

ORVOSI KÉPEK JAVÍTÁSA MÉLYTANULÓ
ÉS TENZOR-FAKTORIZÁCIÓS
MÓDSZEREKKEL



Hatvani Janka
A PhD disszertáció tézisei

Témavezetők Dr. Gyöngy Miklós
Dr. Adrian Basarab
Prof. Denis Kouamé

Pázmány Péter Katolikus Egyetem
Roska Tamás Műszaki és Természettudományi Doktori Iskola
Université de Toulouse III
École Doctorale Mathématiques, Informatique,
Télécommunications de Toulouse

Budapest, 2021

Bevezetés

A pontos diagnózis felállítása érdekében az orvosnak gyakran az emberi test belsejébe kell látnia, hogy megérthesse a mögöttes folyamatokat. Annak érdekében, hogy elkerüljük a sebészi beavatkozások okozta komplikációkat, a non-invazív orvosi képalkotó módszereket részesítjük előnyben.

A fog belső szerkezetét ellátó orvoslás jó példája annak, hogy milyen fontosak a jó minőségű felvételek a sikeres kezeléshez vagy a pontos diagnózis felállításához. A fogászatban a fog 3D struktúráját *cone beam* komputertomográfia (CBCT) segítségével vizualizálják, aminek a tipikus felbontása 500 μm körüli [1]. A gyökérkezeléshez a fogcsatorna pontos feltérképezése szükséges, azonban ezek a felvételek nem megfelelőek a 0.16-1.6 mm-es struktúrák megjelenítéséhez [2]. Annak ellenére, hogy a gyökérkezelés a fogászat egyik leggyakoribb beavatkozása, az eseteknek csak a 60-85%-ban sikeres [3].

A felbontás-javító módszereket a tézis fogászati CT-felvételek segítségével mutatja be, azonban az ismertetett módszerek más képalkotó eljárás esetén is alkalmazhatóak.

Az orvosi képek javításának kihívásai

Jobb felbontású képek felvételéhez vagy új eszközökre van szükségünk (például nagyobb felbontású detektorokra), vagy a meglévő eljárást kell egészségügyi kockázatot eredményező beállítással (például magasabb sugárdózissal) használni, de az is előfordulhat, hogy fizikai korlátokba ütközünk (mint a diffrakciós limit esetében). Mindezek fényében az utólagos képjavító algoritmusokat részesítjük előnyben.

A felvételek degradációs modellje szerint a jó felbontású képnek egy elmosódott, zajos, alulmintavételezett verzióját látjuk. A legáltalánosabb szuper-rezolúciós algoritmusok egyetlen képből becsülik meg a jó minőségű képet (SISR), szemben a kép-sorozatot, vagy több típusú képet használó módszerekkel [4]. 2-dimenziós képek javítására léteznek korszerű, hatékony algoritmusok. Azonban az orvosi képek általában 3-dimenziósak, és a kép mérete ellehetetleníti ezen technikáknak például egy orvosi vizsgálat során történő alkalmazását, hiszen akár azok órákig is futhatnak [5].

A fentiek tükrében a vizsgált központi kérdések a következők:

1. Használható-e mélytanuló módszer a fogászati CT képek SISR javítására?
2. Hogyan alkalmazható a tenzor-faktorizáció a 3D SISR problémában?

3. Gyorsabbak-e a tenzor-implementációs megoldások a 3D SISR esetén, mint a korszerű technikák?
4. Megbecsülhetőek-e a rendszer paraméterei a 3D SISR keretrendszerében?

Új tudományos eredmények

I. Tézispont: *Az SISR probléma megoldására egy mélytanulási módszeren alapuló keretrendszert terveztem, amelyet fogászati CBCT szeletekre alkalmaztam. Az U-net és subpixel típusú neurális hálókat teszteltem, amelyek egyaránt javították a CBCT képek csúcs jel-zaj viszonyát (PSNR) 21-22 dB-lel, és a fogcsatorna szegmentációjának Dice-koefficiensét 1-2.2%-kal – hangsúlyosabban az orvosi kritikus gyökércsúcsnál.*

Kapcsolódó publikáció: [J1]

A konvolúciós neurális hálózatok (CNN) ígéretes eredményeket mutattak a felbontásjavítás terén [6]. Tudomásunk szerint ez volt az első alkalom, hogy mélytanuló algoritmust alkalmaztak az orvosi SISR probléma megoldására.

A U-net hálózat [7] 5 különböző méretskálán valósítja meg a kép tulajdonságainak feltérképezését, és ezeket kombinálja a kimenetén. A jelen implementációban *batch normalizációt* végeztem a háló generalizálásának érdekében, *leaky rectified linear unit* típusú aktivizációs függvényt használtam az inaktív neuronok elkerüléséhez, és a módosított Hubert-*loss* költségfüggvényt alkalmaztam a pontosabb tanításért. A subpixel hálózat [8] közvetlenül az alacsony felbontású kép tulajdonságait térképezi fel 6 rétegen keresztül, és a felbontás-növelést egy csempéző művelettel végzi el, az utolsó rétegben a mélységi információt szomszédságira cserélve. Ezek a módszerek, mivel 2D képekkel operálnak, gyors, de hatékony megoldást kínálnak az SISR problémára.

13 fogból szerzett, 5680 axiális szelet CBCT - μ CT képpárját használtam a hálózatok tanításához, 4 fog 1824 szeletét pedig a tesztelésükhöz. Két létező, rekonstrukció-alapú szuper-rezolúciós módszerhez (ℓ_2 és TV regularizációval) hasonlítottam a kapott eredményeket. Az eredmények értékeléséhez különböző metrikákat alkalmaztam – mint a PSNR, strukturális hasonlósági index (SSI) – , és ellenőriztem a javított képekből szegmentált csatornák pontosságát. A 1. ábrán példaeredmények láthatóak.

Az eredmények szerint a bemutatott CNN-alapú módszerek jobban teljesítenek a fogászati CT felvételeken, mint a jelenleg elterjedt módszerek (1. táblázat). Az orvosilag fontos tulajdonságokat (például a csatorna mérete, formája, görbülete) jobban mérhetővé teszi. A jelenleg alkalmazott költségfüggvény nem feltétlenül tűnt a legjobb választásnak, mert a mért eredmények nem tükrözték teljes mértékben az egyébként szemmel látható javulást (2. ábra).

1. táblázat. *Quantitative DL enhancement results*

A teszt-halmazon mért átlagolt metrikák, a μ CT-hez viszonyítva.

Metrika	CBCT	SRR: ℓ_2	SRR:TV	U-net	Subpixel
PSNR	45.56	64.15	64.80	67.58	66.60
SSI	0.9145	0.8688	0.8830	0.9304	0.9346
DoC	12.39%	12.25%	12.40%	10.12%	6.07%
DC	0.8891	0.8852	0.8913	0.8998	0.9101

II. Tézispont – a.: *A 3D SISR probléma meg-*

oldására olyan algoritmust terveztem, amely a tenzorok kanonikus poliadikus felbontását (CPD) használja. Ez az implementáció megőrzi a 3D objektum strukturális információját, egy költséghatékony és kevés paramétert használó algoritmusban ötvözve a tenzorfelbontáson alapuló zajszűrést, a PSF dekonvolúcióját és a kép felbontásának növelését. Ez a módszer a korszerű rekonstrukció-alapú algoritmusokhoz hasonló eredményeket két nagyságrenddel gyorsabban éri el – a PSNR 1.2 -1.5 dB-lel, a szegmentáció Dice-koefficiense 0.89-0.90-re nőtt.

Kapcsolódó publikáció: [J2]

A 3D-s tenzorok CPD felbontását a közelmúltban multi-, hiper-spektrális fúzióra alkalmazták [9]. Ez a felbontás azt a minimális számú tiszta tenzort (dimenzióként 1-1 vektor külső szorzataként felírható tenzort) találja meg, amelyek összegzése pontosan a keresett tenzort adja vissza. Kevesebb ilyen tenzor összege egy zajszűrt eredményt adhat.

2. táblázat. Kvantitatív CPD-SISR eredmények

	Minta #1	Minta #3
fog típusa	felső metszőfog	alsó örlőfog
μ CT kép mérete	282×266×392	324×306×402
CBCT PSNR	23.17 dB	24.14 dB
LRTV PSNR	24.32 dB	24.61 dB
CPD-SISR PSNR	24.32 dB	25.71 dB
CBCT DC	0.88 dB	0.90 dB
LRTV DC	0.87dB	0.90 dB
CPD-SISR DC	0.90 dB	0.91 dB
LRTV idő	6988 s	10301 s
CPD-SISR idő	71 s	104 s

A javasolt CPD-SISR algoritmus tiszta tenzorok azon halmazát keresi, amelyek a CBCT kép zajszűrte, dekonvolválta, felnagyított változatát közelítik meg. Mindezt egy egyesített egyenlettel teszi, amely csak a dimenziók között iterál. A tenzoros implementációnak köszönhetően – az eddigi algoritmusokkal ellentétben –, elkerülhető a hatalmas rendszer-mátrixok létrehozása ($X \in \mathbb{R}^{IJK \times IJK}$ mátrix egy $\mathbf{X} \in \mathbb{R}^{I \times J \times K}$ tenzor esetén), mégis megőrzi a 3D strukturális információt.

Az eredményeket (3. és 4. ábra) egy korszerű, rekonstrukció-alapú módszerhez hasonlítottam, amely alacsony rang- és TV-regularizációt használ (LRTV). A nagy rendszermátrixok következtében ez a módszer rendkívül lassú, egy példa $282 \times 266 \times 392$ képpontú térfogatot 2 óra alatt javítva fel, ami nagymértékben megnehezíti a nagy számú (6) paraméterkészlet inicializálását. A javasolt algoritmus ugyanezt a térfogatot kicsivel több, mint egy perc alatt javította fel, csupán 3, robosztus paraméter használatával. A PSNR hasonlóan javult a két módszerrel, míg a szegmentáció szignifikánsan jobb lett a CPD-SISR módszerrel (2. táblázat). Ezek az eredmények ösztönöztek engem további kutatásra a tenzor-faktorizáció terén, ahogy azt a következő tézispontokban is kifejtem.

II. Tézispont – b: *Egy algoritmusban egyesítettem a 3D kép CPD-SISR-alapú javítását és az ismeretlen PSF paraméterbecslését. A módszert egy korszerű rekonstruk-*

ció-alapú, 3D-SISR módszerhez hasonlítottam, amelyet az általam javasolt PSF-becsléssel egészítettem ki. Az eljárások hasonló javulást mutattak a PSNR szempontjából mind szimulált, mind valós CBCT képeken, de a javasolt módszer körülbelül 40-szer gyorsabban konvergált, kevesebb, mint 6 perc alatt.

Kapcsolódó publikáció: [C1]

3. táblázat. CPD-SISR-blind kvantitatív eredmények

	Szimuláció	Valós adat
μ CT mérete	$287 \times 266 \times 392$	$274 \times 278 \times 474$
valós $\bar{\sigma}$	[6.0 6.0 6.0]	–
inicializált $\bar{\sigma}$	[8.0 8.0 7.0]	[8.0 8.0 7.0]
$\bar{\sigma}$ LRTV-blinddal	[4.7 4.6 6.3]	[7.6 6.5 7.4]
$\bar{\sigma}$ CPD-SISR-blinddal	[5.0 4.9 4.8]	[8.5 7.8 6.5]
LR–HR PSNR	22.32 dB	19.42 dB
LRTV-blind PSNR	24.39 dB	25.63 dB
CPD-SISR-blind PSNR	26.53 dB	30.07 dB
LRTV-blind idő	9087 s	11823 s
CPD-SISR-blind idő	298 s	354 s

A PSF paramétereinek közvetlen becsléséhez vagy egy alacsony és magas felbontású képpárokat tartalmazó adathalmazra, vagy egy külön erre a célra tervezett fantomon végzett mérésre van szükség. Ezt minden egyes gépre meg kellene ismételni a pontos eredmények érdekében. Máskülönb a PSF-et a magas felbontású képpel együttesen kell megbecsülni.

Ebben a munkában egy félvak (*semi-blind*) módszert valósítottam meg, feltételezve, hogy a Gauss-függvény szórásértékei egy ismert tartományon belül vannak. A

leíró probléma derivált-alapú minimumkereséssel (*gradient descent*) megoldható [10]. Ezen optimalizáció, és a CPD-SISR között iterálva kovergálhatunk a megoldás felé.

A módszer értékeléséhez az eredményeket (5. ábra) a II. a pontban használt LRTV algoritmushoz hasonlítottam, amelyet ugyanezzel a *gradient descent* alapú PSF-becléssel egészítettem ki, hasonlóan iterálva a két lépés között (LRTV-blind). Szimulációban a PSNR 18.9%-kal nőtt a javasolt, és 9.3%-kal az LRTV-*blind* algoritmusban, míg a valós adatokon 54.8% és 31.9% volt a javulás (3. táblázat). Mindazonban a legfontosabb eredmény, hogy a CPD-SISR-*blind* módszer 40-szer gyorsabban konvergált az LRTV-*blind*hoz képest.

II. Tézispont – c: *Java*slatot tettem a 3D-SISR probléma megoldására a Tucker-dekompozíció alkalmazásával (TD-SISR). Először a kép zajszűrését végzem el a Tucker-dekompozíció segítségével, így javítható az ezt követő dekonvolúció rosszul kondicionáltsága. A CPD-SISR módszerhez képest 10-szer gyorsabban fut az algoritmus. A zaj függvényében növekedik a PSNR (0.3 -3.5 dB) és az SSI (0.58 - 2.43%), és a szegmentáció pontosságát mérő Dice-koefficiens (2%). Az algoritmusban használt paraméterek a 2D SVD-alapú zajszűrő algoritmusokéhoz hasonlóak, így könnyebben állíthatóak a CPD-SISR-hez képest.

Kapcsolódó publikáció: [C2]

A TD a magasabb rendű általánosítása a 2D szín-

4. táblázat. TD-SISR szimuláció - kvantitatív eredmények

	Szimulált LR	CPD-SISR	TD-SISR
futásidő	-	17.96 s	1.86 s
	PSNR (dB)		
zaj nélkül	28.56	31.48	34.99
30 dB	28.45	31.17	34.39
25 dB	28.36	31.08	31.40
20 dB	27.98	30.01	29.33
	SSI [0, 1]		
zaj nélkül	0.9623	0.9680	0.9823
30 dB	0.9612	0.9650	0.9763
25 dB	0.9572	0.9595	0.9653
20 dB	0.9463	0.9453	0.9417
	Szegegmentáció 25 dB-n		
Dice	0.8976	0.9242	0.9425

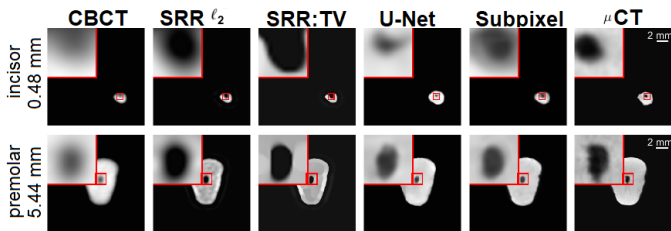
gularis érték felbontásnak (SVD) [11]. A bázisvektorok rendezhetőek a dekompozícióban betöltött súlyuk szerint. Míg a CPD egyetlen rangot definiál, amelyet a CPD-SISR-ban meg kell becsülnünk, a TD három ilyen értéket határoz meg 3D tenzorok esetén. A szinguláris értékek küszöbölésével egy szűrt tenzort kaphatunk. Ebben az algoritmusban a dekonvolúció csak egy külön lépésben valósítható meg.

Bár több paramétert használ az algoritmus, gyorsabb, és minőségileg jobb eredményeket adott a CPD-SISR-hez képest (5. és 4. táblázat, 6. ábra). $280 \times 268 \times 492$ és $324 \times 248 \times 442$ képpontos képeket 2 s alatt javított fel, általános Matlab implementációban. A PSNR 30, 25, 20 dB zajszintnél is nőtt mindkét módszerrel, de a TD-SISR jobban teljesített mind PSNR, mind SSI értékekben, a leg-

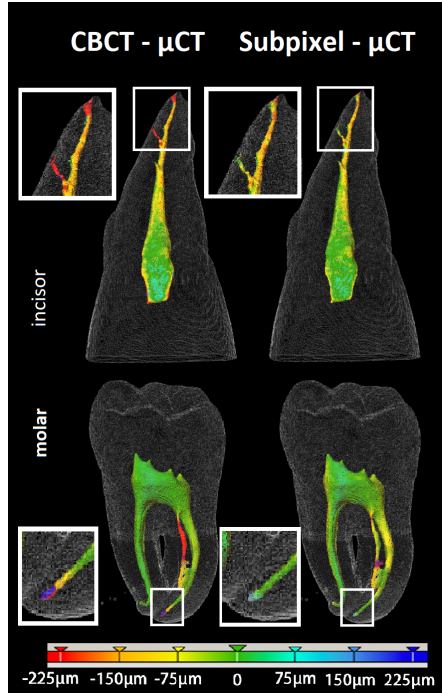
5. táblázat. TD-SISR valós adat - kvantitatív eredmények

	CBCTR	CPD-SISR	TD-SISR
futásidő	-	17.71 s	1.46 s
	PSNR (dB)		
zaj nélkül	19.55	21.25	21.61
30 dB	19.30	20.84	21.57
25 dB	19.10	20.13	21.09
20 dB	18.91	20.21	20.29
	SSI [0, 1]		
zaj nélkül	0.8647	0.8907	0.8935
30 dB	0.8610	0.8870	0.8929
25 dB	0.8478	0.8784	0.8908
20 dB	0.8173	0.8555	0.8814
	Szegmentáció 25 dB-n		
DC	0.8939	0.9189	0.9304

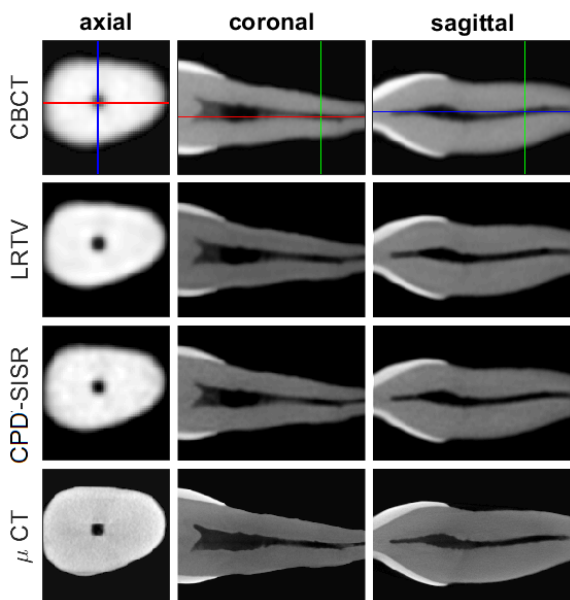
erősebb (20 dB-es) zaj esetét kivéve. A szegmentációt a 25 dB-es zajszintű képeken vágeztem el, és a Dice-koefficiens igazolta: a TD-SISR jobban teljesített. New Scientific Results



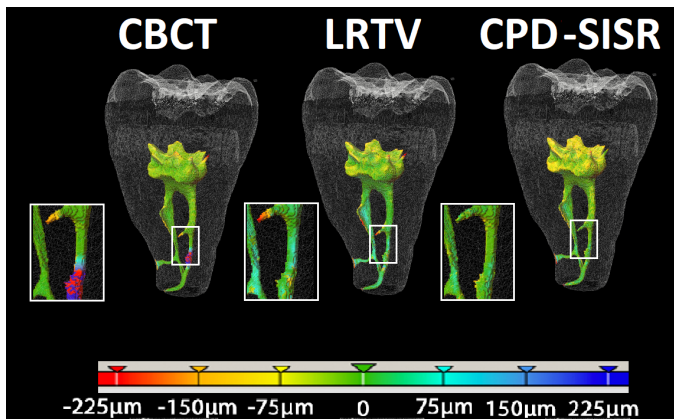
1. ábra. A teszthalmazból származó szeletek felbontásjavításának eredményei. Bal oldalon a fog típusa és a szelet gyökércsúcstól vett távolsága olvasható. A μ CT képeken egy 2mm-es méretskála látható. A dinamikus tartomány 0 és 1 közé lett konvertálva.



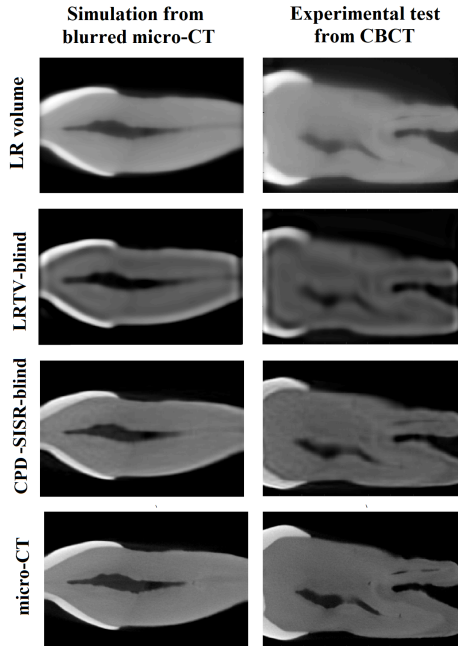
2. ábra. Csatorna-szegmentáció a teszt-halmaz két mintáján. A színezés a CBCT és μ CT (bal oldal), illetve a subpixel CNN és μ CT (jobb oldal) közötti eltérés mértékét mutatja.



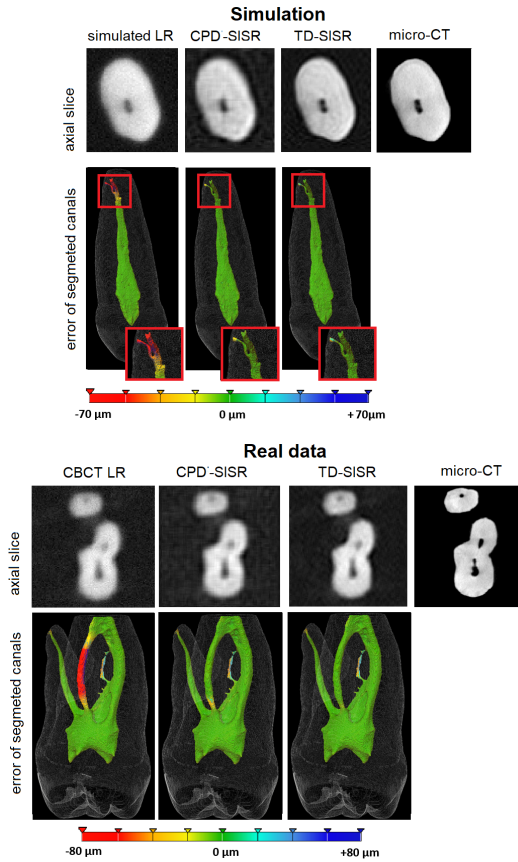
3. ábra. Eredmények az #1-es mintán. A CBCT kép a magasabb felbontáson szerepel a jobb összehasonlíthatóság kedvéért. A színes vonalak a szeletek egymáshoz való viszonyát mutatják a térfogaton belül.



4. ábra. Szegmentációs eredmények CBCT, LRTV és CPD-SISR esetén, a #3-as mintán. A színezés a szegmentációk közti eltérés mértékét reprezentálja.



5. ábra. Kvalitatív eredmények CPD-SISR-blind esetén. Egy koronális szeleten látható javulás. A rossz felbontású képek (mesterségesen rontott μ CT szimulációban, CBCT valós adat esetén) a μ CT képek méretén látszanak, lineáris interpoláció segítségével.



6. ábra. SISR eredmények 25 dB zaj alatt, szimulációban és valós adaton. Az első sorokban egy kiválasztott axiális szelet látható. A második sorokban a szegmentáció pontosságát jelöli a színezés.

A tézishez kapcsolódó publikációk

- [J1] J. Hatvani, A. Horváth, J. Michetti, A. Basarab, D. Kouamé, and M. Gyöngy, „Deep learning-based super-resolution applied to dental computed tomography,” *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 3, no. 2, pp. 120–128, 2018.
- [J2] J. Hatvani, A. Basarab, J.-Y. Tournieret, M. Gyöngy, and D. Kouamé, „A tensor factorization method for 3-d super resolution with application to dental ct,” *IEEE transactions on medical imaging*, vol. 38, no. 6, pp. 1524–1531, 2018.
- [C1] J. Hatvani, A. Basarab, J. Michetti, M. Gyöngy, and D. Kouamé, „Tensor-factorization-based 3D single image super-resolution with semi-blind point spread function estimation,” in *2019 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2019, pp. 2871–2875.
- [C2] J. Hatvani, A. Basarab, J. Michetti, M. Gyöngy, and D. Kouamé, „Single image super-resolution of noisy 3D dental CT images using tucker decomposition,” *arXiv preprint arXiv:2009.08657-eess.IV*, 2020.

A szerző egyéb publikációi

- [Au1] J. Hatvani, „Applicability of computed tomography methods for tomosynthesis problems,” in *PhD Proceedings Annual Issues of the Doctoral School, Faculty of Information Technology and Bionics, Pázmány Péter Catholic University – 2017*, 2017, pp. 48–48.
- [Au2] J. Hatvani, „Iterative optimization techniques for limited angle computed tomography,” *Phd Proceedings Annual Issues of the Doctoral School Faculty of Information Technology and Bionics, Pázmány Péter Catholic University – 2018*, vol. 13, p. 45, 2018.
- [Au3] J. Hatvani, „A review of super-resolution techniques on computed tomography images,” *Phd Proceedings Annual Issues of the Doctoral School Faculty of Information Technology and Bionics, Pázmány Péter Catholic University – 2019*, p. 46, 2019.
- [Au4] J. Hatvani, „A summary of tensor factorisation techniques for 3D image processing,” in *PhD Proceedings Annual Issues of the Doctoral School, Faculty of Information Technology and Bionics, Pázmány Péter Catholic University – 2020 In press*, 2020.

Irodalomjegyzék

- [1] D. Brüllmann and R. K. W. Schulze, „Spatial resolution in CBCT machines for dental/maxillofacial applications — what do we know today?” *Dentomaxillofacial Radiology*, vol. 44, no. 1, p. 20140204, 2014.
- [2] J. Martos, G. H. Tatsch, A. C. Tatsch, L. F. M. Silveira, and C. M. Ferrer-Luque, „Anatomical evaluation of the root canal diameter and root thickness on the apical third of mesial roots of molars,” *Anat. Sci. Int.*, vol. 86, no. 3, pp. 146–150, Sep. 2011.
- [3] Y.-L. Ng, V. Mann, S. Rahbaran, J. Lewsey, and K. Gulabivala, „Outcome of primary root canal treatment: systematic review of the literature—part 1. effects of study characteristics on probability of success,” *International endodontic journal*, vol. 40, no. 12, pp. 921–939, 2007.
- [4] K. I. Kim and Y. Kwon, „Single-image super-resolution using sparse regression and natural image prior,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 6, pp. 1127–1133, 2010.
- [5] F. Shi, J. Cheng, L. Wang, P. Yap, and D. Shen, „LRTV: MR image super-resolution with low-rank and total variation regularizations,” *IEEE Trans. Med. Imag.*, vol. 34, no. 12, pp. 2459–2466, 2015.
- [6] C. Dong, C. C. Loy, K. He, and X. Tang, „Image super-resolution using deep convolutional networks,” *IEEE*

Trans. Pattern Anal. Mach. Intell., vol. 38, no. 2, pp. 295–307, Feb. 2016.

- [7] O. Ronneberger, P. Fischer, and T. Brox, „U-net: Convolutional networks for biomedical image segmentation,” in *Proc. 18th Int. Conf. MICCAI*. Munich, Germany: Springer, Oct.5–9 2015, pp. 234–241.
- [8] W. Shi, J. Caballero, F. Huszár, J. Totz, A. P. Aitken, R. Bishop, D. Rueckert, and Z. Wang, „Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network,” in *Proc. 29th IEEE CVPR Conf.*, Las Vegas, US, Jun. 2016, pp. 1874–1883.
- [9] C. I. Kanatsoulis, X. Fu, N. D. Sidiropoulos, and W. Ma, „Hyperspectral super-resolution: A coupled tensor factorization approach,” *IEEE Trans. Signal Process.*, vol. 66, no. 24, pp. 6503–6517, Dec 2018.
- [10] N. Zhao, Q. Wei, A. Basarab, D. Kouamé, and J.-Y. Tourneret, „Blind deconvolution of medical ultrasound images using a parametric model for the point spread function,” in *Ultrasonics Symposium (IUS), 2016 IEEE International*. IEEE, 2016, pp. 1–4.
- [11] L. R. Tucker, „Some mathematical notes on three-mode factor analysis,” *Psychometrika*, vol. 31, no. 3, pp. 279–311, 1966.

