

GENERIRANJE SINTETIČNIH CT SLIK IZ MR SLIK PODROČJA GLAVE IN VRATU Z UPORABO DIFUZIJSKIH MODELOV

ROK MARKO ŠTER, GAŠPER PODOBNIK, TOMAŽ VRTOVEC

Univerza v Ljubljani, Fakulteta za elektrotehniko, Ljubljana, Slovenija
rs7839@student.uni-lj.si, gasper.podobnik@fe.uni-lj.si, tomaz.vrtovec@fe.uni-lj.si

Računalniška tomografija (CT) je slikovna preiskava, ki se v klinični praksi standardno zajame v okviru načrtovanje radioterapije. V primeru raka območju glave in vratu (HaN) se pogosto zajame tudi magnetno resonančne (MR) slike za natančnejše orisovanje tumorjev in kritičnih organov. V zadnjem času se vse bolj uveljavlja radioterapija na podlagi MR-samostojnega pristopa, ki odstrani potrebo po zajemu CT slik in s tem izpostavljenost ionizirajočemu sevanju, vendar pa zahteva rešitev za generiranje sintetičnih CT slik na podlagi MR slik. Nedavne študije kažejo, da difuzijski modeli nudijo realistično generiranje slik z natančnimi anatomskimi podrobnostmi in manj artefakti kot generativne nasprotniške mreže. V tej študiji smo razvili difuzijski model za pretvorbo MR slik v sintetične CT slike za HaN področje. Naš pristop, ovrednoten na zbirki podatkov HaN-Seg, ki vključuje pare CT in MR slik istih bolnikov, doseže indeks strukturne podobnosti 92,2 %, vršno razmerje signal-šum 33,1 dB ter povprečno absolutno napako 35,3 HU. Model dodatno ovrednotimo tudi za segmentacijo kritičnih organov. Rezultati potrjujejo potencial uporabe difuzijskih modelov pri načrtovanju radioterapije.

DOI
[https://doi.org/
10.18690/um.feri.2.2025.3](https://doi.org/10.18690/um.feri.2.2025.3)

ISBN
978-961-286-960-1

Ključne besede:
radioterapija,
računalniška tomografija,
magnetna resonanca,
difuzijski modeli,
difuzija,
umetna inteligenca,
globoko učenje,
sintetične slike,
računalniško podprta
analiza medicinskih slik

DOI
[https://doi.org/
10.18690/um.feri.2.2025.3](https://doi.org/10.18690/um.feri.2.2025.3)

ISBN
978-961-286-960-1

Keywords:

radiotherapy,
computed tomography,
magnetic resonance,
diffusion models,
diffusion,
artificial intelligence,
deep learning,
synthetic images,
computer-assisted medical
image analysis

GENERATION OF SYNTHETIC CT IMAGES FROM MR IMAGES IN THE HEAD AND NECK REGION USING DIFFUSION MODELS

ROK MARKO ŠTER, GAŠPER PODOBNIK, TOMAŽ VRTOVEC

University of Ljubljana, Faculty of Electrical Engineering, Ljubljana, Slovenija
rs7839@student.uni-lj.si, gasper.podobnik@fe.uni-lj.si, tomaz.vrtovec@fe.uni-lj.si

In cancer radiotherapy, CT images are essential for planning, while MR images accurately delineate tumors and organs-at-risk, especially in the head and neck (HaN) region. MR-only radiotherapy, with which we generate synthetic CT images from MR data, removes patient radiation exposure. Recent studies indicate that diffusion models yield more realistic images with precise anatomical details and fewer artifacts than generative adversarial networks. In this study, we employ a diffusion model to translate MR images into synthetic CT images for the HaN region. Evaluated on the HaN-Seg dataset of paired CT and MR images of the same patients, our approach achieves a structural similarity index of 92.2%, a peak signal-to-noise ratio of 33.1 dB, and a mean absolute error of 35.3 HU, demonstrating its potential in radiotherapy planning. Validation was extended on a downstream task of organ-at-risk segmentation. Results demonstrate the potential of applying diffusion models into the radiotherapy workflow.



1 Uvod

Radioterapija raka na področju glave in vratu (HaN) zahteva natančno lokalizacijo tumorskega tkiva in kritičnih organov (OAR) na podlagi slik za načrtovanje obsevalnega zdravljenja, kar omogoča načrtovanje obsevanja z minimalnim vplivom na zdrava tkiva (Gregoire in dr., 2020). Računalniško tomografske (CT) slike so tradicionalno temelj za takšno načrtovanje, saj vsebujejo informacije o atenuaciji v tkivih, ki so ključne za izračun doznega načrta. V primerjavi s CT ponuja magnetna resonanca (MR) boljši kontrast mehkega tkiva, zaradi česar je neprecenljiva za natančno orisovanje meja tumorjev in OAR, še posebej v zapletenem območju HaN (Podobnik in dr., 2023). Ker je slikanje s CT povezano z ionizirajočim sevanjem (Podobnik in dr., 2024a), se je razvila ideja o tako imenovanem MR-samostojnem (MR-only) načrtovanju obsevanja, ki v celoti sloni le na uporabi MR slik (McDonald in dr., 2024). Ker izračun doznega načrta temelji na informaciji o elektronski gostoti, ki ni enostavno izračunljiva iz MR slike, se je uveljavila ideja o generiranju sintetičnih CT slik na podlagi MR slik, za kar se je uveljavil tudi izraz prevajanje med slikovnimi modalitetama (angl. image-to-image translation, I2I).

Primarni cilj te študije je implementirati in oceniti nov pristop za omenjeni problem, ki temelji na difuzijskih modelih, ki so se v zadnjem času uveljavili kot uporabna metoda globokega učenja. Na podlagi pregleda literature ugotavljamo, da difuzijski modeli predstavljajo znaten napredek v primerjavi s tradicionalnimi generativnimi nasprotniškimi mrežami, saj ponujajo bolj stabilen postopek učenja ter izboljšano realističnost slik in anatomsko natančnost z manj artefakti (Dayarathna in dr., 2024). V tej študiji se osredotočamo na primerjavo dveh različnih strategij I2I: 2D pristop generiranja 2D aksialnih rezin in 3D pristop generiranja manjših 3D koščkov slike. Uporabljamo javno objavljeno zbirko podatkov, ki izhaja iz Onkološkega inštituta Ljubljana: "HaN-Seg: The Head and Neck Organ-at-Risk CT and MR Segmentation Dataset" (Podobnik in dr., 2023). Nabor podatkov vključuje pare CT in MR slik istih bolnikov skupaj s referenčnimi segmentacijami OAR, ki so jih strokovni orisovalci ustvarili ročno. MR slike uporabljamo kot pogoj za generiranje sintetičnih CT (sCT) slik, ki se nato ocenjujejo glede na anatomsko natančnost v primerjavi z dejanskimi ciljnim CT slikami ter celotno kakovostjo slike. Realističnost sCT slik ocenimo na nalogi segmentacije slik, in sicer tako, da primerjamo referenčne segmentacije s segmentacijami, ki jih ustvari orodje TotalSegmentator (Wasserthal in dr., 2023), naučeno izključno na realnih CT slikah iz podatkovnih množic, ki ne vključujejo

zbirke HaN-Seg. V okviru analize primerjamo uspešnost segmentacije na realnih CT slikah iz baze HaN-Seg z rezultati segmentacije na generiranih sCT slikah. Obravnavani pristop je zanimiv z večih vidikov: I2I na podlagi MR slik je ključen gradnik MR-samostojnega načrtovanja radioterapije, ki izniči bolnikovo izpostavljenost sevanju tekom načrtovanja radioterapije in ponuja stroškovno učinkovito alternativo tradicionalnim pristopom slikanja z več modalitetami.

2 Metode

2.1 Zbirka slik

V tej študiji smo uporabili nabor podatkov HaN-Seg (Podobnik in dr., 2023), ki ga sestavljajo 3D pari CT in MR slik 56 oseb. Vsaka oseba je prestala CT in T1w MR slikanje na Onkološkem inštitutu v Ljubljani z namenom zdravljenja raka na področju HaN. Uporabili smo uradno razdelitev slikanih oseb na učno in testno množico, z 42 pari slik v učni in 14 pari slik v testni množici, ter poročali o rezultatih, izmerjenih samo na testni množici, s čimer poročani rezultati služijo kot merilo za primerjavo orodij za I2I na tej javno dostopni zbirki slik.

2.2 Predobdelava slik

Vsak par CT in MR slik je bil poravnan s kombinacijo uporabe toge in elastične poravnave, kot je podrobno opisano v študiji o variabilnosti opazovalcev, izvedeni na naboru podatkov HaN-Seg (Podobnik in dr., 2024a). Intenzitete CT slik so bile linearno oknjene med intenzitetama -1000 in 3000 HU ter nato preslikane na interval $[-1, 1]$. Čeprav preslikava vrednosti izven intervala na mejne vrednosti vodi do izgube dela informacij, je ta izguba zanemarljiva, saj odrezane intenzitete primarno prikazujejo artefakte ali šum v ozadju, ki ni primarnega pomena za našo analizo. Intenzitete MR slik smo linearno preslikali s tako imenovano *z-normalizacijo* (tj. na ničelno povprečje in enotino varianco), saj smo opazili, da takšna normalizacija bolje deluje za MR slike zaradi prisotnosti artefaktov zobnih implantatov, ki povzročajo intenzitete ekstremnih vrednosti in posledično izkrivljajo porazdelitev intenzitet v primeru uporabe normalizacije intervalov. Poravnane in normalizirane CT in MR slike so bile nato ponovno vzorčene na velikost vokslov $1 \times 1 \times 1$ mm³ in središčno obrezane v aksialni ravnini na mrežo velikosti $256 \times 256 \times Z$ slikovnih

elementov, kjer Z predstavlja število aksialnih rezin. Pomembno je omeniti, da se Z med slikami razlikuje zaradi razlik v vidnem polju originalnih slik.

2.3 Verjetnostni modeli difuzijskega odstranjevanja šuma

Uporabili smo difuzijski model, ki temelji na arhitekturi U-Net (Ronneberger in dr., 2015), razširjen z mehanizmi vgrajevanja časovnih korakov in pozornosti, ter razporejevalnik dodajanja šuma DDPM (angl. denoising diffusion probabilistic model) (Ho in dr., 2020), ki sta implementirana v knjižnici MONAI¹ Generative Models² (Cardoso in dr., 2022). Med učenjem model sprejme vhodno sliko šuma in sliko, s katero pogojujemo model pogoja. V našem primeru je to MR slika, ki vodi proces difuzije tako, da je sCT slika v anatomskem ujemanju z MR sliko. V vsakem koraku učenja modela se vsaki sliki CT v paketu učenja dodeli naključni celoštevilski časovni korak. Časovni koraki se vzorčijo iz enakomerne porazdelitve med 0 in N_{train, T_s} ter se uporabljajo za simulacijo različnih količin šuma, dodanega sliki. Vsak učni vzorec v paketu je sestavljen iz originalne CT slike z dodanim Gaussovimi šumom na podlagi naključnega časovnega koraka in načina delovanja razporejevalnika dodajanja šuma. Model se nato uči napovedovanja dodanega šuma na podlagi slike s šumom in MR slike. Čeprav obstaja veliko strategij za posredovanje informacij modelu, smo izvedli združitev slike šuma in MR slike na nivoju slikovnega kanala. Informacije pogojne slike, tj. MR, služijo kot vodilo za model, saj zagotavljajo informacije o anatomski strukturi in podrobnostih mehkega tkiva (Kazerouni in dr., 2023). Med inferenco se uporablja iterativni postopek odstranjevanja šuma: vhod modela sta naključni šum in MR slika. Ta vhod se nato v iterativnem postopku odstranjevanja šuma obdeluje z odštevanjem napovedane slike šuma, kar na koncu privede do realistične sCT slike, ki je anatomsko podobna vhodni MR sliki.

3 Eksperimenti in rezultati

3.1 Eksperimenti generiranja sintetičnih CT slik

Zasnovali smo dva poskusa, pri katerih smo zagotovili čim bolj dosledno strategijo učenja in nastavitve modela, hkrati pa smo omogočili primerjavo med pristopoma 2D generiranja rezin in 3D generiranja koščkov. Za difuzijski model smo uporabili

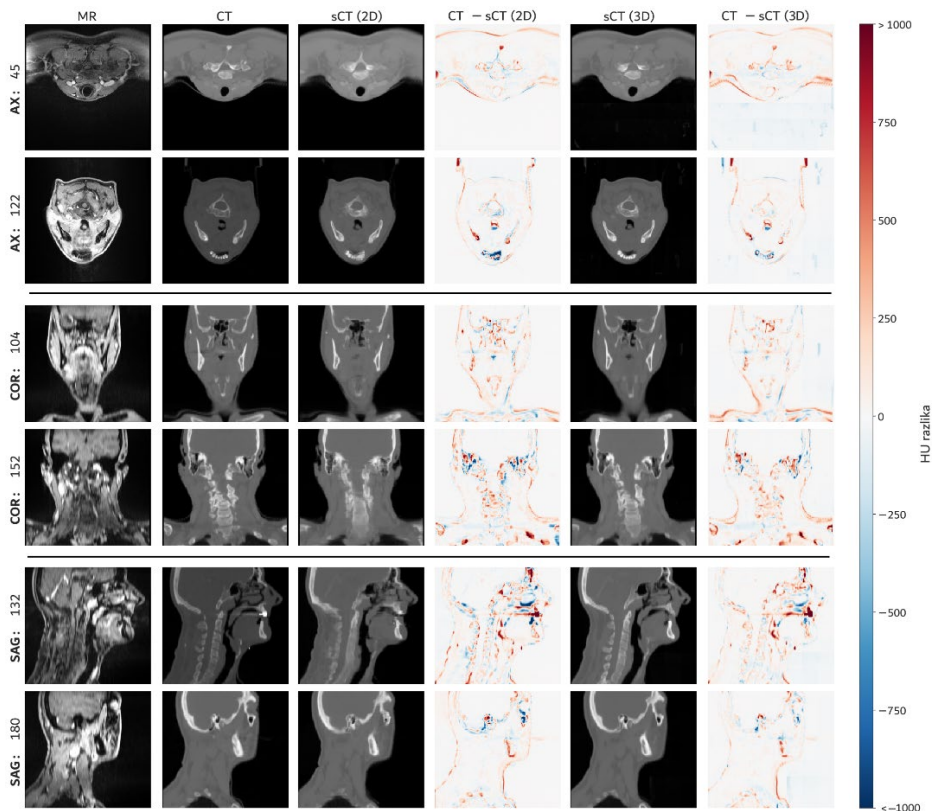
¹ <https://monai.io/>

² <https://github.com/Project-MONAI/GenerativeModels>

arhitekturo U-Net, ki vključuje vgradnjo časovnega koraka in mehanizem tako imenovane *samopozornosti* v dveh najglobljih plasteh mreže. Model 2D je bil naučen z aksialnimi rezinami velikosti 256×256 slikovnih elementov z uporabo mreže U-Net z [32, 64, 128, 256, 512] kanali značilnic, medtem ko je bil 3D model naučen s koščki velikosti $64 \times 64 \times 64$ slikovnih elementov z uporabo mreže U-Net s [64, 128, 256, 512] kanali značilnic. Pri uporabi 2D modela so bile generirane aksialne rezine kasneje enostavno zložene v pravilnem vrstnem redu za namen rekonstrukcije celotne 3D sCT slike. Uporabili smo kombinacijo kriterijskih funkcij L1 in srednje kvadratne napake (MSE) z empirično določenima utežema 0,75 in 0,25. Model smo učili z optimizatorjem AdamW z začetno vrednostjo učene konstante $1e-4$ in z razporejevalnikom dodajanja šuma DDPM z $N_{\text{train}_T_s} = 1000$ časovnimi koraki ter njegovimi sicer privzetimi parametri MONAI modula. V primerjavi z difuzijskimi implicitnimi modeli odstranjevanja šuma (Song in dr., 2021) in psevdonumeričnimi difuzijskimi modeli (Liu in dr., 2022) je razporejevalnik DDPM generiral najbolj anatomsko pravilne in smiselne rezultate za enako število časovnih korakov in epoh. Empirično smo ugotovili, da tehnike bogatenja podatkov, kot so rotacija, skaliranje, translacija in gama transformacija, niso izboljšale učinkovitosti modela, zato med učenjem končnih modelov tehnike bogatenja podatkov nismo uporabili. Med učenjem modela smo redno merili kvantitativne metrike na validacijski³. Ker je celotna inferenca z difuzijskimi modeli razmeroma računsko potratna, smo izvedli dve vrsti vrednotenja med učenjem, ki ju imenujemo enostavno in celovito vrednotenje. Enostavno vrednotenje je bilo izvedeno vsakih 25 epoh, ko smo preprosto izmerili vrednost kriterijske funkcije na slikah testne množice v naključno izbranih časovnih korakih, medtem ko je bilo celovito vrednotenje izvedeno vsakih 100 epoh in je vključevalo celoten postopek obratnega difuzijskega procesa (odstranjevanje šuma), in sicer z reduciranim številom časovnih korakov na 100. Končni model za vsak pristop je bil izbran na podlagi najnižje vrednosti kriterijske funkcije, izračunane pri enostavnem vrednotenju. To je pomenilo skupaj 7.800 in 29.700 epoh, ki so trajale približno 2 oziroma 4 dni za pristop z 2D rezinami in s 3D koščki. Modele z nalogo napovedi 2D aksialnih rezin smo učili na delovni postaji, opremljeni z grafično procesno enoto (GPU) Nvidia RTX 2080Ti z 12 GB pomnilnika, 3D modele pa je bilo treba učiti na GPU Nvidia RTX A6000 z 48 GB pomnilnika.

³ Dve sliki iz originalne učnega množice sta bili premaknjeni v tako imenovano validacijsko množico.

3.2 Vrednotenje sintetičnih CT slik



Slika 1: Primeri sintetičnih CT slik, ustvarjenih z našimi 2D in 3D modeli, pri katerih je število sklepanja nastavljeno na 100 na podlagi ustreznih vhodnih MR slik. V prvih dveh vrsticah so prikazani aksialni prerezi, v naslednjih dveh vrsticah koronalni prerezi, v spodnjih dveh vrsticah pa sagitalni prerezi. Poleg vsake ustvarjene slike je na voljo grafični prikaz razlik med pravo in sintetično CT sliko, pri čemer so razlike v Hounsfieldovih enotah označene z rdečo in modro barvo v skladu z legendo na skrajni desni strani.

Vir: lasten.

Dobljeni modeli so bili vrednoteni na 14 parih 3D slik iz testne množice z uporabo treh kvantitativnih metrik: vršno razmerje signal-šum (PSNR), indeks strukturne podobnosti (SSIM) (Wang in dr., 2004) in srednja absolutna napaka (MAE). PSNR in SSIM smo izračunali tako v 2D, tj. primerjava sintetičnih aksialnih rezin s pripadajočimi ciljnim aksialnimi rezinami, kot v tudi 3D. V nadaljevanju se okrajšavi PSNR in SSIM nanašata na njun 3D izračun, ki je običajen izračun na področju

obdelave 3D slik. Rezultati so navedeni v tabeli 1, na sliki 1 pa so prikazani primeri sCT.

Tabela 1: Metrike PSNR, SSIM in MAE pri generiranju sCT slik z 2D in 3D modeli.

Model	PSNR-2D (dB)	PSNR-3D (dB)	SSIM-2D (%)	SSIM-3D (%)	MAE (HU)
2D	36,9 ± 2,7	32,1 ± 1,2	90,8 ± 2,1	90,3 ± 2,1	33,1 ± 4,9
3D	35,1 ± 1,4	33,1 ± 1,4	93,0 ± 1,4	92,2 ± 1,4	35,3 ± 4,7

3.3 Vrednotenje uporabnosti sintetičnih CT slik

Pomembno merilo uporabnosti sintetičnih slik je njihova realističnost. Medtem ko kvantitativne metrike podobnosti služijo kot učinkovito orodje za primerjavo raznolikih metod, se ob tem pojavijo njihove specifične omejitve. Čeprav vizualno vrednotenje kakovosti ponuja celovitejši vpogled, temelji na subjektivnih ocenah opazovalcev ter je časovno in, v primeru uporabe strokovnjakov iz klinične prakse, tudi precej finančno potratno. V okviru prispevka smo zato uporabili alternativen pristop, kjer smo uspešnost segmentacije OAR na sCT slikah ocenili s pomočjo avtomatske metode, naučene izključno na realnih CT slikah, pridobljenih iz drugih podatkovnih zbirk. Za ta namen smo uporabili segmentacijsko orodje TotalSegmentator in z njim pridobili segmentacije osmih anatomskih struktur: levo oko, desno oko, levi optični živec, desni optični živec, leva parotidna žleza, desna parotidna žleza, hrbtenjača ter ščitnica. Rezultate segmentacije na sCT slikah smo nato primerjali z rezultati, doseženimi na realnih CT slikah.

Uspešnost segmentacije orodja TotalSegmentator smo ocenili z dvema metrikama: Diceov koeficient podobnosti (DSC) in 95. percentil Hausdorffove razdalje (HD95). Kot referenco smo uporabili ročne segmentacije iz uporabljene podatkovne zbirke HaN-Seg, kjer smo za vsako zgoraj navedeno strukturo izračunali razliko metrik, izračunanih na segmentacijah realnih in sintetičnih slik na naslednji način:

$$\Delta DSC = DSC_{sCT} - DSC_{CT}, \quad (1)$$

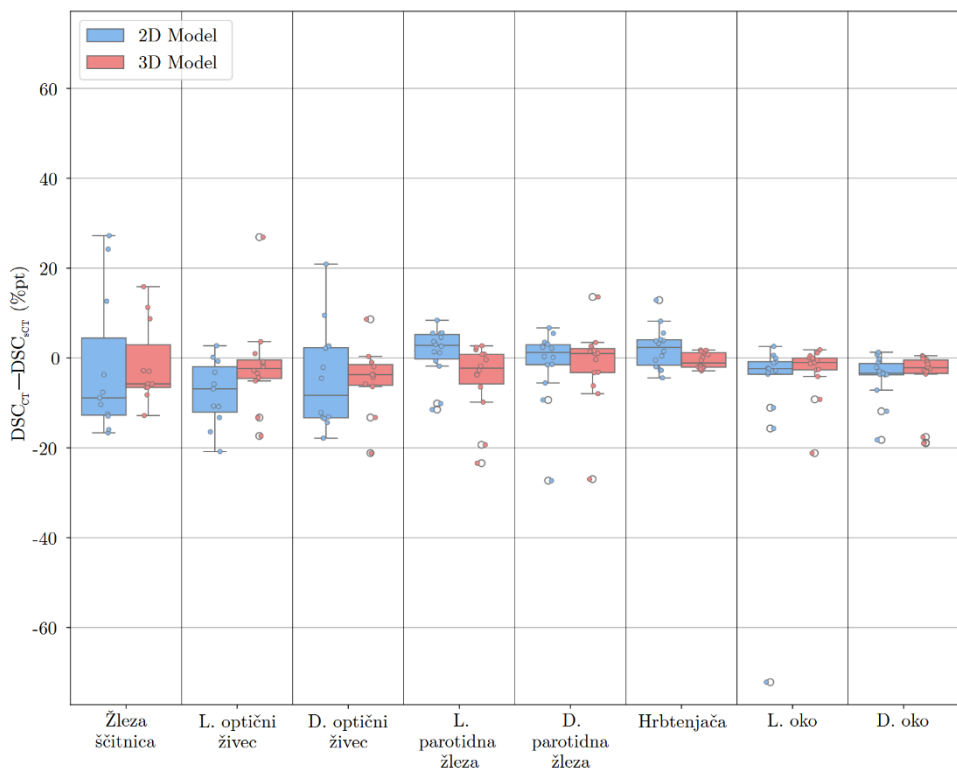
kjer je DSC_{sCT} vrednost DSC med referenčno segmentacijsko masko anatomske strukture in segmentacijsko masko iste strukture, ustvarjeno z orodjem TotalSegmentator na podlagi sCT slike. Vrednost DSC_{CT} pa predstavlja DSC med

referenčno segmentacijsko masko anatomske strukture in segmentacijsko masko, ustvarjeno z orodjem TotalSegmentator na podlagi realne CT slike.

Izračun $\Delta HD95$ poteka na podoben način:

$$\Delta HD95 = HD95_{sCT} - HD95_{CT}, \quad (2)$$

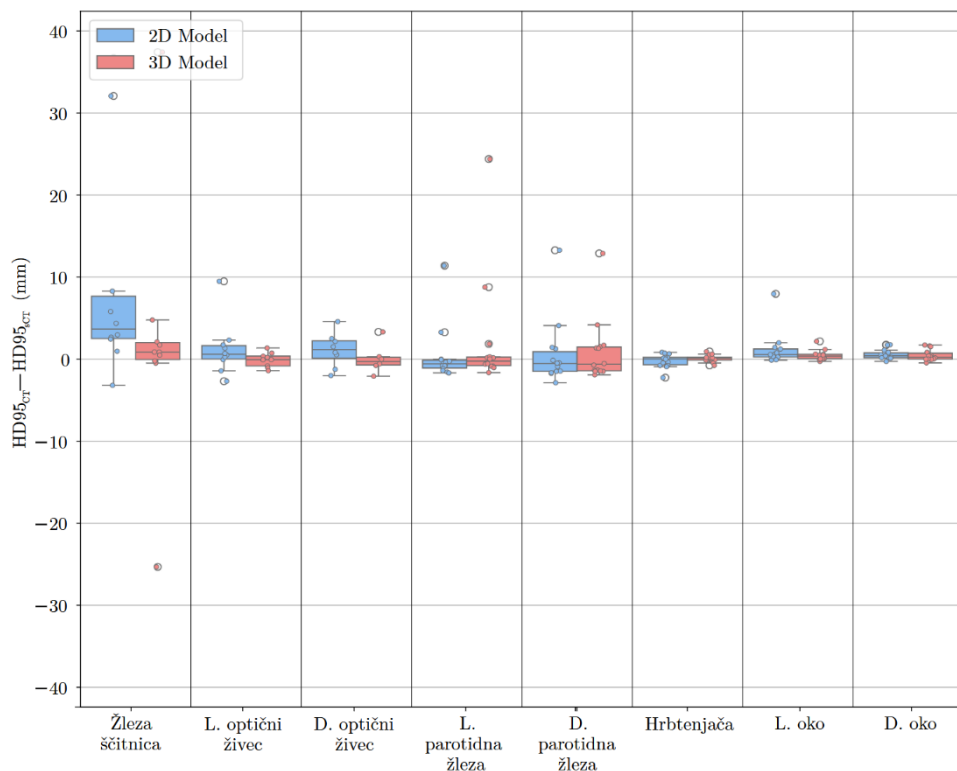
kjer gre za enak princip računanja $\Delta HD95$ kot pri izračunu ΔDSC , oznaki $HD95_{sCT}$ in $HD95_{CT}$ pa sledita istemu principu kot DSC_{sCT} in DSC_{CT} .



Slika 2: Grafični prikaz razlike v vrednosti DSC na segmentacijah sintetičnih in realnih slik, primerjanih z referenčnimi orisi. Barva na škatlastem diagramu prikazuje, kateri model smo uporabili za generiranje sintetičnih slik: modra predstavlja 2D model, rdeča pa 3D model.

Vir: lasten.

Rezultati te analize so prikazani na škatlastih diagramih na slikah 2 in 3. Pozitivne vrednosti Δ DSC pomenijo, da so segmentacije na podlagi sCT slik boljše, tj. bolj podobne ročnim segmentacijam, enako velja za negativne vrednosti Δ HD95. Modra barva označuje metrike 2D modela, medtem ko rdeča barva označuje metrike 3D modela. Vse primere, kjer orodje TotalSegmentator ni uspelo segmentirati anatomskih struktur, smo odstranili iz nadaljnje analize.



Slika 3: Grafični prikaz razlike v vrednosti HD95 na segmentacijah sintetičnih in realnih slik, primerjanih z referenčnimi orisi. Barva na škatlastem diagramu prikazuje, kateri model smo uporabili za generiranje sintetičnih slik: modra predstavlja 2D model, rdeča pa 3D model.

Vir: lasten.

4 Diskusija

Ta študija je prva, ki poroča o rezultatih I2I na javno dostopnem naboru podatkov HaN-Seg (Podobnik in dr., 2023). Čeprav je bil ta nabor podatkov objavljen predvsem za primerjavo algoritmov za segmentacijo OAR iz slik več modalitet na

področju HaN (Podobnik in dr. 2024b), je zaradi parov slik več modalitet primeren za razvoj in vrednotenje nalog I2I med modalitetami. Poročani rezultati lahko zato služijo kot merilo za prihodnje raziskave o ustvarjanju sCT slik v tem anatomskem območju. Naši rezultati potrjujejo potencial difuzijskih modelov za generacijo visokokakovostnih sCT slik področja HaN. Raziskali smo 2D in 3D pristop generiranja slik, pri čemer je naš 2D model dosegel PSNR 32,1 dB in SSIM 90,3 %, medtem ko je naš 3D model dosegel višji PSNR 33,1 dB in višji SSIM 92,2 %. Medtem ko je pristop z 2D rezinami ustvaril sintetične CT prereze visokih kvalitiet, je 3D pristop bolje ohranil anatomsko kontinuiteto v aksialnih prerezih, kot kažeta koronalni in sagitalni prerez na sliki 1. Čeprav je 3D model dosegel boljše vrednosti metrik, pa je zahteval bistveno več računske moči in daljši čas učenja. Kompromis med anatomsko skladnostjo in računsko učinkovitostjo poudarja potrebo po prilagojenih rešitvah glede na prednostne naloge v klinični praksi. Medtem ko lahko na primer 2D modeli zadostujejo za hitro izdelavo prototipov, so lahko 3D modeli nepogrešljivi za celovite delovne postopke, kot je segmentacija hrbtenice, za katero se zdi, da je anatomska struktura, s katero je imel poleg majhnih organov naš 2D model največ težav. Druga ugotovitev je, da kvantitativne metrike, zlasti MAE, morda ne odražajo v celoti kvalitativnih ugotovitev, prikazanih na sliki 1. Slika 1 namreč jasno kaže, da 2D modeli ustvarjajo nejasne prehode med aksialnimi rezinami in ne dosegajo pričakovane anatomske kontinuitete vzdolž sagitalne in koronalne smeri. Predvidevamo, da bi to omejitev lahko ublažili s pristopom z več rezinami, pri katerem so na primer kot vhodni podatki na voljo tri zaporedne rezine, difuzijski model pa ima nalogo ustvariti le srednjo rezino.

V naši implementaciji v eni učni epohi model ne vidi vseh rezin oziroma koščkov vsake slike iz učne zbirke, temveč se iz vsake slike sproti naključno vzorči 20 rezin ali 2 koščka. Prav tako velja omeniti dolgotrajen postopek inference, saj model potrebuje približno 6 minut, da ustvari vse 2D sCT aksialne rezine, potrebne za sestavo ene zložene 3D sCT slike. Ta čas inference je bil izmerjen z uporabo GPU RTX 2080Ti. Obsežno smo eksperimentirali z zmanjšanjem števila časovnih korakov inference in ugotovili, da se kakovost sintetične slike ni bistveno zmanjšala, če smo namesto 1000 časovnih korakov, uporabljenih za učenje, uporabili 100 časovnih korakov. S tem smo lahko čas inference zmanjšali približno za faktor deset.

Neposredna primerjava z drugimi študijami je zaradi razlik v naborih podatkov in anatomskih regijah težja in ni nujno popolnoma reprezentativna. Zavedajoč se teh omejitev primerjave poročamo, da so Yang in dr. (2020), ki so izvedli prevajanja iz MR v CT na podatkih o možganih, poročali o PSNR 25,2 dB in SSIM 80,1 %, Yan in dr. (2023), ki so izvedli prevajanje iz MR v CT na podatkih HaN, so dosegli PSNR 26,2 dB in SSIM 84,9 %, Graf in dr. (2023) so dosegli PSNR 27,9 dB in SSIM 88,7 % na podatkih o hrbtenici, Wolterink in dr. (2017) pa PSNR 32,3 dB na neparnih možganskih podatkih. Iz zgoraj navedenih rezultatov sorodnih študij lahko sklepamo, da so rezultati v naši študiji boljši od rezultatov v prej omenjenih študijah.

Na podlagi rezultatov na slikah 2 in 3 lahko trdimo, da so segmentacijske maske ustvarjene na podlagi sCT slik, generiranih z našimi modeli, močno podobne segmentacijskim maskam, ki jih isto orodje TotalSegmentator ustvari na podlagi resničnih CT slik. Orodje ima največ težav pri segmentaciji majhnih anatomskih struktur, kot so optični živci in žleza ščitnica, kar je razvidno iz same razpršitve točk na škatlastih diagramih, hkrati pa so ti organi pogosto na robu vidnega polja slik, kar orodju dodatno otežuje nalogo segmentacije. Če primerjamo rezultate metrik 2D in 3D modela, je v povprečju mogoče opaziti višja odstopanja pri 2D modelu za večino anatomskih struktur. V nekaterih primerih so sCT slike od resničnih bolj uporabne za segmentacijske naloge, kar sklepamo ne le iz povprečnih vrednosti metrik, ampak tudi iz primerov, ko uporabljeno orodje TotalSegmentator ni uspelo izvesti segmentacije na resničnih CT slikah, na sCT slikah pa mu je to uspelo.

5 Zaključek

Naučen difuzijski model za prevajanje med MR in CT slikovno modaliteto generira kvalitetne slike, kar dokazujejo visoke vrednosti PSNR in SSIM. Ti obetavni rezultati nakazujejo, da lahko sCT slike ohranijo zadostno anatomsko natančnost za nekatere nadaljnje naloge radioterapije, kot sta orisovanje OAR in načrtovanje obsevanja. Za potrditev te hipoteze so potrebne dodatne obsežne raziskave in empirična potrditev. Metodološki napredki difuzijskih modelov, razpoložljivost podatkovnih zbirk s pari CT in MR slik ter zapletenost paradigme MR-samostojnega načrtovanja zdravljenja v delovnem postopku radioterapije predstavljajo številne priložnosti za nadaljnje raziskave.

Opomba

Delo je odobrila Etična komisija Onkološkega inštituta Ljubljana, Slovenija, pod številko ERID-EK/139, podprla pa ga je Javna agencija za znanstvenoraziskovalno in inovacijsko dejavnost Republike Slovenije (ARIS) v okviru projektov P2-0232 in J2-60042.

Viri in literatura

- Cardoso, M. J., Li, W., Brown, R., and et al., “MONAI: an open-source framework for deep learning in healthcare,” *arXiv:2211.02701* (2022).
- Dayarathna, S., Islam, K. T., Uribe, S., Yang, G., Hayat, M., and Chen, Z., “Deep learning based synthesis of MRI, CT and PET: review and analysis,” *Med. Image Anal.* 92, 103046 (2024).
- Graf, R., Schmitt, J., Schlaeger, S., Möller, H. K., Sideri-Lampretsa, V., Sekuboyina, A., Krieg, S. M., Wiestler, B., Menze, B., Rueckert, D., and Kirschke, J. S., “Denoising diffusion-based MRI to CT image translation enables automated spinal segmentation,” *Eur. Radiol. Exp.* 7, 70 (2023).
- Gregoire, V., Guckenberger, M., Haustermans, K., Lagendijk, J. J. W., Ménard, C., Pötter, R., Slotman, B. J., Tanderup, K., Thorwarth, D., van Herk, M., and Zips, D., “Image guidance in radiation therapy for better cure of cancer,” *Mol. Oncol.* 14, 1470–1491 (2020).
- Ho, J., Jain, A., and Abbeel, P., “Denoising diffusion probabilistic models,” in *[34th Conference on Neural Information Processing Systems - NeurIPS 2020], Advances in Neural Information Processing Systems* 33, 6840–6851, NeurIPS (2020).
- J. Wasserthal, H.-C. Breit, M. T. Meyer, M. Pradella, D. Hinck, A. W. Sauter, T. Heye, D. T. Boll, J. Cyriac, S. Yang, M. Bach in M. Segeroth, “Totalsegmentator: Robust segmentation of 104 anatomic structures in ct images,” *Radiology: Artificial Intelligence*, vol. 5, no. 5, str. e230024, 2023.
- Kazerouni, A., Aghdam, E. K., Heidari, M., Azad, R., Fayyaz, M., Hacihaliloglu, I., and Merhof, D., “Diffusion models in medical imaging: a comprehensive survey,” *Med. Image Anal.* 88, 102846 (2023).
- Liu, L., Ren, Y., Lin, Z., and Zhao, Z., “Pseudo numerical methods for diffusion models on manifolds,” in *[10th International Conference on Learning Representations - ICLR 2022]*, 1–23 (2022).
- Li, Y., Xu, S., Lu, Y., and Qi, Z., “CT synthesis from MRI with an improved multi-scale learning network,” *Front. Phys.* 11, 1088899 (2023).
- McDonald, B. A., Dal Bello, R., Fuller, C. D., and Balermipas, P., “The use of MR-guided radiation therapy for head and neck cancer and recommended reporting guidance,” *Semin. Radiat. Oncol.* 34, 69–83 (2024).
- Podobnik, G., Strojjan, P., Peterlin, P., Ibragimov, B., and Vrtovec, T., “HaN-Seg: The head and neck organ-at-risk CT & MR segmentation dataset,” *Med. Phys.* 50, 1917–1927 (2023).
- Podobnik, G., Ibragimov, B., Peterlin, P., Strojjan, P., and Vrtovec, T., “vOARiability: Interobserver and intermodality variability analysis in oar contouring from head and neck CT and MR images,” *Med. Phys.* 51, 2175–2186 (2024).
- Podobnik, G., Ibragimov, B., Tappeiner, E., Lee, C., Kim, J. S., Mesbah, Z., Modzelweski, R., Ma, Y., Yang, F., Rudecki, M., Wodźniński, M., Peterlin, P., Strojjan, P., and Vrtovec, T., “HaN-Seg: The head and neck organ-at-risk CT and MR segmentation challenge,” *Radiother. Oncol.* 198, 110410 (2024).
- Ronneberger, O., Fischer, P., and Brox, T., “U-Net: convolutional networks for biomedical image segmentation,” in *[18th International Conference on Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015], Lecture Notes in Computer Science* 9351, 234–241, Springer (2015).
- Song, J., Meng, C., and Ermon, S., “Denoising diffusion implicit models,” in *[9th International Conference on Learning Representations - ICLR 2021]*, 1–20 (2021).
- Wang, Z., Bovik, A. C., Sheikh, H. R., and Simoncelli, E. P., “Image quality assessment: from error visibility to structural similarity,” *IEEE Trans. Image Process.* 13, 600–612 (2004).

- Wolterink, J. M., Dinkla, A. M., Savenije, M. H. F., Seevinck, P. R., van den Berg, C. A. T., and Išgum, I., “Deep MR to CT synthesis using unpaired data,” in *[2nd International Workshop on Simulation and Synthesis in Medical Imaging - SASHIMI 2017, held in conjunction with MICCAI 2017], Lecture Notes in Computer Science 10557*, 14–23, Springer (2017).
- Yang, H., Sun, J., Carass, A., Zhao, C., Lee, J., Prince, J. L., and Xu, Z., “Unsupervised MR-to-CT synthesis using structure-constrained CycleGAN,” *IEEE Trans. Med. Imaging* 39, 4249–4261 (2020).
- Yang, H., Sun, J., Carass, A., Zhao, C., Lee, J., Prince, J. L., and Xu, Z., “Unsupervised MR-to-CT synthesis using structure-constrained CycleGAN,” *IEEE Trans. Med. Imaging* 39, 4249–4261 (2020).