

Járművek kinematikai mennyiségeinek meghatározása v ideófelvételek alapján

Extracting kinematic quantities of vehicles based on video recordings

KÖPECZI-BÓCZ Ákos Tamás¹, TAKÁCS Dénes²

Műszaki Mechanikai Tanszék, BME, Műegyetem rkp. 3., H-1111 Budapest, Magyarország
HUN-REN-BME Gépek Dinamikája Kutatócsoport, BME, Műegyetem rkp. 3., H-1111 Budapest, Magyarország
kopeczi@mm.bme.hu, takacs@mm.bme.hu

Abstract

Our research aims to create a video processing algorithm suitable for monitoring moving vehicles, determining the speed of vehicles, and counting traffic, as well as providing measurement data for parameter fitting of equations describing individual traffic participants. We solve the problem with classic image recognition methods that process video recordings.

Keywords: Image processing, vehicle dynamics, data-driven techniques, measurement and evaluation

Kivonat

Kutatásunk célja egy olyan videófeldolgozó algoritmus létrehozása, amely alkalmas mozgó járművek megfigyelésére, járművek sebességének meghatározására, forgalomszámlálására, valamint mérési adatokat szolgáltat a forgalom egyes résztvevőit leíró egyenletek paraméterillesztéséhez. A problémát klasszikus képfeldolgozási módszerekkel oldjuk meg mely forgalomról készített videókat dolgoz fel.

Kulcsszavak: Képfeldolgozás, járműdinamika, adat alapú módszer, mérés és kiértékelés

BEVEZETÉS

Az ipari forradalmak eredményeként a közlekedés különböző aspektusai eltérő mértékben törtek előre. Összességében elmondható, hogy a közúti közlekedés a 20. század során erős ütemben fejlődött a belső égésű motorok által meghajtott járművek elterjedésével [1]. A távolságok időben lerövidültek, és az évszázad során a társadalom nagy része számára elérhetővé váltak a személygépjárművek. Napjainkban a közúton való közlekedés olyan szerves részévé vált a mindennapoknak, hogy a gazdasági, társadalmi folyamatok elképzelhetetlenek lennének nélküle.

Sajnálatos módon azonban a fejlődés velejárójaként megjelentek nem kívánatos forgalmi szituációk. Így, az autópályán kialakuló forgalmi helyzetek dinamikájának vizsgálata kutatási területként már régóta jelen van a tudományos életben [2]. A forgalmi dugók egy speciális válfaja a fantom forgalmi dugó. A fantom forgalmi dugó esetében nincs látható oka a torlódás kialakulásnak, nincs baleset, nincs semmi forgalmat akadályozó tényező.

Az ilyen helyzet kialakulásának oka alapvetően a túl sűrű forgalomban gyökeredzhet többek között. Kellően sűrű forgalom esetén az emberi vezetési tulajdonságok könnyen okozhatnak ilyen jelenséget. Például egy erőteljes fékezés képes olyan járművezető reakciókat elindítani, amely egy öngerjesztő fékezés hullámként végighalad a kocsisoron. Ennek eredményeképpen alakulhat ki a forgalmi dugó [3].

A munkánk célja, egy olyan képfeldolgozó algoritmus megalkotása, mely alkalmas a forgalom egyes résztvevőit leíró dinamikai egyenletek paramétereinek identifikációjára. Olyan eszközre van szükség, mely segítségével a járművek sebessége minden mintavételezett időpontban megbízhatóan meghatározható és használatával vizsgálható a közlekedő személyek egymásra gyakorolt hatása. A járművek neurális háló segítségével történő detektálására természetesen léteznek már megoldások [4]. Összességében azonban elmondható, hogy ezen megoldások számítási igénye nagyon magas, és a tapasztalatok alapján a megbízhatósága a jelen esetben támasztott követelményeknek korlátozottan felel meg [5].

ALGORITMUS

A képfeldolgozási algoritmus célja, hogy egy forgalomról felülről és bizonyos szintig oldalról készített felvételen képes legyen a haladó járművek detektálására, időben történő követésére és azok sebességeinek meghatározására. A járművekről gyűjtött adatokat tárolva és felhasználva alkalmas legyen a különböző forgalmi viszonyok vizsgálatát elősegítő diagramok és ábrák generálására.

1.1. Felvételek előkészítése

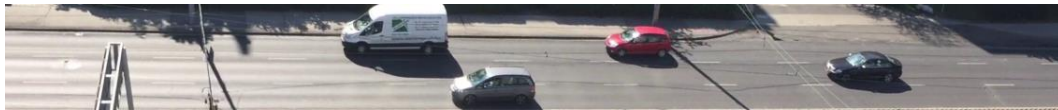
A felvételek elkészítése során nincs lehetőség kizárólag a vizsgálati tartomány felvételére, illetve a járművek mozgásának fő iránya is a legtöbb esetben egy általános helyzetű egyenes mentén fekszik. A könnyebb kezelhetőség és az alacsonyabb számításgigényű vizsgálatok érdekében a videófelvétel elforgatása és kivágása szükséges annak érdekében, hogy a járművek vízszintes irányba mozogjanak és a képmező ne tartalmazzon a vizsgálat szempontjából felesleges részeket.

Az 1. ábrán látható egy saját felvétel, amely Budapesten készült a Petőfi hídra felhajtó járművekről. A felvétel a Kármán Tódor kollégium felső emeletéről készült. Látható, hogy két ellentétes irányba mozgó forgalom van jelen. A két irány közül egyidejűleg csak egy vizsgálat a cél. Ennek érdekében első lépésként a kép elforgatása következik, mely a bezárt szög mértékével és a megjelölt irányban történik.



1. ábra. Nyers képkocka a felvételtől elforgatás előtt és után

Az előkészítés második lépése a releváns térrész kivágása, mely eredménye a 2. ábrán látható.



2. ábra. Releváns képrész metszése után kapott képrészlet

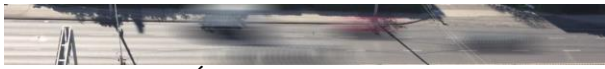
Ezt az előkészítési sorrendet alkalmazva elkerülhető az érdektelen képpontok feldolgozása. Ezen műveleteket minden képkocka esetén el kell végezni legelső lépésként. További műveleteket csak ezt követően alkalmazhatók.

1.2. Háttér szeparáció

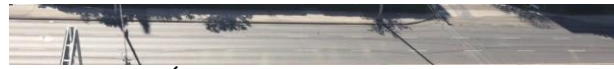
A releváns képrész meghatározása után a következő lépés a háttér szeparálása a mozgó elemektől. A program működésének alapját ez képezi. A háttér szeparálása történhet többféleképpen. A legegyszerűbb módszer egy olyan képkocka keresése, melyen nem helyezkedik el egy jármű sem. Ezen módszer azonban több nehézségbe ütközhet. A választott képkockán megjelenhet egyéb lassan mozgó objektum, például gyalogos, mely a továbbiakban hamis mozgásérzékeléshez vezethet. Ezen felül nem kellően stabil kamerarögzítés esetén pillanatnyi kilengésből kiválasztott háttér az érzékelt zónákat nagy mértékben torzíthatja, illetve egyes esetekben teljesen ellehetetleníti a felvétel kiértékelését. A legjelentősebb probléma azonban ezzel a módszerrel az, hogy nem mindig lehetséges ennek az alkalmazása. Egy nagy terheltségű úton jó forgalomszervezés mellett nem alakul ki olyan holtidő amikor nem tartózkodik egy gépjármű sem a képmezőben.

A beolvasott kép egy mátrixban helyezkedik el, amely minden eleme a kép egy pixelét reprezentálja. A mátrix egy ilyen részegysége egy RGB típusú elemet tartalmaz, mely a három (vörös, zöld, kék) színcsatornát tartalmazza egyenként legalább 0 és legfeljebb 1 értéket hordozva. Ezen reprezentációnak köszönhetően a színcsatornákon minden pixel esetében elvégezhető matematikai műveletek, jelen esetben az értékek átlagolása. Ezen módszer igen jó megoldást nyújt abban az esetben, ha a videófelvétel elég hosszú, és a járművek alapvetően nem araszolnak a felvételen. Az eljárás során a program képkockáinként végighalad a videófelvételen és elvégzi a megfelelő műveletet. Az eljárás során minden cikluslépés az addig átlagolt képpel tér vissza. Az átlagolt képet és a következő képkockát visszatáplálva és képpontokként az átlagban elfoglalt megfelelő súlyértékével megszorozva adódik az eredmény.

A 3. ábrán és 4. ábrán az ily módon kapott háttér eredménye látható a 30. ciklus után és a teljes videóra nézve.



3. ábra. Átlagolt háttér a 30. ciklus után



4. ábra. Átlagolt háttér a teljes videóra nézve

Ezen módszer sokoldalúan alkalmazható és abban az esetben is működik, ha nem létezik olyan képkocka, amely nem tartalmaz járműveket, illetve az egyéb lassan mozgó objektumok is leválaszthatók.

1.3. Perspektivikus torzítás kompenzálása

A forgatási transzformáció segítségével a járművek haladási iránya vízszintes irányba hozható (1. ábra). Azonban a perspektivikus torzításból eredően azonos pixeltávolságok a kép különböző részein más valós távolságot fognak jelenteni. Hiszen a távolabbi jármű kisebb méretűnek fog mutatkozni, mint a közelebbi. A perspektivikus torzítás jelensége többsávos utak esetében is fellép, hiszen a távolabbi sávban haladó jármű szintén kisebbnek fog mutatkozni a közelebbi sávban haladóhoz képest, tekintve, hogy a felvételek nem készíthetők el a járművek felett, csak bizonyos szinten oldalról. A pontos sebességméréshez ennek a kompenzációját el kell végeznünk.

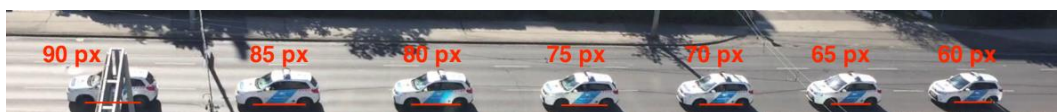
A perspektivikus torzítás kompenzálásának céljából létrehozható egy kompenzációs mátrix. A mátrix mérete megegyezik a videóból kiolvasott képek méretével. Tehát ezen mátrix minden képponthez hozzárendeli az azon pontban érvényes pixelsűrűséget, vagyis azt a számot, hogy hány valós méternek fog megfelelni egy pixel. Élve a feltételezéssel, hogy a kamera által okozott képmezőelhajlás elhanyagolható, illetve a felvett terület teljesen sík, referenciapontok felvételével illeszthető egy sík, mely kijelöli minden pixelpozícióban az ott érvényes pixelsűrűség értéket. Referenciapontoknak olyan képmezőbeli pontok választandók, melyek esetében a pixelsűrűség értéke ismert. A meghatározandó sík illesztése ezen pontok segítségével történik, tehát ez egy közelítő eljárás, mely során a legkisebb négyzetek értelmében vett legjobb illeszkedés elérése a cél.

A keresett síkot az együtthatók illesztésének a segítségével

$$f(x, y) = A_s x + B_s y + C_s \quad (1)$$

alakban keressük. Ez a síkegyenletet folytonos és nem mátrix alakú. Cél, hogy az érintett pixelsűrűség értéke minden pontban gyorsan kiolvasható legyen, ezért diszkretizáció szükséges. Minden pixelpontban a vizsgálati területnek megfelelő méretű tartományban így eltárolandó az egyenletből kapott érték.

A korábban említett referenciapontok felvételéhez szükséges két olyan pont, mely közötti távolság stabilan mérhető. A legtöbb felvétel esetében ez ismert járműtípusok tengelytávolsága lehet. Sávonként körülbelül 4-5 referenciapont felvétele mellett történt az (1) egyenlet által leírt lineáris közelítő összefüggés paramétereinek meghatározása. Az algoritmus a legkisebb négyzetek módszere szerint keresi az illeszkedő síkot. Az 5. ábrán látható egy szerkesztett kép, melyen a piros vonalak a mérendő távolságokat jelölik, felette pedig a mért érték minden esetben. A képen jól felismerhető a rendőrautó, mely típusa Suzuki Vitara (2017-es gyártmány). Tengelytávolságának ismeretében a referenciapontok felvehetők. Jelen esetben ez az érték 2,5 méter.

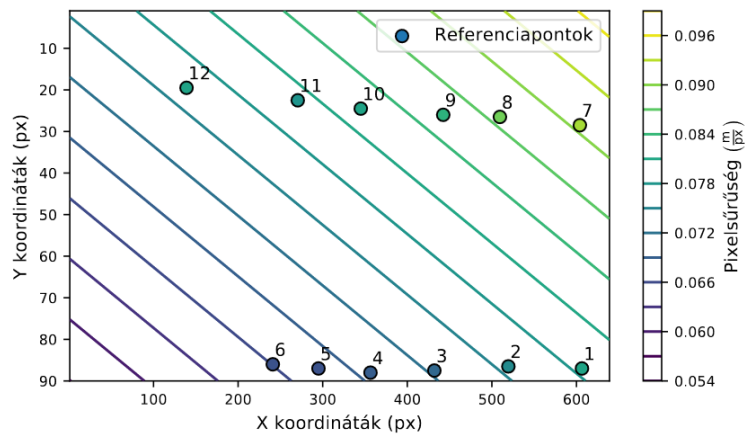


5. ábra. Módosított kép a referenciapontok felvételéhez

Egy referenciapont felvételéhez meghatározandó a két kerék középpontja. Ezen értékeket egy előre megírt függvénybe betáplálva megkapható a megfelelő pixelsűrűség a

$$\frac{\text{tengelytáv}}{\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}} \quad (2)$$

összefüggés alapján. A pixelsűrűség a tárolási ponttal összekötve rögzítésre kerül. A tárolási pont a két kerékközéppontot összekötő szakasz felezőpontja lesz. A referenciapontokat felvéve és a síkot illesztve grafikusán megjelenítve a pontokkal vizsgálható a 6. ábrán az eredmény. A cél, hogy a referenciapontok színe és a sík szintvonalainak a színe jól illeszkedjenek egymáshoz, hisz a szín jelzi a sűrűségértéket.

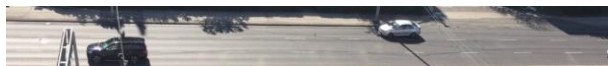


6. ábra. A kompenzációs sík szintvonalai és az illesztéshez használt referenciapontok (körökkel, pixelsűrűség szerint színezve)

Az egyenlet meghatározása után a térbeli diszkretizáció marad csak hátra. Az említett kompenzációs mátrix megkapható végighaladva a vizsgálati tartomány pixelkoordinátáin, minden pontjában eltávolítva az adott koordinátákhoz tartozó függvényértéket a síkegyenlet alapján. Így a mátrix segítségével bármely pontban meghatározható, hogy ott egy pixel megtétele hány méternek felel meg a valóságban

1.4. Járművek detektálása és követése

A vizsgálati tartomány kijelölése, a háttér meghatározása és a torzítás kompenzációjának előkészítése után megkezdődhet a járművek felismerése a képkockákon, illetve az időben történő követése, mely a képkockák közötti egymáshoz rendelését jelenti. A képfeldolgozás részleteit a 7. ábrán mutatott képkockán kísérjük végig.



7. ábra. Beolvasott képkocka a transzformációk után



8. ábra. A háttérrel különbség képzett kép

A járművek felismerése az egyes képkockák és a háttér kivonásán alapul. Ideális esetben ez azt fogja eredményezni, hogy bármilyen zérustól különböző értéket a kép kizárólag azon pontokban vesz fel, ahol a háttértől különböző pixel jelenik meg, tehát minden olyan objektum esetében, amely a háttértől különböző. A képek RGB szintérben vannak tárolva, tehát a művelet elvégzése mind a 3 színcsatornán egyesével szükséges. A kódolás során ezen értékek mind 0 és 1 között szerepelnek alapértelmezetten. A kódolásból eredően előfordulhat az az eset, hogy a különbség során negatív szám adódik. Például egy fehér jármű színekódja (1, 1, 1) közeli értéket vesz fel, míg a sötét háttéré (0, 0, 0) közelit. A kettő különbsége minden színcsatorna esetén negatív számot fog eredményezni. A 7. ábrából kivonásra kerül a 4. ábra. Ennek eredménye látható a 8. ábrán.

Szemmel láthatóan a járművek az emberi szem számára csak nehezebben felismerhetővé váltak. Egyedül az ezüst színű jármű az, amelyik a továbbiakban is jól látható szabad szemmel. Ez annak köszönhető, hogy a negatív számokat a megjelenítés során feketének veszi alapvetően a rendszer.

A szintér egy elemét vektorként kezelve, definiálható rá a vektorhossz, mely arányosan változik a háttértől való eltérés mértékével. Viszont ebben az esetben a három színcsatornából mindösszesen egy adódik, mely szürkeárnyalatos képként vizualizálható. Egy képpont esetében a konverziót a

$$\sqrt{R^2 + G^2 + B^2} \quad (3)$$

összefüggéssel adhatjuk meg, ahol az RGB jelölések a megfelelő színcsatorna értékeit jelentik.



9. ábra. Szürkeárnyalatos kép a (3) egyenlet alkalmazása után



10. ábra. A szürkeárnyalatos kép Gauss-elmosás alkalmazása után

A 9. ábrán látható, hogy a járművek ezzel a módszerrel a későbbi adatok kinyerésének céljából megfelelően azonosíthatók. Mindösszesen pár zajos részt leszámítva, csak a releváns részek kerültek kiemelésre. Látható, hogy nem minden esetben egybefüggő az érzékelt tartomány. Annak érdekében, hogy egy-egy jármű ne szeparálódjon több érzékelt területre Gauss-elmosást alkalmaz a program, ld. 10. ábra.

A Gauss-elmosás során kapott képet tovább alakítva a cél, hogy diszkrét 0 és 1 értékeket tartalmazzon a kép, tehát tisztán fekete-fehérré alakuljon. Ennek érdekében végighaladva pixelenként a képen, egy előre beállított küszöbérték felett a pixel értékét egyre, míg alatta nullára állítva megkapható az érintett kép, ld. 11. ábra.



11. ábra. A detektált járművekhez tartozó pixelek

A kivonás után kapott képpel elvégzett műveletek során két fontos paraméter van, melyet minden alkalommal a felvételnek megfelelően szükséges megválasztani. A Gauss-elmosás értékét és az érzékelési küszöbértéket.

A mozgó objektumokat sikerült elkülöníteni, illetve legenerálni a fekete-fehér képet. A képen szereplő fehér egybefüggő pixelcsoportok egy járműhöz tartoznak. Az egyes pixelcsoportok koordinátáinkénti maximális és minimális értéke meghatározható, ezen koordinátaértékekből pedig előállítható egy befoglaló téglalap minden jármű esetén. Nem minden pixelcsoport tekinthető teljes járműnek. A 11. ábra jobb alsó sarkában látható fehér folt mindösszesen a jármű első lökhárítóját tartalmazza. Sok hasonló objektum detektálható, amelyek valóban mozognak, de irrelevánsak, ilyen például a gyalogos, a képre berepülő madár és egyéb apró mozgó objektumok. Hasonlóképpen jelenhetnek meg zajos foltok is a képeken, ezeknek az eltávolítása sem cél. Összességében elmondható, hogy a detektálás szempontjából releváns és irreleváns objektumok méretben jelentős eltérést mutatnak. Ennek következtében jól meghúzható egy méretbeli határ, mely felett relevánsnak tekinthető egy objektum, és mely alatt pedig nem. Tehát elkülöníthetők a releváns objektumok kiválasztva csak azon ponthalmazokat, melyek legalább az előre meghatározott darabszámú pixelből épülnek fel.

A 12. ábrán kék téglalapokkal körberajzolva láthatók a szűrés után járműnek minősített objektumok. Jól látható, hogy ezen szűrési módszer eltávolította az éppen képbe belépő járművet, mivel annak a területe nem volt elegendő. Jelen esetben azonban csak azon járművek vizsgálata a cél, amelyek teljes egészében a képmézőben tartózkodnak.



12. ábra. Járművek megjelölése méret szerint megszűrve

A követés megvalósítása során a cél két szomszédos képkockán érzékelt befoglaló téglalap közötti hozzárendelés megteremtése. Két egymást követő képkocka látható a 13. ábrán.



13. ábra. Detektált járművek két egymást követő képkockán

A vizsgálat tárgya legyen a továbbiakban a bal oldali jármű. A 14. ábrán látható a két előbb vázolt állapot a járműnek a háttérén egy képen, kinagyítva az érintett térrészlettel.

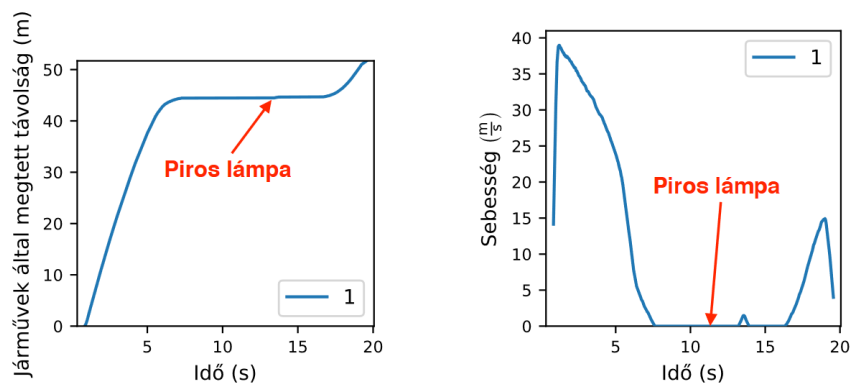


14. ábra. Pozíciók két szomszédos képkockáról a háttéren ábrázolva (képrészlet)

Amennyiben középpontjaik kölcsönösen egymás területén belül helyezkednek el, a hozzárendelés megtehető és tudjuk a járművet időben követni. Ekkor az újonnan érzékelt téglalaphoz hozzárendelhető a már előző képkockáról ismert és összeegyeztetett jármű azonosítószáma. Amennyiben nem található olyan jármű mellyel összeegyeztethető, akkor új járműként tárolható el az érzékelt terület.

A sebesség meghatározása alapvetően két mintavételezett időpont között eltelt idő és az az alatt megtett távolságon alapul. A mintavételezési idő a kamera felvételi sebességének a reciproka. A megtett távolság pedig kiszámítható a korábban meghatározott pixelsűrűségek figyelembevételével. Így a sebesség számításnál kompenzálhatjuk perspektivikus torzítás hatását.

A nyers sebességhatározás során a sebesség időben közel állandó haladási sebesség mellett is változó értéket mutat. Ennek legfőbb oka a térbeli diszkrétizáció és az érzékelési bizonytalanság együttes hatása. A sebességet ábrázoló görbék éppen ezért látványosan zajosabbak. Ennek okán szükséges valamilyen szűrő eljárás bevezetése, melyre a Savitzky-Golay szűrőt alkalmazzuk. A folyamat kimeneteként az eredmények vizualizálhatók a 15. ábrának megfelelően.



15. ábra. A 1. azonosítószámú jármű megtett távolságának és sebességének ábrázolás

KONKLÚZIÓ

A megalkotott algoritmus alkalmas a forgalmi szituációk dinamikai megfigyelésére forgalomról készített videófelvételek alapján. A vizsgálati zóna kijelöléséhez szükséges paraméterek megadása után a program önállóan alkalmas a háttér szeparációjára megfelelő videófelvételek esetén. A perspektivikus torzítás kompenzálható síkillesztési eljáráson keresztül a megfelelő referenciapontok felvételével. Az előkészítő lépések után a videó képkockáiból kivonva a kapott háttérrel elkülöníthetők a mozgó objektumok, melyek méretének és elhelyezkedésének vizsgálata után az egyes járművek detektálhatók. A program alkalmas az érzékelt járművek követésére és adatainak tárolására a felvétel során. A teljes rögzített adatsor függvényében az elmozdulásadatok szűrése és differenciálása után a sebességek meghatározhatók az egyes járművek esetében. Az adatok feldolgozása után a program alkalmas önállóan azok vizualizációjára. A tapasztalatok alapján a program legjobban felhős időben szórt fény mellett működik. Erős napsütésben az árnyékok kis mértékben nehezítik az érzékelés pontosságát. A legnagyobb kihívást esős idő jelenti, ekkor az útfelület tükröződése sok esetben ellehetetleníti a pontos mérést. Eppen ezért fontos a megfelelő időjárási körülmény megválasztása a felvételek elkészítésére. Az eredményeket sofőrmódelllel illesztéséhez, illetve új sofőrmódelllel fejlesztéséhez használtuk. A jövőben tervezzük az algoritmus továbbfejlesztését konvolúciós neurális hálózatokon alapuló képfelismeréssel.

KÖSZÖNETNYILVÁNÍTÁS

Az itt megjelent munka a Bolyai János Kutatási Ösztöndíj és a Nemzeti Kutatási, Fejlesztési és Innovációs Hivatal által megítélt NKFI-146201 pályázat segítségével támogatott. Továbbá a Doktoranduszi Kiválósági Ösztöndíj Program (DKÖP) által támogatott projekt a Kulturális és Innovációs Minisztérium Nemzeti Kutatási Fejlesztési és Innovációs Alapból nyújtott, valamint a Budapesti Műszaki és Gazdaságtudományi Egyetem közös támogatásával, a Nemzeti Kutatási, Fejlesztési és Innovációs Hivatallal kötött támogatási szerződés alapján valósult meg.

IRODALMI HIVATKOZÁSOK

- [1] Statista, Number of motor vehicles registered in the United States from 1990 to 2022, <https://www.statista.com/statistics/183505/number-of-vehicles-in-the-united-states-since-1990/> (Utolsó letöltés: 2024. 03. 30.).
- [2] Orosz G., Wilson R. E., Krauskopf B. *Global bifurcation investigation of an optimal velocity traffic model with driver reaction time* Phys. Rev. E. American Physical Society, 2004, 70(2)
- [3] Orosz G., Wilson R. E., Stépán G. Traffic jams: dynamics and control Phil. Tran. of the Royal Society A, 2010, 368(1928):4455-4479.
- [4] Jocher G.: *YOLOv5*. Szoftver, <https://github.com/ultralytics/yolov5> (Utolsó letöltés: 2024. 02. 22.).
- [5] Junsheng F.: *Vehicle Detection for Autonomous Driving*. Szoftver, <https://github.com/JunshengFu/vehicle-detection> (Utolsó letöltés: 2024. 02. 22.).