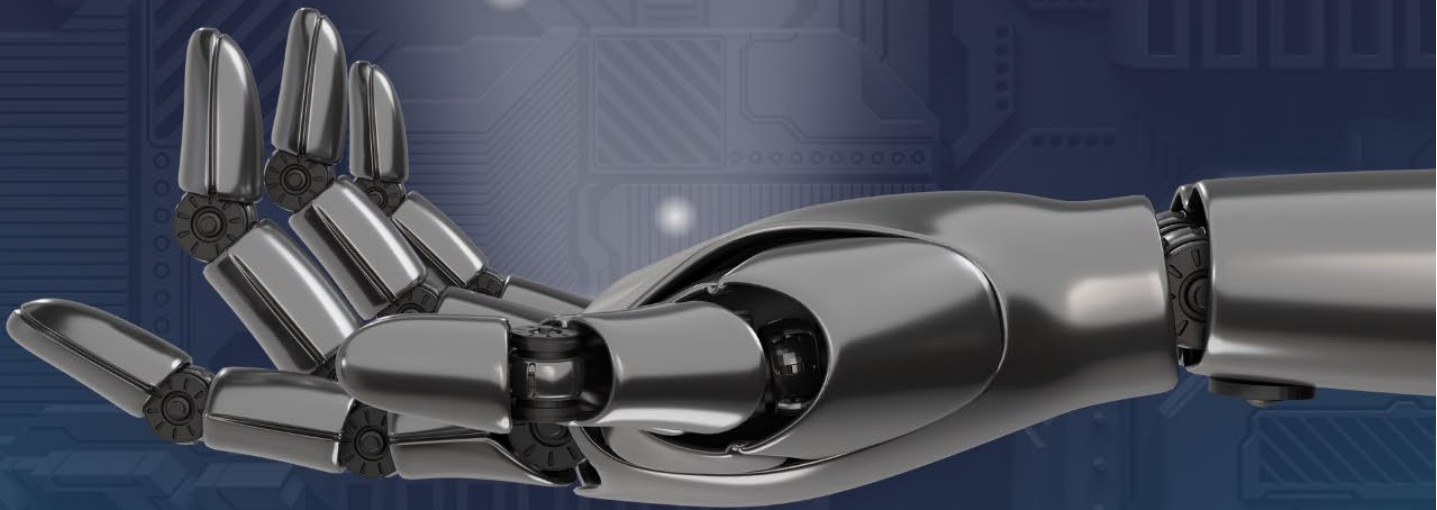


# "MESTERSÉGES INTELLIGENCIA RENDSZEREK ALKALMAZÁSA A MOBILITÁSBAN"

KONFERENCIAKIADVÁNY  
2023. TAVASZ





MESTERSÉGES INTELLIGENCIA  
RENDSZEREK ALKALMAZÁSA  
A MOBILITÁSBAN  
KONFERENCIAKIADVÁNY

2023. tavasz

Lektorok:

Dr. Busznyák Tibor, Dr. Eisingerné dr. Balassa Boglárka,

Dr. Enisz Krisztián, Dr. Gyurián Nagy Nikolett,

Dr. Polák József, Dr. Szénásy István,

Dr. Varga Zoltán

Szerkesztette:

Dr. Szauter Ferenc

Pup Dániel

Csikor Dániel

Földesi Rita

Nagy Bianka Marianna

Széchenyi István Egyetem

2023

ISBN 978-615-6443-29-8



## Tartalomjegyzék

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

<b>STOP VONAL- ÉS PARKOLÓHELY FELISMERÉS KAMERA, ILLETVE LIDAR ADATOKKAL A SHELL ECO-MARATHON AUTONÓM VERSENYÉN .....</b>	<b>9</b>
<i>BÉZI PATRIK</i>	
<b>PUBLIC TRANSPORT SERVICE PLANNING WITH THE SUPPORT OF AI .....</b>	<b>23</b>
<i>ZOLTÁN HIMA</i>	
<b>EMBERI ARC DETEKTÁLÁSÁNAK VIZSGÁLATA KAPSZULA HÁLÓZATOK ALKALMAZÁSÁVAL .....</b>	<b>32</b>
<i>HOLLÓSI JÁNOS<sup>A</sup>, BALLAGI ÁRON<sup>B</sup>, POZNA CLAUDIU RADU<sup>C</sup></i>	
<b>KAMERA DEPROJEKCIÓ FELHASZNÁLÁSA AUTONÓM KÖRNYEZETÉRZÉKELÉS TERÜLETÉN.....</b>	<b>52</b>
<i>MARKÓ NORBERT<sup>A</sup>, KRECHT RUDOLF<sup>B</sup></i>	
<b>FENNTARTHATÓSÁGI KÖZZÉTÉTELEK SZÓGYAKORISÁGI ELEMZÉSE A SCOPE 3 KÁROSANYAG KIBOCSÁTÁSOK MÉRÉSÉRE PYTHON KÖRNYEZETBEN.....</b>	<b>61</b>
<i>MOLNÁR PÉTER<sup>A</sup>, SUTA ALEX<sup>B</sup>, TÓTH ÁRPÁD<sup>C</sup></i>	
<b>KAMERA ALAPÚ INTELLIGENS BUSZVEZETŐ MONITORING RENDSZER TERVEZÉSE .....</b>	<b>73</b>
<i>NAGY VIKTOR<sup>A</sup>, BÉZI PATRIK<sup>B</sup></i>	
<b>ESG MÉRŐSZÁMOK VÁLLALATI GYAKORLATAI ÉS KÖZELÍTÉSE A TECHNOLÓGIAELFOGADÁS-MODELLEKHEZ.....</b>	<b>96</b>
<i>PINTES ORSOLYA<sup>A</sup>, SUTA ALEX<sup>B</sup>, TÓTH ÁRPÁD<sup>C</sup></i>	
<b>AZ EGYETEMI TECHNOLÓGIATRANSZFER JOGI VETÜLETEI KÜLÖNÖS TEKINTETTEL AZ AKADÉMIAI SPIN-OFF VÁLLALKOZÁSOKRA .....</b>	<b>109</b>
<i>STIPKOVITS TAMÁS ISTVÁN</i>	
<b>SZOFTVERARCHITEKTÚRA TERVEZETE EURÓPAI ÜZLETI ADATOK AUTOMATIZÁLT KINYERÉSÉRE ÉS KEZELÉSÉRE .....</b>	<b>121</b>
<i>SUTA ALEX<sup>*</sup>, TÓTH ÁRPÁD, KEDVES LORÁND</i>	
<b>UTÁNZÁS ALAPÚ GÉPI LÁTÁSSAL TÁMOGATOTT AUTOMATA ROBOTPROGRAM GENERÁLÁS.....</b>	<b>134</b>
<i>TÓTH PATRIK<sup>A</sup>, TAKÁCS PATRIK<sup>B</sup>, MONEK GERGŐ DÁVID<sup>C</sup></i>	



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

RRF-2.31-21-2022-00004

Konzorciumvezető: Számítástechnikai és Automatizálási Kutatóintézet (SZTAKI)

Konzorciumi partnerek:

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

A projekt teljes összege: 9 439 666 600 Ft

Támogatási összeg (SZE): 470 000 000 Ft

Támogatás mértéke: 100%

A projekt időtartama: 2022.03.01. – 2025.06.30.

### Projektismertető:

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

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

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

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

# Stop vonal- és parkolóhely felismerés kamera, illetve LIDAR adatokkal a Shell Eco-marathon autonóm versenyén

Stop sign and parking place detection with camera and LIDAR  
datas for Shell Eco-marathon autonom race

**Bézi Patrik**

Széchenyi István Egyetem, Járműipari Kutatóközpont

bezi.patrik@ga.sze.hu

## **Absztrakt**

---

*Bármilyen autonóm funkcióval rendelkező jármű számára fontos a környezetfelismerés. Ebben a cikkben bemutatásra kerül a SZEmission nevű autonóm jármű stopvonal felismerő rendszere, amely az egyes pályaszakaszokat különíti el a Shell Eco-marathon autonóm versenyén. A feladat megoldása során többféle képfeldolgozó algoritmust próbáltunk ki. A végső kihívást a parkolóhely felismerése jelenti, amelyhez az egyes LIDAR adatokat használtuk fel.*

**Kulcsszavak:** lokalizáció, klasszifikáció, detekció, érzékelés

## **Abstract**

---

*Any autonomous vehicle requires a high precision description about its environment. In the following article the stop line detection of the vehicle named SZEmission will be presented. These lines separate the sections of the track of Shell Eco-marathon autonomous racing. During the solutions of the task more image processing algorithms have been tested. The final challenge is detecting a parking place, for that we used LIDAR data.*

**Keywords:** localization, classification, detection, perception

## 1. Bevezetés

A Shell Eco-marathon minden évben egyedülálló kihívás elé állítja a fiatal fejlesztőket, kutatókat világszerte: olyan járművet kell építeniük, amely a lehető leghosszabb távot tudja megtenni egy liter üzemanyag, vagy 1 kilowattórának megfelelő energia segítségével. A verseny célja, hogy inspirálja a résztvevők a saját tudásukat, erőforrásaikat, kreativitásukat használva innovatív megoldásokkal kísérletezzenek, az energiahatékonyság határait feszegetve megtervezzék.

A rendezvényen van külön autonóm kategória, amelyben olyan különböző kihívásokat kell teljesítenie a csapatoknak, mint a környezetérzékelés, útvonaltervezés, akadály elkerülés és parkolás. Ebben a cikkben bemutatásra kerül az egyes pályaszakaszokat elválasztó stopvonal felismerő rendszer, illetve a parkolóhely felismerés.

## 2. Áttekintés

Az autonóm verseny során a járműveknek 3 feladatot kell teljesítenie, amelyek a következők:

- hosszabb útszakasz bejárása
- akadály elkerülés
- parkolás

Az egyes kihívások teljesítése során a kocsinak önállóan meg kell tudnia állnia a kijelölt helyeken, mielőtt elkezdené a következő feladat végrehajtását.

A tavalyi szabályokhoz képest annyi módosítás történt, hogy a teljes pályát megszakítás nélkül kell teljesítenie az autónak. Ebből kifolyólag 3 helyen kell megállnia: 2-szer a már említett stopvonalak előtt, egyszer pedig végre kell tudnia hajtani egy sikeres parkolási manővert, amely során 3 parkolóhelyből 2 foglalt. A jármű szoftverének tehát azt is sikeresen el kell tudnia dönteni, hogy melyik ez az egy hely. Ezt követően olyan útvonalat kell tervezni az autó számára, amely segítségével a jelölés érintése nélkül megfelelően le tud parkolni, különben a teljes feladat érvénytelennek számít. A csapatoknak összesen 4 próbálkozási lehetősége van. A pálya felépítése az **1. ábrán** látható.



**1. ábra:** A Shell Eco marathons versenypályájának felépítése (saját szerkesztés)

### 3. Stop vonal felismerés

Stop vonal felismerés során egy statikus maszkot készítettünk elő, amelynek során levágjuk a horizont vonalat a képről, valamint az autó belógó részeit. A maszk a **2. ábrán** látható.



**2. ábra:** A képi bemeneten alkalmazott statikusmaszk (saját szerkesztés)

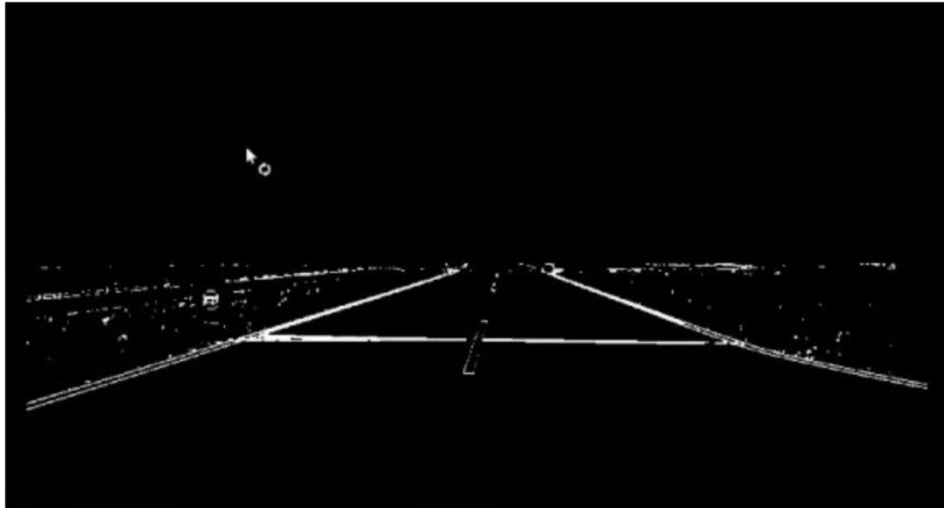
Maszkolás után kellett egy élkeresési eljárást alkalmazni, amelyre a Sobel szűrőt választottuk. Ez az eljárás gradienseket keres pixelintenzitás alapján.[1] A kimenet egy viszonylag zajoktól mentes bináris kép, amelyeken már könnyen el lehetett kezdeni a vonalkeresést. Az alkalmazott kernelt és a konvolúciót tartalmazza az **1. és a 2. képlet**, a kimenetet pedig a **3. ábra**.

$$K_y = \begin{bmatrix} p & q & p \\ 0 & 0 & 0 \\ -p & -q & -p \end{bmatrix}, \text{ ahol } p = 1, q = 2$$

(1)

$$\frac{\partial I}{\partial y} = K_y * I$$

(2)



**3. ábra:** A Sobel szűrő kimenete (saját szerkesztés)

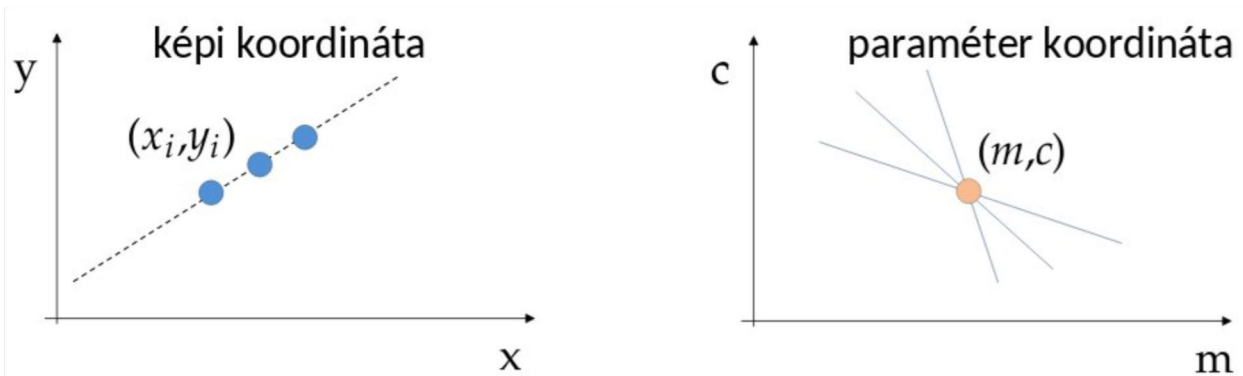
Megfelelő maszk készítése után következnek a megfelelő régió belüli vonalkeresés. Ehhez az OpenCV által biztosított HoughLine algoritmust használtuk. Ehhez írjuk fel az egyenes egyenletét a következő módon:

$$y_i = mx_i + c \quad (2)$$

ahol  $m$  a meredekség,  $c$  pedig az  $y$  metszéspont. Az egyenletet át lehet alakítani a következő formára:

$$c = mx_i + y_i \quad (3)$$

Ennek jelentősége akkor jön elő, ha ezek segítségével Descartes-féle koordináta-rendszerben szeretnénk ábrázolni a kapott értékeket, így fel tudunk venni  $xy$  és  $mc$  koordináta-rendszereket, ahogy az a **4. ábrán** látható. Ha megnézzük az adott egyenesen átmenő, adott  $(x_i, y_i)$  ponthoz tartozó  $m$  és  $c$  értékeket, fel tudunk írni egy másik egyenest. Ez az egyenes tartalmazza az összes olyan  $m$  és  $c$  paramétert, amelyekre felírva az egyenletet megkapjuk azon egyeneseket, amelyek áthaladnak az adott  $(x_i, y_i)$  ponton. Ha az egyenes több pontján is végrehajtjuk ezt, akkor láthatjuk, hogy a paraméter-koordináta-rendszerben ábrázolt egyenesek egy pontban metszik egymást. Ez a pont lesz az az  $(m, c)$  érték, ahol rajta van az összes pont a képi koordináta-rendszer adott egyenesén, amely a vonal.[2]

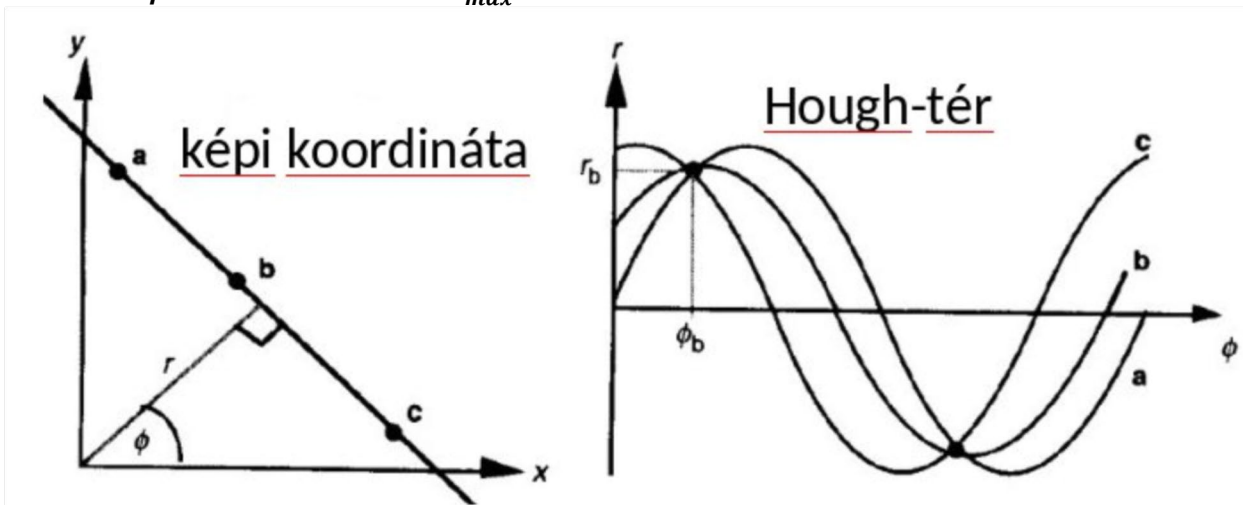


**4. ábra:** A képi koordináta és a Hough-tér az  $m$  és  $c$  paraméterek szerint (saját szerkesztés)

A Hough-transzformáció egy másik és könnyebben finomhangolható megvalósítása, amikor a Houghtérben szinuszoid görbék találhatóak. Az elv hasonló a vonalas megvalósításhoz, vagyis, ahol ezek a görbék metszik egymást a felvett értékeknek megfelelően, ott biztos, hogy található lesz vonal. A szinuszoid görbét a **4. képlet** írja le, az **5. ábra** pedig szemlélteti.

$$r = x_i * \cos \varphi + y_i * \sin \varphi \quad (4)$$

ahol  $0 \leq \varphi \leq 2\pi$  és  $0 \leq r \leq r_{max}$ .



**5. ábra:** A képi koordináta és a Hough-tér szinuszoid görbéi [3]

Önmagában ez az algoritmus nem egy, hanem az összes létező vonalat megtalálja, így a vonalak felismerése után következik a szűrés, melynek során a detektált vonalak közül csak azokat kell megtartani:

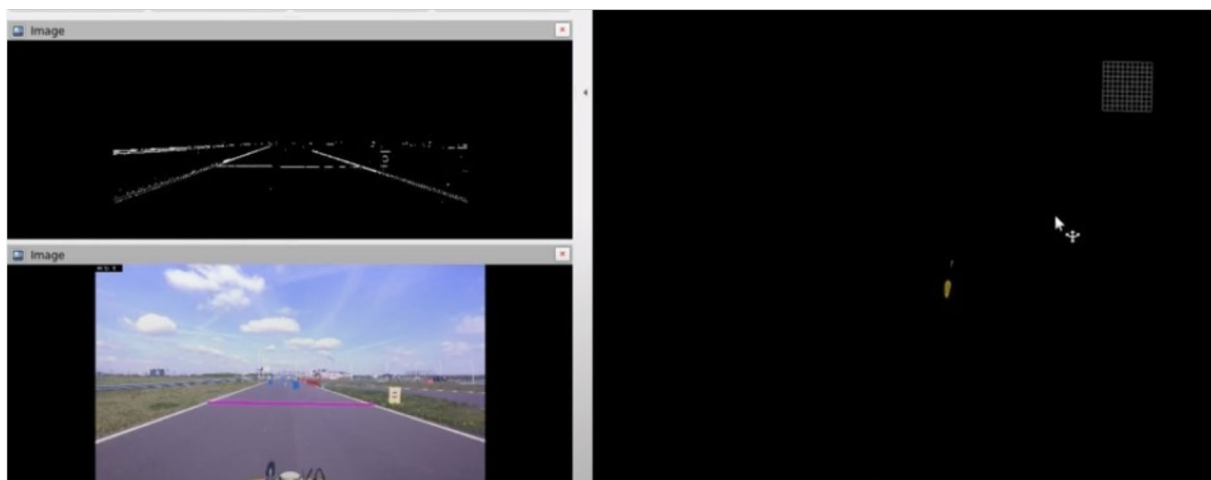
- amelyek minimális hosszúsága az általunk meghatározott értéket meghaladja
- amelyek bizonyos távolságán belül nem található még egy kisebb vonal és ezt a távolságot mi határozzuk meg
- amelyek a kép síkjához képest nem haladnak meg egy általunk meghatározott maximális fokot (közelítve a vízszintes vonalakat)

A szűrés után a vonalakat két csoportra osztottuk bal és jobb oldal szerint, attól függően, hogy a kép melyik oldalára esnek a pontjaik. Végül a két oldal egy-egy vonalának a legszélén elhelyezkedő pontját összekötve megkapjuk a stop vonalat. A hiba szűrője a stop tábla, hiszen, amíg azt nem ismerjük fel, nem kell foglalkozni a fals pozitív detekciókkal. A felismerés után esetlegesen beugró hibás detekciók szűrésére beállítottuk, hogy csak akkor hirdessünk pozíciót, ha legalább annyi képkockán történik egymás után detekció, amennyit meghatároztunk. (Ez az érték alapesetben 2, de dinamikusan állítható). A felépített vonal kimenete a **6. ábrán** látható.



**6. ábra:** A felismert vonal látható a bal alsó ábrán (saját szerkesztés)

Az említett pontokat ez követően egy kameramátrix segítségével átvetítettük világ koordináta-rendszerbe. A deprojekció során természetesen akadnak pontatlanságok, hiszen nem mindegy a kamera dőlésszöge és elhelyezkedése az autón, valamint egy bizonyos távolságon túl nem nevezhető megbízható megoldásnak. Ezt a problémát is hivatott orvosolni a stop tábla felismerő, valamint az egymás utáni detekciók minimális száma. A felismert vonalnak ezt követően vettük a felezőjét és megfelelő szöget állítottunk rá, amelyet ROS topicon keresztül hirdetünk a tervező számára, hiszen a Robot Operatin System nevű köztes szoftvert alkalmaztuk a fejlesztés során.[4] Ezt láthatjuk a 7. ábrán.



7. ábra: A hirdetett pozíció megfelelő orientációval a vonalfelismerés után (jobb oldalt) (saját szerkesztés)

#### 4. Parkolóhely detektálás

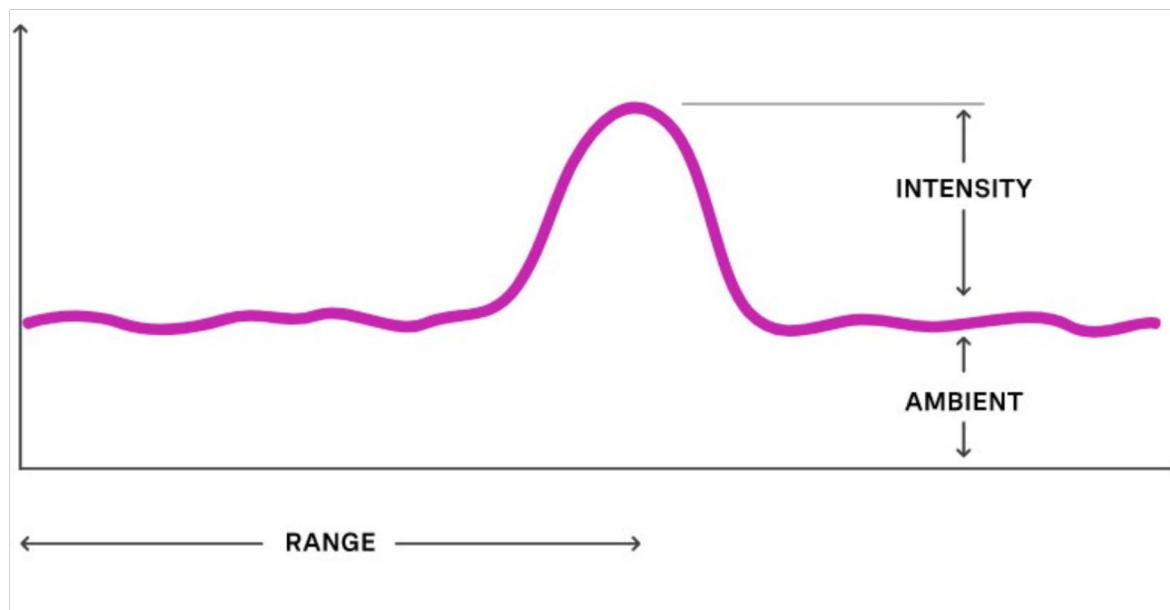
A parkolóhely felismeréséhez két szenzor adatait használtuk fel:

1. A kamera által biztosított képi információkat
2. A LIDAR szenzor által biztosított pontfelhő információit

A LIDAR szenzor által kibocsátott fény visszaverődéséből származó fotonokhoz a következő értékeket tudjuk társítani:

- **Intenzitás:** a legkisebb és az adott pontban mért fotonszám közötti különbség
- **Ambiens érték:** adott pontban mért fotonok számossága.
- **Range:** az adott mérés idejének mértéke
- **Reflektivitás:** az intenzitásérték és a range érték négyzetének szorzata

A felsorolt értékek egymáshoz viszonyított kapcsolatát ábrázolja a **8. ábra**.



**8. ábra:** A LIDAR szenzorból kapott értékek kapcsolata [5]

A megfigyelések alapján az egyes csatornák átlagos értéke a következőképpen alakult aszfalton a **1.**

**táblázat** láthatóak.

**1. táblázat:** Az aszfalthoz tartozó LIDAR értékek (saját szerkesztés)

Helyszín	Időjárás	Intenzitás	Ambiens érték	Reflektivitás
ZalaZONE	napos	100	350	6
Széchenyi István Egyetem	napos	135	320	6
Széchenyi István Egyetem	esős	30	100	6

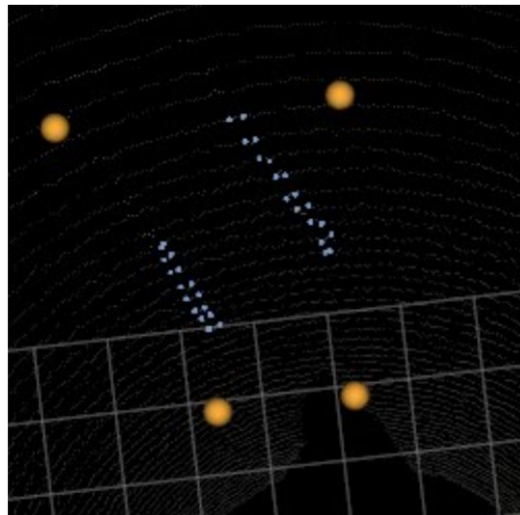
A **2. táblázat** pedig tartalmazza ugyanezen értékeket, amelyek a parkolóhelyet jelölő szalagokhoz tartoznak.

**2. táblázat:** A parkolóhely vonalához tartozó LIDAR értékek (saját szerkesztés)

Helyszín	Időjárás	Intenzitás	Ambiens érték	Reflektivitás
ZalaZONE	napos	40	800	2
Széchenyi István Egyetem	napos	30	450	2
Széchenyi István Egyetem	esős	25	130	4

Ez alapján az intenzitás értékeket és az ambiens értékeket használtuk fel a továbbiakban, hiszen kiugró különbségek ezen két csatornán voltak észlelhetőek, főleg napos időben.

Az első eleme a felismerésnek ugyancsak a már korábban említett deprojekció segítségével történik: a YOLO algoritmus, amelyet különböző objektumok felismerésére fejlesztettek, [6] egy befoglaló geometriával határolja körbe a parkolóhelyet. Kimenetként megadja ennek a bal felső sarkának pixelkoordinátáit, illetve a szélességét és magasságát, valamint a megbízhatósági értéket, vagyis hogy az algoritmus hány százalékban biztos, hogy parkolóhelyet látott. ROS rendszeren belül csak azok az adatok kerülnek hirdetésre, amelyek a legnagyobb megbízhatósági értékkel rendelkező parkolóhelyhez tartoznak. [5] Így ha mindhárom üres parkolót detektálja, az sem jelent problémát.



**9. ábra:** A YOLO által meghatározott befoglaló geometria 4 sarokpontja térbe vetítve (saját szerkesztés)

A sarokpontokat a deprojekció segítségével világkoordinátába vetítettük, és csak az ezen a régióban belüli LIDAR pontokat vizsgáljuk. Ez látható a **9. ábrán**.

A verseny során is ragasztószalagot alkalmaztak ezeknek a kijelölt helyeknek a felépítésére. Az adott trapézon belül a LIDAR pontokat gyűrűnként vizsgáljuk jobbról balra haladva. Utóbbi oka, hogy az indexelés a szenzor felhelyezésének módja miatt jobbról indul. Az eljárás során összefüggő értékű pontokat keresünk és ha szakadás van, akkor megállunk az adott gyűrűn.

Ennek lényege a következő: ha találunk 2 db pontot, amelyek meghaladnak egy előre meghatározott küszöbértéket, akkor megfigyeljük a köztük lévő távolságot. Ha

$$mindist < dist < maxdist, \quad (5)$$

ahol *mindist* a minimális távolság és *maxdist* a maximális távolság, valamint *dist* a két pont közti távolság, akkor elfogadjuk, hogy megtaláltuk a parkoló egyik oldalát. Mivel a LIDAR kifejezetten pontos és tudjuk előre a jelölések méretezéseit, így ez egy viszonylag egyszerű feladat.

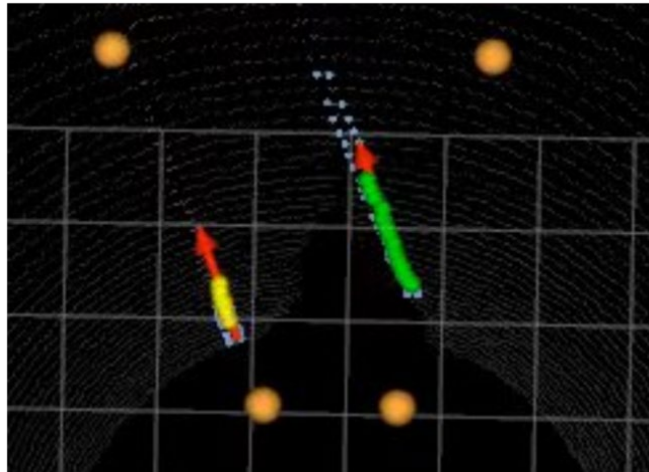
A fő kihívást az egyes pontok megfelelő oldalra való besorolása jelentette. Alapvetően a trapéz két autóhoz közelebbi sarokpontját vettük alapul. Ha a jobb oldalához volt közelebb az első pont, akkor jobb oldali, ha a bal oldalához, akkor a bal oldali osztályba került be. Ennek matematikai leírása a következő módon néz ki: legyen  $T_{jobb}(x, y)$  és  $T_{bal}(x, y)$  a legközelebbi jobb és bal oldalai a kivetített trapéznak, valamint legyen  $p(x, y)$  az aktuálisan vizsgált pont. Ha

$$\sqrt{(x_p - x_{T_{jobb}})^2 + (y_p - y_{T_{jobb}})^2} < \sqrt{(x_p - x_{T_{bal}})^2 + (y_p - y_{T_{bal}})^2} \quad (6)$$

akkor  $p$  pont a jobb oldalra lesz besorolva. Ellenkező esetben bal oldali pontok közé kerül. Ezután – ahogy korábban említésre került,  $T_{jobb}$ , illetve  $T_{bal}$  szerepét átveszi  $p$  pont, az elv pedig ugyanez lesz.

Ezt követően alkalmaztunk egy olyan eljárást, hogy a többi pontot mindig a legutóbb elhelyezett jobb-, illetve bal oldali ponthoz viszonyítottam. Ez segít kiszűrni, ha nagyobb szögben próbálja az autó megközelíteni a parkolóhelyet.

Ha két pontot sikerül találni egy gyűrűn, akkor a fenti módszer egyszerű. Ha azonban csak egyet talált meg az algoritmus, kellett egy módszer az adott pont besorolásának eldöntésére. Ez azért fontos, mert ha mindig csak a két-két ponttal rendelkező gyűrűket vesszük figyelembe, akkor könnyen előfordulhat, hogy túl távoli pontok alapján akar tájékozódni az autó. A **9. ábra** szemlélteti ezt a problémát is, hiszen jól látható, hogy a parkolóhely bal oldala jóval közelebb van, mégsem tudnánk figyelembe venni, mert nem tartozik az adott gyűrűkhöz jobb oldali pont is. A pontok osztályba sorolása látható a **10. ábrán**.



**10. ábra:** A pontok osztályba sorolása (saját szerkesztés)

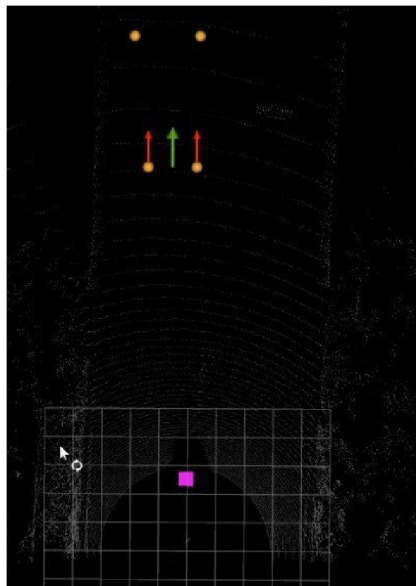
Ha sikerült a pontokat osztályokba sorolni, akkor a kezdő és végpontok segítségével tudunk húzni 1-1 egyenest, megfelelő orientációval.

Pozícióhirdetés már akkor történik, mikor a YOLO felismerte az üres parkolóhelyet. Erre azért van szükség, mert messziről felismerve a célt hamarabb el tud kezdeni az autó a megfelelő szögben ráállni a parkolóra. Ennek módja, hogy  $T_{jobb}$ , és  $T_{bal}$  pontokat összekötő egyenest megfelezzük és merőlegest állítunk rá. A **7. és 8. egyenletek** leírják ennek a módszerét.

$$C = \left( \frac{x_{T_{jobb}} + x_{T_{bal}}}{2}, \frac{y_{T_{jobb}} + y_{T_{bal}}}{2} \right) \quad (7)$$

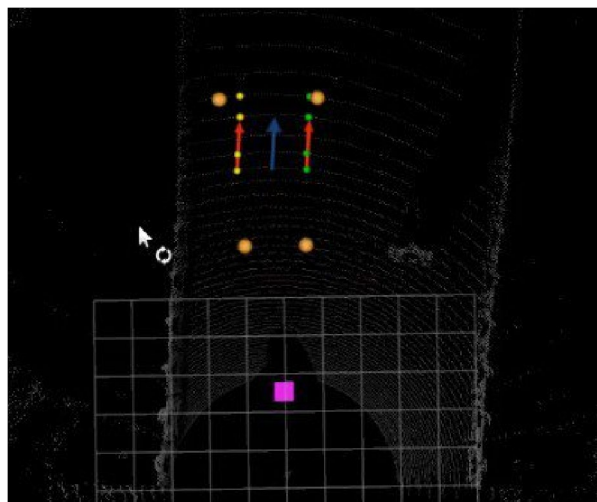
$$m_C = \arctan2 \left( x_{T_{jobb}} - x_{T_{bal}}, y_{T_{jobb}} - y_{T_{bal}} \right) * \frac{180}{\pi} + 90 \quad (8)$$

ahol  $C$  a célpozíció,  $m_C$  pedig a meredekség. Az eljárás vizuális megvalósítását reprezentálja a **11. ábra**.



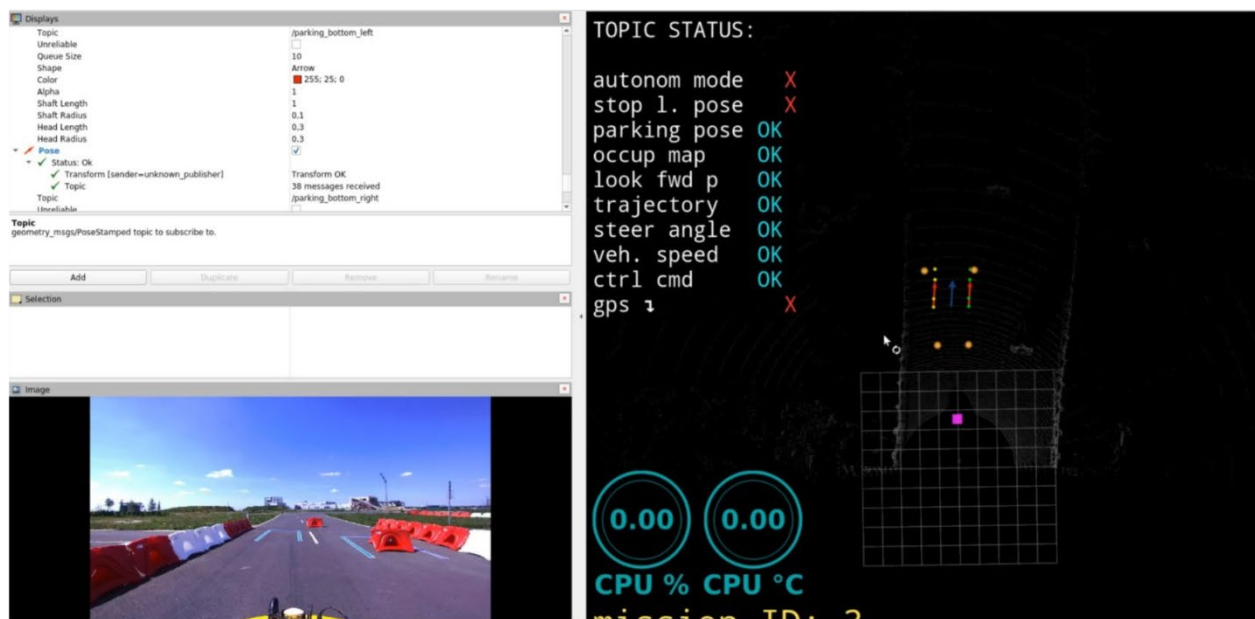
**11. ábra:** Pozícióhirdetés YOLO és deprojekció segítségével (saját szerkesztés)

Hasonló módon történik a pozícióhirdetés, ha megfelelő mennyiségű besorolt pontot találunk. Jelen esetben minimum 4 pontot kell találni mindkét oldalra, hogy megbízhatóan tudjunk pozíciót hirdetni a LIDAR pontok alapján. Ennek vizuális megjelenítése látható a **12. ábrán**. Ha a pontok megvannak, akkor a két egyenes kiindulópontját kötjük össze és az így kapott egyenesre állítunk merőlegest.



**12. ábra:** Pozícióhirdetés LIDAR pontok alapján (saját szerkesztés)

Végül a teljes kép a rendszer működéséről látható, ahol megfigyelhető, hogy adott időpillanatban mit érzékel a környezetből az autó. Jól látható, hogy az autó viszonylag távol – kb 8 méterre – van a parkolóhelytől, a LIDAR pontok segítségével azonban már sikeresen pozicionál a megfelelő helyre.



13. ábra: Szenzoradatparkolóhelydetektálásakor (saját szerkesztés)

## 5. Tesztelés

A tesztelés során a stop vonal felismerést próbáltuk ki először. Ehhez különböző típusú ragasztószalagokat használtunk (különböző színűeket és fényességűeket). Az első tesztek alapján kiderült, hogy a hivatalosan versenyen alkalmazott táblákon a stop táblát a YOLO sokkal nehezebben ismerte fel, mint a normális közutakon használt táblát. A neurális hálózatot ezért újra kellett tanítani újonnan címkézett képekkel, amelynek eredményeként hamarabb felismerte a rendszer az előbb említett táblát. A vonalat időjárástól függetlenül kivétel nélkül mindig felismerte az autó a tesztek során, mivel megfelelő helyen állt meg.

A parkolási feladat komplexebb volt, az általunk alkalmazott megoldást vetettük össze hagyományos, csak képi bemeneteket alkalmazó rendszerrel. A LIDAR pontok ambiens értéke nagyban függ az időjárástól. Tapasztalataink alapján száraz időben nagyon szépen elkülönülnek a vonalhoz tartozó pontok ezen értékei az aszfaltétól. Ha azonban nedves az aszfalt, akkor jóval nehezebb a dolgunk és nem könnyű meghatározni, hogy mekkora ez a küszöbérték.

## 6. Összegzés

A Stop vonal felismerő megbízhatóan teljesítette a feladatát, a YOLO megfelelő tanítása mellett, megbízhatóan sikerült megállni a megfelelő pozícióban.

A LIDAR alapján működő parkolóhely felismerés pontos eredményeket ad megfelelő időjárási körülmények között. A továbbiakban olyan fejlesztéseket kell eszközölni, amelyek lehetővé teszik, hogy kedvezőtlenebb feltételek mellett is jól működjön az algoritmusunk. Ehhez kamera-LIDAR fúziós eljárásokat tervezünk eszközölni a jövőben.

## 7. Köszönetnyilvánítás

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

## Irodalomjegyzék

- [1.] Bernd Jähne - Digital Image Processing 6th revised Edges p331-357 Berlin 2005
- [2.] Bernd Jähne - Digital Image Processing 6th revised Segmentation p449-462 Berlin 2005
- [3.] Kató Zoltán Alakzatok és minták detektálása - Képfeldolgozás és Számítógépes Grafika tanszék SZTE
- [4.] <https://wiki.ros.org> (2023.06.15.)
- [5.] [https://github.com/ousterlidar/ouster\\_example/issues/177](https://github.com/ousterlidar/ouster_example/issues/177) (2023.06.15.)
- [6.] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi: You Only Look Once: Unified, Real-Time Object Detection, arXiv preprint arXiv: 1506.02640, 2015

# Public transport service planning with the support of AI

## Tömegközlekedési szolgáltatás tervezése a mesterséges intelligencia alkalmazásával

**Zoltán Hima**

Széchenyi István University

zoltan.hima@gmail.com

### Abstract

---

*Artificial intelligence fundamentally changes our everyday life. Public transport is a mode of transport whose accessibility is extremely important both for the population and for the companies providing the service. The regularity and quality of the service can only be ensured and sustained in the long term with adequate fee income and cost efficiency. The spread of tickets that can be redeemed electronically opens a new dimension that was not possible with the previous, paper-based payment. Through the fare exchange habits of the traveling public, service providers can form samples of the traveling public, and track the use of redeemed tickets and passes, but finally, they can see the use of public transport in real-time. Since people participating in public transport are entitled to various discounts, such as student discounts, occupational discount, discounts for people with reduced mobility, discount for the elderly, and other social discounts, the discount for each entitlement can be tracked and checked digitally. Currently, transport companies start their flights based on the timetable and their previous experience, based on which both under- and over-planning occurs. In the case of fixed-track transport, the capacity utilization of the required vehicles can be partially planned for rail passenger transport, while for non-fixed-track transport, for road passenger transport, the capacity utilization can be planned more precisely.*

**Keywords:** public transport, fee control, capacity utilization

The Artificial Intelligence Strategy of Hungary gives priority to the issue of the development of transport, it aims to make our country a shaper of innovation in the field of the use of artificial intelligence in the transport industry. In addition to public transport improvements, the domestic plans also include making traffic management and traffic organization more efficient, as well as building digital infrastructure. The widespread introduction of AI technology means challenging comprehensive development programs that use AI technology to create value for citizens. The implementation of these transformative programs aims, on the one hand, at the thorough transformation of the sector concerned, and on the other hand at the everyday use of AI by the wider society. The goal is to create a supportive environment that enables the joint development of the various branches of the self-driving ecosystem. In the field of transport, this includes the introduction of AI-based automatic traffic control, the optimization of urban public transport networks and flights, the counting of passenger traffic, and the development of smart parking systems. With the use of artificial intelligence, in the future, real-time traffic control activities can be implemented in major domestic cities by building image-based traffic management technologies and their corresponding camera systems based on the processed information. In public transport, based on real-time traffic data, it will be possible to introduce fleet management solutions optimized for keeping to the schedule. Real-time vehicle tracking can be made available to users through an application interface that dynamically predicts the expected arrival time by monitoring the traffic situation. (1)

The EU itself also focuses on the issue. Decree 2019/881 on network and information systems and telecommunications networks and services play a vital role in the functioning of society and are the backbone of economic growth. Information and communication technologies form the basis of complex systems that support everyday social activities, ensure the smooth operation of the economy in key sectors such as healthcare, energy, finance, and transport, and above all promote the functioning of the internal market. The directive applies to services provided to consumers after 28 June 2025. These include the following elements of air, bus, rail, and water passenger services, and urban, suburban, and regional transport services. (2)

Directive 2010/40/EU of the European Parliament and the Council deals with the framework for the development of intelligent transport systems in the field of road transport and their connection to other modes of transport. By Intelligent Transport System we mean the unified system made up of information communication technologies used in transport, which can be used to optimize transport modes - improve cost efficiency, - reduce the environmental load, - improve the safety, information, and comfort of transport both from a social and individual point of view. (3) To adapt to the changed shopping habits, and to provide faster and safer service, from August 1, BKK will make new ticket and pass machines cashless. The step reduces the possibility of abuse, and the operation of the devices will be more predictable and reliable. More and more people are buying tickets and pass with a bank card, which is why the Budapest Transport Center has decided to make twenty-five additional machines cashless from August 1st to serve customers quickly and conveniently. Thanks to the company's measures, the operation of twenty-five percent of the current 315 automatic machines will be more predictable and reliable. The service life of the machines can be extended further, so it will be more economical to maintain the network. Thanks to the move, there will be more and more devices in the city that can no longer be manipulated in any way, i.e. they cannot be harmed by criminals who specialize in this. (4) Google will soon monitor several intersections with artificial intelligence to make the green signal as fast and efficient as possible for motorists. A cooperation agreement has been signed between Google and Jakarta City Transport Managers. According to the Indonesian news agency, a traffic system is being developed in the capital that will rely on artificial intelligence to help motorists. According to the report, the new system will be introduced at many intersections. The artificial intelligence will turn the traffic lights green at the optimal rate and time so that the traffic passes through it as smoothly as possible. The cooperation between the Jakarta provincial government and the company is expected to help reduce traffic congestion. Many people will certainly be happy about this, as this is one of the most common complaints from residents. The level of congestion is unbearable, especially during peak hours. (5) It is necessary to analyse processes from different perspectives, such as customer and employee perspectives, as this is essential for the successful implementation of AI-driven innovations in public transport. (6)

Artificial intelligence has a significant impact on the railway industry and is developing day by day. Railways are increasingly relying on digital technologies, and this is having a positive impact on both passenger safety and freight operations. (7) This industry has come a long way since the 1800s. Trains are now run by computers, and digital technologies are used to manage everything from passenger ticketing to railway equipment. The railway industry is facing many changes in the coming years due to the growing role of artificial intelligence. This is expected to have a significant impact on train operations, passenger ticketing, system maintenance, and rail assets. One of the most important aspects of rail transport is safety. Artificial intelligence can play a big role in guaranteeing this. Passengers should feel safe when they are on a train and railway companies should make sure that only legitimate passengers board the train. Is it possible to determine whether a particular person can travel on the train? For this reason, for example, cameras and facial recognition software are used to ensure that only authorized persons are allowed on the trains. This helps ensure that unauthorized people cannot board the train. Railway companies also use smartphone applications to provide accurate information on current news. Smart stations are also becoming more common. These stations use sensors to track the number of people visiting the railway facilities and their movement within the station. This information can be used to improve the design of the stations and make them more user-friendly. (8)

Significant advances in digital technologies such as artificial intelligence, the Internet of Things, cloud and edge computing, or 5G networks are playing a leading role in the digital transformation of the mobility industry.

Autonomous vehicles are on the rise, enabling new transport solutions such as robot taxis, connected trucks, or driverless transport systems. New mobility services enabled by digital platforms are changing from shared mobility, such as cycling, to Mobility Service, which provides access to multiple modes of transport through a single application. Freight transport, on the other hand, relies to a greater extent on a greater degree of synchronization between modes of transport. The data generated and exchanged by vehicles is exploding. Self-driving vehicles can generate about 4 terabytes of data per day. Vehicles are increasingly connected to their physical and digital environment. New habits and behaviours brought about by digitization, such as the rapid development of e-commerce for remote work, are transforming mobility patterns. (9)

Key benefits are that connected and automated mobility can make roads, railways, and waterways safer. By optimizing the use of vehicles and infrastructure, you can increase efficiency, reduce traffic congestion, and help reduce gas emissions. It can also create business opportunities and increase competitiveness and make transport more accessible for the elderly and disabled, transforming travel for years to come (10). Mobility as a service enables seamless multimodal travel and facilitates access to greener alternatives and can also contribute to the decarbonization of transport. Exploiting the complementary nature of different modes of transport would enable more efficient use of resources. Merging and exchanging mobility data can improve synchronization between different transport modes and infrastructures. Access to large data sets would facilitate the development of applications based on artificial intelligence. Better access to mobility data will help authorities monitor transport activities and their impacts, and plan transport infrastructure and services (11).

Access to mobility data can create new sources of value: data from car sensors, for example, can be extremely valuable for insurance, navigation, or road maintenance companies and, finally, support the transition to emission-free mobility. Artificial intelligence can improve the energy efficiency of electric or hydrogen vehicles and optimize the deployment of charging infrastructures. Two-way electric vehicle charging could provide flexibility to the grid, enabling optimization of energy use and production and better integration of renewable energies. Such industry connectivity requires digital capabilities and real-time data sharing. In dealing with technological challenges, intelligent transport systems require huge investments to develop and implement new technologies and infrastructures. To process large amounts of data in real time, we will need a cloud-wide computing continuum, 5G connectivity to ensure fast and reliable data transmission, artificial intelligence to analyse complex information, smart vehicle chips, and new types of automotive operating systems. Unity in diversity: the European mobility ecosystem is made up of a multitude of players with different specializations and sizes. This diversity is strong, but it can also lead to fragmentation. To reap the benefits of digitalization, mobility, and transport, stakeholders need to build partnerships, pool investments, and agree on common standards, infrastructures, platforms, and governance frameworks to reach critical mass and avoid large dependence on dominant actors (12). Security and privacy protection: the digitization and automation of transport pose challenges to the protection of personal data, cyber security, and user acceptance. Cyber-attacks can affect the operation of services, damage systems, data theft, and even cause injuries. Addressing social impacts: social acceptance of

self-driving mobility will be key to its implementation. This transition requires new skills and may lead to the disappearance of certain types of jobs. Therefore, it must be handled well. Some new forms of mobility create types of work characterized by more flexible conditions that must respect EU labour law and ensure social protection.

Today, many applications support the traveling public as well as transport companies. The MÁV Elvira application now not only provides complex train timetables but can also help based on voice, supported by artificial intelligence. (13) The Citi AI development of the Budapest University of Technology works with real-time data and strives to provide city-specific, almost individually tailored help. (14) But applications are already available where the system not only analyses but also predicts the traffic conditions, so it organizes already planned, gives intervention suggestions to the dispatcher, and can even send messages to the public transport passenger information system. Travelers and passengers will thus be able to use a route planner tailored to the needs of the given city, supported by artificial intelligence and effective intervention, instead of the usual route planning applications. The system not only analyses, but also predicts the traffic conditions in advance, so it organizes already planned, gives intervention suggestions to the dispatcher, and can even send a message to the public transport passenger information system (15). Travelers and passengers will thus be able to use a route planner tailored to the needs of the given city, supported by artificial intelligence and effective intervention (16), instead of the usual route planning applications. The city of Szeged has gone even further in terms of urban public transport. The sensor-based passenger and traffic counting system is based on technology patented by the Department of Software Development of the Faculty of Science and Informatics of the University of Szeged.

The sensors use artificial intelligence to recognize boarding passengers, thus determining the number of passengers. Since the system, so to speak only transmits piece numbers, personal data is maximally protected. (17)

And the fact that sometimes the most obvious things can be achieved even without major development projects is no better demonstrated by the Schedules (18) application, which always shows the latest data and even indicates the position of the vehicles on the map. The application can compile complex routes based on accurate, real-time data. In collaboration with Tata Consultancy Services and GFI Informatique (19), the Belfort public transport company equipped its bus fleet of almost 100 vehicles with GPS sensors and ID readers. With the help

of these, the journey time between individual stations could be accurately measured, and the passenger traffic could also be assessed in more detail. (20) For example, they examined where most people get off and on, but they also monitored the average passenger density of each flight. With the help of Big Data systems and artificial intelligence, the cost-effective solutions that can bring the greatest benefit were developed with the data. Extra vehicles can be diverted to routes with higher loads during peak periods, while the frequency of flights can be reduced on routes with lower utilization. (21)

## **Summary**

The effects of digitization can now be found in all forms of transport. New types of data analysis and processing procedures not only help to use forms of transport more efficiently but can also help the development of new transport methods. Not only cost efficiency can be realized on the part of transport companies, but the satisfaction of the traveling public can also increase to a large extent. As a result, the traveling public will be even more open to means of transport that provide not only efficiency or greater utility but also an even more environmentally conscious alternative. The primary purpose of using AI is to improve or improve the quality of service and understand the traveler's behavior. AI is mainly used for three tasks; the most common forecasting, then estimating the current state and allocating resources, including planning and scheduling. Most of the investigated artificial intelligence solutions require a significant amount of data with travelers. Machine learning the most commonly used artificial intelligence technology. There is great potential in using artificial intelligence to improve public transportation that is yet to be discovered, but there are procedures and challenges that need to be considered. However, the observations are often related to data, that are based on high-volume and high-quality data sets, but their production requires significant resources and time.

## **Acknowledgement**

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

## References

- [1.] <https://ai-hungary.com/api/v1/companies/15/files/137203/view>
- [2.] <https://eur-lex.europa.eu/legal-content/HU/ALL/?uri=CELEX%3A32019R0881>
- [3.] <https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32010L0040>
- [4.] Hybrid Decision Support Based on Knowledge Discovery and AI Techniques for the Management of Maintenance Services in the Public Transport Sector | IEEE Conference Publication | IEEE Xplore <https://en.antaranews.com/news/264855/jakarta-plans-to-use-ai-to-ease-traffic-congestion>
- [5.] Bus travel time prediction under high variability conditions on JSTOR
- [6.] Buics, L. & Süle, E. (2020). Service process excellence in public services. Proceedings of ENTRENOVA-ENTerprise REsearch InNOVAtion 6(1), 173–186.
- [7.] <https://digital-strategy.ec.europa.eu/hu/policies/digitalisation-mobility>
- [8.] A review on agent-based technology for traffic and transportation | The Knowledge Engineering Review | Cambridge Core [BDCC](https://www.cambridge.org/core) | Free Full-Text | Ticket Sales Prediction and Dynamic Pricing Strategies in Public Transport (mdpi.com)
- [9.] <https://www.mavcsoport.hu/mav-start/elvira-megszolal>
- [10.] Sustainability | Free Full-Text | Applications of Artificial Intelligence in Transport: An Overview (mdpi.com)
- [11.] Sustainability | Free Full-Text | Applications of Artificial Intelligence in Transport: An Overview (mdpi.com)
- [12.] Bus travel time prediction under high variability conditions on JSTOR
- [13.] Edge Solution with Machine Learning and Open Data to Interpret Signs for People with Visual Disability (ceur-ws.org)
- [14.] Massive Railway Operating Data Visualization; a Tool for RATP Operating Expert - ScienceDirect

- [15.] A comparative study of machine learning classifiers for modeling travel mode choice - ScienceDirect
- [16.] Urban Science | Free Full-Text | Spatial Factor—Using a Random Forest Classification Model to Measure an Internationally Comparable Urbanity Index (mdpi.com)
- [17.] Artificial intelligence for improving public transport: a mapping study | Public Transport (springer.com)
- [18.] (PDF) Artificial intelligence for improving public transport: a mapping study (researchgate.net)
- [19.] Machine Learning Aided Simulation of Public Transport Utilization | IEEE Conference Publication | IEEE Xplore
- [20.] Buics, L. & Eisinger, B. (2020). Analyzing public service processes from customer and employee perspectives by using service blueprinting and business process modelling. Proceedings of ENTRENOVA-ENTerprise REsearch InNOVAtion, 6(1), 195-211.
- [21.] (PDF) Artificial Intelligence in Transportation Systems A Critical Review (researchgate.net)

# Emberi arc detektálásának vizsgálata kapszula hálózatok alkalmazásával

## Human face detection using capsule networks

Hollósi János<sup>a</sup>, Ballagi Áron<sup>b</sup>, Pozna Claudiu Radu<sup>c</sup>

<sup>a</sup>Széchenyi István Egyetem, Informatika Tanszék; Járműipari Kutatóközpont

hollosi.janos@sze.hu

<sup>b</sup>Széchenyi István Egyetem, Automatizálási és Mechatronikai Tanszék; Járműipari Kutatóközpont

ballagi@sze.hu

<sup>c</sup>Transilvania University of Braşov, Automation and Information Technology Department

claudiu.radu63@gmail.com

### Absztrakt

---

*A kapszula hálózat elméletének egyik kulcsfontosságú eleme az úgynevezett útválasztó algoritmus. Ennek a mechanizmusnak köszönhetően a kapszula rétegek közötti információ áramlás optimálisához minél közelebbi iránya és sebessége kerül megvalósításra. Ez a megfontolás igyekszik kiküszöbölni a klasszikus konvolúciós neurális hálózatokban található kiejtő rétegek problémáját, ami akár lényeges adatok veszteséhez is vezethet. Munkánk során egy általunk javasolt, a dinamikus útválasztó algoritmus nyomán megkonstruált új módszer kerül bemutatásra egy gyakorlati alkalmazási példán keresztül. Ebben a megközelítésben a cél digitális fényképeken az emberi arc előre meghatározott komponenseinek a felismerése és a képsíkon vett pozíciójának a meghatározása. Megmutatjuk, hogy a kapszula hálózatok elméletének kidolgozói által javasolt dinamikus útválasztó algoritmus hatékonysága, erre a feladatra nézve, hogyan viszonyul az általunk bevezetett megoldáshoz képest.*

**Kulcsszavak:** neurális hálózat, kapszula hálózat, arc detektálás, útválasztó algoritmus

### Abstract

---

*A key element of the theory of the capsule network is the so-called routing algorithm. Thanks to this mechanism, the optimum approximate direction and speed of information flow between the capsule layers is realised. This approach tries to eliminate the problem of pooling layers in classical convolutional neural networks, which can lead to loss of essential data. In our work, a novel approach we propose, constructed following the dynamic routing algorithm, is presented through a practical application example. In this case, the goal is to detect predefined keypoints of the human face in digital images and to determine its position in the image plane. We show how the efficiency of the dynamic routing algorithm proposed by the theoreticians of capsule networks theory, for this task, compares to the solution we have introduced.*

**Keywords:** neural network, capsule network, face detection, routing algorithm

## 1. Bevezetés

A konvolúciós neurális hálózat egyre nagyobb népszerűsége tesz szert, ahol folyamatosan bővül azon gyakorlati alkalmazási területek halmaza, ahol megfelelő hatékonysággal felhasználhatók az ilyen jellegű megvalósítások. Ilyen hálózatokkal több különféle feladatot végre tudunk hajtani, de talán a legnépszerűbbek a különféle vizuális adatok feldolgozása, úgymint osztályozás [1, 2], detektálás [3, 4], szegmentálás [5, 6] vagy rekonstrukció [7, 8]. Ezzel egy időben látványosan növekedésnek indultak a konvolúciós neurális hálózatok megbízhatóságát vizsgáló kutatások [9, 10]. Hiszen a gyakorlati alkalmazások növekedésével egyre több olyan területen is felhasználásra kerülnek a mesterséges intelligencia alapú megoldások, ahol magas biztonságkritikus szintet képvisel a rendszer. Ilyen rendszerek esetén pedig nem megengedhető, hogy az intelligens rendszer téves döntést hozzon, mivel az emberi, környezeti vagy gazdasági károkhoz vezethet. Ezeket pedig minimalizálni kell. A konvolúciós neurális hálózat megbízhatóságát vizsgáló kutatások során napvilágot látott, hogy az ilyen rendszerek könnyen becsaphatóak, megtéveszthetőek. Ezekre különféle módszerek, például úgynevezett támadási módszerek kerültek kidolgozásra. [11, 12] A támadási módszerek esetén a hálózat belső paramétereit, értékeit alapján olyan módosítások végrehajtása történik a bemeneti adatokon, hogy azt a hálózat tévesen dolgozza fel, ne a megfelelő és elvárt kimenetet szolgáltatassa. Mindezt annak ellenére, hogy a módosítás csupán olyan mértékű, hogy emberi résztvevő számára nem okoz problémát a feldolgozása, megértése.

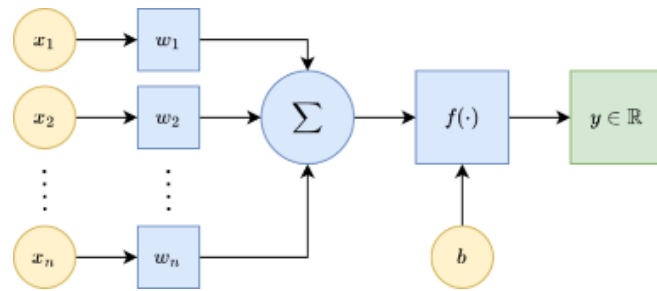
A konvolúciós neurális hálózatok ilyen és ehhez hasonló problémáinak kiküszöbölése céljából vezették be az úgynevezett kapszula hálózatok elméletét. [13-15] Az elmélet kidolgozása Geoffrey Hinton nevéhez köthető, aki a terület számtalan megoldásának az atyja, ma már alapvetőnek minősülő elméletek bevezetője. A kapszula hálózat elméletének célja, hogy jobban hasonuljon a biológiai neurális hálózat működéséhez. Ezzel pedig egy robusztusabb megoldást kínálva a mesterséges neurális hálózatok területén. A kapszula hálózatok kulcsfontosságú mozzanata az úgynevezett útválasztó optimalizáló algoritmus. [13] Ennek célja, hogy a kapszula rétegek közötti információ áramlás fő sodrását realizálja egy iteratív optimalizálási eljárás folyamán. Az elmélet szerzői bevezették az úgynevezett dinamikus útválasztó algoritmust erre a feladatra. Ez a megvalósítás megfelelő hatékonysági szintet képvisel. Ugyanakkor rejteget potenciálokat további fejlesztések tekintetében. A cikkben bemutatott munkánk során az általunk korábban bevezetett egyszerűsített útválasztó algoritmust [16] mutatjuk be emberi arcok detektálása során. Ebben a feladatban úgynevezett kulcsfontosságú pontokat kell

a hálózatnak felismernie, melyek az emberi arc előre meghatározott területeit jelölik, mint például a szemek vagy a száj sarkai. Munkánk során megmutatjuk, hogy milyen módon épül fel a saját megvalósításunk és ennek a feladatnak a tekintetében összehasonlítjuk a saját megoldásunkat a dinamikus útválasztó algoritmussal. A vizsgálat során különféle adatkészlet zajterhelési technikákat is alkalmazunk annak érdekében, hogy a rendszerek robusztusságát minél inkább felszínre hozzuk és láthatóvá váljon a két optimalizáló algoritmus közötti hatékonyságbeli eltérés.

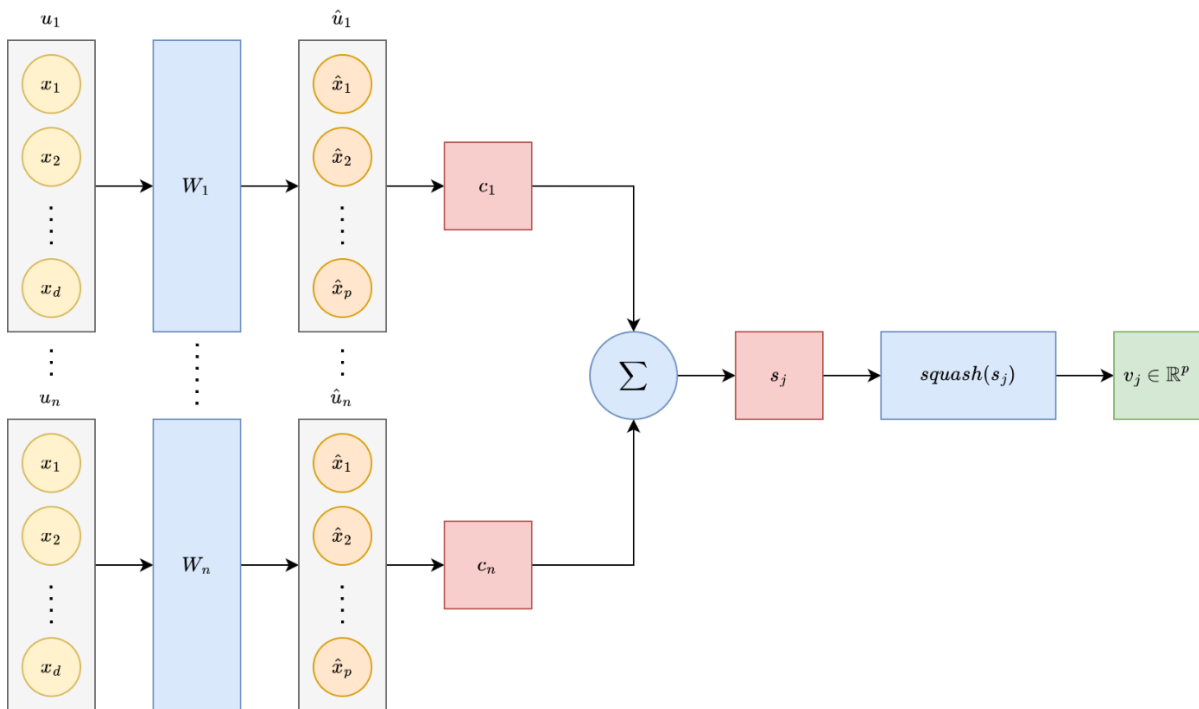
Jelen munkánk felépítése a következő. A 2. fejezetben bemutatjuk a kapszula hálózatok elméletének megértéséhez szükséges minimális szükséges háttérismeretet, rávilágítva a neuron és a kapszula közötti főbb különbségekre. A 3. fejezetben az általunk javasolt útválasztó optimalizáló algoritmust vezetjük be és mutatjuk be a különbségeket a dinamikus útválasztó algoritmushoz képest. A 4. fejezetben a mérések során felhasznált adatokat mutatjuk be, majd az 5. fejezetben a hatékonyság vizsgálatához felhasznált adatkészlet támadási módszereket foglaljuk össze. A 6. fejezetben a kialakított hálózatok tanítását és annak körülményeit tisztázzuk. Majd az ennek nyomán elért tényszerű eredményeket a 7. fejezetben közöljük. Végül a 8. fejezetben összegezzük a munkánkat, kitérve az elért eredményekre, valamint az azok nyomán levonható következtetésekre, végül a jövőt érintő kutatási és fejlesztési terveinket vázoljuk fel.

## **2. Kapszula hálózatok elmélete**

A kapszula hálózatok [14] elméletének kidolgozása Geoffrey Hinton nevéhez köthető, aki sok más alapvető elmélet és algoritmus kidolgozója a mesterséges intelligencia területén belül. Az új elméletet a hagyományos neurális hálózatok nyomán alkották, ugyanakkor igyekeztek olyan új módszert létrehozni, amely a felvázolt problémák esetén robusztusabb, mint a neurális hálózatok. Amíg a neurális hálózatok építőeleme a neuron, addig a kapszula hálózatok alapegysége az úgynevezett kapszula. A fő különbség a neuron és a kapszula között, hogy a neuron skalár értékekkel dolgozik, a kapszula pedig tetszőleges dimenziójú vektorokat használ. Úgy is tekinthetünk a kapszulára, mintha neuronok egy zárt, egybefüggő csoportját képeznék. A hálózatok esetén értelmezett különféle műveletek pedig ennek megfelelően módosulnak. Az 1. ábra a neuron és a kapszula közötti főbb eltéréseket szemlélteti.



Neuron



Kapszula

1. ábra: Neuron és kapszula közötti különbségek (saját szerkesztés)

Legyen  $i$  egy kapszula és  $j$  egy  $i$ -nél magasabb szintű kapszula. Ekkor legyen a súlyozott bemenet

$$\hat{u}_{j|i} = \mathbf{W}_{ij} \mathbf{u}_i \quad (1)$$

ahol  $\mathbf{W}_{ij}$  a súlymátrix és  $\mathbf{u}_i$  a bemeneti vektor az  $i$ -edik kapszula esetén.

Az úgynevezett összekapcsolási együttható megadása a softmax függvény segítségével történik a következők szerint

$$c_{ij} = \frac{\exp(b_{ij})}{\sum_k \exp(b_{ik})} \quad (2)$$

ahol  $b_{ij}$  annak a valószínűsége, hogy az  $i$ -edik kapszula a  $j$ -edik kapszula irányába továbbítja a kimenetét. Az összegzett bemenet a  $j$ -edik kapszula számára az alábbiak szerint adható meg

$$s = \sum_i c_{ij} \hat{u}_{j|i} \quad (3)$$

Kapszula hálózatok esetén az előrejelzés számszerűsítésére a kimeneti vektorok hosszát alkalmazzák. Ehhez az is szükséges, hogy a kapott vektorok hossza a  $[0,1]$  intervallumba kerüljön. Ehhez az elmélet kidolgozói bevezették az úgynevezett squash függvényt, amit a következő módon számíthatunk

$$v_j = \frac{\|s_j\|^2 s_j}{1 + \|s_j\|^2 \|s_j\|} \quad (4)$$

Az egymásra épülő kapszulák közötti információ áramlását a  $c$  súlyértékekkel lehet szabályozni. Ezeknek a számítása egy iteratív módon működő algoritmus segítségével történik. Ez az úgynevezett dinamikus útválasztás. Az útválasztó algoritmus célja, hogy a magasabb szintű és az alacsonyabb szintű kapszulák közötti információ közvetítés fő irányát meghatározza. A dinamikus útválasztó algoritmus működését az 1. algoritmus foglalja össze.

---

**1. algoritmus:** Dinamikus útválasztó algoritmus [13] (saját szerkesztés)

---

- 1        Dynamic routing( $r, i, j, \hat{u}_{j|i}$ ):
  - 2         $\forall i, j \ b_{ij} = 0$
  - 3         $r$  iteráción keresztül:
  - 4         $\forall i \ c_i = \frac{e^{b_i}}{\sum_j e^{b_j}}$
  - 5         $\forall j \ s_j = \sum_i c_{ij} \hat{u}_{j|i}$
  - 6         $\forall j \ v_j = \frac{\|s_j\|^2}{1 + \|s_j\|^2} \frac{s_j}{\|s_j\|}$
  - 7         $\forall i, j \ b_{ij} = b_{ij} + \hat{u}_{j|i} v_j$
  - 8        ki:  $v_j$
- 

### 3. Javasolt optimalizáló megoldás

A Sabour és társai által bevezetett dinamikus útválasztó algoritmus összességében a következő műveletet hajtja végre iteratív módon

$$v_j = \text{squash} \left( \sum_i (\text{softmax}(b_{ij} + \hat{u}_{j|i} v_j) \hat{u}_{j|i}) \right) \quad (5)$$

Vizsgálataink során arra a megállapításra jutottunk, hogy a  $\hat{u}_{j|i}$  súlyozott bemeneti mátrix jelenléte túl erős az optimalizálás során. Az 5. egyenlet esetén látható, hogy az két helyen is szerepel a  $v_j$  kimeneti értékek realizálása során. Ezért a következő egyszerűsített útválasztó algoritmust vezettük be. [16]

A  $j$ -edik kimeneti  $v_j$  kapszula felépítése legyen a következő

$$v_j = \begin{bmatrix} c_{11} & \cdots & c_{1m} \\ \vdots & \ddots & \vdots \\ c_{n1} & \cdots & c_{nm} \end{bmatrix} \quad (6)$$

ahol  $c_{kl}$  a  $k$ -edik osztályhoz tartozó kapszula  $l$ -edik neuronjának az értéke. Ha  $v_j$  egy köztes kapszula réteg, akkor  $n$  a kimeneti kapszulák száma. Ha  $v_j$  egy kimeneti kapszula réteg, akkor  $n$  a lehetséges objektum kategóriák száma. Legyen

$$\|v_j\| = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \quad (7)$$

ahol  $\forall x \in \{1, 2, \dots, n\}$  esetén

$$v_x = \sqrt{\sum_{y=1}^n c_{xy}} \quad (8)$$

és  $v_x \in [0, 1]$  az  $x$ -edik kapszula során vett valószínűségi érték. Végül

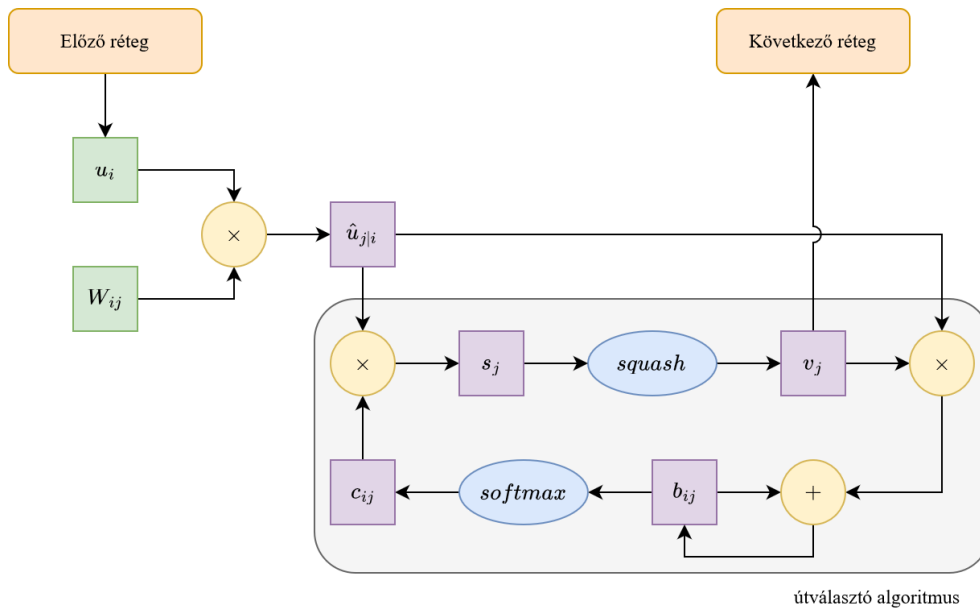
$$b_{ij} = b_{ij} + \|v_j\| \quad (9)$$

ahol  $\forall x \in \{1, 2, \dots, z\}$  esetén  $b_x$  az  $x$ -edik kimeneti kapszulához tartozó súlyérték, melyekből a  $c_{ij} = \text{softmax}(b_{ij})$  útválasztási együttható képződik és  $z$  a bemeneti kapszulák száma.

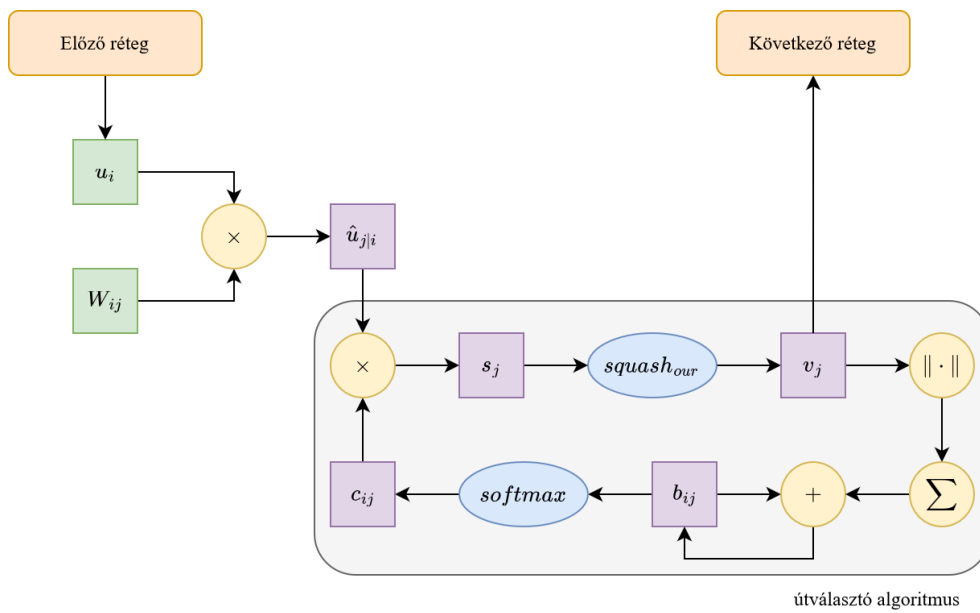
A megvalósításunk során a következő módosított paraméterezhető squash aktivációs függvényt vezettük be.

$$\text{squash}_{our}(s) = \frac{s - e^{-\|s\|_s}}{\|s\| + \varepsilon} \quad (10)$$

ahol  $\varepsilon$  tanuló paraméter, a függvény finomhangolása céljából. A két útválasztó algoritmus blokkvázlatát a 2. ábra mutatja, ahol a két megoldás közötti főbb különbségek jól láthatók. A paraméterezhető squash aktivációs függvényre egy 1-dimenziós esetet, különféle  $\varepsilon$  értékekkel a 3. ábra szemléltet.

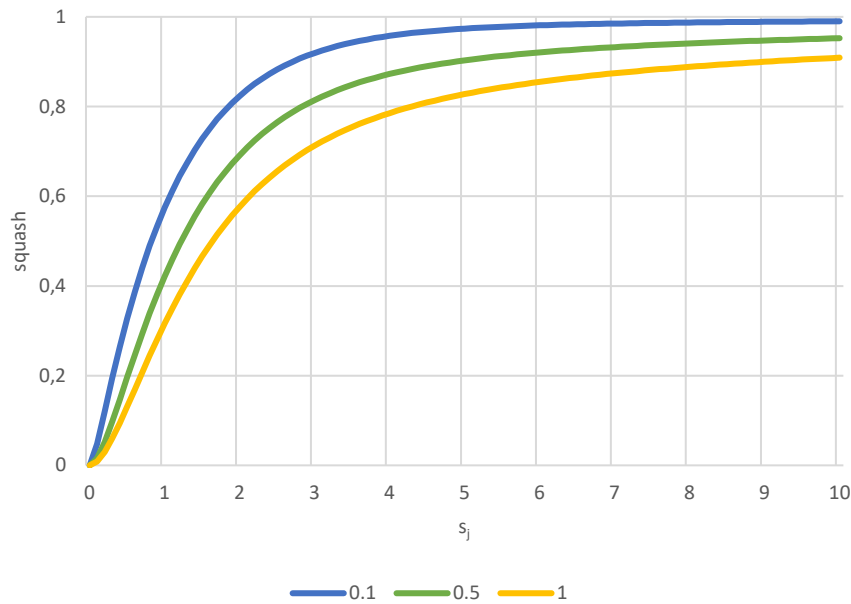


Sabour és társai [13]



Saját megoldás [16]

2. ábra: Útválasztó algoritmusok összehasonlító blokkvázlata (saját szerkesztés)



3. ábra: Javasolt aktivációs függvény, különféle hiperparaméterek esetén (saját szerkesztés)

#### 4. Alkalmazott adatkészlet

A megoldások teszteléséhez és összehasonlításához a Kaggle Facial Keypoints Detection Dataset [17] nevű, ingyenesen hozzáférhető adatkészlet került felhasználásra. Ez az adatkészlet különféle emberi arcokról tartalmaz olyan képeket, ahol szemből nézve, csak az arc látható, változó háttér előtt. Az adatkészlet összesen 7049 darab mintát tartalmaz, melyek 48 képpont szélességgel és magassággal rendelkeznek, valamint 1 színcsatornásak, azaz szürkeárnyaltos képek szereplnek a készletben. Minden bemeneti kép esetén értelmezett 15 darab úgynevezett kulcspont. Ezek a kulcspontok az arc meghatározott területeinek a kamerakép síkján vett pozícióját jelentik. Ezek a kulcspontok a következők: Jobb szemöldök külső sarokpont, jobb szemöldök belső sarokpont, bal szemöldök külső sarokpont, bal szemöldök belső sarokpont, jobb szem külső sarokpont, jobb szem középpont, jobb szem belső sarokpont, bal szem külső sarokpont, bal szem középpont, bal szem belső sarokpont, orrhegy, száj jobb sarokpont, száj bal sarokpont, felsőajak középpont, alsóajak középpont. Ilyen képekre mutat néhány példát a 4. ábra, ahol piros színnel láthatóak az egyes kulcspontok.



**4. ábra:** Mintaképek a kialakított adatkészletből (saját szerkesztés)

Sajnálatos módon a Kaggle adatkészlete hiányos, nem tartalmazza mind a 7049 bemeneti kép esetén az összes 15 darab kulcspontot. A legtöbb kép esetén ez részben hiányos. Ezért az adatkészletből kialakításra került egy olyan részhalmaz, amely csak olyan képeket tartalmaz, ahol mindegyik realizált kulcspont megtalálható. Ez a szűkített adatkészlet összesen 2250 darab mintát tartalmaz. Ezek esetén mind a 15 kulcspont megtalálható. A hálózat tanításához és validálásához ez az adatkészlet került felhasználásra.

## 5. Adatkészlet támadási módszerek

A hálózatok hatékonyságának összehasonlításához adatkészlet támadási módszerek kerültek alkalmazásra a kutatás során. Ezeknek a módszereknek a lényege, hogy valamilyen meghatározott módon zajforrást adnak a bemeneti képhez. A zaj hatására a hálózat téves detektálást hajt végre. Vannak módszerek, amik drasztikus mértékben rontják a kép felismerhetőségének minőségét. Itt a bemenet a zajterhelés következtében már emberi résztvevő számára sem lesz beazonosítható. Azonban a hálózat, a zajterhelés ellenére mégis megpróbál valamilyen kimenettel szolgálni. Ennek a pontosságának a vizsgálata a cél. Más technikák jellegzetessége, hogy a zajterhelés látszólag minimális. Azaz emberi résztvevő számára nem zavaró, a felismerést, kulcspontok bejelölését gond nélkül el tudja végezni. Azonban a hálózatot ennek ellenére mégis félrevezeti a zaj, mert az szándékosan olyan területeken és olyan mértékben jelentkezik, mely a hálózat belső állapotai alapján jól el tudja téríteni a rendszer pontosságát. Az alábbi 7 támadási módszer került felhasználásra a kutatás során.

A Fast Gradient Signed Method (FGSM) nevű módszert Goodfellow és társai vezették be. [18] Ez az egyik legelső és egyben legnépszerűbb módszer mind közül. A megoldás a támadni kívánt neurális hálózat gradienseit használja fel a zajterheléshez. A gradiens értékeket a hálózat veszteségéből határozza meg. Ennek az értékét adja hozzá a képhez, annak érdekében, hogy maximalizálja a hálózat veszteségét, a következők szerint:

$$x_{adv} = x + \epsilon \times \text{sign}(\nabla_x J(\theta, x, y)) \quad (11)$$

ahol  $x_{adv}$  a zajterhelt bemenet,  $x$  az eredeti bemenet,  $y$  az elvárt kimenet,  $\epsilon$  finomhanglást célzó hiperparaméter,  $\theta$  a hálózat paraméterei és  $J(\cdot)$  a veszteségfüggvény.

A Fast adversarial training using FGSM (FFGSM) módszert Wong, Rice és társaik vezették be. [19] Ez a módszer az FGSM megoldásra épül. A fő különbség, hogy az FFGSM véletlenszerű inicializálást használ, ahol egyenletes eloszlás mellett generálódik a zaj, a következők szerint:

$$x' = x + \text{Uniform}(-\epsilon, \epsilon) \quad (12)$$

ahol  $x$  a bemeneti kép,  $x'$  a zajterhelt kép és  $\epsilon$  hiperparaméter.

A Projected Gradient Descent (PGD) módszert Mądry és társai vezették be. [20] Ez a megoldás is hasonló az FGSM támadási módszerhez. Viszont ennél a megközelítésnél iteratív módon többször hajtják végre a támadást az alábbiak folyamán:

$$x'_{t+1} = \prod_{\mathcal{B}(x, \epsilon)} \left\{ x'_t + \alpha \text{sign} \left( \nabla_{x'_t} J(f_\theta(x), f_\theta(x'_t)) \right) \right\} \quad (13)$$

ahol  $\prod \mathcal{B}(x, \epsilon)$  egy projekció a  $\mathcal{N}(0^n, I^n)$  normál eloszlás szerint.

A PGD támadásnak két további variációja is felhasználásra került. Az első elgondolás szerint  $L_2$  veszteségfüggvény alkalmazásával került felhasználásra (PGDL2). [20] Míg a második esetben a TRADES' Projected Gradient Descent (TPGD) került alkalmazásra, mely Zhang és

társai munkája. [21] Ebben az esetben a Kullback-Leibler divergencia [22] került alkalmazásra veszteségfüggvényként az alábbi módon:

$$x'_0 = x + 0.001 \times \mathcal{N}(0^n, I^n) \quad (14)$$

$$x'_{t+1} = \prod_{B(x,\varepsilon)} \left\{ x'_t + \alpha \text{sign} \left( \nabla_{x'_t} \ell_{KL}(f_\theta(x), f_\theta(x'_t)) \right) \right\} \quad (15)$$

ahol  $\ell_{KL}(\cdot)$  a veszteségfüggvény.

A következő támadás egy egyszerű véletlen gauss-zaj terheléssel történt (GN), ahol a zaj a bemeneti képhez került hozzáadásra. Ebben a megközelítésben a hálózat belső tulajdonságai nem kerültek felhasználásra, mindig attól függetlenül, véletlenszerűleg generálódik a zajforrás a következő módon:

$$x' = x + \sigma \mathcal{N}(0^n, I^n) \quad (16)$$

ahol most is  $x$  a bemenet,  $x'$  a zajterhelt kimenet,  $\mathcal{N}(0^n, I^n)$  a normál eloszlás szerinti véletlenzaj és  $\sigma \in (0, 1)$  finomhangolási hiperparaméter.

A Basic Iterative Method (BIM) Kurakin és társai vezették be. [23] Szintén az FGSM módszerre alapozva alkották meg az algoritmusukat. Ennél a módszernél egymás után többször, kis lépésszámokkal alkalmazzák a zajterhelést. Ezen felül egy úgynevezett *Clip* függvénnyel vágják a képpontok értékeit, így azok az eredeti kép  $\varepsilon$ -szomszédságában maradnak. Módszerük a következő:

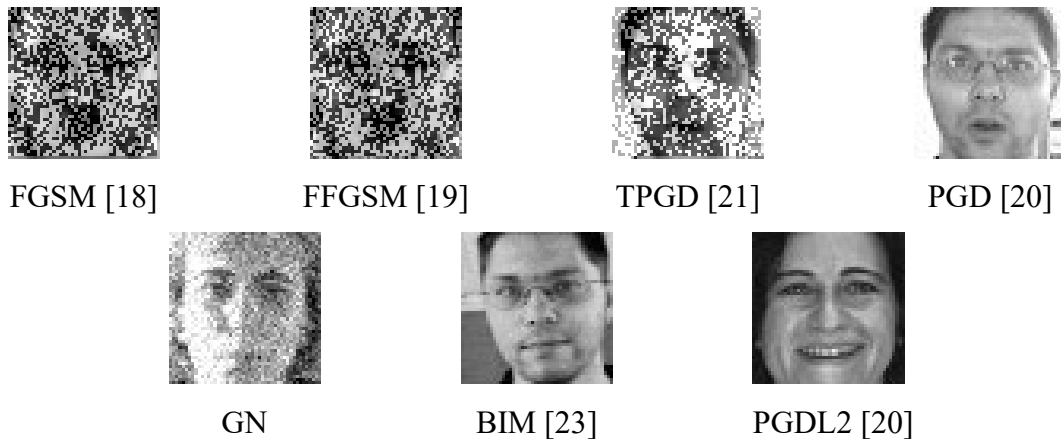
$$x'_{n+1} = \text{Clip}_{x,\varepsilon} \left\{ x'_n + \alpha \text{sign}(\nabla_x J(x'_n, y_{true})) \right\} \quad (17)$$

ahol  $x'_0 = 0$ ,  $x$  a bemeneti kép, továbbá

$$\text{Clip}_{x,\epsilon}\{x'\}(u, v, z) = \min\{255, x(u, v, z) + \epsilon, \max\{0, x(x, y, z) - \epsilon, x'(u, v, z)\}\} \quad (18)$$

ahol  $x(u, v, z)$  az  $x$  bemeneti kép  $z$  csatornájának értéke az  $(u, v)$  pozíción.

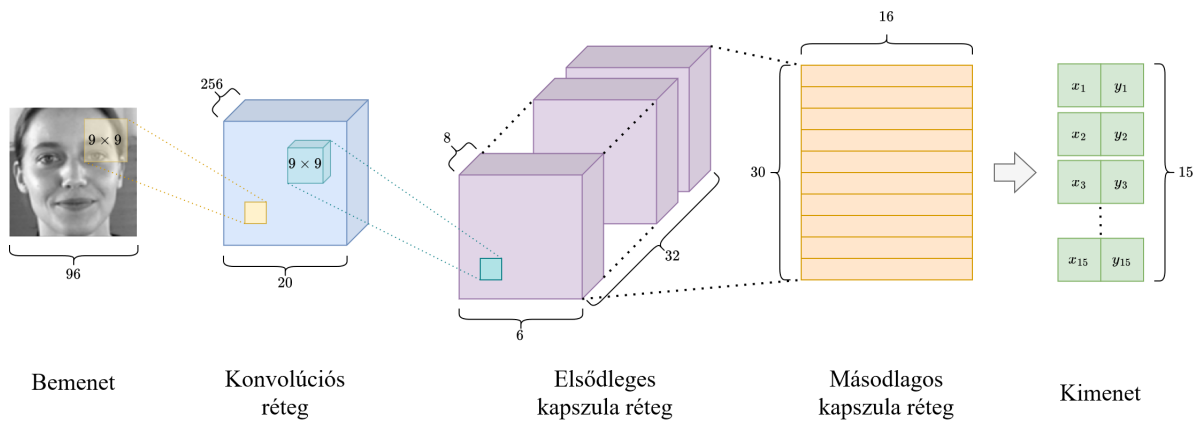
Az 5. ábra az itt tárgyalt támadási módszerek esetén mutat egy-egy példát néhány, az adatkészletből vett minta esetén. Jól látható itt is, hogy bizonyos esetekben a zajterhelés meglehetősen súlyos, míg más esetekben sokkal enyhébb, még felismerhető az emberi arc a képen. Illetve vannak olyan módszerek, ahol szinte észre sem venni a zajterhelést a képen.



5. ábra: Zajterhelt mintaképek a különféle támadási módszerek esetén (saját szerkesztés)

## 6. Tanítás

A kialakított hálózati architektúra a Sabour és társai [13] által javasolt modellt követi. Két módon alakítottuk ki a hálózatot. Az első esetben teljes mértékben az eredeti szerzők megoldását alkalmaztuk. Itt az egyetlen változás csupán a bemenet mérete, ami az esetünkben  $48 \times 48$  képpontot jelent. A második esetben a változás csupán annyit jelent, hogy a dinamikus útválasztó algoritmus és a Sabour és társai [13] által javasolt squash aktivációs függvény helyett a saját megvalósításainkat alkalmaztuk, melyeket a 3. fejezetben részleteztünk. A hálózati architektúrát és főbb egységeit, paramétereit a 6. ábra mutatja be.



**6. ábra:** Tanítás során alkalmazott hálózati architektúra (saját szerkesztés)

A kialakított 2 hálózat 8 különböző módon került tanításra. Az első esetben zajterhelés nélkül, majd az 5. fejezetben részletezett zajterhelési technikák mellett. A tanítás minden más körülményei azonosak voltak minden esetben. A tanítás folyamata során az Adam [24] optimalizáló algoritmus került felhasználásra. A kezdeti tanulási ráta értéke 0,001, ami az előzetes tapasztalataink alapján tűnt optimális kezdőértéknek. A tanulási iterációkon keresztül a tanulási ráta értéke a következő módon alakult:

$$lr_{i+1} = lr_i \times 0,96^{\frac{1}{2000}} \quad (19)$$

ahol  $lr_i$  az  $i$ -edik iterációban vett tanulási ráta értéke és  $lr_0 = 0,001$ . A tanítás 50 iteráción keresztül zajlott minden esetben. Tapasztalatunk azt mutatja, hogy 50 epoch után a tanulási görbék már nem javultak számottevően, a hálózatok közötti eltérés szignifikánsan már nem változott. Mind a tanítási, mind a tesztelési fázisban a kötegméret 16 darab volt. Az L1 veszteség került alkalmazásra, úgy, hogy

$$\ell(x, y) = \{l_1, l_2, \dots, l_n\}^T \quad (20)$$

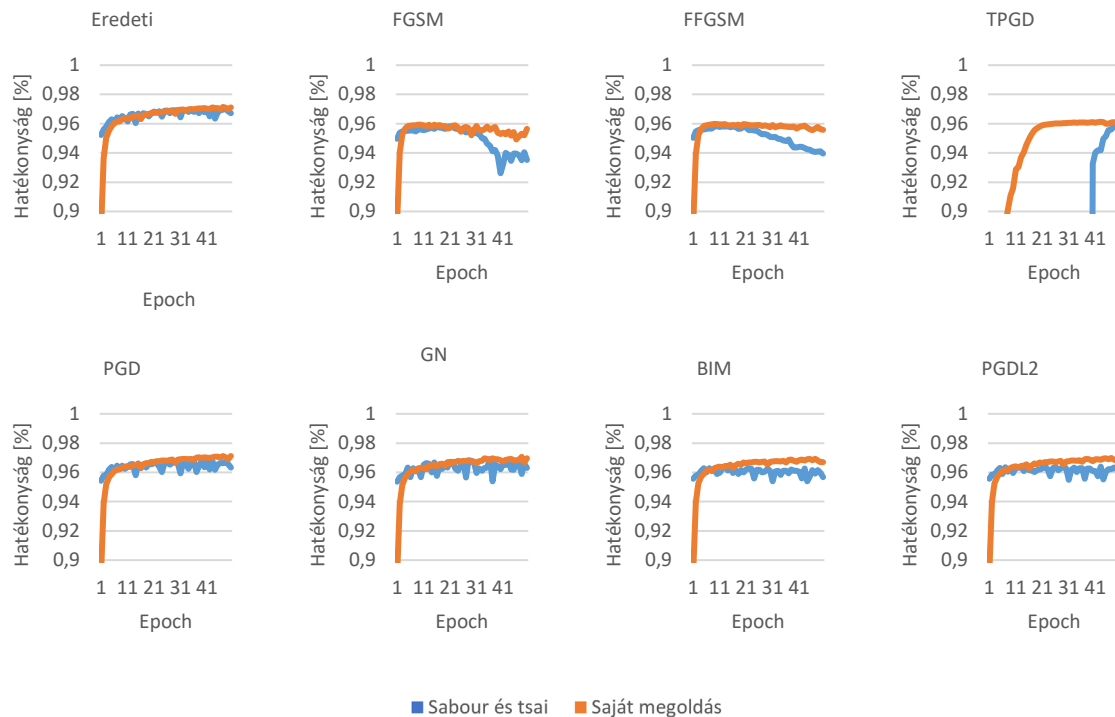
ahol  $\ell$  az adott köteg vesztesége,  $x$  a prediktált kimenet,  $y$  az elvárt és helyes kimenet, valamint  $n$  a köteg mérete. Legyen

$$l_i = \begin{cases} \frac{0,5(x_i - y_i)^2}{\beta}, & \text{ha } |x_i - y_i| < \beta \\ |x_i - y_i| - 0,5\beta, & \text{különben} \end{cases} \quad (21)$$

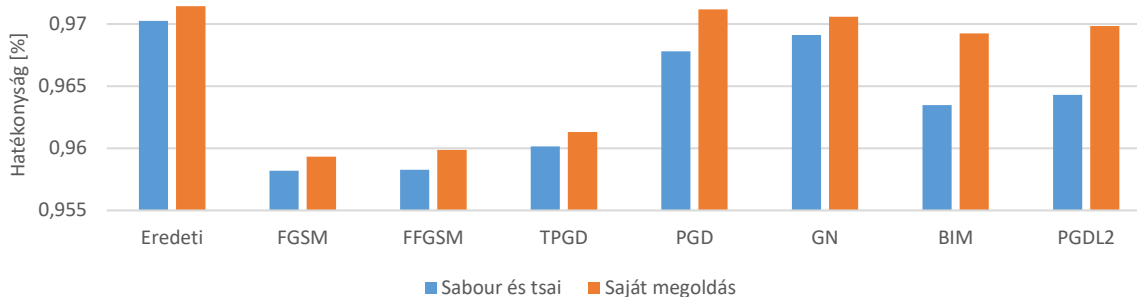
ahol  $l_i$  a vizsgált kötegben az  $i$ -edik bemenetnek a vesztesége.

## 7. Eredmények, kiértékelés

A 6. fejezetben részletezett körülmények esetén végrehajtott tanítási folyamatok eredményét mutatják a 7. és 8. ábrák. A 7. ábrán az 50 iteráción keresztül mért hatékonyság figyelhető meg. Látható, hogy támadás nélkül a két megoldás hatékonysága nagyon hasonló. A szembetűnő különbség inkább az, hogy az általunk javasolt módszer esetén kapott tanulási görbe sokkal zajmentesebb, egyenletesebb, magabiztosabb tanulási folyamatot mutat. Végeredményben pedig valamivel jobb eredménnyel is szolgál, mint a dinamikus útválasztó algoritmus. A 7. támadási módszer esetén a különbség azonban sokkal nagyobb. Mindegyik esetben azt figyelhetjük meg, hogy a saját módszerünk sokkal megbízhatóbb, jobb tanulási görbével rendelkezik.



**7. ábra:** Tanulási hatékonysági görbék zajterhelés nélkül és az alkalmazott zajterhelések esetén (saját szerkesztés)



**8. ábra:** Legjobb eredmények (saját szerkesztés)

A 8. ábra a zajterhelés nélküli, illetve a 7 zajterhelt esetben mutatja a legjobb hatékonyságot, amit sikerült elérnünk kutatásunk során. Itt is jól látszik, hogy az általunk javasolt optimalizáló eljárás ennek a feladatnak a tekintetében is minden esetben jobban teljesített, mint a dinamikus útválasztó algoritmus. Megfigyelhető, hogy vannak olyan esetek, amikor a két módszer hatékonysága megközelítőleg ugyanolyan mértékben romlik. Ilyen az FGSM, FFGSM és a TPGD támadások. Tehát ezek hasonlóan le tudják rontani mindkét megoldás eredményét. Ugyanakkor a saját megoldásunk mindhárom esetben jobban teljesít. Más megoldásoknál az is észrevehető, hogy a dinamikus útválasztó algoritmus nagyobb mértékben veszít a teljesítményéből, mint a saját megoldásunk. Ilyen a BIM és a PGDL2 támadási módszerek.

## 8. Összegzés

Kutatásunk során a kapszula hálózatok elméletének gyakorlati alkalmazhatóságával foglalkoztunk. Bemutattuk a kapszula hálózatokhoz kapcsolódó alapvető elméleti hátteret, kitérve annak fontosabb mozzanataira. Rávilágítottunk a klasszikus neurális hálózatok, kiváltképp a konvolúciós neurális hálózatok, valamint a kapszula hálózatok közötti főbb különbségekre, alapvető eltérésekre. Felhívtuk a figyelmet a konvolúciós neurális hálózatok megtéveszthetőségének problémájára, melyre részben megoldást kínálnak a kapszula alapú megoldások. Bevezettük az általunk javasolt új, egyszerűsített útválasztó algoritmus elméletét, melyet szemben állítottunk a kapszula hálózatok elméletének kidolgozói által ajánlott dinamikus útválasztó algoritmussal. Egy speciális részfeladat esetén, ember archoz tartozó kulcsfontosságú detektálása során megvizsgáltuk, hogy a dinamikus útválasztó algoritmus és az általunk javasolt útválasztó algoritmus hatékonysága között milyen különbségek fedezhetőek

fel. Tettük ezt különféle korszerű zajterhelési technikák alkalmazása mellett is. Ezzel még inkább feltárva a két megoldás között kardinális különbségeket. A tanítás eredményeit részletesen bemutatva rávilágítottunk a két optimalizáló algoritmus közötti hatékonysági különbségekre kulcspont detektálási feladat esetén. Az eredményeink nyomán egyértelműen látható, hogy a vizsgált esetek tekintetében az általunk bevezetett útválasztó algoritmus mindegyik esetben jobb hatásfokot volt képes elérni, mint a dinamikus útválasztó algoritmus. Eredményeink és a kutatás során szerzett tapasztalataink mentén szeretnénk tovább folytatni a kutatásainkat, ahol a kapszula hálózatok elméletének területén még inkább szeretnénk elmélyülni. Ilyen tekintetben további vizsgálatoknak szeretnénk alá vetni az optimalizáló algoritmusunkat, más jellegű feladatok esetén is szeretnénk megvizsgálni a hatékonyságát. További megoldásokkal szeretnénk összemérni az eredményét. A tapasztalatok nyomán pedig további tervezések és fejlesztések során szeretnénk tovább növelni a jelenlegi megoldásunk hatékonyságát. Mindezzel hozzájárulva a kapszula hálózatok elméletének gazdagításához és a mesterséges intelligencia gyakorlati felhasználásának magasabb megbízhatósági fokához.

## 9. Köszönetnyilvánítás

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

## Irodalomjegyzék

- [1.] X. Chen, C. Liang, D. Huang, E. Real, K. Wang, Y. Liu, H. Pham, X. Dong, T. Luong, C. Hsieh, Y. Lu & Q. V. Le, “Symbolic Discovery of Optimization Algorithms.” arXiv, 2023 május 8. doi: 10.48550/arXiv.2302.06675.
- [2.] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit & N. Houlsby, “An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale.” arXiv, 2021. június 3., doi: 10.48550/arXiv.2010.11929.
- [3.] Z. Zong, G. Song & Y. Liu, “DETRs with Collaborative Hybrid Assignments Training.” arXiv, 2023. augusztus 10., doi: 10.48550/arXiv.2211.12860.

- [4.] G. Ghiasi, Y. Cui, A. Srinivas, R. Qian, T.-Y. Lin, E. D. Cubuk, Q. V. Le & B. Zoph, “Simple Copy-Paste is a Strong Data Augmentation Method for Instance Segmentation.” arXiv, 2021. június 23., doi: 10.48550/arXiv.2012.07177.
- [5.] P. Wang, S. Wang, J. Lin, S. Bai, X. Zhou, J. Zhou, X. Wang & C. Zhou, “ONE-PEACE: Exploring One General Representation Model Toward Unlimited Modalities.” arXiv, 2023. május 18., doi: 10.48550/arXiv.2305.11172.
- [6.] W. Wang, J. Dai, Z. Chen, Z. Huang, Z. Li, X. Zhu, X. Hu, T. Lu, L. Lu, H. Li, X. Wang & Y. Qiao, “InternImage: Exploring Large-Scale Vision Foundation Models with Deformable Convolutions.” arXiv, 2023. április 17., doi: 10.48550/arXiv.2211.05778.
- [7.] S. You, N. You & M. Pan, “PI-REC: Progressive Image Reconstruction Network With Edge and Color Domain.” arXiv, 2019. március 25., doi: 10.48550/arXiv.1903.10146.
- [8.] T. Kim, G. Song, S. Lee, S. Kim, Y. Seo, S. Lee, S. H. Kim, H. Lee & K. Bae, “L-Verse: Bidirectional Generation Between Image and Text.” arXiv, 2022. április 6., doi: 10.48550/arXiv.2111.11133.
- [9.] A. Ruospo, A. Bosio, A. Ianne & E. Sanchez, “Evaluating Convolutional Neural Networks Reliability depending on their Data Representation,” in 2020 23rd Euromicro Conference on Digital System Design (DSD), 2020, pp. 672–679. doi: 10.1109/DSD51259.2020.00109.
- [10.] A. Nguyen, J. Yosinski & J. Clune, “Deep neural networks are easily fooled: High confidence predictions for unrecognizable images,” in 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 427–436. doi: 10.1109/CVPR.2015.7298640.
- [11.] Q. Wang, J. Jiang, J. Mo, Q. Kong, Y. Li & S. Feng, “A Survey of Adversarial Attacks on Deep Neural Network in Computer Vision,” in 2022 IEEE 8th International Conference on Computer and Communications (ICCC), 2022, pp. 1879–1886. doi: 10.1109/ICCC56324.2022.10065959.
- [12.] H. Liang, E. He, Y. Zhao, Z. Jia & H. Li, “Adversarial Attack and Defense: A Survey,” Electronics, vol. 11, no. 8, p. 1283, 2022, doi: 10.3390/electronics11081283.

- [13.] S. Sabour, N. Frosst & G. E. Hinton, “Dynamic routing between capsules,” in Proceedings of the 31st International Conference on Neural Information Processing Systems, in NIPS’17. Red Hook, NY, USA: Curran Associates Inc., 2017, pp. 3859–3869.
- [14.] G. E. Hinton, A. Krizhevsky & S. D. Wang, “Transforming Auto-Encoders,” in Artificial Neural Networks and Machine Learning – ICANN 2011, T. Honkela, W. Duch, M. Girolami & S. Kaski, Eds., in Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2011, pp. 44–51. doi: 10.1007/978-3-642-21735-7\_6.
- [15.] S. Sabour & G. E. Hinton, “Matrix capsules with EM routing,” presented at the Sixth International Conference on Learning Representations, Vancouver, Canada, 2018.
- [16.] J. Hollósi, Á. Ballagi & C. R. Pozna, “Simplified Routing Mechanism for Capsule Networks,” Algorithms, vol. 16, no. 7, p. 336, 2023, doi: 10.3390/a16070336.
- [17.] J. Petterson & W. Cukierski, “Facial Keypoints Detection.,” Kaggle, 2013. Elérés: 2023.04.20. [Online] <https://kaggle.com/competitions/facial-keypoints-detection>
- [18.] I. J. Goodfellow, J. Shlens & C. Szegedy, “Explaining and Harnessing Adversarial Examples,” in 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, 2015. május 7-9., Conference Track Proceedings, Y. Bengio and Y. LeCun, Eds., 2015.
- [19.] E. Wong, L. Rice & J. Z. Kolter, “Fast is better than free: Revisiting adversarial training.” arXiv, 2020. január 12., doi: 10.48550/arXiv.2001.03994.
- [20.] A. Madry, A. Makelov, L. Schmidt, D. Tsipras & A. Vladu, “Towards Deep Learning Models Resistant to Adversarial Attacks.” arXiv, 2019. szeptember 4., doi: 10.48550/arXiv.1706.06083.
- [21.] H. Zhang, Y. Yu, J. Jiao, E. P. Xing, L. E. Ghaoui & M. I. Jordan, “Theoretically Principled Trade-off between Robustness and Accuracy.” arXiv, 2019. június 24., doi: 10.48550/arXiv.1901.08573.
- [22.] J. M. Joyce, “Kullback-Leibler Divergence,” in International Encyclopedia of Statistical Science, M. Lovric, Ed., Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 720–722. doi: 10.1007/978-3-642-04898-2\_327.

[23.] A. Kurakin, I. Goodfellow & S. Bengio, “Adversarial examples in the physical world.” arXiv, 2017. február 10., doi: 10.48550/arXiv.1607.02533.

[24.] D. P. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization.” arXiv, 2017. január 29., doi: 10.48550/arXiv.1412.6980

# Kamera deprojekció felhasználása autonóm környezetérzékelés területén

## Utilization of camera deprojection in the field of autonomous environmental perception

**Markó Norbert<sup>a</sup>, Krecht Rudolf<sup>b</sup>**

<sup>a</sup>Széchenyi István Egyetem, Járműipari Kutatóközpont

marko.norbert@ga.sze.hu

<sup>b</sup>Széchenyi István Egyetem, Járműipari Kutatóközpont

krecht.rudolf@ga.sze.hu

### Absztrakt

---

*A gépi látáson alapuló algoritmusok képesek a minket körülvevő világot egy bizonyos fókig értelmezni és erre az értelmezésre alapozva az adott célnak megfelelően szétválasztani, kategorizálni és osztályozni. Az autonóm területen gyakori feladat az útburkolat felismerése, és a terep többi részétől való elválasztása. Az így felismert útfelület irányításhoz való felhasználásához szükség van az útfelület egyes pixeleinek távolságmeghatározására. Jelen cikkben a távolságmeghatározást geometriai problémaként leíró deprojekciós algoritmust szeretnénk ismertetni a feladat megoldásához. Az algoritmus ismeretése után bemutatunk pár felhasználási területet, illetve a felhasználás eredményét egy általunk finomhangolt útburkolat felismerő algoritmuson.*

**Kulcsszavak:** deprojekció, gépi látás, mesterséges intelligencia

### Abstract

---

*Computer vision-based algorithms are able to interpret the world around us to a certain extent and, based on this interpretation, to separate, categorise and classify it according to the given purpose. A common task in the autonomous domain is to recognise the pavement and separate it from the rest of the terrain. In order to use the road surface thus recognized for guidance, it is necessary to determine the distance of each pixel of the road surface. In this paper, we present a deprojection algorithm describing distance determination as a geometric problem to solve this task. After introducing the algorithm, we present some applications and the results of the application on a road pavement recognition algorithm that we have fine-tuned.*

**Keywords:** deprojection, computer vision, artificial intelligence

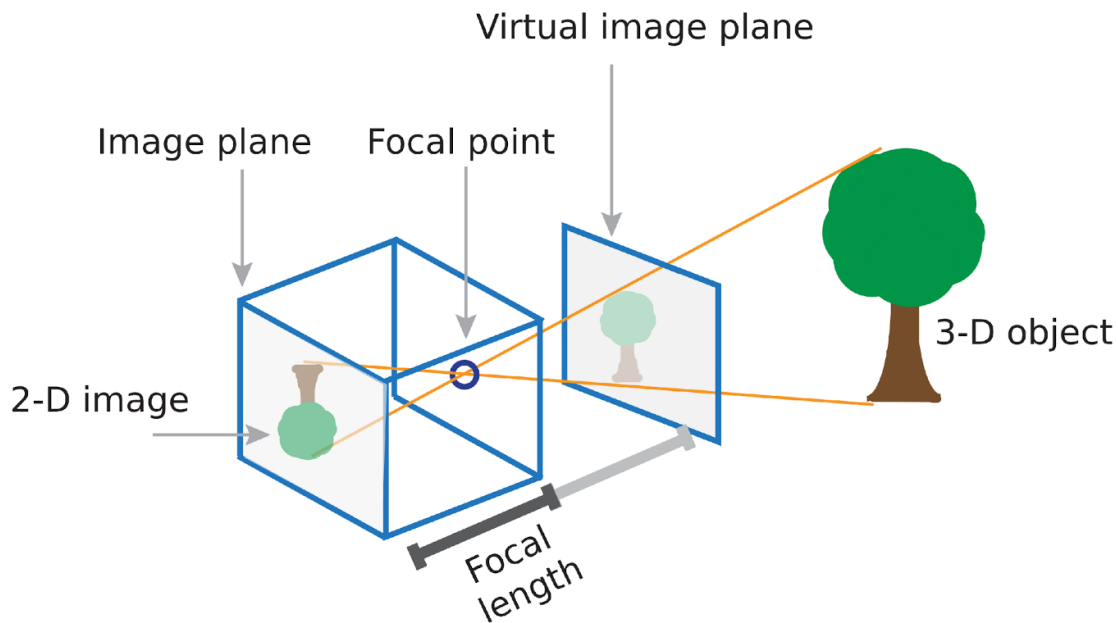
## 1. Bevezetés

A kamerák az autonóm rendszerek és a gépi látás legkritikusabb építőelemei közé tartoznak. Minden kamera a maga egyedi módon fogja fel a térbeli információkat és alakítja át azokat képi adatokká, amelyeket egy számítógép vagy más digitális eszköz könnyen feldolgozhat. Ez az egyedi látásmód a kameramatrixon keresztül ismerhető meg, mely a háromdimenziós világot képezi le két dimenzióra. Ennek a transzformációs folyamatnak a megértése nem csak a képalkotás szempontjából van központi jelentősége, hanem azért is, mert az így keletkező adatokat sok esetben vissza kell alakítanunk a valós, háromdimenziós világ paramétereivé, amit "deprojekciónak" nevezünk. A deprojekció ebben a kontextusban egy kritikus fontosságú feladat, különösen az olyan alkalmazásokban, ahol a gépi látás rendszereknek nem csak érzékelniük kell a környezetet, hanem interakcióba is kell lépniük vele (pl. önvezetés). Ilyen esetekben a képi információknak a valós világ paramétereivé való visszatranszformálása nem egyszerűen opcionális, hanem alapvető követelmény. Jelen cikkünkben részletesen vizsgáljuk a deprojekció matematikai alapjait, valamint példákat hozunk gyakorlati problémák megoldására. A cikk célja egy átfogó képet adni erről a rendkívül összetett, de nélkülözhetetlen folyamatról, amely az autonóm rendszerek és a gépi látás egy alapvető részét képezi.

## 2. Problémafelvetés

A kamerák kritikus szerepet játszanak az autonóm környezetérzékelésben. Ahhoz, hogy a deprojekció folyamatát megértsük, előbb szükséges néhány fogalmat tisztázni a kameraszennorral kapcsolatban. Először is nézzük meg, hogy mi is a kamera pontosan.

A kamera egy olyan eszköz, amely a 3D világot 2D képsíkra képezi le. Ennek a transzformációnak a matematikai leírásához a valós világ kameráját különböző kameramodellekkel közelíthetjük. Ha kiválasztottuk a kameramodelt, akkor ennek paramétereit kalibrációs algoritmusokkal becsülhetjük meg. Ezek a paraméterek egy transzformációs mátrixot alkotnak, amely a 3D világkoordinátákat 2D képkoordinátákra képezi le. A kameraszennorok de facto modellje a lyukkamera modell (pinhole camera model). Ez egy nagyon jó közelítés a lencsés kamerákhoz annak ellenére, hogy ez a modell nem veszi figyelembe a lencse torzító hatásait.



1. ábra: Objektum leképezésének menete kamera segítségével [3].

A kamera paramétereit, amelyeket a fentebb említett kalibrációs algoritmusok segítségével becsülhetünk meg, két fő kategóriába sorolhatók: a külső és a belső paraméterek. A külső paraméterek a kamera helyzetét és orientációját írják le a 3D-s világban. Ezek változnak, ha a kamera pozíciója, például egy mozgó járműben, megváltozik. Ezen paraméterek segítségével alakíthatók át a világkoordináták kamerakoordinátákká. A belső paraméterek a kamera belső karakterisztikáit, például a fókusz távolságot és az optikai középpontot írják le. Ezek a kamerakoordinátákat képp koordinátákká, azaz pixelkoordinátákká transzformálják. E két paramétercsoport segítségével konstruálható egy kameramátrix, amely a világkoordináták és a pixelkoordináták közötti teljes leképezési folyamatot írja le [1].

$$P = \overbrace{K}^{\text{Intrinsic Matrix}} \times \overbrace{[R | \mathbf{t}]}^{\text{Extrinsic Matrix}}$$

(1)

$$= \overbrace{\begin{pmatrix} 1 & 0 & x_0 \\ 0 & 1 & y_0 \\ 0 & 0 & 1 \end{pmatrix}}^{\text{Intrinsic Matrix}} \times \overbrace{\begin{pmatrix} f_x & 0 & 0 \\ 0 & f_y & 0 \\ 0 & 0 & 1 \end{pmatrix}}^{\text{Intrinsic Matrix}} \times \overbrace{\begin{pmatrix} 1 & s/f_x & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}^{\text{Intrinsic Matrix}} \times \overbrace{\begin{pmatrix} I & \mathbf{t} \end{pmatrix}}^{\text{Extrinsic Matrix}} \times \overbrace{\begin{pmatrix} R & 0 \\ 0 & 1 \end{pmatrix}}^{\text{Extrinsic Matrix}}$$

2D Translation
2D Scaling
2D Shear
3D Translation
3D Rotation

(2)

2. ábra: A kameramátrix képlete és komponensei. (saját szerkesztés)

### 3. A deprojekció menete

A deprojekció (vagyis inverse perspective mapping) célja az útfelület pontjainak a távolságának kiszámítása, pontosabban az útfelületről a kamera szenzorba visszaverődött fénysugarak távolságának és irányának meghatározása. A számítás alapjait a következőkben ismertetett feltevésekre építjük fel. Az útfelület sík (nincsenek emelkedések vagy süllyedések), ismerjük a kamera pozícióját és orientációját az útfelülethez viszonyítva, vagyis a transzformációs mátrixot, illetve a kamera matematikai modelljének a lyukkamera modellt választjuk.

A kameramodell alapján alapesetben fel tudjuk írni az említett fénysugár irányát, azonban a fénysugáron kiválasztott pont távolságát nem. Ennek a megoldása, hogy geometriai problémaként írjuk fel a távolságmeghatározást az előbb ismertetett feltevéseket figyelembe véve. A matematikai cél az  $(X_c, Y_c, Z_c)^T$  meghatározása minden egyes  $(u, v)$  pixel esetén, amely az útfelülethez tartozik. Annak érdekében, hogy az útfelülethez tartozó pixeleket előszűrjük a képen, szemantikus szegmentációt használtunk.

Az alapösszefüggés alább látható (1), mely alapján felírjuk a deprojekciót geometriai problémaként [5]. Az egyenlet értelmezése, hogy egy tetszőleges 3D pontot a kamera belső paramétermátrixával megszorozva megkapjuk a pixelkoordinátákat  $(u, v)$ .

$$\lambda \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = K \begin{pmatrix} X_c \\ Y_c \\ Z_c \end{pmatrix}$$

(1)

A deprojekcióhoz az egyenlet inverzére van szükség, ahol a pixelből lesz 3D koordináta. Ahhoz, hogy ez megvalósuljon, mind a két oldalt be kell szoroznunk  $\mathbf{K}^{-1}$  mátrix-szal, melyet követően a (2) egyenletet kapjuk.

$$\begin{pmatrix} X_c \\ Y_c \\ Z_c \end{pmatrix} = \lambda \mathbf{K}^{-1} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix}$$

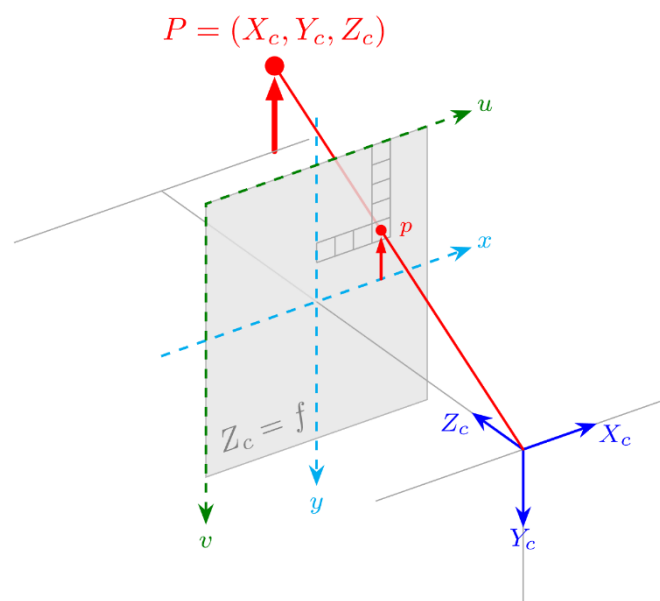
(2)

Az előbbi egyenlettel felírtuk a geometriai probléma reprezentációját, amit meg szeretnénk oldani. A problémát a  $\lambda$  ismeretlen értéke okozza. Azonban annyit már felírhatunk, hogy a 3D koordinátánk valahol az  $(u, v)$  pixelhez tartozó fénysugáron (vagyis egy egyenesen) helyezkedik el. Az egyenes egyenlete az ismeretlen függvényében a következő:

$$r(\lambda) = \lambda K^{-1} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix}, \lambda \in R_{>0}$$

(3)

Az egyenlet alapján a kérdés, amit felteszünk magunknak, hogy mely  $\lambda$  érték adja vissza a képet meghatározó 3D pontot. A lambda érték az egyenletben amiatt van jelen, hogy homogén koordinátákkal számolunk, melyeket az euklideszi koordinátákból kapunk meg, még hozzá egy 1-es érték toldásával a vektor végére. A homogén koordináták definíciójához tartozik, hogy bármely nullától különböző  $\lambda$  érték esetén ekvivalens vektorokról beszélünk. Amennyiben egy tetszőleges homogén koordináta harmadik koordinátájával leosztjuk mindhárom koordinátát, a harmadik komponens 1 lesz, ezt hívjuk kánon alaknak. Az alábbi képen a piros vonal a különböző lambda értékek koordinátáinak összessége, melyek egy egyenest képeznek. Ebből is látható, hogy bármely lambda érték esetén rajta leszünk az egyenesen, azonban csak egy lambda adja vissza a 3D koordinátánkat. A változó lambda kamera esetén pedig a fókusztávolságot jelenti mely a kamera lencse középpontja (koordinátarendszer origója) és virtuális film sík (virtual plane) közti távolság.



**3. ábra:** A kamera koordináta rendszer, a virtuális film sík és a számított fénysugár (piros) [2].

Az egyenes egyenletének megoldásához első lépésben felhasználjuk az első feltevésünket, mely szerint az útfelület egyenletes, vagyis egy síkot képez, így felírhatunk egy sík egyenletet. Az egyenlet felírásához szükség van az útfelület normálvektorára, mely sík útfelület esetén  $n = (0,1,0)^T$  lesz, illetve egy tetszőleges  $r_0$  pontra, mely a síkon található. Amennyiben az egyenletet 0-val egyenlővé tesszük, matematikai előismereteink alapján tudható, hogy az  $r$  koordináta a megoldást követően az egyenesen fog elhelyezkedni.

$$n^T(r - r_0) = 0 \quad (4)$$

A megoldás következő lépése  $n = (0,1,0)^T$  transzformációja, mely az út koordinátarendszeréből átkerül a kamera koordinátarendszerébe egy transzformációs mátrix segítségével ( $R_{cr}$ ).

$$n_c = R_{cr}(0,1,0)^T \quad (5)$$

A sík egyenletében  $r$  meghatározásához szükség van egy  $r_0$  pont meghatározására. Tudjuk, hogy a kamera a saját koordinátarendszerében a  $(0,0,0)^T$  pontban található. Amennyiben a kamera magasságát az útfelülethez képest  $h$  változóval jelöljük, fel tudunk írni egy pontot az út felületén. A kamera origójából irányába indulunk, melyet  $h$ -val megszorozunk, így megkapjuk a pontunkat (6), majd a sík egyenletébe (4) behelyettesítve kapunk egy egyenletet (7), melybe pedig ezt követően az  $r$  helyére behelyettesíthetjük az előzőekben felírt,  $\lambda$ -tól függő egyenesünk egyenletét (3), így kapjuk meg a végső levezetést (8), melyben először rendezzük az egyenletet  $\lambda$ -ra, majd behelyettesítjük az egyenes egyenletébe (9), mely a végső megoldást adja.

$$r_0 = hn_c \quad (6)$$

$$0 = n_c^T(r - r_0) = n_c^T r - h$$

$$h = n_c^T r \quad (7)$$

$$h = n_c^T \lambda K^{-1}(u, v, 1)^T \Leftrightarrow \lambda = \frac{h}{n_c^T K^{-1}(u, v, 1)^T}$$

(8)

$$\begin{pmatrix} X_c \\ Y_c \\ Z_c \end{pmatrix} = \frac{h}{n_c^T K^{-1}(u, v, 1)^T} K^{-1} \begin{pmatrix} u \\ v \\ 1 \end{pmatrix}$$

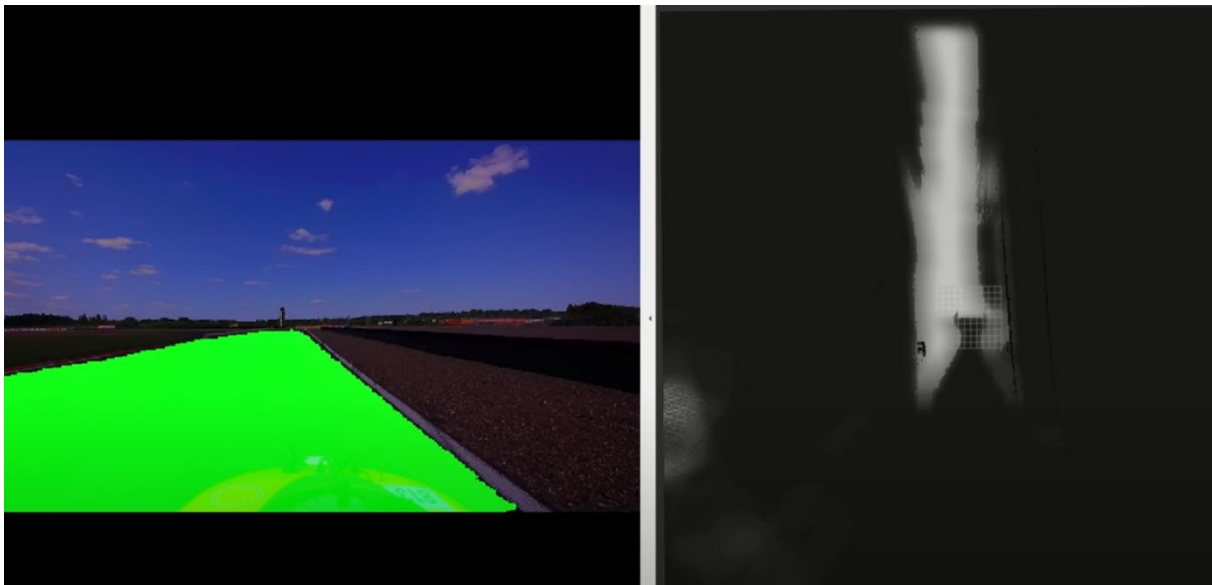
(9)

#### 4. Felhasználási módok autonóm területen

A deprojekció segítségével számos problémát oldhatunk meg, illetve használhatjuk kiindulásként az autonóm járművek és a robotika területén, hogy a környezetérzékelés által meghatározott térelemeket elhelyezzük három dimenzióban. Az objektumfelismerés és -követés algoritmusai gyakran használják a képi adatok 3D koordinátákká történő visszaalakítását, hogy a rendszer pontosabban azonosíthassa és követhesse az objektumokat a térben. Ezen kívül a navigációs rendszerek számára is elengedhetetlen az érzékelt akadályok és útvonalak háromdimenziós modellezése, amely lehetővé teszi az autonóm járművek számára, hogy biztonságosan navigáljanak egy dinamikus környezetben.

A tárgymanipuláció egy másik terület, ahol a deprojekció nélkülözhetetlen. A robotkarok és -kezek pontos mozgásaihoz esszenciális a valós térbeli távolságok és irányok ismerete, amelyek csak akkor számíthatók ki megbízhatóan, ha képesek vagyunk a képi adatokat a valós világ paramétereire visszavezetni. Ez különösen igaz olyan komplex feladatoknál, mint például az ember-robot együttműködés vagy a nagy precizitást igénylő ipari alkalmazások.

Az autonóm járművek területén az útjelzések és útvonalak érzékeléséhez és értelmezéséhez is szükség van deprojekcióra. Az általunk finomhangolt útfelismerő algoritmus [4] is ezen a területen alkalmazza a deprojekciós algoritmust. Mesterséges intelligencián és gépi látáson alapuló algoritmusunk [6] segítségével meghatározzuk a vezethető útfelületet. Ezt követően a meghatározott útfelület minden pixelének távolságát meghatározzuk a deprojekciós algoritmus segítségével, melynek eredményét a 4. ábrán láthatjuk.



**4. ábra:** A fejlesztett szemantikus szegmentációs algoritmus és a 3D-be való leképezése deprojekció segítségével. (saját szerkesztés)

## 5. Köszönetnyilvánítás

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

## Irodalomjegyzék

- [1.] R. Hartley and A. Zisserman: Multiple view geometry in computer vision, 2nd ed., Cambridge, UK; New York, Cambridge University Press, 2003.
- [2.] T. Langner: "Visual Perception for Autonomous Driving," Freie Universität Berlin, 2020, doi: 10.17169/REFUBIUM-26549.
- [3.] R. Szeliski: Computer Vision: Algorithms and Applications, 2nd edition, Cham, Springer Nature, 2022.
- [4.] C. Yu, C. Gao, J. Wang, G. Yu, C. Shen, and N. Sang: "BiSeNet V2: Bilateral Network with Guided Aggregation for Real-time Semantic Segmentation," arXiv, Apr. 05, 2020. Accessed: Aug. 31, 2023. [Online]. Available: <http://arxiv.org/abs/2004.02147>.

[5.] M. Theers and MankaranSingh: "thomasfermi/Algorithms-for-Automated-Driving: zenodo doi release," Zenodo, May 02, 2023, doi: 10.5281/ZENODO.7887756.

[6.] H. Zhong, H. Wang, Z. Wu, C. Zhang, Y. Zheng, and T. Tang: "A survey of LiDAR and camera fusion enhancement," Procedia Computer Science, vol. 183, pp. 579–588, 2021, doi: 10.1016/j.procs.2021.02.100.

# Fenntarthatósági közzétételek szógyakorisági elemzése a Scope 3 károsanyag kibocsátások mérésére Python környezetben

## Sustainability Accounting Scope 3 Emission Measurement Qualitative Modeling in Corporate ESG Reports

Molnár Péter<sup>a</sup>, Suta Alex<sup>b</sup>, Tóth Árpád<sup>c</sup>

<sup>a</sup>Széchenyi István Egyetem, H-9026 Győr, Egyetem tér. 1, Járműipari Kutatóközpont,  
molnar.peter@ga.sze.hu

<sup>b</sup>Széchenyi István Egyetem, H-9026 Győr, Egyetem tér. 1, Járműipari Kutatóközpont  
suta.alex@ga.sze.hu

<sup>c</sup>Széchenyi István Egyetem, H-9026 Győr, Egyetem tér. 1, Járműipari Kutatóközpont  
totha@ga.sze.hu

### Absztrakt

---

*A Scope 3- kibocsátások bemutatása - a jogi kötelezettségek folyamatos fejlődése alapján - magába foglalja a jelentéseket tevő vállalatok “upstream és downstream” közvetett üvegházhatásúgáz-kibocsátásait, amelyek az összes kibocsátás túlnyomó többségét teszik ki. Bár a fenntarthatóság iránt egyre nagyobb az érdeklődés a vállalatirányítás (fenntartható stratégia és ellátási láncok) és a külső felügyelet (az Európai Bizottság által 2022-ben elfogadott CSRD) szempontjából, a vállalati jelentésekben a Scope 3 kibocsátás bemutatása gyakran nem egyértelmű és változó minőségű. Ezek a kulcsfontosságú ESG-mérőszámok egyelőre nem tartalmaznak olyan információkat, amelyek lehetővé teszik a számítási módszertan és az alkalmazott képlet validálását, így hibák és eltérések tekintetében rés marad. Jelen tanulmányban a Scope 3- kibocsátások jelentésének fő technikai és szervezési kihívásait keressük és mutatjuk be, konkrét európai iparágak ESG-jelentéseire vonatkozó, szótáralapú szövegbányászati elemzések segítségével. Az eredményekre vonatkozó részben egy KWIC (kulcsszó a kontextusban) funkciót használtunk, amely alapján összehasonlítás és vizualizáció került bemutatásra. A jelentések hossza és adattelítettsége nem volt állandó, és e tekintetben a javulás kulcsfontosságú lenne a szabványosított, egyértelmű és releváns kibocsátási jelentések teljes mértékben digitalizált formátumban történő biztosításához.*

**Kulcsszavak:** Scope 3 kibocsátási jelentés, szövegbányászati elemzés, fenntarthatósági jelentés, vállalatirányítás, EFRAG, CSRD, ÜHG-csökkentési célok, ISSB, TCFD, NLP-modell, autóipar, energiaipar, környezeti hatások, hírnevet érintő kockázatok.

## **Abstract**

---

*The presentation of Scope 3 emissions—by an evolving set of legal obligations—includes the upstream and downstream indirect GHG emissions of reporting companies, which accounts for the vast majority of total emissions. Although there is a growing interest in sustainability from enterprise governance (sustainable strategy and supply chains) and external supervisory perspectives (the CSRD accepted by the European Commission in 2022), Scope 3 emission presentation in corporate reports is often ambiguous and of variable quality. These key ESG metrics are yet to include information that enables the validation of computational methodology and used formula, leaving a gap for error and discrepancy. In the current study, the main technical and organizational challenges of reporting for Scope 3 emissions are sought and presented using a dictionary-based text-mining analysis of ESG reports of specific European industries. In the results section, a KWIC (keyword-in-context) function was utilized based on which comparison and visualization were presented. The length and data saturation of the reports was not consistent, and improvement in this regard would be key to ensuring standardized, clear, and relevant emission reporting in a fully digitized format.*

**Keywords:** *Scope 3 emissions reporting, text mining analysis, sustainability reporting, corporate governance, EFRAG, CSRD, GHG reduction targets, ISSB, TCFD, NLP model, automotive, energy, environmental impacts, reputational risks*

## 1. Bevezetés

A kézirat egy előzetes kutatás eredménye, amely egy nemzetközi konferenciára készülő teljes kutatásnak ad majd alapot. A kutatás fókuszában a Scope 3-kibocsátások jelentései és az ESG (környezeti, szociális és vállalatiirányítási) jelentések elemzése áll, azzal a céllal, hogy megértsük ezeknek a jelentéseknek a fontosságát és kihívásait az iparági intézkedések és jelentéstétel gyakoriságának meghatározása szempontjából.

Az egyik kiemelt probléma azonban az, hogy a Scope 3-kibocsátásokról szóló jelentések nem mindig egyértelműek és változó minőségűek. Az információk hiánya az alkalmazott számítási módszerekről és képletekről problémát okoz mind a vállalatok, mind a külső ellenőrző szervek számára, miközben a fenntarthatóság iránti érdeklődés növekszik. További nehézségeket okoz a jelentések változó hossza és az adatok hiányossága, ami zavarokat és következetlenségeket szülhet a jelentéstételben.

Az irányadó testületek, mint például az amerikai Értékpapír- és Tőzsdefelügyelet (SEC) vagy az Európai Pénzügyi Beszámolási Tanácsadó Csoport (EFRAG), törekednek a Scope 3 kibocsátási beszámolás szabályozására. Az Európai Bizottság is bevezette a vállalati fenntarthatósági jelentésekről szóló irányelvet (CSRD), amely kötelezővé teszi a Scope 3 kibocsátási beszámolást. Az EFRAG az új irányelvhez kapcsolódóan dolgozza ki a jelentéskészítési szabványokat, amelyek a teljes Scope 3 kibocsátás mintegy 80%-át fogják lefedni.

Az Egyesült Államokban a vállalatoknak kötelező lesz jelenteniük a Scope 3-kibocsátásaikat, különösen akkor, ha jelentős mennyiségről vagy üvegházhatású gázok csökkentésére vonatkozó célokról van szó. Ez azt hangsúlyozza, hogy a vállalatoknak figyelembe kell venniük az egész értékláncukon fellépő emissziókat. A Scope 3-kibocsátások jelentése segíti a környezeti hatások értékelését és a fejlesztendő területek azonosítását, továbbá kifejezi a vállalatok elkötelezettségét az éghajlatváltozás elleni küzdelem és a felelős vállalatiirányítás mellett.

A globális szabványügyi testületek közül az ISSB (International Sustainability Standard Board) játszik kulcsszerepet. Az ISSB az IFRS S1 és S2 fenntarthatósági standardtervezeteket készítette. Az elkövetkező időszakban várható, hogy az ISSB ajánlásai összhangban lesznek a TCFD (Task Force on Climate-related Financial Disclosures) ajánlásaival, amelyek széles körben elfogadottak lettek. A TCFD ajánlásai segítik a vállalatokat az éghajlatváltozással

kapcsolatos kockázatok és lehetőségek nyilvános közlésében, ami a befektetők és a nyilvánosság számára is előnyös.

Összességében a vállalatoknak az éghajlattal kapcsolatos pénzügyi kockázatok és lehetőségek átlátható nyilvánosságra hozatalával való követése elősegíti a fenntarthatóság és a felelős vállalatirányítás iránti elkötelezettségüket, és segít javítani az éghajlatváltozással kapcsolatos döntéshozatalt és hosszú távú ellenálló képességet.

Ezen felül, a kutatás a gépjármű- és energiaiparban működő vállalatok jelentéseit vizsgálta. A szövegelemzési modell alkalmazása a jelentésekben található kulcsszavak és mondatok azonosítására segített. Az eredmények alapján következtetéseket vontak le a minőségi jelentések javítására és a természetes nyelvfeldolgozási technikák további kutatásaira vonatkozóan.

## **2. Szakirodalmi áttekintés**

Az IFRS S2 Exposure Draft alapján a gazdálkodó egységeknek teljesítménymutatókat és célokat kell közzétenniük a klímaváltozással kapcsolatos kockázatok és lehetőségek kezeléséhez és nyomon követéséhez. A Scope 1, 2 és 3 kibocsátásokat az S2 megközelítés ismeri el. A Scope 1-be tartoznak a közvetlen üvegházhatású gázok kibocsátásai a tulajdonban vagy ellenőrzés alatt álló forrásokból, míg a Scope 2-be az indirekt kibocsátások a vásárolt villamos energia előállításából. A Scope 3-ba tartoznak a szervezet értékláncán kívüli közvetett kibocsátások.

A vállalati Scope 3-jelentések elemzésére számos tanulmány készült. Például Hansen et al. (2022) élelmiszer- és italipari vállalatok jelentéseit vizsgálta, rámutatva, hogy a Scope 3-jelentések gyakran hiányosak és következtelenek. Ryan & Tiller (2022) tanulmánya új-zélandi vállalatok Scope 3-jelentéseit elemezte, kiemelve a jelentések hiányosságait és a felkészültség szükségességét az IFRS S2 alkalmazására. Az életciklus-elemzés jelentőségét az üvegházhatásúgáz-kibocsátási jelentéstételben Patchell (2018) is kiemelte.

A tartalomelemzés lehetőséget nyújt az ESG-jelentések minőségének értékelésére. Siew (2017) által meghatározott tartalomelemzés négy típusa közül a szószámlálás, fogalmi elemzés, relációs elemzés és referenciális elemzés alkalmazható. Az ESG-jelentések nyelvezetét, hangnemét és tartalmát vizsgálva megérthető a vállalat fenntarthatósági elkötelezettsége és az

ESG kérdések kezelése. A tartalomelemzés hasznos lehet az ESG-jelentésekben megfigyelhető tendenciák azonosítására, segítve befektetőket és szabályozókat a vállalatok fenntarthatósági teljesítményének értékelésében.

Baboukardos et al. (2022) az ESG-jelentések tartalomelemzésével mérte fel a vállalatok felkészültségét az IFRS S2 alkalmazására. A tanulmány ellenőrző listát alkalmazott a vállalatok megfelelésének értékelésére. Moreno és Caminero (2022) hasonlóan, a TCFD ajánlásai alapján elemzett, a Scope 3 kibocsátásokra koncentrálva, Natural Language Processing (NLP) elemzési technikát alkalmazott az értékelés során, a jelentések tartalom gazdagságának mérésére.

Ezen kutatások alapján az elemzések és a tartalomelemzési módszerek segíthetnek megérteni a Scope 3-jelentések minőségét, hibáit és javulási lehetőségeit, valamint hozzájárulhatnak a vállalatok fenntarthatósági teljesítményének értékeléséhez és fejlesztéséhez.

### **3. Módszertan**

Munkánk célja az volt, hogy az NLP módszertanát felhasználja a vállalati fenntarthatósági jelentések elemzésére. Az EMIS statisztikai nyilvántartást (amely a közép- és kelet-európai (KKE) régió vállalataira összpontosít) használták az elemzésre szánt vállalatok azonosítására. Korábbi kutatások alapján a közvetlen és közvetett kibocsátásokért felelős iparágak közül globálisan az elmúlt 20 éves időszakban az ipari szektorban tapasztalhattunk növekedést a legnagyobb mértékben (75%), amelyet az energiaszektor követ (61%) (Hertwich & Wood, 2018). Így ebben a két globálisan jelentős iparágban, az energiaiparban és az autóiparban (vagy gépjárműgyártásban), mint a Scope 3 kibocsátás szempontjából egyik legjelentősebb termelő iparágban kerestünk vállalatokat, a járművek gyártási és használati fázisai miatt. Ezt követően a két iparágat a 2021. évi legmagasabb nettó árbevétel alapján szűrtük a fontos piaci szereplők kiválasztása érdekében, és mindkét iparágból kiválasztottuk a 10 legnagyobb vállalatot. Megállapítottuk, hogy a vállalatok leányvállalatai korlátozott adatokat tartalmaznak fenntarthatósági jelentéseikben, mivel a korlátolt felelősségű társaságok esetében az ESG-adataik és hozzájárulásuk az anyavállalat állami vállalatának bevonásaként kerül kiemelésre és alapos bemutatásra, ezért ezután visszakövetésre kerültek a vállalatcsoportokhoz, hogy a legtöbb információt ki lehessen nyerni az adatforrásból. Az adatgyűjtés következő lépéseként a 20 kiválasztott vállalat 2021-es fenntarthatósági (ESG) vagy integrált jelentéseit gyűjtöttük össze az egyes vállalatok hivatalos honlapjairól .pdf kiterjesztésű fájlokban.

Ezt követően egy Python szövegbányászati munkafolyamatot fejlesztettünk ki, hogy a dokumentumokból a szöveget mondatok listájaként kivonjuk és megtisztítsuk. Ehhez egy elemzési módszert alkalmaztunk a "tika" csomag használatával, hogy a nyers szöveget mondatokként kivonjuk a fő szövegtestből. A kinyert szöveget ezután az "nltk" (Natural Language Toolkit) csomag segítségével szűrtük és tisztítottuk. Összesen körülbelül 35 ezer mondatot nyertünk ki, ebből 20 ezer az autóiipari vállalatoktól és 15 ezer az energiaipari vállalatoktól származott. A további szövegelemzés támogatása érdekében számos címkét határoztunk meg a releváns és átlátható Scope 3 kibocsátási adatok jelentésének alapvető szükségessége alapján. A címkék két csoportból álltak: (1. csoport) általános címkék, beleértve a "Scope 3", "Életciklus-értékelés" és "Mérés" címkéket, amelyek a fenntarthatósági közzétételben keresendő fő minőségi információkat képviselik, és (2. csoport) a Scope 3 kibocsátáshoz kapcsolódó számviteli kategóriák és tevékenységek esetei, amelyekről az IFRS S2 szabványtervezet (ISSB, 2022) szerint kell beszámolni.

A mondatok kategorizálása érdekében az elemzés során a "kulcsszó a szöveggörnyezetben" (KWIC) megközelítést alkalmaztuk, hogy megfigyeljük, hogy az egyes mondatok tartalmazzák-e a címkékben meghatározott kulcsszavak valamelyikét. A kulcsszavakat egy egyszerű gyakorisági alapú szólista létrehozásával azonosítottuk. A szószám-számítás lehetővé teszi a különböző ESG-jelentések közötti értelmes összehasonlításokat, valamint alapot teremt a jövőbeni elemzési módszerekhez, amelyek a szószám-számítási módszerekkel létrehozott szótárakra épülnek. A szószám-elemzéssel azonosíthatók a különböző vállalatok vagy iparágak által szolgáltatott információk mennyiségének különbségei.

#### **4. Eredmények**

Az „Exposure Draft” a szabvány 116. pontjához tartozó függelékében a javaslat szerint a szervezeteknek a Scope 3 kibocsátások közé kell sorolniuk az értéklánc elején és végén lévő kibocsátásokat, és magyarázatot kell adniuk a jelentésükbe bevont vagy abból kizárt tevékenységekről. E javaslat fő célja, hogy segítse az általános célú pénzügyi beszámolókat felhasználókat abban, hogy megértsék, mely kibocsátásokat vették figyelembe, és melyeket hagyták figyelmen kívül. A javaslatok lehetőséget biztosítanak a gazdálkodó egységeknek arra, hogy szabadon válasszák meg saját számítási módszerüket, feltéve, hogy az egyértelmű és nyilvános. Ugyanakkor elismerik, hogy a Scope 3 kibocsátási adatok összehasonlíthatósága,

lefedettsége, átláthatósága és megbízhatósága kihívást jelent. A Scope 3 kibocsátások bejelentésekor a szervezetek választhatják azt az általános megközelítést, hogy a portfólióállományuk százalékos arányához kapcsolódó kibocsátásokat a teljes kezelt vagyonra vonatkozó egyetlen számadatként teszik közzé.

**1. táblázat:** Iparágak közötti metrikus kategóriák a Scope 3 üvegházhatásúgáz-kibocsátásokhoz (IFRS S2 Exposure Draft)

Forrás (IFRS S2)	Közzététel leírása	Mérési útmutatás
21-a-i-3:	„absolute gross GHG emissions generated during the reporting period, measured in accordance with the Greenhouse Gas Protocol Corporate Standard, classified as Scope 3”	metric t of CO2 equivalent
21-a-ii	„GHG emissions intensity for Scope 3 emissions”	metric t of CO2 equivalent per unit of physical or economic output
21-a-vi-1	„an entity shall include upstream and downstream emissions in its measure of Scope 3 emissions”	not defined
21-a-vi-2	„an entity shall disclose the categories included within its measure of Scope 3 emissions to enable users of general-purpose financial reporting to understand which Scope 3 emissions have been included in, or excluded from, those reported”	not defined
21-a-vi-3	„when the entity’s measure of Scope 3 emissions includes”	
21-a-vi-4	„(4) if the entity excludes those greenhouse gas emissions in paragraph 21(a)(vi)(3), it shall state the reason for omitting them, for example, because it is unable to obtain a faithful measure;”	not defined

Az 1. táblázat azt mutatja, hogy a szabványtervezetben meghatározott hat jelentéstételi elemnek korlátozottak a konkrét elvárásai. Az első két elemnek van számszerűsíthető elvárása, amely a szabályozó által előírt komplex becslési technikák segítségével kiszámítható.

A szövegben bemutatott adatok feldolgozását az alábbi 2., 3. táblázatok tükrözik. A 2. táblázat a két iparági bontásban a címke kulcsszavakat tartalmazó mondatok összesített számát mutatja. Ez azt jelzi, hogy a Scope 3 kibocsátások minőségi közzététele tekintetében a két ágazat között vannak bizonyos különbségek. Konkrétan úgy tűnik, hogy az autóipar inkább a "vásárolt árukra és szolgáltatásokra" és az "eladott termékek feldolgozására" koncentrál, míg az energiaipar nagyobb hangsúlyt fektet az "életciklus-elemzésre" és a "keletkezett hulladéokra".

**2. táblázat:** A címke-specifikus kulcsszavakat tartalmazó mondatok száma iparáganként (saját szerkesztés)

<b>Címkék</b>	<b>Autóipar</b>	<b>Energia- ipar</b>	<b>Összesen</b>
scope 3	66	12	78
lifecycle analysis	126	1,432	1,558
measurement	1,757	817	2,574
purchased goods and services	2,153	615	2,768
capital goods	491	1,789	2,280
fuel and energy	514	223	737
waste generated	313	1,334	1,647
transportation and distribution	1,569	120	1,689
leased assets	115	892	1,007
processing of sold products	1,154	182	1,336
use of sold products	236	115	351
end-of-life treatment	354	25	379
business travel	81	7	88
employee commuting	26	336	362
franchises	568	22	590
investments	8	0	8
<b>Összesen</b>	<b>9,531</b>	<b>7,921</b>	<b>17,452</b>

A 3. táblázat a címkék részletesebb elemzését mutatja be, beleértve a Scope 3 kibocsátásokhoz kapcsolódó első három címke és a többi, pénzügyi és termelési tartalomra utaló címke megkülönböztetését. E két kategória metszéspontjai olyan mondatokat foglalnak magukban,

amelyek mindkét szemponthoz kapcsolódó kulcsszavakat tartalmaznak. Ezek a metszéspontok jelentősebb információt szolgáltatnak, és olyan fontos jellemzőket mutatnak meg, amelyek jobban megfelelnek a szövegbányászaton alapuló elemzéseknek, és a több címkéből származó kulcsszavak mondatban való együttes előfordulását vizsgálják. Ebben a kérdésben a címkék 1. csoportját (3. hatókör, életciklus-elemzés (LCA) és méréssel kapcsolatos szavak) és a 2. csoport címkéit állítottuk szembe egymással. Kiderült, hogy a két iparág között jelentős különbségek vannak. Figyelemre méltó, hogy az előző táblázathoz képest magasabb értékeket találunk, mivel egyes mondatokat potenciálisan több "2. csoportos" címke található.

**3. táblázat:** A több címkén is előforduló mondatok számának kereszttáblázata iparáganként (saját szerkesztés)

Címkék	Autóipar			Energiaipar		
	scope3	lca	measurement	scope3	lca	measurement
purchased goods and services	17	47	284	1	53	32
capital goods	6	14	48	0	106	191
fuel and energy	18	14	68	2	19	25
waste generated	5	24	32	0	114	446
transportation and distribution	14	16	183	0	5	5
leased assets	3	1	8	4	72	59
processing of sold products	6	25	131	0	30	16
use of sold products	11	5	41	2	7	29
end-of-life treatment	5	34	47	0	2	2
business travel	3	3	3	0	0	3
employee commuting	4	1	1	0	30	14
franchises	2	0	69	0	1	2
investments	0	0	0	0	0	0

## 5. Összegzés

Az IFRS S2 Exposure Draft standard hat beszámolási elemet határoz meg, korlátozott konkrét elvárásokkal. Az első kettőnek vannak számszerűsíthető elvárásai, amelyek komplex becslési technikákat igényelnek, de a standardból hiányzik az értékelési módszerekre vonatkozó útmutatás. Ehelyett olyan minőségi elvárásokat fogalmaz meg, amelyek a vállalatok számára alapvető információk közzétételét írják elő. Ezek a közzétételek várhatóan növelik az ESG-jelentések átláthatóságát és hosszát. A jelen tanulmány elemzési eredményei különbséget mutatnak a Scope 3 kibocsátások minőségi közzétételében az autóipar és az energiaipar között: az előbbi a vásárolt árukra és szolgáltatásokra és az eladott termékek feldolgozására, az utóbbi pedig az életciklus-elemzésre és a keletkező hulladékokra helyezi a hangsúlyt. A táblázatok olyan csomópontokat is feltárnak, amelyek a három kategória szövegbányászati elemzése szempontjából relevánsak: Scope 3, az életciklus-elemzés (LCA) és mérés. A bevont vállalatok tehát nem teljesítik széles körben a pontos Scope 3 kibocsátási jelentés követelményeit: több területen szürke mezők fedezhetők fel. Ebből az a következtetés vonható le, hogy szükség van egy jól meghatározott és pontos szabványra, valamint az átláthatóság érdekében például az XBRL jelentéstételi nyelv bevonására. Ezek az eredmények jelentősen hozzájárulnak az elemzés eredményeinek bemutatáshoz, amelyek megadják azokat a kulcsszavakat, amelyeket be kell építeni a vállalatok ESG-jelentéseibe. Ezeket a kulcsszavakat megfelelő értékelés kíséri, amelynek célja, hogy illeszkedjen a jelenlegi előzetes fenntarthatósági szabványoknak. A kutatás jelentősége abban rejlik, hogy elősegítheti az érdekeltek számára, hogy proaktív intézkedéseket tegyenek a jövőbeli szabályozásokra való felkészülésben. A további kutatási területek célja lehet e kutatás hatókörének kiterjesztése az ipari minta növelésével, valamint a címkézési és osztályozási módszertan továbbfejlesztése a gépi tanulási technikák esetleges beépítésével.

## 6. Köszönetnyilvánítás

A kutatást az Európai Unió támogatta a Nemzeti Mesterséges Intelligencia Laboratórium keretében (RRF-2.3.1-21-2022-00004). A cikk előzményét képezi a PRES'23 konferencián prezentált eredményekhez kapcsolódó kutatás.

## Irodalomjegyzék

- [1.] Baboukardos, D., Seretis, E., Slack, R., Tsalavoutas, Y., & Tsoligkas, F. (2022). Companies' Baboukardos D., Seretis E., Slack R., Tsalavoutas Y., Tsoligkas F., 2022, Companies' readiness to adopt IFRS S2 climate-related disclosures. ACCA and Adam Smith Business School, DOI: 10.36399/gla.pubs.278152.
- [2.] Cicchiello A.F., Marrazza F., Perdichizzi S., 2022, Non-financial disclosure regulation and environmental, social, and governance (ESG) performance: The case of EU and US firms. Corporate Social Responsibility and Environmental Management, 30(3), 1121-1128, DOI: 10.1002/csr.2408.
- [3.] EU 2022, Directive (EU) 2022/2464, available: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022L2464>, accessed: 08.08.2023
- [4.] Hansen A.D., Kuramochi T., Wicke B., 2022., The status of corporate greenhouse gas emissions reporting in the food sector: An evaluation of food and beverage manufacturers. Journal of Cleaner Production, 361, 132279, DOI: 10.1016/j.jclepro.2022.132279.
- [5.] Hertwich E.G., Wood R., 2018, The growing importance of scope 3 greenhouse gas emissions from industry. Environmental Research Letters, 13(10), 104013.
- [6.] ISSB 2022, Exposure Draft ED/2022/S2 Climate-related Disclosures. IFRS Foundation available: <https://www.ifrs.org/content/dam/ifrs/publications/pdf-standards-issb/english/2023/issued/part-a/issb-2023-a-ifrs-s2-climate-related-disclosures.pdf?bypass=on>, accessed: 08.08.2023.
- [7.] Moreno A.I., Caminero T., 2022, Application of text mining to the analysis of climate-related disclosures. International Review of Financial Analysis, 83, 102307, DOI: 10.1016/j.irfa.2022.102307.
- [8.] Patchell J., 2018, Can the implications of the GHG Protocol's scope 3 standard be realized? Journal of Cleaner Production, 185, 941–958, DOI: 10.1016/j.jclepro.2018.03.003.
- [9.] Perello-Marin M.R., Rodríguez-Rodríguez R., Alfaro-Saiz J.-J., 2022, Analysing GRI reports for the disclosure of SDG contribution in European car manufacturers. Technological Forecasting and Social Change, 181, 121744, DOI: 10.1016/j.techfore.2022.121744.

- [10.] Stemler, Steve (2001) "An overview of content analysis," Practical Assessment, Research, and Evaluation: Vol. 7, Article 17., DOI: <https://doi.org/10.7275/z6fm-2e34>, Available at: <https://scholarworks.umass.edu/pare/vol7/iss1/17>
- [11.] Ryan J., Tiller D., 2022, A Recent Survey of GHG Emissions Reporting and Assurance. Australian Accounting Review, 32(2), 181–187, DOI: 10.1111/auar.12364.
- [12.] SEC 2021, The Enhancement and Standardization of Climate-Related Disclosures for Investors. available: <https://www.federalregister.gov/documents/2022/05/12/2022-10194/the-enhancement-and-standardization-of-climate-related-disclosures-for-investors>, accessed: 08.08.2023.
- [13.] Siew R.Y.J., 2017. Critical Evaluation of Environmental, Social and Governance Disclosures of Malaysian Property and Construction Companies. Construction Economics and Building, 17(2), 81–91, DOI: 10.5130/AJCEB.v17i2.5328.

# Kamera alapú intelligens buszvezető monitoring rendszer tervezése

## Camera-based intelligent bus driver monitoring system design

Nagy Viktor<sup>a</sup>, Bézi Patrik<sup>b</sup>

<sup>a</sup>Széchenyi István Egyetem, Járműipari Kutatóközpont

nvitkor@ga.sze.hu

<sup>b</sup>Széchenyi István Egyetem, Járműipari Kutatóközpont

bezi.patrik@ga.sze.hu

### Absztrakt

---

*A kamera alapú járművezető monitoring rendszerek elterjedése a járműiparban jelentős, a balesetmegelőzésben betöltött szerepe megkérdőjelezhetetlen. A kutatási projekt célja a közúti közösségi közlekedésre adaptált intelligens megfigyelő rendszer tervezése és fejlesztése. A komplex rendszer funkcionálisan az emberi arc és az ember test elemeinek azonosítására alkalmas, és integrálható egy moduláris, ROS2 keretrendszerre épülő vizsgálati rendszerbe. A buszvezető megfigyelésére alkalmas rendszer tervezése a vezetői és közlekedési környezet speciális paramétereit figyelembe véve történik.*

**Kulcsszavak:** járművezető monitoring, gépi látás, ROS2, buszvezető

### Abstract

---

*The spread of camera-based driver monitoring systems in the automotive industry is significant and their role in accident prevention is unquestionable. The aim of this research project is to design and develop an intelligent monitoring system adapted to public road transport. The complex system is functionally capable of identifying human face and body elements and can be integrated into a modular inspection system based on the ROS2 framework. The design of the system for monitoring the bus driver will take into account the specific parameters of the driving and traffic environment.*

**Keywords:** driver monitoring, machine vision, ROS2, bus driver

## 1. Bevezetés

A biztonságos tömegközlekedés iránti igény egyre nagyobb követelményeket támaszt a járművezetőkkel szemben, mivel növekszik a forgalomsűrűség, amihez komplex forgalmi elem-összetétel és gyors külső ingerek megjelenése társul, különösen a városi környezetben. A fejlett vezetőtámogató rendszerek (ADAS) részben beavatkozhatnak vészhelyzetekben, például fékezéskor [1], de továbbra is a járművezető feladata és felelőssége a jármű biztonságos irányítása. Az ADAS célja a biztonság maximalizálása, ugyanakkor az ilyen összetett rendszereknek mellékhatásaik vannak. Az elfogadás és a felhasználás nehézkes lehet, tanulási folyamatra és bizalom növelésére van szükség [2]. A járművezető figyelmének megfigyelése és a kognitív terhelés elemzése lehetőséget nyújt a munkakörülmények optimalizálására és a baleset megelőzési megoldások javítására.

A közúti tömegközlekedés elterjedtsége Európában még mindig jelentős. Az utasok biztonságának biztosítása prioritás, mivel a járművek egyre több utast szállítanak a növekvő forgalomsűrűség mellett. Bár az autóbuszok (a villamosok és trolibuszok mellett) egyre kisebb arányt képviselnek az összes közlekedési mód mellett, részesedésük a 2020-as adatok alapján még mindig 7,4% az EU-ban [3]. Az EU-ban az autóbuszok utaskilométereinek száma több mint 138 ezer millió kilométer, a COVID-19 miatt bekövetkezett 40%-os csökkenés után (1. táblázat) [4]. A balesetekre vonatkozó történelmi adatok áttekintése azt mutatja, hogy az Európai Unió országaiban a 2017-2019-es adatok alapján az autóbuszos balesetekben meghaltak átlagos száma egymillió lakosra vetítve 1,3 fő. A halálos balesetek áldozatai többnyire gyalogosok, autósok és autóbuszok utasai. A baleseti sérülései statisztikák (2010-2019) azt mutatják, hogy az autóbusz balesetek 2%-a súlyos sérülésekkel jár [3].

**1. táblázat:** Autóbuszok, autóbuszok és trolibuszok Európai Unió - 27 ország [4].

Év	2013	2014	2015	2016	2017	2018	2019	2020	2021
Részarány (%)	10.4	10.0	10.0	10.0	9.6	9.5	9.5	7.4	N/A
Utas-kilométer (millió km)	238 368	230 616	246 661	240 574	224 813	225 958	231 099	135 427	138 262

Az adatok elemzése során számos tényezőt azonosítottak, amelyek befolyásolják a balesetek előfordulását: az autóbusz-üzemeltetők gondatlansága, a járművezetők mulasztása és a külső tényezők (időjárási és útviszonyok) [4]. A járművezetői hiba több tényezőre bontható, jellemzően az aktuális mentális és fizikai állapottól függően, így a járművezetők viselkedésének nyomon követése és a későbbi passzív vagy aktív beavatkozások biztosíthatják a biztonságosabb buszközlekedést [5]. A figyelemelterelő tényezők olyan vizuális, auditív, fizikai vagy kognitív ingerek, amelyek zavarják a vezetési tevékenységet, valamint az olyan másodlagos tevékenységek, mint a mobiltelefon-használat, az utasokkal való kommunikáció és az „álmodozás”. A fáradtság, a monotonitás és az alvásmegvonás növeli a balesetek kockázatát, mivel a csökkent figyelem az információfeldolgozás csökkenéséhez és a döntésképtelenséghez vezethet, így a járművezető vészhelyzetben képtelen reagálni [6]. Az éberséggel kapcsolatos legtöbb kutatás az alvásmegvonásra összpontosít [7-9]. Baleseti adatok és szimulációs vizsgálatok azonban arra utalnak, hogy az éberség csökkenése napközben is előfordulhat, különösen hosszú, monoton utakon [10]. A monoton vezetés során gyakoribbak a járművezetői hibák, mivel a feladatigény és az ingerek szintje alacsony, ami csökkenti a figyelmet.

## 2. Irodalomkutatás

Különböző monitoring rendszereket fejlesztettek ki és alkalmaznak a járművezetők viselkedésének megfigyelésére vagy észlelésére. A különböző megközelítések pszichológiai tesztek és különböző fiziológiai érzékelőket, adatgyűjtési és elemzési technikákat alkalmaznak, az alábbiak szerint: (1) járműadat-alapú mérések; (2) viselkedési mérések és (3) fiziológiai mérések [11]. A járműadat-alapú vizsgálatokat több kutatásban a járművezetői magatartás azonosítása érdekében használták. Többek között kifejlesztettek egy módszert a vezetési szokások felismerésére egyetlen kanyar járműadatainak felhasználásával [12]. Az ember-jármű interakciót elemző rendszerek az üzemanyag-fogyasztás, a CO<sub>2</sub>-kibocsátás, a vezetési stílus és a vezető egészségi állapotának valós idejű, magas előrejelzési megbízhatósággal történő megfigyelésére alkalmasak [13].

Számos viselkedéselemzés – főként a járművezetői magatartás kérdőív segítségével – kimutatta, hogy a hivatásos járművezetők kevésbé kockázatos magatartást tanúsítanak, mint a nem hivatásos járművezetők, de a hosszabb vezetési idő miatt nagyobb valószínűséggel vesznek részt közlekedési balesetekben [14]. Egyes tanulmányok bebizonyították, hogy a

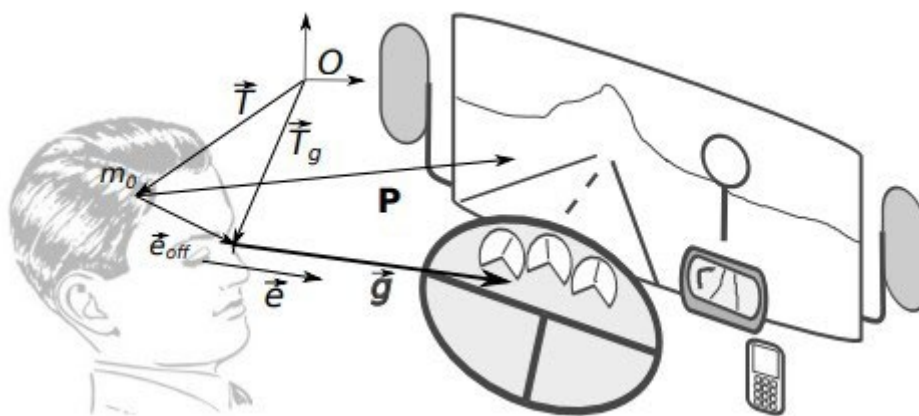
járművezető viselkedése a vezetési körülményekkel (pl. a kényelem szintjével) van összefüggésben, és közvetlenül befolyásolják a vezetési teljesítményt [15]. A helyközi autóbusz-vezetők nagy mintáját elemezték a kérdőívekkel, és a pszichometriai eredmények rávilágítottak a biztonságos és a nem biztonságos járművezetők megkülönböztetésére szolgáló diagnosztika fontosságára [16].

A járművezető fáradtság és disztrakció (figyelemelterelés vagy -zavarás) észlelése alapulhat biometrikus jeleken, a kormányzás érzékelésén vagy a járművezető arcának megfigyelésén [17]. Egyes kutatók úgy találták, hogy a magasabb átlagos szívritmus (HR) azt mutatja, hogy a járművezető megterhelőbb vagy másodlagos feladatokat végez [18]. Mások elektroencefalográfiával validált szívfrekvencia-variabilitási elemzéseket végeztek a járművezető álmoságának észlelésére [19]. Több szerző viselhető Galvanic Skin Responses (GSR) csuklópántot használt a járművezetők figyelemelterelésének azonosítására [20]. Az elektroencefalogramot (EEG) széles körben használják a járművezetői viselkedés felismerésére, de a jelek nem stacionárisak, és az eljárás főként a több érzékelős rendszerek használata esetén zavaró a vizsgálati résztvevők számára [21]. Nemrégiben egycsatornás EEG-alapú álmoság-érzékelő rendszereket teszteltek, és egyetlen jellemző paraméter elemzése lehetővé tette az alacsonyabb feldolgozási és tárolási kapacitást, ami egyszerűbb megvalósítást eredményezett beágyazott rendszerekben [22]. Más fiziológiai érzékelőket - például a járművezető izomfáradtságát mérő elektromiogramot (EMG) - is használtak, és ismertették a megfelelő mérési módszertant [23].

A járművezető monitoring rendszerek (Driver Monitoring System – DMS) legelterjedtebb megoldása a nem viselhető, kameraalapú rendszerek, amelyek gépi látást használnak az arcfelismeréshez. Ez a megközelítés kényelmes a járművezetők számára, mivel nem igényel viselhető eszközök használatát, így nem zavaró megfigyelést tesz lehetővé, és tökéletesen integrálható sorozatgyártású járművekbe is. A 2019/2144/EU rendelet [24] szerint 2024-től minden új európai járművön alkalmazni kell a "figyelemfelhívást a járművezető álmosága vagy figyelemelterelése esetén". A járművezető felé néző, különböző látószögű vizuális érzékelőket, RGB-, IR- vagy egyéb kamerákat használnak a vezetési adatok gyűjtésére - beleértve a járművezető interakcióját, viselkedését és a környezetet (vezetőfülke). Általában a következő adatkategóriákat észlelik: kézmozgás, testhelyzet-becslés, arcfelismerés és

figyelemelterelés/álmoság [25]. Az arcfelismerési módszerek általában a következő kategóriákba sorolhatók: jellemző-alapú és tanulás-alapú módszerek [40]. A tanulás-alapú módszerek általában robusztusabbak, mint a jellemző-alapú módszerek, de többnyire több számítási erőforrást igényelnek. Ezek a módszerek azonban laboratóriumi körülmények között több mint 80%-os felismerési arányt érhetnek el [26].

Az ilyen rendszerek esetén a fő követelmény, hogy a rendszer felhasználótól függetlenül működjön. Bizonyos jellemzők automatikus felismerése történik az emberi arcon egy sztereó kamera segítségével. Harris sarokdetektálás segítségével az algoritmus a két kamerás megoldásnak köszönhetően képes térben elhelyezni ezeket a pontokat, gyakorlatilag egy 3D arcmodell formájában. Ezek a jellemzők követve lesznek az egyes képkockák között, és a várt pozíciójukból kiugró pontok szűrésre kerülnek. A két kamera által felismert jellemzők egymásnak való megfeleltetése történik, amely elősegíti a fejmozgás meghatározását. Levenberg-Marquardt nevű eljárás segítségével történik a fejtartás becslése. Ezek mellett történik egy kiegészítő folyamat is, amely új jellemzőket ad a fejhez, mikor az mozog, ezáltal annak új (eddig ismeretlen) területeit tudja vizsgálni a rendszer. Így a teljes yaw-rotáció (oldalirányú kitérés)  $\pm 90^\circ$ -ig megengedett.



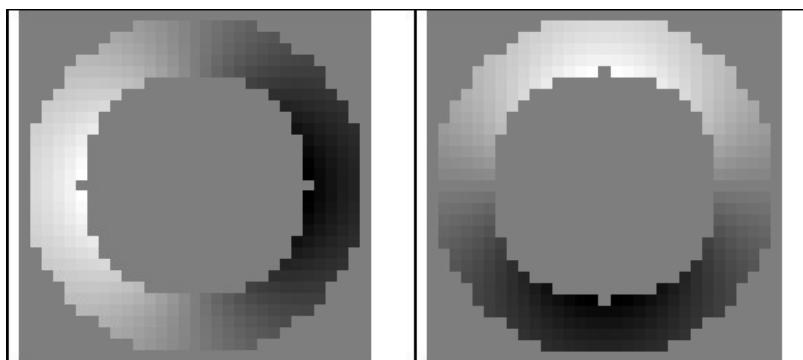
**1. ábra:** Az arc helyzete és a szem irányának kapcsolata [27]

Szem irányát meg lehet határozni az arcmodell koordináta-rendszere alapján, ahol  $\vec{e}$  a szem iránya (1. ábra). A  $G$  tekintet az arc helyzete és a szemek iránya alapján számítható. Ez lehet egy  $T_g$  irányából mutató  $\vec{g}$  vektor. Tehát:

$$G = \{T_g, \vec{g}\}, T_g = \vec{T} + R \cdot \vec{e}_{off}, \vec{g} = R \cdot \vec{e}$$

ahol  $P = \{\vec{T}, R\}$  az arc helyzete,  $\vec{e}_{off}$  pedig a szemközepek és a  $\vec{T}$  arcmodell közepének a távolsága.  $R$  a fej forgási szöge,  $m_0$  pedig a fejmodell közepe.

A valós idejű szemfelismerés, azaz a pupillák felismerése egyszerű küszöbértékeléssel valósítható meg, ahol a sötét- és világos pupilla értékek különbségét kell alapul venni. Az írisz mindig sötétebb a szem ínhártyájánál, szintől függetlenül. Ennek köszönhetően az írisz élei könnyen felismerhetők egy körre helyezett pontok halmazaként. Mivel más dimenziókkal rendelkeznek az emberek íriszei és a kamera-ember távolság sem feltétlenül ismert, így az írisz átmérőjének méretét nehéz meghatározni. Legyen  $[R_{min}, R_{max}]$  intervallum, amely a különböző átmérők felismerésénél lesz felhasználva. Ehhez egy új operátort kell felhasználni, amely az úgynevezett Kör Hough Transzformáció irányára alapszik. Az egyes "éldetektált" képeken konvolúciókat kell alkalmazni.



**2. ábra:** A pupilla felismeréséhez használt maszk [28]

Az 2. ábrán látható maszk reprezentálja minden pontban a radiális vektornak az irányát. A vektor nagysága függ a gyűrűtől, amelyet az  $R_{min}$  és  $R_{max}$  sugarú körök határoznak meg. Ez két távolságérték a belső kör középpontjától számítva. A konvolúció a gradiens képek és a fenti

maszkok között segít kiértékelni, hány pont vág egybe a meghatározott gyűrűvel. A konvolúcióból számított maximális érték mondja meg a képen, hogy hol a kör középpontja.

Az első maximum megtalálása után az egyik szem régióját kell meghatározni. Ebből a területből kiindulva ellenkező irányba haladva található a másik szem, amely megfelel mind távolságban mind orientációban a másik szemnek. Legyen  $(x_1, y_1)$  az első maximum középpontja. A második  $(x_2, y_2)$  pont, ahol a másik szem található. Ezek meghatározása a következő képletek segítségével történik:

$$x_2 = x_1 + /-D \quad D \in [d_1, d_2]$$

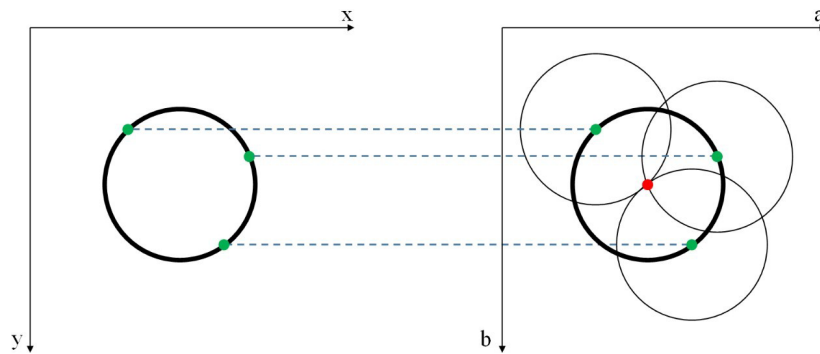
$$y_2 = \tan \alpha \cdot (x_2 - x_1) + y_1 \quad \alpha \in \left[-\frac{\pi}{6}, \frac{\pi}{6}\right]$$

Ez a módszer nagyon jól szűri a fals pozitív eredményeket, ha például az ember haja az arcába lóg és kitakarja a másik szemét, illetve, ha az algoritmus megtalálna más testrészeket (pl. száj, orr stb.). Két régió hasonlóságát egyszerű átlagolással lehet számolni:

$$MAE = \frac{\sum_{i=1}^N \sum_{j=1}^M |a_{i,j} - b_{i(M-j+1)}|}{N \cdot M}$$

ahol a és b a két régiópontot jelöli  $N \times M$  dimenzióval. A második régió tükrözése fontos, hogy a szimmetriának megfelelően lehessen kiértékelni a detektálás hatékonyságát. Ha a hasonlóság egy adott küszöbérték alatt van, akkor elfogadásra kerülnek a szemek, ha nem, akkor további keresés lesz szükséges.

Hough transzformáció segítségével köröket is tudunk keresni. A működés a következő módon történik: bizonyos sugárnagyságú körök rajzolása, amelyek egy bemeneti kép körének élein haladnak át. Az a pont, ahol az összes kör metszi egymást, lesz a kör középpontja (3. ábra). A rádiuszt/méretet nehezebb megtalálni, mint a középpontot, így ez kihívás elé állítja az algoritmust.



**3. ábra:** A Hough transzformáció működése (saját szerkesztés)

A megfelelő algoritmus alkalmazhatóságának tesztelése és előnyeinek bizonyítása kihívást jelent, de hosszú távon az optimalizált arcfelismerés lehetővé teszi egy magasabb megbízhatósági szintű rendszer kifejlesztését, és növeli a közlekedésbiztonságot. Jelen cikk a monitoring rendszer tervezését mutatja be a szoftver és hardver elemek kiválasztása, illetve a vizsgálati célok meghatározása segítségével.

### 3. A rendszer elemei

A monitoring rendszer felépítése az egyes (hardver és szoftver) elemek kiválasztásával kezdődik majd azok integrálása teszi lehetővé a rendszer fejlesztését és az implementációját.

#### 3.1 Hardver

A szoftver és az egyes hardverelemek az NVIDIA Jetson Nano nevű alaplappal segítségével kerülnek összekapcsolásra (4. ábra). Ennek a számítógépnek az előnye a kis mérete mellett, hogy kifejezetten gépi látásos módszerek alkalmazására fejlesztették ki. Támogatja többek között a CUDA platformot, amely a képi adatok feldolgozását nagyban felgyorsítja.



**4. ábra:** A Jetson Nano alaplappal (saját szerkesztés)

A Pupil Core egy szemkövető platform, amely nyílt forráskódú szoftvert és egy szemmozgást figyelő headsetet tartalmaz, amelyet kutatók globális közössége használ. Ennek a működése két szemkövető kamerán és egy homlokkamerán alapszik. Előbbiek 200 Hz-en 192x192 felbontáson működnek, míg utóbbi a 1080p felbontás mellett 30 Hz-en 720p esetében 60 Hz-en, 480p-nél pedig 120 Hz-en képes megjeleníteni az egyes képkockákat [29].



a)



b)

**5. ábra:** A PupilCore Headsetje(a) [29], és az általa becsült tekintet pont, amit zöld körfelület, fixáció esetén egy sárga kör jelöl – a képen a sofőr a külső visszapillantó tükörbe néz éppen (b) (saját szerkesztés)

Mivel a bemutatott PupilCore platform kísérleti célokra alkalmas, de valós vezetési környezetben (autóbusz járat utasokkal) nem alkalmazható, így a szoftver elsődlegesen egy külső kamerával fog működni. Ezek közül is a választás a beépített IMU-val rendelkező Intel RealSense kamerára esett, amely megfelelő felbontás mellett számos hasznos képi információval szolgálhat a későbbiekben (6. ábra).



**6. ábra:** Az Intel RealSense kamera (saját szerkesztés)

A kamera elhelyezése a járművezetőre történő megfelelő rálátás biztosítása mellett fog megtörténni. A vezetőhely statikus felmérése során két potenciális hely került kijelölésre, melyek a 7. ábrán láthatók. Az optimális, szemből történő detektálás érdekében a kormány tengelyének vonalában, a műszerfal tetején kell a kamera modult rögzíteni.



**7. ábra:** Az Intel RealSense kamera elhelyezési lehetőségei a vezetőhelyen (sárga szaggatott keret) piros szaggatott keret a visszapillantó tükröket és a fedélzeti kamerarendszer kijelzőjét jelölik (saját szerkesztés)

### 3.2 Szoftverkövetelmények

A szoftver Python 3.8-as környezetben kerül megvalósításra, OpenCV 4.5.5 verziójú keretrendszerrel. Ennek előnye a platformfüggetlenség, vagyis UNIX és Windows alapú rendszereken is lehet tesztelni, fejleszteni a rendszert. Ez az egyik legszélesebb körben használt kombináció gépi látásos megoldások esetében. A Python lehetővé teszi az egyszerű kódolást és az összes felhasznált keretrendszert könnyű beüzemelni vele. Az OpenCV esetében elengedhetetlen a forrásból való telepítés annak érdekében, hogy a CUDA platform használható legyen. Ennek a platformnak nagy előnye, hogy a neurális hálózatokkal való feldolgozást

nagymértékben gyorsítja. Ez a GPU általános célú feldolgozásait teszi lehetővé párhuzamos módon a szoftverben [30].

A szoftver vázát az ROS2 keretrendszer adja [31]. A Robot Operating System (ROS) egy olyan keretrendszer, amely elsősorban robot applikációk/szoftverek készítésére készült. Programozható C++ és Python nyelvek alatt, számos előnnyel rendelkezik. Többek között az, hogy nyílt forráskódú, és az egyes szenzorok szinkronizálását lehetővé teszi. Számos state of the art jellegű algoritmus fűződik hozzá, és rendkívül sokrétű. Felhasználható ugyanis kommunikációban, vizualizációban, észlelésre, gépi tanulásban, gépi látásos algoritmusoknál, szimulációkban, adat naplózásra, robotok és egyéb okos eszközök irányítására. Ezek közül számos olyan tulajdonsága van, amely az önvezető járművek fejlesztésére is alkalmassá teszi.

Megjelent már az ROS2 is, melynek működése hasonló az ROS1-hez, de számos felhasználást könnyítő funkcióval bővült.

A ROS segítségével az egyes részfeladatok megoldását külön lehet választani úgynevezett node-okra, amelyek felelősek az adatfeldolgozásért, és a kimenetek publikálásáért. Ezekre fel lehet iratkozni, és tovább feldolgozni őket. Előny tehát a modularitás, hiszen így akár a publikált adatok más fejlesztések során is felhasználhatóak lesznek.

#### **4. A rendszer tervezése**

A program alapvetően 2 fő része a testtartásfigyelés, és az emberi arc mimikájának, főleg a szemrégióknak a megfigyelése. Ahhoz, hogy analizálni tudjon a szoftver, szükség van a vizsgált területek felismerésére. Ezen feladatok mind a kamerához tartoznak. A szemkövetéshez szükség van arra, hogy az emberi arc megfelelő területeit, azon belül is magát a szemet felismerjük.

#### 4.1 Funkcionális követelmények

- Rögzített felvételeken a sofőr detektálása szemből
- Felismert alakzat kulcspontjainak (fej, törzs, végtagok) helyzetének lekövetése
- Emberi arcfelismerés, annak kulcspontjainak vizsgálata
- Vizuális megjelenítés a detektálásról
- ROS2 keretrendszer segítségével az egyes adatok kiküldése (folyamat: topic létrehozása, adatküldés, adatfelhasználás)
- .rosvag formátumú állományok rögzítése

#### 4.2 Nem funkcionális követelmények

*A rendszer megbízhatósága:* A rendszer megbízhatósága azon alapszik, hogy a kívánt emberi alakzatot (a sofőrt) kellő pontossággal ismeri fel a szoftver. Ide tartozik még a klasszikus fáradtság-, tekintetfigyelés működése, valamint, hogy az ezzel kapcsolatos adatok biztosan bekerülnek a megfelelő ROS2 topicba.

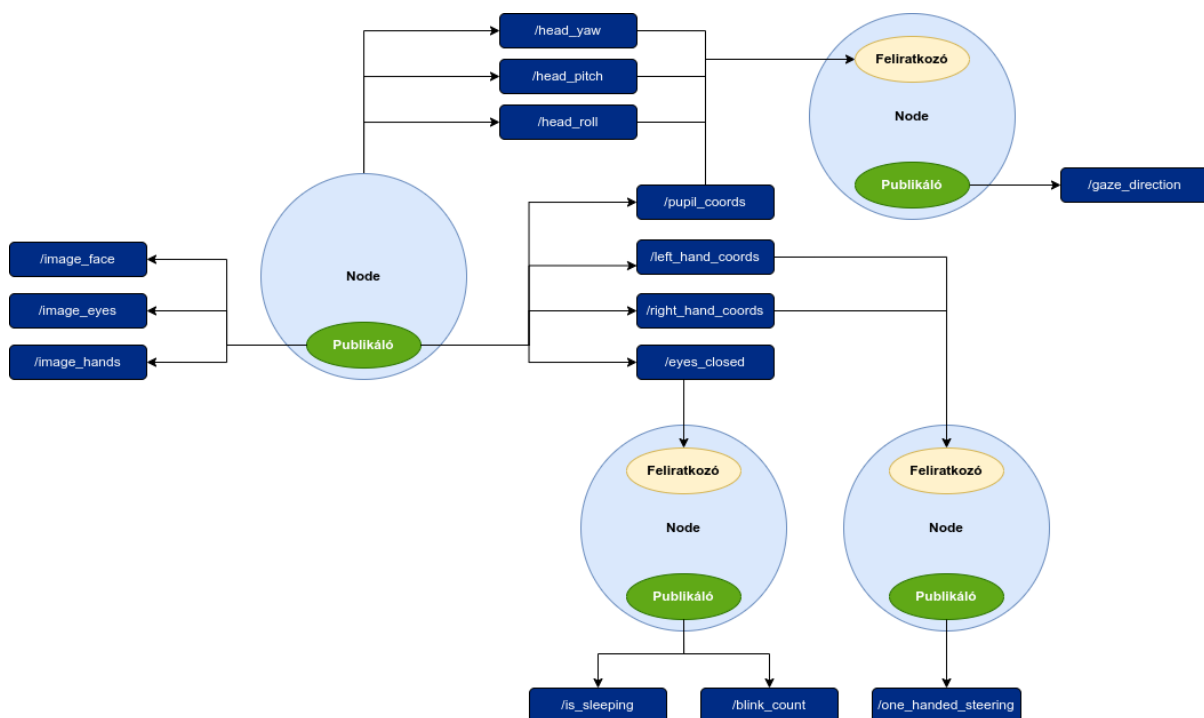
*A rendszer összeomlásával kapcsolatos követelmények:* Az összeomlást mindig megfelelően kell kezelni. A legnagyobb valószínűséggel vagy akkor történik ilyen esemény, ha nincsen bemeneti kép, vagy ha a szoftver egyes elemei nem megfelelő sorrendben indulnak el. Ez utóbbira példa, ha előbb próbál küldeni a szoftver ROS topicra üzenetet, minthogy a hozzá szükséges programcsomagok betöltésre kerültek volna. A fejlesztés fontos része a tesztelés is, így a projekt során minél több esetet meg fogunk vizsgálni a hibák elkerüléséhez.

*State of the art módszerek:* Objektumdetektálásra több state of the art jellegű megoldást alkalmazunk. Nevezetesen a YOLO v7 verzióját, a Google MediaPipe nevű neurális hálózatát és a dLib könyvtárat. Első kettővel a cél a teljes emberi test és annak kulcspontjainak felismerése (fej, törzs, végtagok) és lekövetése. A dLib eszköztár lehetőséget biztosít az emberi arc fontos részeinek a követésére. Ily módon lehetőség nyílik a szemek megfigyelésére, például a pislogás és a fáradtság tanulmányozására.

*Kiértékelés:* Az egyes megoldások hatékonyságát a fejlesztés során folyamatosan teszteljük és kiértékeljük.

### 4.3 A szoftver felépítése

A 8. ábrán az látható, hogyan kapcsolódnak egymáshoz az egyes node-ok, milyen üzeneteket hirdetnek és azok mely más feladat végrehajtása során lesznek hasznosak. A bemenet egy RGB kép, a kiindulópont ezen az emberi alakzat felismerése. Ebben a nodeban nem történik feliratkozás, csak publikálás. Összesen 10 féle adat kerül ki különböző topicokba. Ezek közül 3 ellenőrzés céljából képi adatok, amelyeken megfigyelhetők a felismert alakzat főbb kulcsponjtjai, mint a szemek, az arc és a kezek. A fontosabb rész az egyes testrészekhez tartozó információk publikálása, amelyekre más nodeok is feliratkoznak. A fejtartásból és a pupillakoordináták felhasználásával meghatározhatóvá válik a tekintetfigyelés. A szemnek nemcsak a mozgását lehet lekövetni, hanem azt is, hogy éppen nyitva van-e, vagy sem. Ez lehetővé teszi annak a megfigyelését, hogy a sofőr elalszik-e a volán mögött, vagy csak pislog. Utóbbit lehet számlálni is az egyes felvételeken, ebből megfigyelhető, mennyire koncentrált és éber az illető. A két kéz koordinátájából meghatározható, hogy a sofőr keze a kormánykeréken van-e, illetve, hogy milyen gyakran kormányoz egy kézzel.



8. ábra: A szoftver nodejainak kapcsolata (saját szerkesztés)

#### 4.4 Neurális háló

Az általunk alkalmazott módszer a neurális hálók egy speciális formáját, a konvolúciós neurális hálózatot alkalmazza, amely elsősorban képi feldolgozásra lett kifejlesztve. Előnye a hagyományos hálózatokkal szemben, hogy nem csak globálisan vizsgálja egy-egy kép jellemzőit, hanem lokálisan. Ez annyit jelent, hogy a különböző képi jellemzők segítségével bizonyos transzformáción átesett bemeneteken is megtalálni a keresett osztályokat, hiszen a konvolúciók az adott objektum kulcspontjait, éleit, színét, alakját mind figyelembe veszik, és ez alapján nem jelent számukra problémát, hogy megtalálják az objektumok koordinátáit.

A konvolúciós neurális hálózatok egyszerűsített működési elve a következő: A hálózatok struktúrájában a konvolúciós rétegek vannak, amelyekben végighaladva folyamatosan szűrődnek ki az egyes jellemzők. Az egyik ilyen legfontosabb réteg az úgynevezett *pooling*, amely a redundáns adatokat vonja össze a képeket leíró mátrixokból, és így csökkenti az egyes paraméterek számát [32].

#### 4.5 YOLO – You Only Look Once

Az egyik legismertebb képfeldolgozó algoritmus a YOLO, amely konvolúciós hálózatokat alkalmazva képes megtanulni objektumokat és felismerni őket.

2015-ben készült el az első verziója. Elsődleges cél az objektumok gyors, valós időben történő lokalizálása volt. Nevét onnan kapta, hogy a rendszernek elég egyszer látnia a bemeneti képet ahhoz, hogy azon az összes objektumot észlelje [33].

A YOLO bizonyos nagyságú négyzetrácsokkal osztja fel a bemeneti képeket, és meghatározza, hogy az adott rács, mely objektumhoz tartozik a legnagyobb valószínűséggel. Ezt úgynevezett Intersection Over Union (IOU) nevű dobozok segítségével valósítja meg, amely az általunk elkészített címke, amely az objektumot határoló téglalapot jelenti, illetve az algoritmus által prediktált befoglaló geometria metszetének, valamint a két téglalap uniójának hányadosából lehet kiszámolni [33]. Ennek képlete a következőképpen néz ki tehát:

$$IOU = \frac{doboz_1 \cap doboz_2}{doboz_1 \cup doboz_2}$$

Az adott téglalap tartalmaz egy megbízhatósági értéket is, amely megmondja, hogy a határoló mekkora valószínűséggel tartalmazza az adott objektumot:

$$Conf = P(objektum) \times IOU$$

Természetesen lehetnek az egyes IOU-k között is átfedések, ezt hivatott kivédeni a YOLO a Non Maximal Supression nevű eljárásával [33]. Ennek lényege, hogy a több cellán átnyúló objektumokra nagy valószínűséggel több találatot ad az algoritmus és valahogyan ezen találatok számát csökkenteni kell. A módszer lényege, hogy a meghatározott megbízhatósági érték alá eső találatokat a rendszer kiveszi, majd a legnagyobb értékkel rendelkezőt pedig valódi találatnak minősíti. Ezt követően törölődnek az átfedések is és ez addig ismétlődik, amíg a rendszer nem talált meg minden valódi objektumot [33]. A szoftverben alkalmazott YOLO verzió a 7-es, amelynek elsődleges célja a nagyobb FPS szám elérése még pontosabb modell mellett. Ez a YOLO úgynevezett modell újraparaméterezés, illetve modell skálázás eljárást alkalmaz és tartalmazza már a fontosabb kulcspontjait a felismert emberi alakzatnak. A YOLO alapvető, előre betanított adathalmazát, a COCO adathalmazt fogjuk alkalmazni, hiszen emberi alakzat felismerésére ez megfelelő [34].

#### **4.6. MediaPipe**

A MediaPipe a Google által fejlesztett keretrendszer, amelyet elsődlegesen gépi látással foglalkozó hallgatók, kutatók és fejlesztők számára készítettek el. Erősen támaszkodik a gépi tanulásra, lehetővé téve applikációk fejlesztését, és különböző környezetérzékeléssel kapcsolatos kutatómunkák elvégzését [35]. A MediaPipe inkrementális fejlesztési lehetőséget kínál a felhasználóinak. Ezt úgynevezett kalkulátorok segítségével teszi lehetővé, amelyekre irányított gráfokként kell tekinteni. A gráfok egy protokoll bufferből és egy gráf objektumból állnak. A gráfban ezek a kalkulátorok adat adatfolyamok segítségével kapcsolódnak egymáshoz. Minden ilyen folyam adatcsomagok egy idősorát reprezentálja. A kalkulátorok és folyamok együtt alkotják az adatfolyam gráfot. A csomagok, amik a gráfon keresztül folynak, az időbélyegeik alapján összevetésre kerülnek egy idősoron belül.

#### 4.7 Yolo versus Mediapipe

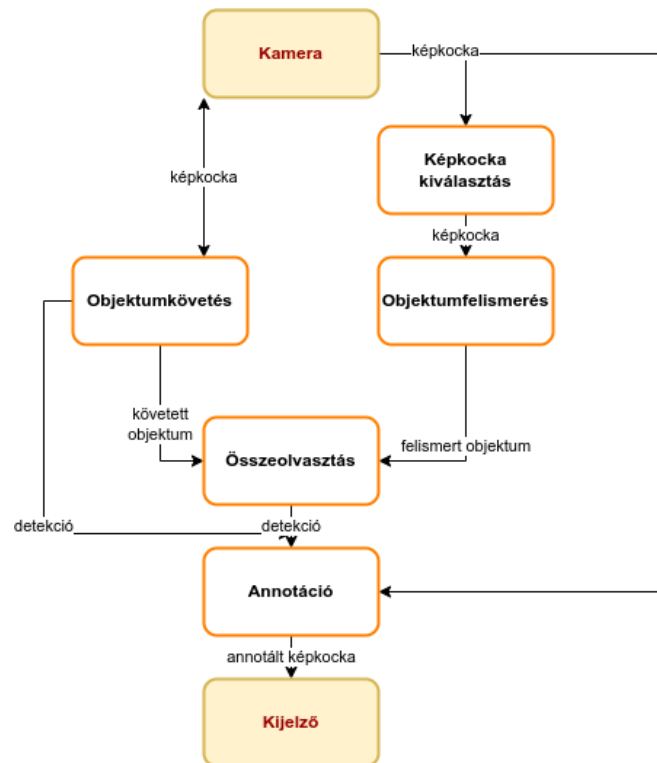
A MediaPipe felhasználásban jelentősen eltér a YOLOv7-től. Ennek egyik fő oka, hogy a két algoritmus fő felhasználási területe is eltérő. A YOLO elsősorban minél több objektum valós időben történő felismerésével foglalkozik, és inkább foglalkozik azzal a kérdéssel, hogy mi a felismert objektum. Nem követ nyomon a felismert alakzatokat, és nincs semmilyen szegmentálási módszer beleintegrálva. A MediaPipe ezzel szemben egy adott objektum részletes leírását teszi lehetővé, és célja, hogy több platformon is stabil sebességgel működjön. Ennek megfelelően nem támaszkodik a GPU-ra. A felismerés után az alakzat lekövetése automatikusan meg van oldva és különböző szegmentálási eljárásokat is megvalósít az alakzaton belül. A főbb különbségeket a két algoritmus között az 2. táblázat tartalmazza.

2. táblázat: A MediaPipe és YOLOv7 összehasonlítása (saját szerkesztés)

Jellemző	YOLOv7	MediaPipe
Topológia	17 kulcspon <b>COCO</b>	33 <b>COCO</b> kulcspon + <b>Blaze Palm + Blaze Face</b>
Folyamat	Minden képkockában detektál	Detektálás egyszer, utána "nyomkövetés"
GPU támogatás	VAN	Csak CPU
Szegmentáció	Közvetlenül a "pose" adatba nincs beintegrálva	Beintegrálva
Emberek száma	Multi-person, vagyis több alakzat felismerésére alkalmas	Egyszerre egy embert lát

Az adott feladatra a MediaPipe objektumfelismerő algoritmusát választottuk ki. Így a tervezett implementáció a 9. ábra szerinti keretrendszert alkalmazza, melynek felépítése a következő:

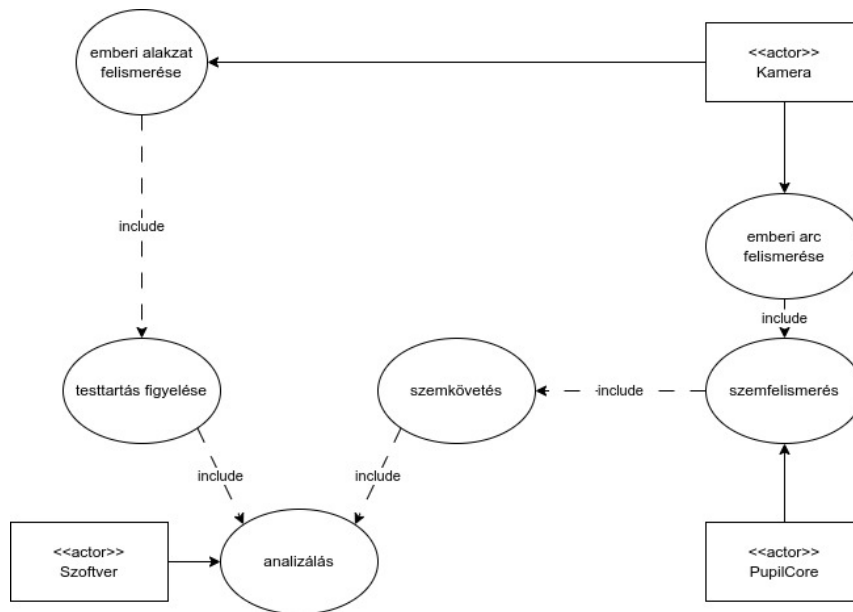
- Objektumfelismerés és követés egy időben
- Újonnan felismert alakzat megfeleltetése korábbi objektumnak
- Nyomkövetett és felismert alakzat „összeolvasztása”
- Kimenet: annotált kép a kijelzőn



9. ábra: A Mediapipe keretrendszer felépítése [35]

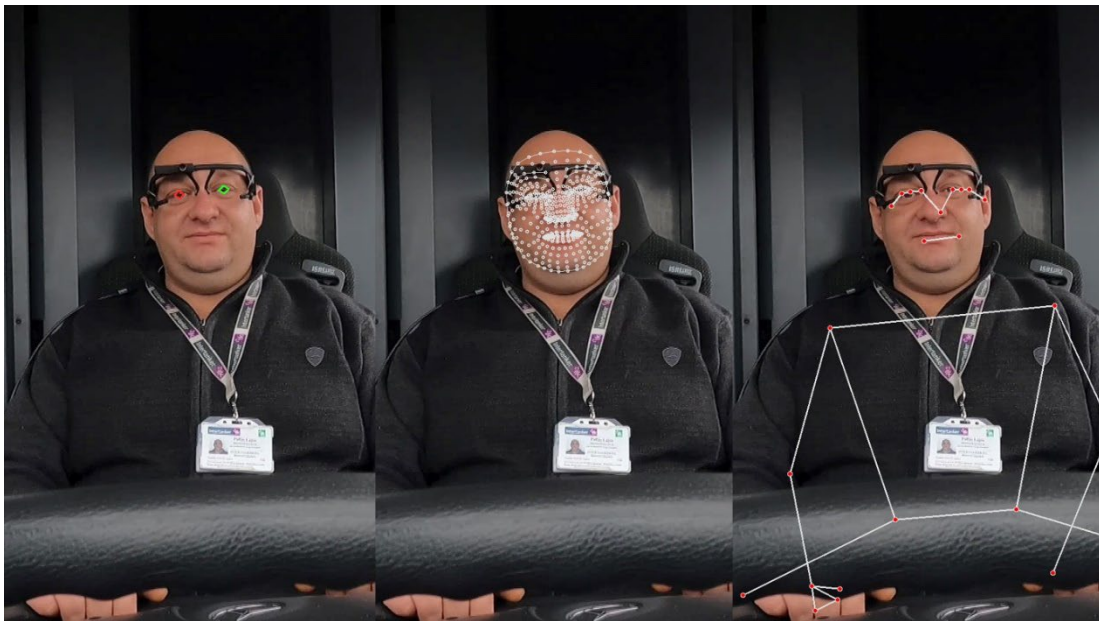
## 5. Összegzés

A 10. ábra összefoglalja a szoftver rendszer Use Case felépítését. Ebben a rendszerben a Pupil Core szemmozgás figyelő, mint kísérleti eszköz szerepel, a komplex adatgyűjtési cél elérése érdekében. Azonban a járművezetőt nem zavaró, fedélzetre telepített, kamera alapú rendszer részét már nem képezi. A javasolt hardver és szoftver összeállítás alkalmas a felsőtest, fej, arc egyes pontjainak és korlátozottan a szemmozgás detektálására.



**10. ábra:** A szoftver Use Case diagramja (saját szerkesztés)

A teszt felvételek során a buszvezető detektálása pilot eredményként megtörtént (11. ábra).



**11. ábra:** A pilot képfeldolgozás eredményei (szembogár felismerése; arcdetektálás; arc és test kulcsponatok felismerése) (saját szerkesztés)

Különösen fontos a tekintet figyelő elem tesztelése, mely magában foglalja fej irány becslését a tekintet vektorok vizsgálatával. A tesztelés során a rendszer megfelelően ismerte fel és hirdette az irány (12. ábra).

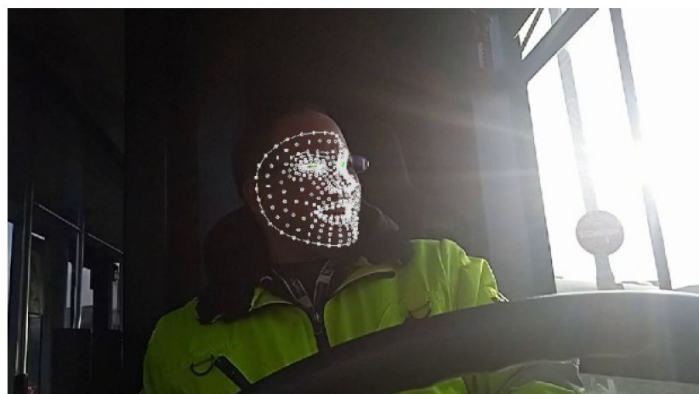


a)

b)

**12. ábra:** A tekintetfigyelő működése. a): a sofőr balra tekint; b): a sofőr egyenest néz (saját szerkesztés)

A gyakorlati alkalmazás során a zavaró hatások (lencse szennyeződése, környezeti fény becsillanás, és egyéb optikai esetek; sofőr szemüveges) is vizsgálatra kerültek. Egyes esetekben a kísérleti kép-halmazon végzett detektálás hibás eredményeket hozott, de kizárólag a szélsőségesen erős és kritikus szögben beeső környezeti fény (napsütés) hozott néhány frame szinten kimutatható hibát (13. ábra).



**13. ábra:** Az erős környezeti fény kritikus esetben megzavarhatja az algoritmus működését (saját szerkesztés)

A tesztelés során az igényként megfogalmazott monitoring feladatokat, mint célokat sikerült bizonyíthatóan elérni, azonban további fejlesztést igényel a rendszer integrációja, valamint a real-time adatfeldolgozás kompakt hardver eszközök segítségével. A jövőbeni feladatok között a rendszer további tesztelése, ROS2 rendszer felépítése, és valós körülmények között történő kipróbálása fog megtörténni.

## 6. Köszönetnyilvánítás

A kutatás a Kooperatív Doktori Program (KDP) keretén belül az Kulturális és Innovációs Minisztérium, valamint a Nemzeti Kutatási, Fejlesztési és Innovációs Alap támogatásával valósul meg.

## Irodalomjegyzék

- [1.] Blades, L.; Douglas, R.; Early, J.; Lo, C.Y.; Best, R. Advanced Driver-Assistance Systems for City Bus Applications. *SAE Technical Papers* 2020, 2020-April, doi:10.4271/2020-01-1208.
- [2.] Nylén, A.B.; Reyes, M.L.; Roe, C.A.; McGehee, D. V. Impacts on Driver Perceptions in Initial Exposure to ADAS Technologies. *Transp Res Rec* 2019, doi:10.1177/0361198119847975.
- [3.] Slootmans, F. European Road Safety Observatory - Facts and Figures - Light Trucks. 2021.
- [4.] Eurostat Passenger Transport by Buses and Coaches by Type of Transport - Vehicles Registered in the Reporting Country. 30/03/2023 23:00 2023.
- [5.] Goh, K.; Currie, G.; Sarvi, M.; Logan, D. Factors Affecting the Probability of Bus Drivers Being At-Fault in Bus-Involved Accidents. *Accid Anal Prev* 2014, 66, 20–26, doi:10.1016/j.aap.2013.12.022.
- [6.] Young; Regan; Hammer Driver Distraction: A Review of the Literature (Report). Monash University Accident Research Centre 2003, 66.
- [7.] Otmani, S.; Rogé, J.; Muzet, A. Sleepiness in Professional Drivers: Effect of Age and Time of Day. *Accid Anal Prev* 2005, 37, 930–937, doi:10.1016/j.aap.2005.04.011.
- [8.] Huhta, R.; Hirvonen, K.; Partinen, M. Prevalence of Sleep Apnea and Daytime Sleepiness in Professional Truck Drivers. *Sleep Med* 2021, 81, 136–143, doi:10.1016/j.sleep.2021.02.023.

- [9.] Felix, M.; Intriago Alvarez, M.B.; Vanegas, E.; Farfán Bajaña, M.J.; Sarfraz, Z.; Sarfraz, A.; Michel, J.; Cherrez-Ojeda, I. Risk of Obstructive Sleep Apnea and Traffic Accidents among Male Bus Drivers in Ecuador: Is There a Significant Relationship? *Annals of Medicine and Surgery* 2022, 74, doi:10.1016/j.amsu.2022.103296.
- [10.] Thiffault, P.; Bergeron, J. Monotony of Road Environment and Driver Fatigue: A Simulator Study. *Accid Anal Prev* 2003, 35, 381–391, doi:10.1016/S0001-4575(02)00014-3.
- [11.] Sahayadhas, A.; Sundaraj, K.; Murugappan, M. Detecting Driver Drowsiness Based on Sensors: A Re-view. *Sensors (Switzerland)* 2012, 12, 16937–16953, doi:10.3390/s121216937.
- [12.] Hallac, D.; Sharang, A.; Stahlmann, R.; Lamprecht, A.; Huber, M.; Roehder, M.; Sosič, R.; Leskovec, J. Driver Identification Using Automobile Sensor Data from a Single Turn. *IEEE Conference on Intelligent Transportation Systems, Proceedings, TSC* 2016, 953–958, doi:10.1109/ITSC.2016.7795670.
- [13.] Campos-Ferreira, A.E.; Lozoya-Santos, J. de J.; Tudon-Martinez, J.C.; Mendoza, R.A.R.; Var-gas-Martínez, A.; Morales-Menendez, R.; Lozano, D. Vehicle and Driver Monitoring System Using On-Board and Remote Sensors. *Sensors* 2023, 23, doi:10.3390/s23020814.
- [14.] Maslač, M.; Antić, B.; Lipovac, K.; Pešić, D.; Milutinović, N. Behaviours of Drivers in Serbia: Non-Professional versus Professional Drivers. *Transp Res Part F Traffic Psychol Behav* 2018, 52, 101–111, doi:10.1016/j.trf.2017.11.020.
- [15.] Fancello, G.; Daga, M.; Serra, P.; Fadda, P.; Pau, M.; Arippa, F.; Medda, A. An Experimental Analysis on Driving Behaviour for Professional Bus Drivers. *Transportation Research Procedia* 2020, 45, 779–786, doi:10.1016/j.trpro.2020.02.095.
- [16.] Karimi, S.; Aghabayk, K.; Moridpour, S. Impact of Driving Style, Behaviour and Anger on Crash In-volvement among Iranian Intercity Bus Drivers. *IATSS Research* 2022, 46, 457–466, doi:10.1016/j.iatssr.2022.07.003.
- [17.] Sigari, M.-H.; Pourshahabi, M.-R.; Soryani, M.; Fathy, M. A Review on Driver Face Monitoring Sys-tems for Fatigue and Distraction Detection. *International Journal of Advanced Science and Technology* 2014, 64, 73–100, doi:10.14257/ijast.2014.64.07.

- [18.] Biondi, F.; Coleman, J.R.; Cooper, J.M.; Strayer, D.L. Average Heart Rate for Driver Monitoring Sys-tems. *Int J Hum Factors Ergon* 2016, 4, 282–291, doi:10.1504/IJHFE.2016.083521.
- [19.] Fujiwara, K.; Abe, E.; Kamata, K.; Nakayama, C.; Suzuki, Y.; Yamakawa, T.; Hiraoka, T.; Kano, M.; Sumi, Y.; Masuda, F.; et al. Heart Rate Variability-Based Driver Drowsiness Detection and Its Validati-on With EEG. *IEEE Trans Biomed Eng* 2019, 66, 1769–1778, doi:10.1109/TBME.2018.2879346.
- [20.] Dehzangi, O.; Rajendra, V.; Taherisadr, M. Wearable Driver Distraction Identification On-the-Road via Continuous Decomposition of Galvanic Skin Responses. *Sensors (Switzerland)* 2018, 18, 1–16, doi:10.3390/s18020503.
- [21.] Li, G.; Chung, W.Y. A Context-Aware EEG Headset System for Early Detection of Driver Drowsiness. *Sensors (Switzerland)* 2015, 15, 20873–20893, doi:10.3390/s150820873.
- [22.] Balam, V.P.; Chinara, S. Development of Single-Channel Electroencephalography Signal Analysis Model for Real-Time Drowsiness Detection: SEEGDD. *Phys Eng Sci Med* 2021, 44, 713–726, doi:10.1007/s13246-021-01020-3.
- [23.] Rahman, N.A.A.; Mustafa, M.; Sulaiman, N.; Samad, R.; Abdullah, N.R.H. EMG Signal Segmentation to Predict Driver’s Vigilance State. *Lecture Notes in Mechanical Engineering* 2022, 29–42, doi:10.1007/978-981-16-4115-2\_3.
- [24.] European Parliament REGULATION (EU) 2019/2144 Of the European Parliament and of the Council. *Official Journal of the European Union* 2019.
- [25.] Koay, H.V.; Chuah, J.H.; Chow, C.O.; Chang, Y.L. Detecting and Recognizing Driver Distraction through Various Data Modality Using Machine Learning: A Review, Recent Advances, Simplified Framework and Open Challenges (2014–2021). *Eng Appl Artif Intell* 2022, 115, 105309, doi:10.1016/j.engappai.2022.105309.
- [26.] Chaves, D.; Fidalgo, E.; Alegre, E.; Alaiz-Rodríguez, R.; Jáñez-Martino, F.; Azzopardi, G. Assessment and Estimation of Face Detection Performance Based on Deep Learning for Forensic Applications. *Sensors (Switzerland)* 2020, 20, 1–21, doi:10.3390/s20164491.
- [27.] P. Jimenez; et al., Gaze fixation system for the evaluation of driver distractions induced by Ivis, *IEEE Transactions on Intelligent Transportation Systems*, 2012, 13(3), pp. 1167–1178.

- [28.] T. D'Orazio, M. Leo, G. Cicirelli and A. Distanto, An algorithm for real time eye detection in face images, *Proceedings of the 17th International Conference on Pattern Recognition*, 2004. ICPR 2004., Cambridge, UK, 2004, pp. 278-281 Vol.3, doi: 10.1109/ICPR.2004.1334521.
- [29.] <https://pupil-labs.com> (2023.04.27.)
- [30.] A. Nischwitz; M. Fischer; P. Haberäcker; G. Socher, *Computergrafik und Bildverarbeitung Band I: Computergrafik 3.*, neu bearbeitete Auflage, p40 2011
- [31.] <http://docs.ros.org/en/foxy/index.html> (2023. 07. 16.)
- [32.] K. Kranthi Kumar; M. Dileep Kumar; Ch. Samsonu; K. Vamshi Krishna, Role of convolutional neural networks for any real time image classification, recognition and analysis. *Materials Today: Proceedings*, 2021
- [33.] J. Redmon; S. Divvala; R. Girshick, A. Farhadi, You Only Look Once: Unified, Real-Time Object Detection, *arXiv preprint arXiv: 1506.02640*, 2015
- [34.] <http://images.cocodataset.org/zips/train2017.zip> (2023.05.08.)
- [35.] C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M. Hays, F. Zhang, Ch. Chang, M. G. Yong, J. Lee, W. Chang, W. Hua, M. Georg; M. Grundmann, MediaPipe: A Framework for Building Perception Pipelines *arXiv:1906.08172*, 2019

# ESG mérőszámok vállalati gyakorlatai és közelítése a technológiaelfogadás-modellekhez

## Corporate practices and alignment of ESG metrics with technology acceptance models

Pintes Orsolya<sup>a</sup>, Suta Alex<sup>b</sup>, Tóth Árpád<sup>c</sup>

<sup>a</sup>Széchenyi István Egyetem - Járműipari Kutatóközpont

pintes.orsolya@ga.sze.hu

<sup>b</sup>Széchenyi István Egyetem - Járműipari Kutatóközpont

suta.alex@ga.sze.hu

<sup>c</sup>Széchenyi István Egyetem - Járműipari Kutatóközpont

totha@ga.sze.hu

### Absztrakt

---

*Az Egyesült Nemzetek Szövetsége (ENSZ) fenntarthatósági céljait kategorizáló Economical, Social, Governance (ESG) rendszer napjaink egyik vitatott témája. Az ESG-mérőszámok fontos tényezők a vállalkozások fenntarthatóságának és társadalmi hatásának értékelésében. Az éghajlatváltozás, a társadalmi egyenlőtlenség és az etikai gyakorlatok iránti aggodalom növekedésével a befektetők, érdekelt felek és fogyasztók egyre inkább felismerik az ESG-tényezők fontosságát a vállalatok teljesítményének és hosszú távú életképességének értékelésében. A tanulmány célja átfogó elemzést nyújtani az ESG-mérőszámokról és azok jelentőségéről az üzleti életben, valamint konkrét mérőszámokat találni az egyes kategóriákra, amelyek alapján nyomon lehet követni egy vállalat fenntarthatósági tevékenységét. Vizsgálja a környezeti, társadalmi és irányítási mutatókat, valamint az ESG-szabványok és keretrendszerek fejlődését. Körüljárja az ESG-mérőszámok elfogadásának motivációit a befektetők és fogyasztók szemszögéből, valamint bemutatja az ESG-gyakorlatok előnyeit a vállalatok számára. Az ESG-szemponatok beépítése a vállalati stratégiába elengedhetetlen a pozitív változások előmozdítása és a fenntarthatósági célok elérése szempontjából. Munkánk hozzájárul az ESG-mérőszámok hatékony eszközként történő megértéséhez a vállalatok fenntarthatóságának és hosszú távú értékteremtésének kiértékelésére.*

**Kulcsszavak:** ESG, fenntarthatósági mutatószámok, vállalati gyakorlat, beszámolóképzés

## Abstract

---

*The Economic, Social, Governance (ESG) framework, which categorizes the United Nations (UN) sustainability goals, is a highly discussed topic today. ESG metrics play a significant role in evaluating the sustainability and social impact of businesses. With increasing concerns regarding climate change, social inequality, and ethical practices, investors, stakeholders, and consumers are increasingly acknowledging the significance of ESG metrics in evaluating the performance and long-term viability of companies. The objective of this study is to present a detailed analysis of ESG metrics and their relevance to business, while identifying particular metrics for each category that can be employed to measure a company's sustainability performance. It analyzes environmental, social, and governance indicators along with the progression of ESG standards and frameworks. It investigates the incentives for implementing ESG measurement from the viewpoints of investors and consumers while highlighting the advantages of ESG practices for corporations. Our research enhances comprehension of ESG metrics as a powerful resource for assessing sustainability and generating long-term value for companies. Integrating ESG considerations into corporate strategy is crucial to promote positive change and attain sustainability objectives and reduce carbon footprint.*

**Keywords:** *ESG, sustainability metrics, corporate practice, reporting*

## 1. Bevezetés

Az emberi tevékenység egyre nagyobb hatást gyakorol az éghajlatváltozásra, ezért a szabályozó állami, és szakmai szervezetek is nagyobb figyelmet fordítanak erre a globális problémára, valamint a nettó szén-dioxid-kibocsátás nullára csökkentésére (ACCA & Adam Smith Business School, 2023). Emellett a befektetők is növekvő nyomást gyakorolnak a vállalatokra, hogy azonosítsák és kezeljék a hosszú távú fenntarthatósági kockázatokat és lehetőségeket.

Az elmúlt években világszerte megnőtt a fenntarthatósággal kapcsolatos jelentéstételi keretrendszerek száma. Ezek a keretrendszerek különböző fókuszú és célkitűzéseket tartalmaznak, és különböző szabványalkotók vagy globális testületek által kerülnek létrehozásra. Ennek eredményeként több, hasonló célt szolgáló, de eltérő követelményekkel rendelkező keretrendszer létezik párhuzamosan.

Az említett keretrendszerek nem mindig kapcsolják össze a fenntarthatósági információkat a pénzügyi információkkal, ezért nem minden esetben tekinthetőek relevánsnak a piacok számára mint például a CRM rendszerek (Reicher, 2016). Ugyanakkor a globális gazdaság több szereplője támogatja a következetes és összehasonlítható keretrendszer létrehozását, amely alapján a szervezetek meghatározhatják saját követelményeiket. A fenntarthatósággal kapcsolatos pénzügyi jelentéseket az iparágak és joghatóságok között megoszthatják, ami hasznos információkat nyújt a piaci döntésekhez. A fenntarthatósági teljesítmény mérésében és a döntéshozatal támogatásában az életciklus-elemzés is egy fontos eszköz lehet (Eisinger et al., 2022).

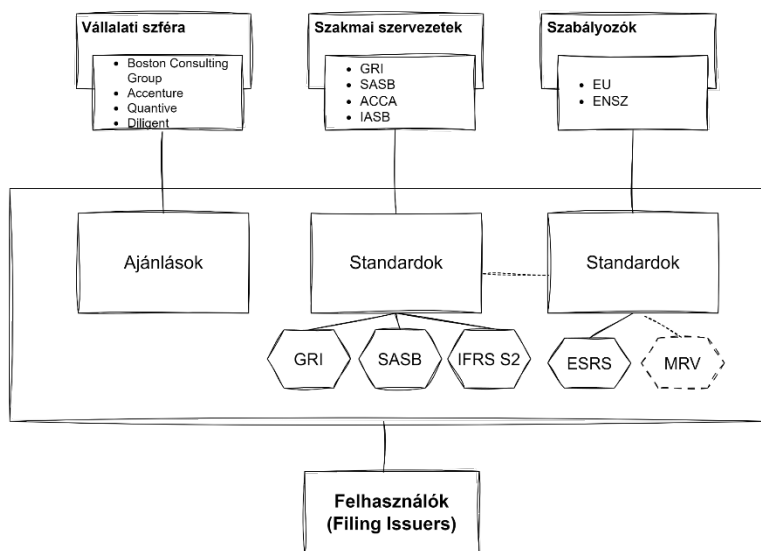
A kutatás elméleti alapja sajnos nem rendelkezik szilárd, előre meghatározott háttérrel, így az általunk készített összevetés elméleti alapon nem előre megalapozott. Fontos lenne tehát egy egységes, átfogó keretrendszer kialakítása, amely összekapcsolja a fenntarthatósági és pénzügyi információkat, hogy a vállalatok és a befektetők számára könnyebb legyen értékelni a fenntarthatósági teljesítményt és kockázatokat. Ez segítene a piacoknak és az érdekelteknek a döntéshozatalban.

## 2. ESG mérőszámok közzétételének motivációi

Az ESG-rendszer egy befektetési döntésekbe integrálható szempontrendszer, amely egy adott vállalkozásnál vagy országnál a környezetre (Environmental), a társadalomra (Social) és az irányításra (Governance) vonatkozó hatásokat és faktorokat értékeli fenntarthatósági szempontból (Diligent, 2023). Az ESG jelentés fontos szempont a vállalatok számára több tényező miatt:

- A kötelezettségvállalások kézzelfoghatósága, konkretizálása
- Átláthatóság növelése
- Tőkéhez (zöld pénzügyek) való hozzáférés
- Jobb márkaképviselés
- Kevesebb szabályozási kockázat

Az 1. ábrán feltüntetésre kerültek a különböző iránymutatások és szakmai tanácsadók elemeire bontva aszerint, hogy melyik szférából származnak és milyen háttérrel rendelkeznek.



1. ábra: Az ESG beszámolást befolyásoló források (Saját szerkesztés)

A választott keretrendszerrel függően akár több mint 20 ESG-mérőszám is jelen van a gyakorlatban, amelyre kivetítve optimalizálni kell a vállalat működési elveit. A szokásos üzleti kötelezettségek mellett kihívást jelenthet egy sikeresen végrehajtható ESG-stratégia kidolgozása. Ezt szem előtt tartva hasznos lehet olyan keretrendszerrel találkozni, amely összehangolja a stratégiát és a végrehajtást, hogy az egyéb üzleti célok mellett az ESG-kötelezettségvállalásokat is teljesítse.

## **2.1. Európai uniós szabályozási környezet**

A nemzetek feletti irányító szervezetek, mint az Európai Unió vagy az Egyesült Nemzetek Szervezete standardok meghatározásával nyújt a vállalatoknak iránymutatást, amelyek felhasználásával a fenntartható működés helyes monitorozása és prezentációja megvalósítható. A standardok között kiemelhető a European Sustainability Reporting Standards (ESRS) véglegesítés előtt áll (draft) verziója, valamint a Measurement, Reporting and Verification (MRV) elve. Az MRV a környezetvédelmi és fenntarthatósági menedzsmentben széles körben használt fogalom (UNFCCC, 2016), a környezeti teljesítmény és a fenntarthatósági mutatók szisztematikus mérésére, jelentésére és ellenőrzésére összpontosít. Az ESG-mérőszámok és a fenntarthatósági célok esetében az MRV alapvető fontosságú a megbízható és átlátható méréshez és jelentéstételhez, amely lehetővé teszi a hatékony döntéshozatalt, valamint elősegíti az elszámoltathatóságot és az átláthatóságot.

Az European Financial Reporting Agency (EFRAG) európai fenntarthatósági jelentéstételi standardokkal foglalkozó munkacsoportja (EFRAG PTF-ESRS) által a 2021. júniusa és 2022. áprilisa közötti időszakban készített ESRS-tervezetek (Exposure Draft, ED) április 30. és 2022. augusztus 8. között álltak rendelkezésre észrevételek benyújtására. Az EFRAG Fenntarthatósági Beszámolási Testület (EFRAG SRB), az EFRAG Fenntarthatósági Beszámolási Szakértői Csoport (EFRAG SR TEG) tanácsára, foglalkozott a konzultáció során kapott visszajelzésekkel, és ennek megfelelően módosította a tizenkét ESRS-tervezetet, amelyeket napjainkban adtak át az Európai Bizottságnak a döntéselőkészítésre.

Az Európai Bizottság konzultálni fog az uniós szervekkel és a tagállamokkal a standard-tervezetekről, mielőtt a végleges standardokat várhatóan 2023. júniusában, felhatalmazáson alapuló jogi aktusként elfogadja, majd ezt követően az Európai Parlament és a Tanács ellenőrzési időszakot tart (EFRAG, 2023).

## **2.2. Vállalati iránymutatások és ajánlások**

A vállalati szféra egyes elemei kifejezetten arra specializálódtak, hogy más vállalatoknak támogassanak a fenntarthatósági felelősségvállalásokat illetően. Ilyen szolgáltatást nyújt például a Quantive, a Boston Consulting Group vagy az Accenture. A példaként hozott vállalatok nem egy egységes keretrendszer alapján hozzák létre tervezeteiket, az általuk kínált tanácsadás, bár segíthet a vállalati szféra tagjainak fenntarthatósági tevékenységük kialakításában, nem járul hozzá az egységes skálán mérhető rendszer kialakulásához.

## **2.3. Szakmai szervezetek kezdeményezései**

Ide tartoznak azok a szervezetek, amelyek az EU vagy az ENSZ iránymutatása mentén igyekeznek megfogalmazni és létrehozni egy egységes rendszert, amely kiterjed minden vállalat tevékenységére. Ilyen az International Accounting Sustainability Board (IASB), a Global Reporting Initiation (GRI), az Sustainability Accounting Standards Board (SASB), vagy hasonló az Association of Chartered Certified Accountants (ACCA). Egy jó példa a keretrendszerek szemléltetésének céljából a nemzetközi számviteli standard, az IFRS S2 Exposure draftjának egy szekciója, amely pontos metrikák szerint határozza meg, hogy a kutatásban szereplő vállalatok közül melyik felel meg a megadott követelményeknek (ACCA & Adam Smith Business School, 2022). A standardben szerepel ezen kívül 11 olyan számszerűsíthető mérőszám, amely iparágtól függetlenül releváns minden vállalat tevékenységének számára. Néhány példát említve: a vállalatok által kibocsátott CO2 mennyiségét pontosan, tonnában fel kell tüntetni az éves vállalati jelentésekben. A kibocsátásokat kategorizálni kell előre meghatározott csoportok alapján: mekkora mennyiségű káros anyag keletkezik a gyártás, a szállítás vagy az értékesítés során. A pénzügyi forrásokat, amelyeket fenntartható tevékenységekre különít el a vállalat, szintén pontosan meg kell határozni.

## **3. Módszertan**

A tanulmány célja volt, hogy megvizsgálja azokat a jelenleg elérhető forrásokat, amelyek hozzájárulnak, hogy egy egységes keretrendszer létrejöjjön a vállalati “zöld” tevékenységek vizsgálatára, mérésére és kategorizálására. A már meglévő standardokat és keretrendszer-kezdeményezéseket vállalati adatlapokon és a Scopus adatbázisban szereplő cikkek alapján

tanulmányoztuk, majd összegzést készítettünk, amelyben összehasonlítottuk a keretrendszerek fő fókuszpontját, és a mérési technikájukat. Ezen felül fontos lépés volt a munkamenet során a számszerűsíthető közzétételek összegzése, majd ezek alapján az egyezések és eltérések vizsgálata. A számok és metrikák után az elméleti háttérrel vizsgáltuk a keretrendszerek mögött, tehát, hogy milyen kutatásra vagy problémára alapozzák a keretrendszer létrehozását. A legtöbb ilyen rendszer mögött azonban komoly elméleti megalapozás nem elérhető vagy megnevezett, ezért szakirodalmi áttekintés alapján különböző releváns elméleti keretrendszereket jelöltünk ki, mint alkalmazható elméleti megalapozást elősegítő módszertant. Ezek alapján a kutatás nem egy szisztematikus áttekintés gondolatával készült, sokkal inkább egy átfogó elemzése a vállalati, szakmai és általános irányítási szervezeti forrásoknak.

## **4. Eredmények**

### **4.1. Globális szabályozók**

Az ENSZ és az EU által megfogalmazott ajánlásoknak és iránymutatások ténylegesen nagy esélyük lenne bekerülni a globális szintű szabályrendszerbe, ezáltal pedig tartós részét képezni a vállalatok által követendő sémának. Az ESRS pontos lépéseket határoz meg az ESG-rendszer minden kategóriáján belül, szem előtt tartva az egyes szegmensek leglényegesebb célkitűzéseit. Viszont addig, amíg ezek használatát nem teszik kötelezővé, nem kaphatnak akkora figyelmet, amely a széles körű elterjedtséghez, valamint a globális változások eléréséhez szükséges.

### **4.2. Szakmai szervezetek: ESG standardok**

A nemzetközi irányítási- és szabályozó szervezetek mellett sok független szakmai szervezet is komolyan foglalkozik a nettó nulla kibocsátás felé való elmozdulás mozgalmával, ők szintén szeretnék irányt mutatni és tanácsot adni a vállalatoknak. Több ilyen szakmai szervezet, mint például az ACCA vagy az IASB bocsájtanak ki keretrendszer-tervezeteket és követendő sémákat. Ezekben gyakran megtalálhatók a szükséges kategorizálások és az egyes rendeleteknek pontos mérőszámokhoz való kapcsolása, viszont mivel az ilyen rendszerekből egyszerre több létezik és kerül használatba párhuzamosan, nem alkotnak egy egyetemes rendszert, amely alapján az összes vállalatot mérni és számszerűsíteni lehet. Kapcsolódó

probléma az is, hogy az ESG-lebontásnak sokféle értelmezése született az évek során, és ezért is különböznek annyira a szakmai szervezetek által készített keretrendszerek.

### **4.3. Vállalati ajánlások**

A fent említett szakmai szervek és a globális irányítók mellett a vállalati szférában is többen szakosodtak arra, hogy irányt mutassanak a további cégeknek, hogyan tehetik tevékenységüket fenntarthatóbbá hosszú távon is. Ide sorolható többek között a Boston Consulting Group, az Accenture, a Quantive vagy a Diligent. A BCG közreműködésével támogatást nyújt bankoknak, vagyonkezelőknek és más pénzügyi intézményeknek versenyelőnyt teremteni a fenntartható finanszírozás és befektetés területén (BCG, 2023). Fő tevékenységeik körébe tartoznak a:

- ESG- és fenntartható finanszírozási célok meghatározása, teljesítményértékelés és forgatókönyvtervezés, valamint ESG- és fenntartható finanszírozási ütemtervek végrehajtása.
- Társadalmi és fenntartható pénzügyi keretrendszerek és fenntartható pénzügyi stratégiák kidolgozása.
- A környezeti és társadalmi kategóriákban a piaci részesedést megragadó, nagy hatású fenntartható befektetési lehetőségek azonosítása és követése.

A Quantive egy konkrét ESG-lebontást készített, amelyben az ESG mindhárom kategóriájának kulcsfontosságú mutatói szerepelnek, magyarázattal és ajánlott mértékegységgel vagy javasolt mérési móddal ellátva (Quantive, 2023). Ezek közül a környezeti mérőszámok átláthatóak és logikailag mérhetőek, azonban a probléma a másik két kategória mutatóinál egyaránt fennáll: a javasolt metrika mérésére nincs egységes és jól átlátható rendszer. Emiatt a vállalati iránymutató szervezetek ajánlásai hosszú távon nem alkalmazhatók ezen kategóriák esetében.

### **4.4. Tudományos háttér**

#### **4.4.1. Technológiaelfogadási modellek**

Fontos, hogy az újonnan létrehozott és bevezetett metrikus rendszert minden vállalat egységesen elfogadja és használja. Ha egy elméleti szempontból megalapozott rendszer állna a különböző keretek mögött, amelyben el lehet helyezni a különböző taxonómiákat, és

metrikákat, a vállalatok számára felismerhetőbb és áttekinthetőbb lenne a követelések rendszere. A cikk következő részében különböző modelleket tekintettünk át, amelyek hozzájárulhatnak a két megjelölt alfejezet elméleti keretrendszerének kialakításához. Ezáltal öt olyan modellt vizsgáltunk, amelyek alkalmasak lehetnek arra, hogy a jelentéstételi keretrendszer elméleti megalapozására használják őket.

A Technology Acceptance Model (TAM) modellezi, hogy a felhasználók miként reagálnak egy-egy új technológiára (Charness & Boot, 2016). Kihangsúlyozza az észlelt hasznosság (*Perceived usefulness*) és az észlelt könnyű használhatóság (*Perceived ease-of-use*) hatását az egyéneknek a technológia használatára vonatkozó szándékával kapcsolatban. Az ESG-mérőszámok és a fenntarthatósági keretrendszerek összefüggésében a TAM segíthet megérteni a kapcsolódó technológiák elfogadását és bevezetését.

A Technology Organization Environment (TOE) a technológiai innovációk eseteiben adaptálható, azonosítja az innovációk elfogadásának és bevezetésének folyamatait (*ISRI, 2023*). Hasznos és adaptív keretrendszer a technológiai innovációk három típusa esetében, melyek: a műszaki feladatokra alkalmazott, valamint az üzleti adminisztrációra vonatkozó innovációk, és a szervezet alapvető üzleti folyamataiba ágyazott innovációk esetében az átvételi magatartás magyarázata (Ramdani & Kawalek, 2007; Swanson, 1994). A TOE-keretrendszer a szervezet kontextusának három olyan aspektusát azonosítja, amelyek befolyásolják a technológiai innováció elfogadási és bevezetési folyamatát, nevezetesen a technológiai, a szervezeti, és a külső környezeti kontextust.

A TOE-keretnek azonban vannak bizonyos korlátai is (TheoryHub, 2023). Az egyik, hogy túlságosan tág értelmezésű és általános lehet, ami megnehezíti a konkrét összefüggésekben való alkalmazását. Továbbá a keretrendszer nem feltétlenül ragadja meg teljes mértékben a technológia átvételének és bevezetésének összetettségét, különösen a gyorsan változó környezetekben, ahol a külső tényezők jelentős hatással lehetnek a technológiai döntésekre.

A TOE alkalmazható annak vizsgálatára, hogy a technológiai tényezők (pl. az ESG-adatok nyomon követésére szolgáló digitális platformok), a szervezeti tényezők (pl. a szervezeti kultúra és struktúra) és a környezeti tényezők (pl. szabályozási követelmények, az érdekelt felek nyomása) milyen kölcsönhatásban alakítják az ESG-mérőszámok és a fenntarthatósági keretek elfogadását és végrehajtását.

A Socio-Technical Theory (STT) a szervezeti viselkedéssel és tervezéssel foglalkozik, a társadalmi és technikai tényezők kölcsönhatására összpontosítva (Abbas & Michael, 2023). Az STT ezeknek a tényezőknek hangsúlyozza az összehangolását a hatékony és fenntartható munkarendszerek érdekében. Az ESG-mérőszámokkal és fenntarthatósági célokkal kapcsolatosan az STT alkalmazható annak mélyebb megértésére, hogy a szervezetek hogyan integrálják a társadalmi és technikai elemeket az ESG-kezdemenyzéseknek megvalósítása és felhasználása érdekében, betekintést nyújtva a fenntarthatósági célok elérésébe és a felelős gyakorlatok előremozdításába. A javasolt új paradigma eltért attól a felfogástól is, hogy az emberek nélkülözhetők, egy olyan szemléletmód felé, amelyben az egyéneket "fejleszhető erőforrásnak" tekintik, és amely az együttműködést, az elkötelezettséget és a kockázatvállalási környezetet ösztönzi, szemben a versengéssel, az elidegenedéssel és a minimális kockázatvállalással.

#### **4.4.2. Mérhetőség, számviteli elszámoltathatóság**

A Natural Resource-Based View (NRBV) elmélete szerint a vállalati siker valódi forrása annak természeti környezettel való kapcsolatától függ, ezzel a módszerrel törekszik a versenyelőny szerzésére és erre ad támogatást és ötleteket (Walls, Phan & Berrone 2011). A stratégiai menedzsment szélesebb körű erőforrás-alapú szemlélet (RBV) alapján a vállalat erőforrásainak és képességeinek stratégiai értékére összpontosít. Ezen belül kifejezetten a természeti erőforrásoknak versenyelőnyök biztosításában betöltött szerepét vizsgálja. Hangsúlyozza az egyedi és értékes természeti erőforrások fontosságát abból a szempontból, hogy a vállalatok fenntartható versenybeli pozíciókat érjenek el. Fenntarthatósági szempontból az NRBV összekapcsolható annak szemléltetésével, hogy a vállalat által felhasznált természeti erőforrások és a környezetvédelem hogyan járulnak hozzá a vállalat fenntarthatósági teljesítményéhez, kiemelt figyelmet fordítva a természeti erőforrások hatékony kezelésének és kihasználásának stratégiai jelentőségére.

A Total Cost Accounting (TCA) lényege, hogy a rejtett környezeti és társadalmi költségek döntéshozatalba való beépítésére összpontosít (Carlsson et al., 2016). Olyan pénzügyi eszköz, amely a közvetlen és közvetett költségek és megtakarítások szélesebb körének figyelembevételével teljesebb képet ad egy gazdálkodó egység valódi jövedelmezőségéről. Hosszabb időablakot határoz meg, amely a projekt teljes gazdasági vagy kereskedelmi élettartamát tükrözi, figyelembe veszi a pénz időértékét, feltárja a rejtett költségeket, és figyelembe veszi a bizonytalan vagy kevésbé számszerűsíthető költségeket.

Az ESG-mérőszámokkal és a fenntarthatósági célokkal összefüggésbe hozva a TCA alkalmazható a vállalat tevékenységeinek környezeti és társadalmi költségeinek és hasznainak értékelésére és integrálására, a fenntartható gyakorlatok és az ESG-célok elérésének tájékoztatására.

## 5. További lépések

Tekintettel arra, hogy már folyamatban van a kutatás a fenntartható számvitel témakörén belül, a további lépések elsősorban a szekunder források alapján való vizsgálódás, majd ezek alapján egy keretrendszer felállítása. A cél, hogy a rendszer az egész ESG beszámolási rendszert (reporting system) magába foglalja, annak üzleti és gyakorlati szabályozói oldalával együtt. A rendszernek szorosan együtt kell működnie a technológiai újdonságokkal, ezért felmerülnek a következő kérdések:

- A különböző jelentési rendszereknél valószínűsíthető, hogy más-más minőségi szintet eredményeznek?
- Mennyire pontosan/milyen hibahatárral mérhetőek és skálázhatók a használt rendszerek?
- Mennyire köthetőek össze már bevezetett pénzügyi számviteli rendszerekkel?

A továbbiakban fontos lesz, hogy a reporting system tudományos keretrendszerekhez legyen kapcsolva, hogy a folyamatos fejlesztés hatására se évüljön el a keretrendszer aktualitása. Az ESG reporting elméleti megalapozottsága még mindig nem áll stabil lábakon, ezért ennek aktualitását primer kutatással kell vizsgálni. Emellett dönteni kell egy teória- vagy gyakorlatialapú reporting system között, hiszen ez a döntés is nagymértékben befolyásolja majd a használt metrikákat és a szükséges elméleti háttér kutatásának módját. Valamint kiemelten jelentős lenne bizonyítékot találni a technológiai elméletek hasznosulására, tehát, hogy az elméletekben megfogalmazott segítő tényezők mennyiben befolyásolják ténylegesen napjaink technológiával kapcsolatos aggályait.

A keretrendszerek egységesítéséhez vezetne, ha a szakmai szervezetek tudása és elkészített alapjai kapcsolatba kerülne a globális irányító szervekkel, mint az ENSZ vagy az EU, és a rendszereket összesítenék és ezek alapján kidolgoznák azt a keretrendszert, amely univerzális és egyetemesen alkalmazható minden vállalatra, méretarányosan, az adott üzleti modelltől függetlenül.

## 6. Köszönetnyilvánítás

A kutatást az Európai Unió támogatta a Nemzeti Mesterséges Intelligencia Laboratórium keretében (RRF-2.3.1-21-2022-00004). A cikk előzményét képezi a PRES'23 konferencián prezentált eredményekhez kapcsolódó kutatás.

## Irodalomjegyzék

- [1.] Abbas, R. & Michael, K. (2023) Socio-Technical Theory: A review. In S. Papagiannidis (Ed), TheoryHub Book <https://open.ncl.ac.uk/theories/9/socio-technical-theory/> Letöltés dátuma: 2023.06.26.
- [2.] ACCA & Adam Smith Business School (2022). Professional insight report: Companies' readiness to adopt IFRS S2 Climate-related disclosures: Are companies ready to adopt the planned climate disclosure standard?. 2022.08.22. <https://www.accaglobal.com/gb/en/professional-insights/global-profession/readiness-for-IFRS-S2.html>. Letöltés dátuma: 2023.05.23.
- [3.] Accenture (2023). Sustainability Services: Net Zero Transitions. <https://www.accenture.com/us-en/services/sustainability-index>. Letöltés dátuma: 2023.05.23.
- [4.] B. Carlsson, M. Meir, J. Rekstad, D. Preiß, T. Ramschak 2016: Replacing traditional materials with polymeric materials in solar thermosiphon systems – Case study on pros and cons based on a total cost accounting approach <https://www.sciencedirect.com/science/article/pii/S0038092X15006647> Letöltés dátuma: 2023.07.03.
- [5.] BCG: Sustainable Finance and Investing: <https://www.bcg.com/capabilities/climate-change-sustainability/sustainable-finance-and-investing> Letöltés dátuma: 2023.05.23.
- [6.] Denkstatt: Mik a GRI-szabványok? 2021.10.10. <https://denkstatt.eu/uj-gri-szabvanyok-2021-az-agazatspecifikus-szabvanyok-celja-hogy-nagyobb-atlathatosagot-biztositsanak/?lang=hu> Letöltés dátuma: 2023.05.23.
- [7.] EFRAG-First set of ESRS <https://www.efrag.org/lab6?AspxAutoDetectCookieSupport=1> Letöltés dátuma: 2023.05.23.

- [8.] Kezia Farnham: ESG Metrics: What Matters, and What Should You Measure? 2022.04.06. <https://www.diligent.com/insights/esg/metrics/> Letöltés dátuma: 2023.05.23.
- [9.] Neil Charness, Walter R. Boot, in Handbook of the Psychology of Aging (Eighth Edition), 2016 <https://www.sciencedirect.com/topics/social-sciences/technology-acceptance-model> Letöltés dátuma: 2023.06.26.
- [10.] Oliveira & Martins, 2010; Ramdani & Kawalek, 2007; Rogers, 1983; Swanson, 1994; Tornatzky & Fleischer, 1990 <https://isri.sciencesphere.org/index.php?o=model&m=TOE> Letöltés dátuma: 2023.06.26.
- [11.] Quantive: What are common ESG metrics and how to measure them? <https://quantive.com/resources/articles/esg-metrics#what-are-common-esg-metrics-and-how-to-measure-them?> Letöltés dátuma: 2023.05.23.
- [12.] Ramdani, B., & Kawalek, P. (2007). SME Adoption of Enterprise Systems in the Northwest of England. In Organizational Dynamics of Technology-Based Innovation: Diversifying the Research Agenda (pp. 409-429). Boston, MA: Springer US
- [13.] Reicher, Regina Zsuzsanna (2016) CRM rendszerek; Budapest, Magyarország: Óbudai Egyetem, 160 p.; ISBN: 9786155460654.
- [14.] TheoryHub: Technology-Organization-Environment Framework <https://open.ncl.ac.uk/academic-theories/23/technology-organization-environment-framework/> Letöltés dátuma: 2023.06.26.
- [15.] UNFCCC: Guidance for MRV systems, 2016: <https://agledx.ccafs.cgiar.org/estimating-emissions/mrv-systems/> Letöltés dátuma: 2023.06.26.
- [16.] Walls, Phan & Berrone 2011: Section 6.3: Natural resource based view framework <https://csmathsg.com/course-content/week-6/section-6-3-natural-resource-based-view-framework/> Letöltés dátuma: 2023.06.26.
- [17.] Eisinger, B., Németh, A., Major, Z., & Kegyes-Brassai, O. (2022). Comparative Life Cycle Analyses of Regular and Irregular Maintenance of Bridges with Different Support Systems and Construction Technologies. Chemical Engineering Transactions, 94, 571-576.

## **Az egyetemi technológiatranszfer jogi vetületei különös tekintettel az akadémiai spin-off vállalkozásokra**

### **The Legal aspects of University Technologytransfer with special remarks on academic spin-off enterprises**

**Stipkovits Tamás István**

Széchenyi István Egyetem

stipkovits.tamas.istvan@ga.sze.hu

#### **Absztrakt**

---

*A felsőoktatási intézmények sikeressége a második akadémiai forradalom óta egy új tényező szemszögéből is mérhetővé vált ez pedig a kutatási eredmények piaci hasznosítása. Jelen tanulmány az egyetemi technológiatranszfer vonatkozásában számba veszi a szakirodalom legfontosabb megállapításait majd azt követően a vállalkozó egyetemek (felsőoktatási intézmények) kialakulását mutatja be. Az akadémiai vállalkozással, mint a hasznosítás egyik lehetséges fajtájával kapcsolatban ismerteti a magyar jog jelenlegi szabályozását és észrevételeket fogalmaz meg annak fejlesztéséhez.*

**Kulcsszavak:** akadémiai forradalmak, egyetemi technológiatranszfer, akadémiai spin-off, Bayh-Dole Act, Innovációs tv.

#### **Abstract**

---

*Since the second academic revolution, the success of higher education institutions has been measured in terms of a new factor - the market exploitation of research results. This paper reviews the main findings of the literature on university technologytransfer and then describes the emergence of entrepreneurial universities (Higher Education Institutions). With regard to academic entrepreneurship as a possible type of exploitation, it describes the current Hungarian legal framework and makes comments on its development.*

**Keywords:** academic revolutions, university technologytransfer, academic spin-off, Bayh-Dole Act, Hungarian Act on Innovation

## 1. Bevezetés

Az egyetemek eredményességét olyan faktorok alapján ítéli meg a társadalom mint, hogy a hallgatók milyen sikereket érnek el, milyen rövid idő alatt tudnak elhelyezkedni a diplomájukkal, vagy hogy milyen színvonalas kutatási projekteken vesz részt az intézmény. Futurisztikusnak hatna-e az egyetemek minősítése kapcsán, ha a sikerességet az egyetemi cégek mutatói vagy az egyetem tulajdonában álló szellemi alkotások hasznosításából vagy értékesítéséből származó bevételek alapján (is) mérnénk? A felsőoktatási intézmények ilyen jellegű tevékenységvégeztése ugyanis már nem a jövő, hanem a jelen.

## 2. Kutatási módszertan

Jelen kutatás az egyetemi technológiatranszfer, azon belül is az akadémiai spin-off vállalkozások hazai szabályozását hivatott bemutatni. A téma teljeskörű prezentálásának érdekében a találmány az egyetemi technológiatranszfer fogalmának ismertetését a nemzetközi szakirodalom megállapításai alapján látja el. Az akadémiai spin-off vállalkozás definiálása után azok kialakulásánál – tekintettel arra, hogy az USA-ban jelentek meg először – elsősorban a nemzetközi szakirodalom megállapításaira alapozva demonstrálja a témát, nem nélkülözve, a releváns külföldi jogszabályok számba vételét és a hazai szakirodalom eredményeit. A kutatás a hazai szabályozás bemutatásának a magyar kutatók által elért konklúziókkal együtt a vonatkozó jogszabályok, különös tekintettel a tudományos kutatásról, fejlesztésről és innovációról szóló 2014. évi LXXVI. törvény ismertetésével tett eleget.

## 3. Technológiatranszfer

A „*technológiatranszfer*” fogalmával kapcsolatban szerteágazó értelemzések születtek,<sup>1</sup> ezért LANE (1991:7.) összegzését átvéve a „*technológia*” valamilyen koncepció, termék vagy eljárás, míg a „*transzfer*” az „A”-ból „B”-be jutás, tehát a mozgás szinonimájaként jelenik meg. AUTIO & LAAMANEN munkája szerint a fő különbség a technológiatranszfer és a technológia terjedés között (technology diffusion), hogy előbbi esetben egy szándékos cselekmény eredményezi a technológia „*mozgását*”. (1995) A technológia transzfer formális csatornái között nevezte meg

---

<sup>1</sup> Egyesek egy meglévő technológia újszerű alkalmazásának kigondolását, mások a kutatás gazdasági fejlődéséért való átkonvertálását értik alatta, lásd: LANE (1991:7.)

RADOSEVIC a közvetlen külföldi befektetést (foreign direct investment), a közös vállalkozás létrehozását (joint venture) és a szellemi alkotás időlegesen átengedett használatát (licensing). (1999:20)

#### **4. Egyetemi technológiatranszfer**

Az egyetemi technológiatranszfer mint jelenség kialakulására ETZKOWITZ (2001:19-20.) az „*akadémiai forradalmak*” elméletén keresztül világít rá. Az *első akadémiai forradalom* a XIX. század elején, Poroszországból indult el, amikor – a humboldti reformok következtében – az egyetem funkciója az oktatás mellett a kutatással bővült ki. A *második akadémiai forradalom* kezdetét pedig a XX. század második feléből és az USA-ból eredezteti ETZKOWITZ (2001:21.), aminek következtében új funkcióval, az oktatás és kutatás mellett *társadalmi (gazdasági) szerepvállalással* bővült az egyetemek feladatköre. ERDŐS (2019:15.) és MAKRA (2019:58.)

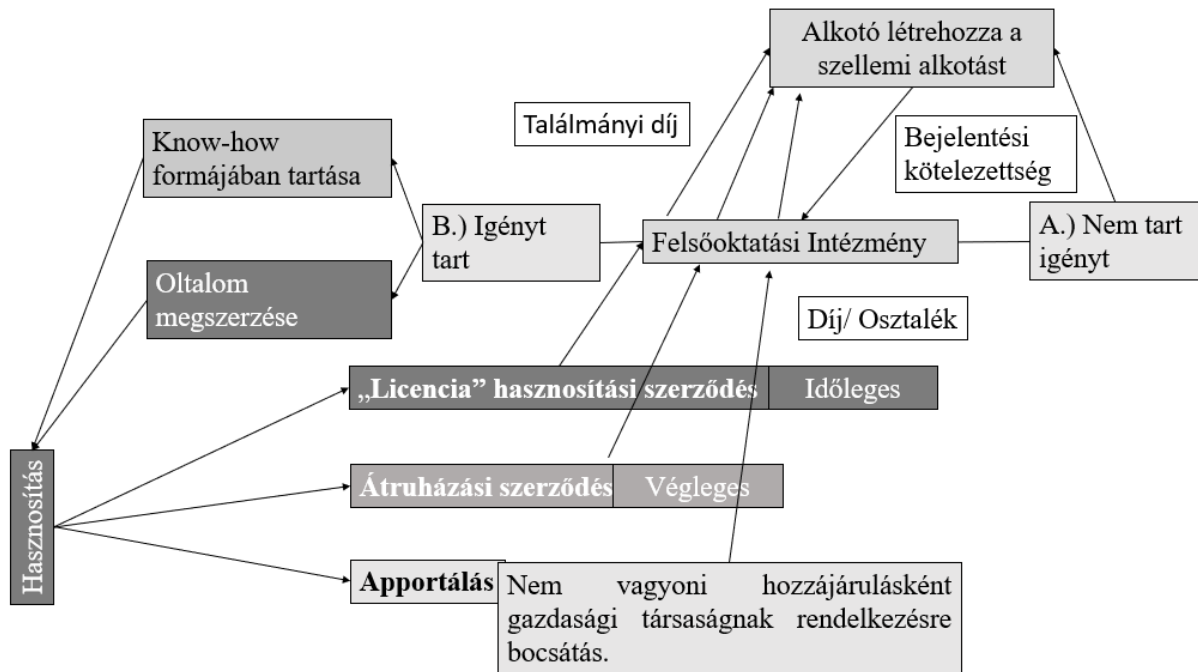
ETZKOWITZ, az Akadémia-Ipar kapcsolatokat – nem taxatív módon – három fajta szervezethez kötve kategorizálta: a.) az *Összekötő Irodák* (Liaison Offices), melyek feladata, hogy arról informálják az érdeklődő piaci szereplőket, hogy milyen kutatásokat végeznek az alkalmazottak; b.) a *Technológia-Transzfer Irodák* (Technology-Transfer Offices), amely az üzletileg hasznosítható szellemi alkotásokat „*pörgeti ki*” az iparba (gazdasági körforgásba) és c.) az *inkubátor programok* (incubator programmes), melyek a szervezeti tanulás megnyilvánulásaként vannak jelen. (2001:25.)

2003-ban CHESBROUGH úgynevezett „*nyílt innovációs*” teóriája arra hívta fel a figyelmet, hogy az innovációs projekteknek nem kell szükségszerűen csak az adott gazdasági szereplő belső „*forrásaira*” építenie. A gazdasági szereplő tulajdonában álló szellemi tulajdon kezelése úgy is jelenthet bevételi forrást, hogy annak hasznosítását nem maga a gazdasági szereplő végzi el. (MOLNÁR 2011:18.) A szellemi alkotások „*piacán*” pedig komoly jelentőséggel bírhatnak az akadémiai környezetben létrejött innovatív megoldások.

#### **5. Akadémiai spin-off vállalkozások**

Az „*akadémiai spin-off*” mint fogalom az alábbi elemekből épül fel: a „*spin off*” és „*spin out*” kifejezések a létrejött újdonság „*kipörgetésére*” utalnak (együttesen: „*spin-off*”), ameddig az „*akadémiai*” jelző funkciója pedig a vállalati spin-off cégektől való elhatárolás. Utóbbi esetben

a vállalatok valamilyen kockázat szem előtt tartása végett adott tevékenység(ek)et vagy projekt(ek)et nem közvetlenül maguk, hanem közvetett módon egy leányvállalaton keresztül végeznek el.<sup>2</sup> (MAKRA, 2019:59.)



1. ábra: Az egyetemi innováció hasznosításának útjai és lehetőségei (saját szerkesztés)

## 6. Kialakulása

Az akadémiai spin-off vállalkozások megjelenése az USA-hoz köthető, ezért a kifejlődésük során szerepet játszó tényezők megértéséhez szükséges a történeti előzmények rövid áttekintése. Az akadémiai spin-off vállalkozások létrejötte szempontjából esszenciális előfeltétel volt a kutató egyetemek (német ihletésre való) elterjedése az USA-ban. Viszont az amerikai egyetemek fenntartásában az állam csak marginálisan vállalt szerepet, ezért önfenntartásuk és az egyetemi alkalmazottak kutatásainak finanszírozása végett saját bevételek szerzésére kényszerültek. (ETZKOWITZ 2001:20-21.) Ennek következtében a felsőoktatási intézetek is élénkebb figyelmet fordítottak az alkalmazottaik által előállított szellemi alkotásokra.

1912-ben FREDERICK COTTRELL, a University of California professzora megalapította szellemi alkotásainak hasznosítása végett a Research Corporation-t. A hasznosításból származó

<sup>2</sup>Jelen cikk tárgyául az akadémiai spin-off vállalkozások szolgálnak.

jövedelmeket újabb tudományos projektek finanszírozására és kutatói ösztöndíjak nyújtására fordították. (CAMPBELL 2019: 92-93.) Vagyonos New York-i üzletemberek megalapították a Smithsonian Institution mellett működő Research Coproration-t, ezáltal nem magának – az üzleti tevékenység végzésétől amúgy is elzárkózó – Smithsonian Institution-nak, hanem az erre a célra létrehozott Research Coproration-nek kellett a szellemi tulajdon menedzsmentjével foglalkoznia. A vállalkozás által termelt nyereségből pedig részesedett a Smithsonian Institution is. (APPLE 1989: 383.) 1925-ben megalakult a University of Wisconsin mellett működő Wisconsin Alumni Research Foundation. A koncepció lényege, hogy az egyetem és a Research Foundation egymástól elkülönülten működött; az Egyetem Kutatási Bizottsága kiválasztotta a projekteket, a Research Foundation pedig a szellemi tulajdon menedzsmenttel foglalkozott egyfajta közvetítőként az akadémia és az üzleti élet között. (APPLE, 1989:383.) Illetve ebben az időszakban jelentek meg az USA egyetemein a szellemi tulajdonkezelési szabályzatok.

A második világháború tapasztalatai nagy mértékben hozzájárultak a második akadémiai forradalom „kirobbanásához.” A BUSH által 1945-ben megírt „*Sicence: The endless forntier*” a k+f támogatásokkal kapcsolatban olyan elvi meglátásokat fogalmazott meg, mint az *alap- és az alkalmazott kutatás* közötti különbségtételt, kiállva amellett, hogy *szükség van olyan kutatások támogatására (alapkutatások) is, melyektől nem vár a támogató – közvetlen – gyakorlati előnyt, vagyis hasznosíthatóságot.* BUSH azzal érvel, hogy az autómobil, a repülés és a rádió területén elért fejlesztések valójában azokon az alapvető áttöréseken alapulnak, melyeket a XIX. századi Európában értek el, tehát a II. világháború alatti nagy előrelépéseknek a zöme valójában ez imént említett európai „*fundamentális*” felfedezések teljes kimerítésén alapult.<sup>3</sup> Alapkutatások (basic research) támogatására pedig azért van szükség, hogy újra létrejöhessenek ilyen alapvető áttörések.<sup>4</sup> (BUSH, 2020:17. és 86.) A National Science Foundation 1950-es létrehozása szintén egy fontos mérföldkő volt a szövetségi k+f támogatások szempontjából.<sup>5</sup>

---

<sup>3</sup> MACLAURIN, BUSH kortársa a rádió feltalálása és elterjedése kapcsán végzett tanulmányában amellett érvelt, hogy az áttörő (fundamentális) felfedezéseket elérő tudósok (pl.: Hertz vagy Maxwell) nem biztos, hogy tudatosan gondolnak az áttöréseik piaci hasznosíthatóságára, ezért szükség van olyan inventorokra (pl.: Marconi), akiknek kifejezetten ez a feladatuk.  
lásd: (GODIN, 2008:352.)

<sup>4</sup> Gondolatainak hatását bizonyítja, hogy napjainkban is például a magyar (új) Innovációs tv. is az alap- és alkalmazott kutatás megosztásán alapul.

<sup>5</sup> Habár a Department of Defense és Atomic Energy Commission már ekkor önállóan is szponzoráltak k+f projekteket. lásd: (USSELMAN, 2013:18.)

	2015-ös Frascati kézikönyv	Új Inno. tv.
Alapkutatás (Basic/ Fundamental Research)	„Kísérleti vagy elméleti munka, amelyet elsősorban a jelenségek vagy megfigyelhető tények háttérével kapcsolatos új ismeretek megszerzésének érdekében folytatnak, anélkül, hogy kilátásba helyeznének azok közvetlen üzleti alkalmazását vagy felhasználását.” F.k. 45.o.	„Kísérleti vagy elméleti munka, amelyet elsősorban a jelenségek vagy megfigyelhető tények háttérével kapcsolatos új ismeretek megszerzésének érdekében folytatnak, anélkül, hogy kilátásba helyeznének azok közvetlen üzleti alkalmazását vagy felhasználását.” Új Inno. tv. 3.§ 1.
Alkalmazott/ Ipari kutatás (Advanced/ Industrial Research)	„Tervezett kutatás vagy kritikus vizsgálat, amelynek célja új ismeretek és szakértelem megszerzése új termékek, eljárások vagy szolgáltatások kifejlesztéséhez, vagy a létező termékek, eljárások vagy szolgáltatások jelentős mértékű fejlesztésének elősegítéséhez, ...” F.k. 45.o.	„Tervezett kutatás vagy kritikus vizsgálat, amelynek célja új ismeretek és szakértelem megszerzése új termékek, eljárások vagy szolgáltatások kifejlesztéséhez, vagy a létező termékek, eljárások vagy szolgáltatások jelentős mértékű fejlesztésének elősegítéséhez, ...” Új Inno. tv. 3.§ 2.
Kísérleti fejlesztés (Experimental Development)	„A meglévő tudományos, technológiai, üzleti és egyéb, vonatkozó ismeretek és szakértelem megszerzése, összesítése, alakítása és felhasználása új vagy javított termékek, eljárások vagy szolgáltatások kidolgozása céljából...” F.k. 45.o.	„A meglévő tudományos, technológiai, üzleti és egyéb, vonatkozó ismeretek és szakértelem megszerzése, összesítése, alakítása és felhasználása új vagy javított termékek, eljárások vagy szolgáltatások kidolgozása céljából ...” Új Inno. tv. 3.§ 7.

**2. ábra:** A K+F kategóriáinak „magdefiníciói” a 2015-ös Frascati kézikönyv és az Új Inno. tv. szerint (saját szerkesztés)

Az USA az 1970-es évek végére kezdte elveszíteni technológiai dominanciáját a világgazdaságban: a japán és nyugat-európai országok – kiheverve a világháborús helyreállítás nehézségeit – az egy főre jutó nemzeti össztermék kapcsán megközelítették az USA számait (MOLNÁR 2004: 66.), az észak-amerikai ország jelentősége csökkent az acél- és autógyártásban, valamint csökkent az újonnan bejegyzett szabadalmak száma az országban stb. (VIGH 2008:158.) A megújulás szükségessége kulcsfontosságú kérdéssé vált, ezért az 1980-as évek során egy törvényalkotási hullám helyezte új alapokra az USA innovációs ökoszisztémáját:

- a) a Bayh-Dole Act<sup>6</sup> (1980),
- b) a Stevenson-Wydler Act<sup>7</sup> (1980),
- c) a Small Business Innovation Development Act<sup>8</sup> (1982) és
- d) a National Cooperative Research and Production Act<sup>9</sup> (1984).

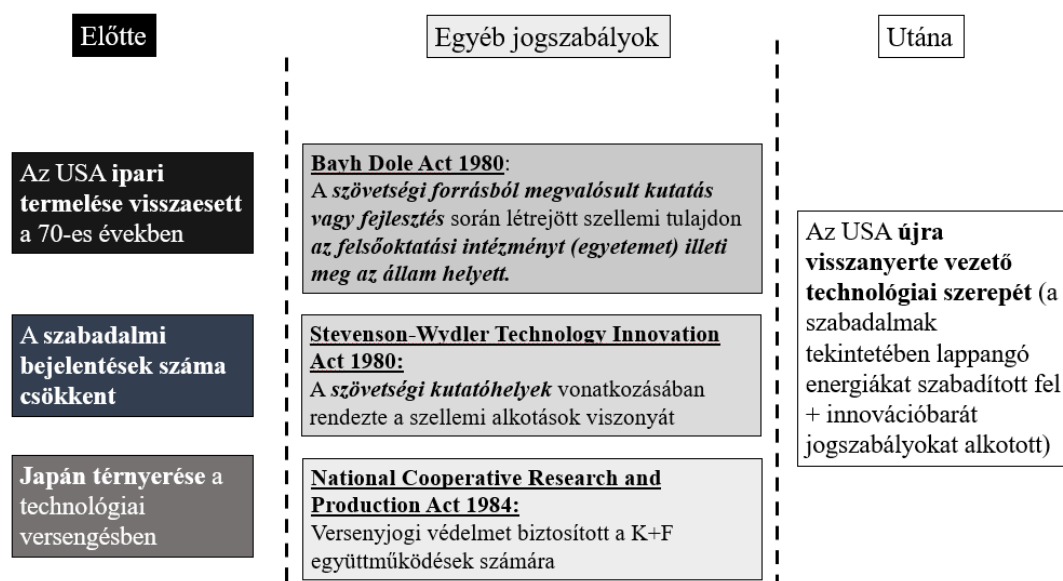
<sup>6</sup> Patent and Trademark Law Amendments Act of 1980, Public Law 96-517

<sup>7</sup> Stevenson-Wydler Technology Innovation Act of 1980, Public Law 96-480

<sup>8</sup> Small Business Innovation Development Act of 1982, Public Law 97-219

<sup>9</sup> National Cooperative Research and Production Act of 1984, Public Law 98-462

A *Bayh-Dole Act* – ami a szabadalmi és védjegy törvényt módosította - abból a szempontból volt forradalmi, hogy a felsőoktatási intézmény tulajdonába „*adta át*” a szövetségi kormány által támogatott k+f projektekből létrejött szellemi alkotás tulajdonjogát, mindamellett, hogy a szövetségi kormány bizonyos feltételeket szabott a felsőoktatási intézmények számára, lásd.: „*march-in right*.” (VIGH, 2006:4.) Ez az intézkedés egy hatékonysági alapú újragondolása volt a szellemi tulajdon kérdésének, arra alapozva, hogy a hasznosítással kapcsolatos döntések „*központosítása*” aránytalanul bürokratikussá teszi a döntéshozatalt, illetve nagyobb motiváltságot is feltételezett a felsőoktatási intézet oldalán.<sup>10</sup> A *Stevenson-Wydler Act*, - melyet átfogó jelleggel módosított az 1986-os Federal Technology Transfer Act<sup>11</sup> - az innovációs politikai közigazgatási rendszerét vázolta fel, valamint rendelkezett a kormányzati kutatóhelyen létrejött szellemi alkotás sorsáról. A *Small Business Innovation Development Act* kisvállalkozások számára nyújtott k+f támogatást (Small Business Innovation Research Program, rövidítve: „*SBIR program*”), ameddig a *National Cooperative Research and Production Act* versenyjogi szempontból biztosított jogszerű lehetőséget a k+f együttműködések számára.



3. ábra: Az USA pozíciója a technológiai versenyben tükrében a 70-es 80-as években (saját szerkesztés)

<sup>10</sup> Ennek hiánya esetén a központi kormányzat „*march-in*” jellegű jogosultságai jöhetnek képbe.

<sup>11</sup> Federal Technology Transfer Act of 1986, Public Law 99-502

Természetesen a nyolcvanas évek elején kialakított innovációs rendszer azóta módosításokon esett át (pl.: az 1984-es változatot az 1993-as felülírta - National Cooperative Research and Production Act<sup>12</sup>) illetve új elemekkel is bővült (pl.: a SBIR program sikerén felbuzdulva megszületett Small Business Technology Transfer Act.<sup>13</sup>)

Magyarországon a 2004. évi CXXXIV. törvény (rövidítve: „*régi Inno.tv.*”) szabályozta elsőként az akadémiai spin-off vállalkozásokat, „*hasznosító vállalkozás*” néven, jelenleg a régi Inno.tv.-t felváltó 2014. évi LXXVI. törvény (rövidítve: „*új Inno.tv.*”) szabályozza a hasznosító vállalkozásokat.

## 7. Hazai szabályozása

Az új Inno.tv. az alábbiak szerint határozza meg a hasznosító vállalkozás fogalmát: „*költségvetési kutatóhelyen létrejött szellemi alkotás hasznosítása céljából alapított gazdasági társaság, amely nem pénzbeli hozzájárulásként rendelkezésre bocsátás, átruházás vagy hasznosítási szerződés alapján vált a költségvetési kutatóhelyen, valamint a felsőoktatási intézményben létrejött szellemi alkotáshoz fűződő jogok jogosultjává vagy hasznosítójává.*<sup>14</sup>” A definíció ugyan tartalmaz némi önellentmondást: nevezetesen, hogyha a hasznosító vállalkozás a költségvetési kutatóhelyen, valamint a felsőoktatási intézményben (együttesen: akadémiai környezetben) létrejött szellemi alkotáshoz fűződő jogok jogosultja vagy hasznosítója, akkor miért „csak” a költségvetési kutatóhelyen létrejött szellemi alkotás hasznosítása céljából alapított gazdasági társaságot tartjuk hasznosító vállalkozásnak. A szellemi tulajdon hasznosításának három módját határozza meg a törvényi definíció: „*nem pénzbeli hozzájárulásként [a gazdasági társaság] rendelkezésre bocsátás, átruházás vagy hasznosítási szerződés.*” A hasznosító vállalkozással szemben nem követeli meg az imént említett meghatározás, hogy a költségvetési szerv vagy felsőoktatási intézmény valamilyen módon (pl.: tagként) közreműködjön a vállalkozás életében. Ellenben mégsem valamennyi olyan gazdasági társaság tartozik a hasznosító vállalkozás kategóriájába, amely akadémiai környezetben létrejött szellemi alkotáshoz fűződő jog jogosultja vagy hasznosítója, hanem csak

---

<sup>12</sup> National Cooperative Research and Production Act of 1993, Public Law 98-462

<sup>13</sup> Small Business Technology Transfer Act of 1992, Public Law 102-564.

<sup>14</sup> Új Inno.tv. 3.§4.

azok, melyeket hasznosítás céljára alapítottak.<sup>15</sup> A törvény előírja, hogy a költségvetési kutatóhelynek és a felsőoktatási intézménynek – más szervekkel együtt - szellemi tulajdonkezelési szabályzattal kell rendelkezniük.<sup>16</sup> A szellemi alkotás hasznosítása vagy értékesítése (a költségvetési kutatóhelyek és a felsőoktatási intézmények számára) a szellemi tulajdonkezelési szabályzatban foglaltak szerint mehet végbe,<sup>17</sup> aminek – többek között – tartalmaznia kell a szellemi alkotás létrehozásában közreműködő kutatóknak - a hasznosítás folyamatával kapcsolatos - jogaikra és kötelezettségeikre vonatkozó szabályozást.<sup>18</sup>

<b>Innovációs (2014. évi LXXVI.) tv. 3.§ 4.</b>		<b>Tágabb értelemben vett hasznosító vállalkozás (Közvetett jelenlét)</b>
<b>Szűkebb értelemben vett hasznosító vállalkozás (Közvetlen jelenlét)</b>		
<b>Állami fenntartású Felsőok. Int.</b>	<b>Nem állami fenntartású Felsőok. Int.</b>	
<b>Innovációs tv. 34. § (2) <u>Költségvetési kutatóhely</u></b> által alapított vagy részesedésével működő hasznosító vállalkozásra a nemzeti felsőoktatásról szóló törvény szerinti <b>intézményi társaságra</b> vonatkozó szabályokat kell megfelelően alkalmazni. <i>Tehát nem állami F.I.-ben nem kell!</i>	<b>Felsőoktatási (2011. évi CCIV.) tv. 88.§ (5) <u>Felsőoktatási intézmény</u></b> üzemeltetéssel, szervezéssel összefüggő tevékenység, valamint <b>kutatás-fejlesztési és innovációs tevékenység ellátására</b> , továbbá a (6) bekezdésben foglaltak szerint, <b>intézményi társaságot</b> hozhat létre.	
<b>Intézményi társaság szabályai</b> 88. § (3) <b>Felügyelőbizottság létrehozása</b> (egy tagot az állami vagyon felügyeletéért felelős miniszter delegálhat) (4) <b>nem sérti</b> [a társaság] a F.I. <b>érdekeit</b> . (8) <b>Személyi korlátok</b> (3) <b>Állami tulajdonban</b> álló gazdasági társaságok szabályai (2) <b>Állami F.I.-re</b> nézve: formakényszer: <b>Kft., Zrt.</b>		
<b>Állami tulajdonban álló G.t.-k szabályai</b> <i>alapítására, a részesedésszerzésére és a vezető tisztségviselőjének felelősségére</i>		

4. ábra: Az akadémiai spin-off vállalkozások hazai jogi szabályozása (saját szerkesztés)

## 8. Összegzés

A technológiatranszfer fogalma alatt tehát valamilyen koncepció, termék vagy eljárás „A”-ból „B”-be való elmozdulását kell értenünk, mindennek egy speciális esete az, amikor az „A” pont egy felsőoktatási intézmény, ezt nevezzük „*egyetemi technológiatranszfernek*.” Az a jelenség, hogy a felsőoktatási intézményeknek ilyen jellegű feladatuk lehet, illetve van, a második akadémiai forradalom következménye. Aminek eredményeként az oktatás és a kutatás mellett

<sup>15</sup> Azt, a dilemmát, hogy végülis a költségvetési kutatóhelyen létrejött szellemi alkotás hasznosítására kell-e létrehozni vagy megfelelő, ha a felsőoktatási intézményben létrejött szellemi alkotás hasznosítása a cél, már a fentiekben ismertettük.

<sup>16</sup> Új Inno.tv. 33.§ (1)

<sup>17</sup> Új Inno.tv. 34.§ (1)

<sup>18</sup> Új Inno.tv. 33.§ (2) d)

a (társadalmi) gazdasági szerepvállalással bővült a funkciók köre. Az Akadémia-Ipar kapcsolatok következménye a Technológia-Transzfer Irodák létrejötte és működése, melyek az akadémiai közegben létrejött üzleti potenciállal bíró szellemi alkotások „kipörgetésére” hivatottak. A nyílt innováció elképzelése pont azt a felfogást testesíti meg, miszerint a szellemi alkotások aktív, más szereplők bevonásával végbemenő felhasználására van szükség. Az akadémiai spin-off vállalkozások az USA-ban először az 1920-as években tűntek fel, mivel az egyetemek motiváltak voltak a kutatások során létrejött szellemi alkotások hasznosításában, ami jellemzően egy egyetemmel párhuzamosan működő gazdasági szereplő tevékenységi körébe esett. Az 1980-as években az USA-nak újra kellett gondolni innovációs politikáját, aminek következtében megszületett a Bayh-Dole Act, ami a szövetségi forrásból támogatott kutatásból megvalósuló szellemi alkotás hasznosítását a felsőoktatási intézményre delegálta (csökkentve a bürokratikus terheket. illetve az intézmény motiváltságát vélelmezve). Magyarországon először a 2004-ben meghozott régi Inno.tv. szabályozta az akadémiai spin-off cégeket „hasznosító vállalkozás” név alatt, jelenleg az új Inno.tv. tartalmazza a hasznosító vállalkozások definícióját.

A felsőoktatási intézményekkel szembeni követelmények átalakulásának vagyunk a szemtanúi. BARTHA szavaival élve a „kommunikációs és információs technológiák, módszerek fejlődésének következtében a tudás privilégiuma megszűnt, így mindenki számára hozzáférhetővé vált.” (2014:320.) A tradicionális feladatkörök átalakulása mellett újszerű funkciók jelennek meg, ami a felsőoktatási intézetben létrejövő tudás másféle megnyilvánulásához vezet.

## **9. Köszönetnyilvánítás**

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

## Irodalomjegyzék

- [1.] 2004. évi CXXXIV. törvény a kutatás-fejlesztésről és a technológiai innovációról (rövidítve: „*régi Inno. tv.*”)
- [2.] 2014. évi LXXVI. törvény a tudományos kutatásról, fejlesztésről és innovációról (rövidítve: „*új Inno. tv.*”)
- [3.] APPLE, R. D. (1989). Patenting University Research Harry Steenbock and the Wisconsin Alumni Research Foundation. *Isis* 1989(3)
- [4.] AUTIO, E. & LAAMANEN, T. (1995). Measurement and evaluation of technology transfer: Review of technology transfer mechanisms and indicators. *International Journal of Technology Management*, 10(7/8), 643-664.o.
- [5.] BARTHA K. (2014) A tudás- és technológiatranszfer értelmezése: technológia, vagy tudástranszfer? *Taylor Gazdálkodás- és szervezéstudományi folyóirat*. 6(1-2)
- [6.] CAMPBELL P. E. (2019). University Inventions Reconsidered: Debunking the Myth of University Ownership. *William & Mary Business Law Review*. (11) 177-125.o.
- [7.] ERDŐS K. (2009). Vállalkozó egyetemek kialakulása az Amerikai Egyesült Államokban és Európában. *A Virtuális Intézet Közép-Európa Kutatására közleményei*, 2009(1)
- [8.] ETZKOWITZ H. (2001). The Second Academic Revolution and the Rise of Entrepreneurial Science. *IEEE Technology and Society Magazine* 20(2)
- [9.] Federal Technology Transfer Act of 1986, Public Law 99-502
- [10.] GODIN, B. (2008). In the shadow of Schumpeter: W. Rupert Maclaurin and the study of technological innovation. *Minerva*. 2008(3) 343-360.o.
- [11.] LANE, J. P. (1999). Understanding Technology Transfer in: *Assistive technology: the official journal of RESNA*, 1999(1)
- [12.] MAKRA ZS. (2019). Az egyetemi spin-off vállalkozások jellegzetességei és alapításának folyamata a nemzetközi szakirodalom tükrében In: INZELT A. & BAJMÓCY Z. (Eds.) *Innovációs rendszerek, Szereplők, kapcsolatok és intézmények*. JATEPress. Szeged

- [13.] MOLNÁR I. (2004). Az Amerikai Egyesült Államok innovációs rendszerének jogi és intézményi vonásai. *Marketing & Menedzsment*. 38(5) 65-75.o.
- [14.] MOLNÁR I. (2011) Szabadalmi anyagi jogi problémák és megoldási elképzelések közfinanszírozású kutatóhelyeken. Szegedi Tudományegyetem
- [15.] National Cooperative Research and Production Act of 1984, Public Law 98-462
- [16.] National Cooperative Research and Production Act of 1993, Public Law 98-462
- [17.] Patent and Trademark Law Amendments Act of 1980, Public Law 96-517 (rövidítve: „Bayh-Dole Act”)
- [18.] RADOSEVIC S. (1999). *International Technology Transfer and ‘Catch Up’ in Economic Development*. Edward Elgar. Cheltenham.
- [19.] Small Business Innovation Development Act of 1982, Public Law 97-219
- [20.] Stevenson-Wydler Technology Innovation Act of 1980, Public Law 96-480, 1980
- [21.] VIGH J. F. (2006). Az USA-egyetemek alkalmazott kutatási tevékenysége a technológiai innováció jegyében Korszakváltás a Bayh-Dole Act nyomán. *Külgazdaság (Jogi melléklet)*, 2006(1)
- [22.] VIGH J. F. (2008). A Bayh–Dole Act megvalósulása és hatása az USA-ra – Az egyetem-  
ipar K+F együttműködés néhány kérdése az EU-ban. *Külgazdaság (Jogi Melléklet)*, 2008(11-  
12) 155-174.o.

## Szoftverarchitektúra tervezete európai üzleti adatok automatizált kinyerésére és kezelésére

### Outline of Software Architecture for Automated Retrieval and Processing of European Corporate Business Information

Suta Alex\*, Tóth Árpád, Kedves Loránd

\*Széchenyi István Egyetem – Járműipari Kutatóközpont

suta.alex@ga.sze.hu

#### Absztrakt

---

*Az éves jelentések a vállalati teljesítményre vonatkozó információk közlésének elsődleges eszközeként szolgálnak. Az Európai Unióban a közelmúltban bekövetkezett szabályozási fejlemények a nyilvánosan jegyzett, beszámolót készítő vállalatok számára digitális szabványok, különösen az eXtensible Business Reporting Language (XBRL) bevezetését írják elő. A nyilvánosságra hozott információkhoz való hozzáférés és azok értelmezése azonban nagy kihívást jelent a szélesebb körű tudományos közösség számára, mivel a mögöttes fájlformátumokban és adatfeldolgozási követelményekben való eligazodáshoz komoly technikai ismeretekre van szükség. A jelen cikk egy új, rugalmas és kiterjeszhető szoftverarchitektúrát fogalmaz meg, amely a legfrissebb adatforrásokból származó üzleti információk lekérdezésére és feldolgozására szolgáló munkafolyamatot mutat be. A tervzet sikeres megvalósítása várhatóan egy olyan nyílt és hozzáférhető technológiai megoldást eredményez, amely tudományos eszközként szolgál majd a különböző XBRL-formátumokban lévő, nagy mennyiségű vállalati adat elemzéséhez.*

**Kulcsszavak:** éves beszámoló, szoftverarchitektúra, XBRL, üzleti adatfeldolgozás, európai szabályozás

#### Abstract

---

*Annual reports are the main method of communicating information regarding company performance. In light of recent regulatory developments within the European Union, publicly listed reporting companies are required to introduce digital standards, specifically eXtensible Business Reporting Language (XBRL). However, accessing and interpreting the disclosed information presents a significant challenge for the scientific community at large, as navigating the underlying file formats and data processing requirements requires considerable technical expertise. This paper introduces a new, adaptable, and scalable software architecture that outlines a workflow for extracting and processing commercial information from the most current data sources. The successful implementation of the design is anticipated to result in an open and accessible technology solution that will serve as a scientific tool for analyzing large volumes of business data in various XBRL formats.*

**Keywords:** annual report, software architecture, XBRL, business data processing, european regulation

## 1. Bevezetés

Ray Ball professzor (2008) több, nagy jeletőségű és eddig megválaszolatlan kérdést vetett fel a pénzügyi beszámolás tényleges gazdasági szerepével kapcsolatban. Hogy miért is merül fel egyáltalán a kérdés – jegyzi azt meg –, az emberek (beleértve a menedzsereket és a könyvvizsgálókat is) nem mindig azt teszik, amit mondanak, vagy amit gondolják, hogy tesznek. Továbbá, a vállalkozások nem azt teszik, amit a menedzserek és a könyvvizsgálók individuálisan tesznek: összetett szervezetek, amelyek tevékenységét egyének és intézmények egész sora befolyásolja, mind a szervezeten belüli struktúrában (vállalati és üzleti egységek vezetői, vállalati jogászok, belső könyvvizsgálók, igazgatóságok), mind a szervezet külső környezetében (könyvvizsgálók, elemzők, sajtó, részvényesek és hitelezők, szabványalkotók, bíróságok, szabályozó hatóságok és szervezetek).

A pénzügyi beszámolásban nagy szerepet kapnak a szabályok (számvetési standardok), de kevésbé kerül magyarázatra a konkrét szerepük a vállalati ökoszisztémában, például nem egyértelmű, hol állnak egyensúlyban a központi szabálytervezés- és az alkalmazás helyi (vállalati) előnyei.

A Nemzetközi Számvetési Standard Testület (IASB) elméleti keretrendszerének (Conceptual Framework) megállapítása szerint az általános célú külső pénzügyi beszámolás célja, hogy olyan információkat nyújtson, amelyek hasznosak a jelenlegi és potenciális befektetők, hitelezők és egyéb stakeholderek számára a befektetési, hitelezési és egyéb erőforrás-elosztási döntések meghozatalában (OB2. §) (Lennard, 2007).

A külső beszámolás (external reporting) a vállalatokra vonatkozó pénzügyi és egyéb információkat egyaránt magába foglalja – az IASB számvetési-elméleti megállapítása szerint – a "döntés-hasznosság (decision-usefulness) és a gondosság (stewardship) céljából. Ezek az információk tartalmukat és formájukat tekintve széles körben elterjedtek és a nemzetközi standardoknak megfelelő beszámolók szerves részét alkotják, ennek értelmében jelentős mennyiségű adatot tartalmaznak. Az adatok kezelésével kapcsolatos fennálló problémák kezelésére a digitalizáció jelentős innovációs lehetőségeket nyit meg a szervezetek előtt: az adatgyűjtés és a vállalati információk feldolgozása lehetővé téve a szervezetek számára a vállalati beszámolók egyszerűbb elkészítését (Lombardi & Secundo, 2020).

Az egyik fő lehetőség az átláthatóság növelésében rejlik. Az átlátható vállalati jelentéstételhez érdemes szétválasztani azon két koncepciót, ami alapján eldől, hogy "mit" akarunk közzétenni, és azt "hogyan" tesszük közzé. Az eXtensible Business Reporting Language (XBRL) egy széleskörben alkalmazott technikai keretrendszer, amely támogatja ezen kérdések megválaszolását. Az XBRL előnye az említett kérdések szétválasztása, mivel a "mi" (szabványok és kapcsolódó technikai taxonómiák) az idővel és az aktuális kontextussal (context element) változik, míg a "hogyan" változatlan, és az XBRL szabvány által szabályzott maradhat. Ezáltal lehetségessé válik:

- Az egyértelmű és stabil dokumentumstruktúra: a közzétételek és az általános felhasználói ismeretek nem válnak elavulttá, és a vállalati XBRL-jelentési gyakorlatok egyre érettebbé válhatnak;
- Egy szabványosított módszer a taxonómiák leírására és az azokra való hivatkozásra. Ez magában foglalja a megengedett címkék meghatározását és leírását, de támogatja az validálási feltételeket is az adott taxonómián alapuló jelentés érvényesítéséhez;
- Az alapértelmezett fájlkiterjesztések és adattípusok készlete;
- A fenti jellemzők részeként magas szintű informatikai eszközök létrehozása az XBRL-jelentésekhez, a taxonómiák létrehozásától és közzétételétől a jelentések létrehozásáig, valamint feldolgozásáig, elemzéséig.

## **2. Az XBRL alkalmazásainak áttekintése**

### **2.1. Európai uniós szabályozási környezet**

Az elmúlt két évtizedben a digitalizáció eredményeként az adatok, különös tekintettel a pénzügyi adatok feldolgozása, felhasználásának módja, értelmezése és ezzel együtt az elektronikus pénzügyi beszámolás jelentős fejlődésen ment keresztül, így napjainkra az XBRL lett az egyik globálisan vezető platform ezen a területen. Az XBRL egy szabadon hozzáférhető, nemzetközi keretrendszer, amelynek célja az üzleti információk összehasonlíthatóságának növelése (XBRL International, 2022). Használatával az elemzett vállalatoknak az eltérő pénzügyi beszámolási szabályozó rendszerekből származó nehézségei kiküszöbölhetők.

Az Európai Unió tőzsdéjén jegyzett vállalatának körében a 2022. január 1-től kezdődő üzleti évekről történő beszámolás során kötelezővé vált az egységes uniós elektronikus formátumú (ESEF, European Single Electronic Format) jelentések készítése (ESMA, 2022), tehát 2023. május végi határidővel már a második teljes üzleti évi ESEF beszámolókat publikálják. Ez korszakváltó előrelépésnek tekinthető az XBRL jelentési nyelv európai használatában, amelynek keretrendszerét az Egyesült Államok Értékpapír- és Tőzsdebizottsága (SEC) már 2005-től használja. A korábbi analóg elemzések helyett a beszámoló teljes tartalmát magába foglaló, sok esetben automatizált megoldásokkal támogatott teljesen új jelentési rendszerek létrehozása válik lehetségessé.

Ezt kiegészíti az International Sustainability Standard Board (ISSB) egy 2021-ben létrehozott szakmai szervezet tevékenysége, amelyek közé tartozik az IFRS-nek megfelelő fenntarthatósági közzétételi szabványok kidolgozása és tanúsítása (Deloitte, 2021). A korábbi nemzetközi együttműködési folyamat eredményeként az ISSB 2022. március 31-i hatállyal konzultációra bocsátotta az IFRS S2 Climate-related Disclosures (Climate Exposure Draft) standardot (IFRS Foundation, 2022). A jelenlegi tesztek várható sikerességével az egységes európai cégbázisok (pl. a <https://filings.xbrl.org>, vagy az European Single Access Point) szintén kiterjesztésre kerülnek a nem-pénzügyi adatok XBRL-alapú lekérdezésével (XBRL International, 2022).

## **2.2. Üzleti adatok felhasználásának eszközei és jelenlegi problémái**

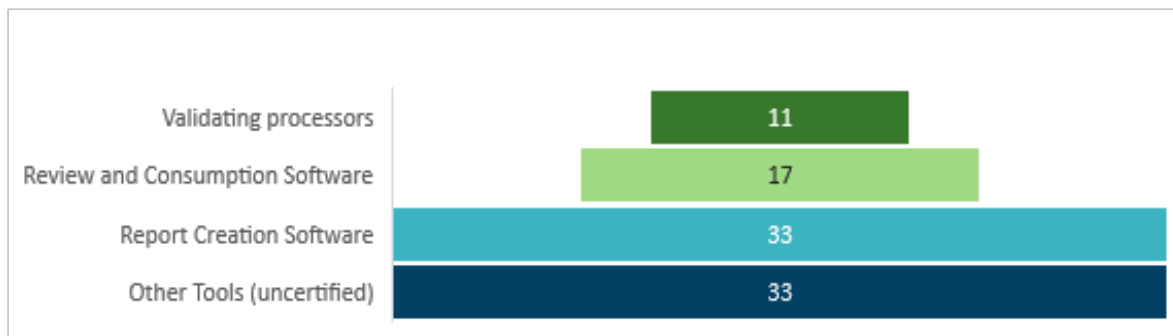
A technikai háttérrel alkotó globális XBRL standard több, mint 20 éves, és maga mögött tudhat több aktív időszakot (2004, 2010-2014), amelyet átmeneti visszaesés követett. A kutatási és fejlesztési eredményekről (a számszerűsíthető kutatási output szintjén is) általánosan elmondható, hogy kapcsolódnak a szabályozók által bevezetett vállalati kötelezettségekhez. Az említett Európai Uniói bevezetés várhatóan a nagymértékben növelni fogja a technológiát felhasználó kutatási eredményeket, és a kutatóközösség általános érdeklődését.

Az elmúlt évtizedekben az XBRL ökoszisztémában számos eszközt hoztak létre profitorientált és nonprofit vállalkozások, szabályozó funkciót ellátó szervek (kormányok és hatóságok), üzleti szereplők és tudományos szervezetek gyakorlati felhasználásának támogatására, amelyekről elmondható, hogy:

- A for-profit eszközök megbízhatóbbak, de az üzleti lehetőségekre összpontosítanak azáltal, hogy eszközöket és szolgáltatásokat nyújtanak a kibocsátóknak (jelentést készítő vállalatok) és a jelentéseket feldolgozó szervezeteknek. Ezek zárt forráskódú és jelentős licenszdíjú termékek, nem általános kutatásra kutatási feladatok kivitelezésére szolgálva.
- A nonprofit és közösségi megoldások az egyéni fejlesztések általánosításából és megosztásából születnek. Ez többféle, egymást átfedő megoldáshoz vezet, amelyek esetében erős függés figyelhető meg a külső komponensektől, és jelentős mennyiségű kódot tartalmaznak. Az ilyen megoldások integrálása vagy javítása időigényes és korlátozott megbízhatóságú.

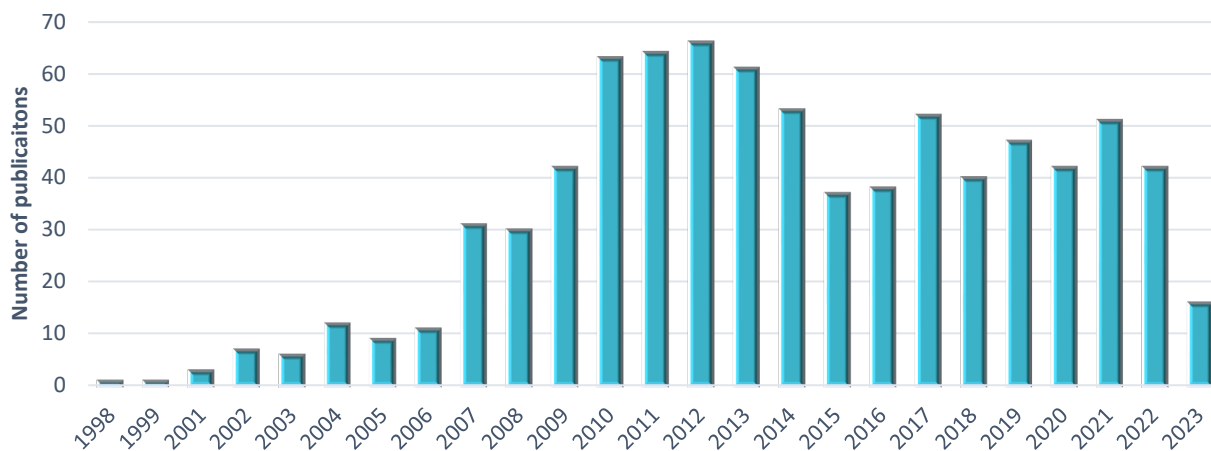
Az XBRL Certified Software™ a szoftverek egy olyan csoportja, amely az XBRL International szervezete ellenőrzött az XBRL specifikációknak megfelelő szempontok alapján. A tanúsítvány jelenleg háromféle szoftvertípusra vonatkozatható, amelyek száma és megoszlása az 1. ábrán került feltüntetésre:

- A **Validating Processor** olyan tanúsított szoftver, amelyet az XBRL vagy „Inline” (böngészőben széleskörben kezelhető) XBRL jelentések érvényesítésére és feldolgozására használnak.
- A **Report Creation Software** vagy jelentéskészítő szoftver az XBRL és iXBRL jelentések elkészítésére szolgál. A tanúsítvány megszerzéséhez az ilyen szoftvereknek hitelesíteniük kell a létrehozott jelentéseket, erre a célra alkalmas hitelesítő feldolgozóval.
- A **Review & Consumption Software**-t az XBRL vagy iXBRL-jelentések felülvizsgálatára, elemzésére, auditálására, vagy egyéb módon történő felhasználására használják. A tanúsítás megszerzéséhez az ilyen szoftvereknek minősített hitelesítő feldolgozóval kell validálniuk a jelentéseket, ezáltal az eszközök nagy része nem minősíthető tanúsítottnak.



**1. ábra:** Az aktív weboldallal rendelkező XBRL szoftverek típusai és száma (saját szerkesztés)

Az XBRL-hez kapcsolódó tudományos kutatások szempontjából a Scopus adatbázisban körülbelül 820 publikáció szerepel a témával kapcsolatban, amelyek időbeli eloszlását a 2. ábra mutatja be. A leggyakoribb dokumentumtípus-kategóriák a tudományos folyóiratcikkek (479), a konferencia-előadások (232) és a könyvfejezetek (56); a legjelentősebb Scopus által besorolt szakterületek az üzleti, menedzsment és számvitel (450), az informatika (393) és a közgazdaságtan, ökonometria és pénzügy (223).



*Megjegyzés: A keresési kifejezés a "TITLE-ABS-KEY ("XBRL" OR "Digital Financial Reporting" OR "Internet Financial Reporting" OR "Web financial reporting" OR "Digital Corporate Reporting")"*

**2. ábra:** Scopus keresési eredmények (saját szerkesztés)

### 3. Módszertan

Az elérhető eszközök felmérése alapján következtetésünk volt, hogy az ökoszisztéma korától függetlenül még mindig hiányzik a megfelelő infrastruktúra a valódi rendszerszintű kutatáshoz. A helyzet összetettségének kezelése érdekében a kijelölt adatkezelési rendszer fejlesztési szempontja, hogy magát az XBRL-szabvány megközelítését kövesse. Célunk, hogy az XBRL-jelentések feldolgozásának egyes feladataira összpontosító, nagy teljesítményű, egymással kommunikáló, moduláris egységek rendszerét hozzuk létre a következő alaprétegekkel:

- A rendszer alsó szintje egy tudásbázis, amely nem tartalmazza a taxonómiák szemantikai információit, hanem a jelentések és taxonómiai definíciók betöltésekor, futás közben "tanulja meg" azokat. Ugyanaz a tudásbázis a tényleges forrástól függetlenül homogén formában tartalmazza a jelentéseket, és egyszerű API-t tesz lehetővé a feldolgozáshoz.
- Minden szabványos XBRL formátumforráshoz (CSV, JSON, XML, XHTML) egy *parser* (olvasó) modult hozunk létre, ami az adott formátumot tudja olvasni, valamint egy *loader* (betöltő) modult, ami elvégzi a homogenizálást, és a tudásbázis API-ját használja a jelentés tárolására. Ez lehetővé teszi egyetlen loader létrehozását, ami bármely bemeneti fájlhoz aktiválja a megfelelő parser-t, és biztosítja a megfelelő hibakezeléssel ellátott, ellenőrzött importot.
- Az XBRL jelentések különböző forrásokból származhatnak, ezért egy konfigurálható és testreszabható forráscsatlakozókkal rendelkező architektúrát építünk fel. Ennek feladata a kapcsolatfelvétel az adott modulokkal, és az aktuális jelentéslista elérése (esetleges automatikus frissítési funkcióval), valamint szükség esetén a tényleges jelentések letöltése és a betöltő aktiválása.
- Ezt az megközelítést követve szükségünk van egy olyan azonosítási és tárolási mechanizmusra, amely lehetővé teszi az új források későbbi integrálását belső hozzáférési problémák nélkül. A platform megvalósíthatóságának validálásához reális becsléssel kell rendelkezniünk a szükséges tárolási és teljesítményigényről.

Ezáltal megteremthetjük egy globálisan támogatott fejlesztői és elemző környezet alapjait, megbízható, jellegükben nem változó szolgáltatásokkal, hogy a platformnak ne kelljen követnie a külső változásokat. A Java környezetet és eszközköszisztémát választottuk, mert ez bizonyult a legmegfelelőbbnek az évtizedek óta nagy jelentőséggel bíró üzleti környezetekben használt, időtálló szolgáltatások biztosítására, mint a beépített XML-elemzés, JSON-feldolgozás, Office-fájlok létrehozása, webszerver vagy akár a szkriptintegráció.

A fejlesztés során a Tracer Bullet Development módszertant követtük (Thomas & Hunt, 2019), ami alapján a mérföldköveket jól meghatározott célok és munkatermékek alapján határoztuk meg az erőforrások és az idő legoptimálisabb kihasználása érdekében. Ezen kívül igyekszünk külső erőforrásokra támaszkodni, ha azok rendelkezésre állnak, és az iteratív fejlesztés érdekében gyors prototípusfejlesztést és visszajelzést végzünk.

## **4. Eredmények**

Az eredmények eléréséhez köszönettel tartozunk azoknak, akik a munkát kívülről támogatták. Az XBRL tudásbázis platformot hosszabb távon egy belső virtuális kiszolgálón szeretnénk üzemeltetni. Az Egyetemi Informatikai Központ munkatársai, Csábi Béla és Kuti Zoltán kérésünkre gyors segítséget nyújtott a milab-srv.sze.hu szerver létrehozásában, amin az első teszteket sikeresen le is futtattuk. A projekt management és verziókövetés kérdéseivel Dr. Horváth Ernőt kerestük meg, támogatásával a Járműipari Kutatóközpont GitHub szervezetén belül (<https://github.com/jkk-research>) létrehoztuk a szükséges forrás repository-t és projektet. A kutatás során kiemelt jelentőségű a megfelelő mennyiségű és megbízható forrás és adat, amit a <https://filings.xbrl.org> portálról töltünk le. A kapcsolódó kérdéseinkre Paul Warren technikai igazgató gyors és pontos válaszokat adott, nagyban megkönnyítve a munkánkat.

### **4.1. Üzleti jelentések adatforrásai**

A független nemzeti nyomonkövetési mechanizmusok (Officially Appointed Mechanism, OAM) intézményét az Európa-szerte a szabályozott pénzügyi információk nemzeti adatbázisainak leírására használják. Az OAM-ok az Európai Unió minden tagállamában megtalálhatók, és a vállalatok számára a pénzügyi beszámolóknak a szabályozó hatósághoz történő benyújtására szolgáló fogadó-, és adatkezelő szervezetként szolgálnak.

Korábban a legtöbb OAM különböző típusú fájlformátumokat engedélyezett a bejelentésekhez, amelyeket általában PDF, szöveges vagy HTML formátumban tároltak. A pénzügyi információkon kívül álló, egyéb szabályozott információkat egyik OAM sem tárolta strukturált formátumban. A legtöbb OAM-adatbázis nyitott a különböző fájlformátumok számára, így opcionálisan az XBRL-fájlok is tárolhatók voltak (CESR, 2010). Ez azonban az Európai Értékpapír-piaci Hatóság (ESMA) hatálya alá tartozó jelentéstételi régiókban a jelenlegi szabályozással megváltozott, ahol az éves jelentések European Single Electronic Format (ESEF) néven ismert jelentéstételi formátumára vonatkozó kötelezettség van érvényben. Az üzleti adatok forrásának egyszerűsített útvonala tehát a tőzsdék; opcionális adatkezelő köztes szereplők (pl. külső jelentéskészítő); OAM-ok; végül az adattár (<https://filings.xbrl.org>); és a jövőben a bejelentések magasabb szintű, hivatalos adattára (European Single Access Point, vagy egyéb pán-európai OAM kezdeményezés). A jelenleg elérhető OAM-ok listáját az 1. táblázat mutatja be.

**1. táblázat:** A szabványosított XBRL-formátumok OAM-forrásainak top 10+1-es listája a bejelentések száma szerint (saját szerkesztés)

Adatrögzítési forrás	Ország	# Fájlok
Financial Conduct Authority (FCA)	GB	988
Polska Agencja Prasowa SA (PAP BIZNES)	PL	709
Finansinspektionen (FI)	SE	702
Virk	DK	627
Nasdaq	FI	597
Info-financiere.fr	FR	538
Oslo Børs	NO	497
Financial Services and Markets Authority (FSMA)	BE	371
Comisión Nacional del Mercado de Valores (CNMV)	ES	306
Oesterreichische Kontrollbank AG (OeKB)	AT	304
...		
Magyar Nemzeti Bank (MNB)	HU	51

A további megfigyelések során a [filings.xbrl.org](https://filings.xbrl.org) adattárt használjuk, amely a legtöbb uniós országból a nemzeti OAM segítségével gyűjti össze az XBRL-szabványos bejelentéseket (a részletekért lásd: <https://filings.xbrl.org/about.html> és <https://filings.xbrl.org/source>). Az adatforrás előnye, hogy a feldolgozott rekordokhoz szabványos JSON:API interfészt biztosítanak. Ezáltal a jelentések folyamatosan bővülő indexe könnyen elérhető kutatási célokra a tényleges jelentéscsomagokra mutató linkekkel, amelyek helyi elemzés céljából le is tölthetők. Kiemelendő azonban, hogy a fejlesztés alatt álló módszertan ettől eltérő adatforrást is támogat.

#### **4.2. Üzleti adatokra vonatkozó feltételek**

Az hatály alá eső európai vállalatokra vonatkozóan az ESEF által meghatározott jelenlegi különböző technikai érvényesítések a következők (XBRL International, 2020):

- A vállalatnak meg kell erősítenie a jelenlegi kibocsátói státuszát;
- Csak XHTML és ZIP fájlok engedélyezettek;
- A MIME-típusnak (Multipurpose Internet Mail Extensions) meg kell felelnie a fájl kiterjesztésének és tartalmának;
- A ZIP-fájlon belüli fájlneveknek és könyvtárneveknek meg kell felelniük bizonyos korlátozásoknak (különleges karakterek nélkül, legfeljebb 128 karakteres teljes hosszúsággal).
- Nincs aktív tartalom (pl. JavaScript, VBScript, Flash);
- Vírusirtó-ellenőrzés;
- Az XHTML-fájlokban hivatkozott összes objektumnak a ZIP-archívumban kell szerepelnie;
- A ZIP-archívumban szereplő összes fájlra legalább egyszer hivatkozni kell egy XHTML-fájlban;

A `filings.xbrl.org` a következő szolgáltatásokat biztosítja minden bejelentéshez:

- A bejelentés egy Inline XBRL megjelenítőben jelenik meg, ami lehetővé teszi a felhasználók számára, hogy a jelentésen belül böngészhessenek a megjelölt adatok között;
- A fenti forrásformátumokból az adatok transzformációjának biztosítása xBRL-JSON formátumban, amely az XBRL-adatok egyszerű, JSON-alapú megjelenítése, és amely számos eszközzel felhasználható és elemezhető;
- Az OAM-hoz benyújtott XBRL jelentéscsomag egy példánya.

A három szabványos fájlformátum kezelése a követelmények miatt így átjárható, az XBRL több rétegei alapján (dimenziókezelés) ezek között lehetséges a konverzió. Az "ideális" XBRL struktúra csak az elsődleges fájlformátumnak tekintett `.xml/.xbrl` fájlokban jelenik meg, amelyből a metaadatok kinyerhetők a tudásbázis kialakítására és a további API kommunikációra.

A `.html/.xhtml` és `.json` fájlokban az XBRL szabvány az eredeti struktúrát az adott formátumnak megfelelő változattá alakítja át (az előbbi esetében speciális iXBRL címkékké, amelyek a feldolgozó szoftverkörnyezetben vizuálisan megjeleníti az információ tartalmát, az utóbbiban pedig egy egyszerűsített struktúrává). Az szoftverarchitektúra részét alkotó `xhtml`- és `json`-kezelők megfordítják ezt az átalakítást, majd végrehajtják az API-hívásokat.

Az XBRL-ökoszisztémában feltételezhetjük, hogy a bemeneti formátumok valamelyikének megfelelően jól formázott bemenetéről van szó, tehát a parser/handler párok minden formátumot lefednek, és ezekhez kötődően egy loader-elem elég az összes szabványos XBRL jelentés feldolgozásához, függetlenül a forrástól és a benne szereplő taxonómiáktól.

## 5. Összegzés

A tanulmány egy új, rugalmas és skálázható szoftverarchitektúra alapját ismerteti, amely az eXtensible Business Reporting Language (XBRL) szabvány alapján készült európai üzleti adatok automatizált kinyerésére és kezelésére szolgál. Az architektúra kifejlesztése az európai uniós szabályozások által előírt pénzügyi jelentések digitalizációjából fakadó új kihívásokra

reagál, amelyek a tudományos közösség és az üzleti szféra számára egyaránt jelentenek technikai akadályokat az adatok hozzáférhetősége és értelmezése terén.

A tervezett rendszer magában foglalja az adatlekérdezést, feldolgozást és tárolást támogató modulokat, amelyek a technológiai adaptálhatóságot és bővíthetőséget szem előtt tartva lettek kialakítva. A moduláris komponensek lehetővé teszik a különböző adatforrásokból származó információk integrálását és kezelését, növelve ezzel az üzleti jelentések elemzésének hatékonyságát és a tudományos kutatások támogatását.

## **6. Köszönetnyilvánítás**

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

## **Irodalomjegyzék**

- [1.] Ball, R. (2008). What is the actual economic role of financial reporting?. *Accounting Horizons*, 22(4), 427-432.
- [2.] CESR (July, 2010). Development of Pan-European Access to Financial Information Disclosed by Listed Companies, Consultation paper, Committee of European Securities Regulators
- [3.] Deloitte, T. T. L. (2021). International Sustainability Standards Board (ISSB). Retrieved from <https://www.iasplus.com/en/resources/ifrsf/issb>
- [4.] ESMA. (2022). European Single Electronic Format. European Securities and Markets Authority.
- [5.] IFRS Foundation. (2022). Climate-related Disclosures—IFRS S2. Retrieved from <https://www.ifrs.org/projects/work-plan/climate-related-disclosures>
- [6.] Lennard, A. (2007). Stewardship and the objectives of financial statements: a comment on IASB's preliminary views on an improved conceptual framework for financial reporting:

the objective of financial reporting and qualitative characteristics of decision-useful financial reporting information. *Accounting in Europe*, 4(1), 51-66.

[7.] Lombardi, R., & Secundo, G. (2021). The digital transformation of corporate reporting—a systematic literature review and avenues for future research. *Meditari Accountancy Research*, 29(5), 1179-1208.

[8.] Thomas, D., & Hunt, A. (2019). *The pragmatic programmer*. Addison-Wesley Professional.

[9.] XBRL International (2020). <https://www.xbrleurope.org/wp-content/uploads/2020/02/3-24hXBRLEuropeDayRoleOAMonESEF-Krzysztof-Jesse-Ulrich-Brass-06-Feb-2020.pdf>

[10.] XBRL International (2022). About [filings.xbrl.org](https://filings.xbrl.org). Retrieved from <https://filings.xbrl.org/about.html>

# Utánzás alapú gépi látással támogatott automata robotprogram generálás

## Mimic-based machine vision assisted automatic robot program generation

Tóth Patrik<sup>a</sup>, Takács Patrik<sup>b</sup>, Monek Gergő Dávid<sup>c</sup>

<sup>a</sup>Széchenyi István Egyetem

tpatrik0905@gmail.com

<sup>b</sup>Széchenyi István Egyetem

patriktakacs00@gmail.com

<sup>c</sup>Széchenyi István Egyetem

monek.gergo@sze.hu

### Absztrakt

---

*A kollaboratív robotok az iparban számos módon alkalmazhatók, technológiájukból adódóan alkalmasak automatizálási megoldásokra, együttműködő jellegükből fakadóan pedig manuális, emberi munkát igénylő környezetben is egyszerűen telepíthetők. Alkalmazásukkal csökkenthetők a gyártási költségek és növelhető a termelékenység. Az ömlesztett áru kommissiózás (bin picking) egy vizuális alapú anyagmozgatás (pick and place), amely napjaink két legjellemzőbb automatizálási technológiájára épül, a gépi látásra és a robotizációra. A gépi látás robotokkal való kombinációja új automatizálási lehetőségeket nyitott meg, A szenzorokkal vagy kamerával felszerelt robotok a véletlenszerűen elhelyezett tárgyak felismerésére és mozgatására lettek képesek. Ennek a cikknek a célja egy kollaboratív robotos, utánzást végrehajtó, szakértő robot programozót nem igénylő digitális képfeldolgozás alapú demonstrációs szortírozó program generálása. A folyamat során színben és alakban is különböző munkadarabok felismerése majd szortírozása valósul meg, az előre definiált paraméterek függvényében. A munkadarabok munkaterületen való elhelyezése során videófelvétel készül, amely a digitális képfeldolgozás segítségével képkockánként feldolgozásra kerül így meghatározva a munkadarab beérkezési sorrendet és a geometriát. A következő lépésben az egyes munkadarabok középponti koordinátái és irányultsága kerül meghatározásra, amely információk alapján megtörténik az automata robotprogram generálás. Ennek eredményeképpen a különböző geometriákhoz előzőleg definiált tárolókba a munkadarabok beadási sorrendjében a kollaboratív robot végrehajtja a szortírozást.*

**Kulcsszavak:** kollaboráció, robotprogram, ipari automatizálás, ipar 4.0

## Abstract

---

*Collaborative robots have a wide range of industrial applications, from automation solutions due to their technology, to simple deployment in manned environments due to their collaborative nature. They can reduce manufacturing costs and increase productivity. Bin picking is a visual-based pick and place solution based on two of today's most common automation technologies, machine vision and robotisation. The combination of machine vision with robots has opened up new automation possibilities, with robots equipped with sensors or cameras being able to detect and move randomly placed objects. The aim of this paper is to generate a collaborative robotic mimicking demonstration sorting program based on digital image processing that does not require an expert robot programmer. The process involves the recognition and sorting of workpieces of different colours and shapes, depending on predefined parameters. During the workpiece placement on the work area, a video image is captured and processed frame by frame using digital image processing to determine the workpiece arrival sequence and geometry. In the next step, the centroid coordinates and orientation of each workpiece are determined, which is used to generate the automatic robot program. As a result, the collaborative robot performs sorting in the order in which the workpieces are entered into the bins previously defined for the different geometries. goes the abstract of the conference paper.*

**Keywords:** *collaboration, robotprogram generation, industrial automation, industry 4.0*

## 1. Bevezetés

Napjainkban a kollaboratív robotok nagy szerepet játszanak az emberi élet minden területén, mert alacsony költségű, magas színvonalú és gyorsan előállított termékek iránt egyre növekszik a kereslet. Ez a munkavállalók számára is új helyzetet teremt, mivel ezek a robotok különösen előnyösek lehetnek, veszélyes munkakörnyezetekben, vagy nehéz környezeti feltételek mellett az ember közelében alkalmazva. Annak érdekében, hogy megértsük a fogalmát és a jelentőségét a kollaboratív robotoknak, fontos megérteni az ipari robotok és környezetük folyamatainak működését. Az egyedi, személyre szabott termékek iránti kereslet egyre inkább növekszik. Ez arra kényszeríti a gyártókat, hogy a mennyiségről a változatosságra helyezték át a hangsúlyt a versenyképességük megőrzése érdekében. Az Ipar 4.0 koncepcióját éppen az ipari környezet ilyesfajta megváltoztatására irányuló igények idézték elő, segítve a vállalatok modern globális piachoz való felzárkóztatását. Az ipar 4.0 a termelési folyamatok olyan hatékony szervezését írja le, melynek keretében az eszközök önállóan kommunikálnak, és összehangoltan működnek. Egyre több gyártóvállalat használ az automatizációhoz kollaboratív robotokat, amelyek akár egyéb eszközökkel együttműködve rugalmas, költséghatékony és biztonságos megoldást nyújtanak gyártási feladatok széles tartományában. Ilyen műveletek lehetnek például a hegesztés, összeszerelés, festés, csomagolás, palettázás, illetve bármilyen feldolgozási folyamat. A jövő gyáraiban a robotok és az emberek nem külön-külön, hanem együttműködésben, közös munkatérben fognak dolgozni, segítve egymás munkáját és a különböző típusú feladatok megfelelő allokációjának köszönhetően a hatékonyságot növelve. [1-5]

## 2. Irodalomkutatás

A robotok bevezetése óta a gyártás területén és az automatizált termelési eszközök fejlődésében mindig ott volt a vágy a teljes automatizálásra és az intelligencia bevezetésére a gyártási eljárások többségébe. A negyedik ipari forradalom tovább fokozta ezt a vágyat az intelligencia iránt és a tanulási képességek fejlesztésére a kollaboratív robotokban. A kollaboratív robotok intelligenciájának a fő bizonyítéka és az egyik legnehezebb probléma, ami megoldásra vár az ipari automatizáció területén az a véletlenszerűen elhelyezett, struktúrátlan alkatrészeknek a pontos észlelése és manipulálása. Különösképp az olyan esetekben mikor a különböző részek változatosak és többféle megfogási technikával kell megoldanunk a feladatot, ami még nagyobb

kihívást jelent. [6] Nagy mennyiségű kutatás és kísérlet történt az iparban történő alkatrész bin picking területén. A Tuan-Tang Le és Chyi Yeu Lin által kifejlesztett rendszerek több mint 99%-os sikerarányt érnek el azáltal, hogy szemantikai szegmentálást és mélységi adatokat használnak fel. Azonban ők csak pneumatikus fejeket használnak a feladat végrehajtására, ami korlátozza a rendszer képességeit. [7, 8] Egy másik, sokoldalúbb rendszert a 2017-es Amazon Robotics Challenge során lett kifejlesztve Andy Zeng által. Ez a rendszer egy multifunkcionális grippert használ, ami lehetővé teszi, hogy a robotkar pontosabb megfogást hajtson végre. Ráadásul ez egy eléggé gyors rendszer,  $0.05 * n$  másodperc feldolgozási sebességgel rendelkezik, ahol  $n$  az a lehetséges megfogási szögek számát jelöli. A rendszer nagy sebességének kárára azonban kisebb, 89%-os pontosságot tudtak csak elérni. A rendszer deep learning használatával elemzi ki azt, hogy milyen megfogási módot válasszon a siker érdekében. [9] A teljes mozgásgeneráló architektúra három részből áll. Ezek a következők: érzékelés, készségek és viselkedés. Az emberi kéz hatékonyságának és rugalmasságának reprodukálása a tárgymanipulációs feladatoknál fontos kihívás. Ehhez alapvetően újra kell gondolni, hogy hogyan lehet kihasználni a számos szenzorból érkező adatainkat és a rendelkezésre álló robotkarnak a mechanikus kézügyességét. Az emberhez vagy főemlőshöz képest a mai robotmegfogók és kezek kézügyessége rendkívül korlátozott. Az érzékelést négy kategóriába sorolják, mint az érzékelő nélküli manipuláció, a képérzékelés, az aktív érzékelés és az interaktív érzékelés. Az érzékelő nélküli manipuláció bár egyszerűbb feladatoknál működhet, azonban összetetteknél már sokszor nem elegendő. [10, 11] A vizuális észlelési megközelítésekben statikus képeket használnak arra, hogy a manipulációs feladatokhoz szükséges érzékelő visszajelzéseket létrehozzák, képeket használnak a fogástípusok és tárgytipusok közötti kapcsolat megértéséhez. [12] Az aktív észlelési megközelítésekben az érzékelő visszacsatolását úgy oldják meg, hogy az emberi figyelmet és tekintetet utánozzák. Ebben a környezetben egy RGB kamera van használva. Egyes esetekben úgy használják ezt a módszert, hogy a robot mozgásával a kamerát a jelenet körül mozgatják és így szegmentálják a felvételt és találják meg a tárgyakat a képen. Összehasonlítva a statikus kamerabeállítással, ez a módszer lehetővé teszi az objektumok jobb modellezését a több nézőpontból származó adatok által. [13] Az interaktív érzékelési megközelítéseket a fizikai interakció és a hagyományos érzékelési módszerek kombinálására fejlesztették ki. E kettő kombinációja szélesebb körű alkalmazásokat tesz lehetővé, például ismeretlen tárgyak manipulálásának megtanulását tárgytulajdonságok tanulását. Ha az érzékelés által a szükséges paraméterekhez hozzájutunk akkor a következő lépés jön. Az objektumok észlelése és azonosítása a megjelenésük és

elhelyezkedésük alapján történik. Objektumkövetés történhet videó alapú megfigyeléssel is, ahol a kamera feladata olyan nyomkövetési stratégiákon alapul, mint a régió alapú és az aktív kontúr alapú eljárás. Az objektum alapú képelemzést olyan földrajzi információs rendszerekben alkalmazzák, amelyek képesek pixeleket úgy elemezni, hogy a képekből könnyen használható objektumok kerülnek letárolásra. Fényintenzitás-elemzésnél felületdetektálásnál alkalmazható. Az objektuminformációk automatikusan az objektumról készített képből vannak származtatva, például az objektum lokalizációjának felismeréséhez, míg a megfogáshoz az objektum súlypontja kerül felhasználásra. [14-22]

### 3. Automata szortírozási folyamat generálás

A fejlesztés célja egy olyan automata folyamat létrehozása kollaboratív munkakörnyezetben, ahol a demonstrációs eszköz inhomogén munkadarabok felismerését és szortírozását végzi. A felhasználó tetszőleges sorrendben bepakolja az előre betanított formájú munkadarabokat a munkaterületre. A rendszer a digitális képfeldolgozás segítségével elvégzi a munkadarabok azonosítását, kategorizálását és az előzetesen definiált, de dinamikusan változtatható paraméterek (munkadarab kontúrhoz és színhez rendelt tároló) függvényében robot programot generál. Majd a munkadarabok bepakolásának sorrendjében elvégzi a szortírozást.

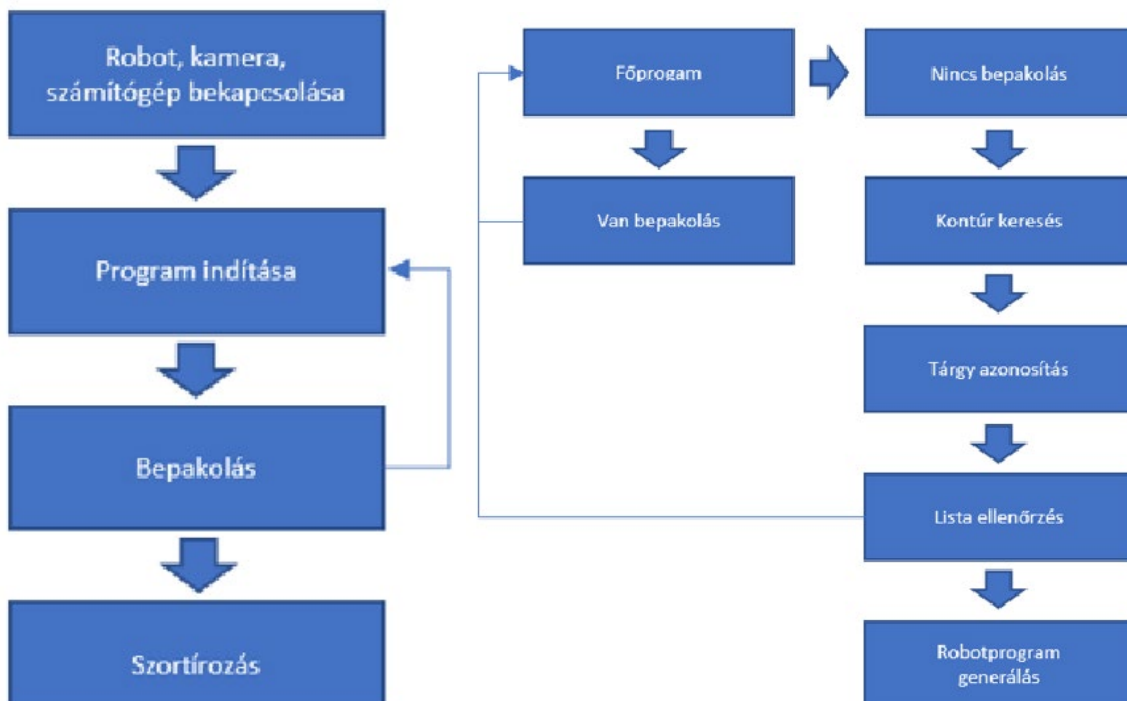


**1. ábra:** Tesztkörnyezet (saját szerkesztés)

A fejlesztés megvalósítási helyszíne a Széchenyi István Egyetem Kiber-Fizikai Gyártórendszerek Laboratóriuma volt. A felhasznált mintakörnyezet, ami a laboratóriumon belül az egyik asztalon van a 1. ábrán látható, az előtérben a robottal. Alatta a fekete munkatér az, ami háttérként szolgál a képvételezéshez. Rajta pedig láthatóak azok a munkadarabok,

amiket a kamera képe alapján felismerve a robotkar bepakolást követően szortíroz a munkatérből egy adott pozícióba.

A program fejlesztését a 2. ábrán látható lépések alapján végeztük. A folyamat a robot, és a számítógép bekapcsolásával és a kamera aktiválásával kezdődik. A programot elindítva a felhasználó párosíthatja az elemeket a kívánt tárolókonténerekkel és elkezdheti a bepakolást. Egymás után tetszőleges sorrendben történhet meg az alkatrészek munkaterületen való elhelyezése. Eközben a rendszer videófelvétel által képes észlelni meddig tart a bepakolás, ez idő alatt még nem kezdi meg a tárgyak konkrét azonosítását, csak a sorrendjüket menti el. Amint a felhasználó végzett az utolsó elem elhelyezésével is, tehát a program már nem észlel mozgást, elkezdődik a videó utolsó képkockájának feldolgozása. A rendszer a munkaterületen található összes tárgyat képes egyszerre észlelni, azonosítani őket, kiszámolni a megfogáshoz szükséges középponti koordinátáikat és irányultságukat. Ezen információk alapján automata robotprogram generálása történik meg, amely lehetővé teszi, hogy a robotkar a megfogó segítségével FIFO elv szerint a kívánt helyre juttassa a munkadarabokat.



**2. ábra:** A robotprogram ciklusai (saját szerkesztés)

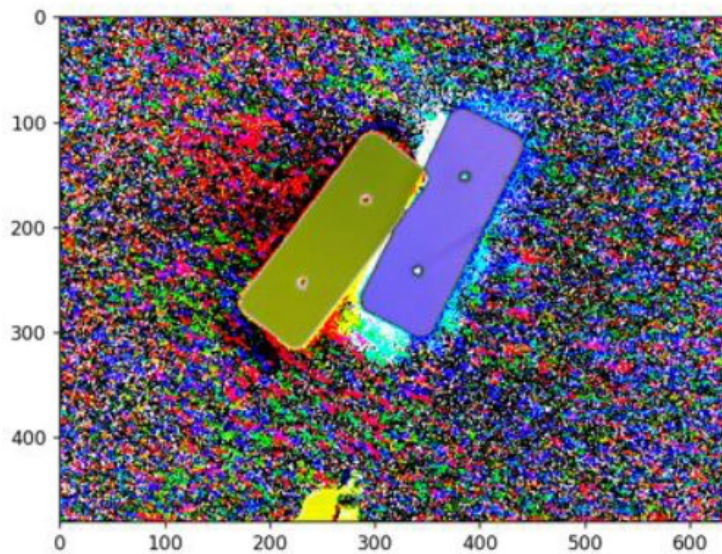
A folyamat részletesebb ábrája a 2. ábra jobb oldalán látható. A feladat során a robotkarral felszerelt kamerától kapja a program a videófelvételt a munkaterületről. Ha nincsen bepakolás – mozgás - a munkaterületen, akkor megkezdődhet az aktuális képkocka kielemezése. Digitális

képfeldolgozással megkeresésre kerülnek a kontúrok, majd ezeknek középponti koordinátáit és a szögelfordulásait egy listában a rendszer eltárolja a sorrend megjegyzéséhez. Ezt követően a kontúr alapján számolt különböző geometriai jellemzők alapján tárgyazonosítás történik. A rendszer képes a munkaterületen elszórtan elhelyezkedő összes tárgyat felismerni és pontosan besorolni a megfelelő kategóriába. Mindegyik csoporthoz rendelve van egy tárolópozíció, ahova az adott elemeket szállítani szükséges. Ezen információk átadásra kerülnek a robotnak, ami azok alapján a sorrend szerint következő alkatrészt biztonságos magassági szinten megközelíti. Ezután a kinyert adatok által kiszámítja milyen szögben kell elforgatnia a harmadik csuklóját, hogy a rajta található 2 ujjas gripperrel meg tudja fogni és a megfelelő helyre mozgatni a tárgyat.

A digitális képfeldolgozás megkezdése előtt fel kellett mérni, hogy vajon a feladathoz szükséges-e valós idejű folytonos videó felvétel elemzése, vagy pedig elégséges egy adott időközönként vételezett kép is. Mint tudjuk a videó is képek halmaza, amit többféle módon is lehet jellemezni, az egyik leginkább elterjedt jellemzője az az fps. Vagyis frame per second, magyarul a másodpercek alatt vételezett képkockák száma. Alapos megfontolás után arra jutottunk, hogy feladathoz valós idejű videó feldolgozást kell használni. Erre azért volt szükség, hogy elkerülhetővé váljanak azok a nemkívánt események elkerülései, mint például a termékek elhelyezésének pillanatai, amikor az azokat elhelyező személy keze is bekerülhet a felvételre. Ezt kiküszöbölve videófelvételezés módszerének alkalmazása került előtérbe egy, a mozgást érzékelni képes fejlesztett programmal. Ezt használva már megállapítható mikor van mozgás, ilyenkor nem elemezzük ki a képkockát. Először a robotkart el kell küldeni a mintavételezési ponthoz, ezt a program elején meghívott függvény teszi meg. Mivel a videót egy fix pontból készíti így ezeket a paramétereket fix értéként elmentve tudjuk tárolni és meghívni mikor szükségünk van rá. Miután a parancsot kiadtuk a felvételt már elkezdti ilyenkor a kamera, késleltetésre nincsen szükségünk hiszen amíg mozgásban van addig a program elején lévő mozgást érzékelő rutin nem engedi, hogy a fő programból tovább lépjen a feldolgozási részre a program.

Ezt a mozgást egy move parancs segítségével listában stringként (szöveggként) összefűzve adjuk meg a robotnak. A robotnak a koordinátákon kívül fontos megadni a sebesség és gyorsulás adatokat is, amiket a mozgáshoz használ, vagy pedig az időt, ami alatt végezze el az adott műveletet. A feladat során a sebesség értéknek 0.1 m/s míg a gyorsulásnak 1 m/s<sup>2</sup> értéket határoztunk meg. A videókapcsolat létrehozása után egyből veszünk egy frame mintát, amit

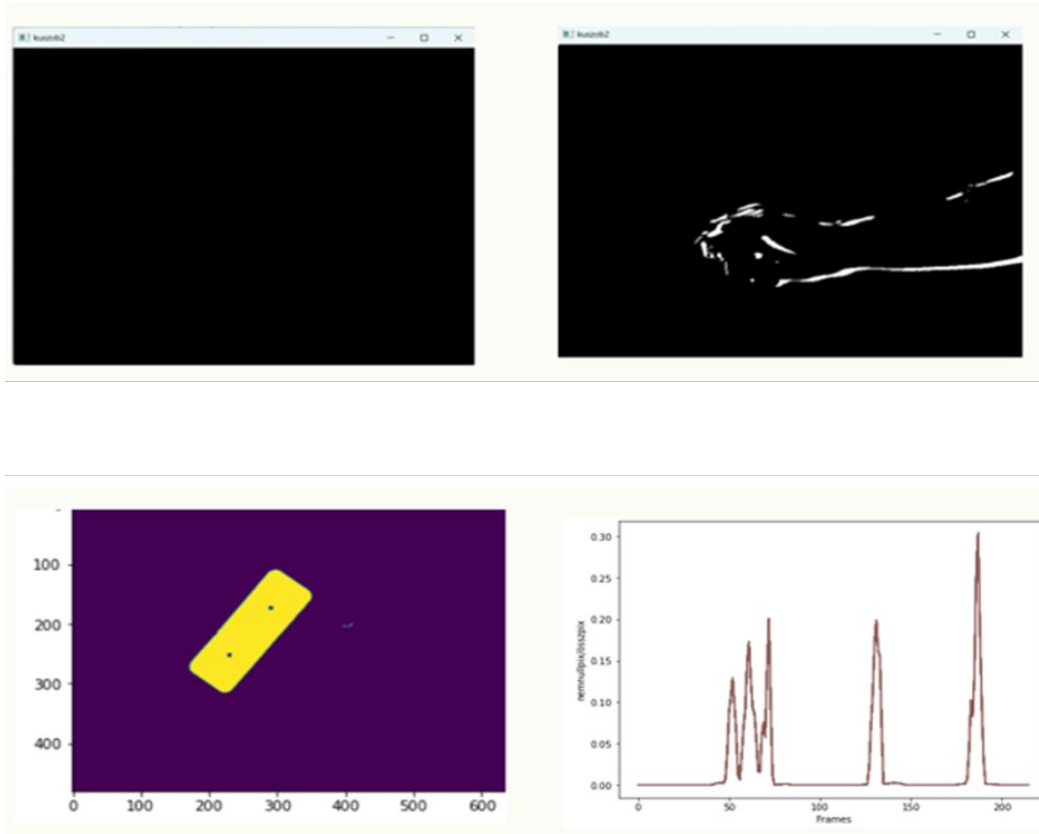
lastframe-nek nevezünk el (előző képkockának). Erre azért volt szükség, hogy a ciklusba való belépés után azonnal tudjunk venni egy aktuális képkockát és ezek különbségét képezni. Ez úgy néz ki, hogy az előző és a mostani képkockának az abszolút értékét egymásból kivonjuk és így megkapjuk a különbségüket. Azért fontos, hogy az abszolútértéküket vonjuk ki egymásból mivel, ha csak simán kivonnánk őket akkor könnyen alulcsordulhatnak a pixelek, ahogy az 3. ábrán is látható.



**3. ábra:** Alulcsordult kép (saját szerkesztés)

A képen látható sok kék, zöld, piros stb. pöttyök mind azt mutatják, hogy az adott pixelek azok alulcsordultak. A színes képek -az esetek döntő többségében- három színcsatornával rendelkeznek. Ez a három színcsatorna a tiszta vörös (Red), a zöld (Green) és a kék (Blue), egyben RGB. Az RGB színrendszer az opencv modulban fordított sorrendben van elmentve, programozásnál a BGR szerint kell hivatkoznunk rá. Ezen színcsatornák adatainak mentésére a képfeldolgozásban egyenként 1 bájt áll rendelkezésre. Vagyis egy pixelen belül 8 vörös, 8 zöld és 8 kék bit lesz, ami tudja menteni az adott komponens értékét. Ez minimálisan a 0, míg maximálisan a 255 érték felvételét biztosítja. Ha az előbb tárgyaltak szerint vesszük a példát, miszerint kivonunk két képet egymásból és választunk egy pixelt, aminek az egymásból kivont értékei nullánál kisebb negatív számot adnak nekünk vissza, akkor ezt hívhatjuk alulcsordulásnak. Ilyenkor a számláló csökkenés közben átlépve a nullát visszaugrik 255-re, majd onnan csökken tovább. Ezt úgy tudjuk kiküszöbölni, hogy abszolútértékét vesszük ennek a kivonásnak és ilyen módon a pixelünk egyes színcsatornái nem csordulnak alul. Ellenkező

esetben beszélhetünk túlsordulásról is. Alulcsordulás nélkül, abszolútértékkel a differencia a 4. ábrán látható. Ezek után a képkockák differenciáját szürkeárnyalatossá konvertáltuk át.



**4. ábra:** Változás detektálás folyamata (saját szerkesztés)

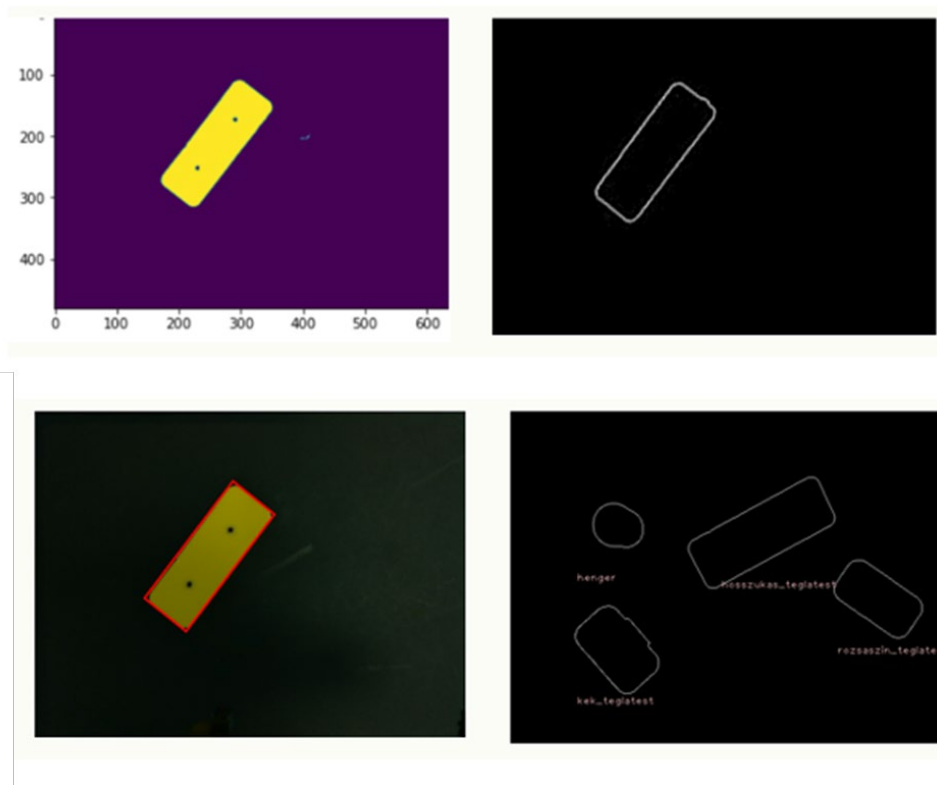
A főprogramban ezután egy változót definiálunk, melyben a számláló helyén a küszöbölt képkockák nemnulla elemeinek száma van, a nevezőben pedig a kép összes pixele. A kép összes pixelét úgy számítottuk ki, hogy beszoroztuk az aktuális frame alakját meghatározó oszlopok és sorok számát, míg a képen lévő nemnulla (vagyis nem fekete, mivel a teljesen fekete pixel az 0 értékű) értékeket parancs segítségével ugyancsak az éppen aktuális frame-ből számítottuk. Ezek hányadosa egy olyan értéket ad vissza számunkra, mely akkor változik mikor valami mozgást érzékel a kamera. Amint a mozgás abbamarad akkor az arányunk változása is lecseng, ahogyan azt a 4. ábrán is láthatjuk. Ennek az érzékenységét változtathatjuk a küszöbérték változtatásával. A programot úgy írtuk, hogy ameddig mozgást érzékel a program addig a fő program hurkában maradunk, nem hívjuk meg a függvényt, ami a képeket feldolgozná.

A függvény az adott képen elvégez egy maszkolást. A küszöbölés a HSV rendszerben történik, S-re és V-re szűrjük a különböző munkadarabokat, megvizsgálásra kerül az összes pixel S~Saturation és V~Value értéke, amelyek a telítettségéről és fényességéről szolgáltatnak

információkat. Ennek megvalósításához több módszer is használható. Az egyik ezek közül a logikai műveletek alkalmazása. Ennél az adott kép pixeleinek valamelyik kiválasztott értékei kerülnek összehasonlításra egy általunk meghatározott számmal. A kritériumnak megfelelő képpontok logikai igaz, tehát 1-es értéket kapnak, a többiek pedig 0-át. A következő lehetőség a threshold parancs használata: működésében hasonló az előzőhöz, csak még lehetőség van az argumentumában megadni azt is, hogy kapásból mennyivel szorzódjanak meg a pixelértékek. Általában ez úgy van megválasztva, hogy 0 és 255 értékű képpontok álljanak elő, amely annyi előnyt jelent, hogy már a művelet közben láthatóvá válik a maszkolás eredménye. Ennél az utasításnál Otsu-módszer is alkalmazható, amelynél nem kell manuálisan megadni egy küszöbértéket, hanem a program megpróbálja automatán felismerni azt.

A program végén még szükség van arra, hogy meghatározzuk a szögét, melyet a munkadarabok az elhelyezésükkor bezárnak a vízszintes tengellyel, illetve a munkadarabok középponti koordinátáit. Erre azért van szükség, hogy a megfogó pofái a megfelelő – a munkadarabbal párhuzamos – állásba kerüljenek és így fel tudja azt venni. Emellett fontos még, hogy a mozgatni kívánt alakzatoknak, ha van hosszabb oldaluk (nem kockák) akkor azoknál fogjuk meg a gripperrel. A kontúrkeresésnél a talált kontúrokat kigyűjti a program egy változóban, amit utána vissza tudunk rajzoltatni az eredeti képre ellenőrzésképpen. Azzal a problémával szembesültünk, hogy sok nulla pixeles hiba van a képeken és a program ezeket találta meg és rajzolta körbe, majd értelmezte kontúrként. Ezután pedig a kiértékelésnél az adott pixeleknek a szögét, illetve középponti koordinátáját kaptuk eredményként. Ezt a problémát úgy oldottuk meg, hogy a kontúrok alapján kiszámítottuk a területet, amit behatárol. Ha ez megvan akkor feltételhez kötjük, hogy egy előre megadott méretnél nagyobbak kell lennie a területnek vagy nem megfelelő kontúrról van szó. Ez úgy történik, hogy a kontúrokon végig iterálva azokat a kontúrokat melyek a megadott méret feltételnek nem adnak eleget átugorja a program. Emellett a kép javítása érdekében ennél a pontnál használtunk morfológiai műveleteket, melyek segítségével kiszűrtük a zajos régiókat. Ilyen egyszerű morfológiai műveletek például az erózió, illetve a dilatáció. Az erózió során a struktúráló elem tagjainak a minimumát számoljuk ki, míg dilatációnál a struktúráló elemek maximumát. Ezeknek kombinációból felépítve létezik még rengeteg más művelet. A dilatáció struktúrálóelem tagjainak a maximumának vételével egyenlő. Ha elképzelünk egy egyszerű struktúráló elemet, mondjuk egy 3x3-as négyzetet csupa egyesekkel. Ha ez egy morfológiai elemnek a struktúráló eleme akkor azt írja le, hogy saját maga és a szomszédjai – esetünkben az öt körülvevő 9 elem közül – a maximummal

helyettesítjük az adott pixel értékét. Ezt úgy is megfogalmazhatjuk, hogy a „fehér” területeket növeljük általa. Ha például mi egy tiszta fekete és fehér képről beszélünk akkor ez azt jelenti, hogy ha a pixelnek a környezetében találunk egy fehér pixelt akkor a pixel fehér lesz a transzformáció után. Az erózió ugyan ezt jelenti, csak abban az esetben a minimumot ellenőrzi. Példánknál maradva egy teljesen fekete és fehér kép esetében ez a módszer a „sötét” területeket fogja növelni. Ha adott pixel környezetében van egy fekete pixel akkor azzal lesz helyettesítve. az erózió, illetve dilatació használatával a maszkolt képünkön a kisebb zajos régiók eltávolítása sikeresen megtörtént, melyet láthatunk is a 5. ábrán. Emellett a méretbeli szűrést alkalmazva a program már megfelelően képes érzékelni a kontúrokat.



**5. ábra:** Kontúrkeresés, alakzatfelismerés (saját szerkesztés)

A kontúrok meghatározását követően a minarearect függvény került meghívásra. Ezzel a függvénnyel képesek vagyunk a kontúrok felhasználásával meghatározni a legkisebb köréírható négyzetet. Ezután a négy sarokpontot a parancs meghatározza és megszámozza őket. Az alapján, hogy melyiknek van a legnagyobb „y” koordinátája, az lesz a nulladik pont, abban az esetben, ha két pont van ugyan olyan „y” értékkel akkor a jobb oldali lesz a nulladik számú. Az „y” koordináta a képfeldolgozásban úgy helyezkedik el, hogy bal felső sarokból növekszik

lefelé, míg az „x” koordináta ugyan onnan csak jobbra növekszik. Ha megvan a legnagyobb „y” akkor ezután a sarokpontok számozása az óramutató járásával megegyezően történik. Ennek segítségével pedig képes a függvény kiszámítani az objektum által bezárt szöget. Azonban a szög, amit visszakaptunk az mindig a nulla és egyes pontok közötti volt. Viszont nem volt megfelelő minden esetben ez a szög, hiszen, ha a munkadarabnak a rövidebb oldalainál lévő szöget mérte meg a program akkor a rövidebb oldal mentén fogta volna meg a gripper, ami nem minden esetben lenne biztonságos. Ezért a programot kiegészítve beleírtunk egy oldalhossz mérő függvényt, ami a meghatározott nulla-egy és nulla-három pontok között számítja ki a távolságot. Ezt használva pedig eldönthető, hogy egyes esetekben melyik oldal hol található. Miután ez megvolt feltételeket szabtuk és aszerint, hogy melyik oldal a hosszabb kikötöttük milyen korrekció kell, hogy megfelelő oldalak mentén fogja meg a munkadarabokat. Ezt 90 fok kivonásával értük el. Emellett beleírtuk a feltételek közé, ha a kamera éppen 0 vagy 90 fokot detektál akkor szintén az oldalhosszúságok alapján maradjon annyi vagy forduljon el. Ezeket az értékeket átváltva radiánba a robotkar értelmezni tudja és be tudja fogadni azokat. A minarearect függvényből még ki tudjuk venni a középponti koordinátákat is, amikkel a megfelelő pozícióba tudjuk mozgatni a robotkart. Ehhez szükség volt egy pontra, amihez viszonyítva hozzászámoljuk a távolságot, amit a fotóról kijelez számunkra a program. Ez a kitüntetett pont a kép sarkában lett meghatározva. A képen mérhető pixel értékeket, amit a program x és y koordinátákban megad átváltottuk méterbe, hiszen a robotkarnak abban kell megadni a távolságot. Ezután pedig hozzáadtuk az értékeket a fix koordinátaponthoz. Olyan módon oldottuk meg ezt, hogy a robot tengelye szerinti x koordinátához adtuk a kép szerinti y tengely átváltott pixel értékeit. A másik robot szerinti y tengelyhez pedig a kép szerinti x értékeket adtuk hozzá. Amiatt kellett keresztbe hozzáadni mivel a kamera abban a helyzetben ahogyan használtuk az ellenkező koordinátatengelyei mentén volt párhuzamos a robotéval.

A feladat ezen részénél a tesztek alatt előjött az a probléma, hogy nem érzékel több kontúrt a program, mivel nem tudja mi az, amit már megtalált és mi az, amit még nem. A program úgy működik, hogy a kamera „y” tengelye mentén halad lefelé az aktuális képnek a kiértékelésekor. Ennél fogva, ha ezen tengely mentén növekvő irányba helyeztük be a munkadaraboknak már most is észlelte a kontúrjait, azonban nekünk az a célunk, hogy bármilyen helyre történő egymás utáni bepakolásnál is megtalálja az új kontúrjainkat. Az alakzat ellenőrző program felelős azért, hogy megvizsgálja szerepelt-e már az adott kontúr előzőleg, ezt olyan módon teszi, hogy elmenti a talált kontúrnak a középponti koordinátáit és elfordulásának szögét és mindig az

aktuálisan kapott új érték alapján megvizsgálja (adott tűréssel), hogy a már elmentett értékekkel egyezik-e az. A függvényben határértékek között kell kezelni a méreteket, hisz mind a képfeldolgozással történő mérésben, mind az élethelyzetekből adódóan – odébb lökünk bepakolásnál egy munkadarabot – lehetnek olyan hibák melyekre nem számítunk. A kisebb elmozdulásokat melyek a mérésből és kisebb lökésből adódnak tudja kezelni a program, de a nagyobb elmozdulásokat új munkadarabnak veszi és újként emeli be a listába. Mindezek után a program első felének a felépítése úgy néz ki, hogy a főprogramba meghívjuk a kontúr kereső függvényt mikor nincsen mozgás. Ebben a függvényben kontúrokat keresünk a képen és azok szögelfordulásait. Ebbe a függvénybe, ha talált kontúrt, meghívjuk az alakzat ellenőrző függvényt, amivel azt figyeljük volt-e már az adott kontúr az előzőekben, és ha igen akkor kerestetünk újabb kontúrt. Ha talál új kontúrt akkor annak középponti koordinátaival és szögelfordulás értékével megy tovább a program, ha nem, akkor üres listával tér vissza. Üres lista visszatérése esetében a főprogramba visszatérve kihagyja a lista ellenőrzést és a while hurok elejére ugrik, ezzel elkezdve a következő képkocka feldolgozását. Ha viszont koordinátákkal és elfordulási szöggel tér vissza akkor azt a sorrendi listába menti, hogy azt később a szortírozási szakaszban fel lehessen használni. Ezután tér csak vissza a főprogramba és kezdi újra a mozgásnak a vizsgálatát majd elemzi ki az elkövetkező képkockákat.

#### **4. Összegzés**

A cikkben bemutatott kutatás és fejlesztés célkitűzése egy olyan gépi látáson alapuló automata demonstrációs eszköz megalkotása volt, mely egy szortírozási feladatot képes megvalósítani inhomogén munkadarabokon úgy, hogy a szortírozási feladat egyes paramétereit, mint például a sorrendiséget és azt, hogy melyik munkadarab melyik tárolóba kerüljön egy utánzás alapú program generálási módszerrel valósítsuk meg. Első lépésében a betanítás alapján a munkaterületen véletlenszerűen elhelyezett inhomogén alkatrészek kontúr és szín azonosítása majd osztályba sorolása történt meg. Ebben az esetben az előre betanított négyféle szint és kontúrt sikerült közel 100%-os hatékonysággal azonosítani. Második lépésben az egymás után a munkatérbe kerülő munkadarabok sorrendiségének azonosítása és rögzítése következett, ami a szortírozás utánzási alapelv szerinti végrehajtásának feltétele. Következő lépésben az automatikus robotprogram generálás valósult meg. Ennek feltétele az előző két pont teljesülése, valamint az egyes munkadarabok középponti koordinátáinak és irányultságának

meghatározása, ami a robot megfogási pozícióját fogja meghatározni. A szortírozás közbeni valós idejű (RTDE) adatkommunikációt is végeztünk a robotkarhoz tartozó control box és a program között, az onnan származó adatokat pedig felhasználtuk a robotprogram generálásához. A szortírozás során a szín és kontúr alapján megkülönböztetett munkadarabokat előre meghatározott tárolókba szortírozzuk, ezen paraméterek dinamikusan változtathatók a felhasználó által a programkódba való beavatkozás nélkül, ezzel a rendszer rugalmas alkalmazását elősegítve.

## **5. Köszönetnyilvánítás**

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

## **Irodalomjegyzék**

- [1.] E. H. Ostergaard (2017): The role of cobots in Industry 4.0, Universal Robots, White paper
- [2.] Varun Gopinath (2016): Risk Assessment Process for Collaborative Assembly – A Job Safety Analysis Approach, 6th CIRP Conference on Assembly Technologies and Systems, pp. 199 – 203
- [3.] Varun Gopinath (2017): Safe Assembly Cell Layout through risk assessment – An Application with Hand Guided Industrial Robot, The 50th CIRP Conference on Manufacturing Systems, pp. 430 – 435
- [4.] George Michalos (2014): ROBO-PARTNER: Seamless HumanRobot Cooperation for Intelligent, Flexible and Safe Operations in the Assembly Factories of the Future, Conference on Assembly Systems and Technologies and systems, pp. 71 – 76
- [5.] F. Sherwani; Muhammed Mujtaba Asad (2020): B.S.K.K. Ibrahim: Collaborative Robots and Industrial Revolution 4.0

- [6.] Albert S. Olesen, B. B. Gergaly, E. A. Ryberg, M. R. Thomsen, D. Chrysostomou (2021): A Collaborative Robot Cell for Random Bin-picking based on Deep Learning Policies and a Multi-gripper Switching Strategy
- [7.] M. Alonso, A. Izaguirre, M. Grana (2018): Current research trends in robot grasping and bin picking, in: The 13th Int. Conf. on Soft Computing Models in Industrial and Environmental Applications, Springer, 2018, pp. 367–376
- [8.] T.-T. Le, C.-Y. Lin (2019): Bin-picking for planar objects based on a deep learning network: A case study of usb packs, Sensors
- [9.] Zeng, S. Song, K.-T. Yu, E. Donlon, F. R. Hogan, M. Bauza, D. Ma, O. Taylor, M. Liu (2018): Robotic pick-and-place of novel objects in clutter with multi-affordance grasping and cross-domain image matching, in: 2018 IEEE Int. Conf. on Robotics and Automation (ICRA), IEEE, 2018, pp. 1–8
- [10.] Thomas Feix, Javier Romero, Carl Henrik Ek, Heinz-Bodo Schmiedmayer, and Danica Kragic (2013): A metric for comparing the anthropomorphic motion capability of artificial hands. IEEE Trans. Rob.
- [11.] Jeannette Bohg, Karol Hausman, Bharath Sankaran, Oliver Brock, Danica Kragic, Stefan Schaal, and Gaurav S. Sukhatme (2017): Interactive perception: Leveraging action in perception and perception in action. IEEE Trans. Rob.
- [12.] Minjie Cai, Kris M. Kitani, and Yoichi Sato (2016): Understanding hand-object manipulation with grasp types and object attributes. In Robotics: Science and Systems
- [13.] Lazaros Nalpantidis, Marten ° Bjorkman, and Danica Kragic (2012): Object segmentation: Exploiting camera movement. In Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst., pp. 2116–2121
- [14.] Battle, J., Casals, A., Freixenet, J., Martí, J. (2000): A review on strategies for recognizing natural objects in colour images of outdoor scenes. Image Vis. Comput. 18(6–7), pp. 515–530
- [15.] Boncyk, W.C., Cohen, R.H. (December 4. 2012.): Image capture and identification system and process., U.S. Patent 8,326,031,

- [17.] Ojha, S., Sakhare, S. (2015): Image processing techniques for object tracking in video surveillance—A survey. In: 2015 International Conference on Pervasive Computing (ICPC), pp. 1–6. IEEE
- [18.] Lang, S. (2008): Object-based image analysis for remote sensing applications: modeling reality—dealing with complexity, Blaschke, T., Lang, S., Hay, G. (eds.) Object-Based Image Analysis, Springer, Berlin, pp. 3–27.
- [19.] Nacy, S.M., Abbood, W.T. (2013): Automated surface defect detection using area scan camera. Innov. Syst. Des. Eng. 4(8), 1–10
- [20.] Boncyk, W.C., Cohen, R.H. (November 19. 2013.): Object information derived from object images. U.S. Patent 8,588,527,
- [21.] van Assen, H.C., Egmont-Petersen, M., Reiber, J.H.C. (2002): Accurate object localization in gray level images using the center of gravity measure: accuracy versus precision. IEEE Trans. Image Process. 11(12), pp. 1379–1384.
- [22.] Danica Kragic, Joakim Gustafdo, Hakan Karaoguz, Patric Jensfelt, Robert Krug: Interactive, Collaborative Robots: Challenges and Opportunities, IJCAI-18, 2018.



