

TDK-dolgozat

Jécsák Fruzsina

Kilácskó Dávid

Időszakos belvizek detektálása multitemporális műholdfelvételeken

EÖTVÖS LORÁND TUDOMÁNYEGYETEM

INFORMATIKAI KAR

PROGRAMOZÁSELMÉLET ÉS SZOFTVERTECHNOLÓGIAI TANSZÉK
INFORMÁCIÓS RENDSZEREK TANSZÉK



Szerzők:

Jécsák Fruzsina

programtervező informatikus BSc

3. évfolyam

Kilácskó Dávid

programtervező informatikus BSc

3. évfolyam

Témavezetők:

dr. Vincellér Zoltán

mesteroktató

Cserép Máté András

egyetemi tanársegéd

Budapest, 2022

Tartalomjegyzék

1. Bevezetés	3
1.1. A felszíni vizek változásának jelentősége	3
1.2. Célkitűzés	4
2. Technológiai és szakirodalmi áttekintés	5
2.1. Multispektrális műholdfelvételek	5
2.1.1. Sentinel-2 műhold	6
2.1.2. Planet műhold	6
2.2. Vízüindexek	7
2.3. Küszöbölő algoritmusok	10
2.3.1. Otsu	11
2.3.2. Minimum	12
2.3.3. Yen	13
2.4. Felhő és felhőárnyék detektálás	15
2.5. Felszíni vizek változáselemzése	16
3. Módszertan	17
3.1. A kiválasztott adatok	17
3.2. A kiválasztott eszközök	19
3.2.1. Preprocesszálás	19
3.2.2. Vízmaszk készítés	21
3.2.3. Változásdetektálás	24
4. Implementáció	26
4.1. Elkészült alkalmazás	26
4.2. Planet felvételek indexelése	27
4.3. Néhány használt algoritmus	28
5. Eredmények	30

5.1. Vízmaszk	30
5.2. Felszíni vizek változásdetektálása	33
6. Összefoglaló	35
Köszönetnyilvánítás	36
Irodalomjegyzék	36
Ábrajegyzék	39
Táblázatjegyzék	41
Forráskódjegyzék	42

1. fejezet

Bevezetés

1.1. A felszíni vizek változásának jelentősége

Felszíni víznek nevezzük a a Föld felszínén lévő vizeket, mint például a tavakat, a folyókat, a vizes élőhelyeket és a víztározókat. Általánosságban a sós vizek, mint az óceánok is ide tartoznának, de a dolgozatunkban a fent említett meghatározást fogjuk használni.

A felszíni vizek számos területen játszanak kiemelkedő szerepet az életünkben. Fontos összetevői a hidrolízis folyamatának, aminek során a víz a Föld felszínéről és a Föld felszínére jut el “mozgásával”. A lefolyó vizek és a csapadék táplálja, míg a talajba szivárgó víz és a párolgás apasztja ezeket a víztesteket. Továbbá emellett, hogy ezek a jelentős édesvízállományok nélkülözhetetlenek az emberek és az ökológiai rendszerek számára, a mezőgazdaságot, az ipart, a várostervezést, a közlekedést és az áramtermelést is meghatározzák.

Általános jelenség, hogy a felszíni vizek időszakosan megerednek vagy elapadnak, aminek hátterében különböző természeti, illetve emberi eredetű hatások állnak. Ezek a dinamikus természetű vizek képesek nagyban befolyásolni a környezetüket, aminek olyan súlyos következményei is lehetnek, mint az árvizek. Az ilyen szélsőséges esetek elkerülésének érdekében kulcsfontosságú a felszíni vizek kiterjedéseinek meghatározása és változásainak nyomonkövetése.

1.2. Célkitűzés

A távérzékelés egyik legsokrétűbb területe a műholdfelvételek multitemporális elemzése. Dolgozatunkban ennek a témakörnek egyik felhasználását kutatjuk. Olyan szoftver létrehozása a célunk, ami képes azokat a területeket detektálni, amelyek belvizes időszakban telítődnek, de aszályos vagy átlagos időszakban kiszáradnak. Például ilyenek lehetnek kiszáradt folyóvölgyek, folyókanyarulatok, és árterületek.

Projektünk elején a motivációt a Lechner Tudásközpont munkatársai fogalmazták meg. Arra kértek fel minket, hogy keressünk módszert az időszakos belvizek detektálásának automatizálására. Ezt eddig szakértők vizuális interpretációja alapján állították elő. Célunk ehhez hasonló minőségű megoldás előállítása volt anélkül, hogy a folyamat során a felhasználói tudásra hagyatkozzunk.

A megoldáshoz a feladatot több részfeladatra bontottuk fel. Elsőként mivel a víz detektálásánál nagy problémát okoznak a felhők és felhőárnyékok, ezekre hozunk létre egy maszkot az fmask algoritmus felhasználásával. Ezután a felvételek indexei, és küszöbölő algoritmusok segítségével vízmaszkot készítettünk a felvételekre, majd ezeken végeztünk változás vizsgálatot.

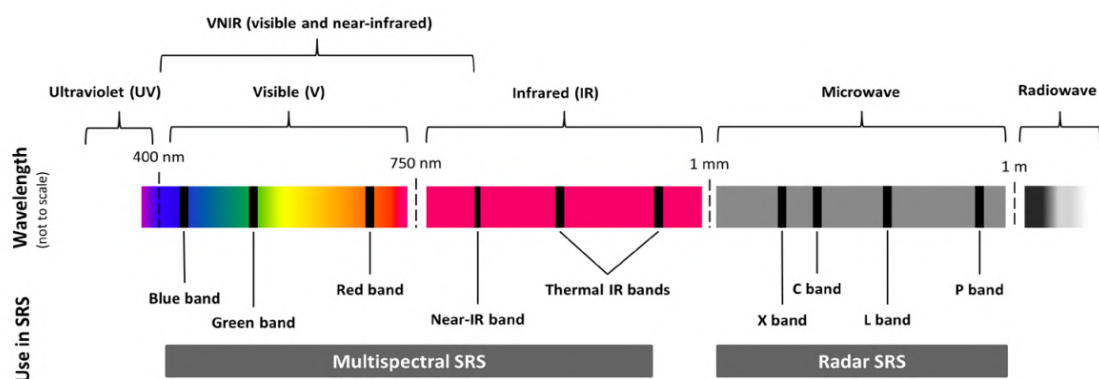
A dolgozat további részeiben bemutatjuk a felhasznált módszereket, technológiákat és az eredményeinket. A második fejezetben áttekintjük a dolgozat tudományos hátterét. A detektáló módszerünk részletes leírását a harmadik fejezet fogja tartalmazni. Az szoftverünk implementációjáról a negyedikben írunk. Az eredményeinket pedig az ötödik fejezetben mutatjuk be.

2. fejezet

Technológiai és szakirodalmi áttekintés

2.1. Multispektrális műholdfelvételek

Multispektrális felvételnek azokat a műholdképeket nevezzük, amelyek az elektromágneses spektrum különböző tartományjaiból foglalnak magukba adatokat. Ezek hullámhosszszávonként csatornába vannak elkülönítve képeken, például: az RGB, a közeli- és termális infravörös sávok. A legtöbb földmegfigyelő műhold így rögzíti az adatait. Különbség abban van közöttük, hogy hány csatornát, és milyen hullámhosszok értékeit tartalmazzák. A dolgozatunkban a Sentinel-2 és a Planet műholdak képeit dolgoztuk fel. Az elektromágneses sugárzás spektrumát a 2.1-es ábrán lehet megtekinteni.



2.1. ábra. Az elektromágneses sugárzás spektruma (nem méretarányos) és felhasználása a műholdas távérzékelésben.

Forrás: <https://www.mdpi.com/2075-4450/12/3/233/htm>

2.1.1. Sentinel-2 műhold

A Sentinel-2 az Európai Űrügynökség Kopernikusz programja keretében működtetett földmegfigyelő műholdpár. Mindkét műhold nagyfelbontású, multispektrális kamerával van felszerelve. Adataikat 13 sávban rögzítik a spektrum látható, közeli- és közepes infravörös tartományaiból, ezeknek a térbeli felbontása 10 és 60 méter között van. A sávokról bővebben a 2.1 táblázatban vannak információk. A Sentinel-2 felvételei sok területen felhasználtak, többek között mezőgazdaságban a termésbecsléshez és káresemények felméréséhez alkalmazzák. Emellett katasztrófavédelmi helyzetekben is jó hasznát veszik az általuk gyűjtött információnak. A képek felhasználását megkönnyíti, hogy akárki hozzájuk férhet és ingyenesen letölthetőek a Kopernikusz program weboldaláról¹. A projektünkhez a Level 1-C szinten feldolgozott adatokat használtuk fel.

Sáv sorszáma	Sáv neve	Felbontás (m)
B01	Coastal aerosols	60
B02	Blue	10
B03	Green	10
B04	Red	10
B05	Red Edge	20
B06	Red Edge	20
B07	Red Edge	20
B08	Near Infrared	10
B8A	Narrow Near Infrared	20
B09	Water Vapour	60
B10	Shortwave Infrared - Cirrus	60
B11	Shortwave Infrared	20
B12	Shortwave Infrared	20

2.1. táblázat. Sentinel-2 sávjainak sorszáma, elnevezése és felbontása

2.1.2. Planet műhold

A Planet Labs-et az amerikai Nemzeti Repülési és Űrkutatási Hivatal korábbi mérnökei alapították 2010-ben, akiknek céljuk, hogy nagyfelbontású, globális lefedettségű műholdhálózatukkal láthatóvá és elérhetővé tegyék bolygónk mindennapi változásait. A Planet három különböző műholdkonstellációval rendelkezik: PlanetScopes, SkySats és RapidEye. Ezek közül mi a PlanetScope felvételeivel dolgoztunk.

A PlanetScope műholdjai mindössze 10x10x30 centiméteresek, csoportosan, úgynevezett rajokban keringenek a Föld körül. Körülbelül 24x16 km-es területről tud egy mű-

¹<https://scihub.copernicus.eu/>

hold egyszerre felvételt rögzíteni 3.7 m-es felbontásban, négy multispektrális sávban: RGB és közeli infravörös (*near infrared, NIR*). Így naponta több mint 300 millió négyzetkilométernyi nagyfelbontású képeket képesek készíteni a Planet archívumába, amiket számos területen lehet felhasználni. A PlanetScope szenzorának fontosabb specifikációi a 2.2 táblázatban vannak. A szántóföldi területek és a termények folyamatos nyomon követésével és elemzésével hatékonyabb és fenntarthatóbbak lehetnek a mezőgazdaságok. Optimalizálja az erdészetben a fakitermelést, könnyebben információhoz juthatnak az erdők egészségi állapotáról, illetve megelőzhetővé teszik az illegális erdőirtást is. Sőt, még az infrastruktúra vagy az energiaipari eszközök felügyeletében is hatásos segítség lehet. Hátránya, hogy csak az ügyfelek számára elérhetőek a felvételek, viszont használatuk kifizetődő.

Műszer	PS2.SD
Pályamagasság (referencia adat)	450 - 580 km (98° dőlésszög)
Szenzor típus	Négysávós képkocka képkészítő butcher-block szűrővel
Spektrális sávok	Blue: 464 - 517 nm Green: 547 - 585 nm Red: 650 - 682 nm NIR: 846 - 888 nm
Földminta távolsága	3.7m - 4.1m (megközelítőleg, magasságfüggő)
Keret mérete	24km x 16km (megközelítőleg)
Képfelvételi kapacitás	200 millió km ² /nap
Képek bitmélysége	12-bit

2.2. táblázat. A PlanetScope konstelláció PS2.SD szenzorának specifikációja

2.2. Vízindexek

Optikai képekből úgy lehetséges a felszíni vizeket kinyerni, hogy a víznek az infravörös csatornákon a visszaverődése alacsonyabb, mint a többi földtakaró típusnak, így el lehet különíteni a vizet. Ebből a tényből kiindulva több módszert is kifejlesztettek, amivel könnyedén lehet ezeket a víztesteket szűrni a felvételekből. Ilyen megoldás egyetlen infravörös sáv sűrűségi szeletelése a víztérkép előállításához [2]. Emellett olyan talajtérképekből állítottak elő könnyedén vízmaszkokat, amiket felügyelt vagy felügyeletlen osztályozási módszerrel hoztak létre [3]. Továbbá döntési fákat is építettek a multispektrális sávokra, hogy elhatárolják a vízborítást a többi földtakaró típustól [4].

Az ilyen típusú módszerekkel azonban az a probléma, hogy az osztályozási szabályaik nehezen felépíthetőek és nem mindig alkalmasak általános használatra. A víz kivonásának egy egyszerű és hatékony módja a vízindexek használata, amelyeket két vagy több sávból számolnak ki, és amelyekkel azonosítani lehet a vizes és nem vizes területek közötti különbségeket. Számos indexet fejlesztettek ki a felszíni vízterületek vagy az árvíz elöntésének mértékének kivonására.

A NDVI (*normalized difference vegetation index*), ami igazából egy vegetációs index, egy régen is és manapság is ritkán, de használt index, amivel detektálták a vizet és az árvizeket néhány tanulmányban [5] [6]. Ennek az indexnek a használatával igazából csak a föld feletti biomasszát tudjuk azonosítani, amiből lehet következtetni a víz jelenlétére, ami viszont nem nyújt információt közvetlenül a felszíni vizekről. Természetesen vannak célzottan a víz és a nem vizes objektumok elkülönítésére használt indexek is, mint például a TCW (*tasseled cap wetness*) index, amelyet a felszíni reflexiós adatok hat sávjából származtatnak [7]. De jobban kitudják szűrni a vizes területeket az NDWI (*normalized difference water index*) és az mNDWI (*modified NDWI*) indexek is, amiket McFeeters 1996-os tanulmánya óta használnak széleskörben [8] [9].

Az NDWI a zöld és a közeli infravörös (NIR) spektrális sávok közötti normalizált rádiókülönbség, amely nemcsak a vízjellemzőket erősítheti fel, hanem egyidejűleg a növényzetet és más talajjellemzőket is lecsökkentheti. Ezzel szemben az mNDWI annyival másabb, hogy a NIR sávot kicserélték a SWIR sávra és úgy képezték a normalizált különbségüket. Így az mNDWI-al egy stabilabb és megbízhatóbb indexet kapunk, mivel tanulmányok bebizonyították, hogy a SWIR sáv kevésbé érzékeny a vízben található üledékekre és egyéb optikailag aktív anyagokra, mint a NIR sáv. Emellett a SWIR sáv képes a víz néhány finom tulajdonságát is tükrözni.

Néhány multispektrális műhold (Landsat, Sentinel 2) felvételeinek 2 féle SWIR sávja van, a SWIR1 és a SWIR2. Az mNDWI egy másik változatát úgy kaphatjuk meg, ha a SWIR1 helyett a SWIR2 sávot alkalmazzuk a zöld sávval. Ez a helyettesítés nagyon alacsony visszaverődési értékeket eredményez a vízjellemzők esetében, így érzékeli a víztesteket és csatornákat, valamint a zaj formájában megjelenő növényzetet. Az mNDWI-2 képből viszont az NDVI index segítségével eltávolítható ez a növényzet, így egy vízdetektálásra alkalmas indexet kapunk [10].

Egy limitációja viszont akad az mNDWI indexnek, hogy nem képes megkülönböztetni a vizet és a havat, mivel a hó esetében a zöld sáv és a SWIR sáv közötti normalizált különbség ugyanolyan nagy, mint a vízé. Tehát hótakaróval fedett területeknél érdemes

más index-szel kiegészítve használni, például az NDSI-vel (*normalized difference snow index*).

Bevezettek továbbá egy új indexet, az AWEI-t (*automated water extraction index*), aminek van egy árnyékos területekre is jól működő (AWEIsh) és egy nem árnyékos területekre jól működő (AWEInsh) változata. Ez a módszer 4 sáv algebrai kombinációjából készíti el a vízindexet: a zöld, NIR, SWIR1 és a SWIR2 sávokból. Először Landsat felvételeken alkalmazták egy 2016-os tanulmányban, amivel kiváló eredményeket értek el [11]. Egy ismét 2016-os tanulmányban tűnt fel a WI2015 nevű index is, ami nagyjából azonosan erős az eddig említett indexekkel. Ezt az indexet a NIR és SWIR sávokon a felszíni reflexióból lineáris diszkriminancia-elemzéssel hozták létre [12].

$$NDVI = \frac{NIR - Red}{NIR + Red}$$

$$NDWI = \frac{Green - NIR}{Green + NIR}$$

$$NDMI = \frac{NIR - SWIR1}{NIR + SWIR1}$$

$$NDMI_d = \frac{B8a - SWIR1}{B8a + SWIR1}$$

$$NDWI_g = \frac{SWIR1 - Red}{SWIR1 + Red}$$

$$NDSI_g = \frac{SWIR1 - NIR}{SWIR1 + NIR}$$

$$RLAI = \frac{B05 - Red}{B05 + Red}$$

$$RENDVI = \frac{B06 - B05}{B06 + B05}$$

$$GRVI = \frac{NIR - Green}{NIR + Green}$$

$$NVIA = \frac{B07 - SWIR1}{B07 + SWIR1}$$

$$mNDWI = \frac{Green - SWIR1}{Green + SWIR1}$$

$$mNDWI2 = \frac{Green - SWIR2}{Green + SWIR2}$$

$$NDSI = \frac{Blue - NIR}{Blue + NIR}$$

$$BSI11 = \frac{(SWIR1 + Red) - (NIR + Blue)}{(SWIR1 + Red) + (NIR + Blue)}$$

$$BSI12 = \frac{(SWIR2 + Red) - (NIR + Blue)}{(SWIR2 + Red) + (NIR + Blue)}$$

$$TCW = 0.1509 * Blue + 0.1973 * Green + 0.3279 * Red + 0.3406 * NIR \\ - 0.7112 * SWIR1 - 0.4572 * SWIR2$$

$$WI2015 = 1.7204 + 171 * Green + 3 * Red - 70 * NIR - 45 * SWIR1 \\ - 71 * SWIR2$$

$$AWEInsh = 4 * (Green - SWIR1) - (0.25 * NIR + 2.75 * SWIR2)$$

$$AWEIsh = Blue + 2.5 * Green - 1.5 * (NIR + SWIR1) - 0.25 * SWIR2$$

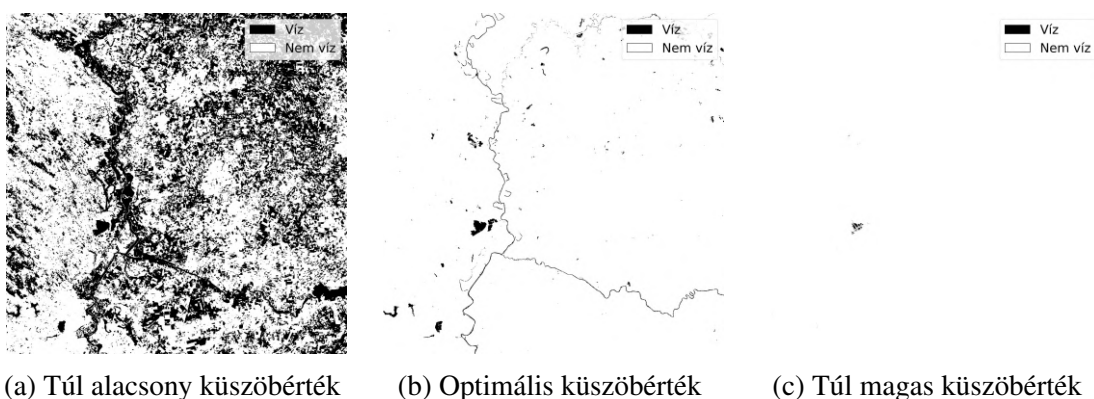
2.3. táblázat. Az összes vizsgált index kiszámításának képletei.

2.3. Küszöbölő algoritmusok

Egy objektumot ki lehet vonni a háttérből egy képből küszöbérték kiválasztásával. Ideális esetben, ha az objektum különböző átlagos szürkeségi szintje eltér a környezétől, a küszöbértékek hatására az eredeti képből egy fehér objektumot kapunk fekete háttérrel, vagy fordítva. Azaz a küszöbölési művelet kimenete egy bináris kép, amelynek egy állapota az előtérben lévő objektumokat jelzi.

A küszöbértékek meghatározása az egyik legkritikusabb kérdés a vízindexek víztestek kivonására történő felhasználása során. A víz visszaverődési jellemzői alapján a vízre vonatkozó NDWI és mNDWI értékek általában 0-nál nagyobbak. Ezért gyakran 0 küszöbértéket alkalmaznak a víz kivonására az indexképekből.

Ehelyett azonban manapság az az egységes és elfogadott vélemény, hogy a küszöbérték beállításával általában jobb extrakciós eredményeket lehet elérni. Ez különösen akkor nehézkes, ha egyazon víztestet lefedő képsorozat vagy egy víztestcsoportot lefedő egyetlen kép küszöbértékének meghatározása történik. Ezekben az esetekben a küszöbnek a "legjobb" küszöbnek kell lennie abban az értelemben, hogy két osztályba kell választania az előtérbe tartozónak tekintett világosabb objektumokat és a háttérbe tartozónak tekintett sötétebb objektumokat. Számos tanulmány foglalkozik ezzel a kérdéssel, amiknek egy része küszöbölő algoritmusokkal keres egy optimális küszöbértéket a megadott képhez [13] [14]. Az optimális és nem optimális küszöbértékek hatásáról a 2.2 ábrán láthatunk példát.



2.2. ábra. Egy Sentinel-2-es tesztfelvétel vízmaszkjai különböző küszöbértékekkel. Túl magas küszöbértékekkel történő küszöbölés információvesztéssel eredményez, míg az alacsony küszöbértékekkel történő küszöbölés zavaró háttérzavarokat okozhat.

A küszöbölési módszereket az általuk kiaknázott információ szerint kategorizáljuk, mint például a hisztogram alakja, a mérési tér klaszterezése, az entrópia, az objektum attribútumok, a térbeli korreláció és a helyi szürkefelület. Megjegyzendő azonban, hogy egy ilyen kategorizálás szükségszerűen homályos, mivel egyes módszerek több kategóriába is tartozhatnak (például Otsu módszere egyszerre tekinthető hisztogram-alakú és klaszterező algoritmusnak). A küszöbölő algoritmusok közül 3-at vizsgáltunk meg a dolgozatunkban: Otsu, Minimum és a Yen. Ezekre láthatunk példát a 2.3 ábrán.

2.3.1. Otsu

Az Otsu algoritmus egy automatikus küszöbérték-választási módszer, amely egy szürkeárnyaltos képet két osztályt tartalmazó bináris képpé redukál. Azt a küszöbértéket keresi, amely minimalizálja az osztályon belüli varianciát, amelyet a két osztály varianciáinak súlyozott összegeként határozzunk meg. Feltételezi a szürkeségi értékek bimodális eloszlását és a képpontok két osztályát a képen. Az algoritmus előnye, hogy csak a szürkeségi szint hisztogramjára van szükség a kép küszöbértékének meghatározásához, minden más előzetes ismeret nélkül.

Működése

Vegyünk egy adott kép képpontjait L szürkeárnyalatban $[1, 2, \dots, L]$ reprezentálva. Az i szinten lévő pixelek számát n_i -vel jelöljük, a pixelek teljes számát pedig $N = n_1 + n_2 + \dots + n_L$. A tárgyalás egyszerűsítése érdekében a szürkeségi szintek hisztogramját normalizáljuk és valószínűségi eloszlásnak tekintjük:

$$p_i = \frac{n_i}{N}, \quad p_i \geq 0, \quad \sum_{i=1}^L p_i = 1$$

Most tegyük fel, hogy a pixeleket két osztályba soroljuk C_0 és C_1 (háttér és objektumok, vagy fordítva) egy küszöbértékkel k szinten; a C_0 a $[1, \dots, k]$ szintű pixeleket, a C_1 pedig a $[k + 1, \dots, L]$ szintű pixeleket jelöli. Ekkor az osztályok előfordulási valószínűségeit és az osztályok átlagos szintjeit a következő képletek adják meg:

$$\omega_0 = \sum_{i=1}^k p_i \quad \omega_1 = \sum_{i=k+1}^L p_i$$

$$\mu_0 = \frac{\sum_{i=1}^k p_i}{\omega_0} \quad \mu_1 = \frac{\sum_{i=k+1}^L p_i}{\omega_1}$$

Ezek alapján az osztály varianciák az alábbi képletekből állnak elő:

$$\sigma_0 = \frac{\sum_{i=1}^k (i - \mu_0)^2 * p_i}{\omega_0} \quad \sigma_1 = \frac{\sum_{i=k+1}^L (i - \mu_1)^2 * p_i}{\omega_1}$$

Így a teljes számítási egyenlet a következőképpen írható le:

$$\sigma_W^2 = \omega_0 \sigma_0^2 + \omega_1 \sigma_1^2,$$

ahol σ_W^2 az osztályon belüli variancia. Az az érték, ami minimalizálja a σ_W^2 -t lesz a keresett küszöbértékünk(k) [15].

2.3.2. Minimum

Az Otsuhoz hasonlóan ez is bimodális hisztogramot feltételez. Egy bimodális hisztogramban a pixelek eloszlása két különböző módusszal és két különböző csúccsal rendelkezik. A hisztogramot iteratív módon simítjuk egy 3 méretű futó átlag segítségével, amíg csak két helyi maximum nem lesz. Ezután a kettő közötti minimumot választja meg a küszöbértéknek. Az olyan képek, amelyek hisztogramja rendkívül egyenlőtlen csúcsokkal vagy széles és lapos völgyel rendelkezik, nem alkalmasak erre a módszerre. [16].

Az algoritmus az alábbi leírás szerint működik:

1. Kiszámoljuk az egyes intenzitástartományokba eső képpontok számát és az egyes intenzitástartományok középpontját.
2. Beállítjuk a simított hisztogram kezdeti értékét.

3. Tovább simítjuk a simított hisztogramot, majd elmentjük egy listába a lokális maximumokat.
4. Addig ismételjük a 3. lépést, amíg a lokális maximumok száma nem kevesebb, mint 3 vagy amíg nem érjük el a maximális iterációk számát.
5. Ha nem egyenlő 2-vel a lokális maximumok száma, akkor nem lehet 2 lokális maximumot találni a hisztogramhoz.
6. Ha túlléptük a maximális iterációk számát, akkor hiba emelkedik.
7. Egyébként pedig kiválasztja a hisztogram 2 lokális maximum közötti területéről a legkisebb értéket és visszaadja azt.

2.3.3. Yen

Ennek az algoritmusnak a leírása egy új kritériumot javasol a többszintű küszöbérték meghatározására, ami két tényezőt vesz figyelembe. Az első a küszöbözött és az eredeti kép közötti eltérés, a második pedig a küszöbözött kép ábrázolásához szükséges bitek száma. A kétszintű küszöbértékelés új maximális korrelációs kritériuma alapján meghatározzuk az eltérést, majd a többszintű küszöbértékeléshez egy olyan költségfüggvényt javasolunk, amely mindkét tényezőt figyelembe veszi. A költségfüggvény minimalizálásával automatikusan meghatározható a szürkeségi szintek osztályozási száma és a küszöbértékek [17].

Az algoritmus az alábbi leírás szerint működik:

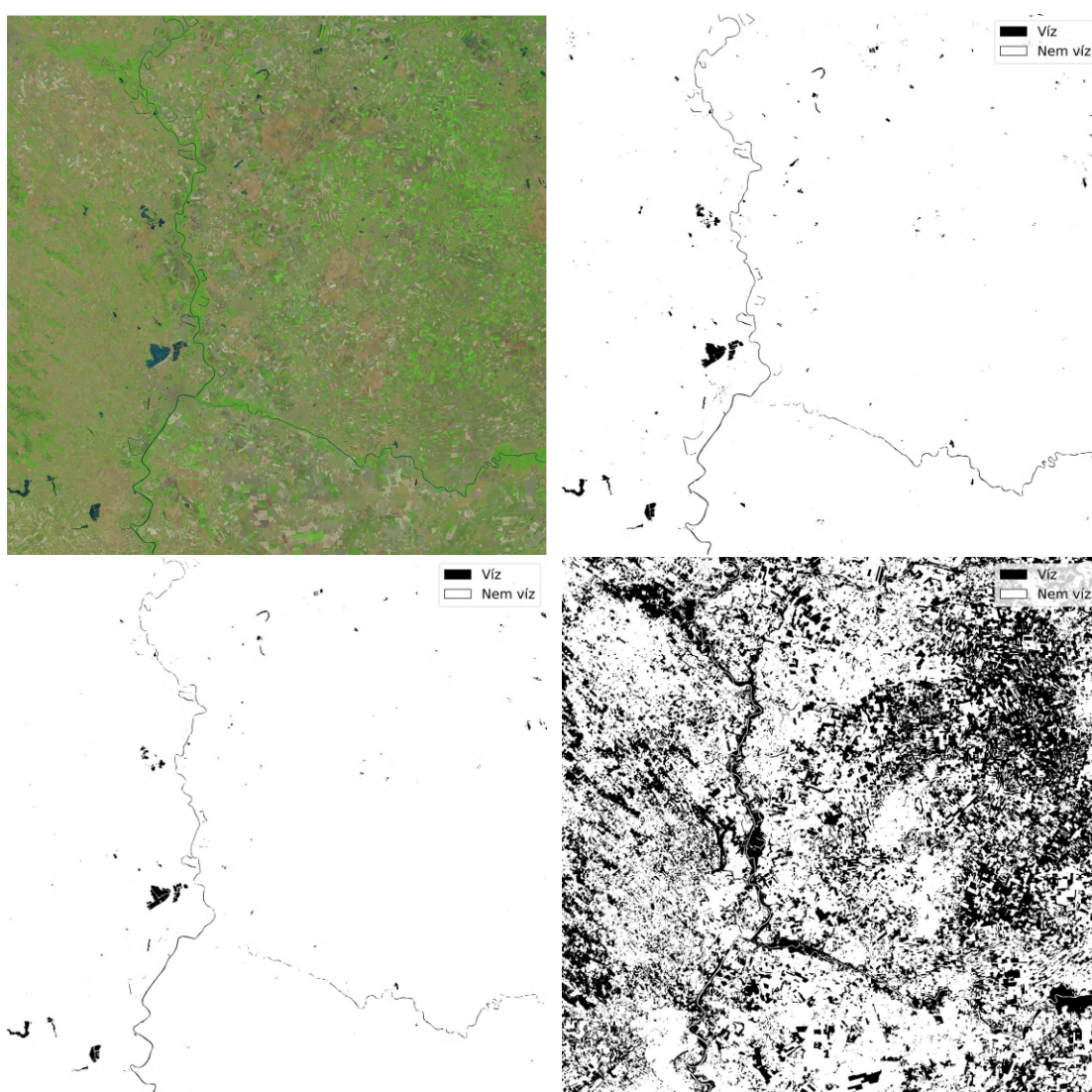
1. Kiszámoljuk az egyes intenzitástartományokba eső képpontok számát és az egyes intenzitástartományok középpontját.
2. Ha int típusú üres kép volt a bemenet, akkor egy elemű lesz az intenzitástartományok középpontjainak száma, így azt az egy darab középpontot adjuk vissza.
3. Ha nem üres kép volt a bemenet, akkor kiszámítjuk a valószínűségi tömegfüggvényt az egyes intenzitástartományokba eső képpontok számára.
4. Majd ennek vesszük az összesített normalizált hisztogramját(P_1), továbbá a négyzetének az összesített normalizált hisztogramját(P_1_{sq}).
5. Végül kiszámoljuk a valószínűségi tömegfüggvény négyzetének a végéről számított összesített normalizált hisztogramját(P_2_{sq}).

6. Ezután az alábbi képlettel kiszámolt tömbnek vesszük a logaritmusát elemenként:

$$(P1_{sq} * P2_{sq})^{-1} * (P1 * (1 - P1))^2$$

7. Majd a logaritmusokat tartalmazó tömbből megkeressük a maximális elemnek az indexét, és a középpontokból visszaadjuk az azon az indexen álló elemet.

A minimum és yen algoritmus működését az *skimage.filters.threshold_all()*² implementáció alapján írtam le. Az implementációk megtalálhatóak a 4.2 és a 4.3 forráskódokban.

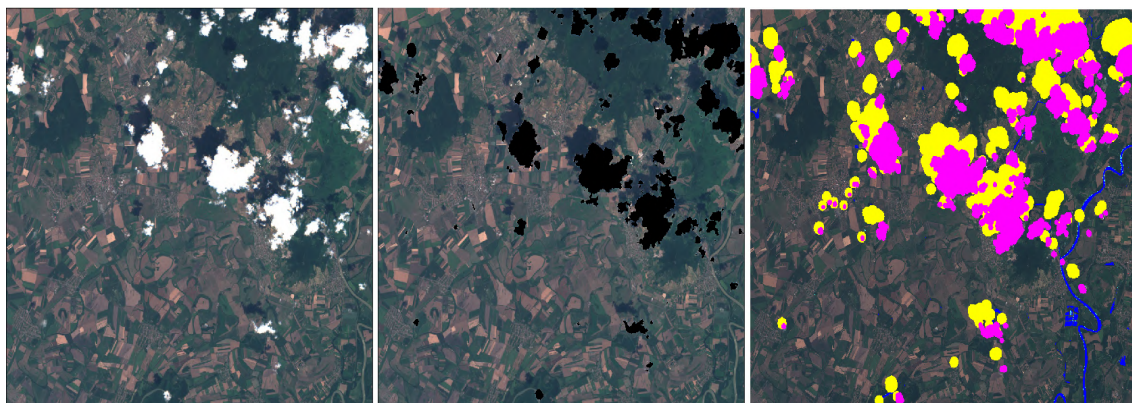


2.3. ábra. A vizsgált küszöbölő algoritmusok eredményei egy Sentinel-2-es tesztadaton Bal felül a tesztkép TCI-je (*true color image*) látható, majd az algoritmusok által elkészített vízmaszkok ebben a sorrendben: Yen, Minimum, Otsu.

²<https://scikit-image.org/docs/stable/api/skimimage.filters.html>

2.4. Felhő és felhőárnyék detektálás

A felhők és árnyékaik detektálása elengedhetetlen lépés a műholdfelvételek feldolgozása során. Egyrészt a változások keresésekor fals pozitív eredményt kaphatunk, ha egyik felvétel felhős a másik pedig tiszta. Másrészt az indexek számításakor is negatív hatásuk van. A felhők szárazföldre vetett árnyékainak vízindexe nagyon hasonlít a belvizekhez, így a küszöbölő algoritmusok nem fogják tudni kiszűrni őket a folyamat során. Ezeknek a hibáknak elkerülése érdekében maszkokat készítettünk a feldolgozott műholdfelvételekre. Erre több módszert is kivizsgáltunk, többek között a világos felhő-pixelek küszöbölő algoritmussal kiszűrését, valamint az s2cloudless³ algoritmust is. Ezek közül az utóbbi jó eredményeket mutatott, de a felhőárnyékok detektálására ez a módszer nem volt alkalmas.



2.4. ábra. Példa az s2cloudless és fmask algoritmussal készült felhőmaszkokra

A végső kiválasztott módszer az fmask [18] lett. Ez egy eredetileg Landsat felvételekre kidolgozott algoritmus, ami képes felhőket és árnyékaikat detektálni. Az alkalmazott implementáció az algoritmus egy továbbfejlesztett verziója [19], ami már működik a Sentinel-2-es képekre is. Példát a 2.4 ábrán láthatunk mindkét algoritmusra.

Az fmask működése nagy vonalakban így írható le: az algoritmus első körben a B09-es vízpára sáv alapján megkeresi a potenciális felhő pixeleket, majd ezt szárazföld pixelektől a spektrális változékonyság valószínűsége alapján, víz pixelektől pedig a fényesség szerint korrigálja.

A felhőárnyékok meghatározásához szükség van a magasságuk kiszámítására. Landsat képeken ehhez a termális sáv van felhasználva, de mivel ez Sentinel-2-ből hiányzik, ezért itt egy fix 200 és 1200 méter közötti keret van meghatározva. Ezt összevetve a felvétel idejében rögzített nap irányszöggel, hozzárendel a felhőpixelektől árnyékpixeleket.

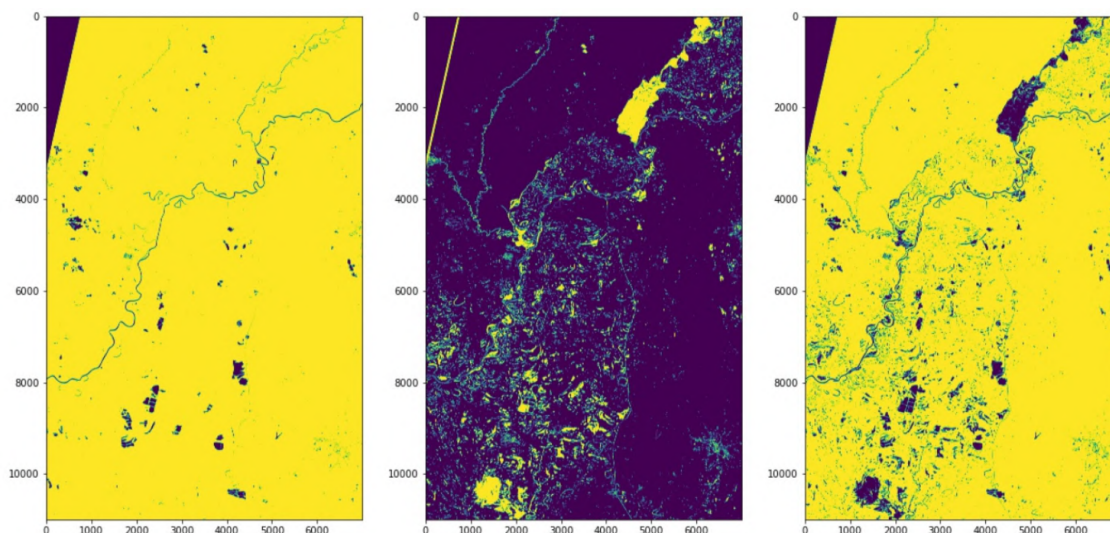
³<https://medium.com/sentinel-hub/cloud-masks-at-your-service-6e5b2cb2ce8a>

A termális sáv hópixelek detektálásához is használt Landsat képeken. Ez a lépés Sentinel-2 képeknél kimarad. Ez hópixelek maszkból kihagyásához vezethet, viszont a nagyobb hőmérsékletű havas területek belekerülhetnek.

Az fmask algoritmusnak a python-fmask⁴ csomagban elkészült implementációját használtuk fel, amit többek között a QGIS Cloudmasking kiegészítője is alkalmaz.

2.5. Felszíni vizek változáselemzése

Műholdfelvételek változáselemzésére sokféle megoldás született már, de mivel a folyamatomunkban ezt a lépést az elkészített vízmaszkokon végezzük más módszert választottunk. A különbségeket a képek strukturális hasonlóságának vizsgálatával kerestük meg. Ez az algoritmus a képek luminanciájának, kontrasztjának és struktúrájának összevetésével működik. A programunkban a scikit-image Python csomagban elkészült implementációt használtuk fel. Ennek alkalmazásával különböző időpontok vízmaszkjaira visszakaptuk a hasonlósági indexeket, emellett létrejött egy hasonlósági kép is. Ennek küszöbölésével kaptuk meg azokat a területeket, ahol lényeges változás történt, amiről a végeredményt a 2.5 ábrán láthatjuk. A strukturális hasonlóság index (SSIM) és az ezt kiszámoló módszer eredetileg egy 2004-es tanulmányban lett publikálva [20] és azóta a képfeldolgozási terület egyik leggyakrabban felhasznált értekezése lett.



2.5. ábra. Két vízmaszk és a rajtuk talált különbségek

⁴<https://www.pythonfmask.org>

3. fejezet

Módszertan

3.1. A kiválasztott adatok

Az elkészített alkalmazásunk Sentinel-2-es adatokat képes feldolgozni. Ezen belül, hogy a saját felhőmaszkot el tudjuk készíteni Level-1C feldolgozottságú adatokra van szükség. A folyamat lefuttatásához kell két felvétel azonos területről, egyik ideálisan egy aszályos, vagy átlagos időszakról, a másik pedig belvizes időpontról. Ezek az adatok elérhetőek a Kopernikusz program oldalán¹.

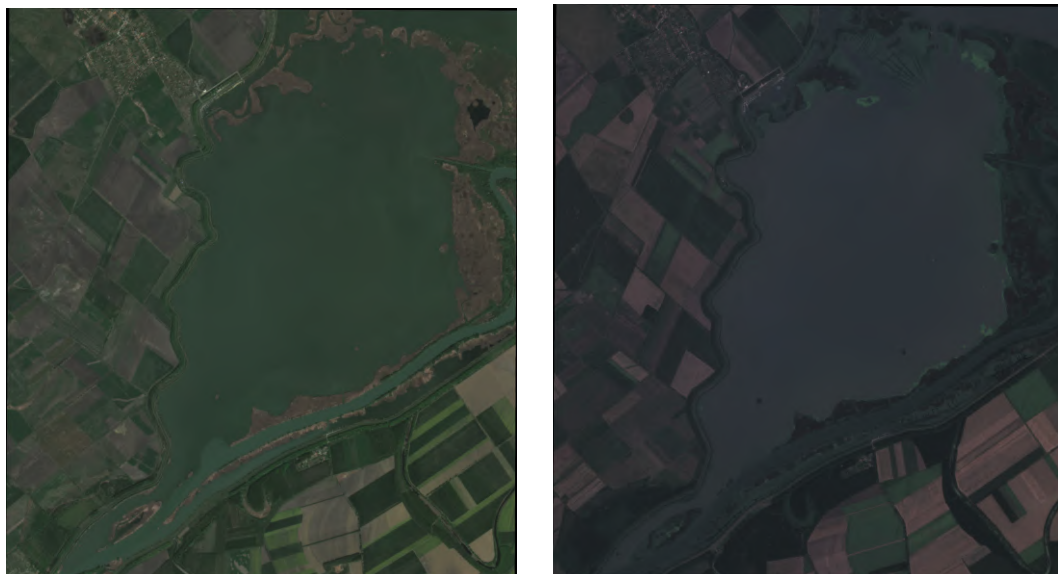
Teszteléshez a Sentinel-2 34UEU azonosítójú csempéjén megtalálható északkelet-magyarországi területet választottunk. Az itt kapott eredmények verifikálásához felhasználtunk a Lechner Tudásközpont által előállított vízmaszkokat 2018. április 20-ai és 2018. augusztus 20-ai időpontokra. Ezek hasonlóan Sentinel-2-es felvételek alapján készültek, de a fedett területük eltér a standard csempétől, ezért a releváns részeket kivágtuk mindkét fajta felvételből és ezeket egyeztetjük össze. Ezen kívül teszteltük a folyamat felhővel fedett felvételekre is (34UEU csempe 2018.03.28. és 2018.07.04.). A Lechner Tudásközponttól kapott felvételek közül a 3.2 ábrán látható kettő felvétel.

Továbbá az indexeket és a küszöbölő algoritmusokat mind Sentinel-2-es, mind PlanetScope-s felvételeken is megvizsgáltuk. A Sentinel-2-es tesztfelvételek meg-egyeznek az alkalmazás tesztelése során használt felvételekkel, kiegészítve néhány az EarthExplorer oldaláról² letöltött felvétellel. Az utóbbi felvételek a Tisza Tisza-ék-ske-Szanád (Szerbia) közötti részeken készültek a Sentinel-2 34TDS azonosítójú csempéjén. A PlanetScope felvételeit az ELTE Informatikai Karától kaptuk, amelyek a Tisza-tó alsó

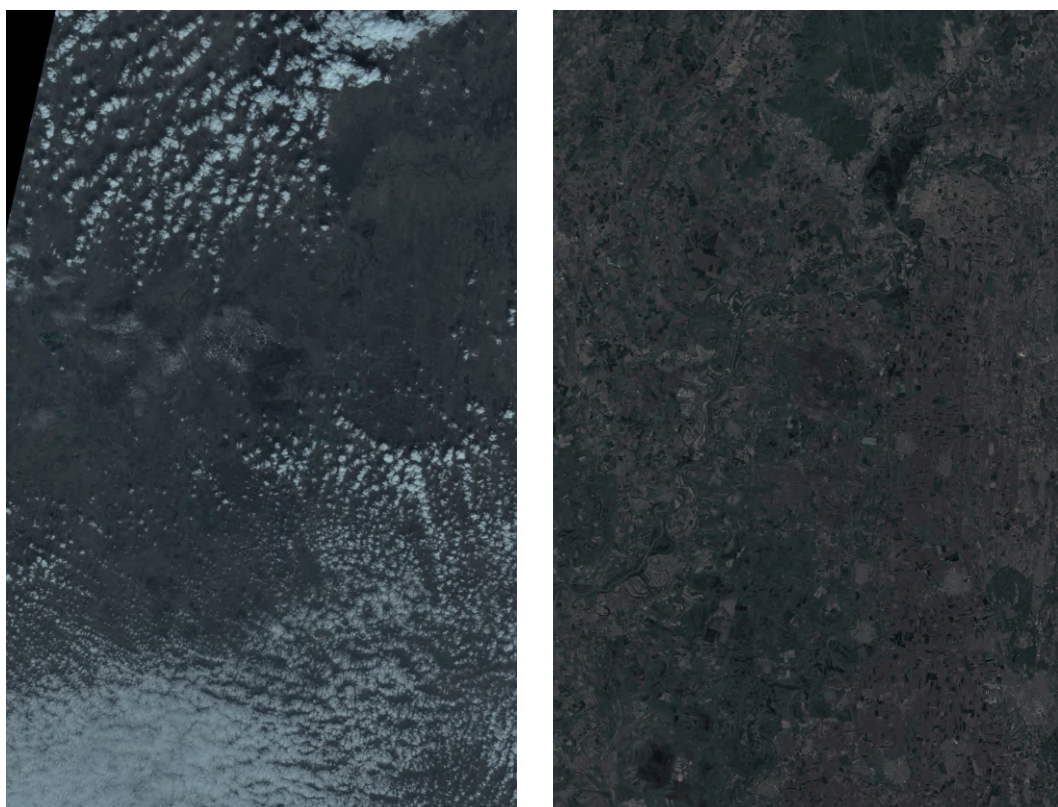
¹<https://scihub.copernicus.eu/>

²<https://earthexplorer.usgs.gov/>

részéről készültek 2019.04.21. és 2019.07.30. dátummal. Ezeket a 3.1 ábrán lehet megtekinteni.



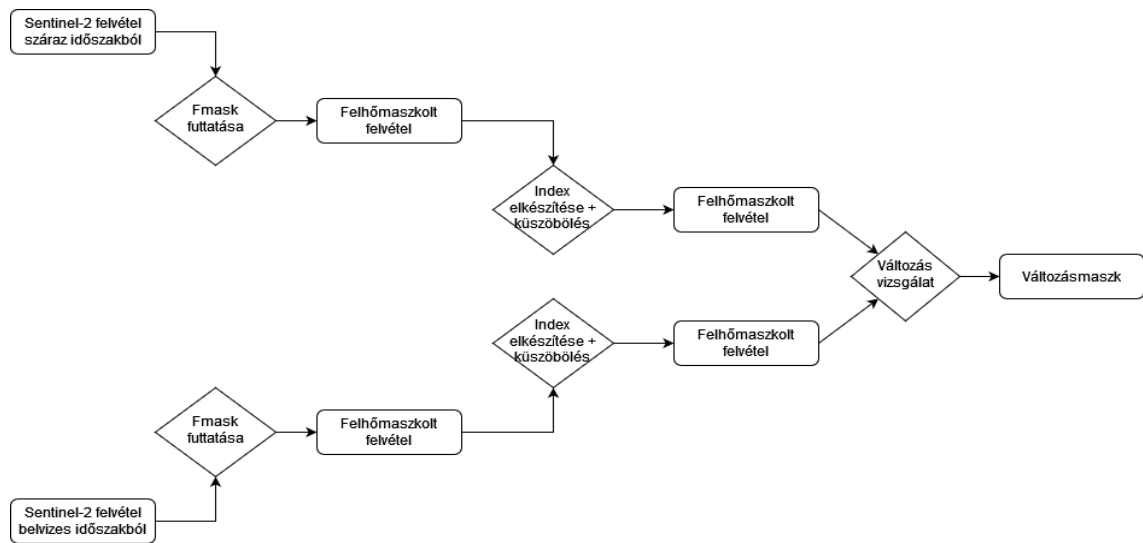
3.1. ábra. PlanetScope tesztadatok egy tavaszi (bal oldal), illetve egy nyári időszakban.



3.2. ábra. A Lechner Tudásközponttól kapott Sentinel-2-es tesztadatok egy felhős (bal oldal), illetve egy napos időszakban.

3.2. A kiválasztott eszközök

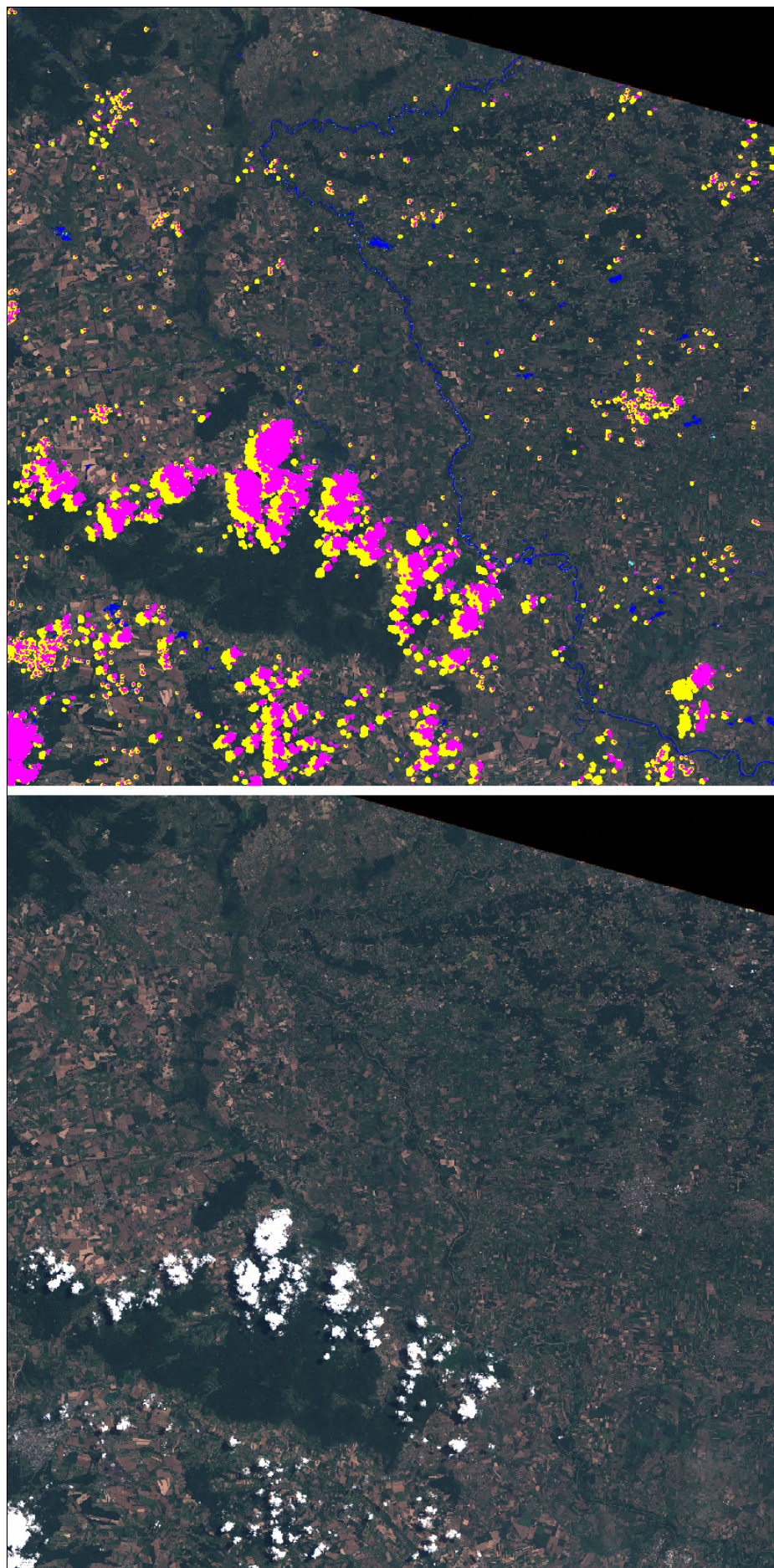
A belvíz detektálás folyamatát megoldásunkban három részfeladatra bontottuk fel. Első lépésként preprocesszálást végzünk el a műholdfelvételeken, hogy kiszűrjük a felhőket és felhőárnyékokat. Második lépésünk a vízmaszk készítése a felhőmaszkolt képekből. Ennek megoldásához kiszámítjuk a felvételek bizonyos indexét, majd ezt küszöböléssel bináris maszkká alakítjuk át. Végül a harmadik lépés változásdetektálás az elkészült vízmaszkokon. Ezeknek a részfeladatoknak egy folyamatábrája a 3.3 ábrán látható, a leírását pedig a következő bekezdések fogják tartalmazni.



3.3. ábra. A folyamat flowchart-ja

3.2.1. Preprocesszálás

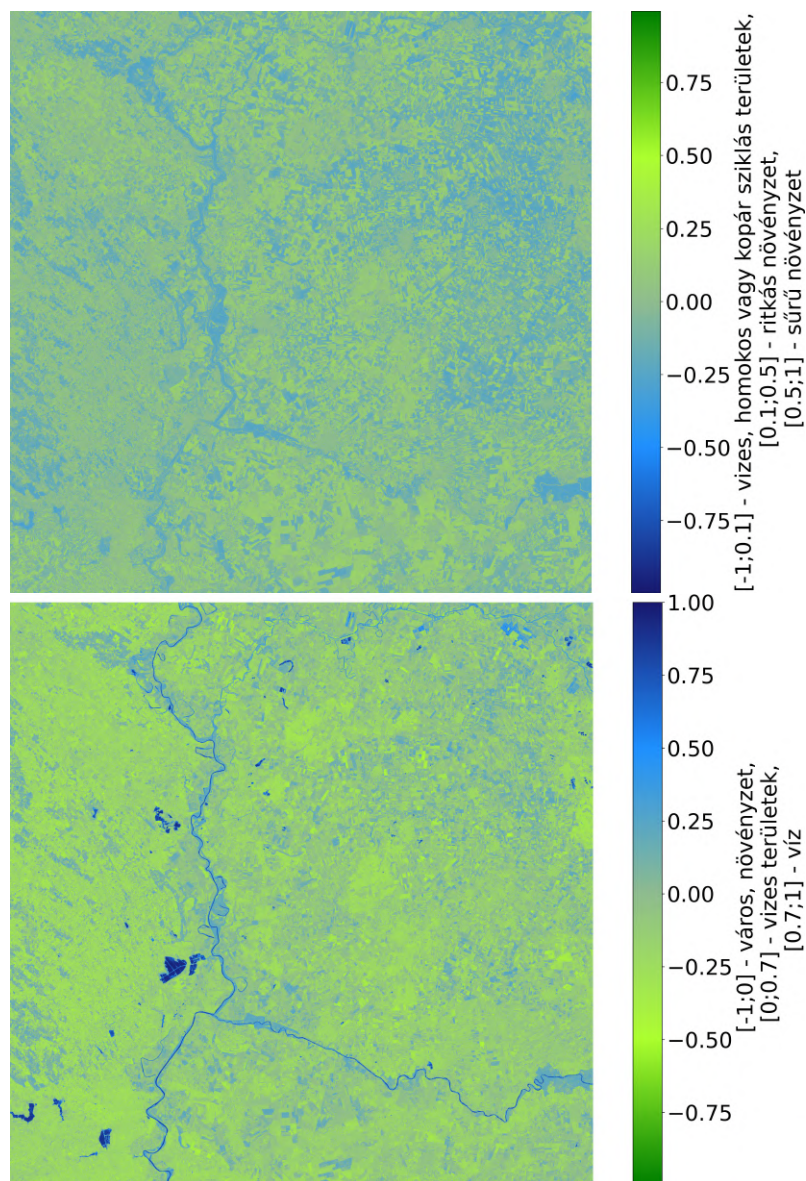
Ebben a lépésben az fmask[19] algoritmust futtatjuk le mindkét Sentinel-2 képre. Az eredmény így két olyan maszk lesz, ami meghatározza a felhő, felhőárnyék, a víz és a hó pixeleket. A feladatunkhoz ez a vízmeghatározás nem elég pontos, és havas területek nem relevánsak. Ezért a felhasznált maszkból ezeket kitöröljük, és csak felhő és felhőárnyék pixeleket töröljük ki a feldolgozandó műholdfelvételekből. Mindegyik képre a folyamat elején ráhelyezzük a saját maszkját, majd a változásvizsgálat elején a felhőmaszkokat összeeszeljük, és mindkét vízmaszkra ráhelyezzük, a felhőkből adódó különbségek detektálásának elkerüléséhez.



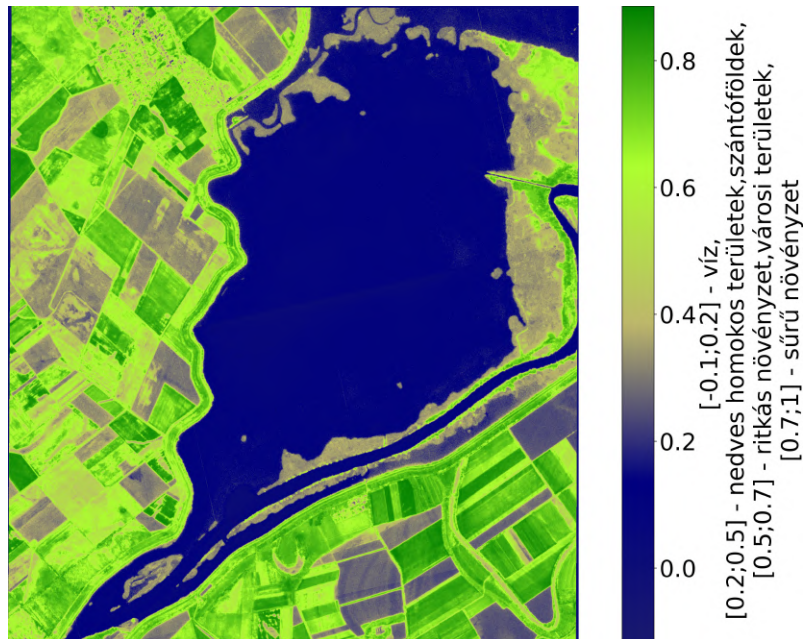
3.4. ábra. 2018.07.04-ei Sentinel-2 felvétel és a hozzá készült fmask (felhő: rózsaszín, felhőármények: sárga, kék: víz)

3.2.2. Vízmaszk készítés

A vizes és nem vizes területek megkülönböztetésének egyik legegyszerűbb és leghatékonyabb módszere a vízindex készítés, hiszen csak magára a nyers geotiff fájlra van szükségünk. Egy geotiff fájl a multispektrális műholdfelvétel típusától függően minimum 3 sávból áll. A sávok száma és típusa határozza meg, hogy milyen vízindexeket tudunk elkészíteni. Míg egy PlanetScope felvétele 4 sávból, addig egy Sentinel-2-es felvétel 13 sávból áll, így az előbbiből 4 féle, az utóbbiból 19 féle indexet tudunk megvizsgálni a kutatásunk során. Az indexeket és előállításainak képleteit a 2.3 -es táblázatban tüntettük fel, példát pedig a 3.5 és a 3.6 ábrákon láthatnak.



3.5. ábra. Különböző vízindexek egy Sentinel-2-es tesztfelvételhez
Felül az NDVI indexről, alul pedig az mNDWI indexről láthatunk példát



3.6. ábra. A GRVI index egy PlanetScope-os tesztfelvételhez

Az eredmények differenciáját a különböző spektrális sávok kombinációja adja, ami miatt egyik index a másikhöz képest más területeket tesz kontrasztosabbá. Ezekre az indexekre felhasználva a három kiválasztott küszöbölő algoritmust (Otsu, Minimum, Yen) eltérő pontosságú bináris vízmaszkokat kaptunk. Az eredmények kiértékeléséhez először minden vízindex értékkészletéből kiszámítottunk egy 40 elemszámú, kéttizedesre kerekített lebegőpontos számokat tartalmazó listát, amiben az elemek egyenletes távolságra vannak egymástól. A listából az ismétlődő elemeket kivettük, hogy minden érték egyedi legyen benne. Majd minden értéket küszöbértékként felhasználva készítettük el a vízindexhez tartozó vízmaszkokat, amik közül a teljesen fehéreket vagy feketéket nem vettük figyelembe. A kapott vízmaszkok közül manuálisan kiválasztottuk a vizuálisan legoptimálisabbat (vagy egy optimális értékekkel teli intervallumot), és ennek a maszknak a küszöbértékéhez hasonlítottuk a vízindexből és az algoritmusokból generált 3 vízmaszk küszöbértékét az algoritmusok pontosságának osztályozásához.

Index neve	Értékkészlete	Megfelelő threshold	Minimum	Otsu	Yen
BSI11	[0.22 ; 0.5]	0.39	0.31	0.39	0.37
BSI12	[0.32 ; 0.51]	0.44	0.36	0.41	0.43
GRVI	[0.12 ; 0.76]	0.28 - 0.38	0.29	0.35	0.42
NDMI	[0.38 ; 0.83]	0.74-0.78	0.76	0.61	0.58
NDMId	[0.34 ; 0.82]	0.6 - 0.7	0.64	0.56	0.54
NDSIg	[0.14 ; 0.57]	0.2 - 0.3	0.23	0.34	0.36
NDVI	[0.13 ; 0.76]	0.3 - 0.4	0.3	0.35	0.4
NDWI	[0.23 ; 0.84]	0.6 - 0.7	0.64	0.61	0.52
NDWIg	[0.11 ; 0.6]	0.2 - 0.3	0.26	0.33	0.45
NVIA	[0.38 ; 0.84]	0.78	0.77	0.62	0.58
RENDVI	[0.35 ; 0.69]	0.44	0.44	0.51	0.46
RLAI	[0.43 ; 0.6]	0.46	0.45	0.49	0.54
TCW	[0.46 ; 0.48]	0.47	0.41	0.47	0.45
mNDWI	[0.34 ; 0.86]	0.8	0.72	0.63	0.49
mNDWI2	[0.22 ; 0.78]	0.56 - 0.66	0.59	0.56	0.46
NDSI	[0.13 ; 0.85]	0.71	0.66	0.61	0.53
AWEInsh	[0.53 ; 0.58]	0.57	0.48	0.55	0.52
AWEIsh	[0.44 ; 0.49]	0.47 - 0.48	0.48	0.47	0.47
WI2015	[0.43 ; 0.51]	0.49	0.5	0.48	0.45

3.1. táblázat. Sentinel-2-es tesztfelvétel küszöbértékeinek vizsgálata

Index neve	Értékkészlete	Megfelelő threshold	Minimum	Otsu	Yen
NDVI	[0.24 ; 0.85]	0.45	0.57	0.55	0.55
NDWI	[0.11 ; 0.8]	0.48	0.54	0.5	0.42
GRVI	[0.17 ; 0.85]	0.48	0.42	0.48	0.59
NDSI	[0.11; 0.88]	0.46	0.52	0.49	0.43

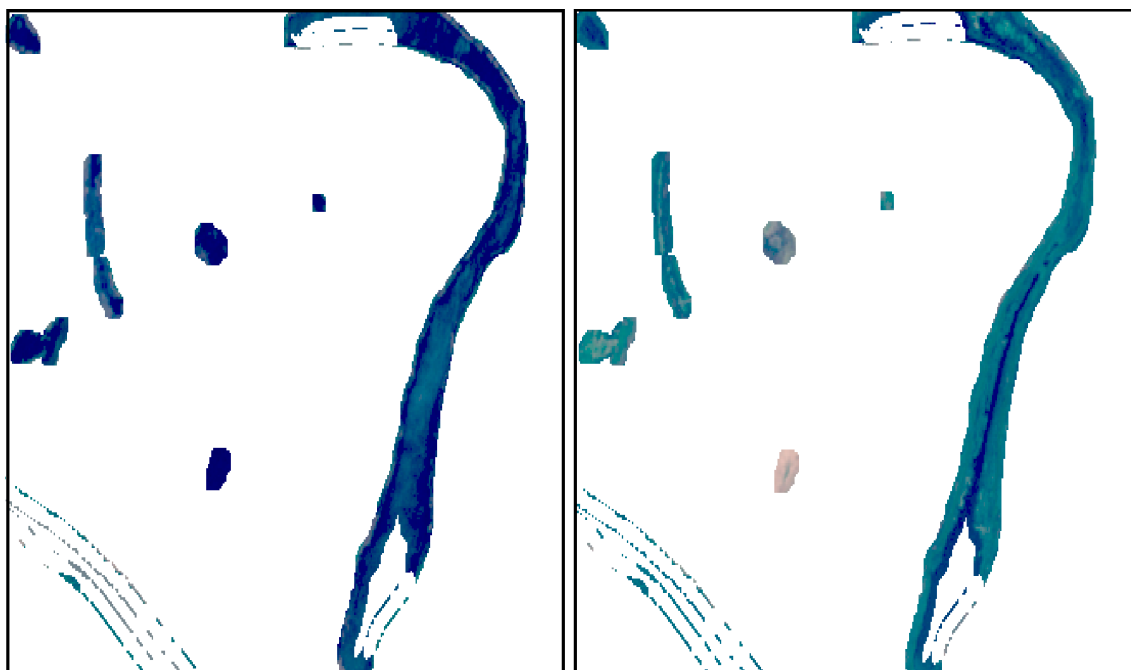
3.2. táblázat. PlanetScope-os tesztfelvétel küszöbértékeinek vizsgálata

A 3.1 és 3.2 táblázatok egy-egy felvétel vízmaszk vizsgálatának adatait mutatják. Egy sora a táblázatnak úgy épül fel, hogy az első oszlopa az index neve, amire vizsgáljuk a vízmaszkot, majd azt követi az index értékkészlete. A megfelelő threshold a manuálisan kiválasztott küszöbértéket vagy az optimális küszöbértékek intervallumát jelenti a harmadik oszlopban. A következő 3 oszlopban a 3 küszöbölő algoritmus által kiszámított küszöbértékek találhatók.

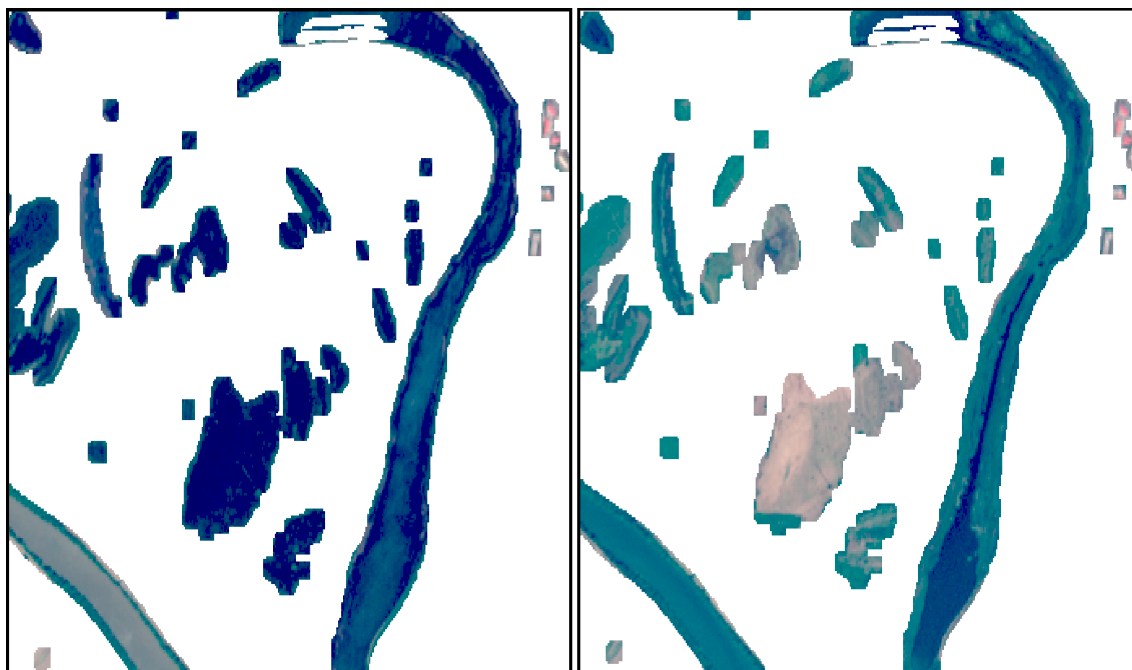
3.2.3. Változásdetektálás

Folyamatunk utolsó lépése a változásdetektálás, ahol az elkészült vízmaszkokat hasonlítjuk össze egymással. Ennek a lépésnek a célja egy olyan maszk elkészítése, ahol csak a két vízmaszk különbségei jelennek meg. Így kiszűrve az állandó vízfelszíneket, például tavakat és folyókat, és meghagyva az időszakosan feltöltődő területeket.

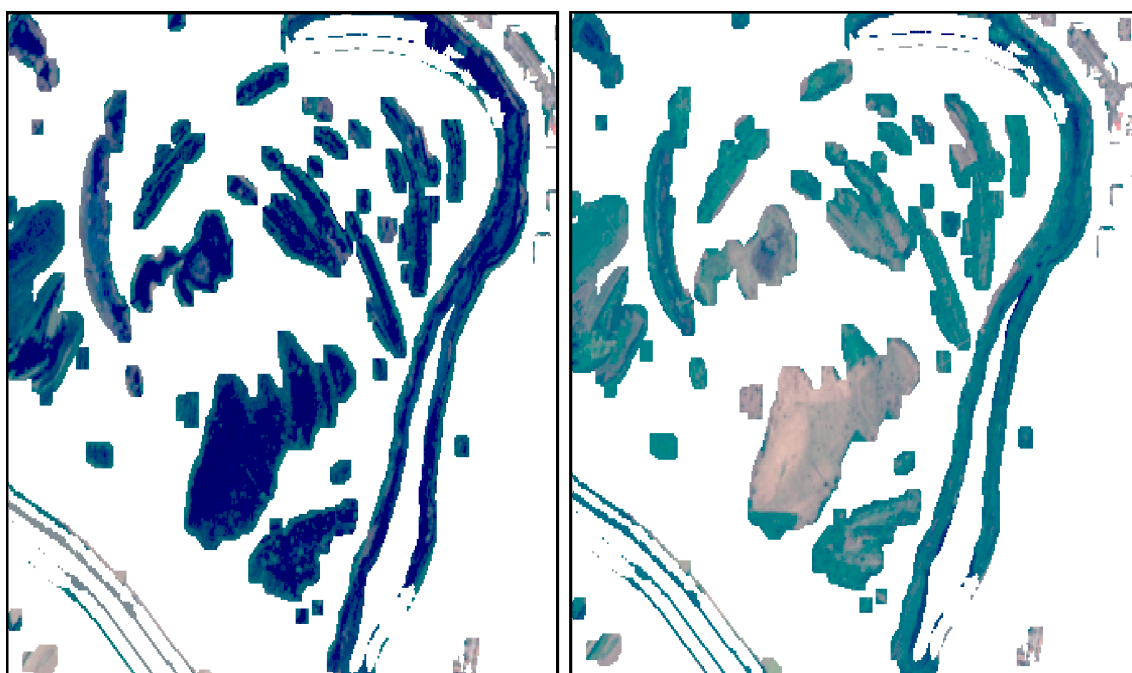
Ez a lépés a strukturális hasonlóságot összehasonlító algoritmussal[20] lett elvégezve. Itt az algoritmus komplikáltsága miatt nem készítettünk saját implementációt, hanem a scikit-image structural_similarity metódusát használtuk fel. A különböző index és küszöbölő algoritmus párosításokkal készített vízmaszkok különbségeire a 3.7, 3.8, illetve a 3.9 ábrákon láthatunk példákat.



3.7. ábra. 2018.04.20-ai (bal) és 2018.08.20-ai (jobb) 34UEU csempe különbség maszkja mNDWI index Yen küszöböléssel.



3.8. ábra. 2018.04.20-ai (bal) és 2018.08.20-ai (jobb) 34UEU csempe különbség maszkja NDWIg index Minimum küszöböléssel.



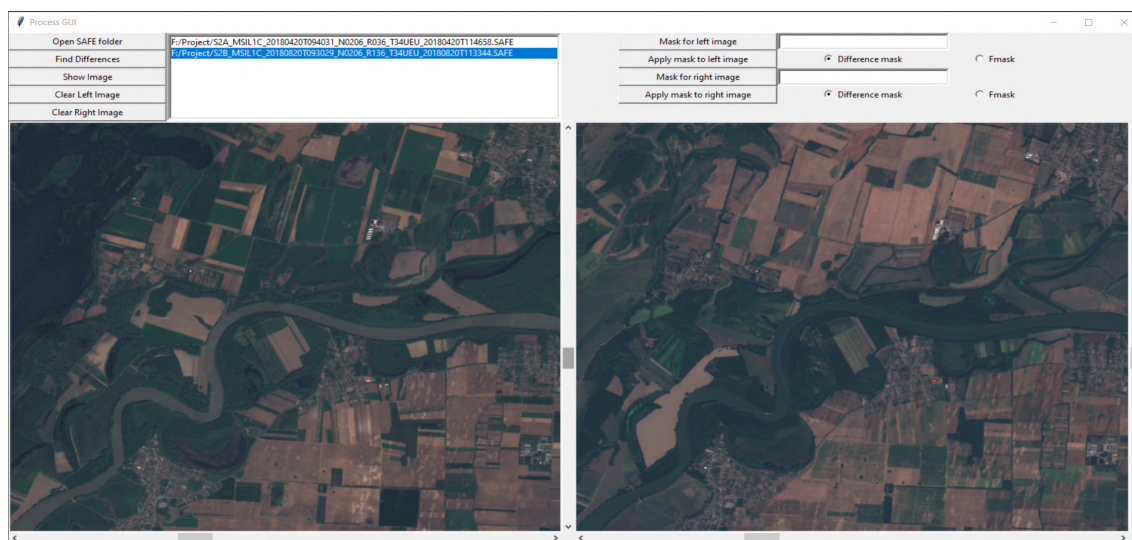
3.9. ábra. 2018.04.20-ai (bal) és 2018.08.20-ai (jobb) 34UEU csempe különbség maszkja NDWIg index Otsu küszöböléssel.

4. fejezet

Implementáció

4.1. Elkészült alkalmazás

A dolgozat készítéséhez létrehoztunk egy grafikus alkalmazást, ami az egész folyamatunkat le tudja futtatni Sentinel-2-es adatokra. Ez Python nyelven lett megvalósítva MVC architektúrával, a GUI Tkinter csomaggal készült. A megoldást segítették a Pythonban implementált képfeldolgozási (pl.: scikit-image¹, OpenCV²), valamint térinformatikai csomagok (pl.: GDAL³, pythonfmask⁴).



4.1. ábra. A GUI kezdőképernyője két Sentinel-2-es felvétel megjelenítésével

¹<https://scikit-image.org/>

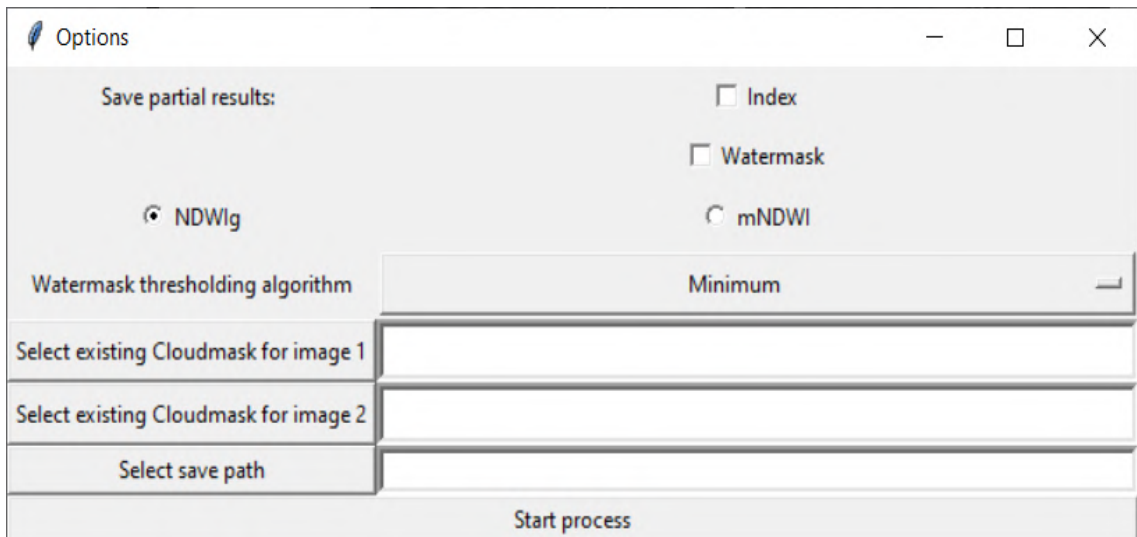
²https://docs.opencv.org/4.x/d6/d00/tutorial_py_root.html

³<https://gdal.org/>

⁴<https://www.pythonfmask.org/en/latest/>

Programunk futtatása során kiválasztható, hogy mNDWI vagy NDWIg indexszel valamint hogy milyen küszöböléssel akarjuk létrehozni a vízmaszkokat. Megadható, hogy melyik részeredményeket akarjuk elmenteni, és felhasználható már korábban elkészített felhőmaszk is. Jelenleg csak a tesztelt módszereket tartalmazza, de a program könnyen bővíthető más vízindexek használatára, és további küszöbölő algoritmusok alkalmazására. Ezen felül lehetséges a kód adaptálása más idősor elemzési feladatok megoldására. Ehhez a vízmaszk készítését kellene kicserélni, például növényzeti, vagy beépítettségi maszkok készítésére.

A projekt forráskódja a kari GitLab szerveren, a következő URL-en elérhető:
<https://gitlab.inf.elte.hu/lechner/summer-lab-2021/project-1>



4.2. ábra. Beállítások a folyamat futtatása előtt

4.2. Planet felvételek indexelése

```

1 def make_index_for_planet(self, index_type, tif_path):
2     dataset = gdal.Open(tif_path)
3     BLUE = dataset.GetRasterBand(1).ReadAsArray().astype(np.
4         float32)
5     GREEN = dataset.GetRasterBand(2).ReadAsArray().astype(np.
6         float32)
7     RED = dataset.GetRasterBand(3).ReadAsArray().astype(np.
8         float32)
9     NIR = dataset.GetRasterBand(4).ReadAsArray().astype(np.
10        float32)

```

```
8     bandA, bandB = None, None
9     if index_type == "NDVI" :
10         bandA = NIR
11         bandB = RED
12     elif index_type == "NDWI":
13         bandA = GREEN
14         bandB = NIR
15     elif index_type == "GRVI":
16         bandA = NIR
17         bandB = GREEN
18     elif index_type == "NDSI":
19         bandA = BLUE
20         bandB = NIR
21     else:
22         print("index_type is not found")
23
24     geotransform = dataset.GetGeoTransform()
25     originX, pixelWidth, empty, finalY, empty2, pixelHeight =
26         geotransform
27     cols = dataset.RasterXSize
28     rows = dataset.RasterYSize
29
30     projection = dataset.GetProjection()
31
32     finalX = originX + pixelWidth * cols
33     originY = finalY + pixelHeight * rows
34
35     index = np.divide(bandA-bandB, bandA+bandB, where=(bandA -
36         bandB)!=0)
37     extentArray = [originX, finalX, originY, finalY]
38
39     return index, extentArray
```

4.1. forráskód. Indexszámítás a PlanetScope felvételeire, saját metódus.

4.3. Néhány használt algoritmus

```
1 def threshold_minimum(image=None, nbins=256, max_num_iter=10000, *,
    hist=None):
```

```

2     counts, bin_centers = _validate_image_histogram(image, hist,
3         nbins)
4
5     smooth_hist = counts.astype(np.float64, copy=False)
6
7     for counter in range(max_num_iter):
8         smooth_hist = ndi.uniform_filter1d(smooth_hist, 3)
9         maximum_idx = find_local_maxima_idx(smooth_hist)
10        if len(maximum_idx) < 3:
11            break
12
13    if len(maximum_idx) != 2:
14        raise RuntimeError('Unable to find two maxima in histogram')
15
16    elif counter == max_num_iter - 1:
17        raise RuntimeError('Maximum iteration reached for histogram
18        smoothing')
19
20    threshold_idx = np.argmin(smooth_hist[maximum_idx[0]:
21        maximum_idx[1] + 1])
22    return bin_centers[maximum_idx[0] + threshold_idx]

```

4.2. forráskód. Minimum küszöbölő algoritmus implementációja az skimage package-ben.

```

1 def threshold_yen(image=None, nbins=256, *, hist=None):
2     counts, bin_centers = _validate_image_histogram(image, hist,
3         nbins)
4
5     if bin_centers.size == 1:
6         return bin_centers[0]
7
8     pmf = counts.astype(np.float32) / counts.sum()
9     P1 = np.cumsum(pmf)
10    P1_sq = np.cumsum(pmf ** 2)
11    P2_sq = np.cumsum(pmf[::-1] ** 2)[::-1]
12
13    crit = np.log(((P1_sq[-1] * P2_sq[1:]) ** -1) *
14        (P1[:-1] * (1.0 - P1[:-1])) ** 2)
15
16    return bin_centers[crit.argmax()]

```

4.3. forráskód. Yen küszöbölő algoritmus implementációja az skimage package-ben.

5. fejezet

Eredmények

A PlanetScope és az EarthExplorer-ről származó felvételekre vizuális összehasonlításon kívül nem tudtunk más pontosságvizsgálatot végezni. Viszont a Lechner Tudásközponttól kapott Sentinel-2-es felvételekkel hozzájutottunk szakértő által készített víz és felhőmaszkokhoz is. Ezeket összevetve egy differencia vizsgáló algoritmus segítségével a saját eredményeinkkel kaptunk egy százalékos összehasonlítási eredményt.

5.1. Vízmaszk

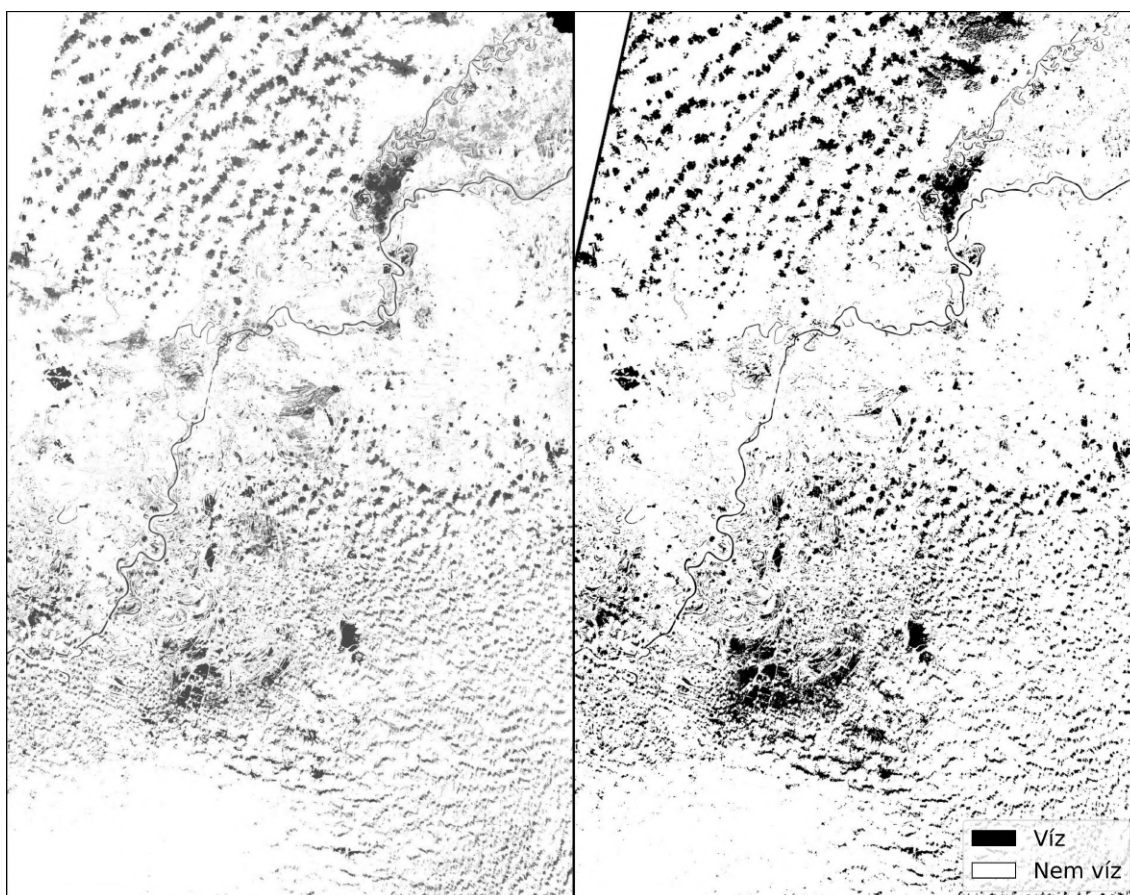
A vízmaszkokat összehasonlítottuk százalékosan felhős és felhőtlen területeken is. Mindegyik indextípus és algoritmus párból készítettünk egy vízmaszkot, majd ezeket az általunk generált vízmaszkokat vetettük össze a szakértőktől kapott vízmaszkokkal. A legjobb eredményeket tartottuk meg és tüntettük fel. Felhős területeken kisebb volt a vízmaszk pontossága, viszont így is mind a 4 vizsgált maszk 95% feletti eredményt mutatott. Ezeket az eredményeket a 5.1 táblázatban lehet megtekinteni. A vizuális összehasonlítási példákat pedig a 5.1, 5.2 és a 5.3 ábrákon lehet látni.

Felvétel ideje	Index neve	Algoritmus neve	A vízmaszk pontossága
2018.03.28.	mNDWI2	Yen	95.70%
2018.04.12.	mNDWI	Otsu	98.72%
2018.04.20.	NDWIg	Yen	99.15%
2018.08.20.	GRVI	Yen	98.79%

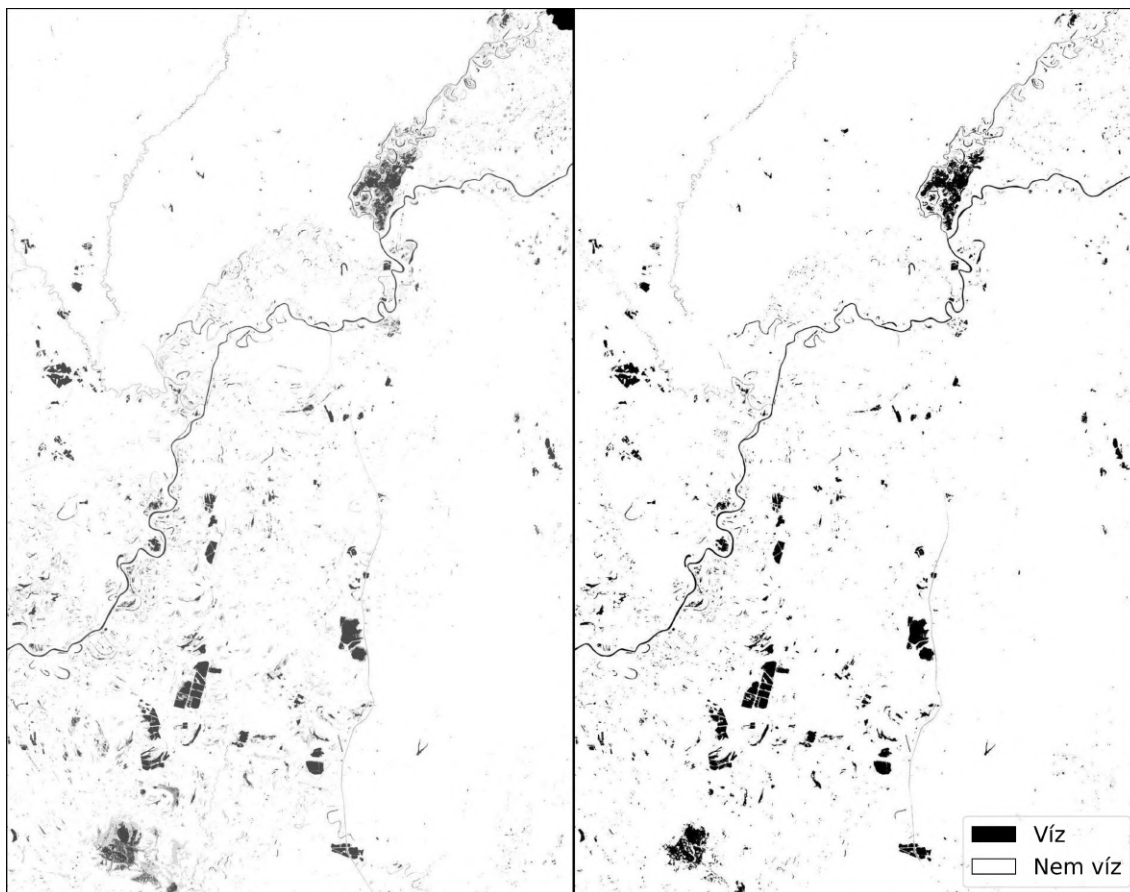
5.1. táblázat. A Lechner Tudásközponttól kapott felvételekre végzett százalékos összehasonlítás eredményei



5.1. ábra. PlanetScope tesztfelvételének eredeti képe (bal oldal) és a vízmaszkja
A vízmaszk a 2019.04.21.-ei felvételre készült a GRVI indexből a Minimum algoritmus
által meghatározott küszöbértékkel.



5.2. ábra. Sentinel-2-es felhős tesztfelvétel pontossága: 95.70%



5.3. ábra. Sentinel-2-es felhőtlen tesztfelvétel pontossága: 98.72%

5.2. Felszíni vizek változásdetektálása

A változásdetektálás tesztelése a Sentinel-2 34UEU és 34TET csempéin lett elvégezve a 2018. április 20-a és 2018. augusztus 20-ai felvételekkel. Ezeken vizsgálva lettek az NDWIg és az mNDWI indexszel készült vízmaszkok Yen, Otsu és Minimum küszöböléssel létrehozott verziói. Ezek validálására felhasználtuk a 34UEU csempe és a Lechner Tudásközpont által biztosított adatok átfedését.

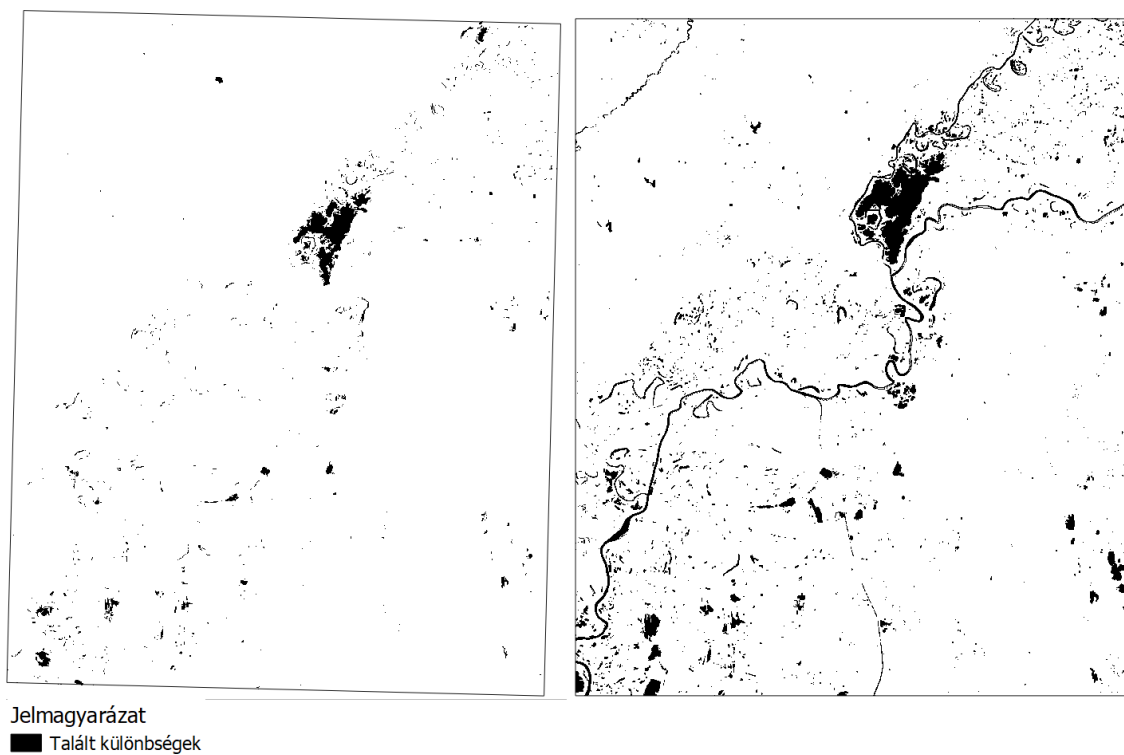
A folyamat során gondot okozott a felhőmaszkolás és küszöbölő algoritmusok egymást befolyásolása. A felhőmaszkolás után futtatott Yen küszöbölés több felvételnél is hibát okozott, az algoritmus nullával való osztásba futott, így itt nem született eredmény. Ez elsősorban az NDWIg-ből származó vízmaszkok létrehozásakor fordult elő, a mNDWI-ből számoltaknál nem okozott gondot. A Minimum küszöbölés viszont pont az utóbbi esetben hibázott, ugyanis a felhőmaszkolt mNDWI indexeknél a határértéket nem tudta jól kiszámolni, és üres vízmaszkokat adott meg eredményként.

A legjobb eredményeket NDWIg küszöbnél a Minimum algoritmus adta, az így készült vízmaszkokon talált különbségek 92,6%-ban egyeztek meg a Lechner Tudásközpont belvíz térképeinek különbségeivel. Ezeknél nagyobb hibát az okozott, hogy a mi vízmaszkjainkban eléggé eltért az állandó vizek (pl.: Tisza) értéke, hogy a változásdetektálásban ezek belekerültek a maszkba. Továbbá néhány szántóföld adott fals pozitív találatot, ahol a belvizes időszakban még a termények látszódtak így magasabb volt az NDWIg értékük. A száraz időszakban viszont már learatták őket, és már nem kerültek bele az ekkor készült maszkba, ezzel változást okozva a végső eredményünkben. Az otsu küszöbölt vízmaszkoknál ezek a hibák még relevánsabbak voltak, és így csak 86,9%-os volt az egyezés.

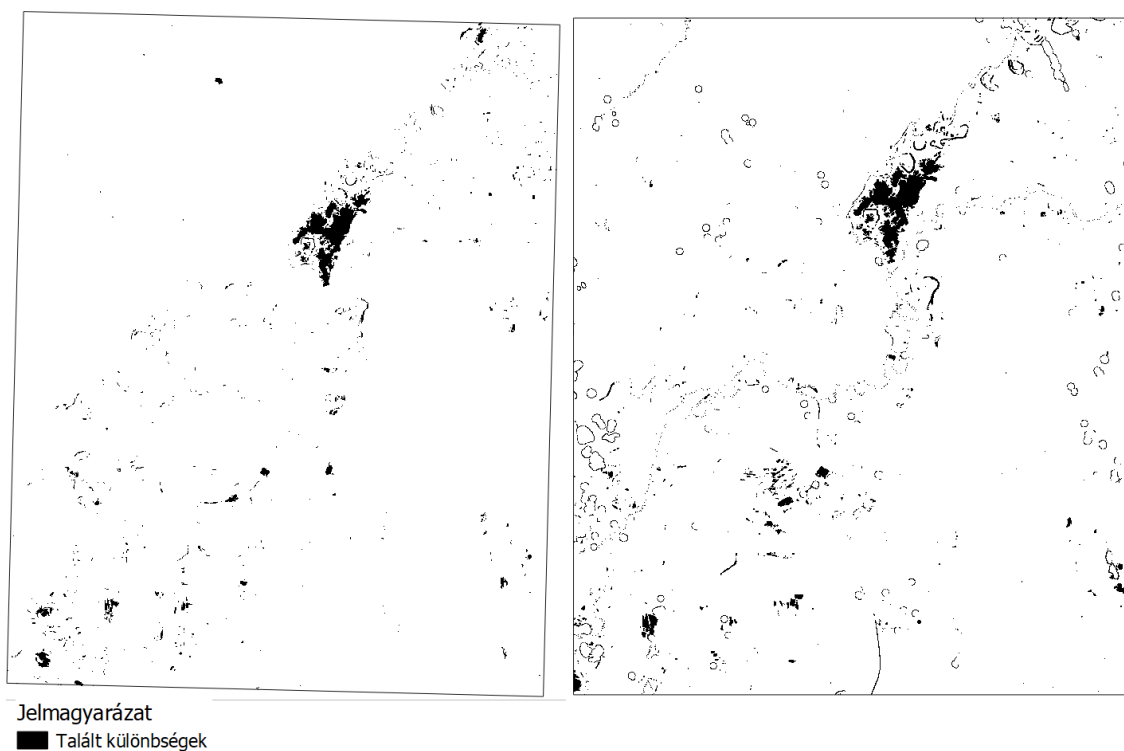
Az mNDWI-s vízmaszkokra legjobb eredménnyel a Yen küszöbölés szolgált. Itt a megegyezés a szakértői adatokkal 95,3% volt. Ez a módszer jobban kiszűrte mind az állandó vizeket, mind a szántóföldeket. A havas és vizes területek meghatározása pedig a kiválasztott időszak miatt nem okozott gondot. Az Otsu küszöbölt vízmaszkok különbségei itt nagyon eltértek a validáló adatoktól, csak 8,9%-os egyezés volt közöttük.

NDWIg	Minimum	92,6%
	Otsu	86,9%
mNDWI	Yen	95,3%
	Otsu	8,9%

5.2. táblázat. Eredmények összefoglalása



5.4. ábra. A Lechner Tudásközpont vízmaszkjainak és az NDWIg Minimum küszöböléssel készült vízmaszkjainak különbségei



5.5. ábra. A Lechner Tudásközpont vízmaszkjainak és az mNDWI Yen küszöböléssel készült vízmaszkjainak különbségei

6. fejezet

Összefoglaló

Tanulmányunk célja időszakosan belvizes területek detektálása volt műholdfelvételeken. Ennek megoldására létrehoztunk egy szoftvert, ami képes ezt a feladatot automatikusan elvégezni Sentinel-2-es képeken. Emellett készítettünk a PlanetLabs műhold felvételeire is hasonlóan alkalmazható módszert.

A kutatás során teszteltünk többféle megoldást a folyamatunk minden lépésére. Kivizsgáltuk a műholdfelvételek preprocessálására használt módszereket, és ezek közül az fmask algoritmust alkalmaztuk. Vízaszkok előállításánál szintén többféle indexet is teszteltünk. Tanulmányoztuk ezeknek előnyeit és hátrányait, majd kiválasztottuk az NDWI-g-t és az mNDWI-t feladatunk megoldására leginkább alkalmasnak. Ezután vizsgáltuk, hogy ezekkel melyik küszöbölő eljárás párosítható a legjobb eredmények érdekében. Itt a Minimum, Yen és Otsu algoritmusok lettek meghatározva legpontosabbnak.

A kiválasztott módszerek tesztelését Sentinel-2 és PlanetScope műholdfelvételeken végeztük. Validáció céljából kiválasztottunk egy Sentinel-2 csemjét, amelynek területe egyezett a Lechner Tudásközpont szakértői által biztosított adatokéval. Itt lefuttattuk a folyamatunkat az összes kiválasztott módszer alkalmazásával, és eredményeiket összevetettük egymással. Legjobb egyezést a validáló adatokkal az mNDWI index és Yen küszöbölés kombinációja adta 95,3%-kal. Emellett jó eredményt mutatott az NDWI-g és Minimum páros is 92,6%-kal.

Összegzésként elmondhatjuk, hogy a módszerünk a szakértői megoldást jól közelítő megoldással szolgált. Mindezt anélkül, hogy a folyamat futtatása során manuális beavatkozást igényeljen, így megkönnyítve belvíz-térképezési feladatokat.

Köszönetnyilvánítás

A kutatási munka és TDK dolgozat az ELTE Informatikai Kar és az InforNess Training Kft. ösztöndíjas pénzügyi támogatásával, továbbá a Lechner Tudásközpont szakmai támogatásával valósult meg. A PlanetScope műholdfelvételek a Planet Labs Inc. *Education and Research Program*-ja keretében kerültek a kutatáshoz biztosításra [21].

Irodalomjegyzék

- [1] Chang Huang és tsai. “Detecting, extracting, and monitoring surface water from space using optical sensors: A review”. *Reviews of Geophysics* 56.2 (2018), 333–360. old.
- [2] P.S. Frazier és K.J. Page. “Water Body Detection and Delineation with Landsat TM Data”. *Photogrammetric Engineering and Remote Sensing* 66 (2000. jan.), 1461–1467. old.
- [3] Stacy Maurer és Marvin Bauer. “Satellite remote sensing of wetlands”. *Wetlands Ecology and Management* 10 (2002. okt.), 381–402.
- [4] Tri Dev Acharya és tsai. “Identification of water bodies in a Landsat 8 OLI image using a J48 Decision Tree”. *Sensors* 16.7 (2016).
- [5] C. Domenikiotis, A. Loukas és N. R. Dalezios. “The use of NOAA/AVHRR satellite data for monitoring and assessment of forest fires and floods”. *Natural Hazards and Earth System Sciences* 3.1/2 (2003), 115–128.
- [6] Xingwang Fan és tsai. “Compositing the Minimum NDVI for Daily Water Surface Mapping”. *Remote Sensing* 12.4 (2020), 1–6. old. ISSN: 2072-4292.
- [7] Eric P Crist. “A TM Tasseled Cap equivalent transformation for reflectance factor data”. *Remote Sensing of Environment* 17.3 (1985), 301–306. old. ISSN: 0034-4257.
- [8] S. K. McFEETERS. “The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features”. *International Journal of Remote Sensing* 17.7 (1996), 1425–1432. old.
- [9] Hanqiu Xu. “Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery”. *International Journal of Remote Sensing* 27.14 (2006), 3025–3033. old.

- [10] S. L. K. Reddy és tsai. “A novel method for water and water canal extraction from Landsat-8 OLI imagery”. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 42.5 (2018), 323–328. old.
- [11] Gudina L. Feyisa és tsai. “Automated Water Extraction Index: A new technique for surface water mapping using Landsat imagery”. *Remote Sensing of Environment* 140 (2014), 23–35. old. ISSN: 0034-4257.
- [12] Adrian Fisher, Neil Flood és Tim Danaher. “Comparing Landsat water index methods for automated water classification in eastern Australia”. *Remote Sensing of Environment* 175 (2016), 167–182. old. ISSN: 0034-4257.
- [13] Junhua Li és Shusen Wang. “An Automatic Method for Mapping Inland Surface Waterbodies with Radarsat-2 Imagery”. *Int. J. Remote Sens.* 36.5 (2015), 1367–1384. ISSN: 0143-1161.
- [14] Dameng Yin és tsai. “Comparison of automatic thresholding methods for snow-cover mapping using Landsat TM imagery”. *International Journal of Remote Sensing* 34.19 (2013), 6529–6538. old.
- [15] Otsu és Nobuyuki. “A Threshold Selection Method from Gray-Level Histograms”. *IEEE Transactions on Systems, Man, and Cybernetics* 9.1 (1979), 62–66. old.
- [16] Judith M. S. Prewitt és Mortimer L. Mendelsohn. “The analysis of cell images”. *Annals of the New York Academy of Sciences* 128.3 (1966), 1035–1053. old.
- [17] Jui-Cheng Yen, Fu-Juay Chang és Shyang Chang. “A new criterion for automatic multilevel thresholding”. *IEEE Transactions on Image Processing* 4.3 (1995), 370–378. old.
- [18] Zhe Zhu és Curtis E. Woodcock. “Object-based cloud and cloud shadow detection in Landsat imagery”. *Remote Sensing of Environment* 118 (2012), 83–94. old.
- [19] Zhe Zhu, Shixiong Wang és Curtis Woodcock. “Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4-7, 8, and Sentinel 2 images”. *Remote Sensing of Environment* 159 (2015), 269–277. old.
- [20] Zhou Wang és tsai. “Image Quality Assessment: From Error Visibility to Structural Similarity”. *Trans. Img. Proc.* 13.4 (2004), 600–612. ISSN: 1057-7149.
- [21] Planet Team. *Planet Application Program Interface: In Space for Life on Earth*. <https://api.planet.com/>. San Francisco, CA. 2017.

Ábrák jegyzéke

2.1. Az elektromágneses sugárzás spektruma (nem méretarányos) és felhasználása a műholdas távérzékelésben.	5
2.2. Egy Sentinel-2-es tesztfelvétel vízmaszkjai különböző küszöbértékekkel .	10
2.3. A vizsgált küszöbölő algoritmusok eredményei egy Sentinel-2-es tesztadaton	14
2.4. Példa az s2cloudless és fmask algoritmussal készült felhőmaszkokra . . .	15
2.5. Két vízmaszk és a rajtuk talált különbségek	16
3.1. PlanetScope tesztadatok egy tavaszi (bal oldal), illetve egy nyári időszakban.	18
3.2. A Lechner Tudásközponttól kapott Sentinel-2-es tesztadatok egy felhős (bal oldal), illetve egy napos időszakban.	18
3.3. A folyamat flowchart-ja	19
3.4. 2018.07.04-ei Sentinel-2 felvétel és a hozzá készült fmask (felhő: rózsaszín, felhőárnyék: sárga, kék: víz)	20
3.5. Különböző vízindexek egy Sentinel-2-es tesztfelvételhez	21
3.6. A GRVI index egy PlanetScope-os tesztfelvételhez	22
3.7. 2018.04.20-ai (bal) és 2018.08.20-ai (jobb) 34UEU csempe különbség maszkja mNDWI index Yen küszöböléssel.	24
3.8. 2018.04.20-ai (bal) és 2018.08.20-ai (jobb) 34UEU csempe különbség maszkja NDWIg index Minimum küszöböléssel.	25
3.9. 2018.04.20-ai (bal) és 2018.08.20-ai (jobb) 34UEU csempe különbség maszkja NDWIg index Otsu küszöböléssel.	25
4.1. A GUI kezdőképernyője két Sentinel-2-es felvétel megjelenítésével	26
4.2. Beállítások a folyamat futtatása előtt	27
5.1. PlanetScope tesztfelvételének eredeti képe (bal oldal) és a vízmaszkja . .	31
5.2. Sentinel-2-es felhős tesztfelvétel pontossága: 95.70%	31

5.3. Sentinel-2-es felhőtlen tesztfelvétel pontossága: 98.72%	32
5.4. A Lechner Tudásközpont vízmaszkjainak és az NDWIg Minimum küszö- böléssel készült vízmaszkjainak különbségei	34
5.5. A Lechner Tudásközpont vízmaszkjainak és az mNDWI Yen küszöbölés- sel készült vízmaszkjainak különbségei	34

Táblázatok jegyzéke

2.1. Sentinel-2 műhold sávjai	6
2.2. A PlanetScope konstelláció PS2.SD szenzorának specifikációja	7
2.3. Az összes vizsgált index kiszámításának képletei.	9
3.1. Sentinel-2-es tesztfelvétel küszöbértékeinek vizsgálata	23
3.2. PlanetScope-os tesztfelvétel küszöbértékeinek vizsgálata	23
5.1. A Lechner Tudásközponttól kapott felvételekre végzett százalékos össze- hasonlítás eredményei	30
5.2. Eredmények összefoglalása	33

Forráskódjegyzék

4.1. Indexszámítás a PlanetScope felvételeire, saját módszer.	27
4.2. Minimum küszöbölő algoritmus implementációja az skimage package-ben.	28
4.3. Yen küszöbölő algoritmus implementációja az skimage package-ben. . . .	29