



ÓBUDAI EGYETEM
ÓBUDA UNIVERSITY

NEUMANN JÁNOS
INFORMATIKAI KAR



SZAKDOLGOZAT

**OE-NIK
2023**

Hallgató neve:
Hallgató törzskönyvi száma:

**Németh Viktor
T/006887/FI12904/N**

Óbudai Egyetem
Neumann János Informatikai Kar
Szoftvertervezés és -fejlesztés Intézet

SZAKDOLGOZAT FELADATLAP

Hallgató neve: **Németh Viktor**
Törzskönyvi száma: T/006887/F112904/N
Neptun kódja: 

A dolgozat címe:

Akadályészlelés városi kötőtpályás közlekedésben
Obstacle detection in urban light rail transportation

Intézményi konzulens: Dr. Vámosy Zoltán
Külső konzulens:

Beadási határidő: 2022. december 15.

A záróvizsga tárgyai: Számítógép architektúrák
Szoftvertervezés és -fejlesztés
specializáció

A feladat:

Tervezzen meg és készítsen el egy olyan rendszert, amely egy villamos elejére rögzített kamera színes képét feldolgozva képes valós időben meghatározni a villamos útvonalát és a körülötte lévő objektumok pozícióját az adott időjárási és fényviszonyok tükrében. A rendszer detektálja, hogy mely objektumok képeznek akadályt a villamos haladása során és ezen objektumokat emelje ki oly módon, hogy értékelhetők legyenek. Ennek érdekében vizsgálja meg a szakirodalom idevonatkozó részeit. A levont következtetések alapján határozza meg a rendszer felépítését, fő elemeit, mely technikákat fogja használni a kivitelezéshez. Megoldását előre rögzített videókkal vizsgálja, eredményeit részletesen mutassa be és értékelje.

A dolgozatnak tartalmaznia kell:

- a feladat leírását,
- releváns fejlesztések ismertetését, valamint azok értékelését,
- a megvalósítandó rendszer tervét,
- tesztelési, mérési adatokat,
- az eredmények értékelését, más megvalósításokkal való összehasonlítását,
- továbbfejlesztési lehetőségek ismertetését.



Vámosy Zoltán

Dr. Vámosy Zoltán
intézetigazgató

A szakdolgozat elévülésének határideje: **2024. december 15.**
(OE TVSz 55.§ szerint)

A dolgozatot beadásra alkalmasnak tartom:

.....
külső konzulens

Vámosy Zoltán
.....
intézményi konzulens



HALLGATÓI NYILATKOZAT

Alulírott hallgató kijelentem, hogy a szakdolgozat saját munkám eredménye, a felhasznált szakirodalmat és eszközöket azonosíthatóan közöltem. Az elkészült szakdolgozatomban található eredményeket az egyetem és a feladatot kiíró intézmény saját céljára térítés nélkül felhasználhatja.

Budapest, 20. 22.12.10.

Ni. azk
hallgató aláírása



KONZULTÁCIÓS NAPLÓ

Hallgató neve:

NÉMETH VIKTOR

Neptun kód:

[REDACTED]

Tagozat:

NAPPALI

Telefon:

Levelezési cím (pl: lakcím):

[REDACTED]

Szakdolgozat / Diplomamunka¹ címe magyarul:

AKADÉLY ÉSZLELÉS VÁROSI KÖTÖTTPÁLYÁS KÖZLEKEDÉSBEN

Szakdolgozat / Diplomamunka² címe angolul:

OBSTACLE DETECTION IN URBAN LIGHT RAIL TRANSPORTATION

Intézményi konzulens:

Külső konzulens:

DR. VÁMOSSY ZOLTÁN

Kérjük, hogy az adatokat nyomtatott nagybetűkkel írja!

Alk.	Dátum	Tartalom	Aláírás
1.	2022.03.17.	RENDSZER STRUKTÚRAJÁNAK ÁTBESZÉLÉSE	Valuany Z
2.	2022.03.31.	SZAKDOLGOZAT FELÉPÍTÉSÉNEK ÁTBESZÉLÉSE	Valuany Z
3.	2022.04.07.	ALKALMAZOTT TECHNIKÁK ÁTBESZÉLÉSE	Valuany Z
4.	2022.05.05.	SZAKDOLGOZAT TARTALMÁNAK PONTOSÍTÁSA	Valuany Z

A Konzultációs naplót összesen 4 alkalommal, az egyes konzultációk alkalmával kell láttamoztatni bármelyik konzulenssel.

A hallgató a Szakdolgozat I. / Szakdolgozat II. (BSc) vagy Diplomamunka 1 / Diplo-mamunka 2 / Diplomamunka 3 / Diplomamunka 4³ tantárgy követelményét teljesítette, beszámolóra / védésre⁴ bocsátható.

Budapest, 2022.05.02.

Valuany Z

intézményi konzulens

¹ Megfelelő aláhúzendó!

² Megfelelő aláhúzendó!

³ Megfelelő aláhúzendó!

⁴ Megfelelő aláhúzendó!



KONZULTÁCIÓS NAPLÓ

Hallgató neve:

NÉMETH VIKTOR

Neptun kód:

[REDACTED]

Tagozat:

NAPPALI

Telefon:

Levelezési cím (pl: lakcím):

[REDACTED]

Szakdolgozat / Diplomamunka¹ címe magyarul:

AKADÉLY ÉSZLELÉS VÁROSI KÖTÖTTPÁLYÁS KÖZLEKEDÉSBEN

Szakdolgozat / Diplomamunka² címe angolul:

OBSTACLE DETECTION IN URBAN LIGHT RAIL TRANSPORTATION

Intézményi konzulens:

DR. VÁMOSSY ZOLTÁN

Külső konzulens:

Kérjük, hogy az adatokat nyomtatott nagybetűkkel írja!

Alk.	Dátum	Tartalom	Aláírás
1.	2022.09.12.	RÖGZÍTETT EREDMÉNYEK ÁTBESZÉLÉSE	Vámoassy Z
2.	2022.10.05.	ÉRTÉKELÉSEK PONTOSÍTÁSA	Vámoassy Z
3.	2022.11.02.	SZAKDOLGOZAT FELÉPÍTÉSÉNEK PONTOSÍTÁSA	Vámoassy Z
4.	2022.11.24.	TARTALOM VÉGLEGESÍTÉSE	Vámoassy Z

A Konzultációs naplót összesen 4 alkalommal, az egyes konzultációk alkalmával kell láttamoztatni bármelyik konzulenssel.

A hallgató a Szakdolgozat I. / Szakdolgozat II. (BSc) vagy Diplomamunka 1 / Diplomamunka 2 / Diplomamunka 3 / Diplomamunka 4³ tantárgy követelményét teljesítette, beszámolóra / védésre⁴ bocsátható.

Budapest, 20.10.30

Vámoassy Z

intézményi konzulens

¹ Megfelelő aláhúzendó!

² Megfelelő aláhúzendó!

³ Megfelelő aláhúzendó!

⁴ Megfelelő aláhúzendó!

Tartalomjegyzék

1. BEVEZETÉS	1
1.1. A KÖZÚTI KÖZLEKEDÉS VESZÉLYEI ÉS KEZELÉSE	2
2. KÖRNYEZET ISMERTETÉSE	4
2.1. VÁROSI KÖTÖTTPÁLYÁS KÖZLEKEDÉS SAJÁTOSSÁGAI	4
2.2. A VILLAMOS ÉS HALADÁSÁT GÁTLÓ TÉNYEZŐK	8
3. HASONLÓ KUTATÁSOK ÉS FEJLESZTÉSEK	10
3.1. A PROBLÉMA MEGKÖZELÍTÉSE	10
3.2. HASONLÓ RENDSZEREK ISMERTETÉSE	11
3.3. RENDSZEREK ÉRTÉKELÉSE	15
4. RENDSZERTERVEZÉS	20
4.1. MODULOK ÉS KÖVETELMÉNYEK	20
4.2. VASÚTI SÍNPÁROK AZONOSÍTÁSA	21
4.2.1. ELŐISMERETEK FELHASZNÁLÁSA	21
4.2.2. PÁLYAAZONOSÍTÁS TECHNIKÁJA	24
4.3. OBJEKTUMOK DETEKTÁLÁSA	30
4.3.1. KONVOLÚCIÓS NEURÁLIS HÁLÓZATOK	31
4.4. AKADÁLYÉSZLELÉS	34
4.4.1. ÉSZLELÉSI FOLYAMAT	36
4.5. RENDSZER ARCHITEKTÚRA	40
5. EREDMÉNYEK ÉS ÉRTÉKELÉSÜK	44
5.1. PÁLYADETEKTÁLÁSI EREDMÉNYEK	44
5.2. AKADÁLYÉSZLELÉSI EREDMÉNYEK	52
5.3. VÉGREHAJTÁSI SEBESSÉG EREDMÉNYEK	58
5.4. TOVÁBBFEJLESZTÉSI LEHETŐSÉGEK	61
6. ÖSSZEFOGLALÁS	64
7. CONCLUSION	66
IRODALOMJEGYZÉK	67
ÁBRAJEGYZÉK	70

1. BEVEZETÉS

Fővárosunk tömegközlekedésének szerves részét képezik a villamosok, mely utasainak biztonságát szakképzett járművezetők biztosítják. Ám vannak olyan veszélyhelyzetek, melyekre nem lehet felkészülni és az ember néha kevés ahhoz, hogy meg tudjon akadályozni egy balesetet.¹ Dolgozatomban egy olyan támogató rendszer megvalósítását mutatom be, mely a közúti vasúti közlekedést célzott biztonságosabbá tenni. A rendszerrel szembeni átfogó követelmény, hogy a veszélyhelyzeteket képes legyen akár a járművezető előtt is felismerni és jelezni mindezt valós időben. A megvalósított rendszert a legkülönbözőbb körülmények között teszteltem, számos felvételen végeztem kiértékelést, továbbá összehasonlítást a szakirodalomban fellelhető hasonló megvalósításokkal. Az eredmények rávilágítanak, hogy egy hagyományos színes kamera segítségével is elvégezhető a feladat, különböző célszenzorok használatát mellőzve.

Fontosnak tartom, hogy mindenekelőtt ismertessem a közúti közlekedés kockázatait, a járművezetőre rótt terheket, illetve az ezekből fakadó veszélyeket. Az előismeretek segítenek a téma kifejtése során megérteni, hogy miért van szükség a tömegközlekedés támogatására gépi eszközökkel. A továbbiakban ismertetem a probléma környezetét, mely a városi közlekedésben nyilvánul meg, illetve a városi kötőtpályás közlekedés speciális tulajdonságait (2. fejezet). Ezen jellemzők tükrében felvázolásra kerül milyen kihívásokkal kell szembenéznie a járművezetőnek, melyek megalapozzák a megvalósításra kerülő rendszerrel szemben támasztott követelményeket.

Rendszertervezés során figyelembe kell venni a hasonló kutatásokat, fejlesztéseket (3. fejezet), illetve a szakirodalom nyújtotta lehetőségeket. Bemutatásra kerül számos kutatás és fejlesztés, melyek hozzájárultak a rendszer megvalósításához.

A továbbiakban ismertetem a rendszer részletes vázlatát (4. fejezet), milyen modulok határozzák meg a rendszer működését és milyen követelményeket állítunk szembe velük. Bemutatásra kerül, milyen technikákkal hajtjuk végre a pálya- és objektum detektálást, illetve hogyan valósul meg az akadályészlelés.

A megvalósított rendszer eredményeit (5. fejezet) pályadetektálási, akadályészlelési és rendszerteljesítményi szempontból ismertetem, összevetem hasonló rendszerek eredményeivel és továbbfejlesztési javaslatokat teszek.

¹ Az állítás személyes szakmai tapasztalaton alapszik

1.1. A KÖZÚTI KÖZLEKEDÉS VESZÉLYEI ÉS KEZELÉSE

Az évszázad kezdetétől számítva hazánkban, a meglévőkön túl több millió közúti gépjármű állt forgalomba [1], melyek fokozott veszélyt jelentenek nem csak a közlekedésben résztvevő emberekre, de a környezetünkre is egyaránt. A megnövekedett közúti gépjárművek számával arányosan a közúti balesetek száma is jelentős mértékben emelkedett [2]. Azon személyek, akik a tömegközlekedésben részt vesznek, mint utasok, kifejezetten nagyobb veszélynek vannak kitéve. Utazáskor ki vannak szolgáltatva a körülményeknek, az adott tömegközlekedési eszközt kezelő személy döntéseinek. A tömegközlekedési eszközöket kezelő személyek magas szintű képzéseken vesznek részt, hogy a veszélyes helyzeteket a lehető legjobban tudják kezelni, ezzel együtt életeket mentsenek. Sajnos vannak helyzetek, amikor nem elég a rutin, mert az ember nem biztos, hogy időben meg tudja hozni a helyes döntést².

A közúti közlekedésben folyton jelen van az emberi tényező³, mely befolyásolja a közlekedés eseményeit. Ezen események egyike akár egy szerencsétlen helyzetben balesetként is megnyilvánulhat. A baleset bekövetkezését előre megjósolni szinte lehetetlen, melyhez leginkább hozzátesz az emberi tényező kiszámíthatatlansága. Annak érdekében, hogy a balesetek számát valamilyen formában csökkenteni tudják, nem csak az infrastruktúrán hajtottak végre módosításokat, fejlesztéseket (mint például veszélyes helyre felhívó útfelfestések, jelzőtáblák), hanem a közlekedési eszközökbe is építettek számos biztonsági berendezést. Ezen berendezéseket két kategóriába lehet sorolni. Az elsőbe soroljuk az aktív eszközöket, melyek célja, hogy a veszélyes közlekedési helyzetek kialakulását csökkentsék. Ilyen eszközök például a blokkolásgátló⁴ vagy a csúszásvédelem⁵. A második kategóriába soroljuk a passzív eszközöket, melyek a baleset bekövetkezését követően próbálják csökkenteni a károkat, személyi sérülést. Ide soroljuk a légzsákokat vagy a biztonsági övet. Az aktív és passzív eszközök együttesen nyújtanak védelmet balesetekkel szemben. A biztonsági eszközök hatékonysága nagymértékben függ, hogy milyen helyzetben léptek működésbe (például a blokkolásgátlónak milyen útfelületen kell teljesítenie), illetve a jármű tulajdonságaitól.

A közlekedésben résztvevő járművek számos tulajdonságban eltérnek egymástól, melyek veszélyességi faktor aránya nagyban differenciált egyes járműfajták és járműtípusok között. A tömegközlekedésben résztvevő járművek fő

² Az állítás személyes tapasztalaton alapszik.

³ Adott helyzetben résztvevő személy hozhat rossz döntést, mely negatív hatással lehet az esemény kimenetére.

⁴ Anti-lock Braking System (ABS)

⁵ Vasúti közlekedésben például a sínfék

tulajdonsága, hogy egy átlagos személygépjárműnél nagyobb mérettel és tömeggel rendelkeznek, melyek miatt nehezebben kontrollálhatók. Egy nem megfelelő időben megkezdett fékezésnek akár súlyos következménye is lehet. Éppen ezért a tömegközlekedésben résztvevő járműveket szakképzett kezelő személyzettel látják el a közlekedési vállalatok.

A járművezetőnek kellő figyelemmel kell defenzív módon közlekednie, mely során előtérbe kell helyezni a biztonságot a menetrend tartásával szemben. A biztonsághoz elengedhetetlen a kellően kipihent elme, mely kellő módon fenn tudja tartani a figyelmet. A szakképzett járművezetők fokozottan ki vannak téve a közúti forgalom okozta feszültségnek, ami intenzív mentális terhet ró az elmére. Továbbá a hosszú munkaórákból adódóan fizikailag és szellemileg is kimerülnek. A fáradt elme nagyobb baleseti kockázattal jár, ami ellen az aktív biztonsági eszközök sem képesek megfelelő védelmet nyújtani. Személy szerint a tömegközlekedés közúti vasúti ágazatában voltam munkavállaló, és munkám során számtalan olyan szituációban találtam magam, ahol a járműbe épített aktív biztonsági eszköz jelentősen csökkentette a baleset bekövetkezésének esélyét. A menetrendtartás és a defenzív vezetés együtt nem tud működni. A forgalomból adódóan nagyon könnyen össze lehet szedni késést, ami később nagyobb gondot tud okozni, ezért a járművezető igyekszik elkerülni őket. Azonban, hogy a menetrendet tudja tartani, szükséges a forgalom tempójával haladni, amely fokozott veszélyt jelent a jármű hosszabb fékútja és kötött haladása miatt.

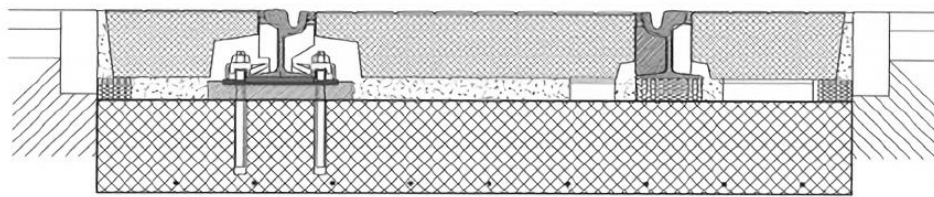
A járművezetőkre rótt intenzív teher enyhítése végett érzett motivációval terveztem meg egy olyan aktív, automata támogató rendszert, ami képes felismerni akár a járművezető előtt adott forgalmi helyzeteket, ahol a jármű haladását akadályozó tényező lép fel. A folyamat során egy kamera színes képét használok fel, mely a jármű elejére van rögzítve, a haladási iránynak megfelelően. A rendszer meghatározza a haladás útvonalát és a környezetében fellépő releváns objektumokat, melyek akadályt képezhetnek.

2. KÖRNYEZET ISMERTETÉSE

Hazánk egyik legmeghatározóbb tömegközlekedési módja a vasút. A mai modern értelemben vett vasút közel kétszáz évvel ezelőtt kezdte meg működését hazánkban. A 16. századbeli bányákban használt csille rendszerek adták alapját a napjainkban is jól ismert nyompályán közlekedő járműveknek. Az eredetileg fából készült nyompályát felváltotta az acél vágányrendszer, mely alkalmas az akár több száz tonnás szerelvények vezetésére. Az eredetileg lóval vontatott vasúti eszközöket az első ipari forradalomban felváltották a gőzmozdonyok, melyek képesek voltak több vasúti kocsiból álló szerelvényeket mozgatni. A technológia fejlődésével a gőzmozdonyok elérték teljesítményük határát, mely vezetett a villamos árammal történő vontatás megvalósításához. A technológia szükségessé tette a meglévő vasúti hálózatok villamosítását, melyeken napjainkig korszerű, nagy hatásfokkal üzemelő villamos motorkocsikat üzemeltetnek. A vasúti hálózatok nagyvárosokat, mellékvonalai kisebb településeket kötnek össze, melyek a közúttól elszeparálva üzemelnek. A vasúti közlekedés egy alfajtája, a közúti vasúti közlekedés, mely lényegében a városokon belül, a közútba épített sínpályán való közlekedést jelenti. Ebben a fejezetben a városi kötöttpályás közlekedés sajátosságait fogom ismertetni. A fejezet Gerlang Tamás, „Általános szakmai ismeretek I.” [3] jegyzete alapján készült.

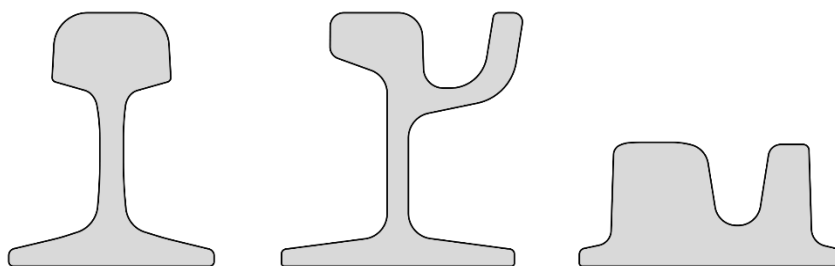
2.1. VÁROSI KÖTÖTTPÁLYÁS KÖZLEKEDÉS SAJÁTOSÁGAI

A városi kötöttpályás közlekedés, vagy másnéven közúti vasút, számos nagyvárosban fellelhető. Tulajdonságaiban és alapjaiban tekintve joggal hasonlíthatjuk a klasszikus nagyvasúthoz, viszont a működési környezetükben eltérést találunk. A közúti vasutak egyik szembevető tulajdonsága, hogy a közúti forgalom szerves részét képezik. Éppen ezért a KRESZ számos szabálya érvényes, mind a közúti vasúti járművekre, mind a hagyományos közúti gépjárművekre egyaránt. A közlekedés a közúti jelzéseknek megfelelően történik, mely alapján véve egy nagy eltérést jelent a nagyvasúthoz képest. Mivel a közúti vasúti gépjármű a közúti forgalomban vesz részt, és a KRESZ szabályai szerint közlekedik, ezért nem alkalmaznak (vagy csak részben) a nagyvasúthoz hasonló biztosítóberendezéseket (például főjelzőket). Hasonló módon a személygépjárművel történő közlekedéskor, a közlekedés látra történik, tehát a járművezetőnek kell meggyőződnie arról, hogy a bejárni kívánt pályaszakasz szabad-e. Mind a biztosítóberendezések hiánya, mind a látra történő közlekedés miatt kifejezetten veszélyes üzemnek kell tekinteni a közúti vasúti közlekedést. A továbbiakban a közlekedés alapjául szolgáló sínpályát ismertetem.



1. ábra. Villamospálya keresztmetszete [4]

Nagyvasúthoz hasonló módon a közúti vasút járművei előre kiépített fém pályán (jellemzően acél) közlekednek. A pálya rendeltetése a rajta haladó járművek alátámasztása és vezetése. A hagyományos nagyvasúti értelemben vett alátámasztás nem állja meg a helyét, ugyanis a pálya kialakítása során a pálya általában az útburkolat részét képezi, hogy a közúti gépjárművek is képesek legyenek a pályán közlekedni (1. ábra). Ezen pályaszakaszok leginkább egy meglévő úthálózatra lettek utólag építve, melyek alapzata nem képes több száz tonnányi terhet átvinni alakváltozás nélkül. A nagyvasútban használt vasúti pályát két fő részre lehet bontani; az alépítményre és felépítményre. A felépítmény célja, hogy a rajta elhelyezkedő vasúti pályán haladó jármű terhet átvegye és ezen terheket egyenletesen ossza el az alépítmény felé úgy, hogy az ne okozzon maradandó deformálódást az alépítményben. A közúti vasútnál ilyen jellegű pálya kiépítésére nincs mód, ezért nem értelmezzük az alátámasztást hagyományos értelemben. Az alátámasztást gyengébb felépítmény végzi, mely teherbírása jóval kisebb a nagyvasúthoz képest. Ezért a rajta haladó járművek tengelyterhelésére létezik egy korlát, mely jóval alacsonyabb a vasúti szerelvényekhez képest.



2. ábra. Balról jobbra: Vignol, Phoenix és Tömb sínprofil

A felépítményen helyezkedik el a jármű haladását biztosító szerkezet, a vágány, melyet két fő részből áll; a sínekből és az aljakból. A sínek közvetlen kapcsolatban állnak a szerelvény vaskerekeivel, vezetik a szerelvényt és átveszik a terhet. Többféle kialakítású sínszálak léteznek. Általánosságban úgynevezett Vignol szélestalpú

kialakítást használják, de a közúti vasútnál leginkább a tömbsín és a Phoenix kialakítást részesítik előnyben (2. ábra). A Vignol sínszál profilja egy „I” betűhöz hasonlítható, melyeket szorítóelemekkel rögzítenek az aljakhoz. A profil „szára” kifejezetten alkalmas a sínkoronára kifejtett teher rugalmas közvetítésére, melyet a profil alsó része közvetít a talaj felé. A Phoenix sínprofi a Vignol-hoz hasonló, ám a sínkorona egy nyomcsatornával van ellátva. A nyomcsatorna biztosítja, hogy a közúti vasúti közlekedés során az útburkolatba építéskor a vasúti kerékpár nyomkarimájának legyen biztosítva megvezetés. A tömb sín nagyban hasonlít a Phoenix sínkoronájához, viszont a sínszál teljes egészében, a nevéből is adódó tömb alakzatot formál meg szár hiányában. A sínszálak alátámasztására szolgálnak az aljak, melyek biztosítják a nyomtávtartást, illetve a terhek egyenletesen eloszlását a felépítmény felé. A hagyományos Vignol sínszálakat fa vagy beton aljakhoz rögzítik, melyeket a felépítmény zúzottkő ágyába helyezik. A tömb sín ezzel ellentétben nem lehet hagyományos értelemben vett aljakhoz rögzíteni, ezért őket előre legyártott betonlemezek vájzatába helyezik a sínszálakat és leszorító gumibetéteket helyeznek el a sínszál és betonlemez közé.

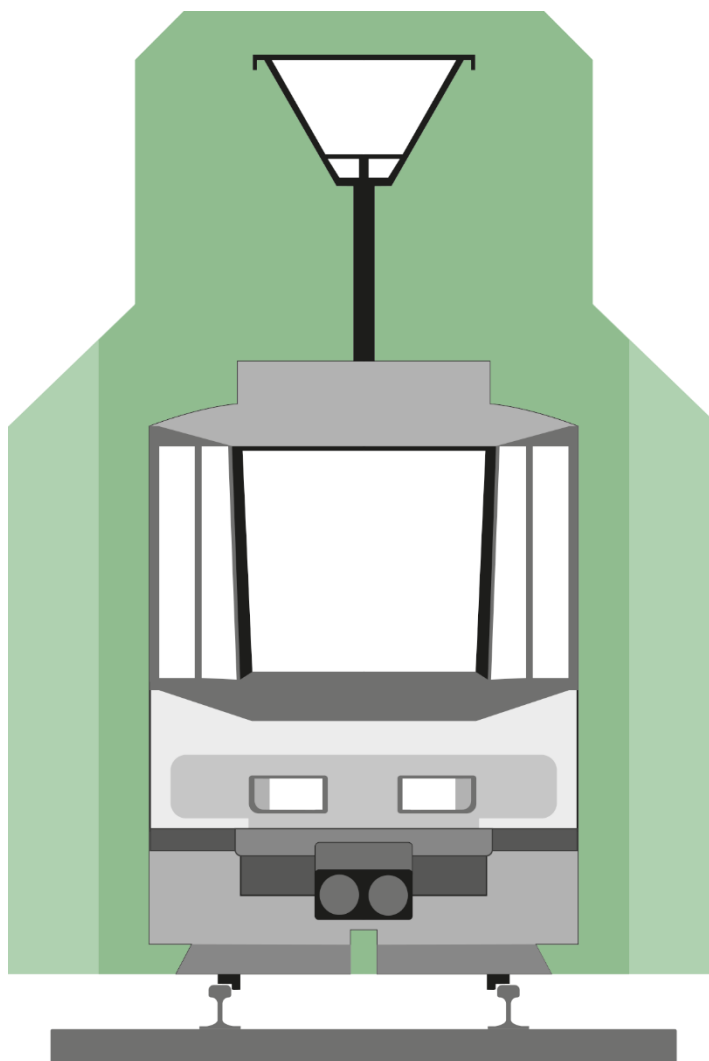


3. ábra. Balról jobbra: Betonlemezes, bazaltkőkockás és aszfalttal burkolt pályaszakasz

Közúti vasúti közlekedésre szolgáló vágányhálózatokat, az adott korszak szerint különféle útfelületbe hivatottak építeni. Fővárosunkban megtalálhatók a hagyományos aszfalttal borított útfelületeken túl a betonlemez, illetve a bazaltkőkocka⁶ burkolattal ellátott útszakaszok (3. ábra). A régi úthálózatoknál alkalmazott bazaltkőkockás burkolat napjainkig fellelhető számos vonalszakaszon, melyek kedvezőtlen időjárási viszonyok esetén kifejezetten nagy veszélyt jelentenek, mind a gyalogos, mind a közúti közlekedésre nézve. Felületük kimondottan csúszásveszélyes, ezért a korhű

⁶ Köznnyelven „macskakő”

útburkolatoknál már nem alkalmazzák az említett technikát. Betonlemez útburkolatokat a legújabb közúti forgalomra szánt pályaszakaszokon alkalmaznak, melyek gyors kiépítést és nagy teherbírást tesznek lehetővé. Általánosságban véve a legelterjedtebb a hagyományos aszfalttal öntött, Phoenix sínszálakkal ellátott vonalszakaszok. A közúti vasútban megtalálhatók a hagyományos nagyvasúti pályaszakaszok is. Azon pályaszakaszokat, melyeken a közúti közlekedéssel egy nyomvonalon halad a jármű, nyíltpályának nevezzük. Azon szakaszokat, melyeken nem zajlik közúti közlekedés, zártpályának hívjuk. Jellemzően itt lelhetők fel a nagyvasútnál is használatos biztosító berendezések (például főjelzők).



4. ábra. A vágányra merőleges űrszelvény (zöld)

Érdeemes megismerni az űrszelvényt (4. ábra), ami a vágány mellett és fölött szabadon tartandó térnek a vágányra merőleges metszete, melyen belül a jármű a legkedvezőtlenebb körülmények között is biztonságosan elfér [5]. Az űrszelvény „szabadon” tartása kritikus része a közúti vasúti közlekedésnek. Az elsodrásos balesetek többnyire az űrszelvényben történő tartózkodásból adódnak, éppen ezért az űrszelvény meghatározása kritikus pontot képez a rendszer megvalósítása során.

2.2. A VILLAMOS ÉS HALADÁSÁT GÁTLÓ TÉNYEZŐK

A villamos olyan jármű, amely az úttestbe épített, vasúti pályaként meg nem jelölt sínpályán való közlekedésre szolgál [6]. A korai járművek kialakítása nagyban hasonlított a vasúti kocsik felépítésére. Forgóváz mentes, több tonnás alvázzal bíró, laprugós felfüggesztéssel rendelkező kéttengelyes konstrukciók voltak az első járművek. Idővel megjelentek a nagyobb tömeg szállítására alkalmas forgóvázas, rugós felfüggesztéssel rendelkező típusok, melyek fő jellemzője a nagyobb teherbírás és a többféle alváz kialakítás. A forgóvázakat felhasználták egyterű csuklós villamosok építésére, melyek többnyire négy forgóvázzal rendelkeznek és az alváz szegmensei közti kapcsolatot csuklópontokkal láttak el. A szőlőkocsik lehetővé tették akár több száz ember egyidejű szállítását. Mára az alvázak kialakítások elavultnak tekinthetők az önhordó vázszerkezettel szemben, melynek lényege, hogy a terhet az egész kocsiszekrény viseli, nem pedig az alváz, ezért csökkent a jármű össztömege, javult a tengelyterhelés. Megjelenésük lehetővé tette az alacsonypadlós villamosok kialakítását, melyek akadálymentes utascserét nyújtanak az idősek és mozgássérültek számára. A jármű tetején megtalálható áramszedő(k) biztosítja a felsővezeték és vontatómotorok közti elektromos kapcsolatot. A villamosok több nagy teljesítményű elektromos vontatómotorokkal vannak ellátva, melyek egyaránt használhatók a jármű gyorsítására és lassítására is. A jármű fékezésére a korábban használatos tuskós légfékrendszert felváltotta az elektrodinamikus fékrendszer, mely a vontatómotorok generátorüzembe kapcsolt állásával érhető el. A mai modern villamosok fékezéskor képesek a nagyfeszültségű hálózatba visszatáplálni generátorüzemben termelt áram egy részét, melyet az azonos hálózati szakaszban lévő más járművek fel tudnak használni gyorsításkor. Az áram másik része a jármű tetején elhelyezkedő ellenállásokon áthaladva hő formájában vész el.

A villamos mérete jóval meghaladja a személygépjárművét. Éppen ezért a mérethez hatalmas tömeg is társul. Jellemzően az alvázak villamosok több tíz tonnát nyomnak egy hagyományos személygépjárművel szemben, mely csupán csak egy-két tonna általában. A nagy tömeghez és mérethez ezért nagyobb fékút is tartozik, mely

sebességtől függően elérheti a 100-150 métert is, relatíve háromszor nagyobb egy személygépjárműnél. Figyelembe véve, hogy a villamosok a közúti közlekedésben vesznek részt, fékút ismeretében fontos a kellő követésitávolság tartása szerelvény és gépkocsi között. Mivel a jármű kötött pályán közlekedik, ezért egy fellépő akadály során nem fog tudni tovább haladni. A vasúti sínek nem teszik lehetővé a pálya elhagyását, a tovább haladás érdekében gondoskodni kell az akadály megszüntetéséről.

Melyek azok az akadályt képző szituációk, amikre folyamatosan figyelni kell a villamossal történő közlekedéskor? Érdemes megkülönböztetni két csoportot. Az egyik csoportba soroljuk azokat az úrszelvény részét képző elemek (például sínpálya és részei, felsővezeték-hálózat) meghibásodását, melyek gátolják a továbbhaladást. Ide soroljuk példának okául a síntörést⁷, mely járművezetői szemmel is csak nagyon nehezen, közlelőre vehető észre. Az imént említett kategóriába sorolható akadályt képző jelenségek vizsgálatával nem foglalkozik a rendszer. A másik csoportba sorolandók azon helyzetek, melyek a közlekedésben résztvevő személyek/eszközök helyzetváltoztatásának okául történnek. Ilyen szituációnak kell tekinteni a hibás parkolást, amikor a jármű egy része az úrszelvénybe ér vagy akár a hibás sávtartás, mely során egy álló kocsisorban tartózkodó gépjármű egy része átlóg a saját sávján túl az úrszelvénybe. Ezen forgalmi helyzetek csúcsidőben rendszeresen jelentkeznek, kiszűrésük nem könnyű feladat. A járművezető munkáját támogató rendszer az ilyen és ehhez hasonló helyzeteket célzott felismerni és megállapítani az akadály fennállását.

⁷ Síntörésről akkor beszélünk, ha a sínszál felépítésében sérülés keletkezik és annak hatására a sínszálban anyagfolytonossági hiba lép fel.

3. HASONLÓ KUTATÁSOK ÉS FEJLESZTÉSEK

Rendszertervezés során szükséges megvizsgálni a célterületre szánt hasonló fejlesztéseket, melyek egy stabil kiindulási alapot nyújthatnak az irodalomkutatás terén. Mivel a közúti vasúti közlekedés egy nagyon specifikus terület, ezért a problémámra konkrét megoldásokat nem sikerült felfedezni publikációk terén. Érdekes más irányból megközelíteni a problémát, miszerint a közúti vasút a vasutak egyik alkategóriájába tartozik. A vasút, mint ősosztály terén viszont fellelhető számos olyan megoldás, melyek részben lefedik az eredeti probléma terét. A továbbiakban pár nagyon hasonló problémára megoldást nyújtó fejlesztést fogok ismertetni, illetve a végén értékelem őket. Felvetésre kerül, mely technikákat érdemes átvenni a tervezés során.

3.1. A PROBLÉMA MEGKÖZELÍTÉSE

A Sensors folyóiratából egy 2021-es cikk⁸ [7] összefoglal számos olyan fejlesztést, melyek az akadályészlelést hivatottak megvalósítani. A cikk kiemeli és kategorizálja a fejlesztések során alkalmazott módszereket, melyek alapján vizsgálom az akadályészlelés megvalósításának lehetőségeit. Nulladik lépésként felbontom az akadályészlelést különböző részproblémákra.

Az első ilyen részprobléma a vasúti pályának meghatározása a szenzor bemenetén. A szenzor esetünkben egy kamera, mely egy színes képet ad vissza. A perspektív nézet feldolgozásának két fő megközelítése van. Az első egy geometriai problémára következtethető vissza, mely során a vasúti pálya karakterisztikáit vizsgáljuk perspektív nézetből. A pályával kapcsolatos ismereteket felhasználva lehet tudni, hogy a sín párok közti távolság fix, előre kalkulálható a képen. Felülnézeti transzformáció során (idegen nevén „bird’s eye view”) a sín párok párhuzamosan haladnak. Kamera szemszögéből a sín párok egy távlatpontba konvergálnak. Ezen ismeretek felhasználásával meghatározható a pálya elhelyezkedése, melyet a 4.2 fejezetben ismertetek.

Egy másik megoldást nyújt az úgynevezett neurális hálózatok használata. A neurális háló képes olyan összefüggések felismerésére, melyekből meghatározható a vasúti pálya elhelyezkedése. A neurális hálózat pontossága részben függ, hogy milyen adatokkal végzik el a tanítást. Jelen esetben ezek az adatok nem mások, mint a vasúti pálya fekvését meghatározó képek. Ahhoz, hogy hatékony legyen egy ilyen rendszer,

⁸ „A Review of Vision-Based On-Board Obstacle Detection and Distance Estimation in Railways” [7]

a tanításhoz megfelelő minőségű és mennyiségű tanítómintát kell biztosítani, melyek képesek lefedni a problémateret a lehető legszélesebb körben. Az adatoknak változatosnak kell lenniük, melyek kivitelezése nehezen valósítható meg. A neurális hálózattal történő pályafelismerés ötletét elvetem a tanításhoz szükséges adatok hiányában.

Akadályészlelés során elengedhetetlen az akadályt képző, vagy képezhető objektumok azonosítása. Miután megtörtént a pálya meghatározása, az észlelés problémája leszűkíthető a pálya körüli területre. Az egyik lehetőség, hogy a pálya mentén keresek akadályokat olyan képfeldolgozási technikák segítségével, melyek a sínszálak körüli élek változását vizsgálják. Ez a technika nem terjed ki a teljes úrszelvény vizsgálatára, mely során az akadály nem feltétlen a sínszálon fekszik, hanem az úrszelvény szélén, ezért ezt a megoldást elvetem. Egy megoldást a mozgó objektumok vizsgálatára az előzményképek felhasználása nyújt. Előnye, hogy ha egy objektum egy korábbi képkockához képest elmozdul, úgy az könnyen azonosítható az aktuális és az előző képkocka különbségén. A probléma, hogy ez a technika álló kamera képén alkalmazható. A villamosra rögzített kamera nem nyújt stabil képet a pályahibák okozta mozgásoktól és kamera képének vonatkoztatási rendszeréből adódóan, mely központjában nem a környezet, hanem a villamos áll, így a kép csak a villamos álló helyzetében lesz használható, ezért a technikát szintén elvetem. A konvolúciós neurális hálózatok egy olyan adaptív megoldást nyújtanak, mely során az objektumok meghatározása a hálózat pontosságán alapul. Napjainkban számos olyan, széles körben elterjedt neurális hálózat létezik, melyek a közlekedésben résztvevő emberek és eszközök azonosítását teszik lehetővé.

A következő alfejezetben olyan megoldásokat mutatok be, melyek az akadályészlelést végző rendszer részproblémáira nyújtanak megoldásokat. A fejezet végén ezen rendszereket értékelem és kiemelem mely technikák járultak hozzá a tervezéshez.

3.2. HASONLÓ RENDSZEREK ISMERTETÉSE

Az első hasonló rendszert, a 2017-es IMECS konferencia, „*Video Analysis for Improving Transportation Safety: Obstacles and Collision Detection Applied to Railways and Roads*” [8] cikkben ismertettek, mely során gépi látással dolgoztak fel a vasúti járműre rögzített színes kamera képét. Az ötlet azon alapszik, hogy számos költséges szenzor csak rövid távon képes hatékonyan érzékelni akadályt, ezért a képi megközelítés egy stabilabb távpontot nyújthat. A rendszer megoldást ad a vasúti pályán elhelyezkedő akadályok észlelésében. Az akadály észlelés a pályaaazonosítás

során történik. A pályaaazonosítást szintén egy geometriai problémára következtetik vissza, viszont nem alkalmaztak felülnézeti transzformációt. Előismereteket felhasználva következtetéseket vontak le, hogy a sínszálak közti távolságnak konstansnak kell lenniük. Ebből következik, hogy a sínszálak perspektív nézetben egy pontba konvergálnak. A sínszálak kezdeti pontjai a kép alsó szegmensében nem mozdulhatnak el. A megvalósított rendszer az alábbi lépésekkel dolgozik; elsőnek elvégzi a képkocka előfeldolgozását, mely során Canny [9] éldetektorral meghatározza a képen szereplő éleket. Ez követően azonosítja a megfelelő sínszálak kezdő pontjait, melynek az alapját a sínszálak közti konstans távolság biztosítja. Csúszó ablakpár segítségével a kép alsó szegmensén végig haladva lokalizálják a sínszálakat alkotó éleket. Ezután iteratív módon létrehozzák a pályát vizsgáló ablakokat⁹ a kép aljától a távlatpont felé. Segítségükkel nem csak a sínparók elhelyezkedését határozzák meg, de minden egyes ablakban külön vizsgálják, hogy szerepel-e akadály. A vizsgálat alapját az ablakban szereplő élek kapcsolata adja, miszerint, ha a sínszálak mentén nem található akadály, akkor az élek az ablak alsó részétől a felső részéig megszakítás nélkül futnak. Ha az ablakban lévő intenzitásösszeg érték átlaga kevesebb, mint az adott küszöbérték, akkor azt törésnek veszik és az ablak további vizsgálatot igényel akadály gyanánt. Ha a töréspont konzisztens hossza meghaladja a sín 20%-át, akkor ott akadály észlelhető. Akadály esetén az ablakok további felépítését nem folytatják tovább, az iteráció megszakad. Ellenkező esetben az iteráció a távlatpont vizsgálatában ér véget. A vágányok közti akadályok vizsgálata során az ablakok közti szakaszokat veszik figyelembe, melyeken az előzmény képkocka segítségével a háttereket kivonják egymásból. A rendszer tesztelését követően megállapították, hogy a tesztvideón elhelyezett mind a 10 akadály azonosítása sikerrel járt. A rendszer szenzitivitása 100%, fals negatív azonosítás nem történt, ezért a rendszer biztonságosnak tekinthető. Viszont történt számos fals pozitív azonosítás is, mely során olyan objektumokat detektált a rendszer, melyek valójában nem voltak jelen a tesztkörnyezetben.

Egy hasonló megközelítés a pályadetektálásra, a 2010-es Szingapúri ICARCV¹⁰ konferencián megjelentetett „*Obstacle-Free Range Determination for Rail Track Maintenance Vehicles*” [10] cikkben olvasható. A rendszer távolságtartás biztosítására szolgál a konvojban közlekedő vasúti pályakarbantartó járművek között. A pályaaazonosítás alapját Hough-transzformáció [11] nyújtja, mely során a képen lévő éleket egy polárkoordináta-rendszerbe képzik le. Első lépésként meg kell határozni a

⁹ „Bottom-up windows”

¹⁰ International Conference on Control, Automation, Robotics and Vision (ICARCV)

vasúti pálya, kamerához közeli szegmensét egy trapézzal, melynek négy sarokpontjával számítják ki azt a transzformációs mátrixot, melynek segítségével képzik le a felülnézeti képet. Az eredményezett képen a vasúti sín párok már párhuzamosan haladnak, konstans távolsággal egymástól. Megtörténik az éldetektálás egy Canny operátor használatával. Az operátor küszöbértékeit egy algoritmus határozza meg, mely során a magas küszöbértéket addig csökkentik, amíg nem lesz elegendő mennyiségű él a képen, hogy elvégezhető legyen a Hough-transzformálás. A pálya meghatározása ezután letről felfelé történik, iteratív módon létrehozott ROI¹¹ ablakokban. Minden iteráció során azon élpárokat keresi az algoritmus, melyek az előző iterációban meghatározott sínszalak irányszögéhez hasonló iránnyal bírnak és az élek közti távolság konstans. Akadályészlelés során az élek gradiens irányát vizsgálják. Ha a gradiens irányja nem merőleges a sín párhoz tartozó élek irányára, akkor a vizsgált ROI területét akadállynak veszik.

Akadályészlelés terén egy sajátos megközelítéssel álltak elő a 2016-os IEEE ICIP¹² konferencián megjelent „*Moving camera background-subtraction for obstacle detection on railway tracks*” [12] cikkben. A megalkotott rendszer a nagyvasúti közlekedésben történő akadályészlelésre javasolt egy megoldást, mely az előzményképek felhasználásával próbálja lokalizálni az objektumokat. Ahhoz, hogy előzményképeket tudjanak felhasználni az aktuális képpel történő összehasonlításához, pixel szinten kell hozzáigazítani a referencia képet az aktuálishoz. Egy mozgó szerelvényre rögzített kamera képét fontos térben és időben is egymáshoz igazítani. A probléma megoldására a rendszer elsőnek az aktuális képhez hasonló referencia képet keres a korábban rögzített képkockák közül. A referencia képen ezután elvégzi minden pixel szerint az igazítást, mely során úgy mond ráigazítja az aktuális képre a referencia képet. Végül a két kép különbségét képzik és az akadályészlelést elvégzik a kimenet vizsgálatokor.

Az elmúlt években nagy népszerűségnek örvendenek a neurális hálózatokkal használata különféle objektumdetektáló problémára. A 2019-es ITSC konferencián bemutattak egy olyan konvolúciós neurális hálózatot, mely célja a nagyvasúti környezetben jelentkező általános akadályt képző objektumok detektálása. A „*Real-time Obstacle Detection Over Rails Using Deep Convolutional Neural Network*” [13] cikkben bemutatott rendszer a Single Shot Multibox Detection (SSD) [14] és a You Only Look Once (YOLO) [15] sikereire építve nyújt egy hasonló, ám a célterületre fókuszáló probléma megoldást. A rendszer egy viszonylag kis Dataset-en lett tanítva

¹¹ Region of interest (ROI)

¹² International Conference on Image Processing (ICIP)

és tesztelve, mely 5107 képet tartalmaz tanításra és 1277-et tesztelésre. A modellt két összetevőre lehet bontani: a fő hálózatra, illetve egy jellemzőket egyesítő összetevőre. A hálózat fő része felel a képen szereplő jellemzők kinyeréséért. A jellemzőegyesítő réteg két különféle konvolúciós kernelből áll, melyek feladata, hogy kinyerje az alacsony és magas szintű jellemzőket, majd egyesítse őket. Utolsó lépésként a nem-maximumok elnyomásával¹³ (NMS) kiválasztják a magasabb konfidenciaszinttel bíró behatároló téglalapokat. A rendszer tesztkörnyezetben kimagasló eredményeket ért el 91,61%-os pontossággal és 26,3 FPS¹⁴-el, míg a YOLO 75,67%-kal és 28 FPS-el teljesített.

¹³ Idegen kifejezéssel „Non-maximum suppression” (NMS)

¹⁴ Másodpercenkénti képkockaszám (Frames per second)

3.3. RENDSZEREK ÉRTÉKELÉSE

Korábban említett fejlesztések mindegyike valamilyen sajátos megközelítésben nyújt egyfajta megoldást a pályadetektálásra, vagy éppen az akadályok azonosítására. A pályadetektáló alrendszer felelős a vasúti sín párok azonosításáért a fényviszonyoktól függetlenül. A fényviszonyok az adott napszaknak megfelelően és az időjárási jelenségek tükrében változnak. Hóval vagy esővízzel fedett pályaszakaszok lokalizálása egy kimondottan nehéz feladat. A hasonló rendszerek ezt a problémát nem, vagy csak részben érintik. Nagyvasúti környezetben a pályát nem fedi esővíz vagy hó, a magasvágány egyértelműen kitűnik, melynek meghatározása jelentősen könnyebb az úttestbe épített pályával szemben. Havas nagyvasúti pályára a [10] cikkben látható példa. A városi környezetet nem fedik le ezen rendszerek, ezért a probléma hagyományos képfeldolgozási technikákkal nagyon nehezen oldható meg, az eltérő időjárási jelenségekben tesztelt rendszer eredményeit az 5.1. fejezet végén ismertetem. Felmerül a kérdés, hogy a hasonló rendszerek mégis milyen javaslattal bírhatnak a pályadetektálás és akadályészlelés megoldására. Továbbiakban ezen szempontok szerint vizsgálom részletesen a megvalósítás lehetőségeit.

A [8] cikkben bemutatott rendszer a hagyományos gépi látás technikáit alkalmazva lokalizálta a sín párokat, közben akadályt keresve a sínszálak adott környezetében és a sín párok között. Hagyományos gépi látás előnye, hogy kiforrott technikák állnak rendelkezésre, végrehajtási sebességük előre, szinte pontosan meghatározható és viszonylag könnyen testre szabhatók. Érdeemes megemlíteni a Canny operátort, mely több, egymást követő képfeldolgozási technikával dolgozik. A bemeneti kép elmosását követően a Sobel [16] operátorral létrehozott gradiens kép képes kiemelni a képen szereplő éleket. A nem potenciális éleket a nem-maximumok elnyomásával (NMS) szűrik és csak a valós éleket tartják meg. Kettős küszöböléssel behatárolják a keresni kívánt éleket, melyeket hiszterézis segítségével összekötnek. Az operátor rengeteg erőforrást igényel és a technikák egy része nem párhuzamosítható. Az eredmény képen fals élek is megjelenhetnek, melyek nem kapcsolódnak a valódi sín pár éleihez, köszönhető a nem-maximumok elnyomásának. A kettős küszöbölés értékeinek automatizálása szükséges a változó fényviszonyokhoz való igazodás érdekében, mely egy újabb alrendszer bevezetését igényelné. Számomra, ami érdekes, az a Sobel operátor által létrehozott gradiens kép, mely a fényviszonyoktól függetlenül is elfogadható eredményt tud nyújtani. A Canny operátor egy bináris képet ad vissza, melyen az élek erőssége elvész és egy esetleges fals él kiemelése nem várt irányba terelheti a sín pálya keresést, melyre a „gyenge” élek megtartása egy megoldást nyújthat, ám a kettős küszöbölés miatt erre nincs mód.



5. ábra. Bazaltkőkockás pályán alkalmazott Canny operátor előtte és utána

Az 5. ábra szemlélteti, hogy a Canny operátor egy bazaltkőkockás burkolattal ellátott pályán milyen zajos eredményt produkál. Jól látható, hogy nem tesz különbséget a „gyenge” és „erős” élek között. A Canny abban esetben megfelelő választás, ha a küszöbértékek változtatása nem szükséges és a sínpárok szemmel is jó látható módon kitűnnek gyengén zajos környezetben is.

Első meggondolásra mégis csak jó megközelítés volna, ha lehetne párhuzamosakat keresni zajos környezetben. Erre a problémára nyújt megoldást a kép felülnézeti transzformációja, melyre példát a [10] cikkben mutattak be. A megközelítés, miszerint az eredeti képen látható sínpárok felülnézetbe történő transzformálása párhuzamosakat fog eredményezni, kifejezetten jó eredményt nyújt. A probléma a teljesítménnyel van. A projective transzformálás során az eredményképet interpolációs technikával szükséges meghatározni, mivel a mátrixszorzás¹⁵ eredménye cellánként nem biztos, hogy egész szám lesz. A pixelenkénti interpolálás erőforrásigényes¹⁶, mely során elképzelhető, hogy a koordináták leképzése többször is azonos eredményt fog adni. Egy párhuzamos számítás során, ha több szál is ugyan azon eredménycellába kívánja a számítás eredményét beírni, úgy versenyhelyzet alakul ki, mely hosszabb számítási időt eredményez a transzformálás során. Egy korai pályadetektáló megvalósítás olvasható a 5.1. fejezet végén, mely implementálja az említett felülnézeti számítást. Az 1. táblázatban összesítésre kerültek a [8] és [10] rendszer előnyei és hátrányai.

¹⁵ A perspektív transzformációs mátrix és a bemeneti kép pixelei alkotta mátrixának szorzata

¹⁶ OpenCV implementáció CPU-n és GPU-n egy GTX 1080-as GPU-val egyaránt ~12 ms végrehajtási idővel rendelkezett a tesztkörnyezetben.

Cikk (rendszer)	Pályaazonosítás módja	Előnye	Hátránya
<i>Video Analysis for Improving Transportation Safety: Obstacles and Collision Detection Applied to Railways and Roads</i> [8]	Canny operátor eredményét csúszó ablakok segítségével elemzi	Viszonylag gyors, Canny eredménye egyértelműen meg tudja határozni a pályát	Canny küszöbértékei fényviszonyoktól és pálya körüli zajtól függenek
<i>Obstacle-free range determination for rail track maintenance vehicles</i> [10]	Felülnézeti transzformáció után Hough operátor segítségével azonosít	Precíz pályaazonosítás, fényviszonyok nehezen befolyásolják	Felülnézeti transzformáció és Hough operátor lassú művelet

1. táblázat, Hasonló rendszerek összegzése pályaazonosítás módja szerint

Akadályészlelés terén részletesen ismertettem a [8], [12] és [13] cikkekben használt módszereket, melyek az ismertett problémákra egy optimális megoldást nyújtottak. Esetemben szükséges megvizsgálni, hogy az alábbi megközelítések mennyire tekinthetők optimálisnak objektum azonosítás során. A [10] cikkben bemutatott pályaazonosítási módszer részben tartalmazott egy minimális akadályészlelést is, mely során a gradiens irányát vizsgálták a sínpálya mentén. Esetemben akár az útburkolatba épített pályát is rengeteg zaj veheti körül, melyeket leginkább az öreg burkolat mentén megjelenő repedések és felületi hibák okozzák, ezért a gradiens iránya többszöri változást fog eredményezni. Esetemben a [10] módszer nem jól alkalmazható és csak magára sínre terjed ki, nem pedig az űrszelvényre. A [8] komplex rendszer nem csak a sínszál kisebb környezetét vizsgálja, nevezetesen a kereső ablakok területét, hanem a sínszálak közti szakaszt is figyelemmel kíséri. Mint korábban a [10] rendszer kapcsán is felvettem, a pálya közvetlen környezetében lévő zaj fals eredményt adhat a keresőablak vizsgálatakor, ezért érdemes nem a keresőablakot vizsgálni, hanem a teljes képet potenciális akadályok kapcsán.

Az űrszelvény és környezetének vizsgálatához egy kézenfekvőbb megoldást nyújt a [12] és [13] rendszere. A [12] cikknél ismertetett előzményképek felhasználására alkalmas módszer futás ideje 10 FPS-re tehető és számos esetben fals negatív eredményt mutatott, melyeket a képkülönbséget képző modul okozta a szerzők szerint. Egybefüggő objektumok problémát jelenthetnek szegmentáláskor, mely egy álló kocsisor esetén nem előnyös. A rendszert leginkább a nem valós idejű végrehajtási sebessége miatt vetem el.

A neurális hálók térnyerésének köszönhetően optimális megoldást nyújtana számomra a [13] cikkben ismertetett rendszer. A vizsgálatot a hálózat nem csak az űrszelvényre, de az egész képre kiterjesztve végzi. Az eredmény pontossága a hálózat tanításának menetétől és a felhasznált tanítómintától függ. A hálózat leírását mellékeltek a cikkben, ami szerint reprodukálható, de mivel a tanítóminta egyedi, publikusan nem elérhető és a hálózat helyes paraméterekkel történő tanítása nagyon sok időbe telhet, ezért esetemben nem kivitelezhető a rendszer implementálása.

Cikk (rendszer)	Objektum detektálás módja	Előnye	Hátránya
<i>Video Analysis for Improving Transportation Safety: Obstacles and Collision Detection Applied to Railways and Roads</i> [8]	Keresőablakok és köztes szakasz vizsgálata gradiens intenzitás alapján	Gyors, fényviszonyoktól független	Csak a sín és sínszálak közti területet vizsgálja
<i>Obstacle-free range determination for rail track maintenance vehicles</i> [10]	Keresőablakon belüli gradiens irányok vizsgálata	Gyors, fényviszonyoktól független	Csak a sínszálakra terjed ki a vizsgálat
<i>Moving camera background-subtraction for obstacle detection on railway tracks</i> [12]	Előzményképek és aktuális kép különbségével végzett szegmentálás	Nagy távolságra alkalmas, fényviszonyoktól független	Lassú, szegmentálás pontatlanságot eredményez
<i>Real-time Obstacle Detection Over Rails Using Deep Convolutional Neural Network</i> [13]	Neurális hálózat segítségével	Relatív gyors, teljes képet vizsgálja	Hosszú tanítási idő, tanítóminta nem publikus

2. táblázat, Hasonló rendszerek összegzése objektumdetektálás alapján

A 2. táblázat eredményeiből következően a neurális hálózat használata tűnik a legjobb megoldásnak objektumdetektálásra. A 6. fejezetben részletesen tárgyalom, hogyan oldottam meg az objektumdetektálás problémáját valós időben. Az 1. táblázatban feltüntetett pályadetektálási technikák csak részben nyújtanak megfelelő lehetőségeket. A [8] cikkben használt Canny operátor részfolyamatát képző Gradiens kép előállítása egy jobb éldetektálást biztosít fényviszonyok, zaj és párhuzamosítás tekintetében, illetve a csúszó ablakok felhasználhatóak az egyes pályaszakaszok azonosítására melyet a 4.2.2 fejezetben ismertetek.

4. RENDSZERTERVEZÉS

Egy kamera képén történő akadály megállapítása egy összetett gépi látással kapcsolatos problémaként értelmezhető, mely során olyan technikákat kell alkalmazni, ami a kép hatékony feldolgozását és értelmezését teszi lehetővé a gép számára. Az algoritmusnak optimális megoldást kell nyújtania a környezeti hatásoktól függetlenül, nem kockáztatva a végrehajtási sebességet. A következőkben ismertetem az rendszer moduljaival szembeni követelményeket, az alkalmazandó technikák kiválasztását, egyes modulok működését és végül a vázlatos rendszertervet.

4.1. MODULOK ÉS KÖVETELMÉNYEK

Az akadályészlelés egy összetett probléma, melyet kisebb alproblémák együttese határoz meg. Két alproblémát kell meghatározni; a pálya azonosításának problémáját és az objektumok detektálását. A kettő eredményét felhasználva határozhatóak meg az akadályok. A program futása során a pálya azonosítás és objektum detektálás két külön folyamatként jelennek meg, melyek közötti kapcsolat egyedül a bemenet, ezért azt lehet mondani, hogy az egyes folyamatok függetlenek egymástól, felmerül a párhuzamos végrehajtás lehetősége. Az implementálás során nem alkalmazok modul szintű párhuzamosítást, melynek okát a 5.3. fejezetben részletezem. A rendszer iteratívan dolgozza fel a kamera képeit, ezért cél a lehető legkisebb iterációs idő elérése, úgy, hogy a feldolgozás minőségének rovására ne menjen.

Szükséges meghatározni az egyes modulokkal szembeni követelményeket. Minden modul esetén elmondható, hogy törekedni kell a legjobb teljesítmény elérésére, biztosítani kell a valós idejű feldolgozást. Általánosságba véve valós idejű feldolgozásról akkor beszélünk, ha az algoritmus futási ideje megegyezik a forrás (bemenet) által szállított képek sebességével. Esetemben a forrást a kamera jelenti, ezért azt lehet mondani, hogy valós idejű feldolgozási sebességgel bír a rendszer, ha a kamera képeit legalább olyan vagy annál nagyobb sebességgel képes feldolgozni, mint ahogy az közvetítve van a rendszer felé.

A pályadetektálás és objektumdetektálás során függetlennedni kell az időjárási jelenségek és napszak tekintetében vett fényviszonyoktól. A két modul közti kapcsolat egyedül az azonos bemenet jelentheti, függőségek nem lehetnek köztük. A pályadetektálásnak eredményül olyan pontthalmazt kell visszaadnia, mely képre vetítve egyértelműen meghatározza a villamos pálya úrszelvényének elhelyezkedését. Az objektumdetektálásnak vissza kell térnie az észlelt objektumok gyűjteményével,

melyek a képre vetítve egyértelműen meghatározhatóak, mind a pozíciójuk, mind a méretük. Az érzékelt objektumokat típusuk szerint eltérő módon vizsgálom, ezért az objektumok klasszifikálása szükséges. Visszatérésként továbbá szükséges az objektum típusa is. Az akadályészlelő modultól elvárt, hogy a két modul eredményét felhasználva azonosítsa az úrszelvényben érzékelt objektumokat és visszatérésül az akadályt képző objektumot szolgáltatassa.

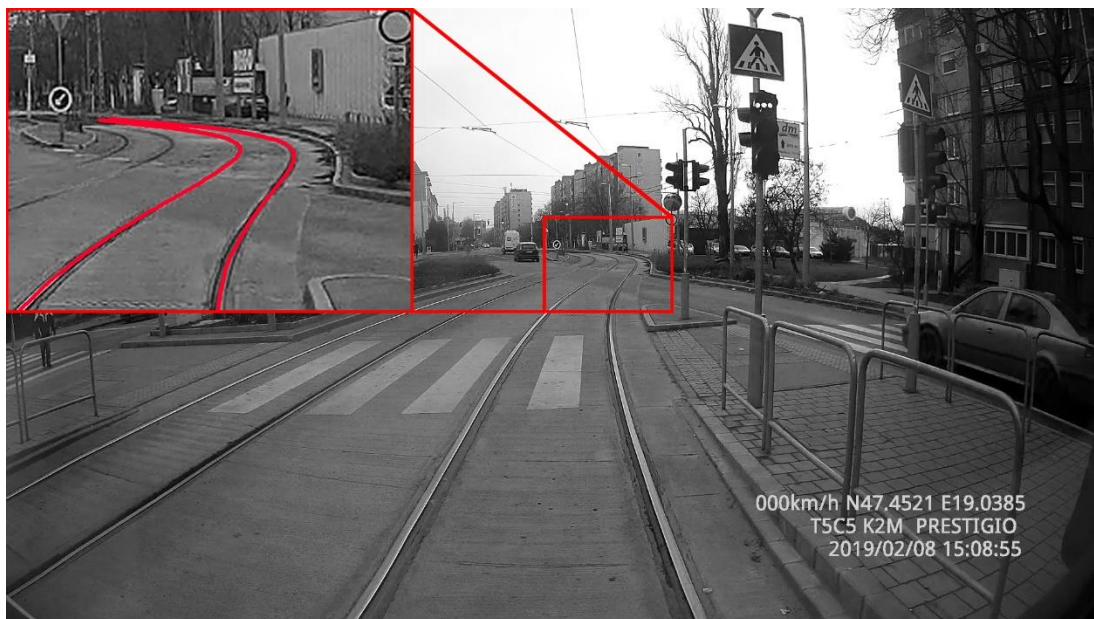
4.2. VASÚTI SÍNPÁROK AZONOSÍTÁSA

A pályadetektálásért felelős modul feladata, hogy a vasúti sínpárokat kamerától mérve a lehető legnagyobb távolságban azonosítsa. A probléma megfeleltethető egy képfeldolgozási és egy gépi látási feladatnak, mely során a digitális kép tartalmát különböző eljárások segítségével manipulálom úgy, hogy a végeredmény alkalmas legyen a sínpárok egyértelmű meghatározására. Ahhoz, hogy az algoritmus képes legyen meghatározni a pálya elhelyezkedését, szükséges definiálni melyek azok a képi jellemzők, amik leszűkítik a keresési teret. Ezen jellemzők meghatározásához felhasználhatók a korábban szerzett vasúti pályával kapcsolatos ismeretek.

4.2.1. ELŐISMERETEK FELHASZNÁLÁSA

A vasúti pálya ismertetésére kitértem a 2.1 fejezetben, melyből a pálya vizuális aspektusait fogom figyelembe venni azonosítás során. A pálya számos jellemzővel bír, melyek alkalmassá teszik a környezetből való kitűnésre. A következőkben ezen vizuális jellemzőket fogom ismertetni és hogy milyen jelenségek befolyásolják a jellemzők „jóságának” mértékét.

Elsősorban elmondható, hogy a vasúti pálya részét képző sínszálak közti távolság konstans, melyet a szerelvény abroncsai közti távolság határoz meg. A pálya ezen tulajdonságára viszont nem minden esetben lehet számítani, mely az adott nézőponttól függ. Ha felülnézetből tekintünk le a pályára, a sínszálak párhuzamos módon haladnak egymás mellett minden esetben. Ha a vezetőfülkéből történik a szemlélet, akkor látható, hogy a sínszálak egy pontba tartanak (vanishing point), mely a horizonton helyezkedik el. A sínszálak közti távolság ebben az esetben is konstans, viszont ezt az értéket a szemlélőtől való pálya távolságához kell kötni. Fizikailag minél távolabb esik a vasúti pálya, annál kisebbnek látni a sínszálak közti távolságot. A pálya ezen tulajdonságára lehet építeni, miszerint a pálya és a szemlélő közti fizikai távolsághoz rendelhető egy változót, mely meghatározza a sínszálak közti távolságot. Ezt a változót egy távolságfüggvény írja le, mely bemenete és kimenete közt egyenes arányosság áll fenn.



6. ábra. Sinszalak összefutása

Érdemes megemlíteni, hogy a távolságot leíró függvény nem nyújt pontos értéket, ha a pályaszakasz ívet ír le. Az ív a felülnézeti képen mért távolságokat nem befolyásolja, viszont perspektív nézetből a távolságfüggvénnyel meghatározott értéknél is kisebb eredmény mérhető (6. ábra). A jelenség modellezéséhez szükséges bevezetni egy szorzót a távolságfüggvény kimeneti értékére, mely szemlélő és az ívhez tartozó irány közt mérhető szög függvényében csökkenti a kimeneti értéket. A valóságban megfigyelhető, hogy ha kellő távolságra helyezkedik el az ívelt pályaszakasz, akkor a görbület mentén a pálya egy sinszalát fog képezni (6. ábra). A jelenség a szemlélőtől mért ív kezdeti távolságának függvényében jelentkezik. A kép feldolgozása során a későbbiekben figyelembe kell venni ezen jelenséget.

A fentiekből következik, hogy a szemlélő nézőpontjából, azaz a kamera képén a sinszalak kiindulási pontjai a kép alján fix távolságra helyezkednek el egymástól. A pozíciók azonban a szerelvény jellegéből adódóan nem mindig egyértelműek. A kamera látószögétől függően a sinszalak kiindulási pontjai a kép bal alsó szegmensétől egészen a jobb alsó szegmenséig elhelyezkedhetnek. A jelenségre a szerelvény felépítése ad magyarázatot, miszerint minden forgóváz a szerelvény eleje és vége között helyezkedik el, ezért az alváz/kocsiszekrény két vége az ívelt pályaszakaszon való közlekedéskor csak részben helyezkedik el a pálya felett (7. ábra).



7. ábra. Ívelt pályán haladó villamos

Egy másik jellegzetesség a sínszalak anyaghasználatából adódó fémes/szürke felület. A fémes hatásból adódóan a sínszalak a rávetődő fénysugarak egy részét visszaverik, ezért a felület színértéke az eredeti szín körüli ε tartományban változhat a betekintési szög és fényviszonyok függvényében. A sínszalak szintartományai egyértelműen meghatározzák a pályát, amennyiben a pálya burkolatát képző felület színe eltér.



8. ábra. Nyomcsatorna, tiszta (fent), homokolt (lent)

Egy másik jellemző a sínszalak között kialakított, nyomkarima futását biztosító nyomcsatorna, mely többnyire egy jól látható sötét területként szerepel a képen. A csatorna jellegéből adódóan, a nap állásától függően, egy árnyékolt terület, mely az esetek nagy részében sötétebb a pályaburkolatnál és a sínkoronánál (8. ábra fenti kép).

A szerelvény csúszásvédelmét képző homokszóró a csatornát képes feltölteni homokkal, mely után a csatorna elveszítheti sötét jellegét (8. ábra lenti kép).

A felsorolt vizuális jellemzők „jóságát” az időjárási viszonyok, mint például az eső vagy a hó képes rontani, ezzel pontatlanabb pályaaazonosítást eredményezve. Míg ezek a jellemzők egyértelműen meghatározzák a sínszálak elhelyezkedését, a gyakorlat során mégis úgy bizonyult, hogy nem lehet rájuk alapozni a detektálás során. A fémes felület esetén a problémát a betekintési szög okozza, mely során a közeli sínszál színe gyakran eltér a messzebb elhelyezkedő szakasztól, illetve sokszor megegyezik az útburkolat színével. A nyomcsatorna a gyakorlat során többnyire olyan mértékben van megtelve homokkal és más szennyeződéssel, hogy nem emelkedik ki kellő képen a sínszál környezetéből. Az említett problémákból kifolyólag nem lehet építeni detektálás során a kép színvilágából adódó jellemzők kinyerésére.

4.2.2. PÁLYAAZONOSÍTÁS TECHNIKÁJA

Pályadetektálási folyamat során a bemeneti képen szereplő sín párok leképezése történik egy pontthalmazba. A folyamat során két egymásra épülő probléma megoldása történik. Az első probléma a kép alkalmassá tétele ahhoz, hogy a pályát leíró részleteket ki lehessen nyerni. Ezt számos alacsony szintű feldolgozási eljárás segítségével valósítható meg. A második probléma az előfeldolgozott állomány értelmezése, mely során az algoritmusnak a pályaleíró pontokat kell meghatároznia a kép pixelei alapján. A következőkben a két problémát megoldó folyamatokat ismertetem.

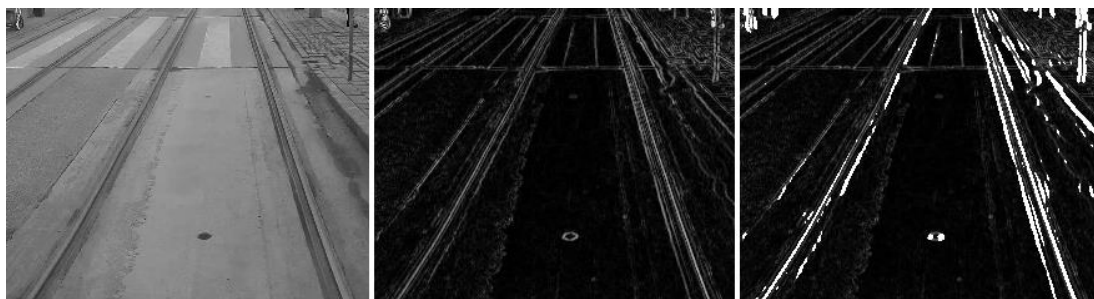
I. ELŐFELDOLGOZÁSI FOLYAMAT

Első lépésben előállítom a kamera képének szürkeárnyaltos változatát. A kamera egy három színcsatornás RGB képet állít elő, melynek pixelértékeit egy $M \times N \times C$ dimenziós mátrixban tárolom, ahol M a kép magassága, N a kép szélessége és C a színcsatornák száma. Szürkeárnyaltos konverzió során a színeket meghatározó csatornákat egyesítem, melynek eredményéül egy $M \times N \times 1$ dimenziós mátrix áll elő, ahol a pixelek értékeinek reprezentálása egy 0-tól 255-ig terjedő diszkrét tartományban történik. A konverzió eredményeképpen a további feldolgozási eljárások komplexitása csökken, ezért jobb futási idő érhető el.

Ahhoz, hogy a sínszálakat egyértelműen meghatározó jellemzőket ki lehessen emelni, elsősorban szükséges az élek meghatározása. A folyamat eltér a klasszikus

éldetektáló eljárásoktól olyan tekintetben, hogy magukat az éleket nem címezzük meg, helyette az intenzitásváltozások nagyságának értékeinek meghatározása történik. Képfeldolgozási szempontból maga az él egy jelentős helybeni intenzitásváltozás, azaz a képpontok egy adott környezetre vetített értékváltozásának mértéke. Ezt az értékváltozást reprezentálja a kép gradiensének a nagysága. Minden pixelre vetített intenzitásváltozás mértéke alkotja a gradiensképet, mely alkalmas a különböző élek által meghatározott jellemzők kinyerésére. A gradienskép pontjait konvolúciós eljárás során állnak elő. Konvolúciós kernelekkel dolgozik számos éldetektálásra szolgáló szűrő. G. M. H. Amer és A. M. Abushaala az „*Edge detection methods*” [17] cikkében ismertetett számos éldetektáló technikát. A technikák során az éleket előállító konvolúciós kernelek eredményeit küszöbölték. Az adott küszöbérték feletti éleket vették valós éleknek, illetve további feldolgozások alá vetették. Esetemben minden élre szükségünk van, ezért a küszöbölés előtti állapot előállítása a cél. Két gradiens alapú éldetektáló eljárás jöhet szóba. Az első a *Laplace* [18], a másik a *Sobel* [16] operátor. A Laplace operátor egy konvolúciós kernellel állítja elő a másodrendű deriváltakat. Hátránya, hogy érzékeny a zajra, ezért célszerű az operátor használata előtt a képet elmosni. A *Sobel* operátor egy X és egy Y kernellel együttesen állítja elő a gradiens képet. Elsőfokú deriválttal történő számítás végett nem érzékeny a zajra, viszont az eredmény előállítása a két kernel miatt lassabb művelet a *Laplace*-al szemben. A *Sobel* nem tesz különbséget a gyenge és erős élek között, ezért fogom alkalmazni.

A *Sobel* X és Y komponense felhasználható finomhangolás céljából is. A vasúti sínszálak perspektív nézetből egy pont felé haladnak letről felfele, ahol megfigyelhető, hogy a sínszálakat alkotó élek leginkább függőleges irányba dominálnak, melyet a vízszintes szűrést végző kernel érzékel. Ezért a gradienskép létrehozásakor az X és Y komponensek súlyozása szükséges, melyek közül az X komponens nagyobb értékkel bír, ezzel jobban kiemelve a sínszálak függőleges éleit.



9. ábra. Balról jobbra: Eredeti kép, Sobel 3×3 , Sobel 3×3 és 5×5 kombinálva

Az 9. ábra közepén látható a *Sobel* operátor eredménye. Megfigyelhető, hogy a sínszálak körüli intenzitásváltozás mértéke nagyon hasonló értékkel bír, mint a sínszál környezete. A jelenséget a sínszálak gyenge kitűnésével magyarázható. Egy megoldást nyújt a *Sobel* operátor módosítása, mely alatt az eredeti 3×3 -as kernel 5×5 -ös kiterjesztett változatát értem. Eredményképpen egy nagyobb területet bejárása történik, mely során a differenciák nagyobb értéket tudnak felvenni, ezzel felerősítve az erős éleket, a gyengéket ezzel egyidőben elnyomva. A technikát önmagában nem érdemes alkalmazni, az élek felerősítése csak abban az esetben fog megfelelő eredményt nyújtani, ha kellően nagy változás történik a konvolúciós kernel területén. A tapasztalatok szerint útburkolattal ellátott felületen nem biztosít pontos élkimelést, ezért a 3×3 -as *Sobel* kiegészítéseként érdemes használni. Az eredmény a 9. ábra jobb oldalán látható.

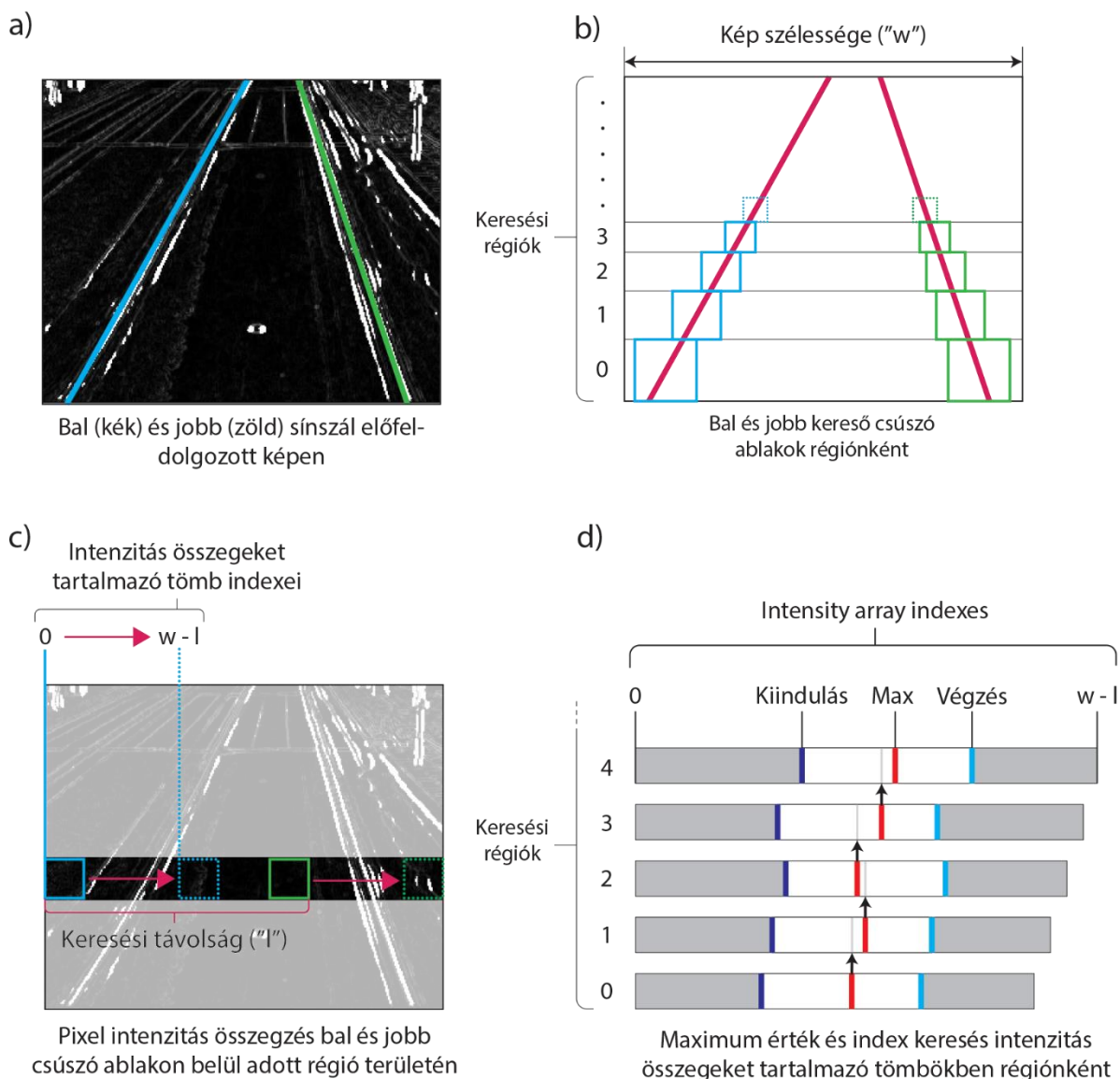
Az 5×5 -ös *Sobel* kernel egy mellékhatása, hogy nem csak a sínszálakat erősíti fel, de a környezetét is, ezzel olyan jelenségeket eredményezve, mint például a zúzottkőves felépítmény túlzott megjelenése. A hatás csökkentése érdekében a szürkeárnyalatos képen homályosítást végzünk az operátor használata előtt. Homályosítással, vagy másnéven életlenítéssel kapcsolatban érdemes megemlíteni két szűrőt. Az egyik a *Gauss*-féle operátor, a másik a *Median* szűrő. A *Median* egy nem lineáris szűrő, mely a konvolúciós ablakban lévő elemek sorrendbe tétele során kiválasztja a tömb középső elemét, mely a szűrés eredményét képezi. A technika célja a zaj csökkentése, miközben bizonyos kondíciók alatt megtartja az éleket. A *Gauss*-féle szűrő egy lineáris, konvolúciós kernelt használó szűrő, mely a *Gauss*-függvényen alapszik. A függvény normáeloszlást követ, ezért a kernel közepétől kifelé haladva az elemek értékei csökkennek. A szűrő biztosítja a kép gyors, egyenletes életlenítését, viszont eredményképpen nem csak a zajt csökkenti, de az éleket is gyengíti. A technikák összehasonlítását bemutatta A. Kumar és S. S. Sodhi a „*Comparative Analysis of Gaussian Filter, Median Filter and Denoise Autoencoder*” [19] cikkében. A tesztkörnyezetben elért eredmények alapján a *Median* szűrő jobb eredményt biztosított kevesebb iterációs idő alatt. Korábbi teszteredményeim alapján a *Median* szűrő valóban képes volt a zúzottkő felépítmény éleit gyengíteni, miközben a sínszálak éleit megtartotta.

Az előfeldolgozás végén egyesítem a *Sobel* 3×3 -as és 5×5 -ös kernelével alkotott gradiens képet, melynek eredménye a 9. ábra jobb oldalán látható.

II. PÁLYALEÍRÓK MEGHATÁROZÁSA

Az előfeldolgozás során módosított kamera képén kiemelésre került minden olyan erős és gyenge él, melyek szemléltethetik a vasúti pályát. A feltételes kijelentés oka, hogy maga a pálya élerőssége nem csak a sínszalak mentén van jelen, hanem a környezetében is, például a sínszalak alátámasztását biztosító aljak, vagy a bazaltkőkockás úttest mentén, melynek felszíne számos esetben erősebb élekkel rendelkezik. Ezeket a nemkívánatos jelenségeket zajnak tekintettem. A zaj kiküszöbölésére előfeldolgozás során a *Gauss*-féle elmosás technikáját vettem igénybe, melyet legfeljebb olyan erősséggel alkalmaztam, hogy magát a sínszalak éleit ne rontsa a felismerhetetlenségig.

A pályaleírók meghatározása során nem konkrét élék keresése történik, hanem élcsoportok vizsgálata. A megközelítésnek két oka van; Az egyik, hogy számos esetben a pálya nehezen vehető ki még szabad szemmel is a képről, ezért gyenge élék formájában jelenik meg. A gyenge élék több esetben is megszakadnak, esetleg zaj formájában jelennek meg a sínszalak körül. Feltételeztem, hogy ahol több gyenge él van jelen, ott található valójában a sínszál. A másik ok, hogy élcsoportokat vizsgálok, visszavezethető az 5×5 -ös kernel egyik problémájára, miszerint csak az erős éleket tartja meg, a gyengéket elveti, mely szintén élszakadások formájában jelentkezik, esetek többségében zajmentesen. A probléma kiküszöbölésére alkalmaztam a 3×3 -as kernel kiegészítésére az 5×5 -öt, mely során az erős élék szakadásait a gyenge élék „feltöltötték”. Mivel a pályát különféle erősségű élcsoportok alkotják, ezért küszöbölést korábban se alkalmaztam élék szűrésére. A *Canny* éldetektor, illetve *Hough* transzformáció használatát ezért is vettem el.

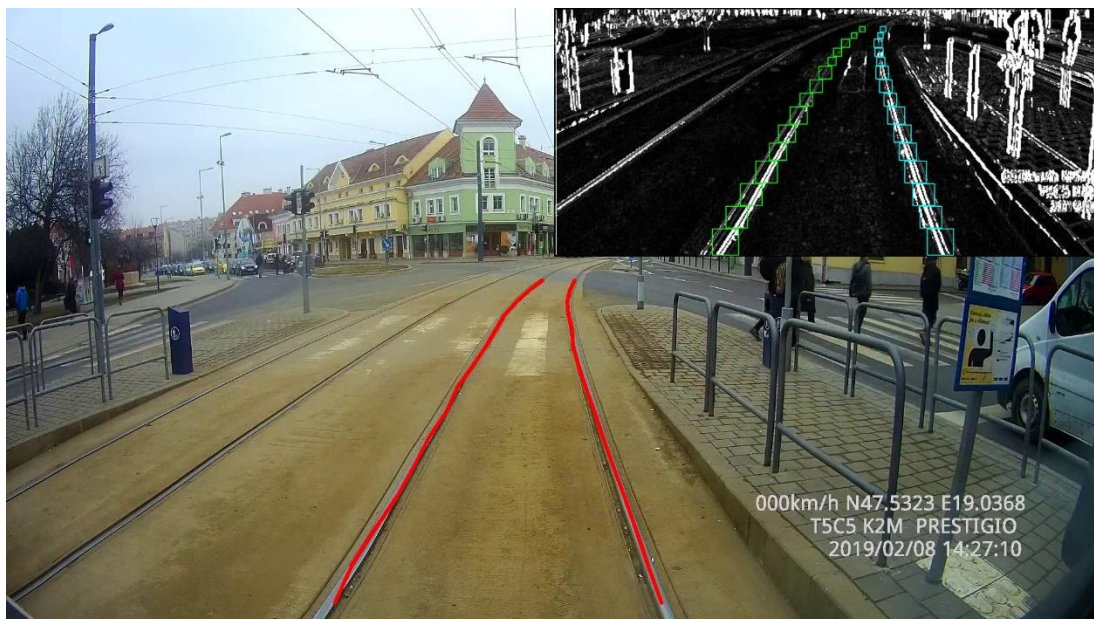


10. ábra, Pályadetektálás lépései előfeldolgozás után

A megfelelő élcsoportokat keresőablakkal lokalizálom, melyekben a gradienskép élintenzitásait összegzem. A keresés alapjául azt vettem, hogy ahol magasabb intenzitásösszeggel rendelkezik a keresőablak területe, ott nagy valószínűséggel sínszál található. A korábban ismertetett távolságfüggvény felhasználásával meghatároztam, hogy milyen magasságban kell keresni a megfelelő sínszálak közti távolságokat. A képet felbontom horizontálisan keresési régiókra és meghatározom a szükséges csúszó ablak méreteket (10. ábra, „b” pont). A régiókon belül előre tudni lehet, hogy milyen régió magassághoz milyen keresett távolság tartozik, melyet később az akadály távolságának meghatározására felhasználok. A régiókon belül a sínszálak keresése párban történnek, fix távolságra egymástól. A

keresés során a csúszó ablakokat végig csúsztatom a régió területén és intenzitás összegeiket eltárolom (10. ábra, „c” pont). Ahol a két ablak intenzitásösszege a legnagyobb, a vasúti pálya nagy valószínűséggel ott található, ezért egy maximum keresést végzek az intenzitás összeget tároló tömbökben, régiónként. A számítási folyamatot iteratív módon kell elvégezni, melynek lépésszáma függ a csúszó ablakok méretétől és a régiók számától (10. ábra, „d” pont). A megtalált maximális értékhez tartozó indexszám jelenti a következő tömbben a keresési terület közepét. Ez egy kifejezetten lassú folyamat tud lenni, ezért a felgyorsításának érdekében az előfeldolgozás végén meghatározom az előfeldolgozott képre vett integrálképet, mely lehetővé teszi az intenzitásösszeg számítását három lépésből. A keresés során a pályát lentről felfele építem fel.

A csúszó ablakokon belül megjelenő gyenge és erős élek néhány esetben hamis eredményhez vezethetnek, melyek a keresést akár teljesen rossz irányba is terelhetik. A probléma kiküszöbölésére egyrészt az előfeldolgozás finomhangolása szükséges, illetve a régiókon belüli keresést egy adott irányba terelni. Az irány meghatározásához az előző– két régió képezte haladási irány fogja megmondani, merre kell folytatni a keresést. A régió belüli keresést korlátozom olyan kisebb területre, mely a csúszó ablak szélességének vett kétszerese feleltethető meg. A korlátozás biztosítja, hogy amennyiben egy vélhetően rossz irányba halad a keresés, úgy vissza tudjon térni az algoritmus a megfelelőnek vélt eredeti irányra. Az előismeretek felhasználása során említésre került, hogy a sínszálak megfelelő nagyságú ívelt pályaszakaszon, horizontálisan egy sínszálként jelennek meg (6. ábra). A probléma kiküszöbölésére a keresés irányának meghatározását követően, az irányszög nagyságától függően a korlátozott keresési területet megnövelem a keresés irányába. Az ismertetett keresési területet a 10. ábra, „d” pontján látható. A megoldás biztosítani fogja, hogy a horizontális megjelenéshez közeledő pályaszakaszra a keresés ki tudjon térni.



11. ábra, Detektált pálya

Sikeres pályaaazonosítást követően a régióként azonosított pályaszakaszok bal és jobb sínszalát meghatározó pontjai visszatérési értéként szolgálnak, melyek egyértelműen behatárolják a pálya területét (11. ábra).

4.3. OBJEKTUMOK DETEKTÁLÁSA

Az objektumdetektáló modul feladata, hogy a képen szereplő objektumokat, alakzatokat felismerje. Elhelyezkedésüket és méretüket olyan formában adja vissza, melyekkel később további számításokat lehet végezni.

Akadályészlelés során szükséges meghatározni, mely objektumok képezhetnek akadályt az űrszelvényen belül. Ezek az objektumok meghatározása történhet a mozgásuk alapján, melyek pixel intenzitás változást okoznak haladásuk során. Az Optical Flow [20] technika ezeket az intenzitás változásokat követi le. A technika fő feladata az aktuális és előző kép különbsége során keletkezett flow vektorok meghatározása, melyek szegmentálását követően képesek meghatározni az objektum alakját, lehetőség nyílik az objektum követésére. A technika hátránya, hogy önmagában csak nem mozgó, fix kamerával történhet objektumdetektálás, hisz könnyen belátható, hogy a kamera mozgása során az egész képre kiterjedő flow vektorok jelenhetnek meg. Egy korábban bemutatott rendszer [12] az Optical Flow technikára építve oldotta meg a mozgó kamerával történő objektum detektálást, ám az említett technikát elvetettem, mert a végrehajtási sebesség nem volt számomra megfelelő.

Amennyiben nem mozgásváltozásból adódóan kell objektumot meghatározni, úgy lehetőség van a pixelek kapcsolatainak vizsgálatára. Egy objektumot számos pixeleyüttes határoz meg. Amennyiben emberi szemmel nézzük a képet, úgy bizonyos jellemzők alapján írjuk le mik lehetők fel a képen. Ezek a jellemzők az ember számára egyértelműen meghatározzák a fellelhető objektumokat. Ezeket a jellemzőket lehetőség van kinyerti a képről különböző konvolúciós technikák segítségével. Minden keresett jellemző meghatározható konvolúciós kernelek segítségével, melyek megadják, milyen jósággal rendelkezik az adott jellemző a vizsgált területen. A megfelelő jósággal rendelkező jellemzők közti kapcsolatot neurális hálózatok segítségével meghatározhatók. Azokat a neurális hálózatokat, melyek a képen szereplő objektumok detektálására képesek a különböző jellemzők kinyerésével és közöttük fennálló kapcsolat megállapításával, konvolúciós neurális hálózatoknak nevezik.

4.3.1. KONVOLÚCIÓS NEURÁLIS HÁLÓZATOK

A konvolúciós neurális hálózatokat (CNN) széleskörben használják minta- és képfelismerési problémák során, mivel a hagyományos technikákkal szemben számos előnnyel rendelkeznek. A CNN egy speciális neurális hálózat típus, mely egy vagy több konvolúciós réteggel rendelkezik, almintavételező (subsampling) réteggel párosítva, melyet egy vagy több sűrűn összekötött réteggel kapcsolnak össze. A CNN alapját az agy részét képező „látókéreg” ihlette. A konvolúciós rétegek felelnek a jellemzők kinyeréséért. Konvolúció során a kép átalakítását végezzük úgy, hogy egy konvolúciós kernelt (ablakot) helyezünk a pixelekre, melyet végig csúsztatunk a képen. A kernel olyan értékek alkotta mátrix, mely mérete és értékei meghatározzák a konvolúciós folyamat eredményét. A konvolúciós réteg szűrőjének értéke, azaz a súlyait, a hálózat tanítása során állítódnak be. A konvolúciós rétegek képesek kinyerni a helyi jellemzőket, mert a helyi neuronok csak a környezetüket vizsgálják, ami a szűrő méretére¹⁷ korlátozódik. Hagyományos technikákkal szemben a szűrők értékeit nem kézzel állítják be, a tanításhoz használt képeken szereplő objektumok felismerésére fog szakosodni a hálózat. Egy megfelelően kialakított és tanított hálózat képes akár ezer különböző objektum klasszifikálására és helyzetmeghatározására.

A bemeneti és kimeneti réteg között helyezkednek el a rejtett rétegek. A rejtett rétegek állnak konvolúciós és egyesítő (pooling) rétegekből. A konvolúciós rétegek eredménye a bemeneti mátrixhoz méretéhez képest egy kisebb mátrix, mely a

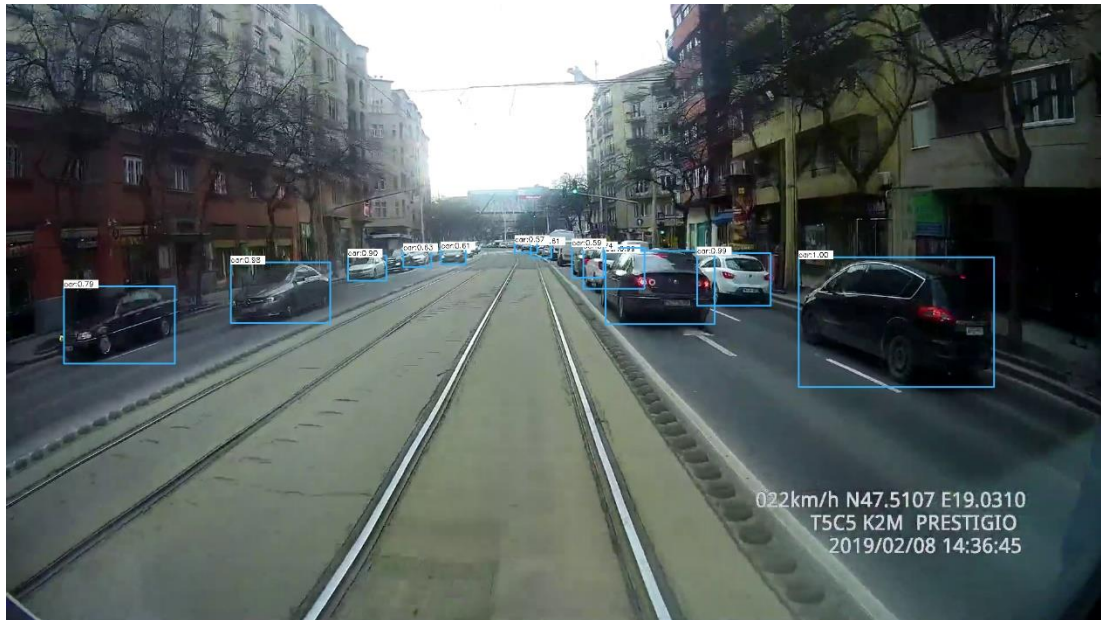
¹⁷ Általában 3x3-as szűrő méretet alkalmaznak.

konvolúciós szűrők sajátossága miatt van. A konvolúciós réteget egy pooling réteg követi, mely feladata a mátrix dimenziójának további csökkentése. Általánosságban elmondható, hogy míg a konvolúciós réteg felel az objektum jellemzőinek a kinyeréséért, addig a pooling réteg az érzékelt jellemzők összekapcsolását látja el. Végül egy teljesen összekötött réteg felel a felismert objektumok klasszifikálásáért. [21]

Képek klasszifikálását követően a következő előrelépést a régió alapú konvolúciós neurális hálózatok [22] (R-CNN) megjelenése jelentette. Az R-CNN szelektív kereséssel generál úgynevezett „javaslatokat”, hogy a képen hol találhatóak a hálózat számára „érdekes” területek, melyek objektumokat tartalmazhatnak. Ezeket a területeket fix méretté konvertálva átadja a CNN-nek, mely egy jellemző (feature) kinyerő funkciót lát el, ami 4096 különböző jellemző előállítására képes. Az eredményt átadja egy SVM klasszifikáló rétegnek. Maga a klasszifikáláson túl négy különböző eltolási értéket is meghatároz az egyes objektumokhoz, melyek együttesen megadják az objektumok körbe foglalo négyzetet. Az R-CNN hátránya, hogy nagyon hosszú végrehajtási idővel bír, mely nem teszi alkalmassá valós idejű alkalmazások használatához. Továbbá pontatlan javaslatokkal térhet vissza a szelektív tanulási képesség hiánya miatt. Ezen hibák kiküszöbölésre továbbfejlesztették az alap modellt mely Fast R-CNN és Faster R-CNN névre hallgat. [23]

2015-ben az objektumdetektálás egy innovációjaként bemutatták a YOLO-t (*You Only Look Once*), egy egységes, valós idejű objektumdetektáló módszert. Az R-CNN-nel ellentétben a YOLO nem használ javaslatok és klasszifikáció kombinációját az objektumok észlelésére, hanem az objektumokat befoglaló téglalapokat (bounding box) és az osztályba sorolás valószínűségének értékét a pixelekből határozza meg. YOLO bemenetként megkapja az egész képet. Egyetlen konvolúciós hálózattal egyidejűleg határozza meg a befoglaló téglalapokat és az osztályvalószínűségeket. Ez a megközelítés számos előnnyel jár. Először is, leegyszerűsíti az észlelési folyamatot egy regressziós problémává, mellyel sikerül elkerülni az R-CNN bonyolult és erőforrás igényes folyamatait. Másodszor, a régiójavaslat alapú megközelítéssel szemben, a predikciók megállapítása során rendelkezésre áll a teljes kép, mint információ, ezzel csökkentve a háttérhibák számát. Az érem másik oldala viszont, hogy a YOLO nagy valószínűséggel fog pontatlanabb helyzetmeghatározásokat végezni. Kiértékelés során a kis és nagy befoglaló téglalap hibákat egyenrangúan kezeli, mely pontatlan helyzetmeghatározást eredményez. Másodszor, objektumdetektálás során képes figyelmen kívül hagyni néhány objektumot, példának okául mikor számos objektum helyezkedik el egy vizsgált cellán belül. YOLO

alapvetően nem tudja a hibát elkerülni, ha minden befoglaló téglalap ugyan abban a cellában helyezkedik el és ugyan azon osztályba vannak sorolva. Idővel megjelentek a YOLO továbbfejlesztett verziói, melyek igyekeznek ezen hibákat kiküszöbölni. A YOLO által prediktált járművekre egy példa látható a 12. ábrán.

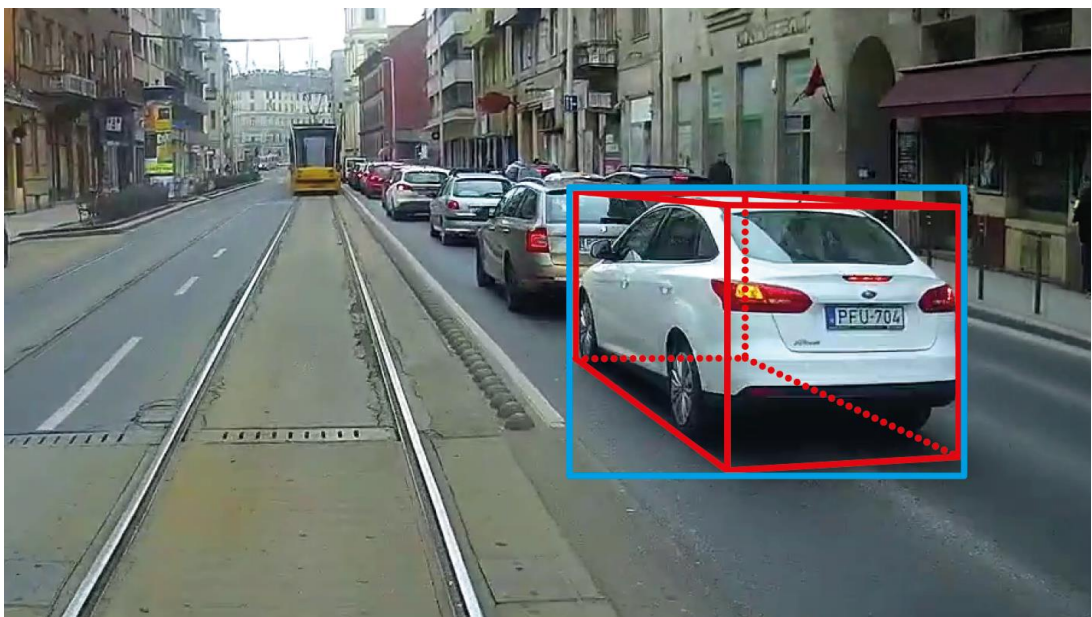


12. ábra. YOLOv5 objektum predikciók

Az objektumdetektálást az ONNX Runtime [24] felhasználásával a YOLO ötödik verziójával [25] (YOLOv5) végzem. Az ONNX Runtime egy nyílt forráskódú projekt, mely különböző keretrendszeren célzott a gépi tanulást gyorsítani. Az ONNX Runtime könyvtára lehetőséget nyújt a YOLO CUDA [26] környezetben történő hatékony implementálására, mely lehetővé teszi a neurális háló GPU-n történő futtatását valós időben. A CUDA az Nvidia vállalat által megalkotott számítási modell és párhuzamos számításokat támogató platform Nvidia GPU-k számára. A CUDA környezet nem csak neurális háló futtatására alkalmas, hanem számos képfeldolgozási technika és egyéb számítás gyorsítására is kiváló lehetőséget nyújt.

4.4. AKADÁLYÉSZLELÉS

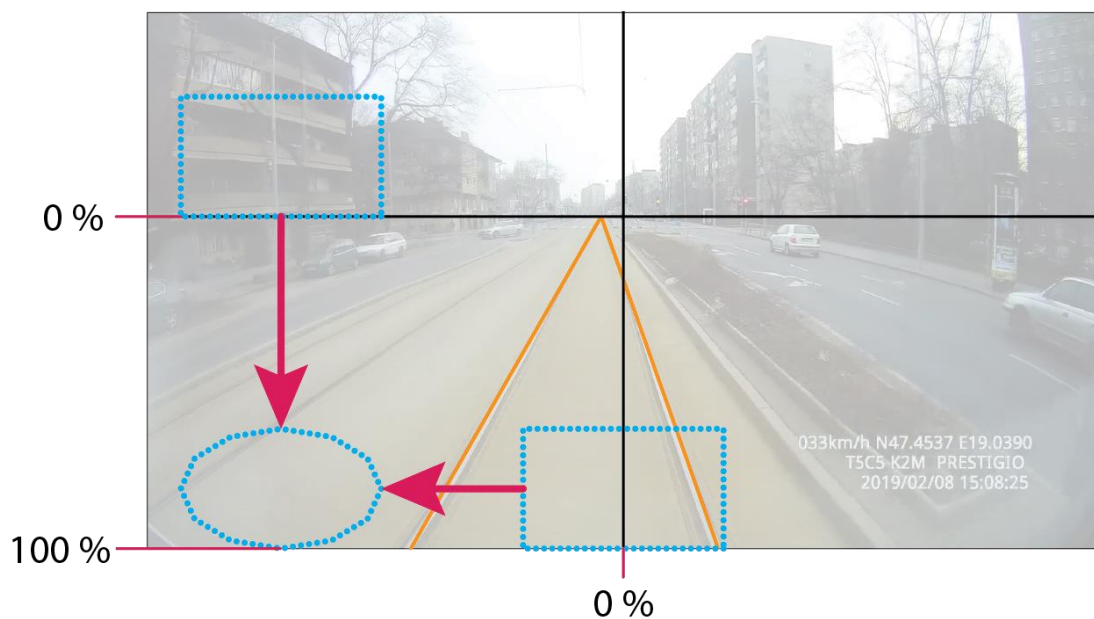
Az akadályészlelést megvalósító modul feladata, hogy az érzékelt objektumokat vizsgálja meg olyan szempontból, hogy az azonosított pályából képzett úrszelvényen belül helyezkednek-e el. Ha igen, akkor térjen vissza az akadályt képző objektummal. A problémát szükséges megvizsgálni elsősorban az emberi logika szempontjából. Ha ránézünk a kamera által rögzített képre, egyértelműen meg tudjuk határozni a vizsgált tárgy helyzetét, melyből következtetésként levonjuk, hogy ránézésre hol helyezkedik például egy autó a pályához képest. Az emberi szem érzékeli a perspektív képen fellelhető mélységeket. Egy gép viszont nem rendelkezik térlátással, szenzorok segítségével határozza meg a távolságokat, melyekből következtetéseket von le. A probléma megoldására egy kamerát alkalmazunk csak, mely önmagában nem képes mélységi adatokat biztosítani. Speciális eljárások segítségével közelíteni lehet a távolságokat. Érdekes megemlíteni a „DisNet” [27] nevet viselő neurális hálózatot, mely relatíve pontos becslést tud adni a befoglaló négyzetbe foglalt objektum távolságára. A hálózat nem térbeli becslést végez, hanem az objektumot meghatározó jellemzők méretével végez egy közelítést, mely az objektum teljes egészére határoz meg egy távolságot. A feladat megoldására viszont további mélységi adatokra volna szükség, melyet a következőkkel lehet magyarázni; Példának vegyünk egy autót, melyet a neurális háló azonosított és meghatározta az autót magába foglaló téglalapot. Ha a téglalapon belül felvesszük az autó pozícióját magába foglaló térbeli, geometriai testet, akkor egy téglatestet kapunk, melyet a 13. ábra szemléltet.



13. ábra. Objektumot reprezentáló téglatest (piros) és 2D befoglaló téglalap (kék)

A 3D hexaéder mérete viszont az autó elhelyezkedésétől függően a 2D képen változik. A 2D-ben reprezentált hexaéder egy poligonnal feleltethető meg, mely pontjait X és Y koordinátákkal lehet reprezentálni. Míg a vasúti pálya, és úrszelvény 3D elhelyezkedését egészen pontosan meg lehet határozni, addig az autót reprezentáló téglalapot 3D pontjait nem lehet egy 2D képről meghatározni, ezért a befoglaló téglalap nem fogja megfelelő pontossággal szemléltetni a jármű térbeli elhelyezkedését. A DisNet csak a jármű középpontjának kamerától vett távolságát fogja megbecsülni, mely bizonyítottan nem elegendő az objektum térbeli elhelyezéséhez, ezért a térben végzett akadályészlelés elméletét elvetem. Továbbá érdemes megemlíteni, hogy a perspektív szemlélet okozta torzulás megfigyelhető a hexaéderen, mely során a befoglaló négyzetbe tartoznak olyan területek is, melyek nem képezik az objektum foglalta teret.

A problémát ezért geometriai úton közelítem meg, mely során egy durva becslést végzek az objektum formájára. Az úrszelvény pontjait 2D koordinátákként kezelem, melyek egy poligont alkotnak. A vizsgálat során azt figyelem, hogy az objektum becsült alakja a poligonon belül található-e vagy sem. Ha az objektumot leíró alakzat kis része is a poligon területén belül helyezkedik el, az objektumot akadállynak tekintem.

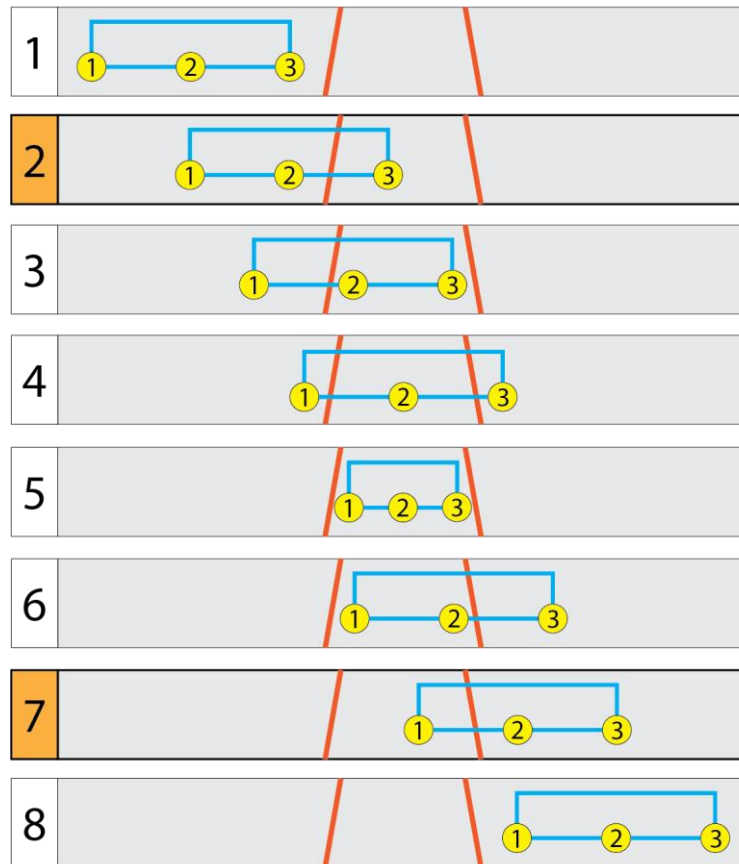
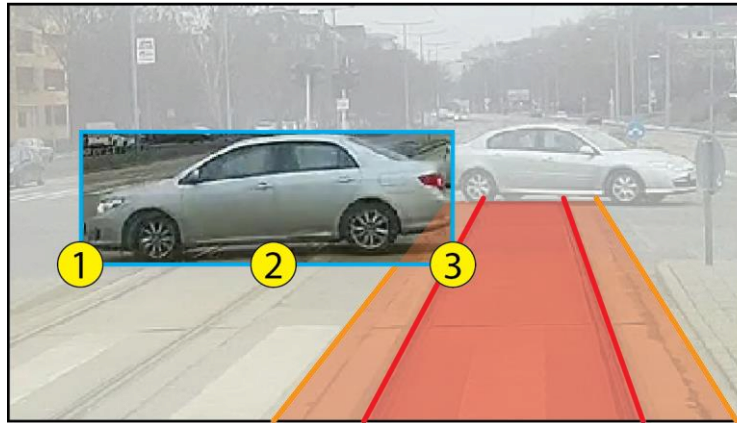


14. ábra, Befoglaló téglalap átmenete ellipszisbe

Az 13. ábra során megfigyelhető, hogy az autót leíró téglatest párhuzamosai a perspektív nézet miatt egy pont felé, a „vanishing point” felé tartanak. Ez a jelenség felerősödik, minél inkább távolabb helyezkedik el a horizont közepétől az objektum. Megfigyelhető, hogy minél inkább a horizont középpontja felé helyezkedik el az objektum, annál inkább feleltethető meg a befoglaló téglalap a téglatest 2D-ben vett alakjának. Minél inkább a kép bal vagy jobb alsó sarka felé helyezkedik el az objektum a horizonttól és a kamera képét felező függőleges vonaltól, annál inkább egy, a befoglaló téglalapot kitöltő ellipszisnek feleltethető meg az objektum. A befoglaló téglalap és az ellipszis közti átmenet az objektum elhelyezkedésétől függ, melyet a 14. ábra szemléltet. Fontos megemlíteni, hogy minél nagyobb az érzékelt objektum és minél közelebb helyezkedik el a kamerához, annál pontatlanabb az ellipszis által meghatározott alak. A probléma kiküszöbölésére a kép egy részét maszkolom, melyen belül történik az objektumdetektálás. Az űrszelvényhez közeli objektumok jellemzőit a YOLO képes felismerni, ezzel egy kisebb befoglaló téglalapot alkotva a kitakart objektumról. A kisebb méretre a becslés tapasztalatok alapján jól működik.

4.4.1. ÉSZLELÉSI FOLYAMAT

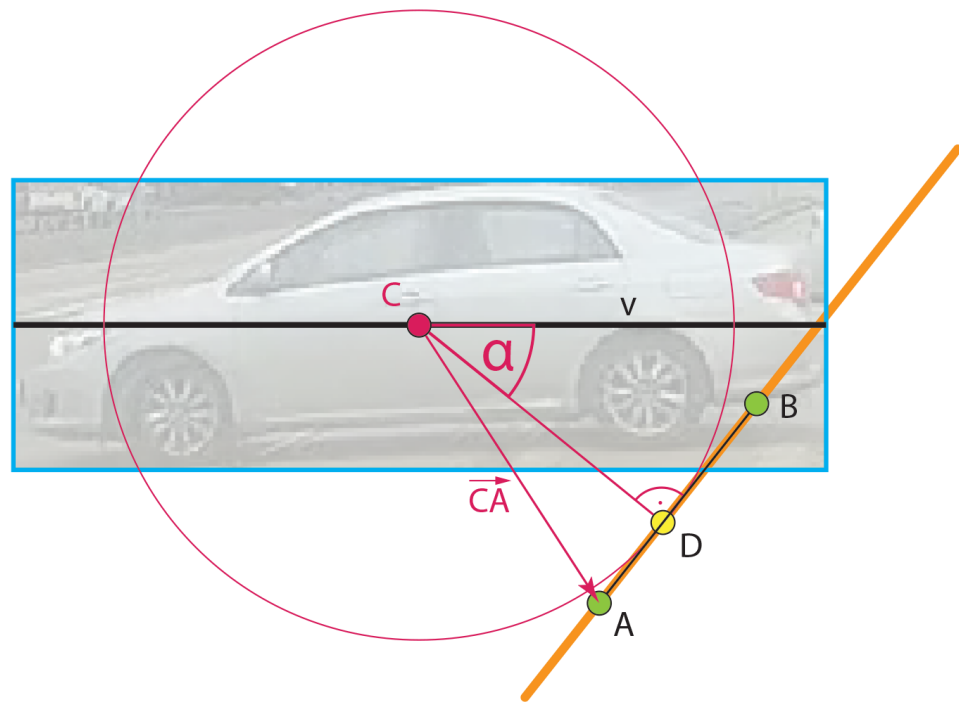
A vizsgálatot minden érzékelt objektum esetén elvégzem, ha szükségessé válik többször is. Az alapötlet szerint minden olyan objektum akadályt képez, aminek a becsült alakja részben vagy teljesen az űrszelvényt behatároló poligon területén belül helyezkedik el. Az űrszelvény régiói szerint felosztva vizsgálom az alakzat pozícióját. Mivel a pálya a horizont alatt helyezkedik el és a horizont fele közeledve egyre nagyobb távolságokat mérünk, ezért az objektumot leíró alakzatot is lentről felfele szükséges vizsgálni.



15. ábra, Objektum helyzetének esetei az űrszelvényhez viszonyítva

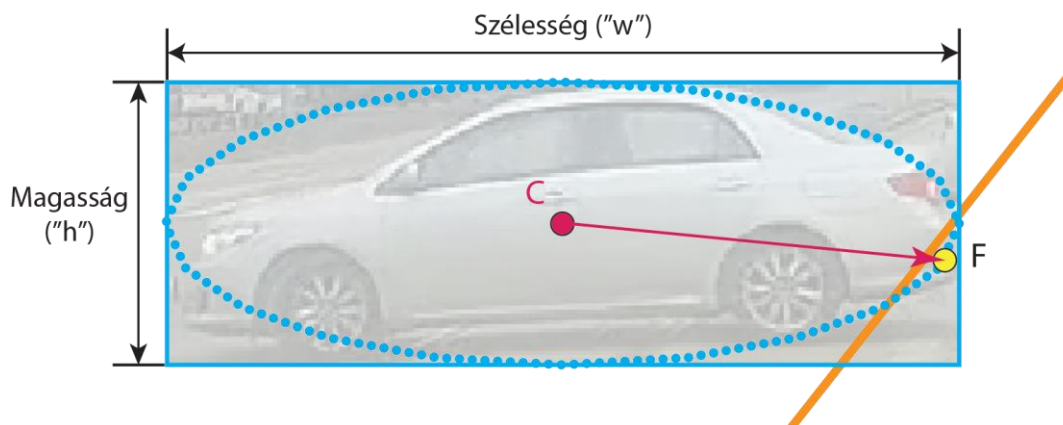
A vizsgálat során definiálok három, az objektumhoz tartozó pontot, melyet a 15. ábra szemléltet. A pontok X tengelyi értékét a befoglaló téglalap két széle (1 és 3) és a közepe adja (2). Az Y értéket a befoglaló téglalap alsó szakasza adja. Amennyiben további vizsgálatot szükséges folytatni, úgy az Y értékét minden további régió alsó szakasza képezi. Az alapfeltevés a következő; Ha a szélső és középső pontok az űrszelvényen kívül helyezkednek el (15. ábra 1. és 8. eset), akkor az objektum nem

képez akadályt, a vizsgálatot nem folytatom tovább. Ha a középső pont (2) az űrszelvényen belül helyezkedik el (3., 4., 5. és 6. eset), akkor az objektum akadályt képez, a vizsgálatot leállítom és a maradék objektumokat kiértékelésére nem lesz szükség. Amennyiben az egyik szélső pont az űrszelvényen belültre esik (2. és 7. eset, úgy az objektumot megközelítő alakot kell vizsgálni.



16. ábra, Objektum legközelebbi pontjának meghatározása az űrszelvényhez 1/2

A vizsgálat első felét a 16. ábra szemlélteti. Az űrszelvényt a régió határai szerint van értelmezve. Esetünkben a szakasz kezdetét az „A” pont, végét pedig a „B” jelenti. Meghatározom a befoglaló téglalap középpontjából („C”) az „A” pontba mutató CA vektort. Kiszámításra kerül a CA vektor vetülete az AB szakaszra nézve. A „C” pont és vetület végpontja („D”) határozza meg a CD szakaszt, mely merőleges az AB szakaszra. A „C” középpontú, CD szakasz hosszú sugárral bíró kör érintője az AB szakasz, mely egyben meghatározza a legrövidebb távolság hosszát az objektum középpontjától az űrszelvény szakaszig. Az objektum középpontja által meghatározott vízszintes („v”) és a CD szakasz által bezárt α szög megadja az űrszelvény szakaszhoz legközelebb eső pont irányát.



17. ábra, Objektum legközelebbi pontjának meghatározása az űrszelvényhez 2/2

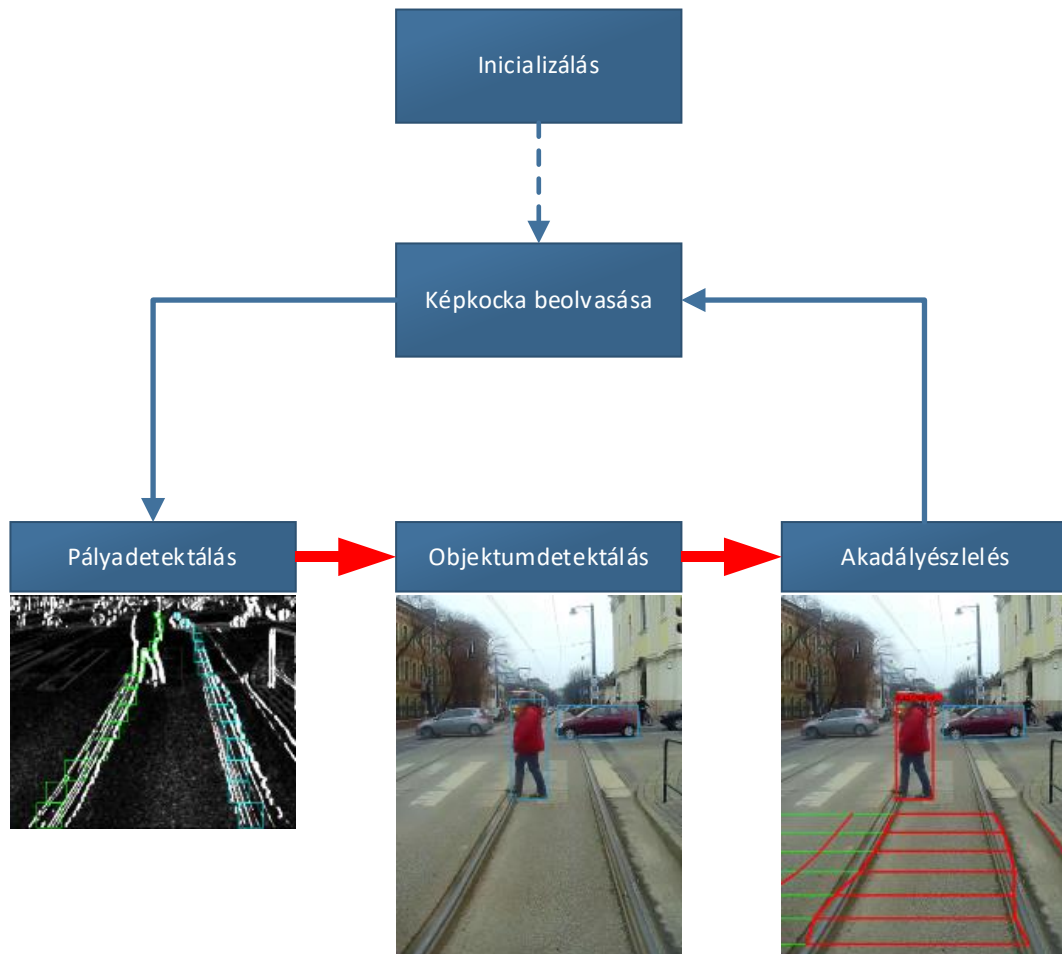
Az azonosítás második felét a 17. ábra szemlélteti. A korábbiakban meghatározott α szög segítségével megkaptam az ellipszisen elhelyezkedő, űrszelvénytárhoz legközelebbeső pont koordinátáit. Ha vesszük a „C” középpontú, $w/2$ sugarú kört, akkor megállapítható, hogy a kör hasonló a korábbi CD sugarúhoz viszonyítva. Ha vesszük a $w/2$ sugarú kört és megpróbáljuk az objektumot befoglaló téglalapon belülre helyezni, akkor egy ellipszist kapunk, mely határai a téglalap szélei (17. ábra, szaggatott kék vonalú ellipszis). A kapott ellipszis feleltethető meg az objektum alakját meghatározó közelítésének, mely a 4.4 fejezetben került ismertetésre. Az ellipszisen elhelyezkedő, űrszelvénytárhoz eső legközelebbi pont („F”) koordinátáit a következők szerint számítom (17. ábra):

Az „F” pont X koordinátáját az α szög koszinusza és a téglalap szélességének felével vett szorzata adja. Az „F” pont Y koordinátáját hasonlóan az α szög szinusza és a téglalap magasságának felével vett szorzata adja. Az „F” pont koordinátáihoz még hozzá kell adni a „C” középpont koordinátáit, így megkapjuk a képre vetített globális koordinátákat. Utolsó előtti lépésként meghatározom, hogy az objektumot reprezentáló téglalap mely sarkához helyezkedik el közelebb az „F” pont. A 4.4 fejezetben a 14. ábra szerint ismertetett átmenetet képezek a kiválasztott sarokpont és az „F” pont között.

Az átmenetképzés során kapott végső koordinátáról eldöntöm, hogy a „C” középpont és űrszelvénytárhoz túl helyezkedik-e el. Ha igen, akkor az objektum akadályt képez, a vizsgálatot nem folytatom le más objektumokra sem. Ha viszont a „C” pont és űrszelvénytárhoz közt helyezkedik el, akkor a következő régióban is megismétlem a vizsgálat lépéseit mindaddig, amíg nem érem el a „C” pont magasságát. Mivel a közelítő ellipszis körvonala a „C” pont magasságától felfelé haladva egyre távolodik a pályától, ezért a vizsgálatot nem érdemes tovább folytatni.

4.5. RENDSZER ARCHITEKTÚRA

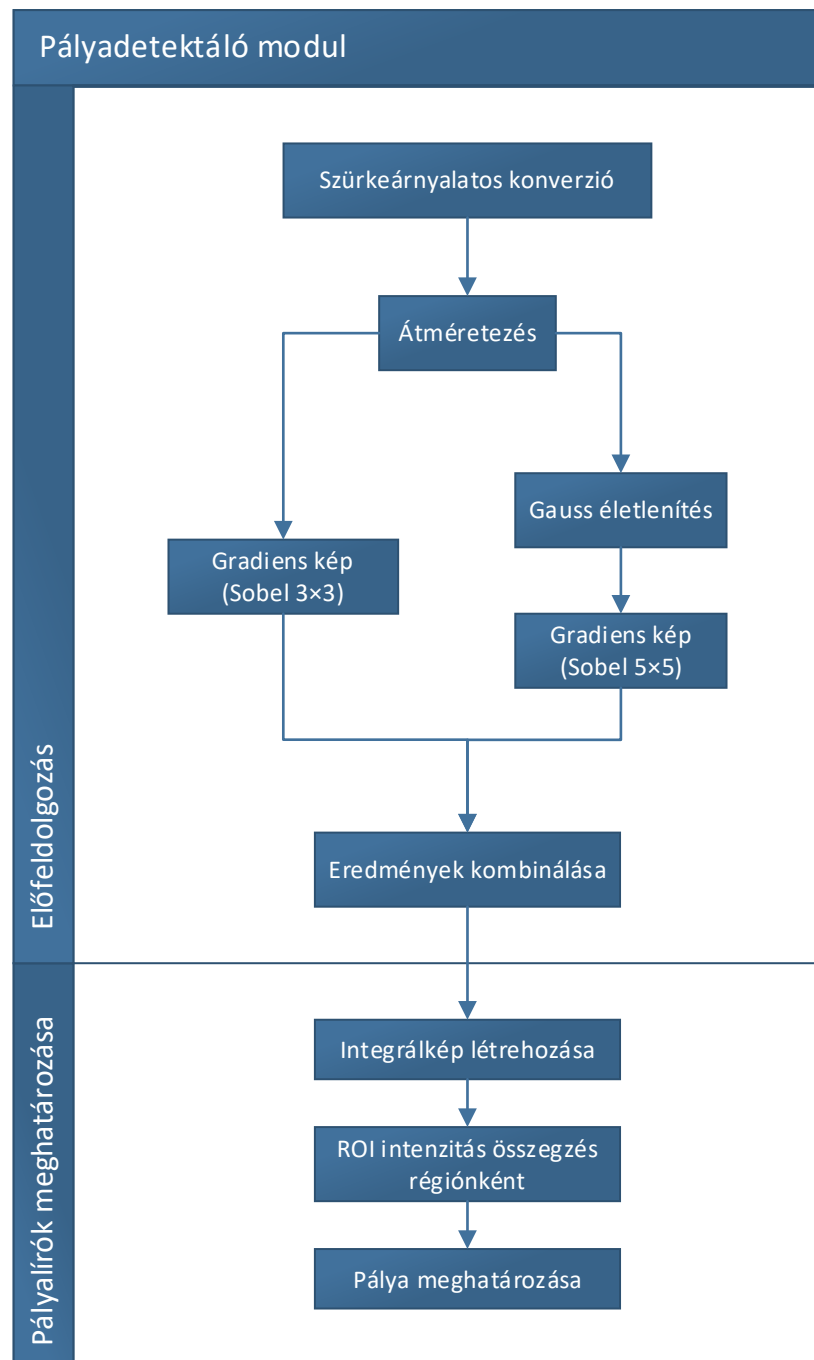
A rendszer felépítése a 18. ábrán látható. Első lépésben szükséges a rendszer inicializálása, mely során megtörténik a szükséges memória területek allokációja rendszer- és video memóriában. Az első képkocka felhasználásával meghatározásra kerül a pálya keresésének kiindulási pontját és az objektumdetektáláshoz szükséges maszkolási terület. Az első képkockát követően történhet meg az objektumdetektálás és akadályészlelés. A 18. ábra szemlélteti a rendszer működését folyamatok szintjén.



18. ábra. Rendszer működésének folyamatábrája

Az inicializálást követően a rendszer iteratívan, minden képkockát feldolgoz. A pályadetektáló és objektumdetektáló modul gyakorlatilag függetlenek egymástól és ráadásra célszerű lehet külön futtatni őket, viszont a gyakorlat azt mutatja, hogy nem érdemes párhuzamosítani. Mivel a folyamatok nagy része ugyan azon GPU-t használja, ami folyamatos kihasználtság alatt áll, ezért, ha közel egyidejűleg szeretnék CUDA kerneleket indítani, akkor ezen kernelek ütemezése nagyobb késleltetéshez

vezetne, mely rontaná a rendszer teljesítményét, ezért maradok a modulok egymás után történő futtatásánál.

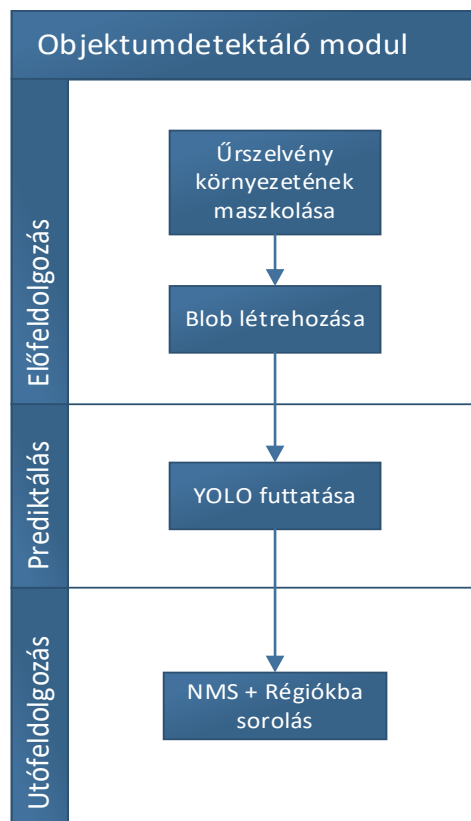


19. ábra. Pályadetektáló modul folyamatábra

A pályadetektálás során a kép feldolgozása (19. ábra, előfeldolgozás) teszi ki a modul legnagyobb részét. A képfeldolgozási lépések során egy művelet hajtók végre a kép összes pixelén. Ugyan azon művelet végrehajtható minden pixel esetén függetlenül egymástól. A végrehajtás során minden pixelre egy külön szál fogja

elvégezni az adott eljárás műveletét. A szálak nyújtotta párhuzamosítási lehetőségek maximális kihasználásáért az eljárásokat CUDA környezetben implementáltam. A szürkeárnyaltos konverziótól egészen a *Sobel* operátorok eredményének kombinálásáig a folyamatnak nem szükséges elhagynia a GPU környezetét. Az integrál kép CUDA környezetben történő megvalósítására egy hatékony módszert mutatott be *Berkin Bilgic, Berthold K.P. Horn és Ichiro Masaki* az „*Efficient integral image computation on the GPU*” [28] cikkükben. A gyakorlat során arra a következtetésre jutottam, hogy a CPU-n futtatott integrálkép sebessége elegendő, a bonyolultabb GPU implementációval elért teljesítmény növekedés nem számottevő.

A régiók vizsgálatának további gyorsításának érdekében az integrálképen elvégzem a csúszó ablakokban számított integrálösszegek meghatározását GPU-n. A folyamat egy $N \times M$ méretű, kétdimenziós tömböt eredményez, ahol „N” a régiók száma és „M” az intenzitásösszegek száma, mely régiónként eltér. A pályaleírók meghatározása során csak az intenzitásösszegek között kell maximum keresést végezni a 4.2.2 alfejezetben ismertetett módon.



20. ábra. Objektumdetektáló modul folyamatára

Az objektumdetektálás (20. ábra) első lépésében megtörténik a képkocka maszkolása, majd a kapott képet átméretezem és létrehozom belőle a neurális háló input tenzorával kompatibilis blob-ot. Ezt követően futtatom a neurális hálót. A prediktálást követő eredménymátrixon elvégzem a küszöbölést, melyt követően csak azokkal a predikciókkal foglalkozok, melyek elérték a megadott küszöbértéket. Az objektumjavaslatok leginkább egymást fedve jelentkeznek különböző konfidencia értékkel. A tényleges objektumot reprezentáló predikció kiválasztása a nem-maximumok elnyomásával történik, mely a pedikciók alkotta javaslatklaszterek közül kiválasztja a legmagasabb konfidencia értékkel rendelkező predikciót. Javaslatklasztereket azok a predikciók alkotnak, melyek befoglaló téglalapjai egy megadott küszöbérték felett fedik egymást. Két befoglaló téglalap fedésének arányát meghatározó számot Jaccard index-nek [29], másnéven Jaccard hasonlósági együtthatónak nevezik. A folyamat végén előállnak a detektált objektumok, melyeket a pályadetektálás során ismertetett (4.2.2 fejezet) régiókba sorolok, gyorsítva későbbi vizsgálatukat.

Az utolsó lépésben megtörténik az akadályészlelés a 4.4.1 fejezetben ismertetett módon, mely során nem végzek párhuzamosítást.

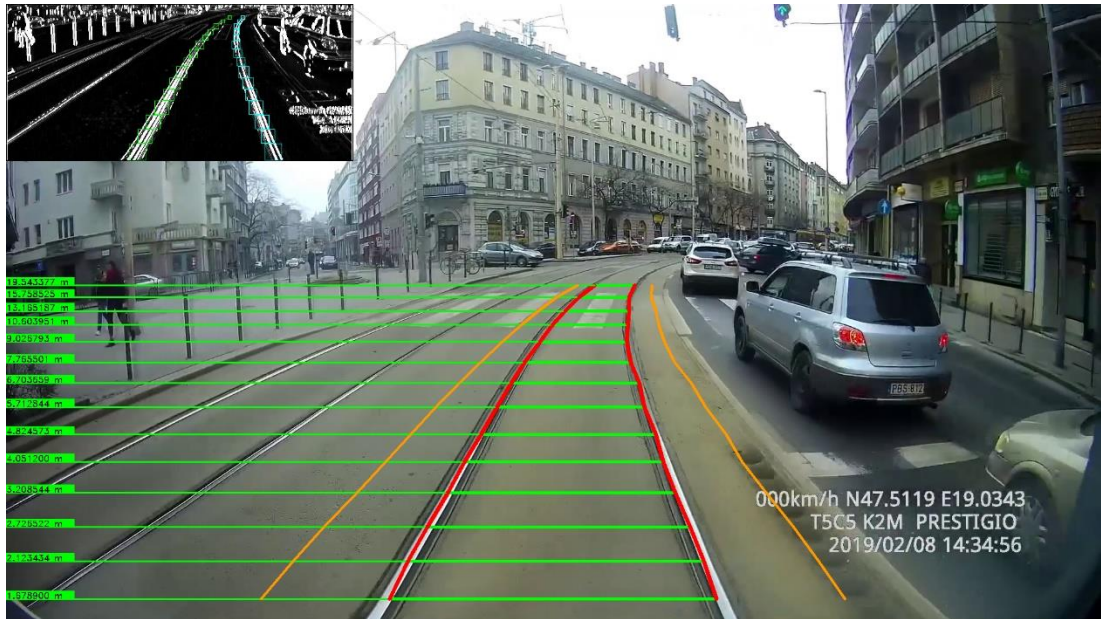
5. EREDMÉNYEK ÉS ÉRTÉKELÉSÜK

A kitűzött feladattal egy olyan problémára kerestem megoldást, mely megfelelő képfeldolgozási technikák használatával és neurális hálózat segítségével képes elvárt módon azonosítani valós időben fellépő akadályokat. A fejlesztés során sikerült megvalósítanom egy olyan rendszert, mely a kitűzött céloknak funkcionalitásban és sebességben is leginkább eleget tesz. Az elért eredmények bemutatása során ismertetem azokat a problémákat, melyek negatívan befolyásolják a rendszer adott területen nyújtott teljesítményét.

Az elért eredményeket három pontba foglalva ismertetem. Az első, a sínpárok azonosításának pontossága, mely befolyásolja a detektálható objektumokat és elhelyezésüket. A második az akadályt képző objektumok azonosításának pontossága, melyek elhelyezésének a sínszalak nyújtanak alapot. Az utolsó pontban a rendszer végrehajtási idejét értékelem egy képkockára vetítve. Az eredményeket a fejlesztés során felhasznált felvételek közül egy megadott felvétel [30] klipjein futtatva rögzítettem és értékeltem ki az általam erre a célra fejlesztett egyszerű szoftverek segítségével.

5.1. PÁLYADETEKTÁLÁSI EREDMÉNYEK

Kiértékelés során csak a sínszalakat azonosító modult futtattam számos videóklipen. A klipek megválasztása során leginkább a sínszalak környezetében lévő, különböző felépítményeket vettem alapul. A 2.1 fejezetben ismertetett számos útburkolati típus mellett egy magasvágányos pályakialakítást, illetve egy alagúton történő áthaladást értékeltem. A különböző útburkolatok különböző méretű és típusú illesztési hibákkal rendelkeznek a felvételeken, melyek néhol erősebb intenzitással bíró körvonalakkal rendelkeznek a sínszalakhoz képest. Ez a jelenség elég erős negatív hatást tud gyakorolni a pályaaazonosítás pontosságára. A magasvágányos pályaszakasz jellegzetes karakterisztikával rendelkezik, melyre a hasonló rendszerek 3.2 fejezetben ismertetett kutatások és fejlesztések építettek. Az alagúton történő áthaladás során azt vizsgáltam, hogy milyen hatást gyakorol a detektálásra a világos környezetből sötétbe és sötétből világosba történő átmenet. A 21. ábrán a pályadetektálás eredménye látható.



21. ábra, Pályadetektálás betonlemezes pályaszakaszon

Minden klipet, minden képkockát egyesével elemeztem, hogy eldöntsem, a pályát helyesen sikerült-e azonosítani vagy sem. A kereső algoritmus már eleve egy kellően szűrt képnek a régiók szerint felbontott sorait teljes mértékben feldolgozva kapja, ahol szinte csak maximumokat kell helyesen meghatározni. Ebből a tényből kiindulva egy nagyon egyszerű értékelési rendszerrel álltam elő. Minden olyan képkockát helyesnek tekintek, ahol a csúszó ablakok teljesen fedik a pályát. Mivel a csúszó ablakokon belül a sínszálak leginkább átlósan helyezkednek el, ezért az algoritmus keresés során találkozhatott olyan összegekkel is, melyek nem csak a pálya élei által generált intenzitást tartalmazzák, hanem a környezeti zajt is, mely nem pontosan tükrözi a sínszál valódi helyzetét. Az említettek miatt egy kisebb tűréshatárt vezettem be, melyen belül még helyesnek mondható a detektált pálya. Minden más esetben helytelennek véltem az eredményt. A vizsgált pályaszakasz hosszát a fejlesztés során határoztam meg, mely 18 méteres távolságot jelent. A 3. táblázaton a 18 méteres kereső szakaszon rögzített eredmények láthatók. A feltüntetett felvételekhez egy-egy feldolgozott kép és eredeti nyers képre vetített eredményét szemlélteti a 22. ábra. A fejlesztés során rögzítésre került egy körülbelül 22 méter hosszú szakasz is, melynek eredményei a 4. táblázatban szerepelnek.

Felvétel	Összes képkocka	Helyes	Helytelen	Pontosság
betonlemez_14_18	480	480	0	100%
magasvagany_32_44	385	385	0	100%
bazaltkockocka_17_44	441	434	7	98,41%
aszfalt_11_09	450	433	16	96,44%
alagut_53_06	510	477	33	93,52%
Összegezve:	2266	2210	56	97,67%

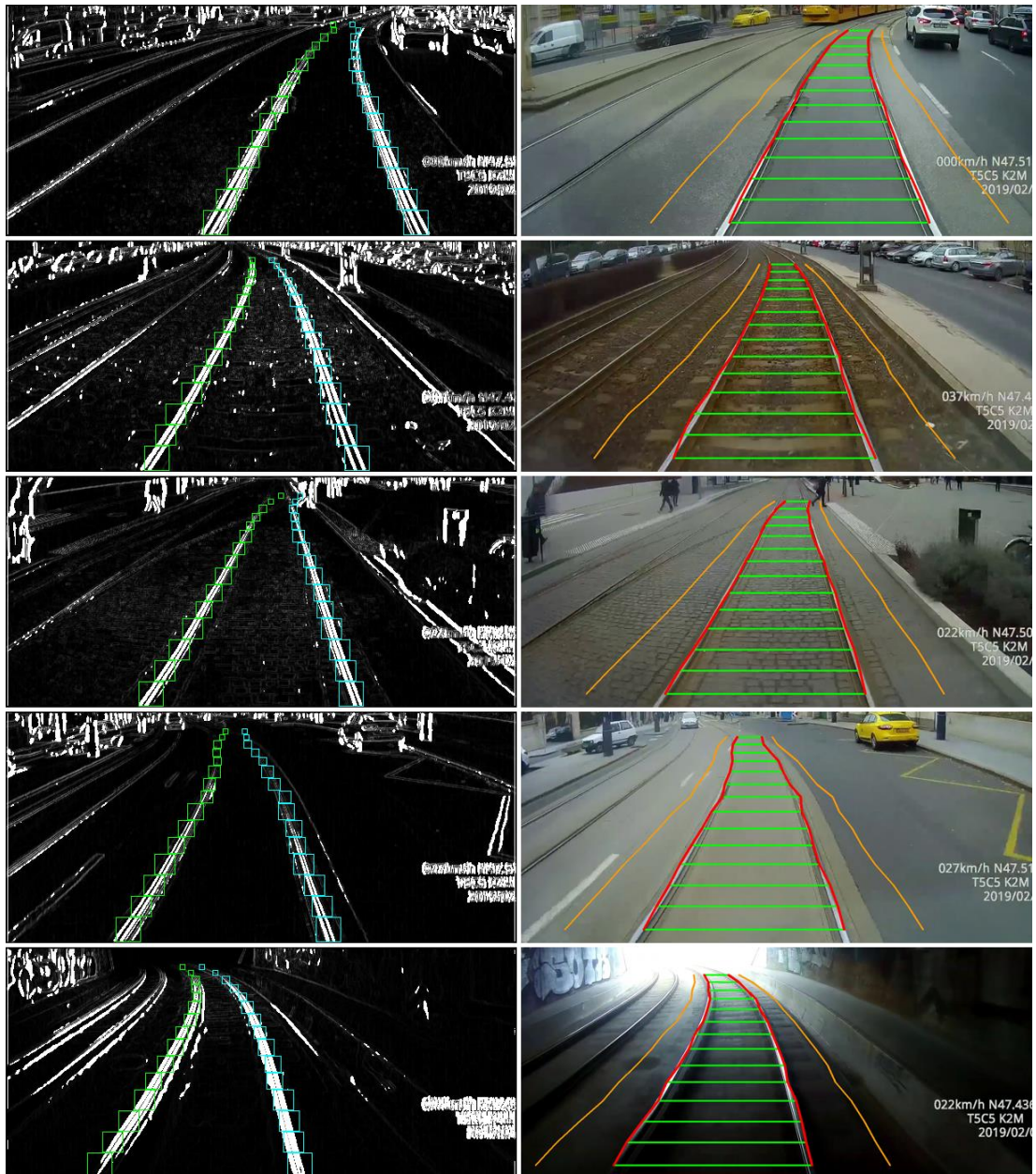
3. táblázat, 18 méteres távolságot bezáró vizsgálat eredményei

Az eredményeken jól látszik, hogy a pályaaazonosítás az esetek több mint 97%-ban helyesnek mondható. Ez az érték kifejezetten jónak mondható, mivel a felvételek minőségén sokat rontott a tömörítés. Két felvételen is 100% pontosságot értünk el, mely annak a ténynek tudható be, hogy ezeken a sínszálak jellegzetes körvonalait igen jól ki lehet venni. A sínszálakat körülvevő beton felülete nagyon kis mértékű intenzitásváltozással rendelkezik, kevés zaj van körülötte, mely egyértelműen meghatározza a közre zárt sínszálakat. A magasvagány esetében a sínszálak környezete nagyobb zajjal bír a zúzottköveknek és aljaknak köszönhetően, viszont a nem csak a sínkorona látszódik, hanem a sínszál teljes egésze, mely több erős élen nyilvánul meg, ezzel egyértelműen behatárolva a pályát.

A bazaltkőkockás pályaszakasznál fellelhető egy kevés pontatlanság, melyet nem a sínszálakat körbevevő bazaltkőkocka intenzitása okozza, hanem a felvételen átsétáló gyalogos. A gyalogos lábai erősebb élekkel rendelkeznek a tompa, egyszínű háttértől elkülönülve, melyeket az algoritmus mindaddig lekövetett, amíg a gyalogos a sínszál közvetlen közelében tartózkodott. Érdemes megemlíteni, hogy a probléma csak addig állt fenn, amíg a gyalogos a tényleges úrszelvényen belül tartózkodott. Az akadályészlelés eredményein látható, hogy az algoritmus a kamerához legközelebbi akadálytól tovább már nem számítja a pályát keresztező többi objektumot, ezért a pályadetektálás az első akadályt követően már nem szükséges figyelembe venni.

Az aszfalttal borított szakaszon szereplő 16 helytelen detektálás a felvétel rossz minőségének tudható be. A tömörített felvételen a pályaszakasz távolabbi része nem volt látható se az ember, se az algoritmus számára. Az útburkolat színe teljesen megegyezett a sínkoronával, ezért az algoritmus teljesen jogosan egy másik erős élen folytatta a keresést. Ez a pontatlanság a pályaszakasz végén jelentkezik, melyre a

legegyszerűbb megoldás nem tömörített felvételek használata, ami a fejlesztés során számomra nem állt rendelkezésre.



22. ábra, Felvételek feldolgozott és nyers képei, fentről lefele haladva: betonlemez, útburkolat, magasvágány, bazaltkőköcs burkolat, aszfalt és alagút

Az alagúton való áthaladás során meglepő módon egész jó eredményeket értem el. Az algoritmus nagyon jól reagált a fényviszonyváltozásokra. A világosból sötétbe történő átmenetnél a távolabbra eső csúszó ablakok is nagyon pontosan tudták azonosítani a gyenge éleket. Ez betudható annak, hogy a keresett pályaszakasz vége sose esett bele a teljesen sötét alagútba, a kamera folyamatos alkalmazkodása az

alagúthoz lehetővé tette, hogy valamennyi gyenge él megmaradjon. A sötétből világosba történő átmenetnél is hasonlókat lehet tapasztalni. Az egyedüli pontatlanságot a keresett pályaszakasz végén lévő ív jelentette. A sínszálak környezetében lévő nagyobb zaj eltérítette a csúszó ablakokat. Ez a jelenség leginkább 15-16 méteres távolságban jelentkezett, mely a keresett pályaszakasz végén található. Ezen a távolságon a csúszó ablakok nagyon kis területeket vizsgálnak, melyekben a zaj sokkal nagyobb negatív hatással bír, mint a kisebb távolságok esetében. A legjobb megoldást az jelentené, ha megnövelném a gradiens kép felbontását, hogy több részlettel tudjon dolgozni a rendszer.

Az elért eredmények ismételten csak arra utalnak, hogy a rendszer zajos környezetben is nagy pontossággal képes helyesen azonosítani a villamos pályáját, viszont ez csak a korábban említett 18 méteres távolságig igaz. Továbbá fontos, hogy a zaj ebben az esetben a fényviszonyok okozta színváltozások, árnyékok tekintetében értendő, melyek nem okoznak kirívóan hatalmas intenzitás változást. 18 méter felett a gradiens kép kis felbontása miatt a rendszer nagyon érzékeny a zajra, melyen túl nem célszerű további vizsgálatokat folytatni. Az így elért eredmények a 4. táblázatban láthatóak. A korábban ismertetett hibák minden felvételre ugyan úgy igazak 18 méter felett.

Felvétel	Összes képkocka	Helyes	Helytelen	Pontosság
magasvagany_32_44	385	385	0	100%
bazaltkokocka_17_44	441	421	20	95,46%
aszfalt_11_09	450	410	40	91,11%
betonlemez_14_18	480	410	70	85,42%
alagut_53_06	510	410	100	80,39%
Összegezve:	2266	2036	230	90,48%

4. táblázat, 22 méteres távolságot bezáró vizsgálat eredményei

Ahogy azt látni lehet, szinte minden felvétel pontossága csökkent 5-10%-al. Minden esetben az utolsó, legtávolabbi csúszó ablak vétette a legtöbb hibát. Érdekes kiemelni a betonlemez felvételt, ahol az ívelt pályaszakaszon való haladáskor a szemből érkező villamos oldallemezei nagyobb intenzitással rendelkeztek, mint a pályaszakasz, ezért a bazaltkőkockás gyalogos esetéhez hasonló módon, a megnövelt távolsággal vizsgáló algoritmus hibásan követte le a tényleges pályát.

A 3. fejezet során bemutatott hasonló rendszerek közül kiemeltem kettőt ([8] és [10]), melyek hasonló formában hajtják végre a pályadetektálást. A [10] megvalósítás közel 100 méteres távolságban képes azonosítani a pályát, mely egy jelentősen nagyobb távolság a saját rendszeremhez képest. Ennek oka, hogy a vasúti szerelvényen magasabban helyezték el a kamerát, mint a villamoson, így nagyobb rálátással rendelkezik a pályára. Továbbá nem érdemes elhanyagolni a tényt, hogy a nagyvasúti sínszálak élei sokkal erősebbek, mint az útburkolattal egy szinten lévő villamos pálya, ezért nagyobb távolságban jobban kivehetőek lesznek. A rendszerem tesztelése során mind 18 méter felett és alatt is nagyvasúti magasságányon 100% pontosságot ért el az általam megvalósított pályadetektálás, mely közel azonos étékkel bír a [10] rendszerrel szemben. A [8] megvalósítás során kiemelték, hogy a rendszer jól kezeli az árnyékos környezetet, ami a Gradiens képen való élkeresésnek köszönhető. Rendszerem hasonló megvalósítással bír. A teszteredményeknél, az alagúton történő áthaladás során, különböző fényviszonyokban 93,5% pontossággal teljesített a rendszerem, melyből levonható, hogy véleményem szerint közel azonos képességgel bír a két rendszer.

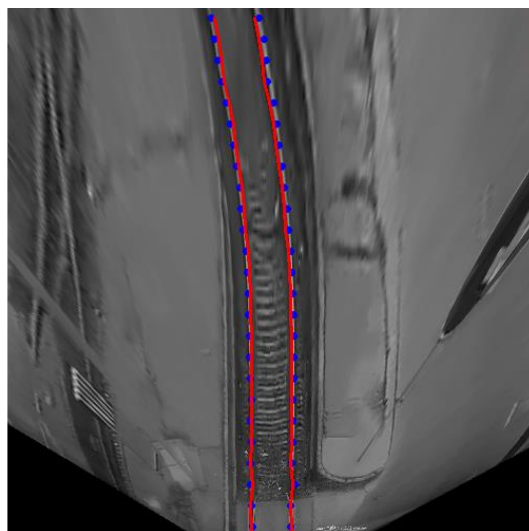


23. ábra, Gradiens és eredeti kép, fentről lefele; elszórtan hóval és vízzel fedett pályaszakasz

A rendszert teszteltem különböző időjárási jelenségek tekintetében, nevezetesen havazást és esőzést követő pályaszakaszokon. A rendszer eredménye teljes mértékben függ a sínszálak környezetében fellépő zaj mértékétől. Az eső okozta kis méretű állóvíz rossz pályaburkolaton nagyon könnyen eltéríti a csúszó ablakokat, mert a foltos felületen gyakran jóval erősebb élek formájában jelenik meg az állóvíz, mint a sínszálak, ezért ezen jelenségeket nagyon intenzív környezeti zajnak tekintem. A 23. ábra szemlélteti ezen intenzív környezeti zajokban előfeldolgozott képet. Az ábrán csak a 3×3-as gradiens kép eredményét szemléltetem, melyen látni, hogy nagyobb távolságban a pályaszakasz már szabad szemmel sem kivehető. A 3×3 és 5×5 gradiens kép kombinálása értelmezhetetlen. A pályadetektálás nem végezhető el, ha a víz teljes mértékben elfedi a pályát vagy az állóvíz kellő távolságban fedi a sínszálakat. Ugyan ez elmondható a havas vagy kevés hóval, esetleg hófoltokkal körbevett pálya kapcsán. A 23. ábra felső képén látható, hogy a hófoltok kevésbé értelmezhető képet eredményeznek, mint az alatta lévő vizes pályaszakasz.

A pályadetektálási eredmények zárása előtt érdemes említést tenni a folyamatot megvalósító modul egy korai fejlesztési változatát, mely más szemszögből igyekezett megoldani a problémát. A pályadetektálási folyamat lépései nem sokban különböztek az aktuális megvalósítástól, viszont egy nagy különbséget jelentett, hogy milyen szemszögből volt feldolgozva a képkocka. A jelenlegi perspektív szemlélettől eltérve korábban a [10] cikkhez hasonlóan fentről tekintett lefele a pályára¹⁸. Hasonló módon első lépésben létre lett hozva a kép szürkeárnyaltos konverziója. Ezután egy perspektív transzformációt hajtottam végre, mely során előállt a pálya felülnézeti képe. Erre azért volt szükség, hogy a pálya sínszálai ne a horizont egy pontjába tartsanak, hanem párhuzamosokként jelenjenek meg. Következőnek előállítottam a gradiens képet és median szűrő alkalmazásával csökkentettem a zajt. Mivel a gradiens képen a sínszálak élei párhuzamosok formájában voltak jelen, ezért a képet hasonló módon régió sávokra bontottam és minden régiónak előállítottam az élerősség tömbjét. Az élerősség tömb minden eleme megadja, hogy az adott régió magasságától függően a vertikális élek milyen erősséggel vannak jelen a megfelelő indexű helyen. Ez az intenzitásösszegző módszernél egy precízebb sínszálkeresést eredményez. Végezetül az élerősségtömbökön régióként lentől felfele haladva maximumkereséssel meghatároztam a sínszálakat (24. ábra, jobb oldali kép).

¹⁸ Általános kifejezéssel élve „madártávlatból” (bird’s eye view)



24. ábra, Felülnézeti pályadetektálás, balról jobbra, eredeti kép és detektált pálya

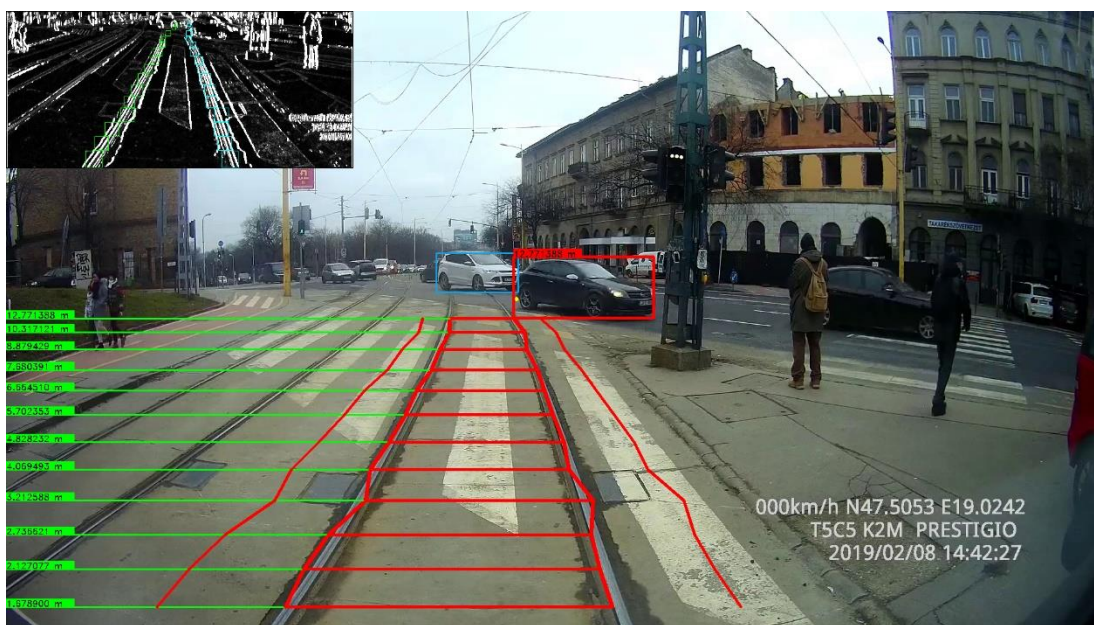
Felmerül a kérdés, hogy ha precízebb sínszáskeresést tett lehetővé az élerősségtömb használata, akkor miért nem lett alkalmazva a jelenlegi pályadetektálási eljárás során? A válasz teljesítményproblémákra és adatvesztésre vezethető vissza. A teljesítmény problémát a 3.3 fejezet [10] rendszer technikájának értékelése során ismertettem, melynek eredményét a pályadetektálási modul ezen korai változata nyújtotta. Az adatvesztés a teljesítményproblémát jelentő a transzformáláshoz vezethető vissza, mivel lényegében a pálya kamerához közelebb eső részét „összenyomjuk”, a távolabbat pedig „széthúzzuk”. Az összenyomás során adatvesztés történik, a széthúzáskor pedig újak létrehozása, ezért a folyamat nem nyújt konzisztens pályareprezentációt. Végezetül élerősségtömböket csak horizontális, vagy csak vertikális élekből lehet előállítani. Amennyiben egy olyan él szerepel ívelt pályán, amely erősen vertikális irányban van jelen, úgy az helytelen pályadetektálást eredményezhet a gyenge sínszálak miatt. Az említett problémák miatt maradtam a kamera szemszögéből történő pályafeldolgozásnál.

Összességében elmondható, hogy 18 méter alatt a pályadetektáló modul különböző fényviszonyokban és különböző útburkolattal rendelkező villamos pályát közel 98%-os pontossággal képes helyesen azonosítani. A hibatartomány leginkább a pályaszakasz végére tehető, ahol a kis felbontás miatt, leginkább zajos környezetben nagyon nehéz helyesen kiemelni a sínszálakat. Az előfeldolgozás során végrehajtott átméretezéskor célszerű egy optimális méretre csökkenteni a képet, mely még nem okozna jelentős teljesítmény veszteséget.

5.2. AKADÁLYÉSZLELÉSI EREDMÉNYEK

Az eredmények rögzítése olyan videóklipen történt, melyeken valamelyest különböző forgalmi helyzetek vannak jelen. Mivel a legtöbb villamos pálya városban belül található, ezért a felvételek városi környezetet mutatnak be és így a rendszert csak útkereszteződéseken és gyalogosátkelőhelyeken tudtam érdemben tesztelni.

A járművet keresztező akadályok szempontjából próbáltam többféle olyan helyzetet keresni, melyben különböző irányban és távolságban haladnak el a járművek és gyalogosok a kamera előtt. A rendszert így három felvételen teszteltem. Az első¹⁹ felvételen a gyalogosok és járművek körülbelül 7 és 18 méter közötti távolságban keresztezik a járművet, a villamospályához képest merőlegesen haladva. A keresztező objektumok távolságai ideálisnak mondhatók a pályadetektálás során meghatározott maximális detektálási távolság szerint, mely 18 méter volt. A második²⁰ felvételen körülbelül 15 méteres távolság felett járművek keresztezik a villamos pályát, srégen haladva az észlelési távolság peremén. A harmadik²¹ felvételen a kamerához közel, 2 és 7 méter távolság között haladnak el gyalogosok a kamera előtt, szintén srégen.



25. ábra, Akadályészlelés a második felvételen, 263. képkocka

A kiértékelés során képkockáról képkockára meghatároztam, melyik objektumot lehetett akadálnak tekinteni. A 25. ábra szemlélteti, hogyan is néz ki a képkocka

¹⁹ obstacle_num2_07_38

²⁰ obstacle_num0_21_18

²¹ obstacle_num1_17_12

feldolgozása során azonosított akadály. Akadálnak kell venni minden olyan objektumot, mely a villamospálya úrszelvényében tartózkodik. Mivel a rendszer lentől felfele vizsgálja az objektumok elhelyezkedését és a legelső akadálydetektálás után nem hajt végre további vizsgálatot, ezért legfeljebb egyszerre csak egy objektum képezhet ténylegesen akadályt, mely mindig a kamerához legközelebb eső lesz. Mivel minden akadály akár több képkockán keresztül is jelen tud lenni, ezért azt vizsgáltam, hogy a tényleges akadályokat hány képkockán keresztül tudta helyesen azonosítani a rendszer. Az objektumok alakját egy durva becsléssel határoztam meg, ezért a közelítésből fakadó helytelen azonosításokat is rögzítettem. Emellett, mivel a YOLO nem biztos, hogy helyesen vagy egyáltalán detektálta az objektumot, ezért az észlelésnek hiányát is rögzítettem. Az akadályészlelés pontosságát ezért az összes képkocka és helyes detektálás százalékos arányában mértem. Az eredményeket az 5. táblázat ismerteti.

Felvétel	Akadály sorszám	Átlagos távolság (m)	Első és utolsó képkocka	Összes képkocka	Helyes	Hibás	Nem észlelt	Pontosság
obstacle_num2_07_38	1.	13,00	4 - 19	16	15	0	1	93,75%
	2.	6,40	20 - 92	72	62	0	10	86,11%
	3.	14,50	92 - 98	6	3	3	0	50,00%
	4.	17,20	99 - 107	8	7	0	1	87,50%
	5.	7,20	108 - 175	67	60	3	4	89,55%
	6.	7,50	176 - 200	25	23	0	2	92,00%
	7.	14,50	214 - 219	6	6	0	0	100,00%
	8.	12,00	220 - 242	23	18	3	2	78,26%
	9.	14,80	243 - 246	4	1	0	3	25,00%
	10.	11,90	282 - 286	5	3	0	2	60,00%
	11.	8,90	287 - 349	63	49	0	14	77,78%
	12.	17,50	379 - 411	32	29	2	1	90,63%
	13.	12,50	422 - 442	21	19	0	2	90,48%
	14.	17,50	443 - 467	24	23	1	0	95,83%
	15.	12,00	481 - 495	15	12	0	3	80,00%
	16.	6,50	496 - 566	71	66	0	5	92,96%
	17.	17,00	567 - 590	24	21	1	2	87,50%
	18.	16,50	613 - 618	6	5	1	0	83,33%
	19.	12,50	619 - 641	23	20	0	3	86,96%
obstacle_num0_21_18	1.	21,50	61 - 69	9	9	0	0	100,00%
	2.	18,00	83 - 120	38	33	0	5	86,84%
	3.	18,00	135 - 168	34	26	2	6	76,47%
	4.	16,60	171 - 200	30	29	1	0	96,67%
	5.	21,40	201 - 202	2	2	0	0	100,00%
	6.	19,45	203 - 234	32	28	2	2	87,50%
	7.	15,20	241 - 265	25	19	2	4	76,00%
obstacle_num1_17_12	1.	5,45	18 - 43	26	26	0	0	100,00%
	2.	3,00	44 - 102	59	47	3	9	79,66%
	3.	7,00	103 - 115	13	10	3	0	76,92%
	4.	5,90	116 - 184	69	61	5	3	88,41%
	5.	5,85	185 - 213	29	21	2	6	72,41%
	6.	7,00	214 - 235	22	21	1	0	95,45%
	7.	2,00	236 - 291	56	36	0	20	64,29%
	8.	8,50	292 - 310	19	12	0	7	63,16%
Eredmények összegezve:				974	822	35	117	84,39%

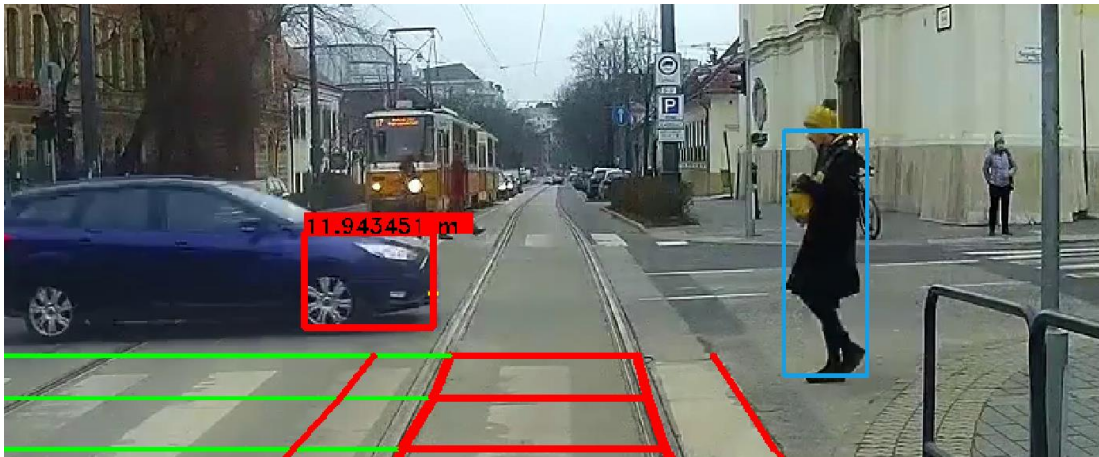
5. táblázat, Akadályészlelés eredmények

Elsősorban érdemes megemlíteni, hogy a közel ezer képkockán keresztül fennálló akadályokat körülbelül 84,5%-os pontossággal sikerült meghatározni, mely véleményem szerint egy kifejezetten jó értéknek mondható. Azonban, ha megnézzük a hibák számát, akkor látható, hogy a helytelen azonosításokhoz képest közel négyszer annyi alkalommal nem sikerült a YOLO-nak felismernie az akadályt. A problémát két szempont szerint célszerű megvizsgálni. Megkülönböztetem azokat az eseteket, ahol az észlelt objektum már a sínszalakon és az általuk közre zárt területen tartózkodik, illetve azt az esetet, amikor még csak az úrszelvény szélén, de a pályán még nem tartózkodik.

Az első esetben elmondható, hogy szinte mindig kellő mennyiségű jellemző áll rendelkezésre az objektumok észlelésére. A YOLO egyik problémája, hogy a több, egymást fedő objektumokat nagyon nehezen tudja megkülönböztetni, sokszor alkalmatlan rá. Ez a probléma jellemzően akkor észrevehető, ha egy többsávos kereszteződésben több autó halad keresztbe a pálya előtt. Sokszor egy csoportba vonva kezeli a különböző objektumokat és van, hogy egy esetleg elhaladó gyalogost pár képkockán keresztül emiatt észre sem vesz, mert a gyalogos jellemzői túl hasonlóak a járművekéhez²². Ezek az esetek a felvételeken jellemzően elszórva nyilvánulnak meg.

Ha az úrszelvény határait nézzük, akkor az objektumdetektálást mindig egy, az úrszelvény környezetét körülzárt képen kell végrehajtani. A széleken jellemző módon az objektumok kis részletei látszódnak, általában srégen ketté vágva. Ha a YOLO kellő mennyiségű jellemző birtokába kerül, akkor a csonka objektumokat a vágás szélén meglepő módon egész jól képes azonosítani. Jellemzően a járművek elejét már egy fényszóróról és egy kerékről is helyesen meg tudja határozni, melyet a 26. ábra szemléltet.

²² Hasonló jellemzők lehetnek olyan körvonalakkal behatárolt alakzatok, melyek például közel azonos színű kitöltéssel rendelkeznek.



26. ábra, Akadályészlelés az első felvételen, 281. képkocka

A vágás szélein történő helyes azonosításhoz kellő mennyiségű jellemzőre van szükség, melyek nem biztos, hogy az az objektum űrszelvénybe történő lépése előtt rendelkezésre fognak állni. A probléma kiküszöböléséhez a vágási határokon kell módosítani oly módon, hogy az űrszelvény környékéből minél több részlet látszódjon. Ez felvet egy új problémát, miszerint a jellemzően hosszabb járművek, esetleg szerelvényt befoglaló téglalapok sarkai oly mértékben metszik az űrszelvényt, melyek nem tükrözik a jármű tényleges elhelyezkedését és helytelen detektálást eredményeznek. A probléma kiküszöbölésére lett bevezetve az objektum alakjának geometriai úton történő becslése, mely kisebb méretű befoglaló téglalapok esetén jól működik, de egy bizonyos méret felett már nem. Az űrszelvény környezetének körül vágása erre a problémára nyújt egy nagyjából jól működő megoldást. Amennyiben túl sok részletet mutatunk a környezetből, akkor a probléma ismét előjön.

Az eredményeket érdemes összevetni két korábban ismertetett hasonló rendszerrel. A [8] rendszer a pályadetektáláson túl akadályészlelési feladatokat is ellát, mely kiterjed a sínszálak közvetlen környezetére és a sínszálak közti területre. A cikkben bemutatott eredményeket aszerint értékelték, hogy a rendszer sikeresen azonosította-e a felvételen elhelyezett akadályt vagy sem. 10 esetből 10 alkalommal sikeresen azonosítottak akadályt, mellyel 100%-os pontosságot jelentettek. Rendszeremet nagyobb részletességgel értékeltem ki, ezért az összehasonlítás érdekében az eredményeket leegyszerűsítettem sikeres és sikertelen akadály azonosítások arányára. Az 5. táblázaton láthatjuk, hogy minden akadályt sikeresen azonosítottunk, ezzel elérve a [8] rendszer pontosságát. Továbbá fontos kiemelni, hogy rendszerem nem csak a sínszálak és köztes területén, hanem a teljes űrszelvényre kiterjesztve hajtja végre az akadályészlelést.

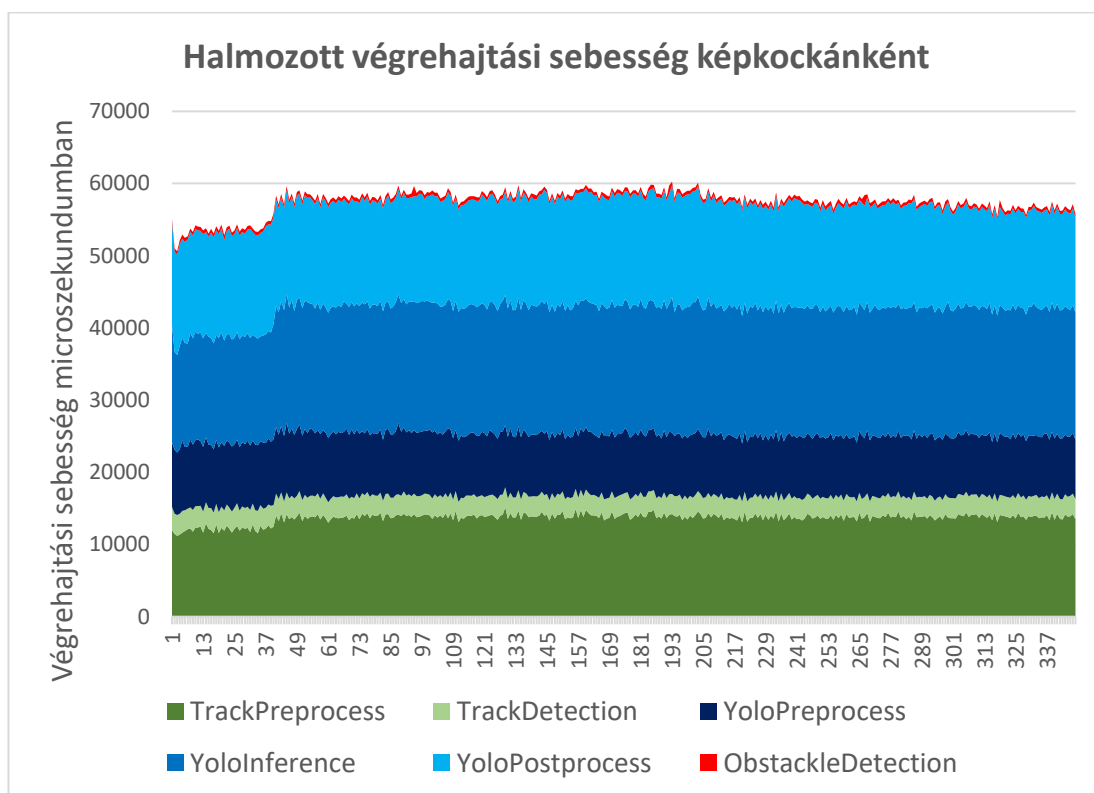
A [13] cikkben nem egy akadályészlelő rendszert mutatnak be, hanem egy objektumdetektáló megoldást, mely egy nagyvasúti környezetben hajt végre egyéni fejlesztésű neurális hálózattal vonat, gyalogos és bőrönd észlelést. Rendszerük átlagosan 91,61%-os pontossággal bír, melyet értékelés során összehasonlítottak más objektumdetektáló megoldásokkal is, köztük a YOLO első verziójával, mely átlagosan 75,67%-os pontosságot tanúsított. Rendszeremben a YOLO ötödik verzióját használok, mely jelentős teljesítmény ugrást jelent mind pontosságban, mind pedig sebességben az első verzióhoz képest. A YOLO precizitásának vizsgálatára az 5. táblázaton feltüntetett „nem észlelt” oszlop összesítését kell vizsgálni az összes képkocka tükrében, mely 87,98%-os pontosságot eredményez. Rendszeremben használt YOLO általános célú, nem kifejezetten vasúti környezetre lett fejlesztve, ezért ilyen szempontból nagyon is jónak tekinthetők az eredmények, mely során kevesebb mint 4%-os differencia mérhető a [13] megvalósításhoz képest.

Az akadályészlelés eredményeit összefoglalva elmondható, hogy a rendszer közel 85%-os pontossága nem elhanyagolható érték, ám figyelembe kell venni azokat a szélsőséges eseteket, melyek során az észlelés nagy valószínűséggel nem fog megvalósulni. Ilyen esetek lehetnek az egy csoportot képző, hasonló jellemzőkkel bíró objektumok, melyek entitásait külön-külön nem, vagy csak nehezen tud behatárolni a YOLO, illetve az úrszelvény környezetének körül vágása során, az úrszelvény szélén fellépő objektumok kevés jellemzői nem mindig elegend az észleléshez.

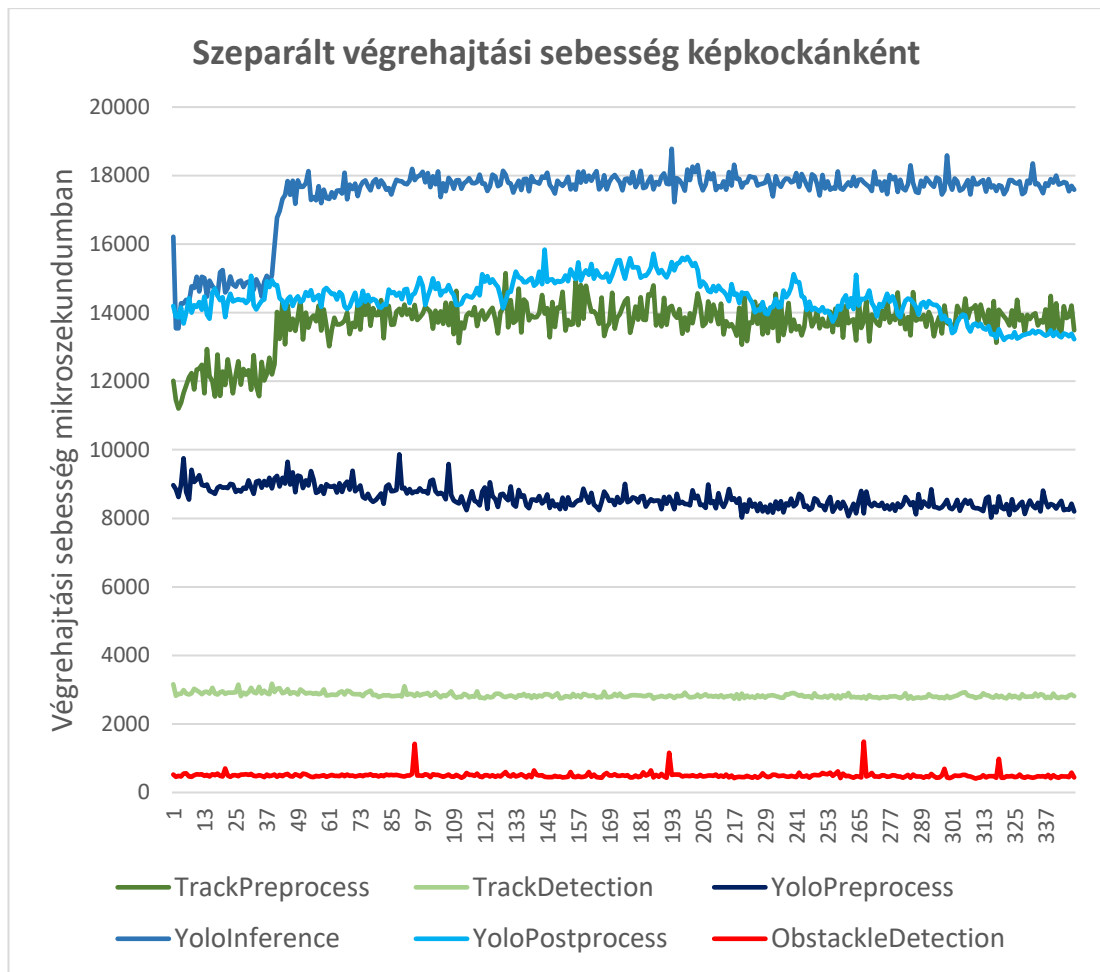
5.3. VÉGREHAJTÁSI SEBESSÉG EREDMÉNYEK

A rendszer teljesítményének vizsgálata során a képkockák feldolgozásához szükséges időt vettem alapul. A végrehajtási idő kritikus pontja az akadályészlelésnek, ezért arra törekedtem, hogy a valós idejű feldolgozásnak eleget tudjon tenni a rendszer. Az akadályészlelés során felhasznált felvételek közül kiemeltem egyet, melyen elvégeztem a futási idők rögzítését.

Az eredmények rögzítése során külön tároltam az egyes modulok egyes komponenseinek futásidejét. A pályadetektálást során különválasztottam a GPU-n futtatott előfeldolgozást és a CPU-n implementált szekvenciális pályadetektáló algoritmust. Az objektumdetektálást három fő részre bontottam. Az előfeldolgozásra, mely során az adott képkockából előáll a YOLO input tenzora, a YOLO-t futtató GPU implementációra, illetve az utófeldolgozásra, mely tartalmazza a YOLO output tenzorának szűrését és a nem-maximumok elnyomását. Az akadályészlelő modult nem bontottam tovább. A sebességrögzítést tíz alkalommal ismételt meg ugyan azon felvételen. Erre azért volt szükség, mert minden futtatással más és más, minimálisan eltérő sebességeket mértem az inkonzisztens ütemezés miatt. Az így rögzített eredmények átlagát véve ténylegesen reprezentatív sebességmutatókat kaptam, melyek ideálisak a rendszer teljesítményének tükrözésére.



27. ábra, Modulok rétegjeinek összesített futási ideje képkockánként



28. ábra, Modulok rétegjeinek futási ideje képkockáknként

Az eredményeket két diagrammon összesítettem. A 27. ábra szemlélteti a különböző modulok részeinek eredményeit halmozottan összesítve képkockáknként. Az ábrán jól látható, hogy az átlagolást követően a különböző képkockák feldolgozási ideje nem konzisztens. A végrehajtási sebesség körülbelül 50 ms és 60 ms között mozog, de a rendszer átlagosan 57,47 milliszekundum sebességgel teljesített. A rendszer megvalósítása során elvettem a pályadetektálás és akadályészlelés párhuzamosan történő végrehajtását, mert az előzetes eredmények szerint is kis teljesítmény növekedést érnék el. A pályadetektálás előfeldolgozása és a YOLO futtatása egy és ugyan azon GPU-n történik. Ezek a folyamatok teljes mértékben kihasználják a GPU-n belüli multiprocesszorokat, melyeken további párhuzamosítás nem hajtható végre, csak a kernelek késleltetésének rovására. Ahhoz, hogy az említett folyamatokat hatékony módon tudjuk futtatni párhuzamosan, úgy célszerű bevezetni egy második GPU-t.

Az 28. ábra szemlélteti, hogy az egyes modulok részei milyen sebességgel rendelkeznek egymáshoz viszonyítva. Érdemes megemlíteni, hogy a pályadetektálás (világoszöld) és akadályészlelés (vörös) CPU-n futó, nem párhuzamosítható folyamat. Sebességük a többi folyamathoz képest elhanyagolhatóan kicsit, képkockánként kicsi sebesség-ingadozás tapasztalható. Ellentétben az objektumdetektálás lépéseivel és a pályadetektálás előfeldolgozásával, melyek ki tudják használni a GPU nyújtotta párhuzamosítási lehetőségeket. Ezzel egyidejűleg jelentős sebesség-ingadozásokat mértem képkockánként, melyek eredetéről konkrét magyarázattal nem rendelkezek.

A 3.2. fejezetben ismertetett hasonló rendszerek nem mindegyike valósítja meg a teljes akadályészlelést, leginkább csak részfolyamatait. A [10] rendszer pályadetektálási sebessége eléri a 25 FPS feldolgozási sebességet, mely 40 milliszekundumnak feleltethető meg. A 27. ábrán látható, hogy a pálya előfeldolgozása (világos zöld) és detektálása (sötét zöld) együtt átlagosan 17 milliszekundum időt tesz ki, mely során megvalósul a teljes pályadetektálás. Elmondható, hogy rendszerem több mint kétszer gyorsabban képes ellátni a feladatot a [10] rendszerrel szemben.

A [8] rendszer teljes akadályészlelést valósít meg pályadetektálással együtt, mely során szintén 25 FPS végrehajtási sebességet tudtak elérni. Saját megvalósításomnak a végrehajtási idő közel 75%-át az objektumdetektálás teszi ki, ezért a [8] rendszernél közel 40%-kal lassabbnak mondható a ~17,5 FPS értékkel. További optimalizálással jobb eredményeket érhetőek el.

Végezetül kizárólag objektumdetektálási szempontból a [13] rendszernél 26,3 FPS végrehajtási sebességet mértek, mellyel rendszerem objektumdetektáló modulja közel azonos idővel fut, majdnem pontosan 25 FPS-el. Sebesség és pontosság tekintetében megközelítőleg azonos eredményeket értem el objektumdetektálás terén.

Összességében elmondható, hogy a rendszer teljesítménye minden modult összesítve jónak tekinthető. A valós idejű végrehajtást korábban a kamera által szállított képek sebességével azonosítottam, mely relatív értelmet nyer az algoritmus teljesítményének meghatározásakor. Ha a kamera körülbelül 20 képkockát szolgáltat másodpercenként, akkor elmondható, hogy a rendszer közel valós idővel tud működni. Ellenben, ha másodpercenként több képkockát szállít a kamera, akkor a teljesítmény egyenesen arányosan csökken.

5.4. TOVÁBBFEJLESZTÉSI LEHETŐSÉGEK

Az eredmények ismertetése során nyilvánvalóvá vált, hogy az 5.2 alfejezetben bemutatott szélsőséges esetek az akadályészlelésre kiemelten negatívan hatnak, mely nem elhanyagolandó probléma. Továbbá, a rendszer teljesítményére számos optimalizálási lehetőséget eszközöltem, melyekkel egy elfogadható feldolgozási sebességet sikerült elérni, viszont további optimalizálással magasabb sebességet érhető el. A rendszer továbbfejlesztési lehetőségére az érzékelés pontosságának növelését és a végrehajtási sebesség növelését találtam optimálisnak, melyek tényleges megoldást nyújtanának a korábban említett problémákra.

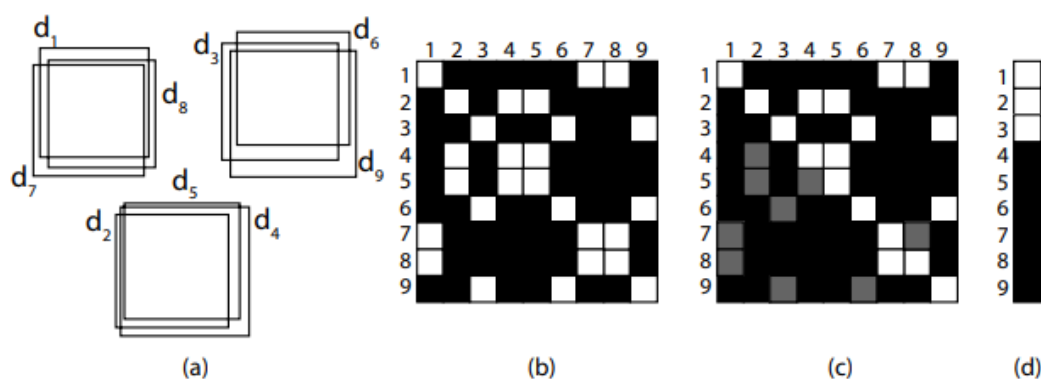
Az akadályészlelés során azokat a helyzeteket tekintem szélsőséges esetnek, ahol az észlelendő objektum az úrszelvény szélén tartózkodik. Problémát az jelenti, hogy az úrszelvény környezete ki van maszkolva és a maszkolás határán szereplő objektumok jellemzőinek egy része le van „nyesve”. Az objektumdetektálás így számos esetben nem történik meg, csak ha rendelkezésre áll kellő mennyiségű jellemző, például egy autó elejének felismeréséhez elegendő az elülső fényszóró és kerék (26. ábra). A maszkolásra azért volt szükség, mert számos esetben a szerelvény mellett elhaladó objektumok alakjának határait a YOLO által szolgáltatott befoglaló téglalapok nem reprezentálják helyesen, mely során fals akadálnak tekinti az elhaladó járművet a rendszer. Az optimális megoldás az volna, ha az objektumokat nem a 2D-s síkon kezelnék, hanem a 3D-s térben. Ilyenkor nem egy síkbeli befoglaló téglalappal kellene dolgozni, hanem az objektumot magába foglaló téglatest sarokpontjai határoznák meg az objektum térbeli pozícióját. A pályadetektálás során detektált pálya szegmenseit térben egyértelműen el lehet helyezni, melyekhez lehetne viszonyítani a hexaéder koordinátáit. Az eredmény elméletben egy jóval pontosabb, kevesebb becslést igénylő akadályészlelő módszer lenne, melyet az úrszelvény teljes környezetére ki lehet vetíteni. A 3D objektumdetektálás egy módszerét fejlesztették ki 2021-ben, melyet a „*M3DSSD: Monocular 3D Single Stage Object Detector*” [31] cikkben ismertetnek. A rendszer képes 3D téglatestek prediktálására (29. ábra) valós időben, mely az ismertetett eredmények alapján körülbelül 6 ms végrehajtási idővel rendelkezik²³. Az implementáció egyszerre nyújtana számottevő növekedést mind az akadályészlelés pontosságának növelésében, mind pedig a fejlesztett rendszer végrehajtási sebességében.

²³ A megvalósított objektumdetektáló modul végrehajtási ideje átlagosan 40ms, összehasonlítva a bemutatott [31] rendszerével, mely közel hétszer gyorsabban teljesít.



29. ábra, Az M3DSSD eredményét reprezentáló 3D téglatestek

Amennyiben a jelenlegi objektumdetektáló modult kívánjuk használni, úgy további optimalizálási lehetőségekkel lehet élni a sebesség növelésének érdekében. A 5.3 alfejezetben ismertetett sebesség eredményeknél látható volt, hogy a YOLO utófeldolgozása átlagosan 14 ms-re tehető. Ez a folyamat teljes mértékben CPU-n fut, melynek jelentős részét a YOLO eredmény mátrixának küszöbérték szerinti szűrését követően történő nem-maximumok elnyomása (NMS) képezi. A NMS hatékony párhuzamosítására 2016-ban kifejlesztettek egy módszert, mely a „*Work-efficient parallel non-maximum suppression for embedded GPU architectures*” [32] cikkben olvasható. A technika eredetileg beágyazott rendszerekre lett fejlesztve, de az ismertetett eredmények szerint hagyományos, nem beágyazott GPU esetén is körülbelül 2 ms végrehajtási idővel rendelkezik függetlenül a detektált objektumok számától, mely egy hétszeres gyorsulást eredményez YOLO utófeldolgozás során. Az NMS GPU-n történő végrehajtásának lépéseit a 30. ábra szemlélteti.



30. ábra, GPU-NMS [32] lépései²⁴

²⁴ (a) detektor által meghatározott egyedek; boolean mátrix (b) klaszterezést követően és (c) nem reprezentatív javaslatok szűrése; az (d) eredmény az AND redukciót követően

Általánosságban elmondható, hogy a képmanipuláló CUDA kernelek teljesítménye a GPU-ban megtalálható SIMD feldolgozó egységek számától függ. A fejlesztés során felhasznált Nvidia GTX 1080-as GPU 20 SM²⁵ multiprocesszorral rendelkezik. A pályadetektálás előfeldolgozása során alkalmazott szűrés és gradienskép létrehozása 608×275-ös felbontáson történik, melyet 164 darab 1024 szálú blokkra tudunk bontani. Az említett blokk számot a GPU-nak közel 9 iteráción keresztül tudja feldolgozni. Könnyen belátható, hogy az SM multiprocesszorok számának növelésével csökken az iterációk száma, ezért célszerű minél nagyobb multiprocesszor (SM) számmal rendelkező GPU-t használni kisebb végrehajtási idő érdekében.

Összességében elmondható, hogy a legtöbb gyorsítási lehetőség mögött a folyamatok GPU-n történő implementálása áll, viszont az optimális megoldást az említett M3DSSD 3D téglatest prediktáló technika implementálása jelentené mind teljesítményben, mind pedig az akadályészlelés pontosságának növelésében.

²⁵ Streaming multiprocessor, az Nvidia GPU-k SIMD feldolgozó egysége

6. ÖSSZEFOGLALÁS

Dolgozatomban egy olyan komplex biztonsági rendszer kidolgozásával és fejlesztésével foglalkoztam, mely aktív védelmet nyújt a tömegközlekedés kötöttpályás ágazatának. A városi környezetre fejlesztett akadályészlelő rendszerem a lehető legkevesebb szenzor felhasználásával nyújt valós idejű pályameghatározást és akadályészlelést a haladás során a vezetőfülkébe felszerelt RGB kamera képének feldolgozásával. Kutatásom során olyan összetett magas és alacsony szintű képfeldolgozási technikákat vizsgáltam, beleértve kortárs irodalomban fellelhető hasonló rendszerek módszereit, melyek segítettek a hatékony pálya és objektumdetektálás megvalósításában.

A megvalósított rendszert architektúráisan három modulra bontottam. Elsőnek a pályadetektálást végző modul meghatározza a kamera képen szereplő megfelelő pálya úrszelvényét számos képmanipuláló technikák összefűzésével. Második lépésben a YOLOv5 konvolúciós neurális hálózat segítségével megtörténik az objektumdetektálás, mely során az akadályt képezhető objektumok észlelése valósul meg. Folyamat végén geometriai úton összevetjük az úrszelvénnyel az objektumok helyzetét. A képkocka feldolgozását követően a rendszer jelzi, mely objektum helyezkedik el az úrszelvényen belül, akadályt képezve.

A rendszert számtalan felvételen teszteltem, különböző fényviszonyok között, leginkább városi környezetben. Az eredmények igazolják rendszerem hatékonyságát. A pályadetektálást 18 méteres távolságig közel 98%-os pontossággal sikerül elvégezni fényviszonyoktól leginkább függetlenül. Jelentős esőzést vagy havazást követően pályaszakasz detektálás nem, vagy csak korlátozva kivitelezhető. Hasonló rendszerekkel összevetve nagy hatékonyságot tanúsít saját rendszerem sebesség és pontosság terén. Míg az objektumdetektálást egy általános célra fejlesztett neurális hálózat végzi, megoldásom így is megközelíti sebességben és pontosságban a hasonló rendszerek különböző technikáit. Mivel akadályészlelés során a mélységi adatok hiánya miatt egy becslést végzünk az objektum alakjára, ezért az akadályészlelés szélsőséges helyzetekben némi pontatlanságot tanúsít, melyet az eredmények alátámasztanak. Továbbfejlesztési javaslataim hatékony megoldást nyújtanak az eredmények során ismertetett problémákra, melyet követően a rendszer pontosságának jelentős- és végrehajtási sebességének drasztikus növekedése várható.

A rendszer éles környezetben való tesztelésére nem volt módom, ezért a felvételekre hagyatkozva vittem véghez a fejlesztést. Mivel a rendszerem csak egy kamera képének feldolgozásával gyűjti össze a kellő mennyiségű információt, ezért

jelentős számítási igényrel rendelkező hardverre van szüksége, mely alkalmatlanná teszi a beágyazott rendszerekre történő implementálását.

Végezetül szeretném megköszönni konzulensemnek, Dr. Vámosy Zoltánnak a dolgozatom elkészítéséhez nyújtott támogatását és segítségét.

7. CONCLUSION

In my thesis, I designed and developed a complex security system for active protection in the urban light rail transportation sector. My obstacle detection system, developed for urban environments, provides real-time track and obstacle detection by processing the images from an RGB camera installed in the driver's cab. In my research, I investigated complex high- and low-level image processing techniques, including methods for similar systems in contemporary literature, that have helped to achieve efficient path and object detection.

I have architecturally decomposed the implemented system into three modules. First, the track detection module determines the corresponding track segments in the camera image by combining several image manipulation techniques. In the second step, object detection is performed using the YOLOv5 convolutional neural network to detect possible obstacles. In the last phase, the position of the objects is geometrically compared with the structure gauge segment. After processing the frame, the system indicates which object is located within the structure gauge, which constitutes as an obstacle.

The system has been tested in numerous clips under different lighting conditions, mostly in urban environments. The results demonstrate the effectiveness of my system. Track detection up to 18 meters can be performed with an accuracy of nearly 98%, mostly independently of lighting conditions. After heavy rain or snowfall, track detection is not or only possible to a limited extent. Compared to similar systems, my own system shows high efficiency in terms of speed and accuracy. While object detection is performed by a general-purpose neural network, my solution still approaches the speed and accuracy of various techniques of similar systems. Since obstacle detection is an estimation of the object shape due to the lack of depth data, in extreme situations obstacle detection exhibits some inaccuracy, which is shown in the results. My suggestions for further improvements provide an effective solution to the problems described in the results, after which a significant increase in the accuracy and a drastic increase in the execution speed of the system can be expected.

I have not had the opportunity to test the system in live environment and have therefore relied on recordings to carry out the development. Since my system collects the necessary amount of information by processing only the camera image, it has significant computational requirements, which makes it unsuitable for implementation on embedded systems.

Finally, I would like to thank my consultant, Dr. Zoltán Vámosy for his support and help in the preparation of my thesis.

IRODALOMJEGYZÉK

- [1] Központi Statisztikai Hivatal, ksh.hu, „24.1.1.23. A közúti gépjárművek száma járműnemenként, az év végén,” 2021. [Online]. Available: https://www.ksh.hu/stadat_files/sza/hu/sza0023.html. [Hozzáférés dátuma: 5 március 2022].
- [2] Központi Statisztikai Hivatal, ksh.hu, „4.1.1.44. Közlekedési balesetek,” 2021. [Online]. Available: https://www.ksh.hu/stadat_files/ege/hu/ege0061.html. [Hozzáférés dátuma: 5 március 2022].
- [3] *BUDAPESTI KÖZLEKEDÉSI ZÁRTKÖRŰEN MŰKÖDŐ RÉSZVÉNYTÁRSASÁG: "Általános szakmai ismeretek I.", 2. kiadás, Összeállította: Gerlang Tamás, Budapest, 2015. 87 p..*
- [4] S. Lakusic and R. Kecur, "The impact of the geometry of rail welds on noise level in urban environments," *ACOUSTICS*, p. 248, November 2005.
- [5] Budapest Közlekedési Vállalat Zrt., „F.1.-F.2. SZÁMÚ JELZÉSI ÉS FORGALMI UTASÍTÁS A KÖZÚTI VASUTAK SZÁMÁRA,” 2016, p. 10.
- [6] *1/1975. (II. 5.) KPM-BM együttes rendelet § (2) bekezdés s) pont.*
- [7] D. Ristić-Durrant, M. Franke and K. Michels, "A Review of Vision-Based On-Board Obstacle Detection and Distance Estimation in Railways," *Sensors*, vol. 21, no. 10, p. 3452, 2021.
- [8] H. Wang, X. Zhang, L. Damiani, P. Giribone, R. Revetria and G. Ronchetti, "Video Analysis for Improving Transportation Safety: Obstacles and Collision Detection Applied to Railways and Roads," *International MultiConference of Engineers and Computer Scientists*, pp. 909-915, 15-17. március 2017.
- [9] J. Canny, "A computational approach to edge detection," *IEEE Transactions on pattern analysis and machine intelligence*, p. 679–698, 1986.
- [10] F. Maire and A. Bigdeli, "Obstacle-free range determination for rail track maintenance vehicles," *2010 11th International Conference on Control Automation Robotics & Vision*, pp. 2172-2178, 2010.
- [11] R. O. Duda and P. E. Hart, "Use of the Hough transformation to detect lines and curves in pictures," *Communications of the ACM*, vol. 15, p. 11–15, 1972.
- [12] H. Mukojima, D. Deguchi, Y. Kawanishi, I. Ide, H. Murase, M. Ukai, N. Nagamine and R. Nakasone, "Moving camera background-subtraction for obstacle detection on railway tracks," *2016 IEEE International Conference on Image Processing (ICIP)*, pp. 3967-3971, 2016.

- [13] X. Yuchuan, G. Chunhai, Y. Lei, T. Simon and W. Guodong, "Real-time Obstacle Detection Over Rails Using Deep Convolutional Neural Network," *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, pp. 1007-1012, 2019.
- [14] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu and A. C. Berg, "SSD: Single Shot MultiBox Detector," in *Computer Vision – ECCV 2016*, Springer International Publishing, 2016, p. 21–37.
- [15] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, *You Only Look Once: Unified, Real-Time Object Detection*, arXiv, 2015.
- [16] N. Kanopoulos, N. Vasanthavada and R. L. Baker, "Design of an image edge detection filter using the Sobel operator," *IEEE Journal of solid-state circuits*, vol. 23, p. 358–367, 1988.
- [17] G. M. H. Amer and A. M. Abushaala, "Edge detection methods," *2015 2nd World Symposium on Web Applications and Networking (WSWAN)*, pp. 1-7, 2015.
- [18] G. E. Sotak and K. L. Boyer, "The laplacian-of-gaussian kernel: A formal analysis and design procedure for fast, accurate convolution and full-frame output," *Computer Vision, Graphics, and Image Processing*, vol. 48, pp. 147-189, 1989.
- [19] A. Kumar and S. S. Sodhi, "Comparative Analysis of Gaussian Filter, Median Filter and Denoise Autoencoder," *2020 7th International Conference on Computing for Sustainable Global Development (INDIACom)*, pp. 45-51, 2020.
- [20] A. Agarwal, S. Gupta and D. K. Singh, "Review of optical flow technique for moving object detection," *2016 2nd International Conference on Contemporary Computing and Informatics (IC3I)*, pp. 409-413, 2016.
- [21] S. L. Hijazi, R. Kumar and C. Rowen, "Using Convolutional Neural Networks for Image Recognition," in *Cadence Design Systems, Inc.*, 2015.
- [22] R. Girshick, J. Donahue, T. Darrell and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," *CoRR*, vol. abs/1311.2524v5, 2014.
- [23] M. Maity, S. Banerjee and S. Sinha Chaudhuri, "Faster R-CNN and YOLO based Vehicle detection: A Survey," *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*, pp. 1442-1447, 2021.
- [24] O. R. developers, *ONNX Runtime*, 2021.
- [25] G. Jocher, A. Chaurasia, A. Stoken, J. Borovec, NanoCode012, Y. Kwon, TaoXie, K. Michael, J. Fang, imyhxy, Lorna, C. Wong, 曾. Yifu), V. Abhiram, D. Montes, Z. Wang, C. Fati, J. Nadar, Laughing, UnglvKitDe, tkianai, yxNONG, P. Skalski, A. Hogan, M. Strobel, M. Jain, L. Mammana and xylieong, *ultralytics/yolov5: v6.2 - YOLOv5 Classification Models, Apple M1, Reproducibility, ClearML and Deci.ai integrations*, Zenodo, 2022.

- [26] M. Garland, S. Le Grand, J. Nickolls, J. Anderson, J. Hardwick, S. Morton, E. Phillips, Y. Zhang and V. Volkov, "Parallel Computing Experiences with CUDA," *IEEE Micro*, vol. 28, pp. 13-27, July 2008.
- [27] M. A. Haseeb, J. Guan, D. Ristic-Durrant and A. Gräser, "DisNet: a novel method for distance estimation from monocular camera," *10th Planning, Perception and Navigation for Intelligent Vehicles (PPNIV18), IROS*, pp. 139-144, 2018.
- [28] B. Bilgic, B. K. P. Horn and I. Masaki, "Efficient integral image computation on the GPU," *2010 IEEE Intelligent Vehicles Symposium*, pp. 528-533, 2010.
- [29] P. Jaccard, *Etude comparative de la distribution florale dans une portion des Alpes et du Jura*, Impr. Corbaz, 1901.
- [30] Lemezes01, „17-es villamos (Bécsi út VÁ. - Savoya park) / Budapest trams: Line 17,” 2019. [Online]. Available: <https://www.youtube.com/watch?v=tIePRa9EZQ0>.
- [31] S. Luo, H. Dai, L. Shao and Y. Ding, "M3DSSD: Monocular 3D Single Stage Object Detector," *CoRR*, vol. abs/2103.13164, 2021.
- [32] D. Oro, C. Fernández, X. Martorell and J. Hernando, "Work-efficient parallel non-maximum suppression for embedded GPU architectures," *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1026-1030, 2016.

ÁBRAJEGYZÉK

1. ábra. Villamospálya keresztmetszete [4].....	5
2. ábra. Balról jobbra: Vignol, Phoenix és Tömb sínprofil.....	5
3. ábra. Balról jobbra: Betonlemez, bazaltkőkockás és aszfalttal burkolt pályaszakasz.....	6
4. ábra. A vágányra merőleges űrszelvény (zöld)	7
5. ábra. Bazaltkőkockás pályán alkalmazott Canny operátor előtte és utána.....	16
6. ábra. Sínszalak összefutása	22
7. ábra. Ívelt pályán haladó villamos	23
8. ábra. Nyomcsatorna, tiszta (fent), homokolt (lent)	23
9. ábra. Balról jobbra: Eredeti kép, Sobel 3×3, Sobel 3×3 és 5×5 kombinálva	25
10. ábra, Pályadetektálás lépései előfeldolgozás után.....	28
11. ábra, Detektált pálya.....	30
12. ábra. YOLOv5 objektum predikciók.....	33
13. ábra. Objektumot reprezentáló téglatest (piros) és 2D befoglaló téglalap (kék)..	34
14. ábra, Befoglaló téglalap átmenete ellipszisbe	35
15. ábra, Objektum helyzetének esetei az űrszelvényhez viszonyítva.....	37
16. ábra, Objektum legközelebbi pontjának meghatározása az űrszelvényhez 1/2 ...	38
17. ábra, Objektum legközelebbi pontjának meghatározása az űrszelvényhez 2/2 ...	39
18. ábra. Rendszer működésének folyamatábrája	40
19. ábra. Pályadetektáló modul folyamatára	41
20. ábra. Objektumdetektáló modul folyamatára	42
21. ábra, Pályadetektálás betonlemez pályaszakaszon	45
22. ábra, Felvételek feldolgozott és nyers képei, fentről lefele haladva: betonlemez útburkolat, magasságány, bazaltkőkockás burkolat, aszfalt és alagút	47
23. ábra, Gradiens és eredeti kép, fentről lefele; elszórtan hóval és vízzel fedett pályaszakasz.....	49
24. ábra, Felülnézeti pályadetektálás, balról jobbra, eredeti kép és detektált pálya ..	51
25. ábra, Akadályészlelés a második felvételen, 263. képkocka	52
26. ábra, Akadályészlelés az első felvételen, 281. képkocka.....	56
27. ábra, Modulok rétegjeinek összesített futási ideje képkockánként	58
28. ábra, Modulok rétegjeinek futási ideje képkockánként.....	59
29. ábra, Az M3DSSD eredményét reprezentáló 3D téglatestek	62
30. ábra, GPU-NMS [32] lépései	62

1. táblázat, Hasonló rendszerek összegzése pályaaazonosítás módja szerint	17
2. táblázat, Hasonló rendszerek összegzése objektumdetektálás alapján	18
3. táblázat, 18 méteres távolságot bezáró vizsgálat eredményei.....	46
4. táblázat, 22 méteres távolságot bezáró vizsgálat eredményei.....	48
5. táblázat, Akadályészlelés eredmények.....	54