

ÓBUDAI EGYETEM
ÓBUDA UNIVERSITY

Alba Regia Kar
Geoinformatikai Intézet

TDK dolgozat

Aratási időpont meghatározása Sentinel-1 idősorokból konvolúciós neurális hálózattal

Szerző:

Simon Máté Péter

Óbudai Egyetem, Alba Regia Kar, Geoinformatika MSc

Témavezető:

Dr. habil Molnár Gábor Péter, egyetemi docens

Óbudai Egyetem, Alba Regia Kar, Geoinformatikai Intézet

Külső konzulens:

Pacskó Vivien, távérzékelési szakértő

Lechner Tudásközpont, Térbeli Szolgáltatások Igazgatósága

**OE-AMK
2024**

Tartalom

1. BEVEZETÉS	3
2. ADATOK	4
2.1 Mintaterület, vizsgált évek	5
2.2 Sentinel-1 felvételek	6
2.3 Referenciaadatok	8
2.3.1 Kérelem-adatbázis.....	8
2.3.2 Aratási időpontok.....	10
3. MÓDSZERTAN.....	13
3.1 Szakirodalmi áttekintés.....	13
3.2 Térbeli szűrés.....	14
3.3 Bemeneti adatstruktúra	16
3.4 Augmentáció.....	17
3.5 Osztályozási modell.....	17
3.6 Regressziós modell	18
3.7 Vizsgálatok	19
3.8 A pontosság vizsgálata	19
4. EREDMÉNYEK ÉS DISZKUSSZIÓ	21
4.1 Osztályozás.....	22
4.2 Regresszió.....	24
4.2.1 Egyéves idősorral tanított modellek azonos éven tesztelve	25
4.2.2 Egyéves idősorral tanított modellek másik éven tesztelve.....	26
4.2.3 Két év idősoraival tanított modell	27
5. ÖSSZEGZÉS.....	28
6. KÖSZÖNETNYILVÁNÍTÁS.....	30
7. IRODALOMJEGYZÉK	30
SAJÁT EREDMÉNYEK	35
FÜGGELÉK.....	36

1. BEVEZETÉS

A klímaváltozással összefüggésben a modern mezőgazdaság új kihívásokkal küzd. Napjainkban a legfőbb cél a termelékenység fenntartása környezetbarát módon, ehhez új technológia és több segédtudomány bevonása szükséges. Ennek érdekében az Európai Unió 2012-ben megalkotta a Közös Agrárpolitikát (KAP), amely tartalmazza a mezőgazdasági területek távérzékeléssel támogatott ellenőrzését is (CAP, 2012). A Központi Statisztikai Hivatal adatai alapján 2023-ban Magyarország területének 55%-a mezőgazdasági terület volt. Ebből 82% (4,2 millió hektár) szántó és 16% (793 ezer hektár) gyep. A gyümölcs- és szőlőültetvények aránya közel 3% (143 ezer hektár). Az erdőterületek és mezőgazdasági területek együtt lefedték az ország területének 80%-át (Agroinform, 2023). A nagy területen végzett jelentős agrártevékenység miatt ezen területek változásának folyamatos nyomon követése és számszerűsítése a KAP alapelvek szerint is kiemelt jelentőségű (Henits et al., 2022).

A műholdas földmegfigyelés nélkülözhetetlen támogatást nyújt a modern mezőgazdaság számára (Weiss et al., 2020). 2014-től, az Európai Űrügynökség (ESA) Copernicus földmegfigyelési programjának köszönhetően nagy mennyiségű műholdas mérési adat áll rendelkezésre nagy térbeli és időbeli felbontásban. Ez lehetővé teszi a dinamikus változó mezőgazdasági területek pontos, naprakész megfigyelését. A két legszélesebb körben használt adattípus a Sentinel-1 szintetikus apertúrájú radar (SAR) űrfelvételek és a Sentinel-2 multispektrális űrfelvételek. Műholdas távérzékelési adatok feldolgozásával, idősoros elemzésével tematikus térképek készíthetők a mezőgazdasági növények térbeli eloszlásáról (Dhumal et al., 2013) vagy a parcellák hozamáról (Roznik et al., 2022), illetve azonosítható és számszerűsíthető a szélsőséges időjárási események által okozott károk mértéke (Birinyi et al., 2024a). A nagy mennyiségű adat folyamatos, automatizált feldolgozásához azonban korszerű informatikai háttér szükséges. Így az említett elemzések módszertanában ma már komoly szerepet kapnak a gépi tanulási eljárások.

A konvolúciós neurális hálózat (CNN) egy rendkívül népszerű és elterjedt gépi tanulási algoritmustípus. A szakirodalom alátámasztja, hogy pontos és elegendő referenciaadattal, valamint a megfelelő modell megalkotásával nagy pontosságú feladatmegoldásra képes (Alzubaidi et al., 2021). Előnye több más gépi tanulási eljáráshoz viszonyítva, hogy sok dimenziót és paramétert képes kezelni a bemeneti adatban, ez kiemelten alkalmassá teszi műholdfelvétel-idősorok kiértékelésére. Nagy mennyiségű és kellően reprezentatív tanítóadattal pedig robusztus előtanított algoritmust hozhatunk, ami általánosan alkalmazható különböző évek idősoraira vagy eltérő területekre létre (*Transfer Learning*, Hosna et al., 2022).

Jelen dolgozatban bemutatom kutatásomat, mely során konvolúciós neurális hálózatokat és Sentinel-1 SAR idősorokat használtam a vetésszerkezet és az aratás időpontjának meghatározására. Kutatásom a Komplex Mezőgazdasági Kockázatkezelési Rendszer (MKR) módszertani fejlesztéseihez kapcsolódik (NÉBIH, 2024), melyben a Lechner Tudásközpont (LTK) Űrtávérzékelési Osztálya is részt vesz. Az aratás idejének meghatározása a mezőgazdasági termelést érintő kárigények elbírálása szempontjából fontos, hiszen a kárigény jogosságát a betakarítás után nem lehet megítélni. Az aratási időpont térképezéséből továbbá statisztikai elemzések készíthetők, amelyekkel számszerűsíthetjük az egyes évek vagy területek közötti különbségeket a növények fejlődésében. A vetésszerkezeti térképezést nyolc kategóriára, míg az aratási időpont azonosítását négy meghatározó növénykategóriára végeztem el.

Dolgozatomban elsőként bemutatom a felhasznált adatokat, a kutatás mintaterületét és a vizsgált éveket, a Sentinel-1 felvételek feldolgozását és a szórási együttható származtatását, valamint a felhasznált referenciaadatokat és a Planet optikai űrfelvételeket (2. fejezet). Ezután következik a módszertani fejezet, melyben áttekintem a Sentinel-1 idősorok jelen feladatokra való alkalmazhatóságának szakirodalmi háttérét, röviden összefoglalom a konvolúciós neurális hálózatok működését, bemutatom a felvételeken végrehajtott Refined Lee és Non-Local Means szűrési eljárásokat, valamint a kutatásban használt CNN modellek felépítését (3. fejezet). Eredményeimet a 4. fejezetben közlöm, melyben kitérek a pontossági statisztikák összehasonlító elemzésére, továbbá az egyes évek és kategóriák közti különbségekre is. Végül összefoglalom kutatásomat és az abból levont következtetéseket a módszertan általános alkalmazhatóságára, illetve további fejlesztésre vonatkozóan (5. fejezet). Dolgozatomat köszönetnyilvánítással, a hivatkozások jegyzékével és függelékkel zárom.

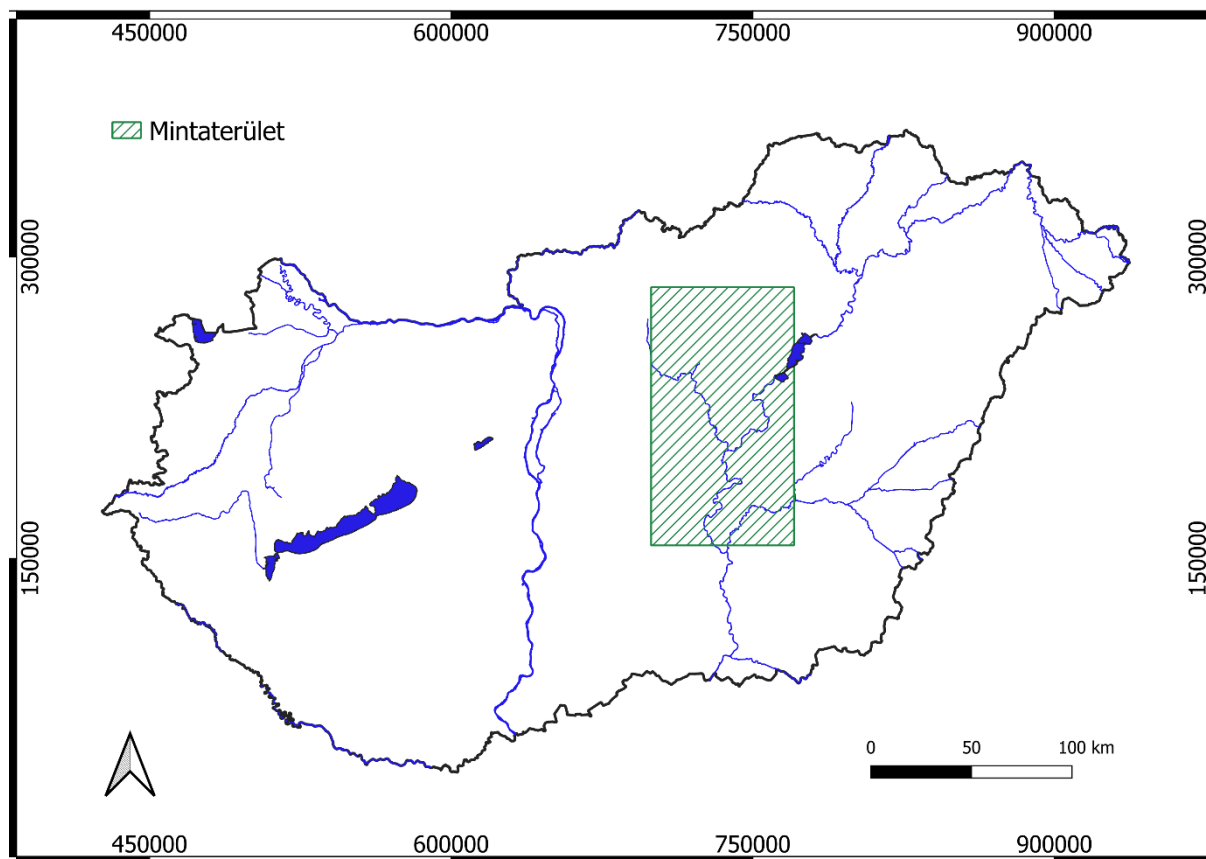
2. ADATOK

Az alábbiakban bemutatom a kutatás során felhasznált adatokat. Elsőként a kutatási mintaterületet, valamint a vizsgált éveket, illetve kiválasztásuk szempontjait ismertetem. Ezután áttekintem a Sentinel misszió kereteit, céljait és az adatszolgáltatás koncepcióját. Ismertetem Sentinel-1 SAR felvételek alapvető jellemzőit: a felvételezés módját, az elérhető termékek különböző feldolgozottsági szintjeit, valamint térbeli és időbeli felbontásukat. Röviden összegzem a feldolgozás lépéseit és megnevezem a felhasznált szoftvereket. Ezt követően ismertetem a referenciaadatokat, az adatbázis szűrésének lépéseit és az ehhez igénybe

vett szoftvereket. Végül kitérek a Planet optikai űrfelvételekre és a betakarítás-detektálás vizuális módszertanára.

2.1 Mintaterület, vizsgált évek

A mintaterület 9115,5 km² kiterjedésű, ez nagyságrendileg az ország területének egytized része. Észak–dél irányban a Bükkaljától Szentestig, nyugat–kelet irányban Hatvantól a Tiszáig húzódik (lásd 1. ábra). Ezen terület nagyrésze agrár hasznosítású, benne a betakarítás szempontjából vizsgált négy főnövény mindegyike jelentős számú parcellán megtalálható. Többségében sík terület, így a radar adatokra a domborzat nem gyakorol számottevő hatást. A mintaterület tekintélyes mérete nagy mennyiségű adat feldolgozását követelte meg, azonban biztosította a módszertan alkalmazhatóságának alaposabb vizsgálatát.



1. ábra: Kutatási mintaterület (HD72 / EOVS / EPSG:23700)

A mezőgazdasági évek eltérő klimatológiai viszonyai a távérzékelte adatok idősorait is nagy mértékben befolyásolják. Ezért, ha egy módszertan alkalmazhatóságát vizsgáljuk, fontos azt több, esetleg eltérő időjárású éven tesztelni. Kutatásom során a kidolgozott módszertant 2021 és 2023 évek idősorán teszteltem, amelyek megfelelnek a fenti szempontnak. A 2021-es év első felében jelentős mennyiségű belvíz volt jelen, ennek hatása a szántóföldi növények fejlődésére

számottevő volt (Birinyi et al., 2024b). A 2023-as év időjárási szélsőségek által nem sújtott, átlagos klímájú év volt. Továbbá a 2.2 alfejezetben később részletezett okokból a 2021-es évben 6 napos, míg a 2023-as évben 12-napos időbeli felbontásban állnak rendelkezésre a Sentinel-1 felvételek, így ennek hatása is vizsgálható.

2.2 Sentinel-1 felvételek

Az ESA 2014-ben indította el a Sentinel missziót a Copernicus földmegfigyelési program keretein belül, mely során különböző műholdak és mérőszensorok megtervezését és felküldését valósítják meg. Céljuk a Föld komplex folyamatainak tanulmányozása, a felszín és a légkör időbeli változásainak állandó nyomon követése. A különböző feldolgozottságú termékeket a Copernicus online adatszolgáltató felületén, a Copernicus Dataspace Ecosystem (CDSE) rendszeren publikálják, a termékek ingyenes regisztrációt követően tölthetők le. Kutatásom kiindulási adatai a Sentinel-1 szintetikus apertúrájú radar (SAR) méréseiből előállított termékek (ESA, 2023).

A Sentinel-1 program alapkoncepciója szerint a megfelelő fedés és időbeli felbontás érdekében egyszerre legalább két működő műholdnak kell(ene) adatot szolgáltatnia. A misszió első műholdját (Sentinel-1A) 2014-ben, a másodikat (Sentinel-1B) 2016-ban állította az ESA Föld körüli pályára. Ezek a műholdak kvázipoláris, azonos napszinkron pályán keringenek a felszín felett 693 km magasan, és 12 naponta térnek vissza ugyanazon földrajzi hely fölé. Annak érdekében, hogy a felvételek időbeli felbontása javuljon, a műholdpár között 180° a fáziskülönbség, így a visszatérési idő hat napra csökken. Azonban 2022 januárjában a Sentinel-1B műhold meghibásodott (ESA Sentinel Online, 2022), a Sentinel-1C felbocsátása pedig a jelenlegi tervek szerint csak 2024 novemberében történik meg, így jelenleg 12 napos az időbeli felbontás (ESA Sentinel Online, 2024).

A műholdakon elhelyezett C-sávú szintetikus apertúrájú radar (SAR) szenzorok 5,405 GHz frekvenciával felvételeznek (ESA Sentinel Online, 2024). A műszer úgynevezett duálpol, ez azt jelenti, hogy a kibocsájtott hullámok esetében meg kell választani, hogy horizontálisan vagy vertikálisan legyen polarizált. A műszerbe beérkező jel esetén azonban mindkét polarizált hullám felvételezésre kerül, így ez a szenzor a HH+HV vagy VV+VH csatornakombinációkat teszi lehetővé. A műholdakkal négyféle felvételezési módban lehet adatot gyűjteni, ezek a *Stripmap* (SM), az *Interferometric Wide Swath* (IW), az *Extra-Wide Swath* (EW) és a *Wave* (WV). Magyarországon az IW mód érhető el, VV+VH csatornákkal, ez a szárazföldök monitorozására a legmegfelelőbb. Ezen módszer esetén a szenzor által megvilágított felszíni

sáv (swath) 250 km széles, és azimuth irányban további három úgynevezett *sub-swath*-ra (“melléksávra”) oszlik. A felvételezés a *Terrain Observation by Progressive Scans SAR* (TOPSAR) technikával történik, melynek köszönhetően az egyes swath-okra egy homogén, jó minőségű kép hozható létre, melynek range irányú felbontása 2,7 – 3,5 m, míg az azimuth irányú felbontása 22,5-22,7 m között változik (De Zan és Monti-Guarnieri, 2006).

Az ESA által publikált termékek feldolgozottsági szintjük alapján háromféle terméktípusba sorolhatók. A nulladik szintű felvételek tömörített, fókuszálatlan adatokat tartalmaznak. Az első szintű termékek már kalibrált, fókuszát adatok. Két fajtája a *Single Look Complex* (SLC) és a *Ground Range Detected* (GRD). A második szintű termékek az első szintű termékekből származtatott, óceánok tanulmányozására alkalmas geofizikai adatok.

Kutatásom során első szintű, SLC adatokból indultam ki. Ezen termékek az amplitúdó értékek mellett a fázisra vonatkozó információkat is tartalmazzák komplex számok formájában. Választott mintaterületemnek (lásd 2.1 alfejezet) megfelelően a 153-as *descending* (leszálló) relatív pálya felvételeit használtam fel. Az SLC termékek feldolgozása a *Sentinel Application Platform* (SNAP) szoftverrel történik (SNAP, 2024). Az első lépés a termikus zaj korrekciója, majd a korrigált SLC termékekből szórási együtthatókat számolunk. Itt kizárólag a VV és VH amplitúdókat használjuk fel, a fázisértékek ehhez a művelethez nem szükségesek. Ezen folyamat során radiometrikus kalibráció történik, melynek lényege, hogy az amplitúdó értékekből szórási együtthatót származtassunk a következő képlettel:

$$\sigma^0 = \frac{DN^2}{A_{dn}^2 K} \sin(\alpha),$$

ahol

- α a beesési szög,
- DN a digitális amplitúdó érték,
- A_{dn} a pixel átskálázási érték,
- K a kalibrációs konstans és
- σ^0 a szórási együttható.

A szórási együttható a felszíni szóró objektumokról reflektált radarjel erősségét adja meg egységnyi felületre vetítve, 0-1 közötti intervallumon. Ezután következik az úgynevezett *multilook* eljárás (De Zan és Monti-Guarnieri, 2006) és a *Refined Lee Speckle filter* (Lee, 1981) azért, hogy csökkenjen a radarfelvételekre jellemző speckle effektus zavaró hatása. A szűrési eljárásokat a 3.2 alfejezetben tovább részletezem. Utolsó lépésként pedig elvégezzük a topográfiai korrekciót az SRTM domborzatmodell alapján (Farr és Kobrick, 2000). Az eddig

részletezett eljárást a Lechner Tudásközpont fejlesztése alapján a Földmegfigyelési Információs Rendszer (FIR, 2024) automatizáltan végzi.

A feldolgozott felvételeket ezután a Python programnyelv *gdal* (GDAL, 2024) és *rasterio* (Gillies S., et. al., 2013) raszterfeldolgozó moduljai segítségével WGS84 (EPSG:4326) vetületből egységes országos vetületbe (EOV, EPSG:23700) vetítettem. Az azonos pályához tartozó képeket egyesítettem, majd a szórási együttható értékeket logaritmikus skálára transzformáltam. Végül kivágtam a felvételeket a kutatási mintaterületemre, és átmintavételeztem 20 méteres térbeli felbontásra.

A vetésszerkezeti osztályozáshoz az idősort a vizsgált években február közepétől október közepéig használtam fel, hiszen ez a mezőgazdasági növények fő vegetációs időszaka. Így 12 napos visszatérés esetén 22, míg 6 napos visszatérés esetén 43 kép került az idősorba. A betakarítási időpont meghatározásához a május eleje és november eleje közötti intervallumot vettem figyelembe annak érdekében, hogy tartalmazza minden vizsgált növény aratási időszakát. Így 12 napos időbeli felbontással 20, míg 6 napos időbeli felbontással 39 radarfelvétel volt az input adat. Mindkét módszertan esetén a VV és VH sigma0 rétegeket használtam fel (1. táblázat).

Időbeli felbontás	Felvételek száma	Rétegek száma (szórási együttható)	Módszertan
6 nap	43	2 (VV, VH)	Vetésszerkezeti térképezés
12 nap	22	2 (VV, VH)	
6 nap	39	2 (VV, VH)	Betakarítás detektálás
12 nap	20	2 (VV, VH)	

1. táblázat: A felhasznált Sentinel-1 adatok idő-és spektrális dimenziói

2.3 Referenciaadatok

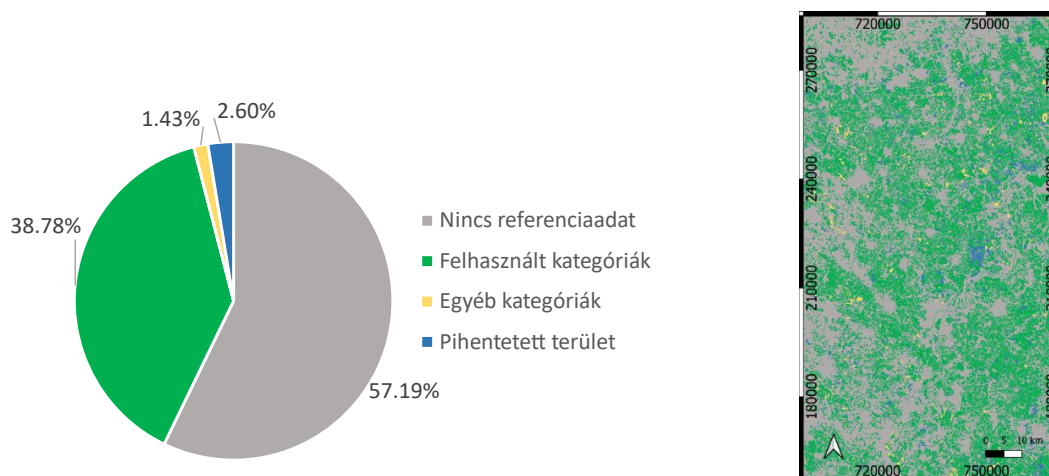
2.3.1 Kérelem-adatbázis

A vetésszerkezeti és betakarítási térképek előállításához egyaránt felhasználtam referenciaadatként egy, az LTK Űrtávérzékelési Osztálya által a rendelkezésemre bocsátott vektoros adatbázist. Ezen adatbázis mezőgazdasági kategóriáit az egységes területalapú támogatásokra vonatkozó kérelem-adatbázis nyújtotta, amely évről-évre a gazdák támogatási igényléseit tartalmazza. A lombhullató erdő és víz kategóriák forrása a Nemzeti Nagyfelbontású

Felszínborítási Réteg (*National High Resolution Layer, NHLR*), amely egy 10 méter térbeli felbontású, országos tematikus raszter (Copernicus Böngésző, 2024).

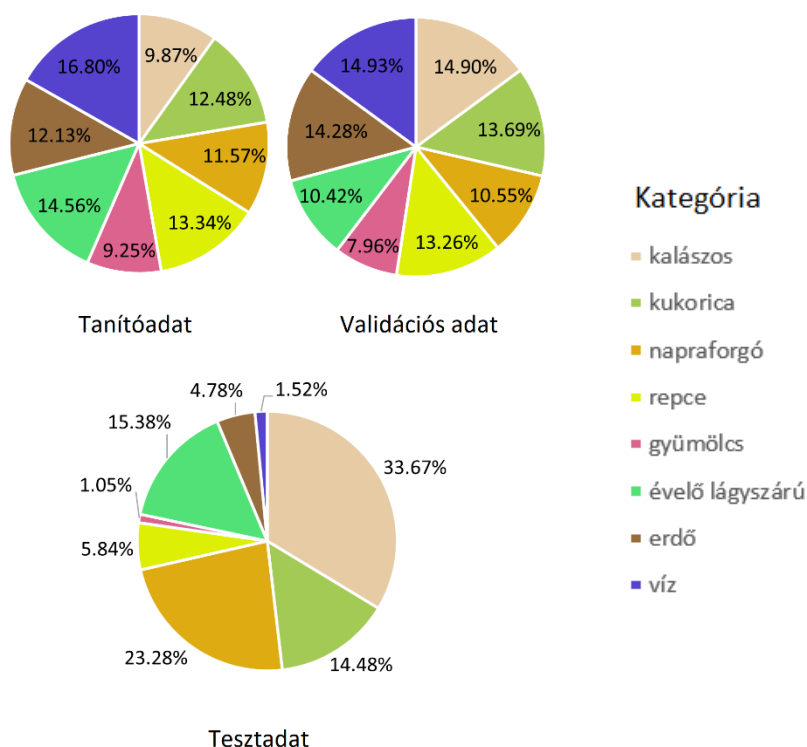
Ezen adatoknál előfordulhat részleges vagy teljes átfedés a poligonok között és geometriai pontatlanság a poligonok határain, amely hibákat topológiai ellenőrzéssel, illetve 10 méteres negatív puffer alkalmazásával korrigáltam. Ezután kiszűrtem az adatbázisból a 4 hektárnál – azaz 100 db Sentinel-1 pixelnél – kisebb, illetve a 0,04-nél nagyobb kerület/terület arányú poligonokat annak érdekében, hogy megtisztítsam az adatbázist azoktól a parcelláktól, amelyek feltételezhetően javarészt kevert pixeleket tartalmaznak. Ez a viszonylag szigorú szűrési feltétel nem okozott jelentős adatvesztést, mert a rendelkezésre álló adatbázis rendkívül nagy méretű volt (lásd 3. ábra). A referenciaadatok előkészítését a Quantum GIS (QGIS, 2024) szoftverrel, illetve Python környezetben a *geopandas* modul (Jordahl et al., 2014) segítségével végeztem.

Az adatbázist a kijelölt mintaterületen a vetésszerkezeti térképezés esetében az alábbi nyolc kategóriára szűrtem: kalászos, kukorica, napraforgó, repce, gyümölcs, évelő lágyszárú (lucerna és gyepek összevonva), erdő, víz. Ez a kategóriarendszer tartalmazza a betakarítási időpont meghatározásánál vizsgált négy legfontosabb szántóföldi növénytypust (kalászos, kukorica, napraforgó, repce), a további osztályok pedig a legfontosabb felszínborítási kategóriák. A 2. ábrán látható, hogy a kiválasztott kategóriarendszerrel a referenciaadat 2021-ben lefedi a teljes terület 38,78%-át, 2,60% pihentetett terület, 1,43% pedig egyéb ritka szántóföldi növény. Az osztályozáshoz a rendelkezésre álló referenciaadat mennyisége megfelelő, a felhasznált kategóriarendszer kellően részletes. A 2023-as évre vonatkozó statisztika hasonló eloszlást mutat, a függelék A) ábráján tekinthető meg.



2. ábra: A vetésszerkezeti térképezéshez felhasznált kategóriák, az egyéb kategóriák és a pihentetett területek aránya a mintaterületre vonatkozó teljes adatbázisban 2021-ben.

A *deep learning* osztályozáshoz ezt az adatbázis 3 részre osztottam, tanító, validáló és teszt halmazokra. A tanításhoz és a modell validációjához használt adathalmazok egyaránt kategóriánként 50 parcellát tartalmaztak. Ennek köszönhetően az osztályozás a kategóriákra nézve a parcellák szintjén kiegyensúlyozott volt. Az osztályozás pontosságának teszteléséhez a teljes fennmaradó adatbázist felhasználtam. A referenciaadatok kategóriánkénti eloszlását és a teljes mintaterülethez viszonyított területarányát a 2021-es évre 3. ábra kördiagramjai és táblázata alapján ismertetem. A 2023-as év statisztikái hasonlóan alakulnak, ezek a függelék B) ábráján szerepelnek.



Adathalmaz	Területarány [%]	Pixelszám [db]	Terület [ha]
Tanítóadat	0,54%	123775	4951,00
Validációs adat	0,53%	122114	4884,56
Tesztadat	37,56%	8590179	343607,16

3.ábra: Statisztika a 2021-es vetésszerkezeti térképezéshez felhasznált referenciaadatokról. Az egyes kategóriák pixelszám szerinti aránya a referenciaadatokban (fent) és a referenciaadatok által lefedett terület aránya a mintaterülethez viszonyítva (lent).

2.3.2 Aratási időpontok

A betakarítás idejét meghatározó mélytanulási algoritmus tanításához a referenciaadatokat vizuális interpretációval állítottam elő. Ehhez a már említett kérelem-adatbázis geometriáit és növénykultúra-kódjait, illetve nagyfelbontású multispektrális Planet

űrfelvétel-idősorokat használtam fel. Ezen felvételek 3 méteres térbeli felbontással és 1 napos visszatérési idővel készülnek. Ennek a kiemelkedően nagy időbeli felbontásnak köszönhetően a betakarítás időpontja akár napra pontosan meghatározható. A felvételek négy spektrális sávot tartalmaznak, ezek a vörös, a zöld, a kék és a közeli infravörös (Planet, 2024).

A vizuális betakarítás-időpont meghatározást a Planet Explorer webfelületen végeztem a kérelem-adatbázissal egyeztetve. A vizsgált optikai idősorok kezdő és záró dátumának meghatározásához a növények aratásáról szerzett általános információkat használtam fel. A korai vetésű növények (kalászos, repce) esetén május végétől július végéig tartó felvételeket vizsgáltam, míg a késői vetésű kapásnövények (kukorica, napraforgó) esetében az augusztus vége és november eleje közötti időszakra vártam a betakarítási eseményt. Az adatgyűjtés során megállapítható volt, hogy a betakarítási esemény a valószínűs (RGB) kompozitokon erős, hirtelen színváltozással jár, így hatékonyan detektálható (lásd 4. ábra).



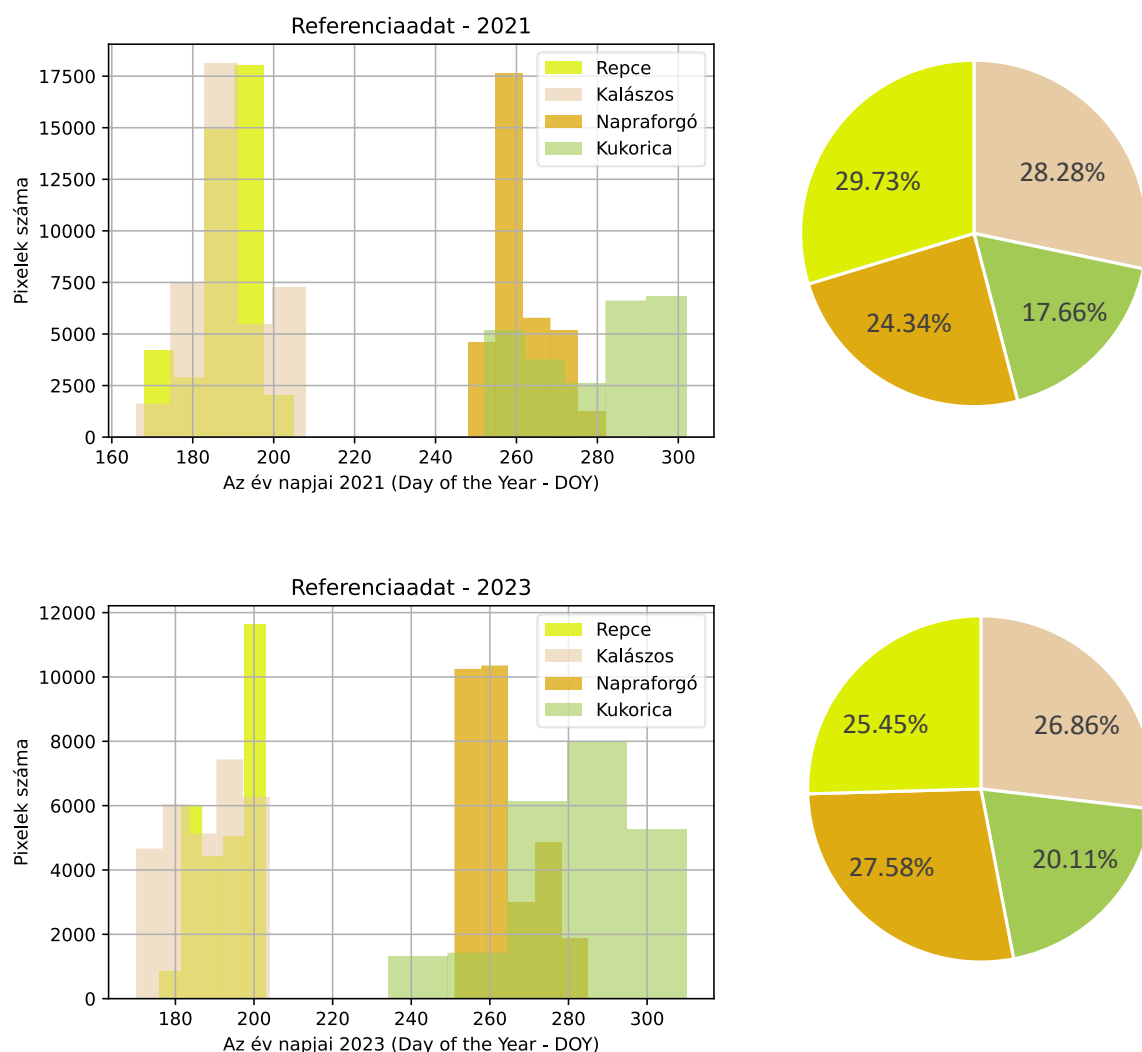
2023.09.07.

2023.09.09.

4. ábra: Napraforgó parcella Planet valószínűs (RGB) optikai felvételen betakarítás előtt (balra) és betakarítás után (jobbra).

A betakarítás idejét meghatározó deep learning algoritmus tanításához a kérelem-adatbázist a négy legfőbb hazai növénykategóriára korlátoztam, melyek: kalászos, kukorica, napraforgó, repce. A 2021-es évre 288, míg a 2023-as évre 296 parcelláról gyűjtöttem betakarítási adatot. A tanításhoz mindkét mintaév esetén kategóriánként 30 parcellát használtam fel, a validálást kategóriánként 10 parcellával végeztem, az adatbázisok fennmaradó részét pedig a tesztelésre szántam. Az 5. ábra hisztogramjain látható a kategóriák aratási

időpontjainak időbeli eloszlása a vizsgált években. A kategóriák pixelszám szerinti arányát pedig az 5. ábra kördiagramjai szemléltetik.



5. ábra: Statisztika az aratási időpontok meghatározásához felhasznált referenciaadatokról: az aratási időpontok időbeli eloszlása és a kategóriák pixelszám szerinti aránya kategóriánként 2021-ben (fent) és 2023-ban (lent).

Ahogy az 5. ábra hisztogramjain megfigyelhető, a négy főnövény aratási időpontjai két fő periódusba esnek. A korai időszakban, június 15-től július 29-ig (DOY 165-210.) a kalászos növények és a repce aratása egymással átfedően zajlik. A késői időszakban, augusztus 25. és november 6. között (DOY 237-310.) takarítják be a kukoricát és a napraforgót, szintén megközelítőleg egyszerre. A korai növények esetén a 2021-es és 2023-as időbeli eloszlás igen hasonló. Az utóbbi periódus esetén azonban komoly különbség tapasztalható a két vizsgált év között, hiszen a kukorica aratása 2023-ban egy jóval szélesebb időszakra esik, mint 2021-ben. A napraforgónál a két év eloszlása megközelítőleg szintén egyezik.

3. MÓDSZERTAN

Ebben a fejezetben ismertetem a kutatás módszertanát. Először megemlítek néhányat a szakirodalomban megtalálható Sentinel-1 idősorok alkalmazásával készült vetésszerkezeti térképezés, illetve aratási monitoring módszerekből. Ezután röviden ismertetem a konvolúciós neurális hálózatok működését és alkalmazási területeit, valamint az osztályozási és regressziós típusú mélytanulási feladatok jellemzőit. Ezután részletezem a kutatásban használt eljárásokat. Bemutatom a felvételek zajszűréséhez felhasznált eljárásokat: a Refined Lee speckle filtert és a Non-Local Means módszert, illetve ezek hatását a speckle effektusra. Ezt követően rátérek a kutatásban szereplő mélytanulási algoritmusokra, bemutatom a bemeneti adatstruktúrát, a felhasznált szoftvereket, az osztályozási és a regressziós modellt. Végül összefoglalom az elvégzett vizsgálatokat különböző szempontok szerint.

3.1 Szakirodalmi áttekintés

A Sentinel-1 SAR idősorok alkalmazása mezőgazdasági monitoring feladatokra igen elterjedt. Előnye az optikai felvételekkel szemben, hogy a felhők zavaró hatása nem érvényesül, így egyenként az idősor, ami számos módszer esetén meghatározó. Szakirodalmi eredmények alapján a VV és VH szórási együttható mennyiségek idősoros vizsgálata alkalmas nagy pontosságú növénykultúra-osztályozásra (Xu et al., 2019) és az aratás időpontjának meghatározására is (Kavats et al., 2019). A legtöbb, széles körben alkalmazott módszer a vetésszerkezet meghatározására Random Forest algoritmust alkalmaz (Dobrinic et al., 2021, Son et al., 2017, Pacskó et al., 2023), az aratási időpont meghatározásához pedig a VV és VH interferometrikus koherencia idősorok elemzése a jellemző (Shang et al., 2020, Amherdt et al., 2021, Sipőcz, 2021).

A konvolúciós neurális hálózat (CNN) (Fukushima, 1988) napjainkban az egyik legnépszerűbb algoritmustípus. A CNN modellek két fő műveletcsoportból, illetve ezek kombinációjából állnak. Elsők általában a konvolúciós rétegek, amelyekkel kiemelhetők az input adatban lévő fontos részletek. A konvolúció művelete egy mozgó ablakkal működik, így ezek a mintázatok az adatban elfoglalt helyüktől függetlenül detektálhatók. A keresett mintázatok számát a filterek száma, míg méretüket a filterek mérete szabja meg. A neurális – vagy más néven *dense* vagy *fully connected* – rétegek minden eleme kapcsolódik az előző réteg minden eleméhez. Ezeket a kapcsolatokat egy-egy súly jellemzi. A neurális réteg kimenete minden neuron esetén a bemenetek súlyozott összege. A réteg sűrűségét a neuronok száma határozza meg. Az utolsó réteg neuronjainak száma meghatározza a kimenet paramétereinek

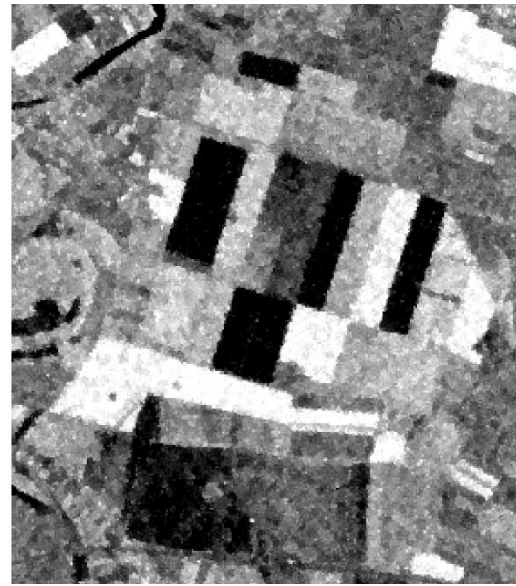
számát. A tanulás folyamata során a neurális rétegek súlyai, illetve a konvolúciós mátrixok értékei folyamatosan frissülnek egy visszacsatolási – úgynevezett *backpropagation* – eljárás szerint. Ezt a műveletet a veszteségfüggvény – azaz *loss* – minimalizálása határozza meg. A veszteségfüggvény definiálásával meghatározható, hogy a hálózat milyen metrikával mérje saját tévesztését. Így az algoritmus a tanulás során folyamatosan optimalizálja önmagát a feladat elvégzésére.

A CNN algoritmustípust először képek automatizált kiértékelésére (szegmentálás, képfelismerés) fejlesztették ki, de mára az adattudomány szinte minden területén hasznosítják, így a földmegfigyelésben is. Az algoritmus mind osztályozási, mind regressziós feladatok elvégzésére alkalmas lehet. Osztályozás esetén a bemeneti adatot előre meghatározott számú, ismert kategóriákba soroljuk, például felszínborítási, vetésszerkezeti kategóriák. Li és társai (2023) szerint CNN algoritmussal 97%-os teljes pontosság is elérhető a vetésszerkezeti térképezésben. A regressziós algoritmus akkor szükséges, ha a bemeneti adatból egy folytonos változót szeretnénk diszkrét értékkel közelíteni, például mezőgazdasági parcellák hozamát. Sayago és Bocco (2018) munkája arra mutat rá, hogy a kukorica és szójabab parcellák hozamának becslése esetén a mélytanulási regresszió akár 95%-os pontosságra is képes.

Kutatásom során a vetésszerkezeti térképezést CNN osztályozással, míg az aratási időpont meghatározását CNN regresszióval végeztem, mindkét esetben VV és VH szórási együtttható idősorokon.

3.2 Térbeli szűrés

A szintetikus apertúrájú radar felvételek a sok felszíni szóró objektumról visszaverődő hullámok koherens interferenciája miatt egy egyedi zajjal terheltek, melyet *speckle* zajnak nevezünk (Lee és Pottier, 2009). Ez a jelenség korlátozza a közvetlen vizuális interpretáció lehetőségét, és az automatizált módszerek eredményére is hatással van. A 2.2 alfejezetben említett feldolgozás során használt Refined Lee szűrési eljárás jelentősen mérsékli a *speckle* zaj jelenlétét az adatban (lásd 6. ábra).

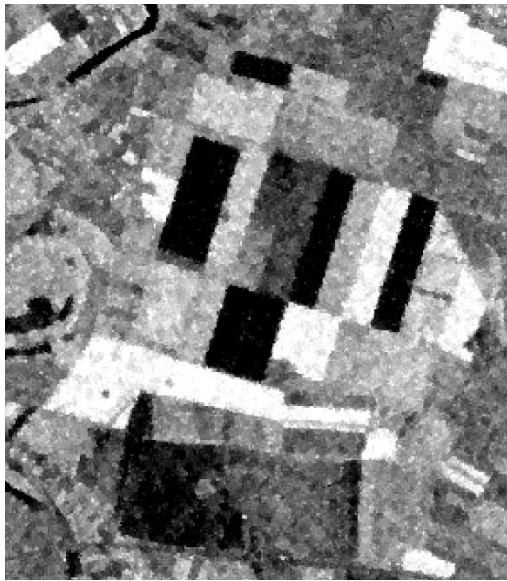


6. ábra: A *Refined Lee* szűrés hatása a radarfelvételekre. A szűrés előtti *VV* szórási együttható (balra) és a szűrt *VV* szórási együttható (jobbra).

A 6. ábrán látható azonban, hogy a zaj hatása a felvételeken továbbra is észlelhető. Annak érdekében, hogy ez az eredmény-rasztereken ne okozzon pixelszintű hibát, a felvételeken további képszűrést végeztem a Non-Local Means eljárással (Buades et al., 2011). Ennek a módszernek a lényege, hogy a képet kisebb részletekben vizsgálja. Ezekben belül hasonló pixelcsoportokat keres, majd ezek súlyozott átlagát helyettesíti be minden csoport helyére. Három legfontosabb paramétere a keresési ablak mérete – azaz *search window size* –, ami az egyszerre megvizsgált képrészlet nagyságát adja meg, a mintaablak mérete – vagyis *template window size* –, ami a keresett pixelcsoportok méretét definiálja, illetve a *h* paraméter, ami a szűrő erősségére vonatkozik. A kutatás során több paraméterbeállítást kipróbáltam, végül a következők bizonyultak a leghatékonyabbnak:

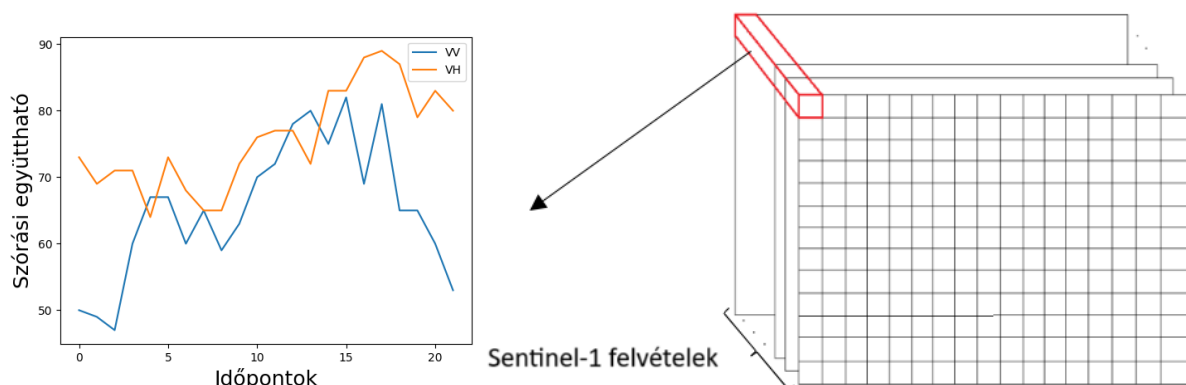
- *search window size* = 21 pixel
- *template window size* = 7 pixel
- *h* = 4.

A 7. ábra szemlélteti, hogy ezekkel a paraméterekkel jelentősen csökkent a zaj mértéke, bizonyos fokú heterogenitás és textúra viszont kivehető maradt a radarfelvételeken.



7. ábra: A Non-Local Means filter hatása a radarfelvételekre. Csak Refined Lee filterrel szűrt VV szórási együttható (balra) és utólag Non-Local Means szűrővel is feldolgozott VV szórási együttható (jobbra).

3.3 Bemeneti adatstruktúra



8. ábra: Bemeneti adatstruktúra.

Az osztályozási és a regressziós feladat során egyaránt a bemeneti adat egy pixel szűrt VV és VH szórási együttható idősora volt (lásd 8. ábra). Mindkét feladathoz egydimenziós, időbeli konvolúciót alkalmazó modellt építettem. Az egydimenziós konvolúció alapja, hogy a pixeleket a mélytanulási algoritmus egymástól teljesen függetlenül kezeli, térbeli kapcsolatokat nem vizsgál. Időbeli mintákat, jellemzőket keres és emel ki az idősorból. Ezek segítségével a neurális rétegek könnyebben képesek meghatározni a kimenet kívánt értékeit, jelen esetben osztályozáskor a vetésszerkezeti kategóriát, regresszió esetén pedig az aratás időpontját.

Az osztályozást és a regressziót végző forráskódot Python programnyelvben írtam Jupyter notebook felületen (Kluyver et al., 2016). A GeoTIFF formátumú fájlok kezeléséhez a

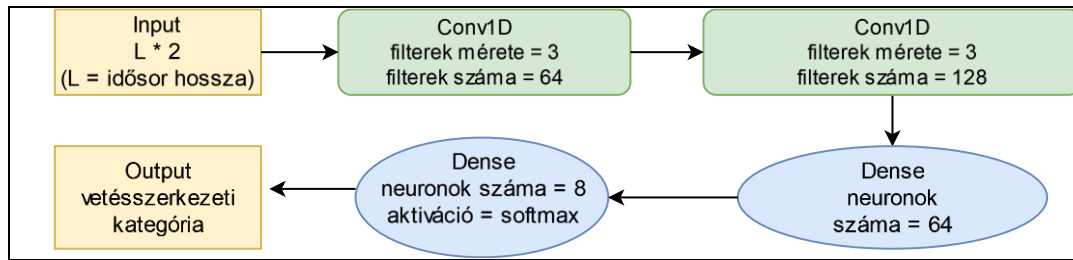
rasterio modult (Gillies et al., 2013), az adatkonvertálási lépésekhez és számításokhoz a *numpy* (Harris et al., 2020) és *scikit-learn* (Pedregosa et al., 2011) modulokat használtam. A modellek megalkotása, a tanítás és az osztályozás a *keras* (Chollet et al., 2015) modul függvényeinek felhasználásával történt, amely az egyik legnépszerűbb mélytanulási platform, a *tensorflow* (Abadi et al., 2015) alkalmazásprogramozási felülete.

3.4 Augmentáció

Általánosságban elmondható, hogy a tanítóadat mennyisége és minősége az egyik legfontosabb tényező a gépi tanulási algoritmusok használata során (Mumuni és Mumuni, 2022). Az aratási időpontok esetében azonban a referenciaadat 2.3.2 alfejezetben részletezett előállítása idő- és adatigényes folyamat. Népszerű módszer a referenciaadat mennyiségének növelésére az augmentáció (Mumuni és Mumuni, 2022). Ez olyan műveleteket jelent, melyekkel az adatmennyiség kiterjeszthető úgy, hogy a már meglévő adatokból új mintákat származtatunk, például tükrözéssel, forgatással vagy időbeli eltolással. Kutatásom során a regressziós algoritmus esetén időbeli eltolást alkalmaztam az adatokon augmentációként úgy, hogy az idősor kezdő és záró időpontját, illetve az aratás relatív időpontját (lásd 3.6 alfejezet) egy, illetve két felvétellel eltoltam. Így minden referencia-pixelt három, egymástól kissé különböző Sentinel-1 idősorral tanult meg az algoritmus. Ezáltal háromszorosára nőtt a tanítóadat mennyisége, a konvolúció pedig a 3.1 alfejezetben leírtak szerint biztosította, hogy az aratási eseményt az idősorban elfoglalt helytől függetlenül felismerje az algoritmus. Az osztályozási modell esetén augmentációt nem használtam, mert ott jóval több referenciaadat állt rendelkezésemre (lásd 2.3.1 alfejezet).

3.5 Osztályozási modell

Az osztályozási algoritmus bemenete a 3.3 alfejezetben említett Sentinel-1 pixelidősorok mellett a 2.3.1 alfejezetben említett, növénykategóriákat tartalmazó vektoros referenciaadat voltak. A modell két időbeli egydimenziós konvolúciós réteget (*Conv1D*) és két *dense* réteget tartalmazott. A második *dense* réteg esetén aktivációs függvényt alkalmaztam annak érdekében, hogy a kimenet értékei a megfelelő skálára kerüljenek. A *softmax* aktivációs függvény (Boltzmann, 1868) biztosította, hogy a háló minden pixelre nyolc darab, 0 és 1 közötti értéket adjon vissza azzal a feltétellel, hogy a nyolc érték összege pontosan 1 legyen. Ezek az értékek reprezentálják, hogy a pixel az egyes osztályokba mekkora valószínűséggel tartozik. Ezen valószínűségek közül a legnagyobb tekinthető az osztályozás eredményének. Az osztályozási modell felépítése a 9. ábrán látható.



9. ábra: Az osztályozási modell felépítése.

3.6 Regressziós modell

Az aratás időpontját meghatározó regressziós algoritmus bemenete a 3.3 alfejezetben említett Sentinel-1 pixelidősorok mellett a Planet idősorok és poligonadatbázis felhasználásával előállított referenciaadat (lásd 2.3.2 alfejezet) voltak. Itt az év napjában (*Day of the Year*, azaz DOY) meghatározott aratási időpontokat 0 és 1 közé normalizáltam a modell kimenetének egyszerűsítése érdekében a következő képlettel:

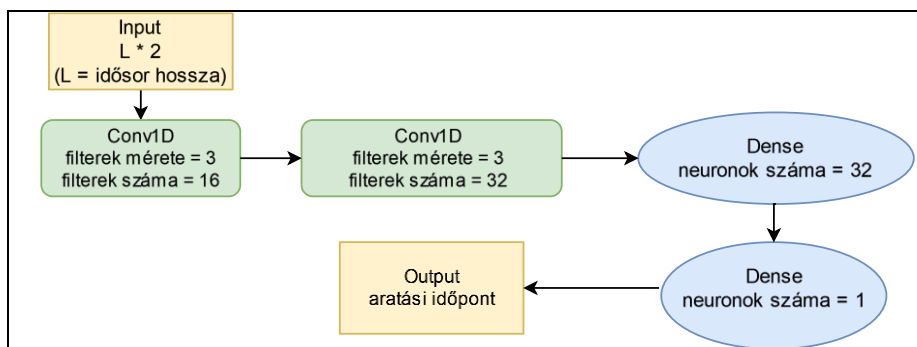
$$N(x) = \frac{DOY_x - DOY_{min}}{DOY_{max} - DOY_{min}},$$

ahol

- DOY_x a pixelhez tartozó aratási időpont [Day of the Year],
- DOY_{min} az aktuálisan felhasznált idősor első időpontja [Day of the Year],
- DOY_{max} az aktuálisan felhasznált idősor utolsó időpontja [Day of the Year] és
- $N(x)$ a pixelhez és az aktuális idősorhoz tartozó normalizált aratási időpont.

Így minden bemenet egy 0 és 1 közötti relatív értéket kapott az alapján, hogy az aktuális idősoron belül hol van az aratási esemény. Ez biztosította azt, hogy a 3.4 alfejezetben említett augmentáció során az időbeli eltolással együtt az aratás relatív időpontja is tolódjon.

Az architektúra két időbeli konvolúciós (*Conv1D*), valamint az előzőhöz hasonlóan két *dense* rétegből áll. Ez minden pixelre egy darab, 0 és 1 közötti érték adott, amelyeket DOY értékévé konvertálva kaptam a modell által meghatározott aratási időpontot. Fontos, hogy mivel a háló kimenete egyetlen időpont, ez a módszertan nem kezeli jól azokat a parcellákat, ahol egynél több betakarítás történt az év során. A regressziós modell felépítését a 10. ábra szemlélteti.



10. ábra: A regressziós modell felépítése.

3.7 Vizsgálatok

A felépített modelleket a két kiválasztott mintaéven teszteltem. A 2021-es évben volt utoljára 6 napos a Sentinel-1 adatok időbeli felbontása, így meg tudtam vizsgálni az idősor sűrűségének hatását az eredményre. A 2023-as év bevonásával pedig lehetőségem nyílt elemezni, hogy az algoritmus azonos pontossággal működik-e különböző évek esetén, illetve, hogy az előtanított algoritmus átvihető-e az egyes évek között. A 2021-es és 2023-as tanítóadatok együttes felhasználásával arra a kérdésre kerestem a választ, hogy létrehozható-e egy olyan előtanított modell, ami bármely évre megbízhatóan működik. Az elvégzett vizsgálatokat a 2. táblázat összesíti.

Év	2021		2023
Vizsgálatok	6 napos felbontás	12 napos felbontás	12 napos felbontás
Osztályozás	x	x	x
Regresszió	x	x	x
Regresszió – a másik évvel tanítva	–	x	x
Regresszió – a két év együtt tanítva	–	x	x

2. táblázat: Az elvégzett vizsgálatok összesítése.

3.8 A pontosság vizsgálata

Gépi tanulási módszerek esetén a 2.3 alfejezetben említett teszt adathalmazok arra szolgálnak, hogy a módszer pontosságát megállapíthassuk. A pontosság mérésére azonban fontos a megfelelő matematikai formulák használata, ami az osztályozás és a regresszió esetén eltérő.

A vetésszerkezeti osztályozás esetén ehhez *tévesztési (konfúziós) mátrixot* használtam, melynek mérete $M \times M$, ahol M az osztályok száma. Oszlopai a modell által megadott, sorai a

valós (tesztadat-szerinti) osztályok pixelszámát ábrázolják. Így a főatlóban szereplő elemek a helyesen osztályozott pixelek mennyiségét jelölik, míg a többi érték összege a tévesztések számát. A tévesztési mátrix összes cellájának összege megadja a vizsgált pixelek számát. Az osztályozásaim összpontosságát a *teljes pontosság* (Overall Accuracy) és a *Cohen-féle kappa pontosság* (Cohen's Kappa) mérőszámokkal állapítottam meg, melyek definíciói a következők:

$$\text{Teljes pontosság} = \frac{\text{Helyesen osztályozott minták száma}}{\text{Összes minta száma}} * 100$$

Teljes pontosság: Megmutatja, hogy a tesztadat mintáinak hány százalékát osztályozta helyesen a modell.

$$\text{Cohen – féle kappa pontosság} = \frac{p_0 - p_e}{1 - p_e},$$

ahol

- p_0 a teljes pontosság,
- $p_e = \frac{1}{N^2} \sum_k n_{k1} * n_{k2}$,

ahol

- N az összes minta száma,
- k a kategóriák száma,
- n_{k1} az osztályozás által a k . kategóriába sorolt minták száma és
- n_{k2} a tesztadatban a k . kategóriába tartozó minták száma.

Cohen-féle kappa pontosság: Robusztus statisztikai mutató, amely összehasonlítja a megfigyelt pontosságot a várható pontossággal. Ellentétben a teljes pontossággal, figyelembe veszi a megegyezések véletlenszerű bekövetkezésének lehetőségét is.

Az osztályozó algoritmus összpontosságán túl a kategóriánkénti pontosságok is mérhetők, én a *felhasználói pontosság* (User's accuracy) és *előállítói pontosság* (Producer's accuracy) metrikákat számítottam ki a következő képletekkel:

$$\text{Felhasználói pontosság} = \frac{c_{ii}}{o_i} * 100,$$

ahol

- O_i a tévesztési mátrix i . oszlopbeli elemeinek összege és
- C_{ii} a tévesztési mátrix i . sorának i . eleme.

Felhasználói pontosság: Megmutatja, hogy az algoritmus által az egyes osztályokba sorolt pixelek hány százaléka tartozik valóban abba az osztályba a tesztadat szerint.

$$\text{Előállítói pontosság} = \frac{c_{ii}}{s_i} * 100,$$

ahol

- S_i a tévesztési mátrix i . sorbeli elemeinek összege és
- C_{ii} a tévesztési mátrix i . sorának i . eleme.

Előállítói pontosság: Megmutatja, hogy osztályonként a tesztadat hányad részét sorolta helyes osztályba az algoritmus.

Az aratási időpont meghatározását végző regressziós algoritmus esetén a hiba mértékét legegyszerűbben az *átlagos abszolút hiba* (Mean Average Error, vagy MAE) adja meg, melynek képlete:

$$\text{Átlagos abszolút hiba} = \frac{\sum_{i=1}^n |e_i|}{n},$$

ahol

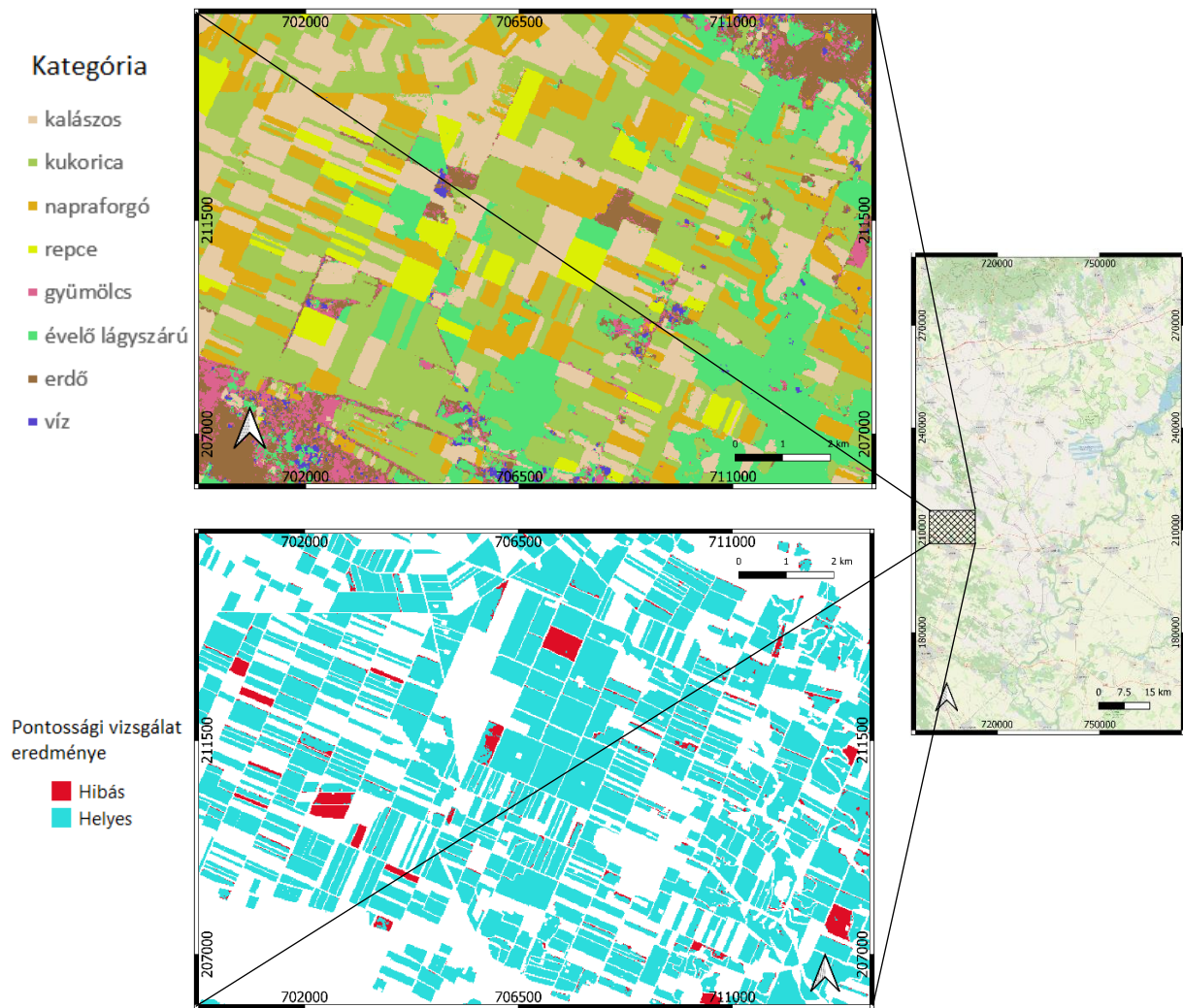
- n a minták száma és
- e_i az i . minta hibája a tesztadathoz képest [DOY].

Az átlagos abszolút hiba jelen kutatásban azt adja meg, hogy az algoritmus átlagosan hány nappal téveszti el az aratás időpontját a tesztadathoz képest. Jelen esetben, mivel a tesztadatban szereplő minták növénytípus szerint is csoportosíthatók, a hibaértéket kiszámoltam kategóriánként is.

4. EREDMÉNYEK ÉS DISZKUSSZIÓ

Ebben a fejezetben bemutatom és értelmezem kutatási eredményeimet. Elsőként az osztályozás eredményeként létrejött térképek vizuális tartalmáról, teljes és kategóriánkénti pontosságáról, valamint a 6 és 12 napos idősorok eredményeinek különbségéről értekezek. Ezután bemutatom a regresszióval előállított aratási időpont térképeket vizuálisan, illetve a teljes és kategóriánkénti hibáikat. Ezt követően diszkutálom az eltérő időjárású évek, illetve az idősor sűrűségének befolyását az évenként tanított modelleken, majd ezen modellek másik éven elért eredményeit mutatom be, végül a két éven tanult algoritmus eredményeit részletezem.

4.1 Osztályozás



11. ábra: A 2021-es, 6 napos idősor felhasználásával előállított vetésszerkezeti térkép (fent) és a tesztadathoz viszonyított tévesztési térképe (lent).

A 11. ábrán a 2021-es év 6 nap felbontású idősorából készült osztályozási eredményt mutatom be. Az eltérő növények határai élesek és jól kivehetők, csak kisebb, pixelszintű zajok láthatók a térképen. A vetésszerkezeti szempontból azonos művelésű területek objektumként is azonosíthatók (lásd 11. ábrán fent). A tévesztési térkép azt mutatja, hogy a parcellák túlnyomó többségét helyesen osztályozta az algoritmus, néhány kivételtől eltekintve (lásd 11. ábrán lent). A parcellánál kisebb, pixel vagy pixelcsoport szintű hibák adódhatnak a radarfelvételeken megmaradt zajból, de valós heterogén zónákat is jelezhetnek. A további két vetésszerkezeti térképet a függelék C) és D) ábráin közlöm, mert a 3. táblázatban később bemutatott statisztikák és a vizuális ellenőrzés is azt mutatja, hogy az egyes térképek között a különbség nem számottevő.

2021 – 6 nap	Felhasználói pontosság	Előállítói pontosság	2021 – 12 nap	Felhasználói pontosság	Előállítói pontosság
kalászos	98,50%	94,35%	kalászos	97,45%	92,02%
kukorica	93,00%	92,65%	kukorica	87,57%	94,51%
napraforgó	95,48%	97,80%	napraforgó	97,08%	96,16%
repce	96,41%	98,32%	repce	97,13%	98,09%
gyümölcs	51,88%	84,01%	gyümölcs	56,52%	75,32%
évelő lágyszárú	93,20%	92,46%	évelő lágyszárú	89,04%	90,40%
erdő	92,62%	93,71%	erdő	91,76%	93,64%
víz	87,64%	99,46%	víz	86,68%	98,71%
Kappa pontosság	93,36%		Kappa pontosság	91,68%	
Teljes pontosság	94,79%		Teljes pontosság	93,45%	

2023 – 12 nap	Felhasználói pontosság	Előállítói pontosság
kalászos	97,91%	95,81%
kukorica	92,22%	91,24%
napraforgó	97,04%	97,46%
repce	94,25%	99,76%
gyümölcs	39,98%	82,00%
évelő lágyszárú	93,33%	87,56%
erdő	87,05%	89,44%
víz	72,73%	99,21%
Kappa pontosság	91,84%	
Teljes pontosság	93,89%	

3. táblázat: Az osztályozási eredmények pontossági statisztikái: 2021 – 6 napos időbeli felbontás (balra fent), 2021 – 12 napos időbeli felbontás (jobbra fent) és 2023 (lent). Az aratási időpont szempontjából is vizsgált kategóriák félkövérrel kiemelve.

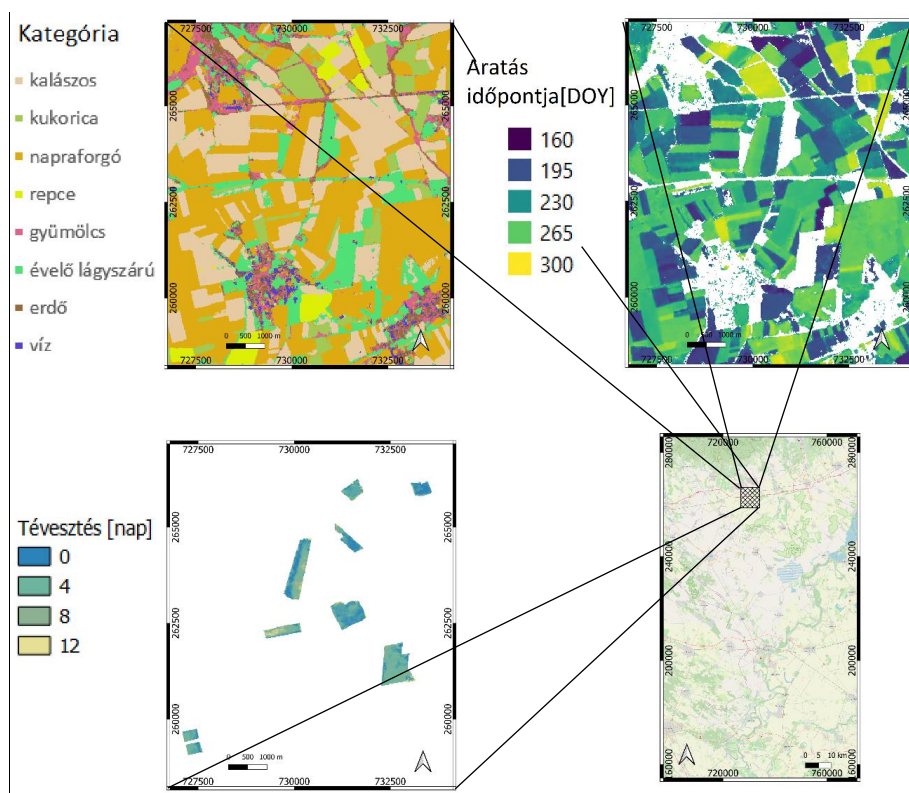
Az osztályozási algoritmus általánosan, minden vizsgálat esetén 90% feletti teljes és kappa pontossággal működik (lásd 3. táblázat). A legalacsonyabb pontosságú a 2021-es év 12 napos időbeli felbontással készült térképe, melynek teljes pontossága 93,45%, a legjobb eredményt pedig a 2021-es év 6 napos felbontású idősorával érte el az algoritmus, ez 94,79%. A két eredmény közti csekély, 1,34% különbség mutatja, hogy a sűrűbb idősor jobb eredményt ad ugyan, de ez a hatás nem olyan jelentős. A 2023-as év teljes pontossági eredménye szintén közel 94%, ami alapján az algoritmus felépítése eltérő időjárású évekre hasonlóan megfelelő.

A kategóriák tekintetében elmondható, hogy az aratási időpont meghatározásánál vizsgált növények közül az algoritmusok mind felhasználói, mind előállítói pontossága stabilan

90% feletti. Kiemelendő a repce 94,25-99,76%-os, illetve a napraforgó állandó, 95-97% körüli megbízhatósága. Az egyéb kategóriák közül a gyümölcsösök esetén teljesítenek a leggyengébben az algoritmusok. Ennek oka feltehetően az, hogy a gyümölcsös a különböző gyümölcsfajták miatt távérzékelési szempontból igen heterogén kategória, valamint a gyümölcsfák között általában gyepterület található, ezért nagy arányban keveredhet például élő lágyszárú növényekkel. Itt a felhasználói pontosság 40 és 56% között, míg az előállítói pontosság 75 és 84% között van. Az, hogy a felhasználói pontosság minden esetben jóval – akár 42%-kal – alacsonyabb, azt jelzi, hogy nagy az úgynevezett *fals pozitív*, azaz tévesen ebbe az osztályba sorolt pixelek száma.

Ezen eredmények alapján az osztályozott vetésszerkezeti térképek mind 6 napos, mind 12 napos sűrűségű idősor esetén alkalmasak arra, hogy az aratási időpont meghatározásához szükséges területeket, azaz a négy főnövényvel borított felszínt megbízhatóan lehatárolják.

4.2 Regresszió



12. ábra: A 2021-es, 6 napos idősorral létrehozott tematikus térképek: vetésszerkezeti térkép (balra fent), aratási időpont térkép (jobbra fent) és a tesztadathoz viszonyított aratási időpont tévesztési térkép maszkolva a vetésszerkezeti térkép kalászos, repce, kukorica és napraforgó kategóriáival (balra lent).

A vetésszerkezeti térképhez hasonló módon az aratási időpont térképen is kirajzolódnak az aratási időpont tekintetében hasonló növények által alkotott zónák (lásd 12. ábrán jobbra

fent). Ez nem feltétlenül jelent különböző növénykultúrát, ugyanis a 2.3.2 alfejezetben látható módon a kalászosok és a repce, illetve a napraforgó és a kukorica aratási időpontjai jelentős átfedésben vannak. A vetésszerkezettel együtt értelmezve megfigyelhető, hogy az aratási időpont térkép következetes, általában nem mond ellent a vetésszerkezetnek: a kalászos és repce aratási időpontja a becslés szerint is jellemzően a 160-210. nap közé, míg a napraforgó és kukorica betakarítása a 240-300. nap közé tehető. A 12. ábrán balra lent található tévesztési térkép azt mutatja, hogy a tesztadathoz képest melyik pixel aratási időpontját hány nappal tévesztette el a modell. Kis mintaterületen a modell hibája alacsony, a teljes mintaterületen viszont a hiba terjedelme nagy. A kiemelkedően magas (30-74 nap) hibák azonban jellemzően szegélypixelek, ahol egyrészt a felszín valós heterogenitása, másrészt tesztadat parcelláinak geometriai hibája is okozhatja ezt a magas értéket. Megfigyelhetők azonban parcellán belüli eltérések is, amelyeket feltehetően a teljes idősorban található fejlődési eltérések okoztak.

4.2.1 Egyéves idősorral tanított modellek azonos éven tesztelve

2021 – 6 nap		2021 – 12 nap		2023 – 12 nap	
Kategória	Átlagos abszolút hiba [nap]	Kategória	Átlagos abszolút hiba [nap]	Kategória	Átlagos abszolút hiba [nap]
kalászos	10,28	kalászos	10,82	kalászos	9,21
repce	6,55	repce	6,01	repce	6,15
napraforgó	5,76	napraforgó	8,07	napraforgó	7,81
kukorica	8,58	kukorica	9,3	kukorica	12,04
Átlag	7,89	Átlag	8,49	Átlag	8,64

4. táblázat: Regressziós aratási időpont becslések pontossági statisztikái: 2021 – 6 napos időbeli felbontás (balra), 2021 – 12 napos időbeli felbontás (középen) és 2023 (jobbra).

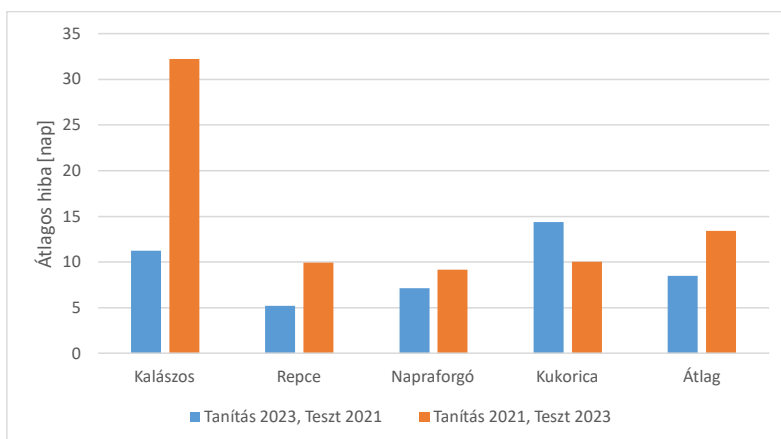
Az aratási időpont térképek átlagos abszolút hibaértékei minden esetben alacsonyak (lásd 4. táblázat). A legkisebb hibát ismét a 2021-es év 6 napos felbontású idősora véti, ez 7,89 nap. Fontos megemlíteni, hogy ez több, mint az időbeli felbontás. Mivel a 6 napos idősor a modell bemeneteként kétszer annyi paramétert jelent, valószínűleg egy szélesebb vagy mélyebb hálózattal a pontosság tovább növelhető lenne. Bár a 2021-es év 12 napos idősorral, illetve a 2023-as év 12 napos idősorral tanított algoritmusok hibája magasabb – 8,49, illetve 8,64 nap – az idősor sűrűségének ismeretében ez mégis jobb eredménynek tekinthető, hiszen ezek a hibák kisebbek, mint 12 nap, tehát ezek a modellek elérték, sőt meghaladták a módszertan pontosságának elméleti maximumát. Ez azzal magyarázható, hogy bár a radar idősor 12 napos

felbontású, a referenciaadat – a Planet műholdak visszatérési idejének megfelelően – akár napi felbontású is lehet. A bemenetekhez tartozó aratási időpont címke a 3.6 alfejezetben részletezett normalizálással egy 0 és 1 közötti lebegőpontos érték, melyre az algoritmus nagy pontossággal képes rátanulni.

A növénykategóriák pontosságai közötti különbség jelentős, és az egyes modellek között nagyjából állandó. Átlagosan a repce hibája a legkisebb, ez 6 és 6,55 nap között változik. A napraforgó esetében szintén alacsony a hiba, azonban itt nagyobb a különbség a vizsgált esetek között: a 2021-es, 6 napos idősor felhasználásával 5,76 napos hibát kaptam, míg a másik két esetben ez az érték 8 nap körüli. Itt megállapítható a sűrűbb idősor egyértelmű pozitív hatása. A legmagasabb hibák minden vizsgált esetben a kalászos és a kukorica esetén voltak. Kalászosnál a hiba egyenletesen magas, 9,2 és 10,8 nap között változik. A kukorica esetén 8,5-12 nap a tévesztés, ez szintén viszonylag nagy terjedelem. Ebben az esetben is a 6 napos felbontású idősor teljesít a legjobban, azonban itt a 8,5 napos hiba ismét kívül esik az idősor felbontásán. Ezen eredmények alapján a 12 napos idősorokkal tanított modellek, mint algoritmusok pontosabbak, de operatív felhasználásra a 6 napos idősorral tanított modellek alkalmasabbak.

Összegezve az évenként tanított modellek eredményeit elmondható, hogy pontosságuk minden esetben minden növényre megfelelő, a módszer jól alkalmazható. Azonban az is megfigyelhető, hogy a napraforgó és repce esetén következetesen pontosabb, mint a kalászosra és kukoricára.

4.2.2 Egyéves idősorral tanított modellek másik éven tesztelve



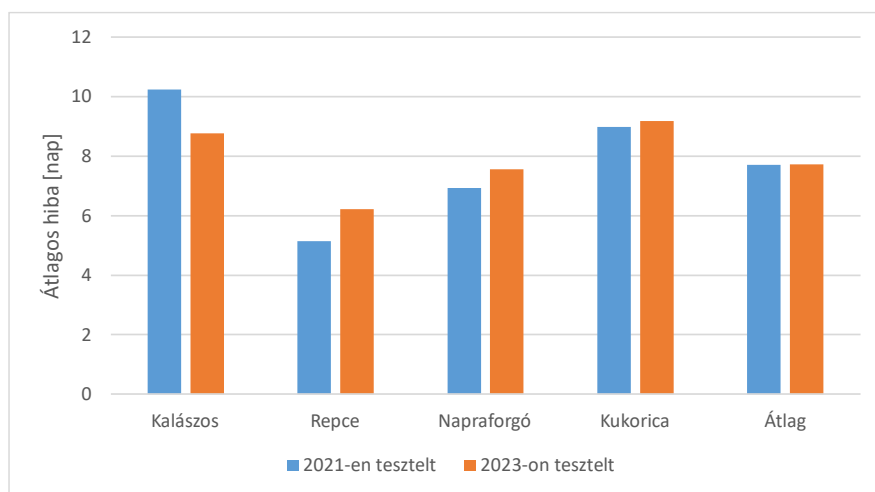
13. ábra: Egy éven tanított modellek másik éven tesztelt hibái kategóriánként.

2023-mal tanított, 2021-gyel tesztelt modell		2021-gyel tanított, 2023-mal tesztelt modell	
Kategória	Átlagos abszolút hiba [nap]	Kategória	Átlagos abszolút hiba [nap]
kalászos	11,22	kalászos	32,26
repce	5,22	repce	9,93
napraforgó	7,15	napraforgó	9,18
kukorica	14,38	kukorica	10,4
Átlag	8,51	Átlag	13,43

5. táblázat: Egy éven tanított modellek másik éven tesztelt hibái kategóriánként.

Az egy éven tanított modellek átültetése egy másik, adott esetben eltérő időjárású évre kiegyensúlyozatlan eredményt hoz, erre a 13. ábra és az 5. táblázat statisztikái rávilágítanak. Egyértelműen pontosabb a 2023-as éven tanított és 2021-es éven tesztelt modell, melynek átlagos abszolút hibája 8,51 nap, és kiemelkedő a repce 5,22 napos hibaértéke. Ez alátámasztja azt, hogy a szélsőségektől mentes éven megtanult karakterisztikák jobban alkalmazhatók általánosan, mint az erősen belvizes év paraméterei. A 2021-es éven tanított és 2023-as éven tesztelt modell átlagos hibája is viszonylag magas, 13,43%, különösen a kalászos esetén tapasztalt 32,3 napos átlagos hiba kiugróan nagy. Ez alapján ezen modell használata nem javasolt az aratási időpont meghatározására kalászos növények esetén. Összegezve, a transfer learning koncepció egy évvel tanítva jelen formájában nagyon megbízhatatlan eredményre vezet és a különböző növényekre nagyon eltérő pontosságú, így további fejlesztést igényel.

4.2.3 Két év idősoraival tanított modell



14. ábra: Két éven együtt tanított modell egyes éveken tesztelt hibái kategóriánként.

Mindkét évvel tanított, 2021-en tesztelt modell		Mindkét évvel tanított, 2023-on tesztelt modell	
Kategória	Átlagos abszolút hiba [nap]	Kategória	Átlagos abszolút hiba [nap]
kalászos	10,25	kalászos	8,77
repce	5,15	repce	6,22
napraforgó	6,93	napraforgó	7,6
kukorica	8,99	kukorica	9,18
Átlag	7,71	Átlag	7,72

6. táblázat: Két éven együtt tanított modell egyes éveken tesztelt hibái kategóriánként.

A kéteves adatokkal előtanított modell átlagos hibája mindkét évre alacsony, 7,71 és 7,72 nap (lásd 14. ábra és 6. táblázat). Ez a mindössze 0,01 különbség azt mutatja, hogy ez az algoritmus átlagosan ugyanolyan jól működik a vizsgált évekre. Különbség mutatkozik azonban a kategóriák között, hiszen a 2021-es tesztelésen a kalászos hibája nagyobb (10,25 nap) mint a 2023-as időszorral (8,77 nap), ugyanakkor az összes többi növény esetén a 2021-es hiba az alacsonyabb. Az előző vizsgálatoknak megfelelően ebben az esetben is a repce hibája a legalacsonyabb, 5,15, illetve 6,22 nap. A napraforgó és a kukorica esetén az évek közötti stabilitást lehet kiemelni, előbbi esetén 0,63, utóbbi esetén csak 0,19 nap különbség van a két vizsgált év hibája között.

Mivel az átlagos abszolút hiba és a kategóriánkénti hibák is alacsonyabbak, mint abban az esetben, amikor a 12 napos idősorokat külön-külön tanítva és tesztelve használtam fel, a vizsgálat szerint az együtt tanítással a módszertan fejleszthető. Bár a kategóriák közötti különbségek az évek között nem jelentősek, valószínűleg még több év együttes felhasználásával az egyensúly tovább növelhető lenne.

5. ÖSSZEGZÉS

A fentiekben bemutatott kutatásom során egydimenziós konvolúciós neurális hálózatokat és Sentinel-1 SAR idősorokat használtam a vetésszerkezet és az aratás időpontjának meghatározására. Módszereimet egy nagy méretű Tisza-menti mintaterületen teszteltem két vizsgált éven. A Sentinel-1 felvételeket VV és VH szórási együttható idősorokká dolgoztam fel. A speckle zaj jelenséget a feldolgozás után a Non-Local Means szűrési eljárással igyekeztem tovább csökkenteni. A vetésszerkezeti osztályozást nyolc kategóriára, a regressziós

aratási időpont meghatározást négy főnövényre végeztem. Vizsgáltam az azonos évvel, eltérő évvel és több éves adattal tanított modellek alkalmazhatóságát.

Osztályozásaim pontossága magas, 93-95%. A négy, regresszióval is vizsgált növény kategóriánkénti pontossága minden esetben 90% feletti. Az osztályozás pontosságát az időbeli felbontás növelése – 12 nap helyett 6 nap – kis mértékben – 1,5%-kal – javítja.

A vizsgált regressziós módszerek abszolút hibája átlagosan 7-8 nap. Minden modell jobb eredményt adott repce és napraforgó aratására, mint kukoricára és kalászosra. A 6 napos időbeli felbontású idősorokkal tanítva a pontosság nagyobb, mint a 12 napos idősor esetén, de a hiba átlagosan 6 napnál nagyobb marad. Az évenként tanított algoritmus pontossága megfelelő, 7,5-8,5 napos átlagos hibával működik, tehát a minden évre külön tanított algoritmus egy megfelelő lehetséges megoldást nyújt a körülírt problémára.

Az egyik éven tanult algoritmus átvihetősége másik évre megbízhatatlan eredményt ad, a hiba szórása a kategóriák között jelentősen megnő. A 2023-as, időjárási szélsőségektől mentes éven tanult modell időben jobban átvihető, mint a 2021-es, belvizes éven tanított algoritmus. Vizsgálataim alapján az egy éven tanult algoritmus időbeli kiterjesztése jelen formájában nem megfelelő, további fejlesztést igényel.

A két éven együtt tanult modell nagyon stabil. Relatív alacsony, 7,7 napos átlagos abszolút hibával működik mindkét vizsgált évre, és a kategóriák hibái közti különbség is mérséklődik. Kutatásom alapján egy több évvel előtanított, általánosan működő regressziós aratási időpont becslő algoritmus létrehozható.

Első eredményeim felhasználásra kerültek az Országos Vízügyi Főigazgatóság árvízi védekezésében 2024 szeptemberében. A jövőben a Mezőgazdasági Kockázatkezelési Rendszer (MKR), valamint várhatóan precíziós mezőgazdasággal foglalkozó cégek is felhasználhatják az eredményeket.

Következő lépésként a két évnyi adattal előtanított modell fejlesztendő, több év és évenként arányosan több tanítóadat felhasználásával a pontosság és a stabilitás jó eséllyel növelhető, az eljárás kiterjeszthető más szántóföldi növényekre. Megvizsgálandó továbbá a módszer területi átvihetősége további, eltérő területekről származó tesztadatok bevonásával. Amennyiben a területi és időbeli hordozhatóság megvalósítható, előállhat egy olyan előtanított algoritmus, ami országosan, bármely évre képes megbízhatóan becsülni az aratás időpontját.

6. KÖSZÖNETNYILVÁNÍTÁS

Munkámat számos személy és szervezet támogatta. Az általam felhasznált Sentinel-1 űrfelvételeket az ESA Copernicus program keretében készültek, a letöltést a Copernicus Dataspace Ecosystem tette lehetővé. A Planet optikai űrfelvételeket ingyenesen biztosították számomra a *PlanetScope Education and Research Program* keretében, ezért köszönet a Planet Labs PBC munkatársainak. Az Egységes Kérelem adatbázist és a MePAR adatokat a Lechner Tudásközpont és a Magyar Államkincstár által kötött szerződés keretei között használhattam fel kutatási célra. Köszönet a Magyar Államkincstár munkatársainak.

Köszönöm Pacskó Viviennek a szakmai tanácsadást és támogatást, a radarfeldolgozási ismereteim és publikációs készségeim fejlesztésében nyújtott segítségét és a folyamatos motivációt. Köszönet Molnár Gábornak a mindig éles, lényeglátó szakmai észrevételeiért, tanácsaiért, az adminisztrációban nyújtott segítségével és a tartalmas, kellemes beszélgetésekért.

Hálás vagyok Kristóf Dániel osztályvezetőnek a szakmai iránymutatásaiért, valamint kollégáimnak, a Lechner Tudásközpont távérzékelési szakértőinek, kiemelten Richter-Cserey Mátyásnak a mélytanulási módszerek megértésében nyújtott segítségével, Birinyi Edinának a téma felvetéséért és az előzetes módszerek bemutatásáért, továbbá Gera Dávid Ákosnak a referencia-adatbázisok előállításáért.

Köszönöm BSc szakdolgozatom témavezetőjének Friedl Zoltánnak a szakmai utam elindításában játszott szerepét és a baráti beszélgetéseket.

7. IRODALOMJEGYZÉK

ABADI, M., BARHAM, P., CHEN, J. ET AL. (2016): Tensorflow: A system for large-scale machine learning. In: *12th USENIX Symposium on Operating Systems Design and Implementation, OSDI 16*, pp. 265–283. <https://doi.org/10.48550/arXiv.1605.08695>

AGROINFORM (2023): Tovább csökkent a gazdaságok száma, átalakulóban a birtokszerkezet Magyarországon. <https://www.agroinform.hu/gazdasag/agrarium-2023-ksh-mezogazdasagi-adatok-67516-001> [Megtekintve: 2024.10.28.]

ALZUBAIDI, L., ZHANG, J., HUMAIDI, A. J. ET AL. (2021): Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of Big Data*, 8(1), 53. <https://doi.org/10.1186/s40537-021-00444-8>

- AMHERDT, S., DI LEO, N. C., BALBARANI, S., PEREIRA, A., CORNERO, C., PACINO, M. C. (2021): Exploiting Sentinel-1 data time-series for crop classification and harvest date detection. *International Journal of Remote Sensing*, 42(19), pp. 7313–7331. <https://doi.org/10.1080/01431161.2021.1957176>
- BIRINYI, E., KERN, A., KRISTÓF, D., HOLLÓS, R., BARCZA, Z. (2024a): Analysis of extreme hydrological events over the Great Hungarian Plain based on Earth Observation data. *EGU General Assembly 2024*, Vienna, Austria, 14–19 Apr 2024, EGU24-21292. <https://doi.org/10.5194/egusphere-egu24-21292>
- BIRINYI, E., KRISTÓF, D., KERN, A. (2024b): Possibilities of Radar-Based Inland Excess Water Mapping in Google Earth Engine During the Single-Satellite Sentinel-1 ERA for a Study Area in Hungary. *IEEE International Geoscience and Remote Sensing Symposium*, pp. 5117–5120. <https://doi.org/10.1109/IGARSS53475.2024.10640894>
- BOLTZMANN, L. (1868): Studien über das Gleichgewicht der lebendigen Kraft zwischen bewegten materiellen Punkten. *Wiener Berichte*, 58, pp. 517–560. <https://doi.org/10.1017/CBO9781139381420.006>
- BUADES, A., COLL, B., MOREL, J-M. (2011): Non-Local Means Denoising. *Image Processing On Line*, 1, pp. 208–212. https://doi.org/10.5201/ipol.2011.bcm_nlm
- CAP (2012): European Commission, Directorate-General for Agriculture and Rural Development, The Common Agricultural Policy – A partnership between Europe and farmers, Publications Office. <https://data.europa.eu/doi/10.2762/31102>
- CHOLLET, F. ET AL. (2015): Keras v1.0.8. Retrieved from <https://github.com/fchollet/keras>
- COPERNICUS BÖNGÉSZŐ (2024): <https://raster.lechnerkozpont.hu/apps/copernicus/> [Megtekintve: 2024.10.28.]
- DE ZAN, F., MONTI-GUARNIERI, A. (2006): TOPSAR: Terrain Observation by Progressive Scans. *IEEE Transactions on Geoscience and Remote Sensing*, 44(9), pp. 2352–2360. <https://doi.org/10.1109/TGRS.2006.873853>
- DHUMAL, R., RAJENDRA, Y., KALE, K., MEHROTRA, S. (2013): Classification of Crops from Remotely Sensed Images: A review. *International Journal of Engineering Research and Applications (IJERA)*, 3(3), ISSN: 2248–9622, pp. 758–761.

DOBRINIĆ, D., GAŠPAROVIĆ, M., MEDAK, D. (2021): Sentinel-1 and 2 Time-Series for Vegetation Mapping Using Random Forest Classification: A Case Study of Northern Croatia. *Remote Sensing*, 13(12), 2321. <https://doi.org/10.3390/rs13122321>

ESA (2023): Copernicus Sentinel data [2021-2023], processed by ESA.

ESA SENTINEL ONLINE (2024): Overview of Sentinel-1 Mission. <https://sentinel.esa.int/web/sentinel/missions/sentinel-1/> [Megtekintve: 2024.10.16]

ESA SENTINEL ONLINE (2022): European Space Agency, Mission ends for Copernicus Sentinel-1B satellite. https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-1/Mission_ends_for_Copernicus_Sentinel-1B_satellite [Megtekintve: 2024.10.16]

FARR, T. G., KOBRICK, M. (2000): Shuttle Radar Topography Mission produces a wealth of data. *Eos, Transactions AGU*, 81(48), pp. 583–585. <https://doi.org/10.1029/EO081i048p00583>

FIR, 2024: Földmegfigyelési Információs Rendszer, eFöld Portál. <https://efold.gov.hu/> [Megtekintve: 2024.11.05.]

FUKUSHIMA, K. (1988): Neocognitron: A hierarchical neural network capable of visual pattern recognition. *Neural Networks*, 1(2), pp. 119–130. [https://doi.org/10.1016/0893-6080\(88\)90014-7](https://doi.org/10.1016/0893-6080(88)90014-7)

GDAL/OGR CONTRIBUTORS (2024): GDAL/OGR Geospatial Data Abstraction Software Library v3.0.2. <https://gdal.org>

GILLIES, S., ET AL. (2013): Rasterio: geospatial raster I/O for Python programmers v1.2.10. Retrieved from <https://github.com/rasterio/rasterio>

HARRIS, C. R., MILLMAN, K. J., VAN DER WALT, S. J. ET AL. (2020): Array programming with NumPy. *Nature*, 585, pp. 357–362. <https://doi.org/10.1038/s41586-020-2649-2>

HENITS, L., SZERLETICS, Á., SZOKOL, D., SZLOVÁK, G., GOJDÁR, E., ZLINSZKY, A. (2022): Sentinel-2 Enables Nationwide Monitoring of Single Area Payment Scheme and Greening Agricultural Subsidies in Hungary. *Remote Sensing*, 14(16), 3917. <https://doi.org/10.3390/rs14163917>

HOSNA, A., MERRY, E., GYALMO, J., ALOM, Z., AUNG, Z., AZIM, M. A. (2022): Transfer learning: a friendly introduction. *Journal of Big Data*, 9, 102. <https://doi.org/10.1186/s40537-022-00652-w>

- JORDAHL, K., VAN DEN BOSSCHE, J., FLEISCHMANN, M. ET AL. (2014): GeoPandas: Python tools for geographic data v0.10.2. Retrieved from <https://github.com/geopandas/geopandas>. <https://doi.org/10.5281/zenodo.3946761>
- KAVATS, O., KHRAMOV, D., SERGIEIEVA, K., VASYLIEV, V. (2019): Monitoring Harvesting by Time Series of Sentinel-1 SAR Data. *Remote Sensing*, 11(21), 2496. <https://doi.org/10.3390/rs11212496>
- KLUYVER, T., RAGAN-KELLEY, B., FERNANDO P. ET AL. (2016): Jupyter Notebooks – a publishing format for reproducible computational workflows. In: Loizides, F., Schmidt, B. (Eds.), *Positioning and Power in Academic Publishing: Players, Agents and Agendas*, pp. 87–90. <https://doi.org/10.3233/978-1-61499-649-1-87>
- LEE, J. S. (1981): Refined filtering of image noise using local statistics. *Computer Graphics and Image Processing*, 15(4), pp. 380–389. [https://doi.org/10.1016/S0146-664X\(81\)80018-4](https://doi.org/10.1016/S0146-664X(81)80018-4)
- LEE, J. S., POTTIER, E. (2009): Polarimetric Radar Imaging: from Basics to Applications. CRC Press, Boca Raton. <https://doi.org/10.1201/9781420054989>
- LI, Q., TIAN, J., TIAN, Q. (2023): Deep Learning Application for Crop Classification via Multi-Temporal Remote Sensing Images. *Agriculture*, 13(4), 906. <https://doi.org/10.3390/agriculture13040906>
- MUMUNI, A., MUMUNI, F. (2022): Data augmentation: A comprehensive survey of modern approaches. *Array*, 16, 100258. <https://doi.org/10.1016/j.array.2022.100258>
- NÉBIH (2024): Mezőgazdasági Kockázatkezelési Rendszer. <https://portal.nebih.gov.hu/-/mezogazdasagi-kockazatkzelesi-rendszer> [Megtekintve: 2024.10.28.]
- PACSKÓ, V., BELÉNYESI, M., BARCZA, Z. (2023): Vetésszerkezeti térképek idősorának kategóriánkénti pontosságvizsgálata. In: *Az elmélet és gyakorlat találkozása a térinformatikában XIV.* (szerk.: Abriha-Molnár, V. É.), Debreceni Egyetemi Kiadó, Debrecen, pp. 191–198. https://giskonferencia.unideb.hu/arch/GIS_Konf_kotet_2023.pdf
- PEDREGOSA, F., VAROQUAUX, G., GRAMFORT, A. ET AL. (2011): Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12(Oct), pp. 2825–2830. <https://doi.org/10.5555/1953048.2078195>
- PLANET (2024): Planet Team, Planet Application Program Interface: In Space for Life on Earth. San Francisco, CA. <https://api.planet.com>

QGIS (2024): QGIS Geographic Information System v3.10.8., Open Source Geospatial Foundation Project. <http://qgis.org>

ROZNIK, M., BOYD, M., PORTH, L. (2022): Improving crop yield estimation by applying higher resolution satellite NDVI imagery and high-resolution cropland masks. *Remote Sensing Applications: Society and Environment*, 25, 100693. <https://doi.org/10.1016/j.rsase.2022.100693>

SAYAGO, S., BOCCO, M. (2018): Crop yield estimation using satellite images: comparison of linear and non-linear models. *AgriScientia*, 35(1), pp. 1–9. <https://doi.org/10.31047/1668.298x.v1.n35.20447>

SHANG, J., LIU, J., PONCOS, V., GENG, X., QIAN, B., CHEN, Q., DONG, T., MACDONALD, D., MARTIN, T., KOVACS, J., WALTERS, D. (2020): Detection of Crop Seeding and Harvest through Analysis of Time-Series Sentinel-1 Interferometric SAR Data. *Remote Sensing*, 12(10), 1551. <https://doi.org/10.3390/rs12101551>

SIPŐCZ J. (2021): Interferometrikus koherencia mezőgazdasági alkalmazhatóságának vizsgálata, TDK dolgozat, ELTE TTK.

SNAP (2024): ESA Sentinel Application Platform v10.0.0. <http://step.esa.int>

SON, N. T., CHEN, C. F., CHEN, C. R., MINH, V. Q. (2017): Assessment of Sentinel-1A data for rice crop classification using random forests and support vector machines. *Geocarto International*, 33(6), 587–601. <https://doi.org/10.1080/10106049.2017.1289555>

WEISS, M., JACOB, F., DUVEILLER, G. (2020): Remote sensing for agricultural applications: A meta-review. *Remote Sensing of Environment*, 236, 111402. <https://doi.org/10.1016/j.rse.2019.111402>

XU, L., ZHANG, H., WANG, C., ZHANG, B., LIU, M. (2019): Crop Classification Based on Temporal Information Using Sentinel-1 SAR Time-Series Data. *Remote Sensing*, 11(1), 53. <https://doi.org/10.3390/rs1101005>

SAJÁT EREDMÉNYEK

1. A kutatásomban megoldott kérdések:

- A Non-Local Means eljárással tovább mérsékelhető a Sentinel-1 felvételekben megtalálható zaj.
- Az aratás időpontja nagy pontossággal meghatározható 4 növényre Sentinel-1 szórási együtttható idősorokból konvolúciós neurális hálózattal.
- Vetésszerkezeti térkép 8 kategóriára nagy pontossággal előállítható Sentinel-1 idősorokból konvolúciós neurális hálózattal.
- A regressziós algoritmus évek között mérsékeltén hordozható, sokszor megbízhatatlan eredményt ad.
- A két évvel tanított regressziós algoritmus mindkét éven nagy pontossággal működik.

2. Az általam alkalmazott módszerek, eljárások:

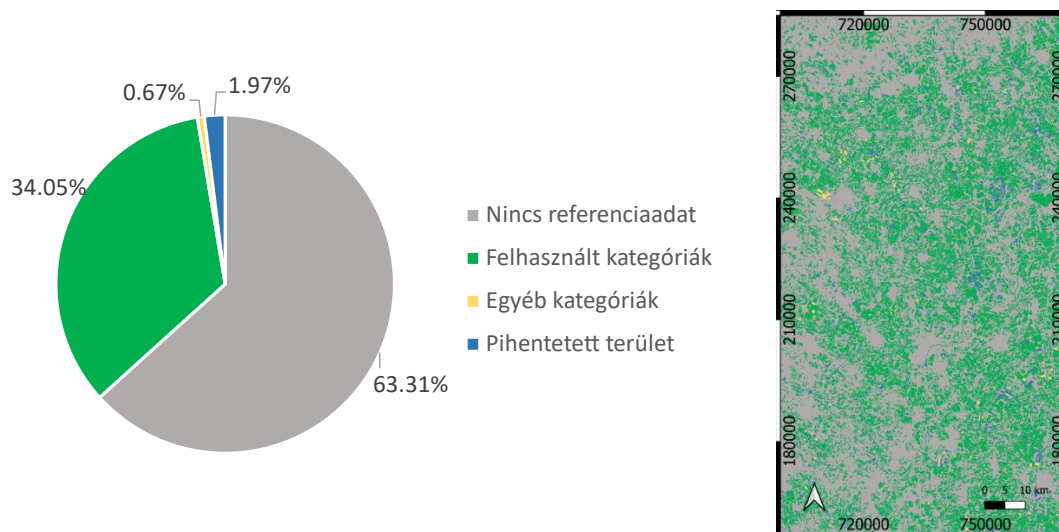
- A Földmegfigyelési Információs Rendszerben előálló Sentinel-1 szórási együtttható termékeken utófeldolgozást végeztem el Python környezetben.
- A szórási együtttható termékeken Non-Local Means szűrési eljárást alkalmaztam Python környezetben.
- Letöltöttem és vizuálisan kiértékeltem a Planet optikai űrfelvételeket.
- Elvégeztem a referenciaadatok adatbázis szintű feldolgozását Quantum GIS szoftverben és Python környezetben.
- Megírtam a vetésszerkezeti osztályozást végző konvolúciós neurális hálózat algoritmust Python környezetben.
- Megírtam az aratási időpont meghatározását végző konvolúciós neurális hálózat algoritmust Python környezetben.
- Pontossági statisztikák elemzésével és vizuális kiértékeléssel értelmeztem az eredményeket és fejlesztettem az algoritmust.

3. Saját kutatási eredményeim:

- Vetésszerkezeti térképek 8 kategóriára 93,45-94,79% pontossággal.
- Aratási időpont térképek 4 főnövényre 7,89-8,64 nap átlagos abszolút hibával.
- Két évnyi adattal előtanított, aratási időpont meghatározására alkalmas algoritmus 7,71-7,72 napos átlagos abszolút hibával.

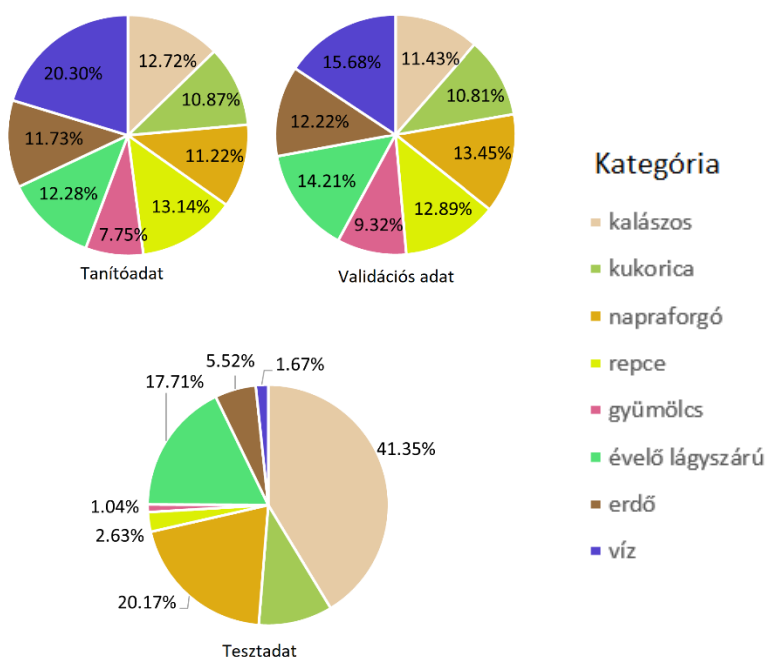
FÜGGELÉK

A)



A vetésszerkezeti térképezéshez felhasznált kategóriák, az egyéb kategóriák és a pihentetett területek aránya a mintaterületre vonatkozó teljes adatbázisban 2023-ban.

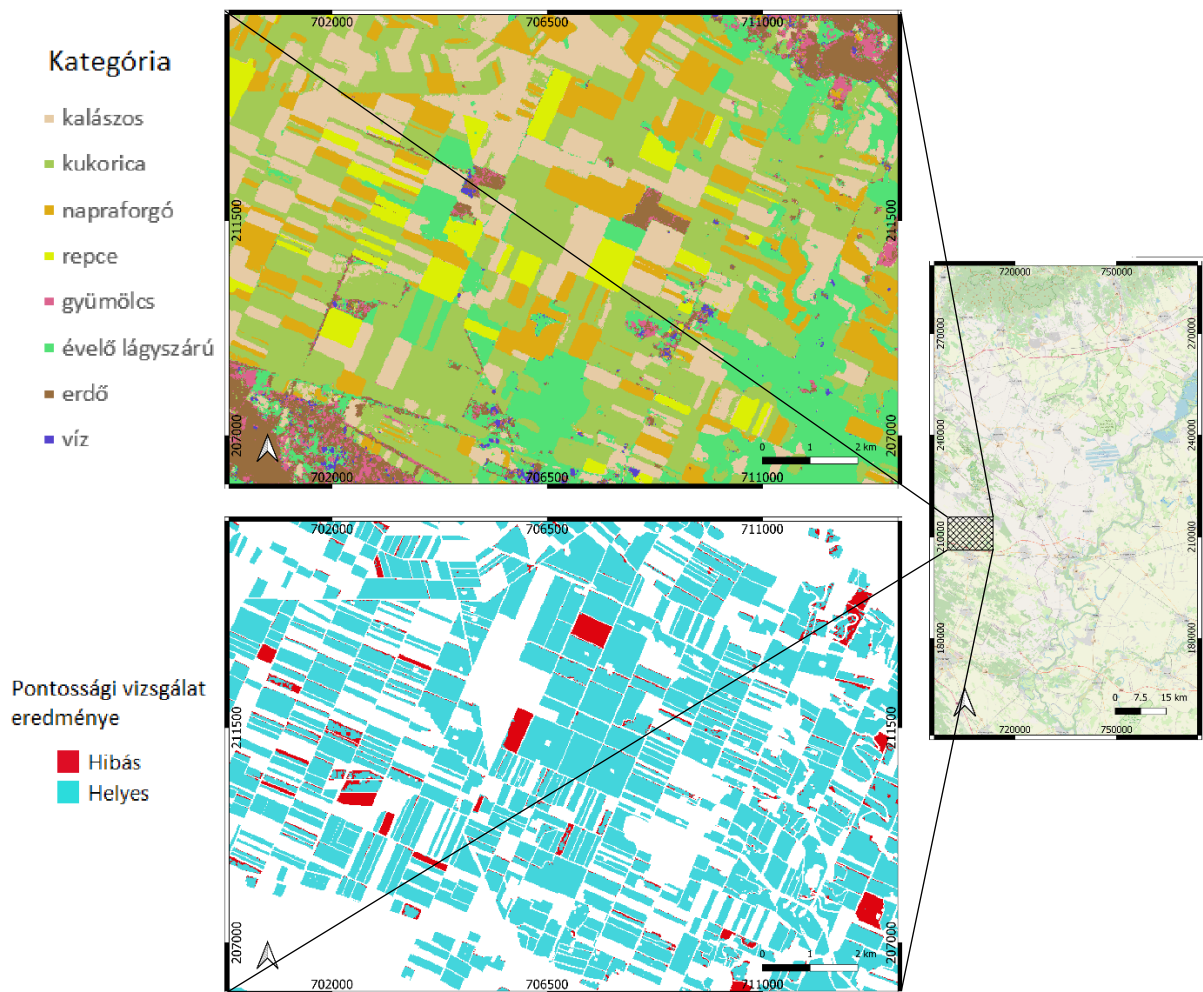
B)



Adathalmaz	Területarány [%]	Pixelszám [db]	Terület [ha]
Tanítóadat	0,51%	116931	4677,24
Validációs adat	0,56%	127851	5114,04
Tesztadat	32,97%	7508406	300336,2

Statisztika a 2023-as vetésszerkezeti térképezéshez felhasznált referenciaadatokról. Az egyes kategóriák pixelszám szerinti aránya a referenciaadatokban (fent) és a referenciaadatok által lefedett terület aránya a mintaterülethez viszonyítva (lent).

C)



A 2021-es, 12 napos idősor felhasználásával előállított vetésszerkezeti térkép (fent) és a tesztadathoz viszonyított tévesztési térképe (lent).

D)



A 2023-as, 12 napos idősor felhasználásával előállított vetésszerkezeti térkép (fent) és a tesztadathoz viszonyított tévesztési térképe (lent).