

Mesterséges intelligencia alapú osztályozó eljárások alkalmazása nagytájszintű területhasználat változás térképezésére

GUDMANN ANDRÁS

Bevezetés

Napjainkra a mesterséges intelligencia (MI) az egyik leggyorsabban fejlődő tudomány, amely kezd a mindennapjaink alapvető részévé válni. Az MI alapú osztályozó eljárások képesek bonyolult feladatokat nagy pontossággal megoldani, továbbá olyan kapcsolatokat képesek észlelni különböző változók között, amelyek az ember számára nem értelmezhetőek. Ezen algoritmusok felhasználása egyre nélkülözhetlenebbé válik a növekvő mennyiségű, típusú, tulajdonságú és egyre gyorsabban előállított adatok miatt. A távérzékelésben a mesterséges intelligencia alapú osztályozó eljárásoknak (gépi tanulás, mély tanulás algoritmusai) köszönhetően a legtöbb alapvető osztályozási feladat (mint például a felszínborítás térképezése) megoldottnak tekinthető. Azonban több komplex osztályozási feladatot még nem sikerült megoldani. Az egyik ilyen probléma a területhasználati kategóriák osztályozása. A területhasználat egy adott terület társadalmi és gazdasági hasznosítását mutatja meg. Mivel sokféle társadalom és gazdaság van, illetve az egyes társadalmak és gazdaságok sokféleképpen képesek egy adott területet hasznosítani, így ezeket a kategóriákat nehezebb meghatározni és lehatárolni is.

A műholdas távérzékelésben különböző problémákat kell együttesen figyelembe vennünk nagy területre kiterjedő területhasználat-osztályozás során. Egyrészt a különböző területhasználati típusok és a földmegfigyelő műholdak felvételeinek spektrális tulajdonságai között a kapcsolat nem egyértelműen meghatározható. Ebből adódóan a hagyományos osztályozási módszerek az ilyen kategóriákat csak kis vagy közepes pontossággal képesek visszaadni. Továbbá ezen eljárások alkalmazása nagy fokú felhasználói beavatkozást igényel, és nagy mennyiségű adat esetében pontosságuk nagy mértékben csökkenhet. A nagy adatmennyiség nemcsak a hagyományos, hanem minden osztályozó eljárás esetében is kihívásokat generál. Az egyes eljárások paraméterbecslési, modellépítési és osztályozási ideje drasztikusan növekedhet, míg az eredményeink pontossága jelentős mértékben csökkenhet. A megfelelő osztályozó algoritmus kiválasztásánál tehát figyelembe kell vennünk azok nagy adathalmazon mért hatékonyságát is.

Ezen kutatás célja Magyarország egész területére területhasználati térképek előállítására mesterséges intelligencia alapú osztályozó eljárások segítségével az 1986-tól 2017-ig terjedő időszakra. Továbbá az elkészült térkép felhasználásával a magyarországi területhasználat változásának vizsgálata: a változás térbeli mértékének meghatározása, a változások lokalizálása nagytáj szinten, illetve az egyes osztályok közötti változásokból a tendenciák megállapítása.

Mintaterület

A kutatás Magyarország teljes, 93 023 km² területére vonatkozik, a kutatás alapegységeit pedig Magyarország hat nagytája jelentette (1. ábra).¹ Magyarország lakossága 9 772 756 fő, akik 346 városban és 2809 faluban élnek (2019. 01. 01. adatok alapján).² Az ország területének több mint 80%-a 200 méter tengerszint alatti alföld, ahol jelentős mezőgazdasági termelés folyik.³ A hivatalos adatok alapján a vizsgált időszakban (1986–2017) Magyarország területének több mint 25%-a alakult át. Az átalakulás főleg a mezőgazdasági területek rovására történt, amelyek aránya Magyarország teljes területéhez képest 70,12%-ról 57,07%-ra csökkent. Ezzel ellentétesen a művelés alól kivett területek aránya 11,33%-ról 21,33%-ra emelkedett. A mezőgazdasági területek csökkenése és a művelés alól kivett területek növekedése nem egy okra vezethetőek vissza: a társadalmi változások mellett a politikai okok is közrejátszanak (magyar és EU-s agrárpolitika miatti terület kivonások). Mindazonáltal a legfőbb oka az átalakulásoknak főleg a városok és mesterséges felszínek növekedéséhez köthető (urbanizáció).⁴



1. ábra: Magyarország földrajzi nagytjai

1 Kocsis 2018.

2 KSH 2019.

3 GÁBRIS ET AL. 2018.

4 MAGYARORSZÁG FÖLDTERÜLETE

Felhasznált adatok

Referencia adat (CORINE adatbázisok)

A kutatáshoz és az osztályozások alapjául szolgáló referencia adatként a CORINE (Coordination of Information on the Environment) Land Cover (CLC) adatbázisokat használtam. A CLC program 1985-ben indult azzal a céllal, hogy megbízható, összevethető és részletes adatokat állítson elő az Európai Közösség és a programban résztvevő más országok felszínborításáról.⁵ Az adatbázisok legkisebb térképezési egysége 25 hektár, vonalas objektumok esetén a minimális szélesség 100 méter.⁶ Az adatbázis háromszintes nomenklatúrával rendelkezik, amely első szintjén még alapvető felszínborítási kategóriákat találunk, azonban a legalsó, harmadik szinten már főleg területhasználati kategóriák vannak. A részletes nomenklatúrájának, a nagy térbeli lefedésnek és a tematikus pontosságának köszönhetően ezen adatbázisok több különböző kutatás alapjául szolgáltak, mint például a területhasználati változás vizsgálatok,⁷ klímamodellelés és városi hőszigetek vizsgálata,⁸ aszály- és belvíztérképezés.⁹ A CLC adatbázisok 1990, 2000, 2006, 2012 és 2018-as időpontokra érhetőek el. Az osztályozás referenciájaként mindegyik adatbázis felhasználásra került a nomenklatúra harmadik szintjén.

Műholdképek

Az osztályozáshoz olyan műholdfelvételeket választottam melyek azonos térbeli felbontással, azonos vagy közel hasonló szenzorral kerültek felvételezésre. Ezen felvételeknek legjobban az 1970-es évektől működő Landsat program felvételei feleltek meg.¹⁰ A Landsat programot úgy alakították ki, hogy a benne résztvevő műholdak képei kompatibilisek legyenek egymással, így hosszú idősoros elemzésekre is alkalmasak legyenek.¹¹ A Landsat műholdképek a Föld felszínének 185×180 km-es részét fedik le, közepes térbeli felbontással (30–60 méter) a látható fény spektrális tartományán kívül infravörös és termális sávokkal is rendelkeznek. A vizsgált 30 éves időszakból 4 évre vonatkozó (1986, 2003, 2015, 2017) műholdfelvételeket dolgoztam fel. Magyarország teljes lefedéséhez tíz darab műholdkép szükséges, és arra törekedtem, hogy a felvételek időbeli eltérése minimális legyen. A felszíni reflektanciaértékeket tartalmazó műholdképeket az Earth Resource Observation and Science (EROS) Center Processing Architecture (ESPA)¹² rendszerén keresztül rendeltem

5 HEYMANN ET AL. 1994.

6 MARI – MATTÁNYI 2002.

7 SZILASSI 2017; LISKA ET AL. 2017.

8 MOLNÁR 2017.

9 TOBAK ET AL. 2019; BEZDAN ET AL. 2019.

10 EMERY – CAMPS 2017.

11 MUCSI 2013.

12 <https://espa.cr.usgs.gov/> (Letöltés: 2021. 06. 13)

meg és töltöttem le. A műholdképek a WRS katalógus 187, 188, 189-es sorának 026, 027, 028-as oszlopának csempéi, melyek vetületi rendszere Universal Transverse Mercator (UTM WGS84 N34).

Származtatott adatok

A Landsat műholdképek sávjai sok esetben a komplex elemzések, osztályozások során önmagukban már nem nyújtanak elegendő információt. Ennek kiküszöbölésére lehetőség van a vizsgálatunkba további adatforrásokat bevonni, vagy a meglévő adatainkból újabbakat kiszámítani. Idősoros elemzéseknel új információt bevonni igen problematikus, hiszen kevés olyan adatforrás van, amely mindegyik vizsgált időpontban rendelkezésünkre állna és összekapcsolható a meglévő adatainkkal. Emiatt az ilyen elemzések során a meglévő adatainkból kell származtatnunk új információval rendelkező adatot. Ebben a kutatásban az egyes képelemek környezetéről, a képek textúrájáról információt szolgáltató adatokat vontam be az elemzésekbe. A textúra adatok az egyes képelemek környezetének változékonyságát mutatják meg, azaz milyen értékű pixelek vannak egy megadott távolságon belül. A képelem környezetének textúrájának meghatározása többféle módon történhet, mint például a variancia (variance), ferdeség (skewness), kurtosis vagy átlagos euklideszi távolság számításával.¹³ A textúra értéket a pixel környezetének egy felhasználó által meghatározott övezetéről (számítási ablak) számítjuk ki. Ahhoz, hogy megfelelő mennyiségű plusz információt szerezzünk a textúra értékének számításával, célszerű a számítási ablakot minél nagyobb értékre állítani, illetve több különböző nagyságú ablak értékeit együttesen felhasználni a vizsgálatban. Továbbá a különböző ablakméret alkalmazása lehetővé teszi a kiugró pixelértékek torzításának csökkentését.

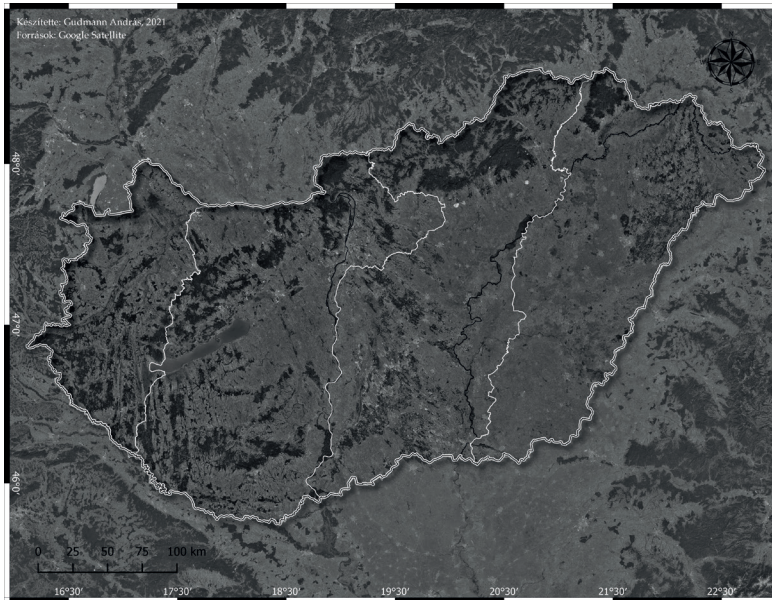
Módszerek

Mozaikok és származtatott adatok előállítás

A területhasználati térképek előállításához az első lépés a teljes országot lefedő műholdképmozaikok előállítása jelentette. A mozaikok előállításához tíz darab, Landsat műholdképet használtam fel. A tíz felvétel négy különböző műholdpásztában készült, ami miatt a felvételek csoportjai között időbeli és az átfedő területeken értékbeli eltérések adódtak. Emiatt az egyes pásztákban lévő képeket mozaikoltam elsőnek össze úgy, hogy a képek között egy geometriai alapon automatikusan generált vágóvonalat használtam és a vágóvonal környezetének értékeit átlagoltam (simítás). Ezután a pászták összevont képeit illesztettem össze, megkapva a végleges mozaikképet. A pászták képeinek összeillesztésénél egy általam készített vágóélt használtam fel, amely a természetes határvonalakat (folyók, csatornák) követte. Utolsó lépésként a mozaikok pixelei közül kiszűrtem a hamis értékeket (a lehetségesnél na-

¹³ IRONS – PETERSEN 1981.

gyobb vagy kisebb értékek, melyek mérési hibákból származnak) tartalmazókat egy általam készített modell segítségével. Az elkészült mozaikok hat sávot tartalmaznak és felszíni reflektanciaértékkel rendelkeznek. Az elkészült mozaikokból varianciaértékeket számoltam ki az egyes képelemekre különböző pixel nagyságú környezettel. A varianciaértékeket 7x7-es és 17x17-es ablakméretre számítottam ki. A spektrális sávokat és a textúraértékeket egy képbe egyesítettem, majd a létrehozott képet négy részre vágtam, a mozaikok előállításánál használt vágóélek mentén (2. ábra).



2. ábra: A feldarabolás során létrehozott négy részmintaterület

Tanító adathalmaz előállítása

Az osztályozáshoz szükséges tanítópontokat azokon a területeken választottam ki véletlenszerűen, ahol a felszínborítás nem változott a vizsgált időszakban. Ehhez összevettem egymással a CLC adatbázisokat és a nem változó területeket kiválogattam, majd felosztottam a mozaikoknál is használt metszőélek mentén ezeket a területeket. Az egyes kivágatokra tanítópontokat jelöltem ki véletlenszerűen elszórva, 20 000 pontot osztályonként (a tanító adathalmazok nagysága a teljes területhez képest ~1–2%) megadva, hogy az egyes pontok között minimálisan 34 méternek kell lennie, így elérve, hogy egy pixelt egy pont fedjen le. Végül ezekhez a tanítópontokhoz hozzáfűztem a műholdképek spektrális sávjainak értékeit és a textúra adatokat. Az adatok transzformálásához, az adattisztításhoz és osztályozáshoz python programnyelven írt szkriptjeimet használtam fel. A szkriptek megírásához az erre a

céla leggyakrabban használt függvénykönyvtárakat alkalmaztam: gdal,¹⁴ numpy,¹⁵ pandas,¹⁶ scikitlearn,¹⁷ tensorflow.¹⁸ Ezekkel a szkriptekkel gyorsan és rugalmasan tudtam átalakítani a raszteres adataimat osztályozásra kész adattömbökké, továbbá automatikusan tudtam elvégezni a modell felépítéséhez szükséges paraméterbecslést és végül az osztályozást.

Osztályozó algoritmusok

A kutatásomhoz két mesterséges intelligencia alapú algoritmust választottam ki, a véletlen erdő (random forest) osztályozó eljárást, illetve a mesterséges neurális hálót (artificial neural network). A véletlen erdő egy együttes osztályozó, ami azt jelenti, hogy egy egyszerű osztályozó eljárás sokaságát használja fel a döntéshozásra. A random forest algoritmus döntési fák halmazát alkalmazza, amelyeket véletlenszerűen kiválasztott adatokon épít fel, így valamilyen mértékben eltérnek egymástól. Ennek az eltérésnek köszönhetően, egy ugyanazon adatra, más és más választ adnak. Ezeket a válaszokat kombinálva, többségi szavazási sémát, – a döntési fák eredményei között legtöbbször szereplő eredményt – használva becsüli meg a véletlen erdő az eredményeket. Ezt az eljárást a gépi tanulás és az adatbányászat is sűrűn alkalmazza. Előnyei, hogy robusztus az adathalmazban lévő hibákra, könnyen kezelhető és nagy adathalmazon jó teljesítményt nyújt.¹⁹ Hátránya, hogy megfelelő működéséhez nagy adatmennyiségre van szükség: mind a tanításhoz, mind a teszteléshez. A másik általam választott osztályozó eljárás a mesterséges neurális háló. A mesterséges neurális háló az egyik legrégebbi eljárás (a kezdeti kutatások még az 1940-es években történtek),²⁰ amely az agyban lévő idegrendszer struktúráját és működését utánozó számítási mechanizmus.²¹ A modell neuronok és az azokat tartalmazó rétegek előre meghatározott sorrendjéből épül fel. A számítási kapacitása miatt ez az eljárás komplex problémák megoldására is alkalmas, osztályozási feladatoknál a többi eljáráshoz képest nagy pontosságot produkál. Hátrányai, hogy a modell felépítéséhez nagy szintű felhasználói háttértudás szükséges, a létrejövő modell az ember számára értelmezhetetlen ('black box') és a modellépítő folyamatokban lévő véletlenszerűségek miatt ugyanazon modell előállítása nem biztosított. Az eddigi összehasonlító kutatások rámutattak, hogy a neurális háló kis mértékben jobb pontosságot tud nyújtani kis adathalmazok esetén, mint a véletlen erdő.²²

Az osztályozó modellek felépítéséhez szükséges optimális paraméterek meghatározását (paraméterbecslés) automatizált módszerekkel hajtottam végre. A vé-

14 GDAL

15 HARRIS ET AL. 2020.

16 MCKINNEY 2010.

17 PEDREGOSA ET AL. 2011.

18 ABADI ET AL. 2015.

19 BREIMAN 2001.

20 MCCULLOCH – PITTS 1943.

21 ALTRICHTER ET AL. 2006.

22 LEEUWEN VAN ET AL. 2020.

letlen erdő esetében egy rácskeresés nevű paraméterbecslési módszert használtam, kereszt-validáció alkalmazásával. A módszer a megadott paraméterek előre meghatározott lehetséges értékeinek minden kombinációját leteszteli és az általunk meghatározott mérőszám alapján visszaadja a legmegfelelőbb paraméter kombinációt. A paraméterbecslés során a döntési fák számát (erdő nagysága) és a létrehozásukhoz véletlenszerűen kiválasztott attribútumok számát is vizsgáltam. A neurális háló esetében a keras-tuner²³ függvénykönyvtár hyperband algoritmusát használtam. Ez egy, a sportvilágból átvett paraméterbecslési módszer, amelynek lényege, hogy több fordulóban teszteli le a különböző paraméterekkel felépített modelleket és minden körben a legjobb x számú esetet viszi át a következő körbe úgy, mint a sportban a többfordulós események. A fordulókön belül minden paraméter kombinációt x -szer épít fel és az átlagos pontosságuk alapján rangsorol. Emiatt a paraméterbecslés megbízhatóbb lesz, mivel a jobb eredményt adó modellek többszörösen lesznek tesztelve (körön belül és körönként is), azonban emiatt lassabb is a folyamat. A paraméterbecslés során a rejtett rétegek számát, a neuronok számát, a tanulási rátát vizsgáltam. A paraméterbecslésnek köszönhetően automatikusan kaphattam meg a legjobb modellépítő paramétereket, és így a legjobb modell segítségével tudtam az osztályozásokat elvégezni.

Eredmények

Modellezés

A paraméterbecslést a véletlen erdő és a neurális háló esetén is a vizsgált terület négy részén külön-külön lefuttattam. A becslések eredményeképp a véletlen erdő esetén az erdő nagysága 501 fa, míg a véletlen változók száma 6 és 7 között adta a legjobb modellt. A neurális háló esetében a legjobb eredményt az 5 réteg, 2048 neuron és 0,0001 tanulási ráta paraméterek esetén kaptam. A paraméterbecsléseket az 1986-os adatokon futtattam le. Az eredmények alapján a véletlen erdő mind a négy részmintaterületen hatékonyabbnak bizonyult, nagyobb pontosság (5–10%) mellett rövidebb futási idővel végezte el az osztályozási feladatot. Ezek alapján a többi osztályozást a véletlen erdővel hajtottam végre.

Osztályozás eredményei

A megfelelő eljárás és a paraméterek meghatározása után minden adathalmazra lefuttattam az osztályozásokat. Az eredményeket python szkript segítségével visszaalakítottam térképekké. Az egyes évekre vonatkozó osztályozott térképeket egyesítettem, így megkapva a teljes országot lefedő területhasználati térképeket. A pontosságbecsléshez a tanítópontok előállításánál is használt, véletlenszerűsített pontelhelyezéssel pontthalmazt generáltam, osztályonként 20 000 ponttal. A gene-

23 O'MALLEY ET AL. 2019.

rált ponthalmaz pontjainak helyén megvizsgáltam az eredeti adatbázis (CLC) és az osztályozás eredményének értékét, az összevetés alapján megállapítottam az egyes térképek összpontosságát. Az eredmények alapján minden térkép legalább 75%-os összpontossággal rendelkezik, a legalacsonyabb pontossága az 1986-os térképnek van (78,69%), míg a legmagasabb pontossággal a 2015-ös rendelkezik (90,67%) (1. táblázat).

Év	1986	2003	2015	2017
Pontosság	78,69%	80,54%	90,67%	90,65%

1. táblázat: Az osztályozások pontossága a különböző időpontokban

Ezeket az eredményeket több szempont figyelembevételével kell kiértékelni. Egyrészt a CLC adatbázisok készítésének módszertana az évek során finomult, ezáltal a tematikus pontossága az egymás utáni adatbázisoknak növekedett. A CLC90-es adatbázis esetén még a kívánt (>=85%) tematikus pontosságot nem biztos, hogy elérték a készítők, addig a CLC legújabb változatainál ez már biztosan garantált. A referenciaadat bizonytalansága nagyban befolyásolhatta az osztályozás eredményét, ami megmagyarázza, hogy az újabb osztályozott képek miért rendelkeznek egymáshoz hasonlóan magas pontossággal. A referencia-adatbázis bizonytalanságán kívül figyelembe kell venni a kisebb (pixel szintű, 900 m²) térképezési egységet, ami a CLC adatbázis 25 hektáros egységéhez képest jelentősen kisebb. Ez az eltérés magában is jelentős pontosságbeli eltéréseket okozhat. Továbbá azt az alapvető tényt sem szabad figyelmen kívül hagyni, hogy az osztályozott térképek időpontjai nem egyeznek meg a CLC készítésének időpontjaival, amelyek így eleve veszíthetnek a pontosságukból. Ezeknek a tényezőknek a figyelembevételével az osztályozás eredményei kiemelkedően jónak tekinthetők.

Területhasználat-változás vizsgálat

A területhasználat-változás vizsgálatok a különböző időpontokban készült terület-használati térképeket hasonlítottam össze egymással. A térképek összevetése után az egyes osztályokról terület mennyiség alapján tudtam követni a változásokat, illetve a pixel szintű összehasonlítás alapján meg tudtam határozni, hogy hol és milyen változás történt. A változásokat hat csoportba kategorizáltam Feranec munkája alapján:²⁴ urbanizáció (mesterséges felület növekedés), mezőgazdaság intenzív növekedése (a mezőgazdasági területet jobban, hatékonyabb kihasználó művelési ágra való átállás), mezőgazdaság extenzív növekedése (új vagy rég nem használt területek művelés alá vétele), erdősítés (fateleptetés), erdőritkulás (fakivágások), vízfelszín-létrehozás. Az urbanizáció a városok, beépített területek, mesterséges felszínek növekedését takarja. Ez főleg egyirányú területhasználat változás, mivel nagyon kis mértékű a mester-

²⁴ FERANEC ET AL. 2010.

séges felszínnek újra művelés alá vétele, inkább parkosítással történhet ezen felületek csökkentése. A mezőgazdaság intenzívvé válása a már meglévő mezőgazdasági területen történő technológia váltásra utal, amellyel hatékonyabb termelés érhető el. Ezzel ellentétben a mezőgazdaság extenzív növekedése új területek bevonására utal a termelés növekedése érdekében. Az erdősödés és erdőritkulás az erdők területének növekedését és csökkenését jelenti. Ezen kategóriáknál figyelembe kell venni a tervezett erdőrezszt kivágásokat és azok pótlására az erdőtelepítéseket. Ebből adódóan manapság az olyan erdőkivágás, amely után nincs pótolva a faállomány már szinte nincs. A vízfelszín létesítések halastavak, rekreációs és sport komplexumok, víztározók létrehozásához köthető változások. Ezen kategorizálásból kimaradt, de nem szabad eltekinteni a művelés alóli kivétel, amellyel a mezőgazdasági területek csökkenhetnek. Az eredmények alapján az 1986 és 2003 közötti időszakban, az egész ország területének 36,42%-a változott meg és a változásokon belül a mezőgazdaság intenzívebbé válása dominál minden nagytájon (2. táblázat).

1986-2003	Urbanizáció	Mezőgazdaság intenzívebbé válása	Mezőgazdaság extenzívebbé válása	Erdősödés	Erdőritkulás	Vízfelszín telepítés
Dunántúli-dombság	4,54%	11,78%	8,22%	7,07%	4,61%	0,12%
Alpokalja	4,66%	10,00%	7,85%	8,82%	4,76%	0,18%
Északi-középhegység	5,64%	8,52%	3,69%	9,34%	5,09%	0,20%
Dunántúli-középhegység	5,80%	10,89%	6,46%	9,22%	5,99%	0,06%
Kisalföld	7,03%	10,87%	6,78%	6,37%	2,95%	0,72%
Alföld	6,38%	12,51%	7,41%	7,44%	2,86%	0,63%

2. táblázat: Az 1986 és 2003 közötti területhasználat változások nagytájanként

Ennek mértéke 8,52% és 12,51% között mozgott az egyes tájakon, a táj nagyságához mérten, a legjelentősebb az Alföldön és a Dunántúli-dombságon. Ezen kívül még az erdősödés és a mezőgazdaság extenzívebbé válása van jelen majdnem minden nagytájon átlag fölötti mértékben. A mezőgazdaság extenzívebbé válása a Dunántúli-dombságon a legerőteljesebb (8,22%), míg az Északi-középhegységben a legkisebb (3,69%). Az erdősödés terén kis eltérés van a nagytájak között az átlaghoz képest (7,79%), a legnagyobb változásokat a középhegységeinkben tapasztalhatjuk (9,22–9,34%). Az urbanizáció egy fokkal kisebb ütemben változtatta meg a területhasználatot, mértéke 4,54 és 7,03% között mozog. Az 1986 és 2003 közötti trendek megmaradnak a 2003–2015 közötti időszakban is, azaz a mezőgazdaság intenzívebbé válása dominál (3. táblázat). Az erdősödés minden nagytájon átlag fölötti mértékű, főleg a Dunántúli-középhegységben (9,96%). Az urbanizáció és a mesterséges felszínnek növekedésének folyamata főleg a mezőgazdasági nagytájakon (Alföld,

Kisalföld) és a Dunántúli-középhegység területén jelentős mértékű változást okoztak (8,6–11,55%). A mezőgazdaság extenzívebbé válása és az erdőritkulás kisebb mértékben csökkentek (~0,5%). Az egész országot figyelembe véve az urbanizáció hatásának növekedése volt a legnagyobb az előző időszakhoz képest (+2,52%), majd a mezőgazdaság intenzívebbé válása (+1,12%) végezetül az erdőszedés (+0,57%). A 2015–2017-es időszakban ugyanúgy, mint a többi időszakban a mezőgazdasági intenzívebbé válása jelentős változást okozott minden nagytáján (4. táblázat).

2003-2015	Urbanizáció	Mezőgazdaság intenzívebbé válása	Mezőgazdaság extenzívebbé válása	Erdősödés	Erdőritkulás	Víz-felszín telepités
Dunántúli-dombság	5,35%	11,23%	7,55%	6,94%	5,13%	0,27%
Alpokalja	4,91%	10,38%	7,46%	7,52%	4,87%	0,48%
Északi-középhegység	5,89%	11,86%	4,79%	8,96%	4,24%	0,25%
Dunántúli-középhegység	7,55%	13,07%	5,86%	9,96%	4,73%	0,12%
Kisalföld	11,55%	10,54%	6,87%	8,46%	1,99%	1,01%
Alföld	10,01%	13,67%	7,07%	8,48%	2,24%	1,02%

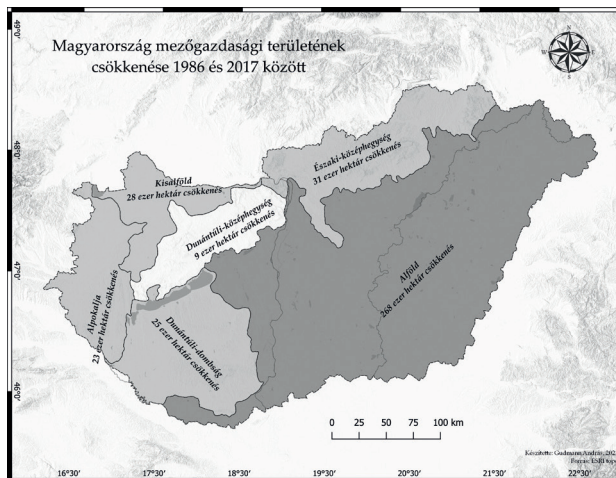
3. táblázat: A 2003 és 2015 közötti területhasználat változások nagytájanként

2015-2017	Urbanizáció	Mezőgazdaság intenzívebbé válása	Mezőgazdaság extenzívebbé válása	Erdősödés	Erdőritkulás	Víz-felszín telepités
Dunántúli-dombság	5,22%	13,33%	7,62%	6,16%	3,55%	0,37%
Alpokalja	5,63%	10,44%	6,78%	7,24%	4,82%	0,82%
Északi-középhegység	8,60%	7,26%	4,67%	8,44%	3,85%	0,45%
Dunántúli-középhegység	5,38%	10,85%	6,32%	7,05%	4,25%	0,12%
Kisalföld	11,21%	11,25%	7,00%	5,48%	3,81%	1,37%
Alföld	7,41%	13,58%	7,58%	6,48%	3,01%	1,19%

4. táblázat: A 2015 és 2017 közötti területhasználat változások nagytájanként

A legerősebb hatás az Alföldön és a Dunántúli-dombság területén tapasztalható. A mezőgazdaság intenzívebbé válása mellett az extenzívebbé válás is jelentős mértékben (6,32–7,62%) okozott változást a nagytájakon, ez alól csupán az Északi-középhegység a kivétel. Az erdősítés közepes mértékben csökkent, országos szinten -1,65%-kal kisebb mértékű változás sorolható ide. A legnagyobb csökkenés az erdősítés ütemében az Alföldön és a Kisalföldön történt. Az urbanizáció folyamata az Alföldön és a Kisalföldön továbbra is jelentős mértékű maradt (7,41–11,21%) és az Északi-középhegységben is jelentős mértékű lett. Az erdőritkulás az előző időszakokhoz hasonlóan 3 és 5% között mozognak az egyes nagytájakon. A vízfelszínek telepítése mindhárom időszakban 1% körüli mértékben mozgott, ez a Központi Statisztikai Hivatal adatai alapján főleg a halastó létesítésekhez köthető. A teljes időszakot megfigyelve az urbanizáció az Alföld és a Kisalföld területén volt a legjelentősebb, a 2003 és 2015 időszakban elérve csúcspontját. A mezőgazdaság intenzívebbé és extenzívebbé válása minden nagytájon jelentős volt, kiemelkedően az Alföldön és a Dunántúli-dombság esetében. Az erdősítés a középhegységeinkben volt a legjelentősebb, azonban ezeken a területeken az erdőritkulás által okozott területhasználat változás folyamatosan csökkent.

Összességében Magyarország területhasználatának változásának utóbbi 30 évében a mezőgazdasági szerkezetváltások, az urbanizáció és az erdősítés voltak a hangsúlyosak. Ezeknek a változásoknak köszönhetően az elkészített térképek alapján a mesterséges felszín nagysága 217 ezer hektárral, míg az erdők nagysága 94 ezer hektárral nőtt. Ugyanakkor a szántóföldek nagysága 307 ezer hektárral, míg a mezőgazdasági területek nagysága 384 ezer hektárral csökkent (3. ábra). Ez annak fényében igazán negatív, hogy Magyarország legnagyobb természeti kincse az igen jó minőségű termőföldje, amellyel a Világ egyik legtermékenyebb és legtöbb termőterülettel rendelkező országa (az ország területéhez mérten).



3. ábra: Magyarország mezőgazdasági területének csökkenése

Összefoglalás

A kutatásom célja az utóbbi 30 év területhasználat változásának detektálása Magyarország területére. A térképezéshez teljes Magyarországot lefedő műholdkép mozaikokat állítottam elő az általam kiválasztott négy időpontra vonatkozóan. A területhasználati térképek előállításához mesterséges intelligencia alapú osztályozó eljárások felhasználásával a mozaikokból területhasználati térképeket készítettem négy időpontra vonatkozóan: 1986, 2003, 2015, 2017. A kutatásom során kiválasztottam a legmegfelelőbb módszert az osztályozások elvégzéséhez, amely a véletlen erdő eljárás lett. Ezzel a módszerrel sikeresen tudtam előállítani a területhasználati térképeket, amelyek igen magas pontossági értékekkel rendelkeznek (78,69–90,67%), így megfelelő alapot biztosítottak a változások vizsgálatához. A változás vizsgálata során nagytáj szinten megállapítottam az egyes osztályok területét és pixel szinten a változások irányát. A változásokat hat kategóriába sorolva általános képet alkottam a területhasználati változások trendjéről. A változás vizsgálata alapján Magyarország területén a vizsgált időszakban (1986–2017) a mezőgazdasági szerkezetváltások, az urbanizáció és az erdőszedés miatti területhasználat változások voltak a meghatározóak. Ezen változások eredményeképp a mezőgazdasági területek jelentősen csökkentek és a mesterséges felszínek és erdők területe nőtt.

Irodalom

ABADI ET AL. 2015 = Abadi M. – Agarwal A. – Barham P. – Brevdo E. – Chen Z. – Citro C. – Corrado G. S. – Davis A. – Dean J. – Devin M. – Ghemawat S. – Goodfellow I. – Harp A. – Irving G. – Isard M. – Jia Y. – Jozefowicz R. – Kaiser L. – Kudlur M. – Levenberg J. – Mané D. – Monga R. – Moore S. – Murray D. – Olah C. – Schuster M. – Shlens J. – Steiner B. – Sutskever I. – Talwar K. – Tucker P. – Vanhoucke V. – Vasudevan V. – Viégas F. – Vinyals O. – Warden P. – Wattenberg M. – Wicke M. – Yu Y. – Zheng X.: TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems, 2015.

ALTRICHTER ET AL. 2006 = Altrichter M. – Horváth G. – Pataki B. – Strausz Gy. – Takács G. – Valyon J.: *Neurális hálózatok*. Budapest 2006.

BEZDAN ET AL. 2019 = Bezdán A. – Vranešević, M. – Blagojević, B. – Pejić, B. – Bezdán, J. – Milić, D. – Tica, N. – Zekić, V.: Mezőgazdasági aszálykockázat-becslés a Vajdaságban. In: Ladányi Z. (szerk.) – Blanka V. (szerk.): *Aszály és belvíz monitoring és menedzsment, valamint a kapcsolódó kockázatok a Dél-Alföldön és a Vajdaságban*. Szeged: Szegedi Tudományegyetem, Természeti Földrajzi és Geoinformatikai Tanszék, 2019.

BREIMAN 2001 = Breiman L.: Random Forests, *Machine Learning*. 45 (2001) 5–32.

EMERY – CAMPS 2017 = Emery W. – Camps A.: The history of satellite remote sensing. In: *Introduction to Satellite Remote Sensing*. Elsevier 2017.

FERANEC ET AL. 2010 = Feranec J. – Jaffrain G. – Soukup T. – Hazeu G.: Determining changes and flows in European landscapes 1990–2000 using CORINE land cover data. *Applied Geography* 30/1 (2010) 19–35.

GÁBRIS ET AL. 2018 = Gábris Gy. – Pécsi M. – Schweitzer F. – Telbisz T.: Domborzat. In: Kocsis K. (főszerk.): *Magyarország nemzeti atlasza: természeti környezet*. Budapest: MTA CSFK Földrajztudományi Intézet, 2018.

HARRIS ET AL. 2020 = Harris C.R. – Millman K. J. – van der Walt S. J. – Gommers R. – Virtanen P. – Cournapeau D. – Wieser E. – Taylor J. – Berg S. – Smith N. J. – Kern R. – Picus M. – Hoyer S. – van Kerkwijk M. H. – Brett M. – Haldane A. – del Río Jaime F. – Wiebe M. – Peterson P. – Gérard-Marchant P. – Sheppard K. – Reddy T. – Weckesser W. – Abbasi H. – Gohlke C. – Oliphant T. E.: Array programming with NumPy, *Nature* 585/7825 (2020) 357–362.

HEYMANN ET AL. 1994 = Heymann Y. – Steenmans C. – Croisille G. – Bossard M. – Lenco M. – Wyatt B. – Weber J.-L. – O’Brian C. – Cornaert M.-H. – Sifakis, N.: *Corine Land Cover Technical Guide*. Office for Official Publications of the European Communities, 1994.

IRONS – PETERSEN 1981 = Irons J. – Petersen G.: Texture transforms of remote sensing data. *Remote Sensing of Environment* 11 (1981) 359–370.

KOCSIS 2018 = Kocsis K.: Magyarország dióhéjban. In: Kocsis K. (főszerk.): *Magyarország nemzeti atlasza: természeti környezet*. Budapest: MTA CSFK Földrajztudományi Intézet, 2018.

KSH 2019 = Központi Statisztikai Hivatal: *Magyarország közigazgatási helynévkönyve* Budapest: KSH, 2019.

LEEUVEN VAN ET AL. 2020 = Leeuwen van, B. – Tobak Z. – Kovács F.: Machine Learning Techniques for Land Use/Land Cover Classification of Medium Resolution Optical Satellite Imagery Focusing on Temporary Inundated Areas. *Journal of Environmental Geography* 13/1–2 (2020) 43–52.

LISKA ET AL. 2017 = Liska Cs. M. – Mucsi L. – Henits L.: Hosszútávú felszínborítás-változások vizsgálata Csongrád megyében idősoros adatok felhasználásával, Random Forest módszerrel. *Földrajzi közlemények* 141/1 (2017) 71–83.

MARI – MATTÁNYI 2002 = Mari L. – Mattányi Z.: Egységes európai felszínborítási adatbázis a CORINE Land Cover program. *Földrajzi Közlemények* 76/50 (2002) 31–38.

MCCULLOCH – PITTS 1943 = McCulloch W. – Pitts W.: A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics* 5/4 (1943) 115–133.

MCKINNEY 2010 = McKinney W.: Data Structures for Statistical Computing in Pyt-

hon. In: van der Walt, S. – Millman, J. (eds.): *Proceedings of the 9th Python in Science Conference*, Austin 2010. 56–61.

MOLNÁR ET AL. 2017 – Molnár G. = Gyöngyösi A. Z. – Gál T.: A városi hősziget vizsgálata meteorológiai modell segítségével Szegeden. *Légtér* 61 (2017) 130–135.

MUCSI 2013 = Mucsi L.: Műholdas távérzékelés: Elmélet és gyakorlat. 2013.

PEDREGOSA ET AL. 2011 = Pedregosa F. – Varoquaux G. – Gramfort A. – Michel V. – Thirion B. – Grisel O. – Blondel M. – Prettenhofer P. – Weiss R. – Dubourg V. – Vanderplas J. – Passos A. – Cournapeau D. – Brucher M. – Perrot M. – Duchesnay, E.: Scikit-learn: Machine Learning in Python, *Journal of Machine Learning Research* 12 (2010) 2825–2835.

SZILASSI 2017 = Szilassi P.: Magyarországi kistájak felszínborítási változékonysága és felszínborítási mozaikosságuk változása. *Tájökológiai lapok* 15/2 (2017) 131–138.

TOBAK ET AL. 2019 = Tobak Z. – Leeuwen van, B. – Kovács F. – Szatmári J.: Belvíz előntések nagy pontosságú térképezése és monitorozása. In: Ladányi Z. (szerk.) – Blanka V. (szerk.): *Aszály és belvíz monitoring és menedzsment, valamint a kapcsolódó kockázatok a Dél-Alföldön és a Vajdaságban*. Szeged: Szegedi Tudományegyetem, Természeti Földrajzi és Geoinformatikai Tanszék, 2019.

Internetes források

GDAL = GDAL/OGR Geospatial Data Abstraction Library: <https://gdal.org> (Letöltés: 2021. 06. 15.)

MAGYARORSZÁG FÖLDTERÜLETE = https://www.ksh.hu/stadat_files/mez/hu/mez0008.html (Letöltés: 2021. 06. 11.)

O'MALLEY ET AL. 2019 = O'Malley T. – Bursztein E. – Long J. – Chollet F. – Jin H. – Invernizzi L: Keras-tuner. In: <https://github.com/keras-team/keras-tuner> (Letöltés: 2021. 07. 05.)

Application of artificial intelligence-based classification algorithms to map large-scale land use change

ANDRÁS GUDMANN

The land use change properly shows the transformation of a certain area. This study aims to create and analyse land-use maps from Hungary in 4 different acquisition times (from 1986 to 2017). For the land-use map creation, Landsat mosaics with texture data were created. Two different classification algorithm was tested: artificial neural network and random forest. The test shows that the random forest procedure was more suitable for this task, because its speed and high accuracy. After the procedure selection random forest was applied for the classification task. For reference data CORINE Land Cover databases (CLC90, CLC00, CLC06, CLC12, CLC18) were applied. The accuracy assessment had shown the land-use maps high overall accuracy (78,69 – 90,67%). A land-use map were compared with each other, and the main change flows were determined. According the land-use change results, the main land-use processes were the intensification and extensification of agriculture, the urbanisation and afforestation. The intensification and extensification of agriculture were high in the whole country area during the investigated period. The urbanisation was significant in the main agricultural areas and the afforestation in mountains. Thanks of these processes the area of the artificial surfaces were increased with 217 thousand hectares and the size of the forests with 94 thousand hectare. However, the size of the non-irrigated arable lands and with that, the agricultural lands were decreased with significant (307 thousand and 384 thousand hectare).