	
Maximális áram:	140 [A]	(15 másodpercig)
Névleges fordulatszám:	4000 [rpm]	(@48V)
Névleges nyomaték:	3,5 [Nm]	
Maximális nyomaték:	15 [Nm]	
Fázis ellenállás:	29 [mΩ]	
Induktivitás:	87 [μH]	
Hall elemek eltolása:	120 [fok]	
Fáziseltolás:	-120 [fok]	
Póluspárok száma:	5	
IP osztály:	IP54	
Nyomaték konstans	0.0875 [Nm/A]	
Feszültség konstans	83.33 [rpm/V]	

A fenti táblázat összefoglalja a legfontosabb műszaki jellemzőit a meghajtómotornak választott BLDC-108 motorról. Névleges teljesítménye 1500 W, amelyet mi 1900 W-ra emeltünk fel annak reményében, hogy a motor körül áramló levegő – a jármű mozgásából eredően – visszahűti annyira, hogy minden körülmény között elviseli a motor ezt a teljesítmény szintet. Mivel a belsőégésű motor jellemző maximális fordulatszáma jóval 10.000 [rpm] felett van, így a 4.000 [rpm] maximális fordulatszám indokolta a viszonylag nagy hajtó lánckerék fogszámokat. Az alkalmazás során nem kívántunk mezőgyengítést alkalmazni. Célunk volt a hajtásnál, hogy az IAME M1 motor alternatíváját adja a mi hajtásunk. A belsőégésű kategóriában a manapság alkalmazott technikai szabályban meghatározott áttétellel számoltunk.



2. ábra: IAME M1 motor adatok (saját szerkesztés)

A fenti ábra bemutatja az IAME M1 motorral szerelt Bambino kategória jellemző teljesítmény és vonóerő konfigurációját. A vonóerő diagram kb. 30 km/h felett folyamatosan csökken és 65 km/h felett drasztikusan csökken. A célunk tehát az volt az áttétel és teljesítmény konfigurálásnál, hogy ezt a görbét minél pontosabban közelítsük az elektromos hajtásláncunkkal.

3. A hajtáslánc adatlapi értékeiből számolt menetdinamikai jellemzők

Az előző fejezetben kifejtettem, hogy az elektromos hajtásláncunk az IAME M1 motor technikai jellemzőit kívánta teljesíteni, így létrehoztam egy excel táblázatot, ahol a meghatározott paraméterek mellett meghatározható az a pár műszaki paraméter, amellyel az elektromos hajtás egyenértékű hajtás lehet a Bambino gokartok legerősebb belsőégésű motorjával szemben. A hajtásláncához ennek a táblázatnak a segítségével határoztuk meg a hajtó fogaskerék fogszámát, amelyet a következő táblázat foglal össze:

2. táblázat: Alkalmazható áttételek a BLDC-108 motorral (saját szerkesztés)

Ratio	25	27	29
66	2.640	2.444	2.276
68	2.720	2.519	2.345
70	2.800	2.593	2.414
72	2.880	2.667	2.483
74	2.960	2.741	2.552
76	3.040	2.815	2.621
78	3.120	2.889	2.690
80	3.200	2.963	2.759

A hajtott lánckerekek méreténél figyelembe vettük, hogy a gokartokhoz mely lánckerekeket lehet beszerezni és alkalmazni. A 80-as méret fölé nem tanácsos menni, mert a lánckerék fogai közel kerülnek a talajszinthez, így megnő a lánccsakadás veszélye. Választásunk a 80-as lánckerekre esett – ami a 2.963-as lassító áttételt határozza meg –, mert ez egy gyakran alkalmazott és gyárilag szállított hajtott fogaskerék méret a Bambino kategóriában. Jól látható, hogy a három hajtó lánckerékkel és a 60-80 méretig variálható hajtott lánckerékkel igen nagy áttétel tartományt meg lehet valósítani – ami 2.276 és 3.2 közé adódik. A számításokat Excel-alapú adatmodellezéssel támasztottuk alá, ahol figyelembe vettük az elérhető vonóerőt, gyorsulást és a teljesítményhatárokat. Ezek alapján határoztuk meg a vezérlőbe programozott beállításokat is (pl. 1900 W teljesítménykorlát, 120 A fázisáram).

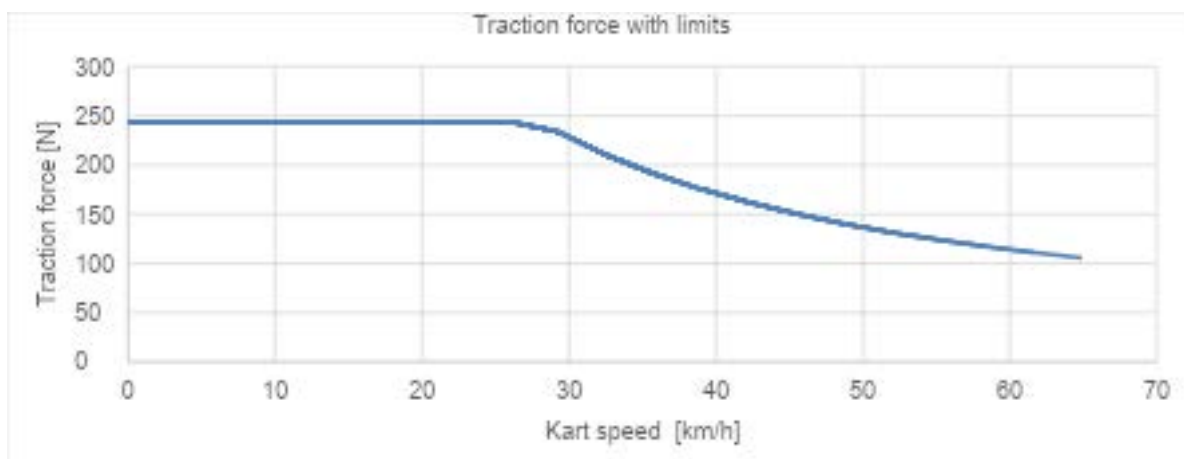
3. táblázat: Hajtáslánc konfiguráló táblázat BLDC-108 motorhoz (saját szerkesztés)

Power limit	1900	[W]
Phase current limit BLDC 108	120	[A]
Applied gear	2.963	
Max speed @46V	4000	[rpm]
Torque constant	0.0875	[Nm/A]
Maximum Traction force	244.01	[N]
DC Current limit must be	47	[A]
Mass of the kart (overall)	85	[kg]
Mass of the kart	54.7	[kg]
Calculated maximum speed	64.9	[km/h]

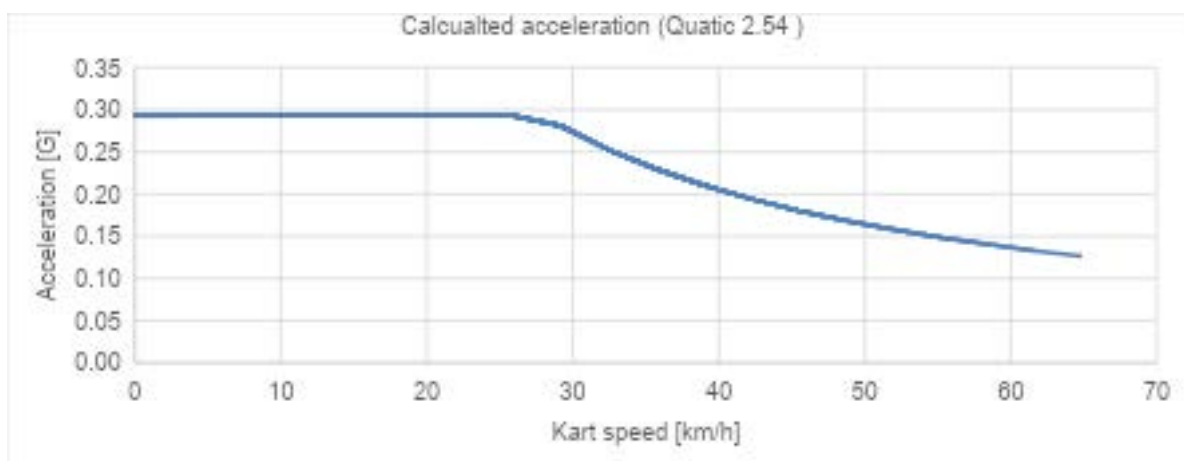
Az alábbi diagramokat kapjuk a fenti táblázatot alkalmazva:



3. ábra: Hajtáslánc teljesítmény diagramja (saját szerkesztés)



4. ábra: Gokart vonóerő diagramja (saját szerkesztés)



5. ábra: Gokart gyorsulás diagramja (saját szerkesztés)

A fenti diagramok bemutatják, hogy mi a hajtás elméleti menetdinamikai jellemzői, természetesen a gokart ellenállás erejét figyelmen kívül hagyva. A gokart ellenállása nagy mértékben függ a gokart össztömegétől, mechanikai beállításán át a pilóta magasságától – talán a legnagyobb mértékben. A tervezés pontosságát csak a versenypályán logolt adatok segítségével lehetséges validálni. A hajtáslánc a Quatic 2.54 modell nevet kapta, mivel az 1900 W teljesítmény 2.54 LE.

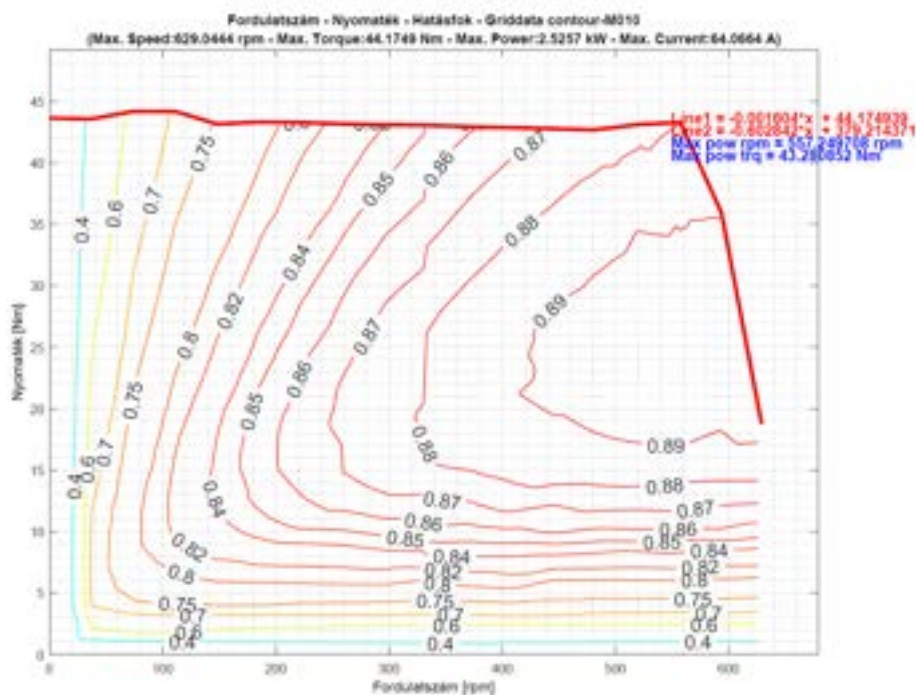
4. A hajtáslánc mérése

A hajtáslánc működésének elsődleges ellenőrzését fékpadi mérésekkel végeztük. Mivel a rendelkezésre álló fékpád 700 fordulat/perc fölött nem tudott elegendő fékerőt kifejteni, a teszteléshez alternatív áttételt alkalmaztunk: a hajtó lánckerék 12, a hajtott pedig 80 fogas volt, így 6.66-os lassító áttételt kaptunk. A jelenlegi vizsgálatok elsősorban paraméterezési célokat

szolgáltak, azonban a közeljövőben további méréseket tervezünk, hogy minden hajtáskonfigurációról részletes hatásfokmező álljon rendelkezésre – különösen a csökkentett nyomatékú munkaterület vonatkozásában. A méréseket egy teljesen automatizált fékpadon végeztük, amely .tdms formátumú logfájlt generál, így az adatok Excelben és MATLAB-ban is feldolgozhatók. Minden rögzített munkapont 200 mérési értéket tartalmaz, és a rendszer kizárólag stationer állapotban rögzít, így a fájl tranziensmentes és pontos eredményeket ad. A BLDC motor jellegzetessége, miszerint 0 rpm közelében az egyenlőtlen nyomatékeloszlás – kis mechanikai szögeltérés mellett – jelentős eltéréseket okozhat a kívánt nyomatékhoz képest. A maximális nyomaték lengése fordulatszám-függő, és jól mutatja a motor szinuszos gerjesztésének hatását. A teszt során alkalmazott fázisáram 50 A volt, így a mérés lefedte a névleges tartományt, valamint annak körülbelül 25%-os túlterhelési zónáját is, teljesítménykorlátozás nélkül.

5. A hajtáslánc műszaki jellemzőinek pontosítása a mérések alapján

A fékpad mérések lehetőséget adtak arra, hogy a motor adatlap paramétereit pontosítsuk, különösen a nyomatékkonstans értékét, amely a gyakorlatban kissé alacsonyabbnak bizonyult, mint a gyártói specifikációban szereplő érték. A kapott adatokat visszaszámoltuk a motor tengelyére vonatkozóan, kizárva a mérőtengely és a lánchajtás közötti veszteségeket, amelyek becsléseink szerint 97–99% közötti hatásfokkal működnek a vizsgált tartományban. Ez lehetővé tette, hogy a tervezési számításokhoz a lehető legpontosabb adatokat használjuk. A fékezőnyomaték akkor lép fel, amikor a motorvezérlő nem kap nyomatékreferencia értéket. A 0A nyomatékreferencia mellett a motor és a lánchajtás mechanikai veszteségei és a motor fognyomatéka „fékezi” a hajtást, ami ugyan kis teljesítményű, de mindig fennálló sebességfüggő fékhatást eredményező jellemző. A kifejtett nyomaték mellett fontos mérési adat a hatásfokmező, ahol az adott munkatérben a hajtás energia jellemzőit szemléltethetjük a diagramon. A hatásfokmező mindig adott hajtáslánc, motor, motorvezérlő összeállításra jellemző és alakja C görbéket mutat. A C görbék felső részének függőleges való közelítése jelenti a jelentősen túlterhelhető hajtásokat, míg a C görbék erősen „hegyezett” alakja az adott munkapontra optimalizált hajtást jelentik.



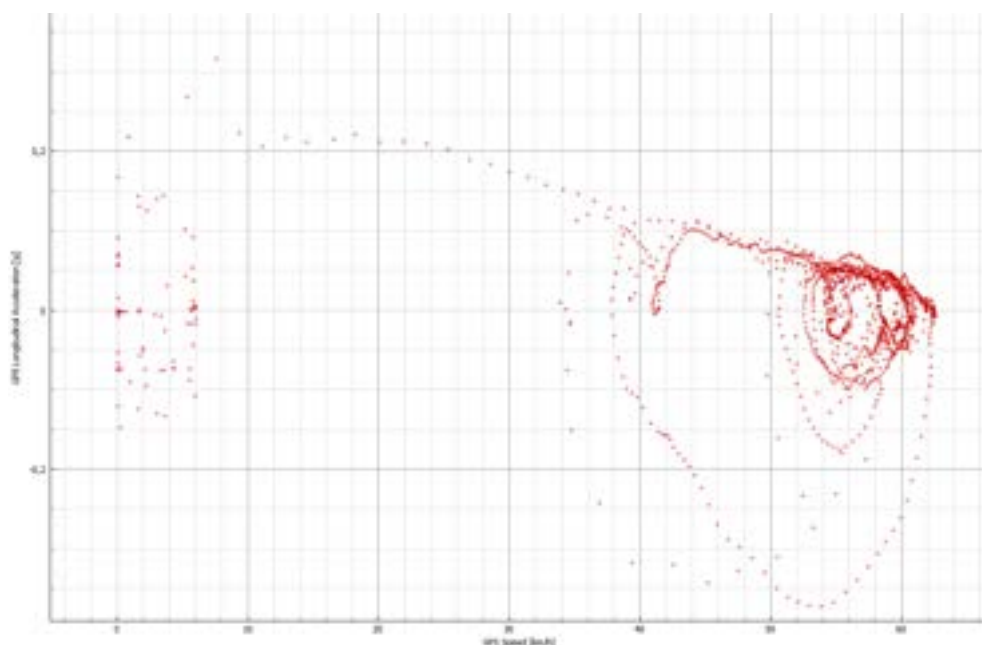
6. ábra: Hatásfokmező C kagylógörbéi (más motor paraméterekkel futtatott tesztmérés) (saját szerkesztés)

Az ábra C-alakú kagylógörbéi alapján megállapítható, hogy a hajtás a 1400–1600 W közötti tartományban éri el a legjobb hatásfokát, amely elérheti a 89%-ot. A 1900 W-os teljesítménykorlát is ebbe az optimális zónába esik, így a hajtás a maximális sebesség közelében is megfelelő energetikai hatékonysággal üzemel. A motor ebben az állapotban körülbelül 200 W veszteségűt termel, amelyet figyelembe kell venni a hűtési megoldások tervezésekor. A mérések által pontosított nyomatékkonstans érték 0.0804 [Nm/A].

6. A hajtáslánccal szerelt gokartok menetdinamikai jellemzői

A gokartok menetdinamikai tulajdonságai azok a fizikai és szerkezeti jellemzők, amelyek meghatározzák a jármű gyorsulását, fékezési képességét, kanyarodási stabilitását és általános vezethetőségét. Ezek a viselkedési jellemzők elsősorban a hajtáslánc, futómű, kormányzás és abroncsok kölcsönhatásából adódnak, és különösen fontosak versenykörülmények között. Motor teljesítménye és nyomatéka: Ezek határozzák meg a jármű gyorsulását és végsebességét. Bár az áttétel módosításával lehetne finomítani a karakterisztikát, a versenyszabályok ezt nem engedik, ezért az általunk alkalmazott áttétel minden versenyen kötelezően használandó. Az eBambini kategóriában ritkán alkalmaznak klasszikus fékezést; a fiatal pilóták többnyire a kormányzás révén lassítják a járművet. Tipikus fékezési pontok a szűk visszafordítók és

derékszögű kanyarok, ahol a sebesség 60 km/h-ról akár 40 km/h alá is csökkenhet. A kormánygeometriai paraméterek (pl. utánfutás, összetartás, csapdőlés) határozzák meg, mennyire gyorsan és pontosan irányítható a gokart. A kategória technikai előírásai szerint a beállítások egységesek, így minden versenygokarton azonos konfigurációt alkalmazunk. Ezeket a jármű tömegeloszlása és a futómű beállításai befolyásolják. A fejlesztés során törekedtünk arra, hogy a nehezebb alkatrészek elosztása révén a tömegközéppont kb. 60/40%-os arányban helyezkedjen el az első és hátsó tengely között, ez javítja a kanyarodási stabilitást és az egyenesfutást. Mindezek mellett a hajtáslánc beállításai verseny közben nem módosíthatók – ez csak külön engedéllyel, minden járműre egységesen történhet meg. A korábbi mérések alapján elméleti vonóerő- és gyorsulási diagramokat is készítettünk. Az eredmények alapján a Quatic 2.54 hajtáslánc teljesítménye nemcsak eléri, hanem bizonyos sebességtartományokban meg is haladja az azonos kategóriában használt belsőégésű motoros hajtásokét. Az adatgyűjtéshez minden gokartot felszereltünk UniGo One típusú adatrögzítővel, amely képes sebesség-, gyorsulás- és nyomvonaladatokat rögzítésére GPS-alapon, fordulatszámjel nélkül. A rendszer kifejezetten gokartos alkalmazásra lett fejlesztve. Az UniGo One adatgyűjtő rendszer rendelkezik saját diagnosztikai programmal, amellyel könnyen és gyorsan tetszőleges diagramok ábrázolhatók. A legfontosabb ilyen diagram a hajtásláncot vizsgálva a sebesség függvényében ábrázolt hosszirányú gyorsulás diagram.

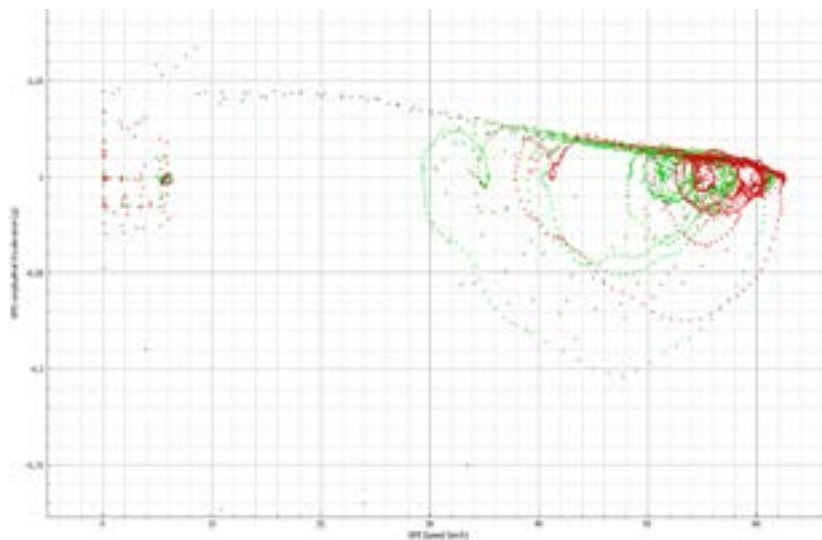


7. ábra: Kecskeméti versenyen rögzített hosszirányú gyorsulás diagram (saját szerkesztés)

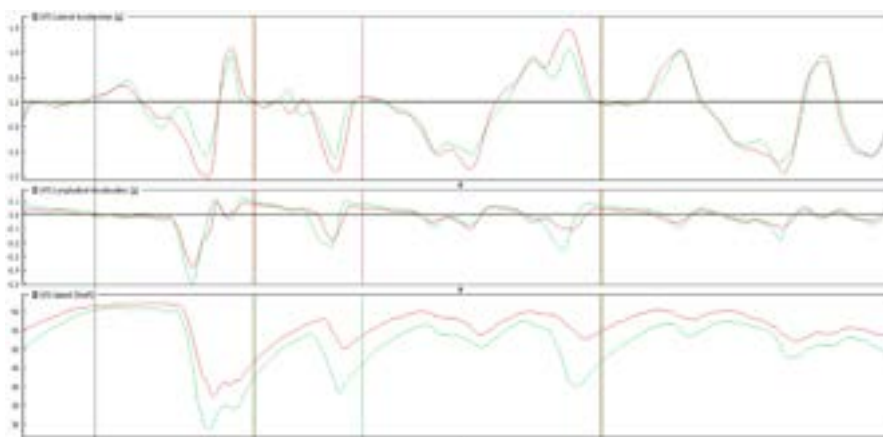
A mért hosszirányú gyorsulás jól mutatja, hogy a gyorsulás csökken, amint a jármű ellenállásai (pl. légellenállás, abroncs tapadás) érvényesülni kezdenek. A kezdeti gyorsulás értéke 0.24–0.25 G, de ez hamar 0.21 G körülire esik. Ezen túl, kb. 25 km/h-nál már jelentkezik a motorvezérlő által beállított 1900 W-os teljesítménykorlát, amit az akkumulátor aktuális feszültsége is befolyásol.

7. Két hajtáslánc menetdinamikájának összehasonlítása, mentett adatok alapján

Az eBambini kategória egyik alapvetése, hogy minden versenyző azonos műszaki állapotú járművel induljon. A logolt adatok alapján vizsgáltunk két, azonos beállítású hajtást – a gyorsulási karakterisztikájuk megegyezett, de a köridőben akár 6 másodperc eltérés mutatkozott. A 8. és 9. ábrák alapján ez a különbség egyértelműen a vezetési stílusból adódik: a gyengébb köridőt futó pilóta nagyobb és gyakoribb fékezésekkel, lassabb kanyarsebességgel haladt, még akkor is, ha technikailag nem lett volna szükség rá. Ez alátámasztja, hogy az adatrögzítés nemcsak a technika, hanem a pilótaképzés szempontjából is fontos eszköz.



8. ábra: Két hajtás összehasonlítása (köridőben 9%-os különbség) (saját szerkesztés)



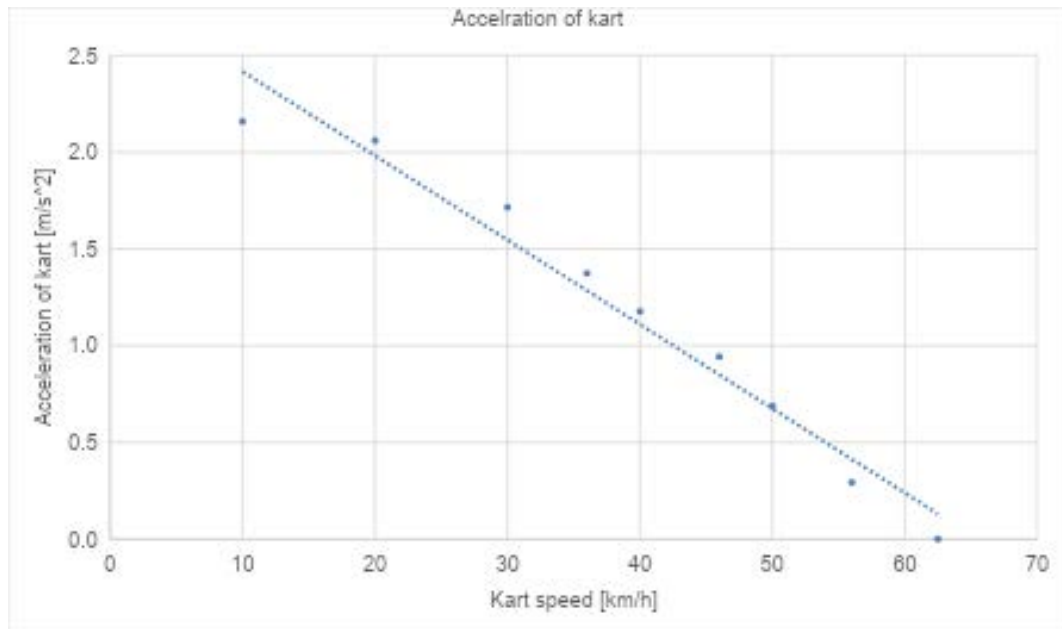
9. ábra: Gyorsulás és sebesség adatok két hajtásláncot összehasonlítva (saját szerkesztés)

A 9. ábra két azonos műszaki konfigurációval felszerelt gokart gyorsulás- és sebességadatait hasonlítja össze. Bár a gyorsulási karakterisztikák szinte teljes egészében megegyeznek, a zöld vonallal jelölt körítő 9%-kal lassabb. Ez a különbség nem műszaki, hanem vezetési tényezőkre vezethető vissza. A grafikon alapján egyértelmű, hogy a zöld adatsorhoz tartozó pilóta gyakrabban és intenzívebben fékezik, különösen a kanyarokban. Bár a technika lehetővé tenné a nagyobb tempót, a kanyarsebesség 5–10 km/h-val alacsonyabb, mint a másik esetben. Ezt erősíti meg az is, hogy a gyors kanyarokban az oldalgyorsulás így is eléri az ± 1 G-t, tehát a jármű képességeit nem használja ki teljes mértékben. Összességében kijelenthető, hogy a két hajtáslánc azonos teljesítményre képes, a kialakuló időkülönbség viszont a vezetési stílusból és a vonalvezetés pontatlanságából adódik.

4. táblázat: Gokart gyorsulás értékei adott sebességeken (saját szerkesztés)

Kart speed [km/h]	Speed [m/s]	Acceleration of kart [G]	Acceleration of kart [m/s ²]
10	2.78	0.22	2.1582
20	5.56	0.21	2.0601
30	8.33	0.18	1.7168
36	10.00	0.14	1.3734
40	11.11	0.12	1.1772
46	12.78	0.10	0.9418
50	13.89	0.07	0.6867
56	15.56	0.03	0.2943
62.5	17.36	0.00	0.0000

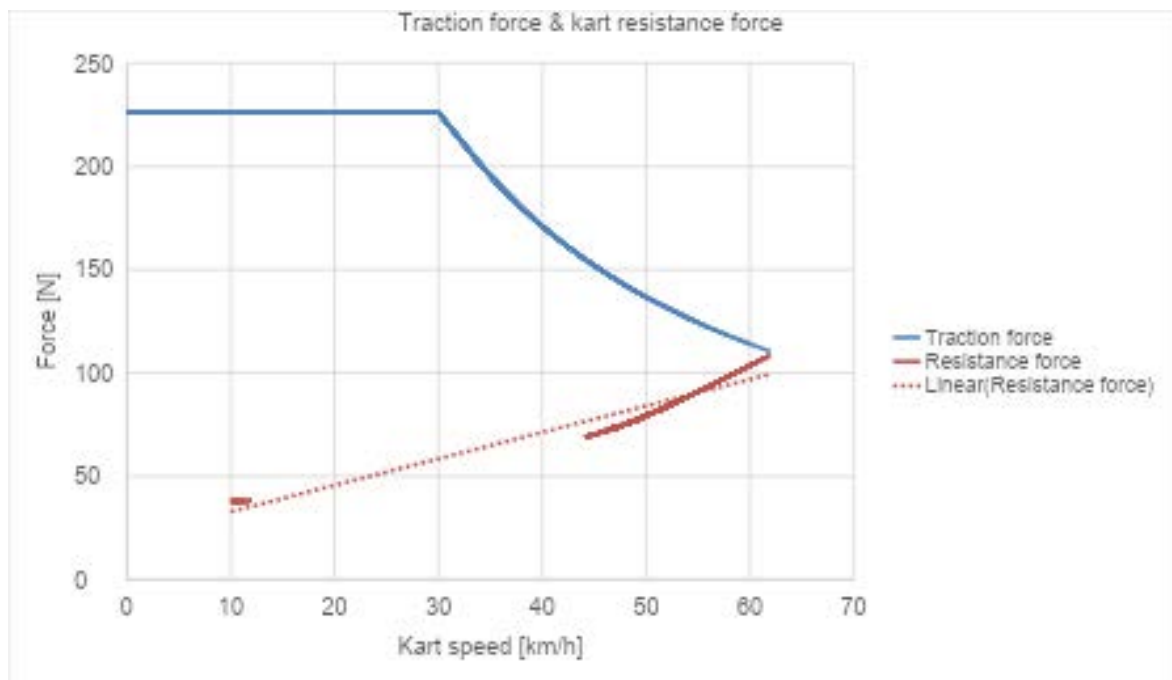
A táblázat értékei alapján felvettem egy harmadfokú polinomot, ami jól leírja az adott gokart és hajtáslánc páros gyorsulását. A diagram és a polinom a következő ábrán olvasható.



10. ábra: Gokart gyorsulását leíró harmadfokú polinom meghatározása (saját szerkesztés)

8. Számított gokart menetellenállás erő

A gokart gyorsulását és hajtás által kifejtett vonóerőt ismerve meghatározható, hogy a gokart milyen ellenállás erőt képez adott sebességen. Ezzel azért érdemes számolni, mert ennél kisebb teljesítményekkel nem nagyon fog üzemelni a hajtás, motoros üzemmódban. A hajtáslánc vonóerejéből kivonjuk a gyorsulás diagramból származtatott gyorsítási erőt, akkor az alábbi diagramot kapjuk.



11. ábra: Számított menetellenállás függvény (saját szerkesztés)

A diagram megmutatja, hogy a gokart hajtásához kb. 110 N vonóerő szükséges 63 km/h járműsebesség mellett. A menetellenállás tagjai vizsgálva az is megállapítható, hogy a gokart ellenállásának legnagyobb rész a légellenállása. Ez a mechanikai kialakításából és a teljesen nyitott felépítésből adódik.

9. Összegzés

A tanulmány célja egy intelligens, moduláris elektromos hajtáslánc fejlesztésének bemutatása gyermek gokart platformra, amely a QUATIC Technologies Kft. megbízásából valósult meg. A projekt során egy 1500 W névleges teljesítményű, állandómágneses BLDC motoron (BLDC-108) alapuló hajtásláncot terveztünk, amelynek vezérlését egy Trampa VESC MK6 vezérlőegység látja el. A rendszer teljesítménye eléri a 1900 W-ot, így versenyképes alternatívát kínál az eBambini kategóriában elterjedt IAME M1 belső égésű motoros hajtásláncokkal szemben. A hajtáslánc konfigurációját numerikus szimulációk és Excel-alapú adatmodellezés segítségével optimalizáltuk, figyelembe véve a teljesítmény-, vonóerő- és gyorsulás-karakterisztikákat. A menetdinamikai célértékeket versenypályán rögzített GPS-alapú adatrögzítéssel validáltuk. A rendszer képes alkalmazkodni különböző képességszintű versenyzőkhöz a beállítható áttételek és teljesítményhatárok révén. A fejlesztés során fékpadi méréseket végeztünk a motor karakterisztikáinak pontosítása érdekében, különös tekintettel a

hatásfokmezőre és a nyomatékkonstansra, amelyet 0,0804 Nm/A értékben határoztunk meg. A motor 1400–1600 W közötti tartományban érte el legjobb, akár 89%-os hatásfokát. A rendszer egyszerű, de hatékony hűtéssel rendelkezik, és nem alkalmaz rekuperációt. Az eredmények alapján a hajtáslánc nemcsak energetikai szempontból bizonyult versenyképesnek, hanem a versenyzői visszajelzések és adatgyűjtések alapján is képes támogatni a fiatal tehetségek fejlődését. A járműbe integrált adatrögzítő rendszerek (UniGo One) lehetővé tették a vezetési stílus és a technikai konfiguráció közti különbségek elemzését, rámutatva a vezetési technika domináns szerepére a köridő alakulásában. A projekt egyúttal rámutatott a mesterséges intelligencia és automatizált adatfeldolgozás lehetőségeire a motorsportban, a karbantartási igények előrejelzésétől kezdve a pilóták viselkedésének elemzéséig. A Quatic 2.54 hajtáslánc így nemcsak műszaki, hanem oktatási és fejlesztési célokat is kiszolgál.

10. Köszönetnyilvánítás

A 2021-1.1.4-GYORSÍTÓSÁV-2022-00002 számú projekt a Kulturális és Innovációs Minisztérium Nemzeti Kutatási Fejlesztési és Innovációs Alapból nyújtott támogatásával, a GYORSÍTÓSÁV pályázati program finanszírozásában valósult meg.

Irodalomjegyzék

- [1.] L. Prabhu, MD Shadan, Harsh Raj.: Design and analysis of electric Go-Kart Sangeetha Krishnamoorthi, Nadeem Akram Department of Mechanical Engineering, Aarupadai Veedu Institute of Technology, Vinayaka Mission's Research Foundation, Deemed to be University, Tamil Nadu, India, 2020
- [2.] Árpád Fehér, Szilárd Aradi, Tamás Bécsi, Péter Gáspár.: Highly Automated Electric Vehicle Platform for Control Education, Department of Control for Transportation and Vehicle Systems, Budapest University of Technology and Economics, Budapest, Hungary, 2020
- [3.] Carlos Renato Vazquez, Diego Cardenas-Fuentes, Alejandro Guajardo-Cuellar.: Project-based Learning on Automotive Engineering: Building an Electrical Vehicle,
- [4.] Mr. Shankar Duraikannan, Dr. Lim Wee Han, Mr. Lioe De Xing, Ms. Subhashini Gopal Krishnan, Dr. Vinesh Thiruchelvam, Dr. Firas B Ismail, Dr. Raed Abdulla, Ms. Vickneswari,

Mr. Tan Gim Heng, Mr. Lioe De Xing, Mr. Ali Ahmadian, Mr. Wilson You Wei Lim, Mr. Eason Choo, Mr. Lim Chee Cheow.: The Influence of Robotics Engineering, Asia Pacific University of Technology & Innovation, April 2013

[5.] Jiale Zhu.: Racing Esports in China-COI's Business Expansion Plan, B.S. Information and Computer Science Fudan University, 2015

[6.] Stephanie Gagnon Rodriguez Peter Trousdale.: ENERGIZING THE FUTURE MOBILITY WORKFORCE IN MICHIGAN, Center for Climate and Energy Solutions, August 2024

[7.] Ron Frielinck.: Learning Factory Configuration Tool: An Approach for Preserving the Value of Educational Learning Factories, University of Twente, Faculty of Engineering Technology, Mechanical Engineering, MSc. Assignment, 30 April 2022 - 15 February 2023

[8.] Carlos Cardoso, Júlio Ferreira, Vítor Alves, Rui Esteves Araújo.: The design and implementation of an electric go-kart for education in motor control, Faculty of Engineering of the University of Porto, International Symposium on Power Electronics, Electrical Drives, Automation and Motion, SPEEDAM 2006

[9.] Mohd Fitri Arshad, Muhammad Faris Ahmad, Amir Khalid, Izuan Amin Ishak¹, Shaiful Fadzil Zainal Abidin, Mohd Azahari Razali.: Investigation The Influences of Electric Motor and Electric Go Kart Performance, Department of Mechanical Engineering Technology, Universiti Tun Hussein Onn Malaysia, 84600 Pagoh, Johor, MALAYSIA, 20 July 2022

[10.] EMMA GRUNDITZ, EMMA JANSSON.: Modelling and Simulation of a Hybrid Electric Vehicle for Shell Eco-marathon and an Electric Go-kart, Master of Science Thesis in Electric Power Engineering Department of Energy and Environment CHALMERS UNIVERSITY OF TECHNOLOGY Performed at Division of Electric Power Engineering and QRTECH AB Göteborg, Sweden, 2009

[11.] Tihomir Mihalić, Josip Hoster, Vladimir Tudić, Toni Kralj.: Concept Design and Development of an Electric Go-Kart Chassis for Undergraduate Education in Vehicle Dynamics and Stress Applications, Department of Mechanical Engineering, Zagreb University of Applied Sciences, 10000 Zagreb, Croatia 2023

„Mesterséges intelligencia rendszerek alkalmazása a mobilitásban”

2025. tavasz
Konferencia kiadvány