

TDK-dolgozat

Boga Zsombor

XXVII. reál- és humántudományi Erdélyi Tudományos Diákköri Konferencia (ETDK)

Informatika I.: elméleti módszerek, kísérletek, szimulációk szekció

Kolozsvár, 2024. május 15–18.

Tumor szegmentálás MRI képeken többdimenziós részcseke-raj alapú optimalizációval és klaszterezéssel

Szerző:

Boga Zsombor

Babeş–Bolyai Tudományegyetem, Matematika és Informatika Kar, Informatikai matematika
szak, III. év

Témavezetők:

dr. Kovács Péter, egyetemi docens,

Eötvös Loránd Tudományegyetem, Numerikus Analízis Tanszék, Informatikai Kar

Sándor Csanád, tanársegéd,

Babeş–Bolyai Tudományegyetem, Magyar Matematika és Informatika Intézet, Matematika és
Informatika Kar

Kivonat

A részecske-raj alapú optimalizációs (Particle Swarm Optimization – PSO) algoritmus széles körben alkalmazott az optimalizációs problémák megoldásában különböző területeken, ideértve az képfeldolgozást is.

A dolgozatban bemutatásra kerül egy újszerű, klaszterezésen alapuló szegmentációs algoritmus, amely a PSO algoritmus többdimenziós általánosítását használja kiindulópontként. Ellentétben a legtöbb hagyományos szegmentációs módszerrel, ez a megközelítés automatikusan meghatározza a különböző képeken az optimális klaszterszámot, figyelembe véve a csoportokon belüli és azok közötti hasonlóságokat. A módszer hatékonyan szegmentál agyi daganatokat mágneses rezonancia képalkotásból (Magnetic Resonance Imaging – MRI) származó szürkeárnyalatú képeken. A korszerű mélytanuló módszerekkel ellentétben, a dolgozatban bemutatott optimalizálási stratégia nem igényli nagyméretű, orvosok által előszegmentált tanítóhalmazok használatát. A hiperparaméterek beállításához elegendő néhány MRI kép felhasználása.

A dolgozat második felében a kifejlesztett eljárás tesztelésének eredményei, összehasonlítása korszerű gépi tanulási algoritmusokkal, és ezek ötvöztetésével elért eredmények kerülnek bemutatásra. A bemutatott módszerek teljesítménye egy valós MRI képadatbázison lett kiértékelve, melyet szakértő neuroradiológusok már előzetesen szegmentáltak.

Tartalomjegyzék

1. Bevezető	1
1.1. Problémafelvetés	1
1.2. Általános célkitűzések	2
1.3. A dolgozat felépítése	3
2. MRI képképzés és szegmentálás	4
2.1. MRI képek és jellemzőik	4
2.2. A felhasznált adathalmaz	6
2.3. Automatikus képszegmentálás jelentősége	8
2.4. Korábbi megközelítések és problémák	9
3. A felhasznált metaheurisztikus algoritmus	12
3.1. Részecske-raj intelligencia optimalizációs (PSO) algoritmusok	12
3.2. Többdimenziós részecske raj optimalizációs (MDPSO) algoritmus	15
4. Agytumor szegmentálás	18
4.1. Motiváció és kutatási célkitűzések	18
4.2. Az MDPSO algoritmus kiterjesztése képszegmentálásra	18
4.3. Az javasolt metrikák a büntető attribútumokra	21
4.4. Képfeldolgozási lépések	22
4.5. Utófeldolgozás	25
4.6. Tanítás és validálás	27
4.7. Eredmények, numerikus szimulációk	28
5. Összegzés	31

1. Bevezető

1.1. Problémafelvetés

A WHO adatai alapján egyre több agy és a központi idegrendszert érintő daganatos megbetegedést diagnosztizálnak világszerte. 2019-ben összesen 347.992 új agydaganatos esetet és 246.253 ahhoz kapcsolódó halálesetet regisztráltak [1]. A valós esetszám valószínűleg ennél is magasabb, mivel az egyes országok fejlettségi szintje jelentősen befolyásolja a statisztikákat. Például, a legmagasabb korhoz igazított előfordulási és halálozási arányokat a magas jövedelmű régiókban (Európa és Amerika) találták, míg a legalacsonyabbakat a szegényebb (például afrikai) régióban figyelték meg. Európában 2019-ben közel 7 esetet diagnosztizáltak 100.000 főre vetítve, de például Dániában ez az érték meghaladta a 17-et. Ezzel szemben, Afrikában az éves új esetek diagnosztikájának rátája 2 alatti volt.

Az automatizált agydaganat szegmentálásnak kiemelkedő szerepe van a korai fázisú diagnózisban, mivel ezáltal sokkal több embert lehet megvizsgálni. Ez a megközelítés jelentősen hozzájárulhat a betegség korai diagnózisához és ezáltal a páciensek gyógyulási esélyeinek növeléséhez. A szegmentálás pontosításának fontosságát jelzi, hogy számos kutatócsoport próbált már hatékony módszereket kidolgozni erre a célra. Az agydaganatok szegmentálásának módszereit célzó fejlesztések terén, egy 2023-as tanulmány [2] szerint, 2015 és 2022 között összesen 5.086 felügyelt vagy nem felügyelt tanuláson alapuló és 2.511 mély tanuláson (Deep Learning – DL) alapuló publikáció jelent meg. Ezek a kutatások az agydaganatok pontosabb azonosítására és szegmentálására törekednek, hozzájárulva ezzel a betegség korai észleléséhez, a megfelelő kezelés kiválasztásához, és így a gyógyulási esélyek növeléséhez.

A pontos agydaganatok szegmentálása körüli intenzív kutatásokat indokolja az is, hogy számos kihívással kell szembenézni annak érdekében, hogy megbízható eredményeket kapjunk. Elsődleges kihívást jelent a különböző MRI eszközök által készített felvételek sokszínűsége [3]. Ez azt jelenti, hogy a különböző készülékekkel készített felvételek eltérőek lehetnek a felbontásban, zajszintben, kontrasztban és más technikai paraméterekben, ami nehezítheti a pontos és általános feltételek melletti szegmentálást. Emellett a tumorok is jelentős változatosságot mutathatnak méretben, elhelyezkedésben, morfológiában és a különböző szövettani tulajdonságokkal rendelkező daganatok eltérően viselkedhetnek az MRI felvételeken [2]. Mindezek a kihívások megkövetelik, hogy a szegmentálási módszerek

robosztusak és sokoldalúak legyenek, hogy tudjanak alkalmazkodni a változó körülményekhez és a daganatok sokféle jellemzőjéhez.

Ezek a kihívások arra utalhatnak, hogy a mélytanulási (Deep Learning – DL) algoritmusok ideális megoldások lehetnek. Azonban ez egy túlzottan egyoldalú megközelítés, mivel ezek a modellek jelentős mennyiségű orvosok által annotált adatot igényelnek, aminek előállítása idő- és költségigényes. Ez az adatigény a gyakorlatban korlátozhatja a DL algoritmusok alkalmazhatóságát. Az orvosok által, kézi szegmentációra fordított rengeteg időt betegellátásra is lehetne fordítani, így az erőforrások hatékonyabb felhasználása érdekében érdemes megfontolni más alternatív megoldásokat is. Emellett a jól teljesítő mélytanuló rendszerek nagy számítási kapacitást is igényelnek, amely nem minden intézmény számára elérhető. Például, a 2021-es RSNA-ASNR-MICCAI Brain Tumor Segmentation (BraTS) Challenge (BRATS2021) verseny győztese egy módosított U-Net mélytanulási modellt használt [4], amelyben a tanítható paraméterek száma 15,729 millió volt. Ugyanakkor ennek a modellnek a tanításához 1251 darab agyról készült szegmentált voxel-képet használtak fel.

1.2. Általános célkitűzések

A kutatás a helyi neuroradiológiai intézmények technológiai színvonalának és az ott dolgozó szakértők által szegmentált és címkézett adatok mennyiségének felmérésével kezdődött, mivel fontos szempont, hogy a fejlesztés során figyelembe vegyük a helyi intézményi adottságokat és erőforrásokat. A figyelem főként a sepsiszentgyörgyi "Dr. Fogolyán Kristóf" Megyei Sürgősségi Kórházra és a kolozsvári „Prof. Dr. Ion Chiricuță” Onkológiai Intézetre irányult. Az előbbi (2023 szeptemberében) még nem rendelkezett szemantikusan szegmentált, annotált adatokkal, míg az utóbbi intézményben (2023 októberéig) egy körülbelül 80 darabos, szakértők által szegmentált agyi MRI adatbázis volt elérhető.

A helyi orvosi képzőintézmények fejlettségi szintjének és adatbázisának megfigyelése, valamint a problémafelvetésben (1.1) felsorolt kihívások figyelembe vétele alapján célként tűztük ki egy olyan módszer fejlesztését, amely minél kevesebb (legfeljebb 80 darab) szakértők által annotált 3D-s szegmentált minta felhasználásával is hatékonyan és megbízhatóan működik, mindeközben alacsonyabb számítástechnikai kapacitást igényel. A továbbiakban a helyi adatokhoz való hozzáférés lehetőségét a Babeş-Bolyai Tudományegyetem és az említett onkológiai intézet között létrejött együttműködési megállapodás tenné lehetővé. A rendelkezésre álló adatok felhasználásával a távolabbi cél egy átfogó tumorszegmentáló

rendszer felépítése. A dolgozatban kifejlesztett szegmentáló módszer fő motivációja, hogy elősegítse a technológia széleskörű alkalmazását és megkönnyítse a klinikai környezetben való alkalmazását. Ez különösen fontos tekintettel a szakorvosok hiányára, valamint a fejlettebb és kevésbé fejlett régiók közötti jelentős statisztikai különbségekre az agydaganatok diagnosztizálása terén.

E célok megvalósítása nem csupán technológiai innovációt jelent, hanem hozzájárul a helyi és globális egészségügyi szolgáltatások fejlesztéséhez és a betegellátás minőségének javításához is.

1.3. A dolgozat felépítése

A dolgozat a tumorszegmentálás problémájával kapcsolatos kihívásokra egy innovatív módszert kínál, amely minimalizálja az előre szegmentált adatok mennyiségét és lehetővé teszi a kisebb teljesítményű számítási eszközök használatát. A módszer magába foglalja a képek hagyományos feldolgozási technikákkal történő előkészítését, melyek egy utófeldolgozási lépésben gépi tanuló módszerek segítségével tovább finomíthatóak.

A dolgozat következő fejezete (2. fejezet) a mágneses rezonancia képalkotás alapjait, a rendelkezésre álló adatbázis jellemzőit és a tumorszegmentálás terén elért eddigi eredményeket ismerteti. Ezután (3. fejezet) a rajintelligencia alapú algoritmusokat, különös tekintettel a részecske-raj optimalizációs algoritmust és annak többdimenziós változatát mutatja be. A 4. fejezet egy új képszegmentálási eljárást tárgyal, amely az előfeldolgozás során végzett hisztogramkiegyenlítést és morfológiai operációkat követően kiterjeszti és alkalmazza az MDPSO algoritmust a képek szegmentálására. Ugyanebben a fejezetben bemutatásra kerülnek az alkalmazott utófeldolgozási lépések, ahol a döntési fákon alapuló véletlen erdő osztályozó (Random Forest Classifier– RFC) algoritmus kerül felhasználásra, az elért eredmények, majd azok összehasonlítása más eljárásokéval. Az utolsó rész (5. fejezet) röviden összefoglalja a kutatás főbb eredményeit, és felvázolja a továbbfejlesztési lehetőségeket is.

2. MRI képképzés és szegmentálás

2.1. MRI képek és jellemzőik

Az MRI az egyik leghatásosabb és leggyakrabban alkalmazott orvosi képképzési és diagnosztikai eszköz, különösen az agyi és egyéb lágy szövetek vizsgálata során. Az MRI eljárás lehetővé teszi a különböző szövetek közötti kontrasztok finom megkülönböztetését anélkül, hogy ionizáló sugárzást kellene használni, így a páciensek számára biztonságosabb alternatívát kínál, ellentétben például a CT (komputertomográfia) vizsgálattal, ami röntgensugarakat használ a képek előállításához [5].

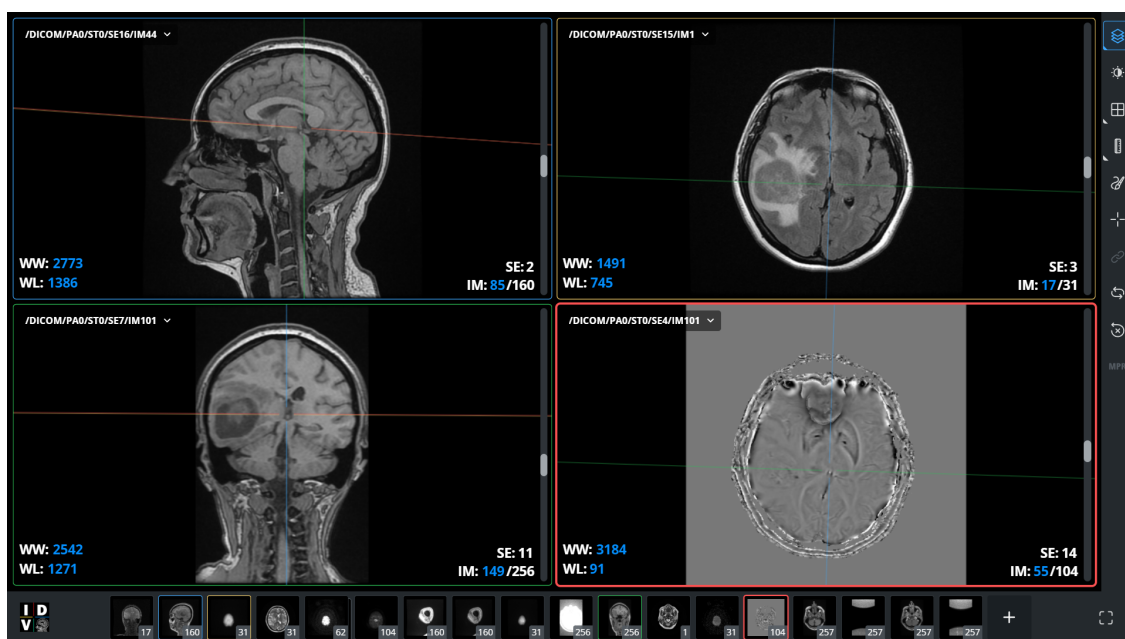
Kutatásom során lehetőségem nyílt betekinteni a sepsiszentgyörgyi "Dr. Fogolyán Kristóf" Megyei Sürgősségi Kórház adatbázisába, ahol részletesebb felvilágosítást is kaptam a képek elkészítéséről és megvizsgálásának folyamatáról. Az adatvédelmi előírásoknak megfelelően, további elemzések céljából, öt páciens anonimizált agyi MRI-vizsgálatának felvételét kaptam meg.

Az MRI alapelve az erős mágneses mező és a rádiófrekvenciás hullámok alkalmazásán alapul, amelyek lehetővé teszik a test belső szerkezetének nem invazív vizsgálatát. Az eljárás során a páciensek testét erős mágneses mezőbe helyezik, ami hatással van a hidrogén protonokra, amelyek a rádiófrekvenciás impulzus hatására rezonálni kezdenek, majd visszatérnek eredeti állapotukba, miközben elektromágneses jeleket bocsátanak ki. Az MRI képképzés során ezeket a jeleket dolgozzák fel és jelenítik meg [6].

Egy MRI vizsgálat során több felvétel készül, amelyeket ha egymásra igazítunk, akkor úgy is tekinthetünk, mint egy képnek a különböző csatornáira. Azért fontos a különböző protokollok alkalmazása, mert ezek a testünk különböző részeit, szöveteit, esetenként pedig azok elváltozásait emelik ki [5, 2, 6]. A leggyakrabban alkalmazott protokollok, MRI modalitások a következők:

- **T1-súlyozott képek** – ezek a képek kiválóan alkalmasak zsírszövetek és friss vérzések megjelenítésére.
- **T1CE** (Contrast-enhanced weighted-T1 – T1-súlyozott, kontrasztanyaggal erősített képek) – T1-súlyozott képképzéshez hasonló, de intravénás kontrasztanyagot adnak a páciensnek, amely fokozza bizonyos típusú daganatok és gyulladásos folyamatok megjelenítését.

- **T2-súlyozott képek** – a T2 súlyozás a víztartalmú szöveteket, mint az ödéma vagy más gyulladt területek, világosabban ábrázolja, ami segít a patológiai területek könnyebb azonosításában.
- **FLAIR** (Fluid Attenuated Inversion Recovery) – hasonló a T2-súlyozotthoz, de egy inverziós impulzust alkalmaznak a folyadékok, különösen az agyvíz (központi idegrendszer folyadék), jeleinek kiszűrésére, ami javítja az agyi ödémák és egyéb kóros elváltozások detektálását.



1. ábra. A sepsiszentgyörgyi kórház egyik páciensének MRI felvételének vizualizációja a IMAIOS DICOM Viewer platform segítségével. A felnagyított felvételek sorfolytonosan a következő síkok mentén vannak megjelenítve: szagittális, axiális, koronális és ismét axiális. A képkalkotási protokollok között találunk kontrasztanyag nélküli T1-súlyozott, FLAIR, kontrasztanyaggal erősített T1-súlyozott, valamint fáziskontrasztos képeket.

Az általam vizsgált helyi adathalmaz egy MRI vizsgálatról 18 különböző felvételsorozatot tartalmazott, ahogy az 1. ábrán is láthatjuk, különböző protokollok szerint és különböző síkok mentén:

- **szagittális** sík – a testet balról jobbra (vagy fordítva) kettéosztva ábrázolja, gyakran használják az agy elülső és hátulsó részeinek vizsgálatára.
- **axiális** (vagy transzverzális) sík – a testet fentről lefelé kettéosztva ábrázolja, jól mutatja az agy felső részét, és segítenek a struktúrák horizontális elrendezésének megértésében.

- **koronális** (vagy frontális) sík – a testet előlről hátrafelé kettéosztva ábrázolja, gyakran használják az agy függőleges elrendezésének vizsgálatára.

Automatikus feldolgozás esetén ezek a keresztmetszetek csak akkor játszanak szerepet, ha csak 2D-s információkat használunk fel. Ezek a megközelítések általában kevésbé számításigényesek [7]. Azonban ha már 3D-s információkkal dolgozunk [4], akkor ezek az irányok kevésbé lesznek relevánsak.

Világszinten a legtöbb intézmény még nem tért át az automatikus képszegmentálás alkalmazására mivel azok megbízhatósága nem mindig kielégítő, noha ezzel nagyon sok időt és orvosi energiabefektetést meg lehetne spórolni. Ebből adódóan a különböző intézmények között sem alakult ki egy egységes protokoll, hogy milyen előírások szerint valósítsák meg ezeknek a felvételeknek az elkészítését. A felvételek változatosságát főként az okozza, hogy a különböző intézmények különböző cégek által készített eszközöket használtak, amelyek gyakran nagyon eltérő eredményeket adnak [8].

Az 1. ábra MRI képei a színes helyzetjelző vonalak segítségével mutatják a különböző síkokban megjelenített felvételek kapcsolatát, ami lehetővé teszi a radiológusok számára, hogy pontosan lokalizálják az anatómiai struktúrákat. Mivel a vonalak nem feltétlenül merőlegesek egymásra, és a különböző síkok képszeleteinek száma eltérő, ez nehezíti a képinformációk összehasonlítását és az adott pozíciókhoz tartozó információk kinyerését. Ezért az automatikus szegmentáló algoritmusok fejlesztése során figyelembe kell venni a modalitást is, azaz az MRI felvételek elkészítésének módszertanát. Ebből adódóan egy nemzetközi verseny előfeldolgozott és előkészített adatbázisát használtam fel a kutatásom során.

2.2. A felhasznált adathalmaz

Az eredmények összehasonlíthatóságát és reprodukálhatóságát szem előtt tartva egy nemzetközi verseny, a BraTS 2021 publikusan elérhető adathalmaz felvételeit használtam a dolgozatban bemutatott módszer tesztelésére. A tanítási adathalmazt nem tették nyilvánossá, hogy a versenyben résztvevők számára a tesztelés teljes mértékben ismeretlen adatokon történjen. Mivel az általam fejlesztett módszer amúgy sem igényel sok tanítási adatot így az összesen 1251 darab felvételtől a 80 darab tanításra felhasznált felvételen kívül bőven maradt képi adat a validációra is.

Az adathalmaz létrehozásához több intézményből származó felvételeket gyűjtöttek össze, így minimalizálva az egyes képalkotó eszközök specifikus jellemzőinek hatását [9, 10, 11].

A felvételekről eltávolították az agyszöveteken kívüli részeket, mint például a koponya vagy más fejen lévő szervek. Ez egyrészt megkönnyíti a versenyzők feladatát is, hogy jobban összpontosíthassanak a tumor szegmentálásának problémájára. Ugyanakkor ennek az eljárásnak etikai vonzatai is vannak, mivel teljes fejre kiterjedő MRI-felvétel alapján elég jól rekonstruálható az emberi arc [12], ami etikai és jogi kérdéseket vet fel a személyes adatok védelmének kontextusában.

Általában ezeket a felvételeket a kórházak DICOM (Digital Imaging and Communications in Medicine – digitális képalkotás és kommunikáció az orvostudományban) típusú fileokban mentik le (lásd 1. ábra), azonban a felhasznált adathalmazban egységesen $240 \times 240 \times 155$ -voxelből álló 3D-s NIfTI (Neuroimaging Informatics Technology Initiative) képekként vannak lementve a felvételek. Mindkét formátumot széles körben használják az orvosi képalkotásban. Míg a DICOM formátum sokkal bővebb metaadat-készletet tartalmaz, amely részletezi a páciens információit, a képalkotó eljárás paramétereit, berendezés adatait és sok egyebet, addig a NIfTI formátum egyszerűsített metaadat-struktúrát használ, amely főként a képek dimenziójára, orientációjára és voxel méretére koncentrál, kevesebb hangsúlyt fektetve a képalkotó berendezés specifikus adataira vagy a páciens információira. Egy adott anatómiai sablonhoz négy különböző képalkotási protokoll szerint készített MRI felvétel tartozik, nevezetesen T1, T1CE, T2 és FLAIR képtípusok (lásd az 4.7. fejezet, 1). táblázat) Az adatok azonos felbontásra vannak interpolálva, ahol minden voxel 1 mm^3 térfogatú információt tárol. A felvételek esetében nem szükséges a test különböző irányokból történő vizsgálata, mert ezeket azzal szabályozhatjuk, hogy melyik indexet rögzítve járjuk be a 3D-s adattömb egy 2D-s szeletét.

Az annotált adatok a következő osztályokba vannak besorolva:

- **0-ás címke:** azokat a területeket jelöli, amelyek nem tartalmaznak daganatos szövetet, illetve a háttér képpontjai is ezzel a címkével vannak ellátva.
- **1-es címke:** a nekrotizált daganatmag (NCR – Necrotic Tumor Core) címkéje, amely a tumor elhalt, nekrozissal érintett központi részét jelöli.
- **2-es címke:** a peritumorális ödémás/támadott szövet (ED – Peritumoral Edematous/Invaded Tissue) címkéje, amely a tumor által érintett, gyulladt vagy daganat által megtámadott szöveteket jelöli.
- **4-es címke:** az erősödő daganat (ET – Enhancing Tumo) címkéje, amely azokat a

daganatrészeket jelöli, amelyek kontrasztanyag felhalmozódást mutatnak.

A címkézéseket kombinált automatizált és manuális módszerekkel készítették, amelyek több iterációban lettek finomítva. Szakértők és előző versenyeken jól teljesítő algoritmusok segítségével minimalizálták az megfigyelők közötti hibákat. Ez a hibrid megközelítés garantálja az adatok minőségét és klinikai relevanciáját [9].

Kutatásom során csak egy bináris osztályozást végzek el: megjelölöm pozitív pixelenként azokat a részeket, amelyeket tumor által érintettek és negatívnak az egészséges szöveteket reprezentáló képpontokat (lásd [13]). Ez azonban nem zárja ki annak a lehetőségét, hogy a módszerem eredményeit összehasonlítsam olyan tanulmányok eredményeivel, amelyek az alrégiókat is külön csoportba sorolják, mivel a ezekben is kiértékelik, hogy milyen pontossággal sikerült detektálni az egész tumort (lásd [14, 4]).

2.3. Automatikus képszegmentálás jelentősége

Az automatikus képszegmentálás jelentősége vitathatatlan a modern orvosi képalkotásban, különösen a tumorszegmentálás területén. A világszerte tapasztalható orvoshiány különösen érzékelhető a rákszűrés és -kezelés területén. Az automatizált képszegmentáló rendszerek bevezetése csökkentheti az orvosi személyzet terhelését, lehetővé téve számukra, hogy több időt szenteljenek a betegek közvetlen ellátására. Továbbá, mivel a kézi szegmentálás rendkívül időigényes [9] folyamat, az automatizálás jelentősen növelheti a diagnosztikai folyamat hatékonyságát.

Egy 2020-as tanulmányban [15] kimutatták, hogy már egy olyan rendszer is, ami félig automatizált, azaz ha a számítógép által adott eredményeket egy szakorvos átvizsgálja, jelentősen felgyorsítja a munkafolyamatot (átlagosan 167 másodpercre) a manuális MRI-szegmentáláshoz képest (átlagosan 479 másodperc), ugyanakkor kevésbé igényel fizikai erőfeszítést. A pontosság tekintetében a két módszer hasonló eredményeket ért el.

Ugyanakkor a manuális képszegmentálás során előforduló megfigyelők közötti és megfigyelőn belüli változatosság (interobserver and intraobserver variability) csökkentése jelentős előrelépést jelentene a diagnosztika pontossága és következetessége szempontjából. Egy másik tanulmány [16] eredményei szerint az alkalmazott automatikus szegmentációs módszer (nnU-Net alapú neurális háló) képes volt arra, hogy csökkentse az inter-observer változatosságot a két radiológus között. Ez azt jelenti, hogy az algoritmus által végzett automatikus szegmentáció eredményeinek felhasználása után kevesebb eltérés volt észlelhető

a két orvos által manuálisan végzett szegmentáció eredményei között. Az automatikus szegmentáció ebben az értelemben segít a képszegmentálási folyamat egységesítésében és a konzisztens eredmények elérésében is.

Egy másik orvosi tanulmány [17] rámutat, hogy a gliómák növekedési mintái lehetővé teszik a korai felismerést. Nyilvánvalóan az időben történő tumorfelismerés és a korai beavatkozás kulcsfontosságú a rákbetegségek kezelésében, és jelentős mértékben javíthatja a betegek túlélési esélyeit. Az automatizált képszegmentáló rendszerek használata ebben is segíthet, mivel általa több páciens lehet hatékonyabban megvizsgálni, és így gyakoribbá válhatnak a szűrővizsgálatok is.

2.4. Korábbi megközelítések és problémák

A képszegmentáció alapvető feladat a számítógépes látás területén, amely a képek szerkezetileg különböző részeinek elkülönítését valósítja meg. Ez a feladat megvalósítható felügyelt és nem felügyelt módszerekkel is. A módszerválasztást több tényező is befolyásolja, mint például a rendelkezésre álló hardver, a kívánt vagy szükséges szegmentálási pontosság, az előre szegmentált adatok (ground truth data) megléte, a kép színmélysége és csatornáinak száma (színes vagy szürkeárnyalatos) [5].

Egy 2023-as összegző tanulmányban [2] részletes elemzést végeztek az agydaganatok MRI képeken történő szegmentálásának legkorszerűbb módszereiről, kiemelve az egyes technikák teljesítményét, előnyeit és hátrányait. A cikk alapos áttekintést nyújt a terület legígéretesebb technológiáiról, elemezve a különböző szegmentációs eljárásokat, a hozzájuk kapcsolódó adatbázisokat, az orvosi képalkotás kulcsfogalmait, de még az agydaganatok leggyakoribb típusait is. A szegmentáció az orvosi képalkotásban, az agydaganatok vizsgálata során is, számos kihívással néz szembe, mint például a zaj, a szöveti változékonyság, valamint a daganatos és nem daganatos szövetek közötti határok elmosódása. Ráadásul a különböző protokollok szerint készített felvételeken azonos koordinátájú voxelek nem mindig ugyanarról a pontos helyről tárolnak információt, hanem valahonnan a közeli, szomszédos területről [14].

Az automatikus agydaganat-szegmentálás jelentőségét hangsúlyozza a BraTS nemzetközi verseny éves megrendezése is, ahol a résztvevőket arra ösztönzik, hogy évről évre a lehető legpontosabb eredményeket ériék el az orvosi képfeldolgozás területén. Az elmúlt két évben bevezetett nehezítések, mint például az a tesztorozat, amely során hiányzott egy-egy képtípus (T1, T1CE, T2, FLAIR) a négy közül, tovább fokozták a verseny kihívásait. 2023-ban a

kihívás különösen a gyermekkori agydaganatokra összpontosított, amelyek jellegzetességei és megjelenési formái jelentősen eltérhetnek a felnőttkori daganatoktól [18].

A 2021-es kihívás célkitűzései közvetlenül kapcsolódnak a dolgozatban bemutatott eljárások céljaihoz, mint például a minél pontosabb szegmentálás elérése általános adathalmazon. A legjobb eredményeket egy módosított nnU-Net architektúra érte el, amely csoportos normalizálást és ún. attention rétegeket alkalmazott a dekóderben, jelentősen javítva ezzel a szegmentációs pontosságot [4]. A verseny adathalmazán a modell legeredményesebb konfigurációja a teljes tumor szegmentálásában 0,9277 Dice-értéket ért el. Felületesen nézve az eredményeket bizakodóak lehetünk, de például az Európai Parlament és Tanács orvostechikai eszközökről szóló rendelete szerint még nem alkalmazható teljesen automatizált diagnózisra. Ennél is nagyobb probléma, hogy még ezekhez az eredményekhez is hatalmas mennyiségű adatra volt szükség, és a klinikai környezetbe való bevezetés is nagy valószínűséggel további beruházásokat követel meg a felhasználó intézményektől, mivel egy átlagos számítógép nem rendelkezik megfelelő grafikus processzor egységgel (GPU – Graphics Processing Unit) a kifejlesztett technika alkalmazásához. Az említett nnU-Net modell fejlesztéséhez egy NVIDIA RTX 3090 GPU-val, 24 GB VRAM-mal (Video Random Access Memory) rendelkező számítógépet használtak fel. Ez a konfiguráció jelzi, hogy a hasonló méretű és komplexitású architektúrák erőteljes hardveres támogatás szükséges, amely képes kezelni a nagy memória- és számítási igényt. Ugyanakkor nnU-Net használatakor az ugró kapcsolatok (skip connectionok) révén továbbított jellemzőtérképek jelentős memóriát igényelhetnek, a dekóder szakaszban a hálózat tanítása és inferencia futtatása közben is.

A hagyományos képfeldolgozási módszerek nemcsak akkor értékesek, amikor a címkézett adatokból nagyon kevés áll a rendelkezésünkre, hanem jelentősen hozzájárulnak az új technológiák teljesítményének a javításához és a memória- illetve számítási költségek csökkentéséhez is. Például, egy mélytanulás alapú, ún. attention rétegeket alkalmazó felügyelt szegmentációs módszer [19] megköveteli a célterület (ROI – region of interest) meghatározását, amit a tanulmány hisztogram Z-normalizálással, különböző MRI-képtípusokon alkalmazott küszöbérték-technikák és morfológiai operátorok kombinációjával ér el.

Egy marosvásárhelyi kutatócsoport is foglalkozott a tumorszegmentálás problémájával, ahol főként a RF algoritmus kapta a főszerepet, de számos variációt teszteltek, mint például a Gábor-szűrőket és az Atlasz módszereket[20, 13, 14, 21]. A BraTS 2019-es adathalmazon különböző típusú RF algoritmusokkal és adatmennyiségekkel végzett kísérleteik során 0,8232

és 0,8516 közötti Dice-értéket értek el [14].

A nem felügyelt módszerekre összpontosítva számos képszegmentáló algoritmus létezik, beleértve a küszöbérték-alapú megközelítéseket [22, 23] (pl. Otsu módszer, K-means klaszterezési algoritmus) vagy a metaheurisztika-alapú optimális küszöbértéket kereső algoritmusokat [24, 25]. Általánosságban ezek a módszerek nem alkalmasak objektumfelismerésre, mert kizárólag a pixelintenzitásra támaszkodnak, ami gyakran a képpontok szórt, töredezett klaszterezéséhez vezet, amelyek nem, vagy alig feleltethetők meg objektumoknak. Egy másik hátránya az említett módszereknek, hogy ezek általában fix számú szegmensek vagy csoportok meghatározására vannak tervezve, amelyek nem feltétlenül alkalmazkodnak könnyen a különböző képekhez. A szegmensek számára tulajdonképpen úgy is tekinthetünk mint hiperparaméternek, amit a legtöbb algoritmus esetében előre, manuálisan kell beállítani [26]. A klaszterszám helytelen megállapítása az objektumok elégtelen elkülönítéséhez vezethet. Így például a túl alacsony klaszterszám az objektumok összeolvadásához, a túl magas pedig a kép felaprózódásához vezet.

3. A felhasznált metaheurisztikus algoritmus

3.1. Részecske-raj intelligencia optimalizációs (PSO) algoritmusok

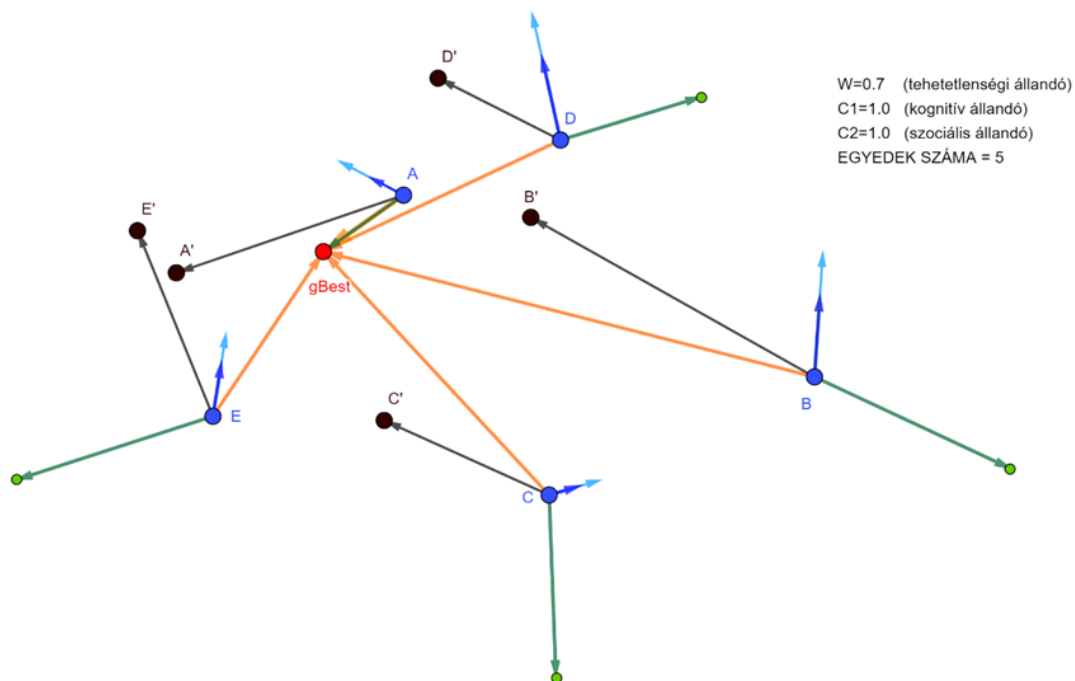
A PSO algoritmust először James Kennedy és Russell Eberhart mutatta be 1995-ben [27]. Az elmúlt évtizedekben a PSO számos változatát kifejlesztették, emiatt már algoritmus-családról is beszélhetünk [28, 29]. Ahogy a nevük is sugallja, a természetben előforduló rajok, csoportok mozgását, stratégiáit imitálva egy probléma optimális megoldásának megkeresésére hivatottak. Általában ezek az algoritmusok feladattól független, metaheurisztikán alapuló stratégiát követnek. Ebből kifolyólag ugyanaz az algoritmus (esetleg különböző hiperparaméterekkel) több, különböző optimalizációs feladatban is alkalmazható, viszont nem garantált, hogy a ténylegesen optimális eredményt találja meg. A hiperparaméterek módosításával az algoritmus alkalmazkodik a különböző kihívásokhoz, de így sem garantálja a globálisan optimális megoldást minden esetben. Egy másik előnyük a determinisztikus algoritmusokkal szemben, hogy sokdimenziós keresési térben is képesek keresni, és viszonylag rövid idő alatt elfogadható, szuboptimális eredményt adnak [30].

A rajintelligenciát a kollektív döntéshozatal, az egyedek közti kommunikáció, a hatékony önszerveződés, illetve a munkamegosztás jellemzi. A raj tagjainak együttműködése olyan intelligenciát eredményez, amely túlmutat a tagok egyéni képességein. A rajban a feladatok lehetnek egységesen kiosztva, viszont különböző egyedek különböző szerepköröket is betölthetnek. Mindkét eset a globális cél hatékony elérését szolgálja. A standard PSO-ban [27] egyenértékű részecskék (egyedek) munkáját figyelhetjük meg, melyekre mint megoldásjelöltekre tekinthetünk. Egy egyed rendelkezik egy aktuális pozícióval a keresési térben (a probléma egy lehetséges megoldása), sebességgel, illetve legjobb pozícióval (az eddigi iterációk során megtalált leghatékonyabb megoldás/állapot). Ezekre az értékekre úgy tekinthetünk, mint vektoriális mennyiségek, melyek dimenziója megegyezik a keresési tér dimenziójával.

Az aktuális pozíciók általában véletlenszerű (egyenletes eloszlású) pontokkal vannak inicializálva a keresési térből, az algoritmus lépései során ezeket az előző pozíciók, illetve aktuális sebességek fogják meghatározni. A részecskék sebessége kezdetben a pozíciókhoz hasonlóan véletlenszerű, majd a későbbi lépések során a következő három komponens súlyozott összegével számolunk: a tehetetlenségből származó mozgás, a szociális befolyás és a kognitív összetevő. Az összegzésben megjelenő súlyok az algoritmus futása közben legtöbbször nem

változnak, éppen ezért megválasztásuknak nagy szerepe van az algoritmus hatékonyságában és gyorsaságában. A részecskék megjegyzik az eddigi iterációk során lokálisan megtalált, legoptimálisabb (általában minimális) fitnessértékű pozíciójukat. Az algoritmus futása közben tárolunk egy globális eredményt is, az eddigi (fitnessérték alapján kiválasztott) globális legjobb megoldás, amely az aktuális iterációban, a valamilyen szinten (a hiperparaméterektől függően) meghatározza minden egyed elmozdulását.

Az egyedekhez tartozó fitnessértéket a fitnessfüggvény segítségével határozhatjuk meg. Mivel ez a függvény a keresési tér egy pontjának, azaz egy lehetséges megoldás optimalitását méri, ezért általában egy többváltozós, valós értékű függvényről beszélhetünk. Az értelmezési tartományának szükségszerűen magába kell foglalnia a keresési teret, ami igazából az összes lehetséges megoldást jelenti.



2. ábra. A hagyományos PSO algoritmus működési elvének szemléltetése egy 2D-s keresési térben

Egy kétdimenziós keresési térben a következő módon képzelhetjük el az egyedek mozgását (2. ábra): A, B, C, D, E pontok jelölik egy iterációban a megoldásjelöltek elhelyezkedését a keresési térben. Ebben az iterációban az A egyed saját legjobb megoldása az eredeti problémára, egyben az eddigi megtalált megoldások közül is a legjobb ($\hat{y}$). Az ábrán megfigyelhető a többi egyed eddigi legjobb pozíciója is (zöld pontok). Az iteráció elején az algoritmus aktualizálja az egyedekre vonatkozó lokális és globális legjobb pozíciókat. A részecskék sebességvektora

(fekete vektorok) három, az aktuális pozícióból kiinduló vektor összegéből fog állni. Az első az előző iterációban vett sebességvektor (világoskék vektor), beszorozva az inercia állandóval. A második minden egyednek az eddigi legjobb pozíciója felé mutató vektora (zöld vektor), súlyozva a kognitív állandóval. A harmadik a globális legjobb megoldás felé mutató vektor (narancssárga vektor) beszorozva a szociális állandóval. A sebességvektort hozzáadva az előző lépésben kiszámolt pozíciójához, megkapjuk az új pozíciót (A' , B' , C' , D' , E' pontok). Az algoritmus lényege, hogy iteratívan próbálja javítani a megoldást, és addig ismétli ezt a folyamatot, amíg valamilyen leállási feltétel nem teljesül. A leállási feltétel lehet az, hogy elértük az optimálisnak tekinthető értéket, vagy hogy a részecskék nagy része egy megoldás közelébe konvergált. Gyakran előfordul, hogy egyszerűen csak egy rögzített iterációs számot határoznak meg, amely után a folyamat befejeződik, függetlenül attól, hogy mennyire közel állunk az optimális megoldáshoz.

Ezek alapján egy egyed sebességvektora a

$$\mathbf{v}_i = w \cdot \mathbf{v}_{i-1} + c_1 \cdot r_1 \cdot (\tilde{\mathbf{y}}_i - \mathbf{x}_i) + c_2 \cdot r_2 \cdot (\hat{\mathbf{y}}_i - \mathbf{x}_i) \quad (1)$$

képlet segítségével aktualizálható.

Jelölések:

- $\mathbf{v}_i$ – az egyed sebességvektora az i . iterációban;
- w – a tehetetlenségi állandó, amely meghatározza az előző iterációban számított sebességvektor súlyát az új sebességvektor kiszámításakor;
- c_1 – a kognitív tényező súlya, amely meghatározza az egyén saját legjobb pozíciójának hatását a sebesség változására;
- c_2 – szociális tényező súlya, amely meghatározza a globális legjobb pozíció hatását a sebesség változására;
- $\mathbf{x}_i$ – az egyed i . iterációban vett aktuális pozíciója, az optimalizálni kívánt probléma egy lehetséges megoldása;
- $\tilde{\mathbf{y}}_i$ – az egyednek az i . iterációig megtalált saját legjobb pozíciója;
- $\hat{\mathbf{y}}_i$ – az i . iterációig megtalált, globálisan legjobb pozíció.

- r_1 és r_2 – egyenletes eloszlással vett véletlen számok a $[0, 1]$ intervallumról, amelyek a részecskék mozgásának nem determinisztikusságát biztosítják.

3.2. Többdimenziós részecske raj optimalizációs (MDPSO) algoritmus

A PSO többdimenziós változatát, az MDPSO algoritmust 2011-ben mutatták be [31]. Az algoritmus megőröklí a hagyományos PSO előnyös tulajdonságait és működési elveit, miközben jelentősen bővíti az algoritmus képességeit, hogy megfeleljen a gyakorlati problémák speciális elvárásainak, ahol gyakran a probléma optimális dimenziója is ismeretlen. A PSO logikát követve automatikusan navigál a különböző dimenziójú megoldási tereken, anélkül, hogy elvégezze a teljes optimalizációt minden lehetséges dimenzióra vonatkozóan. Ezeket az algoritmusokat széles körben alkalmazzák különböző területeken, többek között képfeldolgozásban [32], paraméterhangolásban és gépi tanulásban [33], hálózatoptimalizálásban [34], jeltömörítésben [35, 36] és számos más területen.

Működési elv

Az MDPSO algoritmus úgy lett kialakítva, hogy képes legyen kezelni olyan problémákat is, amelyek dimenzionalitása nem rögzített, hanem egy meghatározott tartományban változik. Ebben a módszerben a dimenziók egy intervallumon belül azonosíthatók, $d_{\min} \leq d \leq d_{\max}$, lehetővé téve az algoritmus számára, hogy dinamikusan alkalmazkodjon a különböző dimenziós kihívásokhoz. Az MDPSO minden részecskéje két folyamaton megy keresztül: az első a hagyományos PSO-nak a pozicionális optimumkeresése, ami a szokásos sebességfrissítéseket és a helyzetbeli elmozdulásokat jelenti egy adott (d indexű) dimenziós térben. A második folyamat a dimenzionális PSO, ami lehetővé teszi a részecske számára, hogy a különböző dimenziók között navigáljon.

Az optimalizáció során mindegyik részecske eltárolja az adott dimenzióban az utolsó pozícióját ($\mathbf{x}^d$), sebességét ($\mathbf{v}^d$) és lokális legjobb helyzetét ($\tilde{\mathbf{y}}^d$), így amikor később visszatér ugyanabba a dimenzióba, képes folytatni a keresést az ott korábban eltárolt információk alapján. A dimenzionális PSO folyamat során a részecske átmehet egy másik indexű dimenzióba, ahol szintén emlékszik a korábbi pozíciós állapotára, és ebben a dimenzióban is folytatja az optimalizálást. A raj másrésről nyomon követi minden dimenzióban a legjobb részecskék helyzetvektorát ($\hat{\mathbf{y}}^d$), amelyek mindegyike jelzi a legjobb globális pozíciót, amit eddig a d dimenzióban felfedezett, és ezeket használják a sebességfrissítési egyenletben az adott

dimenzióban. Hasonlóképpen, a dimenzionális PSO folyamat minden részecskéje használja a saját lokális legjobb dimenzióját ($\tilde{d}_k$), ahol eddig a legjobb fitnessértéket érte el. Végül a raj nyomon követi a legjobb globális dimenziót ($\hat{d}$) az összes lokális legjobb dimenzió között. Az utolsó iterációban pedig a leghatékosabbnak megítélt dimenzióban ($\hat{d}$) lévő legjobb részecske ($\mathbf{x}_{gBest\hat{d}}^{\hat{d}}$) képviseli az optimális megoldást.

Az algoritmus

Az bemutatott pszeudókodban (1. Algoritmus) az alábbi elnevezési konvenciókat alkalmazzuk:

- N jelöli a keresési folyamat során végrehajtott iterációk számát.
- w a sebességvektor tehetetlenségi súlyát jelenti, amely az iterációk során lineárisan csökken 0.9-ről 0.4-re. Ez a lineáris csökkenés hozzájárul az algoritmus konvergenciájának elősegítéséhez [37, 35].
- $[d_{\min}, \dots, d_{\max}]$ az optimalizáció során vizsgált dimenziók alsó és felső határát jelöli, melyek között a megoldást keressük.
- $\mathbf{x}$ az egyed aktuális pozícióját jelölő vektor, amely a probléma egy lehetséges megoldását reprezentálja, $\mathbf{y}$ a részecskék által elért legjobb pozíciót jelöli (amit az iterációk során mért minimális fitnessérték jellemez), $\mathbf{v}$ pedig a sebesség- vagy mozgásvektort jelöli a továbbiakban bemutatott konvenciók szerint.
- A hullámvonal ($\sim$) az egyes részecskék által elért lokális legjobb megoldást jelöli.
- A kalap jel ($\hat{\cdot}$) az összes részecske között elért globálisan legjobb megoldást jelöli.
- A $gBest$ változóban a legjobb eredményt elérő részecske sorszáma kerül eltárolásra.
- d a problémater dimenzionalitását jelöli. Ha d felső indexként jelenik meg, az azt jelzi, hogy a hivatkozott eredmény, legjobb megoldás vagy sebesség kifejezetten ezzel az indexsel jelölt dimenzióra vonatkozik.
- S a rajban lévő részecskék teljes számát, k pedig az aktuális részecske sorszámát jelöli.
- $\rho(\cdot, \cdot)$ egy vektorértékű valós veszteségfunkcionál, amely az adott probléma megoldásának optimalizálását végzi a változók figyelembevételével, miközben számításba veszi a keresési tér dimenzionalitását is (2. Algoritmus).

1. Algoritmus

```
1: function MDPSO
2:    $w_0 := 0.9, w_N := 0.4$  ▷ Kezdeti és végső inercia súlyok beállítása
3:    $c_1 := 1.5, c_2 := 2$  ▷ Kognitív és szociális súlyok inicializálása
4:   for all  $d \leftarrow d_{min}, \dots, d_{max}$  do
5:     for all  $k \leftarrow 1, \dots, S$  do
6:       Randomly initialize  $\mathbf{x}_k^d, \mathbf{v}_k^d$ 
7:       Initialize  $\tilde{\mathbf{y}}_k^d := \mathbf{x}_k^d$ 
8:       Randomly initialize  $d_k, vd_k$ 
9:       Initialize  $\tilde{d}_k := d_k$ 
10:    end for
11:    Initialize  $gBest^d := \arg \min_{1 \leq k \leq S} \rho(\mathbf{x}_k^d, d)$ 
12:    Initialize  $\hat{\mathbf{y}}^d := \mathbf{x}_{gBest^d}^d$ 
13:  end for
14:  Initialize  $\hat{d} := \arg \min_{d_{min} \leq d \leq d_{max}} \rho(\hat{\mathbf{y}}^d, d)$ 
15:  for  $i \leftarrow 1, \dots, N$  do
16:     $w := w_0 - \frac{i}{N}(w_0 - w_N)$  ▷ Súlyok frissítése
17:    for  $k \leftarrow 1, \dots, S$  do
18:      if  $\rho(\mathbf{x}_k^{d_k}, d_k) < \rho(\tilde{\mathbf{y}}_k^{d_k}, d_k)$  then
19:         $\tilde{\mathbf{y}}_k^{d_k} := \mathbf{x}_k^{d_k}$  ▷ A részecske saját legjobb pozíciójának frissítése
20:        if  $\rho(\mathbf{x}_k^{d_k}, d_k) < \rho(\tilde{\mathbf{y}}_k^{\tilde{d}_k}, \tilde{d}_k)$  then
21:           $d_k := d_k$  ▷ Saját legjobb eredményt elérő dimenzió indexének frissítése
22:        end if
23:      end if
24:       $\rho_{d_k} := \min\{\rho(\hat{\mathbf{y}}_k^{d_k}, d_k), \min_{1 \leq l \leq S, l \neq k} \rho(\mathbf{x}_l^{d_k}, d_k)\}$ 
25:      if  $\rho(\mathbf{x}_k^{d_k}, d_k) < \rho_{d_k}$  then
26:         $gBest_{d_k} := k$  ▷ Globális legjobb pozíciónak frissítése a  $d_k$  dimenzióban
27:         $\hat{\mathbf{y}}^{d_k} := \mathbf{x}_k^{d_k}$ 
28:        if  $\rho(\mathbf{x}_k^{d_k}, d_k) < \rho(\hat{\mathbf{y}}^{\hat{d}}, \hat{d})$  then
29:           $\hat{d} := d_k$  ▷ A legjobb megoldás dimenziójának indexének frissítése
30:        end if
31:      end if
32:    end for
33:     $\mathbf{v}_k^{d_k} := c_1 r_1 \cdot (\tilde{\mathbf{y}}_k^{d_k} - \mathbf{x}_k^{d_k}) + c_2 r_2 \cdot (\hat{\mathbf{y}}^{d_k} - \mathbf{x}_k^{d_k}) + w \cdot \mathbf{v}_k^{d_k}$ 
34:     $\mathbf{x}_k^{d_k} := \mathbf{x}_k^{d_k} + \mathbf{v}_k^{d_k}$  ▷ A részecske elmozdul a keresési térben
35:     $vd_k := c_1 r_1 \cdot (\hat{d} - d_k) + c_2 r_2 \cdot (\hat{d} - d_k) + vd_k$ 
36:     $d_k := d_k + vd_k$  ▷ A részecske szükség esetén dimenziót vált
37:  end for
38:  return  $\mathbf{x}_{gBest^{\hat{d}}}^{\hat{d}}$ 
39: end function
```

4. Agytumor szegmentálás

4.1. Motiváció és kutatási célkitűzések

A legfrissebb kutatási eredmények és a különböző konvolúciós architektúrák összehasonlító elemzése alapján megállapítható, hogy a különböző U-net architektúrák kiemelkedő hatékonyságot mutatnak orvosi képek feldolgozásában más konvolúciós háló alapú modellekkel szemben is [38, 39]. Ennek oka, hogy az U-net struktúra egyaránt figyelembe veszi a kontextuális és a texturális jellemzőket, ami növeli a diagnosztikai pontosságot és javítja a szegmentációs eredmények minőségét. A jó eredmények eléréséhez azért van szükség ennyi adatra, mivel önmagában egy összetett feladatról beszélhetünk. Az adathalmaz jobb kihasználhatósága érdekében erre az lehet egy megoldás, ha külön próbálunk meg kontextuális és külön textúra jellemzőket kinyerni, majd ezeket egyesítve egy végleges szegmentációt biztosítani.

A kontextuális információ kinyeréséhez egy olyan szegmentáló algoritmust fejlesztünk, amely képes automatikusan adaptálódni a különböző képek sajátosságaihoz, optimalizálva a szegmensek számát és a képi elemek kategorizálását. Megmutatjuk, hogy a TDK-ban kidolgozott eljárás a nem felügyelt algoritmusokhoz képest javítja az objektumfelismerés pontosságát, kezeli a pixelintenzitásokra támaszkodó szétszórt pixelek problémáját, és hatékonyan elkülöníti az objektumokat, mindezt úgy, hogy az algoritmus paramétereinek beállításához minimális mennyiségű adatra van szükség. Az algoritmus fejlesztése során kiemelt figyelmet fordítunk arra, hogy minimalizáljuk az objektumok összeolvadását és felaprózódását, hogy az eredmények minél pontosabban tükrözzék a képen látható strukturális elemeket és azok valódi térbeli elhelyezkedését.

4.2. Az MDPSO algoritmus kiterjesztése képszegmentálásra

Az általunk kifejlesztett szegmentáló algoritmus az MDPSO-n alapul, ami maga után vonja a keresési tér diszkretizálását, de inspirációt merít a "K-means" [40] és a "mean shift" algoritmusokból [41] is. Egy részecske helyzetvektora $\mathbf{x}^d$, a kép különböző szegmenseinek (vagy pixel klasztereinek) középpontjait jelöli, amely egy potenciális megoldást képvisel a szegmentálási problémára [26]. A klaszterközéppontokat x és y koordináták határozzák meg, így klaszterenként két paramétert kell optimalizálni. Ha tehát egy képet szeretnénk szegmentálni d klaszterszámmal, akkor tulajdonképpen egy $2d$ dimenziós ($2dD$) keresési teret járunk be.

A szegmentálási problémában a d dimenziós indexet egy $2dD$ keresési térhez rendeljük. A szegmentálás kezdetén ezek a középpontok egyenletesen és függetlenül helyezkednek el a képen.

A szegmentációhoz az eredeti MDPSO (2. Algoritmus) módszer ρ veszteségfukcionálját is meg kell adni. Ezért egy saját költségfüggvényt definiáltunk (2. Algoritmus), mely jól illeszkedik az MRI képek szegmentálásához. Ez a függvény minden képpontot a legközelebbi szegmenshez rendel, figyelembe véve a középpontokhoz relatív csoportosítási attribútumok transzformált és súlyozott összegét. Büntetjük a pixelintenzitás különbségét, a pixelek euklideszi távolságát, és még azt is, ha egy besorolandó pixel szomszédjai már más klaszterbe kerültek. Az említett büntető tagok mellett a fitnessfüggvény értékeli a szegmentálás jóságát is. Esetünkben csak a pixelek intenzitásbeli különbségéből származó összeget tekintjük fitnessértéknek. Ezt az összeget $\sqrt{d}$ -vel szorozzuk, ezzel megakadályozva a túlszegmentálás létrejöttét [42].

2. Algoritmus Fitnessfüggvény

```

1: function  $\rho(\mathbf{x}^d)$ 
2:    $fitness = 0$ 
3:   for all  $\mathbf{p}$  pixel in  $I$  image do
4:      $c = \arg \min_{1 \leq i \leq d} \left( \alpha_1 \cdot f_{color}(|I_{FLAIR}(\mathbf{x}_i^d) - I_{FLAIR}(\mathbf{p})|) \right.$ 
5:        $\quad + \alpha_2 \cdot f_{color}(|I_{T1CE}(\mathbf{x}_i^d) - I_{T1CE}(\mathbf{p})|)$ 
6:        $\quad + \beta \cdot f_{Euclidean}(\|\mathbf{x}_i^d - \mathbf{p}\|)$ 
7:        $\quad \left. + \gamma \cdot Neighbor(\mathbf{p}, \mathbf{x}_i^d) \right)$ 
8:      $\triangleright$  Választunk egy  $c$  klasztert  $\mathbf{p}$ -nek
9:      $\mathbf{p}_{segmentIndex} = c$ 
10:     $fitness += f_{color}(|I_{FLAIR}(\mathbf{x}_i^c) - I_{FLAIR}(\mathbf{p})|)$ 
11:   end for
12:   return  $\sqrt{d} \cdot fitness$ 
13: end function

```

A 2. Algoritmus esetén a következő elnevezési konvenciókat követjük:

- $\mathbf{x}_i^d$ az i . középpontot jelöli a helyvektor $\mathbf{x}^d$ -ben.
- α_1 , α_2 , β , és $\gamma > 0$ súlyok, amelyek meghatározzák az attribútumkülönbségek fontosságát a szegmens középpontjától.
- $f_{color}(\cdot)$ és $f_{Euclidean}(\cdot)$ transzformációs függvények, amelyet a 4.3. alfejezet ismertet.

- Az $I_{FLAIR}(\cdot)$, illetve $I_{T1CE}(\cdot)$ függvény visszaadja a megadott koordinátákon lévő pixel szürkeárnyaltos értékét a megadott típusú felvételeken.
- $Neighbor(\mathbf{p}, \mathbf{x}_i^d)$ egy olyan függvény, megadja $\mathbf{p}$ -nek az $\mathbf{x}_i^d$ középpontú, különböző szegmensbe sorolt szomszédos pixelek számát.
- $|\cdot|$ és $\|\cdot\|$ az euklideszi távolságmértékeket jelölik, rendre 1D és 2D-ben.

Az iterációk során a PSO logikát követve frissítjük a szegmenseket meghatározó középpontok pozícióit. A középpontok eltolási vektorát a saját legjobb megoldás, a globális legjobb megoldás (az aktuális dimenzióban), és az előző mozgás irányát jelölő tehetetlenségi vektor súlyozott összege alapján számítjuk ki.

3. Algoritmus A részecskék sebességkorlátozása

```

1: function LIMIT( $\mathbf{v}^d$ ,  $speedLimit$ )
2:   for each centroid velocity  $\mathbf{v}_i^d$  in  $\mathbf{v}^d$  do
3:     if  $\|\mathbf{v}_i^d\| > speedLimit$  then
4:        $\mathbf{v}_i^d = \frac{speedLimit}{\|\mathbf{v}_i^d\|} \cdot \mathbf{v}_i^d$ 
5:     end if
6:   end for
7:   return  $\mathbf{v}^d$ 
8: end function

```

A dolgozatban bemutatott MDPSO algoritmusban a részecskék túl nagy ugrását a sebességvektor nagyságának korlátozásával ($speedLimit = 4$) akadályozzuk meg (3. Algoritmus). Ez csak a sebességvektor komponensenkénti küszöbölését jelenti, a középpontok elmozdulásainak irányát nem változtatja meg. Ez a szabályozás javítja az algoritmus optimalizáló képességét, mivel megakadályozza, hogy a részecskék nagy lépésekkel ugorjanak át a keresési tér távoli pontjaira. Ezáltal az egyedek részletesebben megvizsgálják a keresési teret a már megtalált globális vagy lokális optimum körül. Ez azért is fontos, mert nagy elmozdulásokkal egy egyed könnyen átugorhatja a keresett optimumot, valamint nagyon sokszor kiléphet a keresési térből. Az utóbbi eset ugyancsak lassíthatja az optimális megoldás megtalálását, mivel általában a keresési térből kiszaladt egyed egy véletlenszerű, keresési téren belüli értékkel folytatja tovább a keresést.

A dimenzióváltozások is korlátozás alá esnek, hogy megakadályozzuk a részecskék dimenziójának túl hirtelen változását. Emiatt egy részecske egy iteráció során az aktuális dimenziójához képest legfeljebb két távolságra lévő dimenzióba ugorhat át. Az eredeti algoritmusban [33] maximum 15 részecske léphet be egy dimenzióba, a többlet részecskék

az előző dimenzióban maradnak. Egy másik megközelítést alkalmazunk az egyedek populációszámát vizsgáló tanulmány [30] alapján, melyben a 20 és 50 közötti populációméreték bizonyultak optimálisnak a különböző PSO változatokban. Ebből a megfontolásból 25-ös populációméret-korlátot állítunk be minden dimenzióban. A publikáció [30] alapján nyilvánvaló, hogy a magasabb dimenziók nem igényelnek jelentősen több részecskét az optimális megoldás eléréséhez.

4.3. Az javasolt metrikák a büntető attribútumokra

Szignifikáns javulást értünk el klaszterező algoritmusunkban azáltal, hogy az attribútumkülönbségeket transzformáló függvény használatával a nem lineárisan súlyoztuk az azonos csoportba tartozó pixelek és klaszterközéppontjuk közötti távolságokat. Az alábbi függvény bizonyult a legjobbnak:

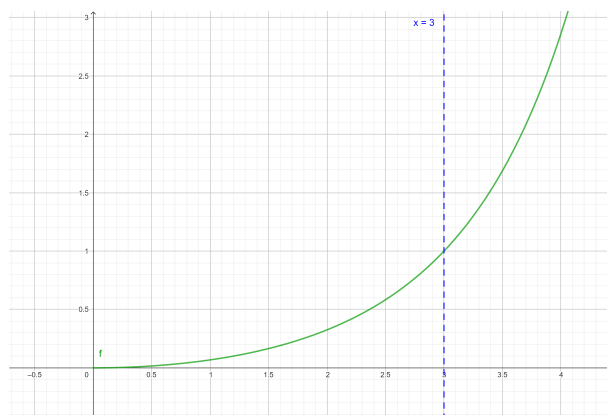
$$f : [0, \infty) \rightarrow [0, \infty), \quad f(x) = \frac{e^{\left(\frac{x}{\alpha}\right)^2} - 1}{e - 1}, (\alpha > 0) \quad (2)$$

A büntetőtag transzformáló függvényét a normál eloszlásból merített inspiráció alapján alakítottuk ki, mégpedig úgy, hogy egy nulla várható értékű és tetszőleges $(\sqrt{2} \cdot \alpha)$ szórású normál eloszlás sűrűségfüggvényét vettük alapul. Az $e^{-\left(\frac{x}{\alpha}\right)^2}$ kifejezés, a $\frac{1}{\alpha \cdot \sqrt{\pi}}$ konstanstól eltekintve, megegyezik az előbb említett normál eloszlás sűrűségfüggvényének képletével. Ennek a függvénynek a reciprokát fogjuk felhasználni, mivel ezáltal megfordul a monotonitás: a 0-hoz közeli eltéréseket alig, míg a 0-tól távoli értékeket már nagyon bünteti. Ezt még tovább toljuk -1 -el, és normalizáljuk $e - 1$ -el, ami segít a több attribútum hibatagjainak kiegyensúlyozásában, mivel így a $[0, \alpha)$ intervallumba eső hibákat a $[0, 1)$ intervallumból vett értékekkel bünteti. A (2) függvény azért is előnyös, mert lassan növekszik az $[0, \alpha)$ intervallumon, de $[\alpha, \infty)$ -ban felgyorsul az exponenciális növekedés. Ez azt jelenti, hogy amikor egy pixel egy közeli középpontú szegmensbe sorolódik, akkor az onnan származó hibaérték éppen csak ösztönzi a középpontot, hogy minél közelebb kerüljön a kialakított klaszter attribútumok súlyozott átlagérték vektorához. Ezzel szemben, ha egy pixelnek α -nál nagyobb numerikus attribútumkülönbsége van egy középponthoz képest, akkor ezt a függvény nagyobb mértékben bünteti, további iterációkban esetleg más klaszterbe sorolja.

A fitnessfüggvényben (2. Algoritmus) például két változatát használjuk fel az f transzformációs függvénynek, majd a súlyozott összege fogja képezni a részecske

fitnessértékének főbb részét. Az $\alpha = 300$ a színek esetében, míg $\alpha = 30$ a távolság esetében. Ezeket az értékeket az előfeldolgozott tesztadathalmaz pixelintenzitás hisztogramjai alapján határoztuk meg, figyelembe véve, hogy általában a tumort tartalmazó képpontok jelentős része általában egy 300 sugarú intervallumon belül helyezkedik el. Megjegyezzük, hogy nem garantált, hogy egy centroid a megfigyelt attribútumok alapján az intervallum közepéről vesz fel értéket.

A daganat méretét illetően orvosi szakcikkek [43] és a kísérleti eredmények (5. fejezet) alapján az $\alpha = 30$ beállítás bizonyult megfelelőnek az euklideszi távolság transzformálása során.

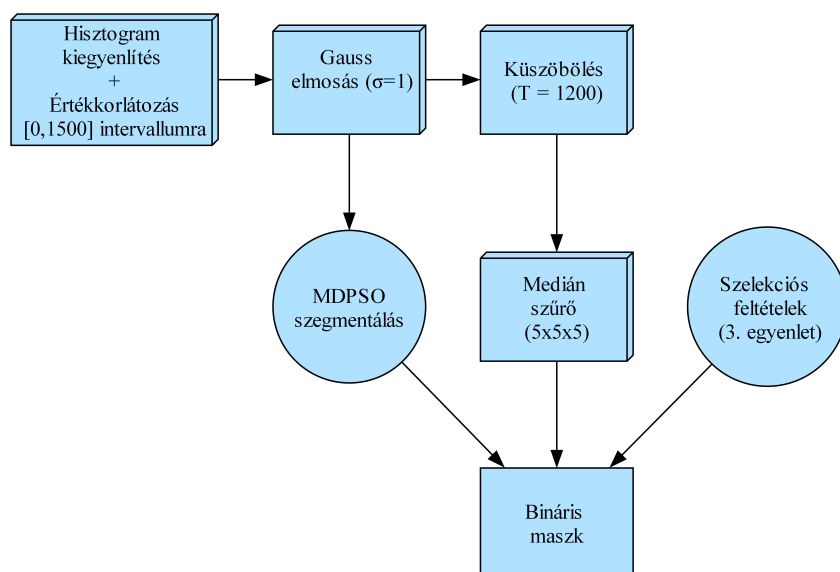


3. ábra. Az f függvény grafikus ábrázolása (zölddel, $\alpha = 3$ esetén) és az $x = \alpha$ egyenes (kékkel). Utóbbi egy tulajdonképpeni határvonal, ami jelzi hogy a függvény melyik tartományában alkalmaz 1-nél kisebb vagy nagyobb büntetést egy attribútumot véve figyelembe.

4.4. Képfeldolgozási lépések

Az MRI képeken a pixelértékek az adott szövetek mágneses tulajdonságaitól függően változnak, ezért a pixelintenzitásokat inkább relatív értékeknek, mintsem abszolút mérőszámoknak tekinthetjük. Ahogy korábban (a 2.1 alfejezetben) is említettük az MRI felvételeken intenzitás értékek az adott mágneses térben lévő hidrogén atomok viselkedését tükrözik, ami változik a különböző körülmények között. A betegek egyéni biológiai adottságai, például a víztartalom és a zsírszövetek aránya, befolyásolják az MRI jel intenzitását. Az MRI adatoknak nincs egységes intenzitási skálája, csak a színértékek nemnegativitása garantált. Ez a változékonyság tovább bonyolítja az adatok automatikus feldolgozását [13].

A nem felügyelt tanulás során, amit C++ programozási nyelvben valósítottunk meg (ennek lépéseit a 4. ábra szemlélteti) elsősorban a FLAIR szekvenciák voltak a legnagyobb hatással,



4. ábra. A nem felügyelt képszegmentálási folyamatára a dolgozatban ajánlott módszertan alapján. Az egyes lépéseknél használt alakzatok dimenzionalitása jelzi, hogy hány dimenziós információkat vettünk figyelembe.

mivel általában ezeken különböznek el legjobban és egyenletesebben a tumoros régiók. Azonban a szegmentálás robusztusságának érdekében figyelembe vettük a T1CE szekvenciákat is. Ennek köszönhetően már jól körvonalazódtak a különböző alrégiók is (lásd az 4.7. fejezetben: 1-es táblázat). A T1 és T2 szekvenciák nem játszottak szerepet a klaszterezés során, mivel a tumoros régiók színintenzitása a különböző felvételeken változatosak, vagy egyáltalán nem elkülönülők, emiatt csak rontották a másik két szekvencia felhasználásával a teszt adathalmazon elért szegmentálási pontosságot. A képek normalizálása, hisztogramkiegyenlítése érdekében a legtöbb kutatás során az általában jól bevált Z-normalizálást alkalmazzák [9, 19], azonban tekintve az MRI képek változatosságát egy stabilabb hisztogram-kiegyenlítést választottunk [13, 14]. A voxel-képek azon pozícióin lévő képpontokat vettük figyelembe, ahol a 4 modalitású felvétel közül legalább az egyik érték nagyobb mint 0, mivel a többi képpont a háttérre reprezentálta. Ezeknek a képpontoknak a színérték szerinti hisztogramját úgy állítottuk be, hogy a 25-percentilist 250-es értékre, míg a 75-percentilist 750-re rögzítettük a szürkeárnyalati skálán, ezáltal meghatározva a lineáris transzformáció két szabadságfokát. A transzformációt követően a 0 és 1500 intervallumon kívüli pixelértékeket az intervallum végpontjainak megfelelő értékre állítottuk, aminek következtében egy rögzített intenzitási skálát kaptunk.

Ezután egy 3D-s $\sigma = 1$ szórású Gauss-elmosást alkalmaztunk. Ennek kiemelt szerepe van a szegmentációs folyamatban, mert a 4.2. alfejezetben bemutatott algoritmusban fontos szerepe

van az ún. centroidoknak az egyes szegmensek azonosításában. Az elmosásnak köszönhetően viszont eltávolíthatjuk a képeken megjelenő zajos képpontokat, ami a szegmentáció stabilitását is eredményezi.

A következő lépésben a bemutatott szegmentációs algoritmust (4.2-es alfejezet) alkalmaztuk, 50-es rajmérettel, 30 iterációval, $d_{\min} = 4$ és $d_{\max} = 12$. A fitnessfüggvényben (2. Algoritmus) $\alpha_1 = 4$, $\alpha_2 = \beta = 1$ és $\gamma = 2$ értékekkel dolgoztunk. A paraméterek beállítását kezdetben 20 MRI felvétel felhasználásával végeztük el, majd a stabilabb és általánosabb értékek kiválasztásának érdekében különböző paraméter konfigurációkat is lefuttattunk a rendelkezésre álló 80 voxel-képből álló adathalmaz 80-as képszeleteire.

Az $[0, 1500]$ intervallumra lineáris transzformációval normalizált FLAIR képeken egy 1200-as értéket használó bináris küszöbölést és egy 3D-s medián szűrést alkalmazunk, $5 \times 5 \times 5$ -ös kernel ablakmérettel. Ennek célja egy olyan maszk létrehozása, amely kiemeli a nagyobb, világosabb szegmenseket az MRI felvételeken, segítve ezzel a potenciális daganatos egymás mellé tömörülő, tumoros szövetek azonosítását a szegmentáció során. Bár ez a megközelítés általában hatékonyan kiemeli a tumor magját, mivel a legtöbb esetben az világosan jelenik meg az említett modalitású felvételeken [2], azonban jelentős mennyiséget eltávolít a tumor többi részéből, ezért szükség van egy finomabb szegmentációra is. Ha az MDPSO alapú szegmentáló algoritmusunk eredménye szerint egy szegmens több mint 30%-a tartalmazza az előbbi módszerrel létrehozott maszkot, akkor azt daganatszegmensként azonosítjuk. Mivel ez a szelekciós eljárás az esetek nagyrésztében, de nem mindig jelöli ki a tumoros régiókat, ezért más feltételeket is bevezettünk a tumorok és azok alrégióit jelölő szegmensek kiválasztására. Ha a kiválasztott szegmens képpontjainak átlagértéke nagyon kicsi, vagy éppen nagy (3) (a különböző szekvenciákon különböző a megjelenés [2]), ami azt jelenti, hogy feltűnően sötét vagy világos a különböző modalitásokon, akkor azt is tumornak tekintjük. Ez az eljárás kissé megnövelheti a hamisan tumornak jósolt pixelek számát, de a probléma súlyosságát figyelembe véve ez még mindig kisebb gond, mint a tévesen egészségesnek jelölt szöveti területek megjelenése. Általános megfigyelések és saját tesztelések alapján a következő feltételek

vezethetőek be a tumorok azonosítására MRI képeken:

$$\text{Tumort jelölő szegmensek} \approx \begin{cases} T1CE < 200 & \text{és} & T2 < 200 \\ \text{FLAIR} > 1000 & \text{és} & (T1CE > 1000 \text{ vagy } T2 > 850) \\ \text{FLAIR} > 850 & \text{és} & T2 > 1000 \end{cases} \quad (3)$$

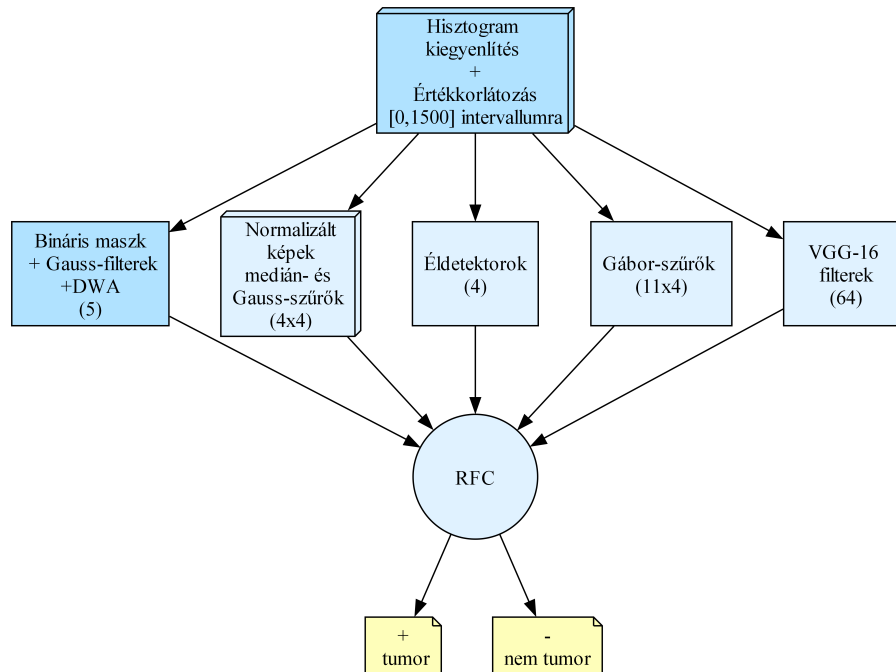
A FLAIR, T1CE és T2 paraméterek a kialakított szegmensek képpontjainak színintenzitásának átlagát jelölik. Ezek az értékek segítenek azonosítani azokat a területeket, ahol a szöveti szerkezet eltér a normálistól, ami daganat jelenlétére utalhat.

4.5. Utófeldolgozás

A bemutatott szegmentációs utófeldolgozás Python nyelvben valósult meg, amely magában foglal egy hibrid megközelítést: a véletlen erdő osztályozót (RFC – Random Forest Classifier) hagyományos szűrők által kinyert attribútumokkal és a VGG16 konvolúciós hálót. Ez a hibrid megoldás kihasználja mind a hagyományos gépi tanulás, mind a mélytanulási technikák előnyeit. Az RFC mintahatékonysága és strukturális egyszerűsége lehetővé teszi a modellek értelmezhetőségét és megbízhatóságát, ami kulcsfontosságú a klinikai döntéshozatali folyamatokban [44]. A Gábor-szűrők alkalmazása segít az emberi látáshoz hasonló textúraelemzésben, ami növeli a szegmentációs pontosságot különösen textúrában gazdag képeken [13]. A VGG16 konvolúciós hálózat segítségével pedig mélyebb jellemzőket tudunk kinyerni a képekből [45], ami javítja a szegmentáció általános minőségét, különösen kevés tanítási adat esetén, ahol csak a mélytanulási modellek alkalmazása túltanulást eredményezhet [46].

A véletlen erdő (RF – Random Forest) egy összetett gépi tanulási módszer, amely több döntési fa összességéből áll, és az ún. ensemble learning technikát alkalmazza. Ennek lényege, hogy az algoritmus több különböző adathalmazon tanít több döntési fát, majd ezek eredményeit összefogja, így csökkentve a modell túltanulásának esélyét és növelve az általánosító képességét. A Python scikit-learn könyvtárában elérhető `RandomForestClassifier` [47] osztálya számos paramétert kínál, amelyek segítségével finomhangolhatjuk a modell működését. Az `n_estimators` paramétert 300-ra állítottuk, hogy növeljük a modell robusztusságát és csökkentsük az egyes fákra gyakorolt túlzott illeszkedés esélyét, így a modell általánosító képessége javul. A `max_features` paramétert az alapbeállításon hagytuk, azaz

$\sqrt{n}$, ami azt jelenti, hogy minden egyes döntési fánál a rendelkezésre álló n jellemzőből $\sqrt{n}$ kerül figyelembe vételre, ez segít elkerülni a fák közötti korrelációt és növeli a modell diverzitását. A `max_depth` értékét 30-ra állítottuk, ami korlátozza a fák mélységét, így megakadályozva a túltanulást. A `min_samples_split` és `min_samples_leaf` paraméterek, amelyeket rendre 5 és 2 értékre állítottuk, tovább finomítják a fa növekedését: az előbbi meghatározza, hogy minimum hány minta szükséges egy új csomópont létrehozásához, míg az utóbbi a levélcsomópontokban minimálisan szükséges minták számát szabályozza.



5. ábra. A felügyelt szegmentáló algoritmus folyamatábrája, amely összesíti a különböző módszerekkel kinyert információkat és elvégzi a végső osztályozást. A különböző típusú attribútumok alatt azok számossága jelenik meg. A pontos attribútumok és azok paraméterei, valamint az osztályozásban a relevanciájának mértéke megtalálható a csatolmányban.

A RFC modell tanításához alkalmazott attribútumokat három fő csoportból válogattuk: először is, a hagyományos Gabor-szűrőket [48] használtuk fel mind a négy MRI szekvenciára ($4 \times 11 = 44$ jellemző). Ezeket a szűrőket a képfeldolgozásban textúraelemzésre és az emberi látás szimulálására fejlesztettek ki. Ezek a szűrők egy Gauss-függvénnyel kombinált trigonometrikus hullámokat használnak, amelyek a frekvencia-érzékenységet és a térbeli lokalizációt egyaránt biztosítják. Ezáltal képesek különféle texturális mintázatok, mint vonalak és élek detektálására. A Gabor-szűrők magyarázhatósága növeli a technológia klinikai elfogadását is.

Második csoportként, további hagyományos szűrőket, mint a Prewitt, Sobel, Scharr, Roberts az élek detektálására a 2D-s képszeleteken valamint 3D-s medián (5×5 és 7×7 -es

kernelmérettel) és Gauss-szűrőt ($\sigma = 3$ szórással) alkalmaztunk a FLAIR típusú MRI felvételekre (7 jellemző). Az egyes szűrők paramétereit és az általuk generált attribútumokat részletesen elemeztük kisebb adathalmazokon, majd a végső, a legjobb eredményeket elérő modell tanítására legrelevánsabbakat választottuk ki. Az egymással korreláló attribútumokat kiszűrve törekedtünk az optimális attribútum készlet összeállítására.

A harmadik típusú jellemzőkinyerő eszközként a VGG16 konvolúciós neurális hálózatot használtuk, amelyet az ImageNet adatbázissal előzetesen betanítottak. Ezen belül a modell első blokkjának, második konvolúciós rétege után kapott attribútumokat használtuk fel a szegmentáláshoz. A T1CE, FLAIR, és T2 szekvenciákból álló háromcsatornás képeket a hálózat bemenetére méreteztük, majd a konvolúció után a kapott 64 csatornás kimenetet visszaállítottuk az eredeti méretre, ami által minden képponthoz 64 újabb attribútumot rendelhettünk. Ez lehetővé teszi a mélyebb textúrák figyelembevételét a szegmentálási folyamat során, amit a hagyományos szűrők nem feltétlenül képesek detektálni.

A bemutatott hibrid szegmentáló algoritmust a szegmentálás utófeldolgozásaként is használtuk, amelyhez további jellemzők hozzáadásával bővítettük a RFC modellt. Az MDPSO szegmentáló algoritmus kimenete mellett három különböző szórású ($\sigma \in \{1, 3, 6\}$) Gauss-homályosítást alkalmaztunk, amelyek segítségével az RFC fel tudja mérni a kép éleinek bizonytalanságát. Ezenfelül a főként CNN-ek esetében alkalmazott [19], távolságarányos figyelemréteget (DWA – Distance-Wise Attention) is integráltunk, amely az MDPSO alapú szegmentáló algoritmus által generált maszk szerint észlelt objektum súlypontjától számított távolságot méri.

4.6. Tanítás és validálás

A jellemzők kinyerése és az RFC modellek tanítása, majd tesztelése egy AMD Ryzen 7 4700U CPU-val, 8 GB RAM-mal felszerelt alaplapon történt, amelyet egy beépített AMD Radeon Graphics GPU egészít ki, 512 MB dedikált memóriával. A rendelkezésre álló adathalmazból, amely 1251 felvételt tartalmaz, véletlenszerűen kiválasztottunk 80-at tanítási célokra. A megmaradt felvételeket a módszerünk tesztelésére használtuk. Mind a tanítás, mind a tesztelés során minden felvételből kiválasztottuk a 80-as képszeleteket az [1, 155] intervallumból. Ennek alapján célunk a tumordetektálás hatékonyságának növelése, miközben a módszerünk képes a tumor jelenlétének kizárására is a felvételeken. A tanító adathalmaz összeállításához a helyes szegmentálási adatok alapján minden tumornak címkézett pixelt

felhasználtunk. Ezenfelül, azonos számban mintavételeztünk nem tumorként címkézett, de nem minden szekvencián nulla értékű képpontokat is. Amennyiben egy voxel-kép szelete nem tartalmazott tumort, annak érdekében, hogy a modellünk változatossága és stabilitása növekedjen, ebből a képszeletből is véletlenszerűen kiválasztottunk 20 képpontot.

Mivel más módszerek általában jelentősen nagyobb adathalmazokat igényelnek a tanításhoz, a mi megközelítésünk célja, egy erőforrás-hatékony, kevés adatot igénylő szegmentáló módszer kidolgozása.

4.7. Eredmények, numerikus szimulációk

Az eredmények kvalitatív értékelése

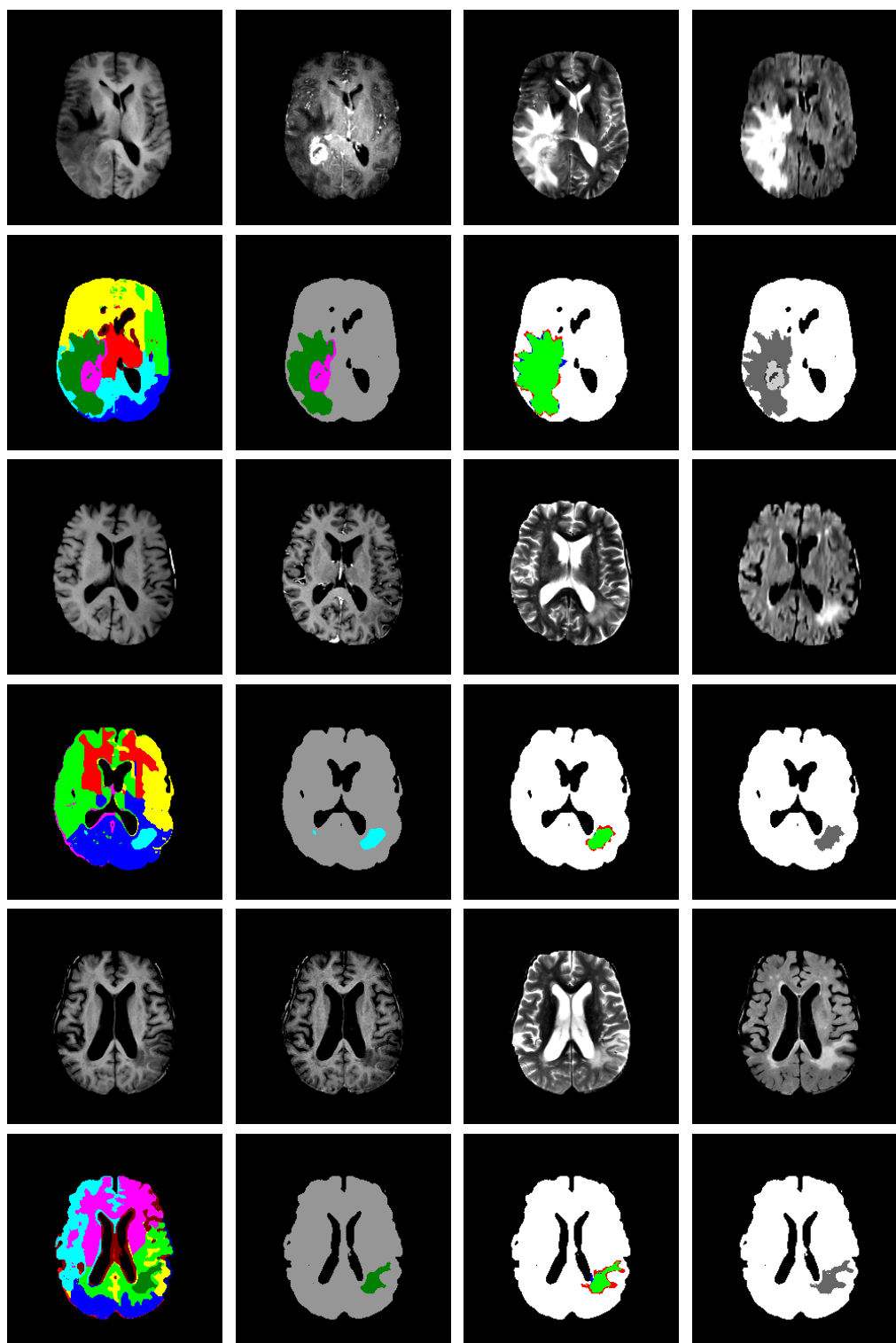
A dolgozatban tárgyalt eredmények 1171 voxel-kép 80-as képszeleteire vonatkoznak, amelyek kívül esnek a 80 tanítási és paraméterhangolási mintán. Ezen képszeletek közül 1027 db (87.70%) tartalmaz tumoros pixeleket, melyeknek képenkénti mennyisége széles skálán mozog. A 300 pixel alattiak (235 db), ami körülbelül 2 cm-nél és annál kisebb átmérőt jelent, míg a legnagyobb 6175 pixel. Az eredmények stabilitását növeli, hogy a MDPSO alapú szegmentációs algoritmus képes a különböző alrégiók pontos elkülönítésére és még a szabad szemmel is alig látható tumorok detektálására is (lásd: 1. táblázat).

Az eredmények kvantitatív értékelése

Képszegmentálás esetén számos pontosság-mérő metrikát különböztethetünk meg. Ezek általában az igazságmátrix (confusion matrix) értékein alapulnak, amely a következőket tartalmazza: valós pozitív (TP – true positives), valós negatív (TN – true negatives), hamis pozitív (FP – false positives), hamis negatív (FN – false negatives). Tumorszegmentálás esetén leggyakrabban a Dice-együtthatót használják, mert a kisebb objektumok szegmentálásának hatékonyságát tartja szem előtt [2].

A klaszterezéssel megvalósított szegmentálás javítására (lásd a 2. táblázatot), a textúra jellemzők bevonásával, a már korábban agytumor szegmentálására is alkalmazott Gábor-szűrők által kinyert attribútumokat felhasználva, az RFC algoritmust vettük alapul (RFC(GF)) [14, 13]. Ebben a tanítási folyamatban a négyféle MRI képtípusra összesen 64 Gábor-szűrőt alkalmaztunk. Ezeket a szűrőket kisebb adathalmazokon választottuk ki rácskereséses módszerrel és az RFC által generált attribútumok fontossági listája alapján, majd a korreláló

1. táblázat. A hisztogram-kiegyenlítéssel feldolgozott különböző MRI modalitások (sorrendben T1, T1CE, T2, FLAIR) és az ezekhez kapcsolódó szegmentációs algoritmusok eredményei. Az eredményeket mutató sorokban az első képen az MDPSO alapú szegmentációs algoritmus eredménye, utána az az algoritmus által kiválasztott tumornak megítélt szegmensek vannak kiemelve, ezt követi a RFC által végzett szegmentálás (a helyesen tumorosnak ítélt pixeleket zölddel, a tévesen egészségesnek ítéltet pirossal, a tévesen tumorosnak ítéltet pedig késsel jelöljük) és végül pedig az orvosok által elvégzett szegmentáció (alrégiókra lebontva) van bemutatva.



szűrőket vizuális elemzéssel választottuk ki.

További fejlesztésként a VGG16 modellből származó két konvolúciós réteg által generált szűrőket integráltuk az RFC(GF + VGG16) modellbe. Bár ezek a szűrők alacsonyabb fontossági százalékot kaptak, jelentős különböző információkat nyertek ki a Gábor-szűrőkhöz képest.

Szignifikáns javulást az MDPSO által adott szegmentációval értünk el, több mint 6%-al magasabb pontosságot mutatva az előző modellekhez képest. Ez a modell 133 attribútumot generál a pixelek kategorizálásához. A klaszterező algoritmus által készített maszk egy $\sigma = 3$ szórású Gauss-elmosással az attribútumok fontossági listájának élén 7,42%-al szerepel, míg az MDPSO eredményét felhasználva generált DWA réteg a lista 13. helyén, 2,62%-os szignifikanciával. A klaszterezéssel készített bináris maszk után a lista első felében különböző paraméterű Gábor-szűrők szerepelnek, a második felében pedig főként a CNN rétegei, melyek nem voltak MRI jellemzők kinyerésére optimalizálva.

2. táblázat. A különböző modellek által elért szegmentációs teljesítmény összehasonlítása a teljes illetve kisebb tesztelési adathalmazokon. Az utolsó két sor bemutatja az adathalmazban egymást követően átlagosan a legjobb, 260, illetve 100 képből álló teszteredményeket.

Módszer	Dice	Pontosság	Precizitás	Érzékenység	Specifitás	Tanítás	Tesztelés
MDPSO	0.7187	0.9897	0.8255	0.7665	0.9957	80	1171
RFC(GF)	0.7207	0.9925	0.8202	0.7668	0.9983	80	1171
RFC(GF + VGG16)	0.7406	0.9923	0.8373	0.7729	0.9983	80	1171
MDPSO + RFC(GF + VGG16)	0.8014	0.9930	0.9021	0.8006	0.9983	80	1171
	0.8773	0.9956	0.9018	0.8986	0.9980	80	260
	0.8991	0.9961	0.9210	0.9200	0.9981	80	100

Ahogy más kutatások is kimutatták [14], a tanítási és tesztelési adathalmazok nagysága és diverzitása nagy mértékben befolyásolja a teljesítménymutatókat. Ez nyilvánvaló a fenti táblázatból is. Ennek következtében nem is várható el, hogy módszerünk jobb eredményeket érjen el, mivel jelentősen kevesebb adatot használtunk fel. Ugyanakkor megmutattuk, hogy a TDK-ban kidolgozott módszerrel és más algoritmusok ötvözésével az irodalomban ismert szegmentáló eljárások pontossága jelentősen javítható. Továbbá, ahogy más kutatások is kimutatták [14], a tanítási és tesztelési adathalmazok nagysága és diverzitása nagy mértékben befolyásolja a teljesítménymutatókat, ami nyilvánvaló a 2. táblázat utolsó soraiból is. Jelentősen kisebb méretű adathalmaz felhasználásával is hasonló eredményeket értünk el, mint más erőforrásigényesebb eljárások. Például egy RFC alapú megközelítés a rendelkezésre álló 335 felvételnek 5/6-át használta fel tanítási célokra, és a legjobb modell 0,8516 Dice-értéket ért el [14]. A 2021-es BraTS nyertes modellje [4] 0,9277 Dice-értéket ért el, de itt jelentősen több tanítási adatot (1251 felvételt) és magasabb hardverigényű módszert alkalmaztak.

5. Összegzés

A dolgozatban bemutatott innovatív és hagyományos jel- és képfeldolgozási, illetve gépi tanuló módszerek ötvözésének eredményei alapján elmondható, hogy az MDPSO algoritmus különösen hatékonyan alkalmazható olyan komplex szegmentációs feladatokra is, mint amilyen az MRI felvételeken történő tumorszegmentálás. A TDK-ban bemutatott eljárás átmenetet képez a teljesen, a részben, és a nem felügyelten tanuló módszerek között, ami különösen fontos az orvosi adatok feldolgozásában, mivel ezáltal csökken a szegmentálás pontosságának függősége a tanítóhalmaz méretétől.

Az MDPSO algoritmusnak egy megfelelő, erre a célra kifejlesztett fitnessfüggvénnyel való ötvözése a hagyományos küszöbölési technikáknál sokkal fejlettebb objektumfelismerési képességekkel rendelkezik, amely autonóm módon képes meghatározni az optimális szegmensek számát. Ennek köszönhetően kevesebb adat felhasználásával is lehetővé teszi a szegmentáló algoritmus hiperparamétereinek finomhangolását, miközben a felügyelt tanuláson alapuló módszerekkel összehasonlítható eredményeket ér el.

Ezen túlmenően, a bemutatott módszer képes növelni az irodalomban jól ismert szegmentáló technikák pontosságát is. A szegmentálási eljárás során alkalmazott stratégia olyan információk kiemelését teszi lehetővé, amelyeket más módszerek esetleg figyelmen kívül hagynak. Az algoritmus eredményeit ötvözve más jellemzőkkel és az RFC algoritmussal, a helyi kórházak címkézett adatbázisainak méretéhez alkalmazkodva sikerült egy hibrid modellt kidolgozni, ami egy kis méretű tanítóhalmazzal (80 felvétellel) felhasználva egy teljesen általános validációs adathalmazon (1171 felvétellel) is jól teljesít. Utóbbi jól mutatja a szegmentáló algoritmus általánosítóképességét és robusztusságát is. Ez azzal magyarázható, hogy a módszerünk nem csupán a tanítási adatokból származó információkra hagyatkozik a döntések meghozatalakor, hanem a bemenő adatokon végzett matematikai számításokra is. Így olyan általános érvényű feltételeket sikerült bevezetni a fitnessfüggvénybe, melyek hozzájárulnak a pontosság növeléséhez, de kevésbé függenek a tanítóadatoktól.

Kitekintés

A közeljövőben célunk a pontosság és teljesítménymutatók további finomítása a klaszterizáció optimalizálásával, különös figyelmet fordítva a daganatos pixelek precízebb azonosítására. Tervezzük további attribútumok integrálását, hogy hatékonyabban tudjunk

különböző daganattípusokat felismerni. Érdeemes lenne kiterjeszteni az MDPSO algoritmust 3D szegmentációhoz is, kihasználva a PSO algoritmus robusztus teljesítményét magas dimenziós keresési terekben. Ez a kiterjesztés várhatóan gyorsabb és pontosabb eredményeket biztosít a hagyományos szeletenkénti szegmentációval szemben, figyelembe véve a térbeli információkat is. Közben a paraméterek automatizált finomhangolásával szeretnénk a módszert a teljes térfogat szegmentálására általánosítani és más orvosi vagy nem orvosi feladatokra is alkalmazni.

A felügyelt tanuláson alapuló lépések finomítása és tesztelése is szerepel a jövőbeli célkitűzések között. Például, a már most is rendelkezésre álló adathalmazon betanított CNN szűrőit tervezzük kipróbálni a jelenleg alkalmazott, teljesen általános adathalmazon betanított szűrők helyett, felhasználva a transzfer tanulási technikákat. A fejlesztések célja egy olyan végleges szegmentációs módszer kifejlesztése, amelyben a nem felügyelt számítások paraméterei már a helyileg, egy-egy kórházban elérhető adathalmazon is stabilan finomhangolhatók. Ez a módszer esetleg több paramétert is tartalmazhat a jelenleginél, miközben megőrzi hatékonyságát, és pontosabb eredményeket nyújt. Mindezek következtében módszer alkalmazása hagyományos kórházi számítógépeken is lehetséges lesz, ami jelentősen hozzájárulhat a pontos diagnózisok elvégzéséhez és az orvosi munka megkönnyítéséhez is.

Köszönetnyilvánítás

Köszönettel tartozom a Balassi Bálint-ösztöndíjprogramnak, hogy a Márton Áron Tehetséggondozó Ösztöndíjával lehetővé tették kutatómunkám elkezdését (2022-2023), Vekov Géza Károly tanár úr segítségével. Hálás vagyok a Márton Áron Szakkollégiumnak az ösztöndíjbeli támogatásért (2023-2024) és dr. Kovács Péter egyetemi docensnek a magas színvonalú tutorálásért, amely nélkül kutatásom nem érte volna el jelenlegi eredményeit. Nagy köszönet illeti Sándor Csanád kiváló témavezetői munkáját az egész 2023-2024-es tanév folyamán. Hálás vagyok a sepsiszentgyörgyi "Dr. Fogolyán Kristóf" Megyei Sürgősségi Kórház és a kolozsvári „Prof. Dr. Ion Chiricuță” Onkológiai Intézet csapatainak a betekintésért munkafolyamataikba és az eszközök megismeréséért. Végezetül szeretnék köszönetet mondani családomnak, barátaimnak, valamint minden tanáromnak és tanítómnak az inspiráló támogatásért és motivációért tanulmányaim során.

Hivatkozások

- [1] Irena Ilic és Milena Ilic. „International patterns and trends in the brain cancer incidence and mortality: An observational study based on the global burden of disease”. *Heliyon* 9.7 (2023). URL: <https://www.sciencedirect.com/science/article/pii/S2405844023054300>.
- [2] Ramin Ranjbarzadeh és mások. „Brain tumor segmentation of MRI images: A comprehensive review on the application of artificial intelligence tools”. *Computers in biology and medicine* 152 (2023), 106405. oldal.
- [3] Juha I Peltonen és mások. „Inter-and intra-scanner variations in four magnetic resonance imaging image quality parameters”. *Journal of Medical Imaging* 7.6 (2020), 65501–065501. oldal.
- [4] Huan Minh Luu és Sung-Hong Park. „Extending nn-UNet for Brain Tumor Segmentation”. *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*. Cham: Springer International Publishing, 2022, 173–186. oldal.
- [5] Geoff Dougherty. *Digital image processing for medical applications*. Cambridge University Press, 2009.
- [6] Catherine Westbrook és John Talbot. *MRI in Practice*. John Wiley & Sons, 2018.
- [7] Pawel Mlynarski és mások. „3D convolutional neural networks for tumor segmentation using long-range 2D context”. *Computerized Medical Imaging and Graphics* 73 (2019), 60–72. oldal.
- [8] Jean-Philippe Fortin és mások. „Removing inter-subject technical variability in magnetic resonance imaging studies”. *NeuroImage* 132 (2016), 198–212. oldal.
- [9] Ujjwal Baid és mások. „The rsna-asnr-miccai brats 2021 benchmark on brain tumor segmentation and radiogenomic classification”. *arXiv preprint arXiv:2107.02314* (2021).
- [10] Bjoern H Menze és mások. „The multimodal brain tumor image segmentation benchmark (BRATS)”. *IEEE transactions on medical imaging* 34.10 (2014), 1993–2024. oldal.
- [11] Spyridon Bakas és mások. „Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features”. *Scientific data* 4.1 (2017), 1–13. oldal.
- [12] Duc-Phong Nguyen és mások. „Fast 3D Face Reconstruction from a Single Image Using Different Deep Learning Approaches for Facial Palsy Patients”. *Bioengineering* 9.11 (2022), 619. oldal.
- [13] László Szilágyi és mások. „Low and high grade glioma segmentation in multispectral brain MRI data”. *Acta Universitatis Sapientiae, Informatica* 10.1 (2018), 110–132. oldal.

- [14] Ágnes Győrfi, László Szilágyi és Levente Kovács. „A fully automatic procedure for brain tumor segmentation from multi-spectral MRI records using ensemble learning and atlas-based data enhancement”. *Applied Sciences* 11.2 (2021), 564. oldal.
- [15] Hari McGrath és mások. „Manual segmentation versus semi-automated segmentation for quantifying vestibular schwannoma volume on MRI”. *International journal of computer assisted radiology and surgery* 15 (2020), 1445–1455. oldal.
- [16] Diana Veiga-Canuto és mások. „Comparative multicentric evaluation of inter-observer variability in manual and automatic segmentation of neuroblastic tumors in magnetic resonance images”. *Cancers* 14.15 (2022), 3648. oldal.
- [17] Patrick J Kelly. „Gliomas: Survival, origin and early detection”. *Surgical neurology international* 1 (2010).
- [18] Hongwei Bran Li és mások. „The brain tumor segmentation (BraTS) challenge 2023: Brain MR image synthesis for tumor segmentation (BraSyn)”. *ArXiv* (2023).
- [19] Ramin Ranjbarzadeh és mások. „Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi-modalities brain images”. *Scientific Reports* 11.1 (2021), 1–17. oldal.
- [20] Zoltán Kapás és mások. „Automatic brain tumor segmentation in multispectral MRI volumes using a random forest approach”. *Pacific-Rim Symposium on Image and Video Technology*. Springer. 2017, 137–149. oldal.
- [21] Szabolcs Csaholczi, Levente Kovács és László Szilágyi. „Automatic segmentation of brain tumor parts from MRI data using a random forest classifier”. *2021 IEEE 19th World Symposium on Applied Machine Intelligence and Informatics (SAMi)*. IEEE. 2021, 471–000476. oldal.
- [22] Prasanna K Sahoo, SAKC Soltani és Andrew KC Wong. „A survey of thresholding techniques”. *Computer vision, graphics, and image processing* 41.2 (1988), 233–260. oldal.
- [23] Hui Zhang, Jason E Fritts és Sally A Goldman. „Image segmentation evaluation: A survey of unsupervised methods”. *computer vision and image understanding* 110.2 (2008), 260–280. oldal.
- [24] Diego Oliva és mások. „Multilevel thresholding for image segmentation based on metaheuristic algorithms”. *Metaheuristic algorithms for image segmentation: theory and applications* (2019), 59–69. oldal.
- [25] Minghai Xu és mások. „Application of swarm intelligence optimization algorithms in image processing: A comprehensive review of analysis, synthesis, and optimization”. *Biomimetics* 8.2 (2023), 235. oldal.

- [26] Mahamed Omran, Andries P Engelbrecht és Ayed Salman. „Particle swarm optimization method for image clustering”. *International Journal of Pattern Recognition and Artificial Intelligence* 19.03 (2005), 297–321. oldal.
- [27] James Kennedy és Russell Eberhart. „Particle swarm optimization”. *Proceedings of ICNN’95-international conference on neural networks*. 4. kötet. iee. 1995, 1942–1948. oldal.
- [28] Meetu Jain és mások. „An overview of variants and advancements of PSO algorithm”. *Applied Sciences* 12.17 (2022), 8392. oldal.
- [29] Francisco Martínez-Álvarez és mások. „Coronavirus optimization algorithm: a bioinspired metaheuristic based on the COVID-19 propagation model”. *Big data* 8.4 (2020), 308–322. oldal.
- [30] Adam P Piotrowski, Jaroslaw J Napiorkowski és Agnieszka E Piotrowska. „Population size in particle swarm optimization”. *Swarm and Evolutionary Computation* 58 (2020), 100718. oldal.
- [31] Serkan Kiranyaz, Jenni Pulkkinen és Moncef Gabbouj. „Multi-dimensional particle swarm optimization in dynamic environments”. *Expert Systems with Applications* 38.3 (2011), 2212–2223. oldal.
- [32] Upendra Kumar Acharya és Sandeep Kumar. „Particle swarm optimized texture based histogram equalization (PSOTHE) for MRI brain image enhancement”. *Optik* 224 (2020), 165760. oldal.
- [33] Serkan Kiranyaz, Turker Ince és Moncef Gabbouj. *Multidimensional particle swarm optimization for machine learning and pattern recognition*. Springer, 2014.
- [34] Saugata Roy, Nabajyoti Mazumdar és Rajendra Pamula. „An energy optimized and QoS concerned data gathering protocol for wireless sensor network using variable dimensional PSO”. *Ad Hoc Networks* 123 (2021), 102669. oldal.
- [35] Péter Kovács, Sándor Fridli és Ferenc Schipp. „Generalized rational variable projection with application in ECG compression”. *IEEE Transactions on Signal Processing* 68 (2019), 478–492. oldal.
- [36] Péter Kovács, Serkan Kiranyaz és Moncef Gabbouj. „Hyperbolic particle swarm optimization with application in rational identification”. *21st European Signal Processing Conference (EUSIPCO 2013)*. IEEE. 2013, 1–5. oldal.
- [37] Yuhui Shi és Russell Eberhart. „A modified particle swarm optimizer”. *1998 IEEE international conference on evolutionary computation proceedings. IEEE world congress on computational intelligence (Cat. No. 98TH8360)*. IEEE. 1998, 69–73. oldal.
- [38] Jiawei Zhang és mások. „Mdu-net: Multi-scale densely connected u-net for biomedical image segmentation”. *Health Information Science and Systems* 11.1 (2023), 13. oldal.

- [39] Olaf Ronneberger, Philipp Fischer és Thomas Brox. „U-net: Convolutional networks for biomedical image segmentation”. *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III* 18. Springer. 2015, 234–241. oldal.
- [40] K Krishna és M Narasimha Murty. „Genetic K-means algorithm”. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 29.3 (1999), 433–439. oldal.
- [41] Yizong Cheng. „Mean shift, mode seeking, and clustering”. *IEEE transactions on pattern analysis and machine intelligence* 17.8 (1995), 790–799. oldal.
- [42] Jianqing Liu és Yee-Hong Yang. „Multiresolution color image segmentation”. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 16.7 (1994), 689–700. oldal.
- [43] Mary F Dempsey, Barrie R Condon és Donald M Hadley. „Measurement of tumor “size” in recurrent malignant glioma: 1D, 2D, or 3D?”. *American journal of neuroradiology* 26.4 (2005), 770–776. oldal.
- [44] Hyunseok Seo és mások. „Machine learning techniques for biomedical image segmentation: an overview of technical aspects and introduction to state-of-art applications”. *Medical physics* 47.5 (2020), e148–e167.
- [45] Mohammadreza Soltaninejad és mások. „Multimodal MRI brain tumor segmentation using random forests with features learned from fully convolutional neural network”. *arXiv preprint arXiv:1704.08134* (2017).
- [46] Xiangbin Liu és mások. „A review of deep-learning-based medical image segmentation methods”. *Sustainability* 13.3 (2021), 1224. oldal.
- [47] *RandomForestClassifier documentation in Scikit-learn*. <https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html>. Accessed: 2024-04-25.
- [48] *Gabor Filters in OpenCV*. https://docs.opencv.org/4.x/d4/d86/group__imgproc__filter.html. Accessed: 2024-04-26.