

KÉPFELDOLGOZÁS ALAPÚ MESTERSÉGES INTELLIGENCIA ALKALMAZÁSA A SERTÉSTENYÉSZTÉSBEN

APPLICATION OF IMAGE-PROCESSING BASED ARTIFICIAL INTELLIGENCE IN PIG FARMING

Kovács Tamás^{0000-0003-3947-8696*}, Kovács Lóránt⁰⁰⁰⁹⁻⁰⁰⁰⁵⁻¹³⁰⁴⁻⁴⁵⁰⁹

¹ Informatika Tanszék, GAMF Műszaki és Informatikai Kar, Neumann János Egyetem, Magyarország
<https://doi.org/10.47833/2026.1.CSC.006>

Kulcsszavak:

mesterséges intelligencia;
képfeldolgozás;
sertéstenyésztés

Keywords:

artificial intelligence, image
processing,
pig farming

Cikktörténet:

Beérkezett 2025. november 25.
Átdolgozva 2026. február 6.
Elfogadva 2026. április 5.

Összefoglalás

Ebben a munkában egy nagyobb, a mesterséges intelligencia sertéstenyésztésben való hasznosításával foglalkozó projekt egy részét mutatjuk be. Ebben az alprojektben a célkitűzés az, hogy képfeldolgozási módszerekkel a sertés-karámok felett elhelyezett kamerák felvételei alapján automatizált monitoring rendszert hozunk létre, amely képes a karámban elhelyezett egyedek általános aktivitását mérni. Ezen a részen belül is ez a cikk a betanított neurális hálózat alapú egyed-lokalizáció lehetőségeit és képességeit részletezi. A YOLO rendszer két kisebb hálózatát használtuk fel a rendszer továbbtanítására, hogy képes legyen elfogadható pontossággal lokalizálni a sertések teljes testét és fejét. A betanított rendszer validációja alapján elmondható, hogy a helyes test felismerés aránya az összes egyed százalékában 95% feletti, a helyes fej objektumok felismerési aránya 83% feletti.

Abstract

In the present work the basis of an image processing based monitoring system introduced that is capable of measuring the activities of the pigs in a pen. This work is a part of a greater project dealing with the application of artificial intelligence in pig farming. We applied two of the pre-trained YOLO neural networks (nano and small) to train them further so as they are able to recognize and localize pig bodies and pig heads in our environment. We trained the network with images taken by cameras fixed above the pens. The validation of the trained system has shown that the trained network successfully recognized 95% of the pig bodies and 83% of the pig-heads.

1. Bevezetés

Az elmúlt évtizedben a digitalizáció, a gazdaság egyéb területeihez hasonlóan, a mezőgazdasági szektorokban is fókuszba került, mint elsődleges fejlesztési terület. Ezeknek a fejlesztéseknek egyik kiemelt területe a mesterséges intelligencia alkalmazása, amely bizonyítottan képes növelni a termelékenységet. Ez a munka a képfeldolgozáson alapuló mesterséges

* Kapcsolattartó szerző.
E-mail cím: kovacs.tamas@nje.hu

intelligencia alkalmazására mutat be ilyen fejlesztési lehetőségeket a sertéstenyésztési ágazaton belül.

Ez a téma az alkalmazott informatikai kutatás területén is számos eredményt hozott. Több, mint tíz éve Kashiha és munkatársai a hízók vízfogyasztási szokásinak vizsgálatára alkalmaztak kamerás képfeldolgozást [1]. Ezzel a karámonkénti vízfogyasztás mérésén túl az ivási gyakoriság napi eloszlását is meg tudták határozni. Ezekon a felvételeken a pillanatnyi itató-használat felismerése volt a feladat, amit előtér pixel szegmentációval és az állatok test-kontúrájának felismerésével oldottak meg.

2017-ben Nasirahmadi és munkatársai felülnézeti kamerák állóképeit vizsgálták [2], amelyek 22 egyedes karámról készültek, és az elemzés célja a sertések csoportos fekvési mintázatainak és szokásainak elemzése volt. Az elemzéshez Delaunay háromszögelést alkalmaztak a sertésekre szegmentált képeken. Nem sokkal később Yang és munkatársai már neuronhálós tanítási módszert alkalmaztak (Faster R-CNN) [3], amellyel képesek voltak egy adott karámon belül az egyedek azonosítására az egyedek fej-alakjának és mintázatának megtanulása révén. Ezzel a módszerrel nyomon lehetett követni azt, hogy az egyes példányok mikor tartózkodnak az etetőnél, ami lehetővé tette egy egyedekre lebontott táplálkozási statisztika készítését.

Ugyanabban az évben Hansen és munkatársai [4] szintén betanított neurális háló segítségével azonosítani tudták az egyedek pófáját az etetőnél, az állatokkal szemközt elhelyezett kamera felvételein és ezzel hasonló információhoz jutottak, mint Yang és munkatársai munkájában.

Nemrégiben Li és munkatársai [5] a konvolúciós neuronháló alapú tanítást már nem állóképeken, hanem rövid video szegmenseken alkalmazták, és így képesek voltak néhány tipikus viselkedési elemet (veszekedés, vakarózás, egymásra hágás, stb.) azonosítani. Ennél a feladatnál a neuronháló bemeneti adat-vektorát már nem csak az állóképek pixeleinek síkbeli pozíciói alkották, hanem a megtanult egyszerű kép-elemek (features) időbeli változása is.

Az alább bemutatott munkában az eredeti kitűzött cél az, hogy a Kéleshús Sertéstenyésztő és Hizláló Kft. jánoshalmi telephelyén a Vitafort Kft. által folytatott takarmányozási-kísérletben a Neumann János Egyetem kutatói egy kamerás képfeldolgozásra épülő kiértékelő rendszert fejlesszenek ki, amely a sertések illetve sertés-csoportok aktivitását jellemzi a kísérlet szempontjából releváns módon. A következő fejezetben ismertetjük a konkrét feladatot és a kísérleti környezetet. A 3. fejezetben részletezzük az alkalmazott neuronháló alapú mesterséges intelligenciát, majd beszámolunk az eddig elért eredményekről. Végül az utolsó fejezetben összegezzük az elvégzett munka konklúzióit.

2. A feladat részletezése

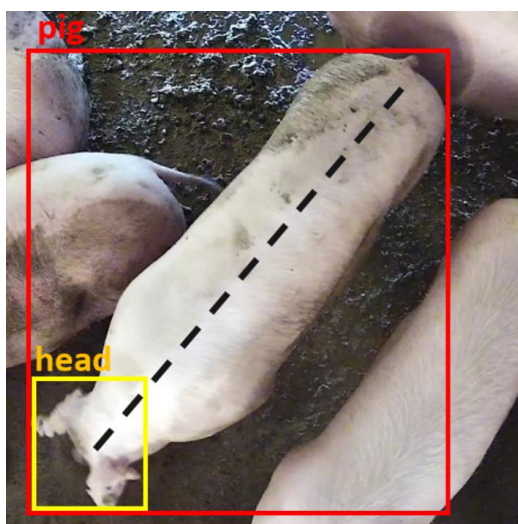
Ebben a fejezetben felvázoljuk, hogy milyen informatikai célt tűzünk ki a takarmányozási kísérlet kiszolgálásaként, de a cikk további részeiben nem ennek a feladatnak a teljes megoldásáról számolunk be, hanem a megoldás felé tett lépéseink részeredményeiről. A takarmányozási kísérletben a mérhető egyedenkénti súlygyarapodás mellett fontos kérdés még az állatok általános közérzete, amelyre a viselkedési minták vizuális megfigyeléséből következtetni lehet. A legalapvetőbb viselkedési paraméter az egyes egyedek és a karámban lévő csoport egészének mozgási aktivitása. A fentebb idézett munkákban ez az egyik paraméter, amelynek mérését képfeldolgozási módszerekkel megoldották. A csoport mozgás jellemzésének tekintetében azt a célt tűztük ki, hogy meghatározzuk az egyedek mozgási aktivitási szintjének eloszlását egy adott fix T_w időablakon belül. A mozgási aktivitás szintjét az időablakon belül az egyed középpontja által megtett út hosszával definiáljuk. Az eloszlás meghatározása tehát azt az információt adja meg a csoportra vonatkozóan, hogy a csoport hány százaléka volt nagyon aktív (azaz sok utat járt be), kevésbé aktív illetve gyengén aktív. Az aktivitási szintek száma tetszőleges, de túl finom felbontást nem érdemes alkalmazni.

Ennél a megközelítésnél az egyedek mozgását külön mérjük, azonban az egyedek azonosítását nem tűzzük ki célként. Egy másik, későbbi időablakban egy adott megfigyelt egyed nem azonosítható egy korábbi időablakban megfigyelt egyeddel. Lehet, hogy egy adott időablakban a legaktívabb egyedek egy későbbi időablakban a legkevésbé aktívak közé tartoznak, de erre vonatkozóan ebben a kitűzött feladatban nem kapunk információt. Tehát az időablakbeli aktivitás-eloszlás alapvetően a csoport általános aktivitását jellemzi egy adott időpontban.

Ennek a feladatnak a megoldásához informatikai szempontból szükséges a karámok megfelelő felülnézeti bekamerázása, a kamerák képeinek folyamatos rögzítése a kísérlet megelőző és a kísérlet alatti időszakban. Ez a mi esetünkben legalább két hónapnyi rögzítést jelent, mivel a kísérlet kb. egy hónapnyi időtartamot vesz igénybe. Két karám felett helyeztünk el kamerát a hizlaldai részlegen, ugyanígy két karámnál az anya-koca részlegen.

Második lépésként szükséges az egyedek automatikus objektum szintű, azaz nem egyed-azonosítás szintű felismerése a képeken, pontosabban fogalmazva egy sertés-objektum azonosítására alkalmas mesterséges intelligencia szolgáltatás kidolgozása az adott környezetben. Ezt a feladatot kétféle módszerrel tervezzük megoldani: az egyik egy objektum-pixel szegmentáció (más szóval előtér-szegmentáció) alapú módszer, ami a sertés objektumok pixeleit ismeri fel színtartomány alapján, majd egy feladatra szabott modell alapján próbálja megadni az adott egyed kontúrját. A másik módszer egy neuronháló alapú, azaz egy nagyobb méretű kép-adathalmazzal betanított objektum lokalizáló és felismerő rendszer, amely a felismert sertés-objektumokra illetve sertés-testrész objektumokra egy téglalapot („box”-ot) illeszt. A cikk további részében csak ez utóbbi módszer részeredményeit ismertetjük.

A számunkra fontos objektum kategóriák a sertések egésze továbbá olyan testrészek, amelyek könnyen felismerhetők a betanított hálózat számára. Az egyedek lokalizációjához és a test orientációjának megállapításához legszükségesebb és ugyanakkor elegendő kategóriák: az egész test (category: „pig”), és a fej (category: „head”). Ha a hálózat helyesen lokalizálja és egy box-szal megjelöli ezt a kettőt, azaz az 1. ábrán vázolt példához hasonló információt szolgáltat, akkor a test-középpont és az orientáció közelítőleg meghatározható.



1. ábra: Példa a fej és egész-test objektum azonosítására (sárga és piros téglalap), és ez alapján a test orientációjának megállapítására.

Szóba jöhet még kategóriaként a láb és a farok, de a láb csak akkor látszik a felülnézeti képeken, ha a sertés oldalt fekszik vagy áll, de a kamera tengelytől távol van. A farok azonosítása redundáns információt ad a fej meghatározása mellett, viszont növelné a középpont és az orientáció meghatározás robusztusságát. Első fázisban két legszükségesebb kategóriát, a testet és a fejet választottuk a betanításhoz. Ennek a munkának a célja, hogy a fentebb részletezett test-fej objektumfelismerő rendszer létrehozását és teljesítményének értékelését bemutassa. A jelen objektum felismerési séma újdonsága a bevezetőben bemutatott kutatási munkákhoz képest az, hogy a felülnézeti kamerák képein az objektum-azonosítás alapja nem az él-detektálásra és a kontúr illesztés, hanem a manapság jóval modernebb technikának számító deep learning módszer, amellyel a test és fej azonosításával nemcsak az egyed középpontját kaphatjuk meg, hanem a test pontos orientációját is. Emellett az objektum azonosítás a fő cél, nem pedig az egyed azonosítás (mint a [3]-as vagy [4]-es irodalomban megadott munkáknál), mivel az objektum azonosítás jóval robusztusabban működik. Mivel a deep learning objektum detektálás egy fontos eszközként szolgál egy nagyobb, egyed-követő rendszerhez, az objektum felismerés pontossága ebben az esetben

kiemelten fontos paraméter. Bár még az egyed követő rendszer csak a jövőben fog tesztelhető formában megvalósulni, jelenleg úgy becsüljük: ahhoz, hogy a követés például egy 10 perces időablakban az egyedek legalább felénél sikeres legyen, 90% feletti „recall” és „precision” értékeket kell elérnünk az objektum felismerés eredményeinél.

3. Az eddig elért eredmények

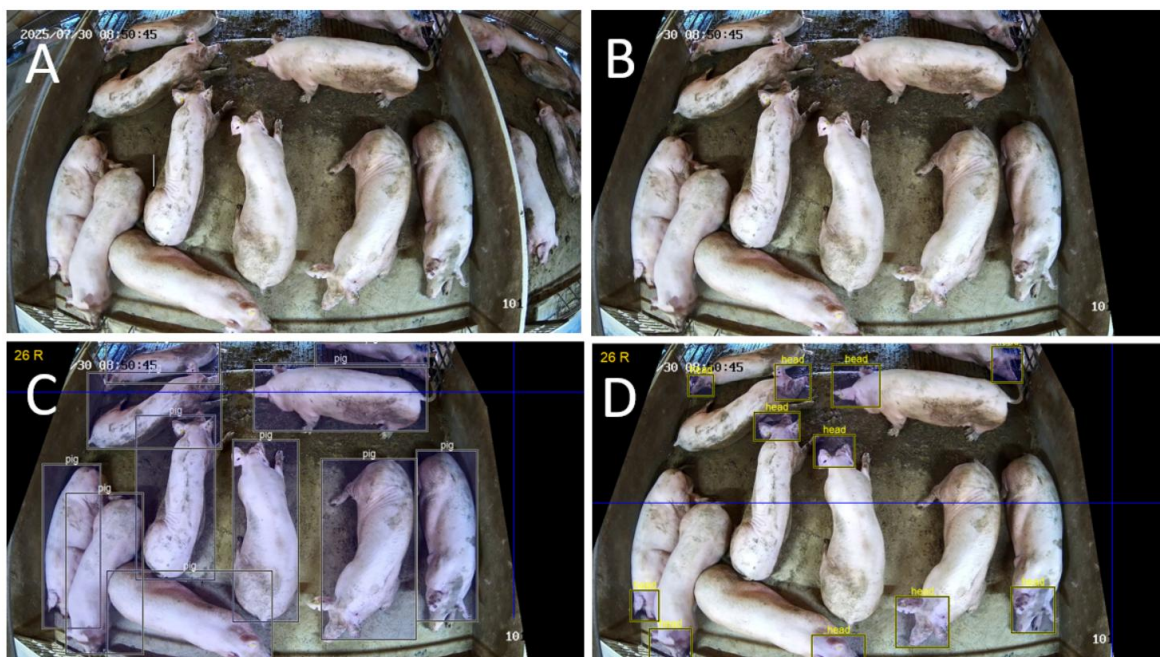
A képen található objektumok lokalizációja, azaz annak meghatározása, hogy hol van a képen bármi érdekes, és a megtalált objektumok kategóriájának meghatározása két különböző feladat a neurális háló alapú mesterséges intelligencia rendszerek számára. Az általunk a munkához választott, az utóbbi másfél évtizedben sok kutató által fejlesztett YOLO rendszer (You Only Look Once) [6] ezt a két problémát egyetlen „inferenciával” oldja meg, azaz az objektumok lokalizációját és kategorizálását is megadja úgy, hogy a kép pixeleit csak egyszer futtatja át a neurális hálón. Ennek a rendszernek már létezik egy több százezres kép-adathalmaz (COCO128 dataset) betanított hálója, ami alkalmas 128 különféle, a környezetünkben gyakran előforduló objektum azonosítására.

A YOLO rendszer emellett tovább is tanítható egy adott speciális feladatra a felhasználó igényei szerint. Ennek az az alapja, hogy a használt és betanított neuronhálózat sok neuron-réteget tartalmaz: az alsóbb, bemenethez közelebbi rétegek a kisebb pixel kontextusokat (kép-részleteket) tanulnak meg. Ilyenek például az élek, sarkok, különféle alakú foltok, stb., amelyek a nagyobb kontextusban objektummá állnak össze. A hálózat fentebbi rétegei tanulják meg a nagyobb méretskálán lévő objektumok jellemzőit. A YOLO hálózatának újratanítása egy speciális feladatra a tudás-transzfer elvén alapul, amelyről a [7] publikációban részletes leírást lehet találni. A jelen esetben a YOLO v8-as verziójú rendszer „nano” és „small” hálózatát használtuk a tudás-transzfer alapján a sertés test és fej objektumok megtanítására. Ez a két hálózat nem túl nagy mérete miatt viszonylag gyors inferenciára (objektum lokalizációra) képes. Az elérhető hálózatok architektúrájáról bővebb leírást kaphatunk a [8] publikációban.

A betanításhoz használt kép-adathalmaz előállítása a következő lépésekben történt:

- Mintavételezés: a kiválasztott videó-felvételekből 3 percenként állókép mintát vettünk;
- Maszkolás: A képekre egy maszkot illesztettünk, amely kitakarja a nem a karámhoz tartozó részt;
- Annotáció: A képeken kézzel bejelöltük a sertés-test és a felismerhető fej objektumokat egy saját fejlesztésű szoftver segítségével;
- Tanító és validációs (tesztelő) halmaz szétválasztása: az annotált képek kb. 10%-át elkülönítettük a tanító halmaztól, ezekkel a képekkel validáljuk a betanítás eredményességét;
- Tanítás: a YOLO rendszer erre szolgáló parancssori moduljával;
- Kiértékelés: elsősorban a konvergencia jószágát és a tévesztési mátrix (confusion matrix) elemeit alapul véve értékeljük az eredményeket.

A fenti tanítási programnak a lépéseit illusztrálja a 2. ábrán látható képsorozat.



2. ábra: Példa a kép adathalmaz elkészítési fázisaira: az eredeti kép (A), a maszkolt kép (B), a sertés-test objektumokat annotált kép (C) és a fej objektumokat annotált kép (D).

A betanítást elsőként egy az anyakoca karámban készült felvételen végeztük el, mert ezen a felvételen a testek és fejek jobb felbontásban látszottak mint a hizlaldai részleg felvételein.

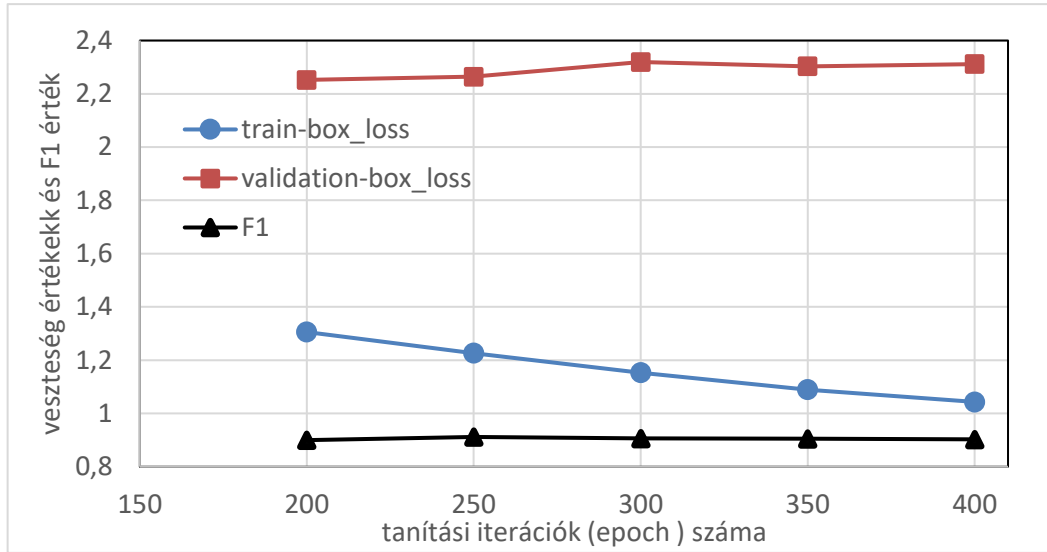
A mintavételezést három különböző nap (2025.07.30, 2025.09.30 és 2025.10.30) nappali óráiban készült videofelvételeiből végeztük, ami kb. 30 órányi videóanyag. A mintavételezést 3 percenként végeztük, azaz óránként 20 képet kaptunk, ami összesen kb. 600 képet jelent az adott karámban. Mivel a karámban 11 egyed tartózkodott, ezeken a képeken összesen kb. $600 \times 10 = 6000$ sertés-test objektum és $600 \times 6 = 3600$ fej objektum látható. Azért nem 600×11 , mert a takarások és a rejtett fejek miatt az átlagos felismerhető test 10, az átlagos felismerhető fej 6 volt képenként. A tanításhoz ebből a halmazból 500 képet, a validációhoz 48 képet használtunk, amelyek fele a tanító halmaz képeinek időszakából lett kiválogatva egyenletes eloszlással, a másik fele olyan videókból lett kiválasztva, amelyek időben távol estek tanító adathalmaz képeitől.

A betanításhoz a következő hardver-összeállítást használtuk:

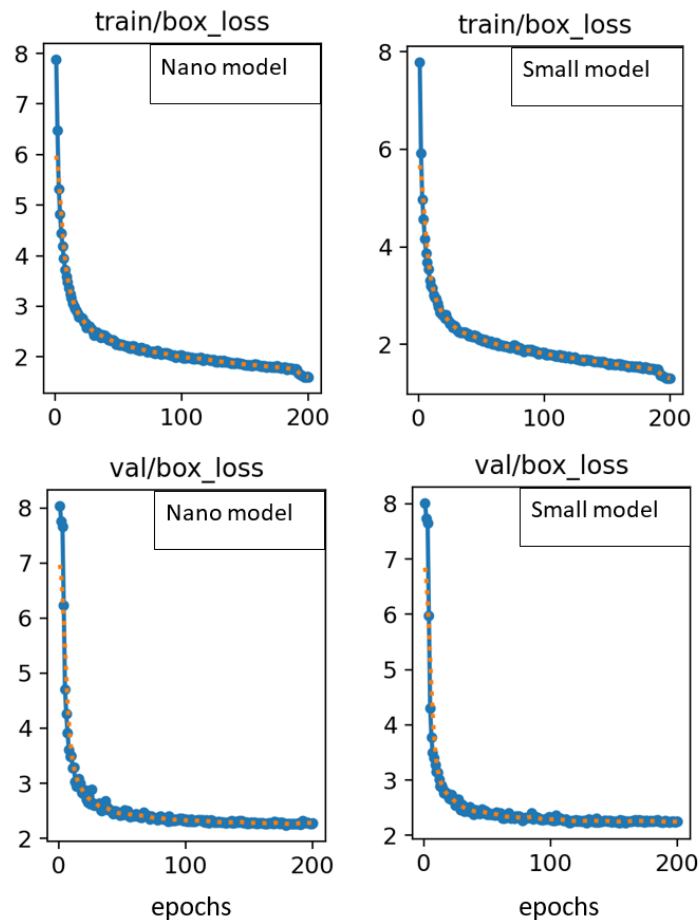
- Videókártya: ASUS GeForce RTX 5060 Ti 16GB Dual OC Edition videokártya (DUAL-RTX5060TI-O16G)
- Processzor : Intel Core i7-14700K 3.4Ghz LGA1700 dobozos (BX8071514700K)
- Alaplap: MSI MAG Z790, (2x48GB) Intel memóriával

A tanítást 200, 250, 300, 350 és 400 tanítási ciklusszám (epochs) mellett is elvégeztük, hogy képek kaphassunk arról, hogy mekkora az az epochs szám, amely a konvergált rendszer elérhető pontosságát biztosítja és elkerüli a túltanulást. A 3. ábra mutatja a validációs F1 érték alakulását és a tanítási és a validációs halmazon mért `box_loss` veszteségi érték alakulását a növekvő epochs szám mellett a small hálózati modell esetén. (A `box_loss` veszteségi hiba a becsült objektum téglalap és a valós téglalap eltéréseinek összegzett hibáját jellemzi és arányos a téglalap sarokpont és a téglalap méreteinek lineáris eltéréseivel.) A veszteségi értékek azt mutatják, hogy a 200 epochs feletti régióban a tanítási halmazon javulás mutatkozik, a validációs halmazon pedig enyhén romlik a teljesítmény, míg a validációs halmazon mért F1 érték érdemben nem változik. Ebből következik, hogy a 200 epochs feletti értékektől nem várhatunk javulást a pontosság terén, ugyanakkor a túltanulás miatti teljesítmény romlás sem jelentős. Ennélfogva a 200-as epochs számot alkalmaztuk a későbbi teszteléseknél a nano és a small hálózatnál is.

A képek mérete minden esetben 1280x720 pixel volt, a tanulási ráta a 0.01-től folyamatosan csökkent a 0.0001 záró-értékig a tanítás alatt. A beállított optimalizáló algoritmus az AdamW volt, a batch méret: 16.



3. ábra: Az F1 érték, továbbá a box_loss veszteségi értékek alakulása az epoch szám növelésével a small hálózat esetén.



4. ábra: A box_loss veszteségek alakulása a tanítási (train) és a validációs (val) halmazokon a nano és a small hálózatokra a tanítási iterációk előrehaladásával.

A 4. ábrán bemutatott box_loss konvergencia görbék megerősítik a feltételezést, hogy a 200 epochs feletti régióban a hálózat már a túltanulás fázisában van: a 200 epochs körüli tartományban a validációs halmazon a veszteség-értékek már konvergálnak, míg a tanító halmaznál még közel sem vízszintes a konvergencia-görbe. Ez igaz mind a nano, mind a small hálózatra.

A 200 epochs melletti tanítási ciklus metrikái az 1. táblázatban láthatók.

1. Táblázat

Mutatók	Nano modell	Small modell
precision	0.876	0.885
recall	0.896	0.915
F1	0.886	0.900
mAP@0.5	0.924	0.945
mAP@0.5:0.95	0.624	0.645

A validáció során kapott, oszlopokra normált tévesztési mátrix (confusion matrix) a nano hálózat esetén az 5. ábrán látható. A mátrix adatait az alábbiakban számokkal is megadjuk.

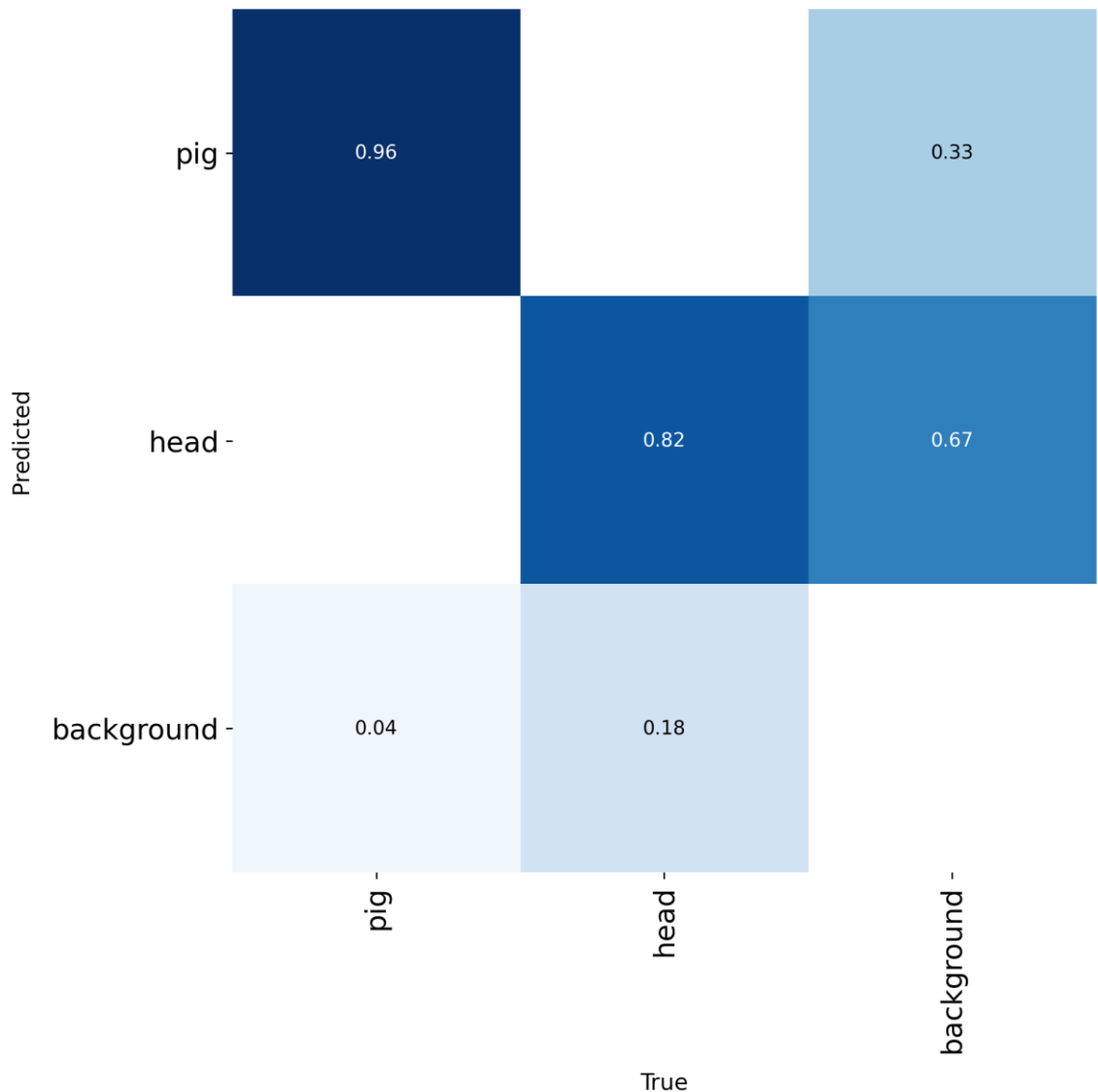
A pig objektumra vonatkozó eredmények:

- A 48 db validálási képen szereplő pig objektum összesen: 521.
- Ebből a tanított rendszer helyesen felismer: 500, azaz felismerte a pig objektumok 95.7%-át.
- Tévesen head-nek azonosított pig: 2
- Fel nem ismert pig: 19, azaz 3.7%
- Tévesen pig-nek azonosított háttér objektum: 29, ami a helyesen azonosított pig-ek 5.6%-a

A head objektumra vonatkozó eredmények:

- A 48 db validálási képen szereplő head objektum összesen: 367
- Ebből a tanított rendszer helyesen felismer: 300, azaz felismerte a head objektumok 81.7%-át.
- Tévesen pig-nek azonosított head: 0
- Tévesen head-nek azonosított háttér objektum: 58, ami a helyesen azonosított head-ek 19.3%-a

Az egyik fontos kérdés ezekben a kezdeti kísérletekben, hogy a neurális háló, azaz a modell méretének növelése mennyire javítja a becslés pontosságát. Ezért végeztük el a betanítást egy nagyobb hálózaton is (YOLO Small). A betanítás paraméterei ugyanazok voltak. Amint látható a 3. ábrán 200 epochs betanítási ciklusszám mellett ennél a hálózatnál is hasonló konvergenciát értünk el. A kapott tévesztési mátrix a 6. ábrán látható.



5. ábra: A betanított nano hálózat validációja során kapott tévesztési mátrix.

A nagyobb modell esetén a mátrix adatai a számokkal kifejezve:

A pig objektumra vonatkozó eredmények:

- A 48 db validálási képen szereplő pig objektum összesen: 521
- Ebből a tanított rendszer helyesen felismer: 511, azaz felismerte a pig objektumok 98.1%-át.
- Tévesen head-nek azonosított pig: 0
- Fel nem ismert pig: 10, azaz 1.9%
- Tévesen pig-nek azonosított háttér objektum: 19, ami a helyesen azonosított pig-ek 3.7%-a

A head objektumra vonatkozó eredmények:

- A 48 db validálási képen szereplő head objektum összesen: 367
- Ebből a tanított rendszer helyesen felismer: 320, azaz felismerte a head objektumok 87.2%-át.
- Tévesen pig-nek azonosított head: 0

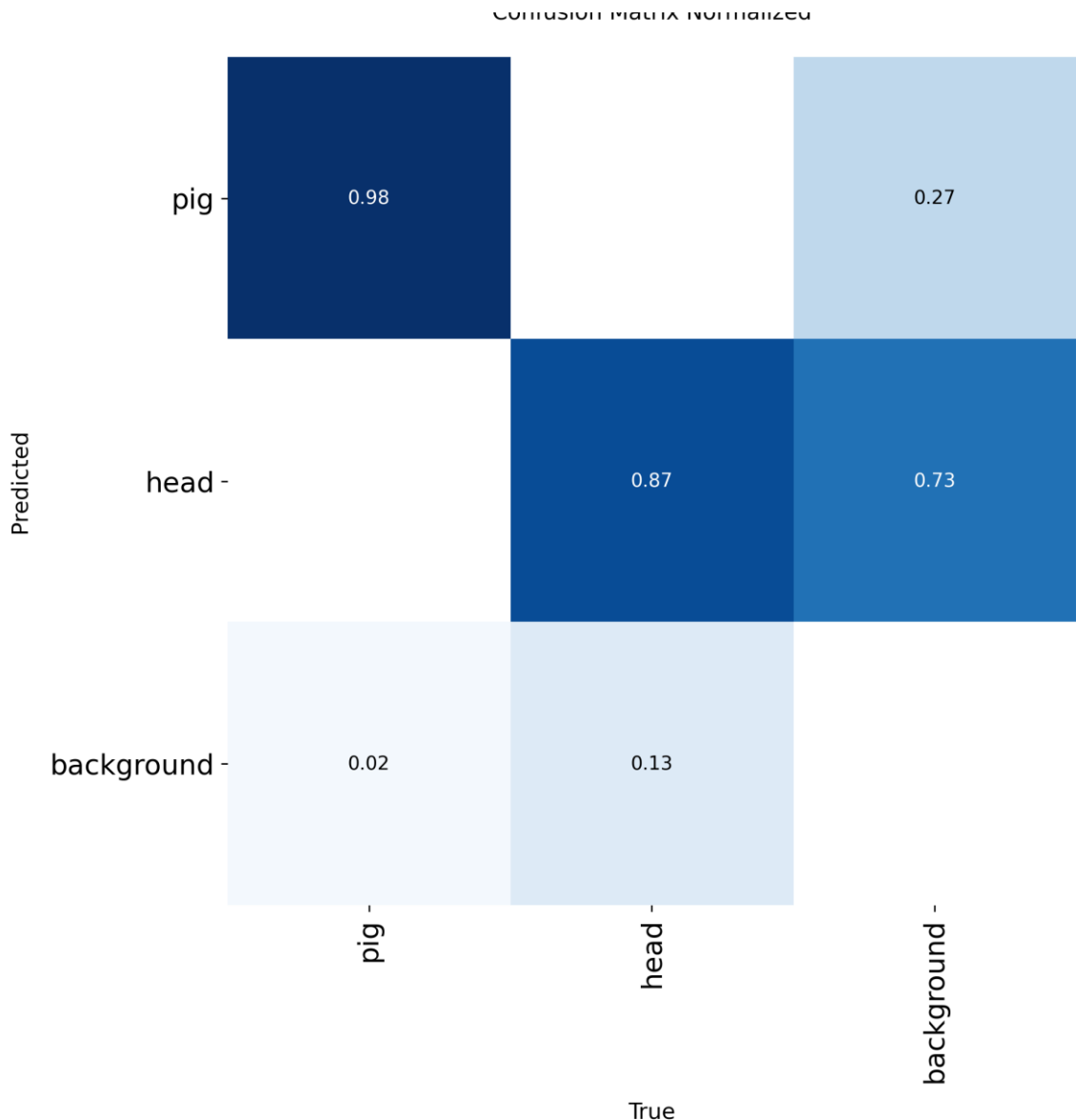
- Tévesen head-nek azonosított háttér objektum: 52, ami a helyesen azonosított head-ek 16.3%-a

Amint korábban említettük, a 48 annotált validációs kép fele a tanításra használt képek idősávjából, a másik fele a tanítási képek idősávjától időben távoli sávból kerültek válogatásra. Emiatt teszteltük azt is, hogy ha csak időben távoli képekkel végezzük a validációt, akkor ez mennyit ront a betanított rendszer mért pontosságán. A tesztet csak a „small” hálózat végeztük el. Ennek a tesztnek az eredményeként az adódott, hogy a helyesen felismert pig objektumok aránya 95.1%, a tévesen pig-nek azonosított háttér objektum a helyesen azonosított pig-ek 4.1%-a. A helyesen felismert head objektumok aránya 83.0%, a tévesen head-nek azonosított háttér objektum a helyesen azonosított head-ek 16.0%-a. Ezek határozottan gyengébb eredmények, mint a kevert validációs halmazzal mért képek.

A könnyebb összevethetőség kedvéért ezeket az eredményeket összefoglaltuk a 2. táblázatban.

2. Táblázat

Mutatók	Nano modell	Small modell	Small modell, csak időben távoli képekkel validálva
helyesen felismert pig objektumok aránya	95.7%	98.1%	95.1%
tévesen pig-nek azonosított háttér objektumok aránya	5.6%	3.7%	4.1%
helyesen felismert head objektumok aránya	81.7%	87.2%	83.0%
tévesen head-nek azonosított háttér objektumok aránya	19.3%	16.3%	16.0%



6. ábra: A betanított small hálózat validációja során kapott tévesztési mátrix.

4. Összegzés és az eredmények értékelése

Az itt bemutatott kutatási munka a teljes projektnek azt a részét ismerteti, ami a karámban tartott sertés egyedek kameraképeken való CNN betanítás alapú lokalizációjával foglalkozik. A kiindulási hálózat a széles körben alkalmazott betanított YOLO hálózat volt. A sertés egyed lokalizációjához és a test orientáció meghatározásához a test és a fej objektumok felismerésének betanítását végeztük el egy kb. 500 képes és kb. 5000 sertést tartalmazó tanító halmazon. A validációs tesztek azt mutatták, hogy a használt nano és small hálózatok (amelyek elég gyors inferenciát biztosítanak) ekkora méretű tanító adathalmaz esetén 95% feletti felismerési pontosságot adnak a pig objektumok esetén és 83% feletti pontosságot adnak a head objektumok esetén. Ez egy ekkora méretű tanító halmaz esetén kifejezetten jó eredménynek számít és a teljes projekt-feladatban célként kitűzött egyed követés megvalósításához vélhetően elegendő pontosságot biztosítanak. A nagyobb hálózat alkalmazása (nano helyett small) a pig objektumok esetében már nem javít sokat a pontosságon, de a head objektumok esetén a javulás már számottevő. Amennyiben az egyed követésnél nem lesz feltétel a valósidejű megvalósítás, azaz az inferenciánál kevésbé fontos a gyorsaság, javasolt a nagyobb hálózatot választani a nagyobb elérhető pontosság

miatt. Emellett, a későbbi kutatási munka során fontos lesz ugyanezt a betanítási munkát és tesztelést elvégezni a YOLO által kínált jóval nagyobb hálózatoknál is.

Fontos megjegyezni, hogy a tanító halmaz és a teszt halmaz képei ugyanazon karámból és körülbelül ugyanolyan méretű egyedekről készültek, ennél fogva még nem tudjuk, hogy a rendszer teljesítménye mennyire romlik, ha a betanított hálózatot más karámok és más méretű egyedek esetén készült felvételeken teszteljük. További munkaként ennek elvégzése is szükséges lesz, és amennyiben ez a betanított hálózat egy elvárt szint alatt teljesít más karámoknál, akkor bővíteni kell a tanítóhalmazt a többi karám képeivel egy újabb, jóval nagyobb halmazzal történő tanításhoz.

Köszönetnyilvánítás

A szerzők köszönetet mondanak a projektben résztvevő Neumann János Egyetem GAMF Műszaki és Informatikai Kar - kollégáinak. Köszönettel tartozunk a kutatás támogatásáért, amely a 2023-1.1.1-PIACI_FÓKUSZ azonosítószámú pályázathoz tartozó „Fenntartható termelésből származó alapanyagok felhasználása a magas minőségű élelmiszer-előállítás támogatására és hatásuk a sertések termelési, állatjóléti, szaporodásbiológiai és gazdaságossági eredményességére” című pályázat keretében valósult meg.

Irodalomjegyzék

- [1] Kashiha, M., Bahr, C., Haredasht, S. A., Ott, S., Moons, C. P., Niewold, T. A., ... & Berckmans, D. (2013). The automatic monitoring of pigs water use by cameras. *Computers and Electronics in Agriculture*, 90, 164-169.
- [2] Nasirahmadi, A., Hensel, O., Edwards, S. A., & Sturm, B. (2017). A new approach for categorizing pig lying behaviour based on a Delaunay triangulation method. *Animal*, 11(1), 131-139.
- [3] Yang, Q., Xiao, D., & Lin, S. (2018). Feeding behavior recognition for group-housed pigs with the Faster R-CNN. *Computers and electronics in agriculture*, 155, 453-460.
- [4] Hansen, M. F., Smith, M. L., Smith, L. N., Salter, M. G., Baxter, E. M., Farish, M., & Grieve, B. (2018). Towards on-farm pig face recognition using convolutional neural networks. *Computers in Industry*, 98, 145-152.
- [5] Li, D., Zhang, K., Li, Z., & Chen, Y. (2020). A spatiotemporal convolutional network for multi-behavior recognition of pigs. *Sensors*, 20(8), 2381.
- [6] Wang, C. Y., & Liao, H. Y. M. (2024). YOLOv1 to YOLOv10: The fastest and most accurate real-time object detection systems. *APSIPA Transactions on Signal and Information Processing*, 13(1).
- [7] Drenyovszki, R. (2024). Solving a classification problem using transfer learning on small image datasets. *Gradus*, 11(1).
- [8] Terven, J., Córdova-Esparza, D. M., & Romero-González, J. A. (2023). A comprehensive review of yolo architectures in computer vision: From yolov1 to yolov8 and yolo-nas. *Machine learning and knowledge extraction*, 5(4), 1680-1716.