

Mély tanulási algoritmusok a csontszcintigráfia javítására



Kovács Ákos

A doktori disszertáció tézisei

Témavezetők:
Dr. Horváth András,
Dr. Bükki Tamás

Roska Tamás Műszaki és Természettudományi Doktori Iskola
Pázmány Péter Katolikus Egyetem

Budapest, 2023

1. Bevezetés

Az orvosi képalkotó készülékek rendkívül fontosak a mai, tudományos eredményeken alapuló betegközpontú ellátásban. Ezek a készülékek az elmúlt 30 évben teljesen megváltoztatták az orvoslás gyakorlatát. Lehetővé tették a betegségek korai stádiumban történő felismerését (különösen a szűrővizsgálatok során), és a korai kezelésnek köszönhetően javultak a betegek gyógyulási esélyei. A röntgen, az ultrahang, a CT, az MRI és más képalkotó technológiák használata mindennaposá vált a klinikákon, és így széles körben ismertté váltak a társadalomban. Az orvosi képalkotás különösen hasznos a rák felismerésében és azonosításában, mivel a korai felismerés jobb gyógyulási esélyeket jelent. Az egyik ilyen célra rutinszerűen használt eszköz a gammakamera, amely különböző radioaktív izotópokat használ, hogy információt nyerjen a szervezetben zajló élettani folyamatokról.

Az egyik leggyakrabban végzett vizsgálat, a csontszcintigráfia tökéletesítésén dolgozom, amely egy gyakori, viszonylag olcsó és széles körben elérhető technika, amely érzékenysége miatt felbecsülhetetlen értékű számos kóros állapot diagnosztikai értékelésében. Magyarországon átlagosan 5-600MBq MDP izotóp aktivitást adnak be a betegnek, és 15-20 perces vizsgálatot végeznek a készülékkel. A gammakamerás felvételek jellemzően alacsony jel-zaj aránnyal rendelkeznek, és jelentős Poisson-zajnak vannak kitéve. Ezek a körülmények arra ösztönöztek, hogy doktori tanulmányaim során olyan képjavító eszközt fejlesszek ki, amely lehetővé teszi a

beadott aktivitás (radioaktív dózis) csökkentését és a képalkotási idő lerövidítését.

Kivételesen magas jel-zaj arány érhető el zajszűrő neurális hálózatok használatával, amelyek úgy konfigurálhatók és taníthatóak, hogy speciális zajszűrőként működjenek. A neurális hálózat alapú zajszűrők fontos jellemzője, hogy zajos bemeneti adatokon és jelentős zajjal rendelkező referencia (elvárt) képpárokon is taníthatóak. Kimutatták, hogy mindazonáltal egy megfelelően felépített, ily módon képzett neurális hálózat alapú szűrő képes a referenciaképeknél jobb jel-zaj aránnyal rendelkező szűrt képeket szintetizálni, és felülmúlja a hagyományos zajszűrőket, például a BM3D módszert, akár Gauss-, akár Poisson-zaj esetén [1]. Tekintettel arra, hogy a szűrők tanításához használt képadatbázis a valós betegadatok nehéz hozzáférhetősége miatt általában igen korlátozott számú képből áll, különösen fontosnak tartjuk a neurális hálózat alapú képfeldolgozó algoritmus robusztusságának vizsgálatát, azaz a képek zajtartalmára és annak különböző szempontok szerinti eloszlására való érzékenységet, beleértve a beteg korát, nemét, testtömegindex-értékét, valamint a jellegzetes patológiai struktúrák jellegét és eloszlását a képen. Ez az elemzés feltárhatja egy ilyen képfeldolgozó algoritmus robusztusságát, akár önmagában, akár egy nagyobb CADx (számítógépes diagnosztika) rendszer részeként, amelyet a világ klinikáin alkalmaznak.

A csontszcintigráfia gyakori indikációja a különböző daganatok csontmetasztázisainak kimutatása és nyomon követése, ezért továbblépésként egy olyan szoftver kifejlesztésébe kezdtünk, amely a kóros halmozódások keresésére szolgál. Ez az eszköz lehetővé teszi majd a képjavító eszköz diagnosztikára gyakorolt hatásának számszerűsítését is, így a képalapú metrikák mellett egy alkalmazásspecifikus, lézió alapú értékelés is elvégezhető. A kóros dúsulás felderítésére és előrejelzésére szolgáló szoftverkomponens létrehozásának egyik módszere a konvolúciós neurális hálózatok használata. Ehhez kapcsolódóan egy új szegmentációs metrikát, a "wave loss"

nevű módszert fejlesztettük ki. Disszertációmban bemutatom, hogy a veszteségfüggvénybe beépített topológiai információ felhasználható a szegmentáló hálózatok pontosságának növelésére.

A fentiek fényében a következő kutatási kérdésekre keresek választ:

- Használhatunk-e mély tanulást a jó minőségű, megbízható zajszűrésre a síkbeli csontszcintigráfiában?
- Mennyire robusztus egy ilyen megoldás a való életben?
- Hogyan kell kiértékelni egy ilyen eszköz teljesítményét, ha nem rendelkezünk zajmentes, tökéletes képekkel, amelyek összehasonlítási alapként szolgálnak?
- Lehetséges-e olyan veszteségfüggvényt létrehozni a neurális hálózatok tanításához, amely figyelembe veszi a szegmentációk topográfiáját is, ahelyett, hogy csak pixel-szintű összehasonlításokat végezne?

2. Új tudományos hozzájárulások és tézisek

Tézis I a: *Sikerült kifejlesztenem egy robusztus, magas minőségű zajszűrő módszert a planáris csontszcintigráfiához, ami bemutatja a mélytanulás hatékonyságát ebben a kontextusban. Valós páciens adatbázison történő validáció során, melyet teljesítményértékelés céljából választottunk ki, a legjobban teljesítő neurális hálózat átlagos RMSE értéke 1,15 volt normál statisztikai feltételek mellett, túlszárnyalva a legjobb nem neurális hálózatalapú BM3D-nek hívott megoldást, ami 1,29-es átlagos hibát ért el. Továbbá, 1/3-os statisztikájú felvételek mellett a neurális hálózat átlagos RMSE értéke 1,38 volt, meghaladva a legjobb nem neurális hálózatalapú Gauss 9mm megoldást, ami 2,07 RMSE-t ért el. Ezek az eredmények bizonyítják a neurális hálózatok kiváló zajszűrő képességeit az egészségügyi gyakorlatban alkalmazott hagyományos, nem neurális hálózatalapú módszerekkel szemben. Kapcsolódó publikáció: [2]*

Ebben a disszertációban egy neurális hálózat alapú zajszűrőt javaslok, amely többféle zajszintű csontszcintigráfiai felvételek esetén is használható, ahelyett, hogy minden egyes zajszinthez külön hálózatot kellene kifejleszteni. A javasolt zajszűrési megoldás egy konvolúciós neurális hálózat (CNN), amelyet az U-NET architektúra ihletett. A zajszűrő hálózatot a szabványos protokoll szerinti valós statisztikájú

csontszcintigráfiai felvételek segítségével képeztük ki, zajmentes felvételek nélkül.

A neurális hálózatok tanítására alkalmas adatok kiválasztása a Q-Bot szoftverrel történt. [3] A fejlesztéshez 2430 anonimizált felvételt használtunk (1215 páciénstől, elülső és hátsó), amelyeket AnyScan^{tiny®} DUO/TRIO SPECT/CT (Mediso Ltd.) és InterViewTM feldolgozó SW (Mediso Ltd.) segítségével készültek. Minden betegnek 5-600MBq Tc-99m metilén-difoszfónátot (MDP) (Isotope Institute LTD, Budapest, Magyarország) adtak intravénásan, 2-5 órás felhalmozódási idővel. A mátrix mérete 256x1024 volt, 130 mm/perc pásztázási sebességgel.

Minden felvételt binomiális eloszlás alapján újramintavételeztünk, mesterségesen létrehozva a valóságnak megfelelően degradált felvételeket, mintha azok egyharmad, egynegyed, egynolcad stb. felvételi idővel készültek volna.

Ha ezt a degradált képet kivonjuk az eredeti mérésből, akkor egy független felvételt kapunk, amelynek statisztikája jobb, mint a bemeneti oldalon lévő adatoké. Ez egy nagyon hatékony augmentációs technika is egyben, amellyel nemcsak a bemeneti oldalt, hanem a kimeneti oldalt is megváltoztatjuk.[4]

A neurális hálózat feladata a két generált felvétel közötti transzformáció becslése. A [5] cikkhez hasonlóan a tényleges szűrt és a referencia kép közötti tanulási veszteség függvényeként a Mean Absolute Error (MAE) értéket használtuk.

A neurális hálózatok tanításához az általánosan alkalmazott módszert használtuk: a backpropagation algoritmust és a gradiens alapú optimalizálást [6].

Az így kapott megoldás a vizsgált határértékeken belül robusztusnak bizonyult a képek zajszintjével szemben. Az értékelés során a hálózatok teljesítményét a Gauss-szűrőkkel és a mediánszűrőkkel, valamint a Block-matching and 3D filtering (BM3D) szűrővel hasonlítottuk össze. Kimutattuk, hogy a zajszűrő neurális hálózatokkal különösen magas jel-zaj arány érhető el, amelyek robusztusabbak a hagyományos módszereknél, és segíthetik a diagnózist, különösen a nagy zajtartalmú képek esetében.

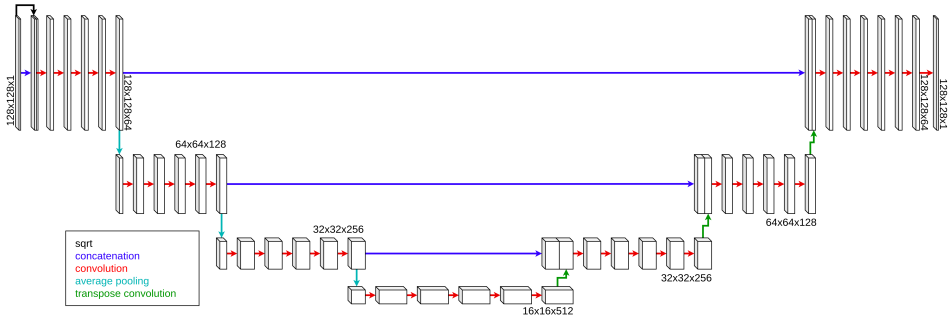


Figure 2.1: A zajszűréshez használt hálózati architektúrák mindegyike U-NET alapú volt. Az architektúrák közötti különbségek a szintenkénti konvolúciós rétegek számában és a szűrők számában voltak. A *L-NN* nevű hálózat 4 szintet tartalmazott, és mindig 5 konvolúciós réteg követte egymást. Az egyes szinteken használt konvolúciós blokkok szűrő darabszámai 64, 128, 256 és 512 voltak. A *S-NN* nevű neurális hálózat szintén 4 szintet tartalmazott, mindig 3 konvolúciós réteg követte egymást, és az egyes szinteken használt szűrők száma 16, 32, 64 és 128 volt.

Az eredményeket RMSE-mérőszám alapján bemutató 2.1 táblázatból látható, hogy minden statisztika esetében a neurális hálózaton alapuló megoldások érték el a legjobb eredményeket. Megjegyzendő, hogy normál és 1/3 statisztika mellett, ezen a metrikán a BM3D és a Gauss-szűrők teljesítménye összehasonlítható a neurális hálózatéval, de rosszabb statisztikák esetén e megoldások teljesítménye használhatatlan szintre romlik.

Tézis I b: *A planáris csontszcintigráfia mélytanuláson alapuló zajszűrő megoldásom megbízható valóélet-beli alkalmazás esetén is. Különböző homogén és torzított validációs adathalmazokon (beleértve a különböző korcsoportokat, BMI-tartományokat és nemi kategóriákat) végzett átfogó vizsgálattal értékeltem ki a zajszűrő algoritmus teljesítményének változékonyságát. A normál statisztikákkal rendelkező adathalmazokban az átlagos RMSE hiba 1,05 és 1,28 között mozgott különböző részhalmazokban, a vegyes adathalmaz átlagos hibája pedig 1,15 volt. Emellett az egyes részhalmazokon belüli szórás 0,4 és 0,48 között változott, amely kiemeli a szűrő algoritmus stabilitását*

	Statistics									
	normal		1/3		1/9		1/16		1/32	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
RMSE:										
BM3D	1.29	0.36	2.07	0.33	4.50	0.36	7.34	0.46	13.76	0.72
Gaussian 11mm	1.99	0.93	2.21	0.91	2.76	0.83	3.27	0.82	4.21	0.81
Gaussian 13mm	2.34	1.21	2.48	1.19	2.85	1.12	3.23	1.10	3.94	1.05
Gaussian 3mm	2.81	0.33	4.85	0.57	8.38	0.97	11.16	1.31	15.80	1.85
Gaussian 5mm	1.69	0.31	2.66	0.37	4.44	0.53	5.87	0.70	8.26	0.97
Gaussian 7mm	1.56	0.47	2.15	0.45	3.33	0.48	4.30	0.57	5.97	0.73
Gaussian 9mm	1.71	0.68	2.07	0.65	2.87	0.60	3.58	0.63	4.82	0.69
L-NN	1.15	0.40	1.38	0.41	1.80	0.47	2.09	0.54	2.54	0.63
Median 3px	1.79	0.47	2.48	0.45	4.36	0.42	6.81	0.37	11.70	0.82
Median 5px	2.64	0.91	3.06	0.87	4.50	0.72	6.71	0.55	11.51	0.82
Median 7px	3.57	1.30	3.91	1.26	5.15	1.08	7.17	0.83	11.82	0.93
Median 9px	4.57	1.67	4.90	1.63	6.01	1.45	7.85	1.17	12.32	1.13
S-NN	1.21	0.35	1.56	0.38	2.09	0.48	2.45	0.56	3.00	0.67

Table 2.1: A különböző szűrők teljesítménye az RMSE alapján számítva. A táblázatból látható, hogy minden statisztika esetében a neurális hálózaton alapuló megoldások érték el a legjobb eredményeket (legkisebb átlag és SD). Megjegyzendő, hogy normál és 1/3 statisztika mellett, ennél a metrikánál a BM3D és a Gauss-szűrők teljesítménye összehasonlítható a neurális hálózatéval, de rosszabb statisztikák esetén e megoldások teljesítménye használhatatlan szintre romlik.

és megbízhatóságát. 1/3 statisztikán is kiértékeljük, ahol az átlagos hiba 1,28 és 1,54 között volt, a szórás értékei 0,41 és 0,49 között mozogtak. Ezek az eredmények alátámasztják a mélytanuláson alapuló zajszűrőnk alkalmazkodóképességét és következetes teljesítményét különféle betegcsoportokban, megerősítve annak alkalmazhatóságát a való életben, a planáris csontszcintigráfiai képalkotásban. Kapcsolódó publikáció: [2]

Az értekezés állításainak alapjául szolgáló mérési eredményeket a 2.2 táblázat tartalmazza. A különböző halmazokon a statisztikák romlásának függvényében mért teljesítménytendenciák megegyeznek a vegyes halmazon megfigyeltekkel.

Thesis I c: *Sikerült kifejlesztem egy hatékony kiértékelési módszert a mélytanuláson alapuló, planáris csontszcintigráfia felvételek zajszűrésére szolgáló eszközhöz anélkül, hogy szükség lett volna zajmentes referencia képek használatára az összehasonlításhoz. A fejlesztés késői szakaszában kiválasztottunk egy neurális*

RMSE:		Statistics									
		normal		1/3		1/9		1/16		1/32	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
L-NN	Age High	1.28	0.48	1.54	0.49	2.00	0.54	2.32	0.58	2.82	0.67
	Age Low	1.05	0.41	1.28	0.41	1.68	0.47	1.96	0.54	2.41	0.64
	BMI High	1.09	0.43	1.31	0.44	1.68	0.47	1.94	0.50	2.33	0.57
	BMI Low	1.28	0.48	1.54	0.48	2.02	0.54	2.35	0.61	2.89	0.72
	Female	1.20	0.43	1.43	0.44	1.83	0.48	2.12	0.53	2.55	0.61
	Male	1.20	0.45	1.46	0.46	1.91	0.54	2.23	0.61	2.71	0.73
	Mixed	1.15	0.40	1.38	0.41	1.80	0.47	2.09	0.54	2.54	0.63
S-NN	Age High	1.35	0.43	1.74	0.46	2.31	0.54	2.70	0.60	3.29	0.71
	Age Low	1.12	0.35	1.47	0.38	1.99	0.46	2.34	0.54	2.88	0.65
	BMI High	1.16	0.39	1.47	0.41	1.93	0.48	2.26	0.53	2.73	0.61
	BMI Low	1.34	0.42	1.76	0.45	2.36	0.55	2.78	0.62	3.42	0.74
	Female	1.27	0.38	1.62	0.41	2.14	0.48	2.50	0.55	3.02	0.64
	Male	1.27	0.41	1.65	0.44	2.21	0.56	2.59	0.63	3.17	0.77
	Mixed	1.21	0.35	1.56	0.38	2.09	0.48	2.45	0.56	3.00	0.67

Table 2.2: A neurális hálózaton alapuló szűrők teljesítménye RMSE alapján különböző validációs halmazokon. A nagyobb neurális hálózat teljesítménye minden halmaz és statisztika esetében jobb, mint a kisebb neurális hálózaté. A különböző halmazokon a statisztikák romlásának függvényében a teljesítmény tendenciái meg-egyeznek a egyes halmazon mutatkozó tendenciákkal.

hálózatot, amely kielégítő teljesítményt mutatott alacsony zajtartalmú mérések esetében. Ezzel a kiválasztott neurális hálózattal létrehoztunk egy 544 mérést tartalmazó zajmentes validációs adathalmazt, majd ezeket a képeket orvosokkal vizsgáltattuk ki, hogy azonosítsák az esetleges rendellenességeket, szokatlan struktúrákat, felhalmozódásokat vagy műtermékeket az eredeti képekhez viszonyítva. Kiértékelési keretrendszerünkben ezeket a szűrt képeket zajmentesnek tekintettük, ideális képekként. Ezeket a "zajmentes" képeket normál statisztikai feltételek mellett Poisson-zajjal terhelt képek létrehozására használtuk, amelyek a megoldásaink bemenetül szolgáltak. Továbbá, kihasználva ezeket a mesterségesen létrehozott, szabványos méréshez hasonló képeket, binomiális mintavételezéssel rosszabb statisztikájú felvételeket állítottam elő az összehasonlító értékelésekhez. Kapcsolódó publikáció: [2]

A teljes folyamat és a folyamat által előállított képek példái a 2.2 ábrán és a

2.3 ábrán láthatók.

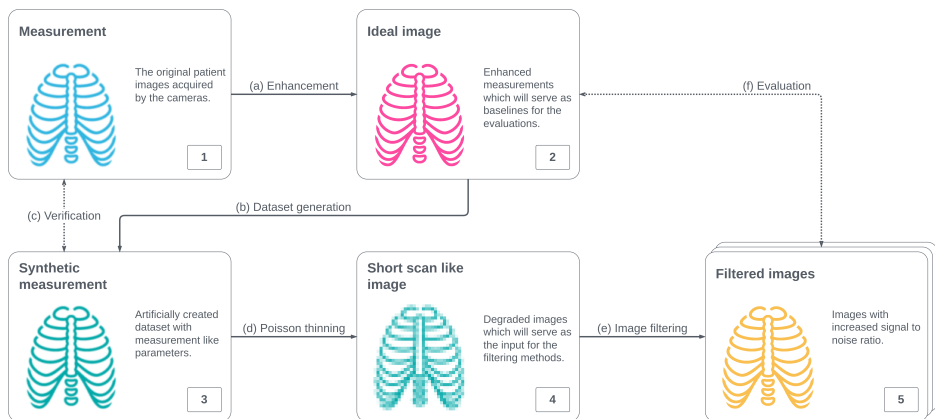


Figure 2.2: Kiértékelési folyamat: A kamera által készített valós mérésekből indulunk ki (1). A második lépés egy zajmentes kép (2) létrehozása egy referencia javító megoldással (a), amely esetünkben egy neurális hálózat alapú zajszűrő volt [4]. Az ideális képet ezután az orvosok megvizsgálják, hogy volt-e a képen bármilyen szokatlan struktúra, felhalmozódás vagy artefaktum. Ebből a zajmentes ideális képből szintetikus mérést (3) generálunk poisson-zaj hozzáadásával (b), amelyet statisztikai tesztekkel ellenőrizünk (c). A következő lépés a rosszabb statisztikájú mérések (4) konstruálása Poisson ritkítással (d). Végül ezek a képek lesznek a különböző szűrőeszközök bemenetei (e), amelyek eredményeit (5) összehasonlítjuk (f) az ideális képekkel (2).

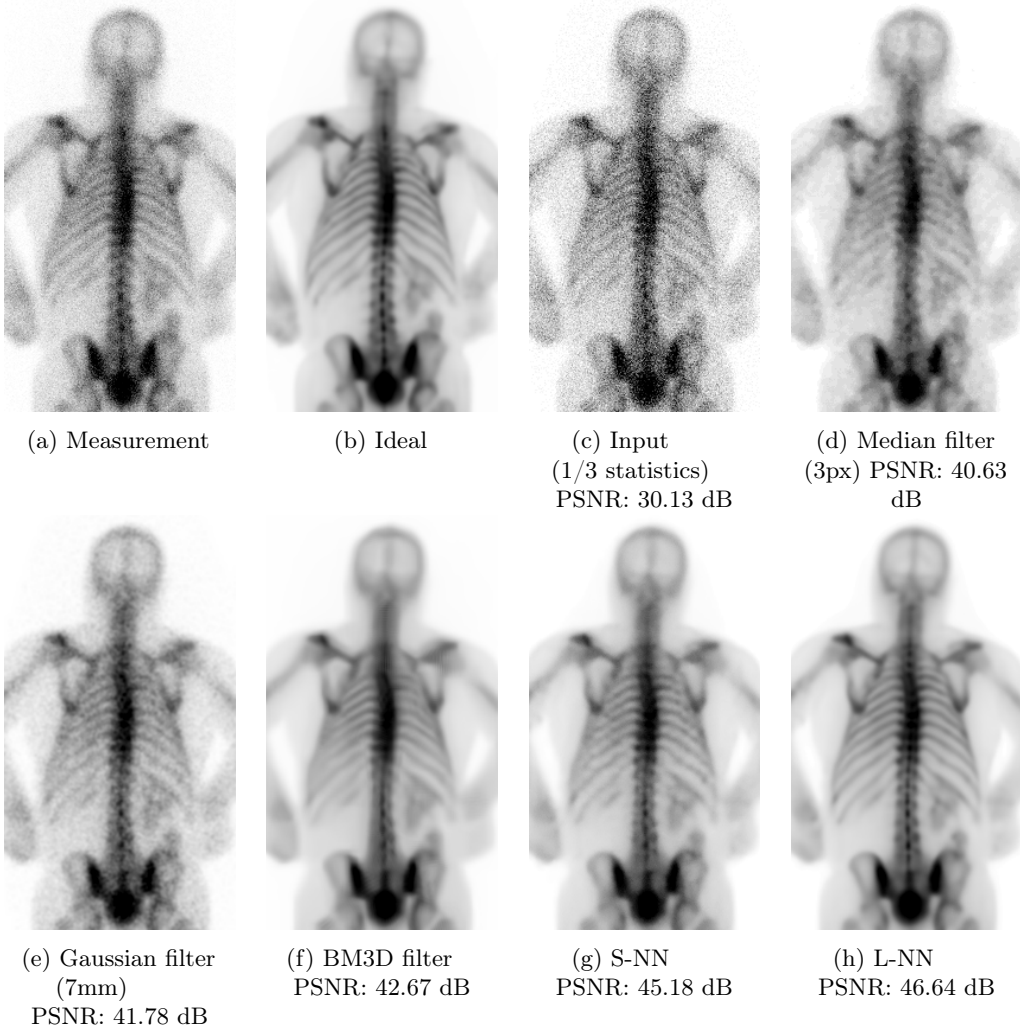


Figure 2.3: Kiértékelési folyamat: Valós mérésből indulunk ki (a), amelyből egy referencia szűrővel (b) zajmentes képet készítünk. Ezt azután az orvosok felülvizsgálják, és viszonyítási alapnak tekintik. Ebből létrehozunk egy mesterségesen degradált zajos képet (c). A (d), (e), (f), (g), (h) képek a különböző szűrők eredményeit mutatják. Mivel rendelkezésünkre áll a zajmentes referenciakép, a metrikák segítségével megfelelően tudjuk kiszámítani az egyes módszerek hibáit.

Tézis II: *Kifejlesztettem egy olyan neurális hálózatokhoz használható tanítási hibafüggvényt, amely figyelembe veszi a szegmentációk topográfiai szerkezetét, szemben a csak pixel-szintű összehasonlításokkal. A javasolt módszer mind a Cityscapes, mind az MS-COCO adathalmazokon a kereszt-entrópiához képest 3%-kal növelte a szegmentálási pontosságot, különböző hálózati architektúrákat használva. Kapcsolódó publikáció: [7]*

A szegmentálási problémák mély neurális hálózatokkal történő megoldása jól definiált veszteségfüggvényt igényel az összehasonlításhoz és a hálózat tanításához. A legelterjedtebb megközelítésekben jellemzően csak a terület alapú különbségeket veszik figyelembe, míg a topológiát, azaz a pixelek térbeli eloszlását nem veszik figyelembe. Agyunk könnyedén képes összetett objektumok összehasonlítására, és egyszerre veszi figyelembe a pixelszintű és a topológiai különbségeket. Az objektumok összehasonlításához egy megfelelően definiált metrikára van szükség, amely az alak és értékbeli változásokat egyaránt figyelembe véve határozza meg a hasonlóságot. Az elmúlt években a topográfiai szempontokat veszteségfüggvényekbe építették be, ahol vagy a határpixelet, vagy a területek arányát alkalmazták a különbségszámításban. Munkánk során bemutattuk, hogy a neurális hálózatok tanítása során hogyan alkalmazható egy topográfiai metrika, az úgynevezett "wave loss", és hogyan növelhető a hagyományos szegmentáló algoritmusok pontossága.

Az 1 algoritmus két szürkeárnyalatos kép hullámveszteségét számítja ki. A bemeneti értékek $Img1$ és $Img2$, az algoritmus kimenete pedig egy $WaveLoss$ skalárváltozó.

Algorithm 1: A hullámveszteség kiszámítása.

```

Data: Img1, Img2
Parameters : ValInc, SpaInc, SpaW, ValW
Result: WaveLoss
1 Union ← max(Img1,Img2);
2 CurrentWave ← min(Img1,Img2);
3 NewWave ← min(Img1,Img2);
4 WaveLoss = 0;
5 i ← 0;
6 num_iter ← int(1/ValInc);
7 while i ≤ num_iter do
    /* Loss for intensity differences */
8     NewWave += ValInc;
9     NewWave = min(NewWave,Union);
10    ValueChange = sum(NewWave-CurrentWave);
11    WaveLoss += ValW[i]*ValueChange;
12    CurrentWave = NewWave;
    /* Loss for spatial differences */
13    NewWave = maxpool(CurrentWave,[SpaInc,SpaInc], [1,1]);
14    NewWave = min(NewWave,Union);
15    SpatialChange = sum(NewWave - CurrentWave);
16    WaveLoss += SpaW[i] * SpatialChange;
17    CurrentWave = NewWave;
18    i += ValInc;
19 end

```

Az algoritmus paramétereit a következők:

- A *ValInc* elem határozza meg, hogy a hullám milyen gyorsan terjed az intenzitáskülönbségek mentén; minden egyes pixel intenzitása minden egyes iterációban ezzel az értékkel növekszik. Ez a paraméter fogja meghatározni a szükséges iterációk maximális számát is, és ez által a metszéspont legnagyobb távolságát is, ahol a topológiai különbségeket figyelembe vesszük. A hálózatban lévő neuronok maximális receptív mezőjénél nagyobb távolságot tartani logikátlan, mert így a hiba visszavezethető egy olyan neuronra, amelynek nem volt szavazata az adott bemeneti pixel osztályozásában. Kísérleteinkben ez az érték 0,05 és 0,1 között volt, ami azt jelenti, hogy egy kiválasztott pontból kiinduló hullám 20 és 10 pixelre terjedhetett.
- A *SpaInc* határozza meg a hullám térbeli terjedési sebességét. A térbeli terjedés egy max pooling művelettel valósul meg *SpaInc* ablakmérettel és egy lépésközzel. Szimulációinkban ezt az értéket mindig 3-ra állítottuk.
- A *ValW* az intenzitáskülönbségekre vonatkozó büntetések vektora. Ha ez az érték állandó, akkor a súlykülönbségek lineárisan arányosak lesznek a veszteségben szereplő büntetésekkel. Ha ez az érték növekvő, akkor ez azt jelenti, hogy a nagyobb különbségek (ahol több iterációra van szükség a kívánt érték eléréséhez) egyre nagyobb büntetéseket kapnak. Szimulációinkban állandó értékeket használtunk a *ValW*-ban.
- A *SpaW* egy vektor, amely a topográfiai különbségekért járó büntetéseket tartalmazza. A *SpaW*[0] azokat a pontokat súlyozza, amelyek egy térbeli terjedéssel elérhetők, és amelyek a metszéspont közvetlen szomszédságában vannak. *SpaW*[*k*] büntetést kapnak azok az értékek, amelyeket a *k*-edik iteráció során érünk el. Szimulációinkban lineárisan növekvő értékeket alkalmaztunk, amelyek mind alacsonyabbak voltak, mint a *ValW* értékei. A

legtöbb hálózatban átlagosan jó eredményeket szeretnénk elérni, de kisebb hibák az objektum alakját illetően elviselhetők. Az intenzitásúyoknál (*ValW*) alacsonyabb értékek alkalmazása azt jelenti, hogy a szegmentált objektum alakjának fontossága csökken. A monoton növekvő *SpaW* azt jelenti, hogy minél távolabb vagyunk az objektumtól, annál nagyobb költséggel jár a téves besorolás. A monoton csökkenő *ValW*-nél nagyobb súlyok alkalmazása azt jelentené, hogy a határok valóban fontosak, és egy határ körüli pixel osztályozása nagyobb problémát jelent, mint egy, az objektumtól távol eső pixel téves osztályozása.

Table 2.3: Ez a táblázat öt független futtatás átlagos pontossági eredményeit tartalmazza a Cityscapes adathalmazon négy különböző hálózati architektúra (sorok) és hat különböző veszteségfüggvény használatával a szemantikus szegmentáláshoz.

Model	L1 Loss	CrossEnt	Dice	Boundary	ShapeAware	Wave
SegNet	54.2%	57.0%	57.3%	57.7	58.6%	59.5%
DeepLab	59.7%	63.1%	64.1%	64.3%	65.4%	66.7%
DeepLabv3	77.6%	81.3%	81.4%	81.5%	81.7%	82.2%
HRNET	77.4%	81.6%	81.8%	81.8%	82.1%	83.4%

Table 2.4: Átlagos pontossági eredmények a COCO 2017 validációs halmazon ugyanazon hálózati architektúrák esetén, három különböző oszlopban lévő veszteségfüggvény (ℓ_1 , cross entropy, Dice veszteség, aktív határveszteség, shape aware loss és wave loss) használatával. A sorokban két különböző architektúra (ResNet-50 és ResNet-101) található. FPN esetekben a feature pyramid networks (FPN), C4 esetén a negyedik konvolúciós réteg (C4) aktiválását használtuk a régió-javaslatokhoz. Az eredmények az átlagos pontosságot mutatják minden objektumra, kivéve az utolsó három sort, ahol a legjobban teljesítő hálózat pontossági eredményei részletesen szerepelnek a kis, közepes és nagy méretű objektumok esetében is.

Model	L1	CrossEnt	Dice	Boundary	Shape	Wave
R50-C4 mAP all	28.75%	32.2%	32.83%	32.9%	34.721%	35.93%
R50-FPN mAP all	29.43%	35.2%	36.14%	36.12%	37.53%	38.11%
R101-C4 mAP all	30.17%	36.7%	37.2%	37.4%	38.86%	38.23%
R101-FPN mAP all	31.67%	38.6%	38.8%	39.3%	40.25%	41.7%
R101-FPN mAP s	14.25%	17.37%	18.18%	18.35%	19.33%	22.24%
R101-FPN mAP m	37.53%	39.23%	39.74%	40.52%	41.27%	43.26%
R101-FPN mAP l	50.14%	51.64%	51.83%	52.17%	52.22%	53.27%

A javasolt módszer mind a Cityscapes, mind az MS-COCO adathalmazok esetében 3%-kal növelte a szegmentálási pontosságot a kereszt-entrópiához képest, különböző hálózati architektúrákat használva.

2.1 A csontszcintigráfia zajszűrőjének perspektívái

A robusztussági vizsgálatot követően klinikai előtesztelést végeztünk orvosok bevonásával (ScanoMed Kft., Debrecen, Magyarország). A vizsgálat célja az volt, hogy a sok hasonló képpel dolgozó orvosok rámutathassanak az esetleges hibákra, műtermékekre, és véleményt mondhassanak az eszköz használhatóságáról. A ScanoMednél végzett 412 rutin csontszcintigráfiai egésztest-vizsgálat képein zajszűrést végeztek az disszertációban bemutatott AI-alapú alkalmazással. A betegek rutinszerűen 550-600 MBq ^{99m}Tc -MDP-t kaptak intravénásan, és az egész testről készült felvételeket 2 óra felhalmozódási idő után vették fel. A felvétel elkészülte után 1-2 percen belül elkészült a szűrt kép, amely segített az orvosoknak a további vizsgálatokról való döntésben, így ha bármelyik képen metasztázisra gyanús elváltozást találtak, SPECT/CT volt javallott. Amint arról a Referencia [4]-ban beszámoltunk, az orvosok párhuzamosan nézték az eredeti, normál statisztikájú, nem módosított képet és a kép zajszűrt változatát, és így értékelték a képeket. Az orvosok megállapították, hogy a neurális hálózaton alapuló szűrő nem törölt vagy generált új elváltozásokat, és nem azonosítottak műtermékeket a képeken. Arra a következtetésre jutottak, hogy könnyebb volt lokalizálni a rendellenességeket (bordákat, csigolyákat számolni), eldönteni, hogy szükség van-e további vizsgálatokra (SPECT/CT), és mindez felgyorsította magát a diagnózist.

Ez a kísérlet azt sugallja, hogy a zajszűrő használata hasznos a normális statisztikájú képek esetében, de további vizsgálatokra van szükség, hogy mennyire lehet csökkenteni a mérési időt vagy a beadott aktivitást az eredeti, megbízható diagnosztikai képesség megőrzésével. Ezért dolgoztunk a zajszűrő algoritmus komplex klinikai értékelésén, melybe integráltuk az elváltozás detektáló és osztályozó szoftverkomponenseket, hogy optimalizáljuk a teljesítményt a ROC (receiver operation curve) elemzés tekintetében. Jövőbeni célunk, hogy a bemutatott zajszűrő

megoldással a klinikai diagnosztikai értéket az érzékenység és a specificitás tekintetében még jelentősen alacsonyabb beadott aktivitás vagy mérési idő mellett is biztosítsuk.

3. A szerzőnek a szakdolgozathoz kapcsolódó folyóiratcikkei

- [4] Á. Kovács, G. Légrádi, A. Wirth, F. Nagy, A. Forgács, S. Barna, I. Garai, and T. Bükki, “A mesterséges és emberi intelligencia értéke a csontszcintigráfia példáján keresztül,” *Magyar Onkológia*, vol. 64, no. 2, pp. 153–158, 2020
- [2] A. Kovacs, T. Bukki, G. Legradi, N. J. Meszaros, G. Z. Kovacs, P. Prajczner, I. Tamaga, Z. Seress, G. Kiszler, A. Forgacs, S. Barna, I. Garai, and A. Horvath, “Robustness analysis of denoising neural networks for bone scintigraphy,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 1039, p. 167003, sep 2022.
- [7] Á. Kovács, J. Al-Afandi, C. Botos, and A. Horváth, “Wave Loss: A Topographic Metric for Image Segmentation,” *Mathematics* 2022, Vol. 10, Page 1932, vol. 10, p. 1932, jun 2022.

References

- [1] J. Lehtinen, J. Munkberg, J. Hasselgren, S. Laine, T. Karras, M. Aittala, and T. Aila, “Noise2Noise: Learning image restoration without clean data,” in *35th International Conference on Machine Learning, ICML 2018*, vol. 7, pp. 4620–4631, International Machine Learning Society (IMLS), 2018. 1
- [2] A. Kovacs, T. Bukki, G. Legradi, N. J. Meszaros, G. Z. Kovacs, P. Prajczner, I. Tamaga, Z. Seress, G. Kiszler, A. Forgacs, S. Barna, I. Garai, and A. Horvath, “Robustness analysis of denoising neural networks for bone scintigraphy,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 1039, p. 167003, sep 2022. 2, 2, 2, 3
- [3] F. Nagy, A. K. Krizsan, K. Kukuts, M. Szolikova, Z. Hascsi, S. Barna, A. Acs, P. Szabo, L. Tron, L. Balkay, M. Dahlbom, M. Zentai, A. Forgacs, and I. Garai, “Q-Bot: automatic DICOM metadata monitoring for the next level of quality management in nuclear medicine,” *EJNMMI Physics*, vol. 8, pp. 1–13, dec 2021. 2
- [4] Á. Kovács, G. Légrádi, A. Wirth, F. Nagy, A. Forgács, S. Barna, I. Garai, and T. Bükki, “A mesterséges és emberi intelligencia értéke a csontszcintigráfia

- példáján keresztül,” *Magyar Onkológia*, vol. 64, no. 2, pp. 153–158, 2020. 2, 2.2, 2.1, 3
- [5] N. Yuan, J. Zhou, and J. Qi, “Half2Half: deep neural network based CT image denoising without independent reference data,” *Physics in Medicine and Biology*, vol. 65, p. 215020, nov 2020. 2
- [6] C. C. Aggarwal, *Neural Networks and Deep Learning*. Springer International Publishing, 2018. 2
- [7] Á. Kovács, J. Al-Afandi, C. Botos, and A. Horváth, “Wave Loss: A Topographic Metric for Image Segmentation,” *Mathematics 2022, Vol. 10, Page 1932*, vol. 10, p. 1932, jun 2022. 2, 3