

TÉZISGYŰJTEMÉNY

Zeleny Klaudia Éva

**Gépi látáson alapuló termékazonosítási keretrendszer
dinamikus környezetekben**

Témavezetők:

Dr. Kő Andrea

egyetemi tanár, intézetvezető

Dr. Ruppert Tamás

egyetemi docens

BUDAPEST, 2026

Adatelemzés és Informatika Intézet

Tézisgyűjtemény

Zeleny Klaudia Éva

**Gépi látáson alapuló termékazonosítási keretrendszer
dinamikus környezetekben**
című Ph.D. értekezéséhez

Témavezetők:

Dr. Kő Andrea egyetemi tanár, intézetvezető

Dr. Ruppert Tamás egyetemi docens

© Zeleny Klaudia Éva

Tartalomjegyzék

1. Kutatási előzmények és a téma indoklása	4
2. A felhasznált módszerek	6
3. Az értekezés új tudományos eredményei	9
3.1 Moduláris, bizonytalanságkezelő azonosítási architektúra.....	9
3.2 Strukturált adathalmaz és réteges kiértékelési protokoll kontrollált változékonyság mellett	10
3.3 Empirikusan validált képkocka- és epizód szintű azonosítási teljesítmény	10
3.4 Az azonosítási megbízhatóságot leginkább korlátozó feltételek azonosítása	11
3.5 A fennmaradó hibaforrás lokalizálása a helyreállítási mechanizmuson belül	12
3.6 Koncepcionális keret a bizonytalanságkezelő készletgazdálkodási döntéstámogatáshoz	12
4. Főbb hivatkozások.....	13
5. A témában született saját publikáció	15

1. Kutatási előzmények és a téma indoklása

A disszertáció azt vizsgálja, hogyan támogathatja a vizuális percepció a termékazonosítást és a készletgazdálkodáshoz kapcsolódó bizonyítékok előállítását zárt téri raktári környezetekben, különös tekintettel a nagy termékváltozatosságú, kis volumenű (high-mix low-volume, HMLV) belső logisztikai rendszerekre. Ezekben a rendszerekben a termékek elhelyezése dinamikusan változik, a csomagolás nem egységes, a termékkínálat pedig gyakran cserélődik, ami korlátozza a hagyományos, jól strukturált azonosítási munkafolyamatok alkalmazhatóságát.

A raktárkezelő rendszerek (Warehouse Management System, WMS) a fizikai termékek és a digitális nyilvántartások összekapcsolásához jellemzően explicit, géppel olvasható azonosítókat használnak, például vonalkódot, QR-kódot vagy Data Matrix kódot. Ezen adatok pontossága és frissessége azonban a manuális szkennelésen és a kontrollált adatrögzítésen múlik. A termék helyének meghatározása és az azonosítása két, egymással összefüggő, de különálló feladat: a rögzített tárolóhely elérése önmagában nem garantálja, hogy a keresett termék ott van, egy pontatlan vagy elavult helyadat pedig továbbra is manuális keresést igényelhet.

A HMLV raktárakat sokféle, egymástól eltérő terméktípus, alacsony ismétlődési arány, nem rögzített vagy csak ideiglenes tárolóhelyek, valamint a termékek és azonosítók folyamatosan változó elhelyezkedése jellemzi. Ezek a körülmények megnehezítik az olyan automatizált vizuális azonosítási módszerek alkalmazását, amelyek stabil elrendezést, egységes megjelenést vagy előre rögzített objektummodelleket feltételeznek. Emellett jelentős vizuális változékonysággal is számolni kell: a méretarány, a nézőpont, a megvilágítás, a takarás és a mozgási elmosódás egyaránt eltérhet felvételtől felvételre.

Bár a HMLV termékeket többnyire már egyedi, géppel olvasható Data Matrix azonosító jelöli, a jelenlegi munkafolyamatokban a kódleolvasás inkább egy elszigetelt, manuálisan indított esemény, semmint egy folyamatos érzékelési folyamat része. A szakirodalom két, egymástól elszakadt irányt mutat: egyfelől a vizuális detektálás és dekódolás módszertani fejlődését, másfelől a dinamikus HMLV logisztika operatív igényeit. Nem található olyan korábbi kutatás, amely a mélytanulás-alapú detektálást, a dedikált Data Matrix dekódolást és a hasonlóság-alapú validálást egyetlen, valós HMLV körülmények közötti változékonyságra kiértékelt rendszerbe integrálná.

A disszertáció fő célja ezért egy gépi látáson alapuló termékazonosítási folyamat kifejlesztése és kiértékelése, amely megbízható Data Matrix felismerést tesz lehetővé, és amely a termékek lokalizációjának támogatásán keresztül megalapozhatja a készletgazdálkodási döntéstámogatást HMLV raktári környezetekben. A kutatást három kutatási kérdés vezeti: az első a korábbi ismeretek integrálását lehetővé tevő rendszerszintű architektúra kialakítására irányul (K1), a második a javasolt folyamat megvalósítására és empirikus validálására kontrollált környezeti változékonyság mellett (K2), a harmadik pedig az időben validált azonosítási eredmények készletgazdálkodási döntéstámogatási bizonyítékként való operatív értelmezésére (K3).

2. A felhasznált módszerek

A kutatás módszertani keretét a CRISP-DM (Cross-Industry Standard Process for Data Mining) folyamatmodell adja, amely hat, jól elkülönített fázisban strukturálja az adatvezérelt projekteket. Bár a kutatás egy új, gépi látáson alapuló azonosítási folyamat és egy célzottan létrehozott adathalmaz megalkotását is magában foglalja, és bizonyos vonásaiban a design science kutatáshoz is közel áll, a munka elsősorban empirikus, adatvezérelt vizsgálatként értelmezhető,

ezért a CRISP-DM-et választottam elsődleges módszertani keretnek. A hat fázist megtartottam, de a hangsúlyt az üzleti bevezetés helyett a módszertani szigorra, a reprodukálhatóságra és az új tudományos eredmények létrehozására helyeztem.

Az Üzleti megértés fázisában határoztam meg a kutatás kereteit: a mobil kamerával, változó körülmények között végzett Data Matrix azonosítást vizsgálom, a raktárkezelő rendszerbe való integrációt és a készletállapot frissítését azonban nem. A szisztematikus szakirodalmi áttekintés, valamint a kutatási kérdések és célkitűzések alapján a modularitás, a kontrollált változékonyság, a referencia-alapú validálás, illetve a képkocka-szintű érzékelés és az epizódszintű azonosítás elválasztása vált vezérelvvé.

Az Adatmegértés és az Adat-előkészítés fázisában határoztam meg az adathalmazban megjelenítendő adatgyűjtési tényezőket: a megvilágítást, a nézőpontot, az elmosódást, a szimbólum látszólagos méretét, a részleges láthatóságot (takarást), a több látható szimbólum egyidejű jelenlétét és a negatív képkockákat. A felvételeket ezt követően annotált képkocka- és epizódjegyzékekké alakítottam, poligon-alapú objektumannotációval, valamint normál, túl kicsi és takart objektumállapot-kategóriákkal és egy külön elmosódás-

jelzővel kiegészítve. A tanító-, validációs és tesztthalmazt makroszcéna-szinten választottam szét, hogy ugyanaz a fizikai kódrendezés és azonosítókészlet ne jelenhessen meg több partícióban is.

A Modellezés fázisában nem végponttól végpontig tartó, hanem moduláris rendszert építettem fel: YOLOv8 példányszegmentáló modellt a szimbólumok detektálásához, maszk-alapú geometriai normalizálást (kivágás, torzításmentesítés, perspektívakorrekció), dedikált Data Matrix dekódert, egy előre rögzített azonosítólistán alapuló referencia-validálást, hibaesetén aktiválódó helyreállítást klaszter-dekódolással és soft-Hamming hasonlósági rangsorolással, valamint egy dominancia-alapú temporális döntési szabályt a képkocka- és epizódszintű bizonyítékok összegzésére.

A Kísérleti terv és kiértékelés fázisában a tesztthalmazt, a viszonyítási alapot (ground truth) és a jelentési egységeket még az eredmények értelmezése előtt rögzítettem. A detektálást, a dekódolást, a képkocka-szintű azonosítást és az epizódszintű azonosítást külön-külön értékeltem ki, kiegészítve komponensszintű ablációval (a hasonlósági mérték és az akkumulációs szabály tekintetében) és feltétel szerinti elemzéssel (megvilágítás, nézőpont, elmosódás, objektumállapot, szimbólumméret szerint), hogy pontosan

lokalizálni lehessen, a folyamat mely pontján keletkeznek a veszteségek.

A Bevezetés fázisát tudományos disszeminációként és koncepcionális operatív értelmezésként fogtam fel, nem pedig tényleges rendszerbevezetésként. A disszertáció dokumentálja az adathalmazt, a folyamatot, a döntési szabályokat, az eredményeket és a korlátokat, és bemutatja, hogy az elfogadott, kétértelmű és feloldatlan epizódszintű kimenetek hogyan szolgálhatnak korlátozott érvényű bizonyítékként a későbbi készletgazdálkodási döntéstámogatáshoz – anélkül, hogy a kutatás során ténylegesen megvalósítottam volna a raktárkezelő rendszerrel való integrációt, a készletadatok frissítését vagy az éles üzembe helyezést.

3. Az értekezés új tudományos eredményei

A kutatás legfontosabb tudományos eredményei az alábbi pontokban foglalhatók össze.

3.1 Moduláris, bizonytalanságkezelő azonosítási architektúra

A disszertáció egy integrált azonosítási architektúrát dolgoz ki, amely élesen elválasztja egymástól a képkocka-szintű érzékelést és a szekvencia-, illetve epizódszintű

azonosítást. A detektálás, a geometriai normalizálás, a dedikált Data Matrix dekódolás, a referencia-validálás, a hibaesetén aktiválódó helyreállítás és a temporális bizonyítékösszegzés egyetlen moduláris folyamatban egyesül, amelyben a képkocka-szintű kimenetek bizonyítéknak számítanak, nem pedig végleges döntésnek.

3.2 Strukturált adathalmaz és réteges kiértékelési protokoll kontrollált változékonyság mellett

Kiértékelés céljára létrehozott adathalmazt építettem fel, amelyben a megvilágítás, a nézőpont, az elmosódás, a látszólagos szimbólumméret, a takarás és a negatív minták változékonysága kontrollált, a tanító-, validációs és teszhalmaz elválasztása pedig makroszcéna-szinten, kódfüggetlen módon történik. A detektálást, a dekódolást, valamint a képkocka- és epizódszintű egységeket egyaránt lefedő, réteges kiértékelési protokoll az azonosítási megbízhatóságot komponensenként bontja szét, egyetlen összesített mutató helyett.

3.3 Empirikusan validált képkocka- és epizódszintű azonosítási teljesítmény

Az 1015 annotált objektumot és 34 epizódot tartalmazó teszhalmazon a detektálás az objektumok 98,5%-át lokalizálta. Az elsődleges dekódolás önmagában az objektumok 35,5%-át

azonosította, a helyreállítási ág pedig 59,1%-ra emelte a képkocka-szintű azonosítási arányt, az elfogadott azonosítók 98,7%-os pontossága mellett. Epizód szinten a teljes folyamat az epizódok 91,2%-ára adott azonosítót, ebből 85,3% bizonyult helyesnek, 93,5%-os pontosság mellett – szemben a visszamenőlegesen számított, egyetlen képkockára épülő dekódolási felső korláttal (73,5%). A temporális integráció négy olyan epizódot is fel tudott oldani, amelynél egyetlen képkocka sem eredményezett helyes dekódolást, és egyetlen esetben sem rontott a végeredményen.

3.4 Az azonosítási megbízhatóságot leginkább korlátozó feltételek azonosítása

A rétegzett eredmények szerint a korlátozott szimbólumláthatóság a legerősebb korlátozó tényező: a takartként vagy túl kicsiként annotált objektumokat az esetek mindössze 9,2%-ában sikerült azonosítani, szemben a normál objektumok 67,9%-os arányával. A nézőpont 30,8, az elmosódás pedig 25,4 százalékpontos különbséget okozott, az azonosítási pontosság pedig a szimbólum látszólagos területével együtt nőtt: a legkisebb méretkategóriában 37,6%, a legnagyobbban 86,0% volt.

3.5 A fennmaradó hibaforrás lokalizálása a helyreállítási mechanizmuson belül

Validált dekódolásból egyetlen esetben sem született helytelen azonosító, sem képkocka-, sem epizódszinten. A megfigyelt hibák mindegyike egyetlen mechanizmusra vezethető vissza: a dominancia-alapú hasonlósági elfogadásra olyan esetekben, amikor csak egyetlen, gyengén alátámasztott jelölt állt rendelkezésre. Ez arra utal, hogy a döntési szabályt érdemes lenne kiegészíteni egy explicit minimális bizonyíték-feltétellel, a már alkalmazott relatív elkülönülési teszt mellett.

3.6 Koncepcionális keret a bizonytalanságkezelő készletgazdálkodási döntéstámogatáshoz

A disszertáció olyan koncepcionális keretet dolgoz ki, amelyben az elfogadott, kétértelmű és feloldatlan epizódszintű azonosítási kimenetek a készletgazdálkodási döntéstámogatás bizonytalanságkezelő bizonyítékaként értelmezhetők. Ez a keret a validált vizuális azonosítást a készletállapot-becslés egyik lehetséges bemeneteként pozicionálja, ugyanakkor világosan kijelöli a határt e bizonyíték és egy tényleges raktárkezelő rendszer-integráció vagy térbeli lokalizáció között – ezek ugyanis a megvalósított rendszeren kívül maradnak.

4. Főbb hivatkozások

Almeida, T., Santos, V., Mozes, O. M., & Lourenco, B. (2021). Comparative analysis of deep neural networks for the detection and decoding of data matrix landmarks in cluttered indoor environments. *Journal of Intelligent & Robotic Systems*, 103(1), 13.

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C., & Wirth, R. (1999). The CRISP-DM user guide. In 4th CRISP-DM SIG Workshop in Brussels in March (Vol. 1999).

Edula, V., Ammisetty, K., Kotha, A., Ravipati, D., Davuluri, A., et al. (2023). A novel framework for QR code detection and decoding from obscure images using YOLO object detection and Real-ESRGAN image enhancement technique. In 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT) (pp. 1–6). IEEE.

Gan, Z. L., Musa, S. N., & Yap, H. J. (2023). A review of the high-mix, low-volume manufacturing industry. *Applied Sciences*, 13(3), 1687.

Karrach, L., & Pivarčiová, E. (2020). Recognition of data matrix codes in images and their applications in production processes. *Management Systems in Production Engineering*, 28(3), 154–161. <https://doi.org/10.2478/mspe-2020-0023>

Karrach, L., & Pivarčiová, E. (2021). Comparative study of data matrix codes localization and recognition methods. *Journal of Imaging*, 7(9), 163.

Liao, L., Li, J., & Lu, C. (2022). Data extraction method for industrial data matrix codes based on local adjacent modules structure. *Applied Sciences*, 12(5).

Liu, Y., Song, Y., Gu, G., Luo, J., Wang, T., & Jiang, Q. (2024). A data matrix code recognition method based on L-shaped dashed edge localization using central prior. *Sensors*, 24(13), 4042.

- Losada del Olmo, J. J., de Teruel, P. E. L., Ruiz, A., & García-Clemente, F. J. (2025). Computer vision on the edge: A scalable auto-ID solution for industrial logistics. *Procedia Computer Science*, 265, 276–284. <https://doi.org/10.1016/j.procs.2025.07.182>
- Matuszczyk, D., & Weichert, F. (2023). Reading direct-part marking data matrix code in the context of polymer-based additive manufacturing. *Sensors*, 23(3), 1619.
- Muallim, T., Kucuk, H., Bareket, M., & Kahraman, M. (2025). Lightweight deep learning model and novel dataset for restoring damaged barcodes and QR codes in logistics applications. *Computer Modeling in Engineering & Sciences*, 143(3), 3557.
- Rybakova, E. O., Limonova, E. E., & Bezmaternykh, P. V. (2025). Improving data matrix mobile recognition via fast Hough transform and adaptive grid extractors. *Computer Optics*, 49(6), 1182–1190. <https://doi.org/10.18287/COJ1804>
- Simeth, A., Kumar, A. A., & Plapper, P. (2025). Flexible and robust detection for assembly automation with YOLOv5: A case study on HMLV manufacturing line. *Journal of Intelligent Manufacturing*, 36(5), 3447–3463.
- Tadjine, C., Ouafi, A., Benlamoudi, A., & Taleb-Ahmed, A. (2025). Computer vision in warehouse management automation: A survey on implemented methods with prototyping hardware. *Engineering Applications of Artificial Intelligence*, 160, 111886. <https://doi.org/10.1016/j.engappai.2025.111886>
- Vaccari, L., Coruzzolo, A. M., Lolli, F., & Sellitto, M. A. (2024). Indoor positioning systems in logistics: A review. *Logistics*, 8(4), 126. <https://doi.org/10.3390/logistics8040126>
- Winkelhaus, S., Grosse, E. H., & Morana, S. (2021). Towards a conceptualisation of order picking 4.0. *Computers & Industrial Engineering*, 159, 107511. <https://doi.org/10.1016/j.cie.2021.107511>

Xu, L., Kamat, V. R., & Menassa, C. C. (2017). Automatic barcode extraction for efficient large-scale inventory management. In *Computing in Civil Engineering 2017* (pp. 340–348). ASCE.

5. A témában született saját publikáció

Zeleny, K., Ruppert, T., & Kó, A. (2026). A Proposed Framework for Robust Visual Product Identification in Dynamic Environments. *Acta Polytechnica Hungarica*, 23(8), 211–240. <https://doi.org/10.12700/APH.23.8.2026.8.12>