

Rekonstrukcija oblika organa iz para rendgenskih slika korištenjem neuronskih mreža

Sekula, Dominik

Master's thesis / Diplomski rad

2024

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Electrical Engineering and Computing / Sveučilište u Zagrebu, Fakultet elektrotehnike i računarstva**

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:168:368863>

Rights / Prava: [In copyright](#) / [Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2025-03-26**



Repository / Repozitorij:

[FER Repository - University of Zagreb Faculty of Electrical Engineering and Computing repository](#)



SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

DIPLOMSKI RAD br. 50

**REKONSTRUKCIJA OBLIKA ORGANA IZ PARA
RENDGENSKIH SLIKA KORIŠTENJEM NEURONSKIH
MREŽA**

Dominik Sekula

Zagreb, veljača 2024.

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

DIPLOMSKI RAD br. 50

**REKONSTRUKCIJA OBLIKA ORGANA IZ PARA
RENDGENSKIH SLIKA KORIŠTENJEM NEURONSKIH
MREŽA**

Dominik Sekula

Zagreb, veljača 2024.

Zagreb, 2. listopada 2023.

DIPLOMSKI ZADATAK br. 50

Pristupnik: **Dominik Sekula (0036519524)**
Studij: Informacijska i komunikacijska tehnologija
Profil: Automatika i robotika
Mentor: izv. prof. dr. sc. Matko Orsag

Zadatak: **Rekonstrukcija oblika organa iz para rendgenskih slika korištenjem neuronskih mreža**

Opis zadatka:

Potrebno je istražiti mogućnosti i dostupna rješenja za rekonstrukciju 3D modela objekta iz dvodimenzionalnih slika korištenjem Neural Radiance Fields metode (NeRF). Među odabranim rješenjima treba odabrati najpogodnije za primjenu na paru rendgenskih slika s ciljem rekonstrukcije oblika organa. Za učenje NeRF neuronske mreže potrebno je stvoriti sintetski skup podataka tehnikom digitalno rekonstruiranog radiografa (DRR), za što treba koristiti skup označenih CT snimki iz skupa AbdomenCT-1K. Kod izrade skupa za učenje DRR tehnikom potrebno je osigurati sličnost podataka za učenje sa stvarnim rendgenskim slikama. Mogućnosti konačnog rješenja potrebno je ispitati na DRR slikama i na rendgenskim slikama.

Rok za predaju rada: 9. veljače 2024.

Istinski se zahvaljujem mentoru prof. dr. sc. Matku Orsagu na svim prilikama za sudjelovanja na projektima i svim savjetima kroz moje studijsko obrazovanje. Hvala na uvijek pozitivnom stavu i velikom povjerenju zbog kojeg sam neizmjereno puno naučio i napredovao. Hvala mojem asistentu dr. sc. Frani Petricu na iskrenoj pomoći u kasnim nedjeljnim satima.

Od srca se zahvaljujem mojim roditeljima na stalnoj podršci i saslušanju svih mojih problema koliko god oni zamorni ili kompleksni bili. Vaša neprekidna potpora kroz život služila mi je kao ključni motivator da mogu više i bolje nego što sam ikada očekivao. Hvala mlađem bratu i sestri, što su uvijek bili tu za mene, odvrćajući mi misli od stresnih trenutaka.

Hvala nije dovoljna riječ za moju Josipu, za njen nevjerojatan trud i predanost u svemu što radi za mene. Njene nesebične geste i motiviranost kroz sve situacije su izvor inspiracije i podrške koji me neprestano guraju naprijed i zbog tog sam neizmjereno sretan i zahvalan, znajući da imam takvu osobu uza sebe.

Hvala vam, dragi prijatelji, vaša prisutnost učinila je ovaj put posebnim i ugodnijim. Zahvaljujem se od srca što ste sa mnom dijelili zabavne trenutke (i loše), bilo da smo trenirali ili eksperimentirali s kuhanjem, igrali kvizove (ne baš dobro) ili se upuštali u avanture.

Od srca hvala mojim Zagrepčanima i ostatku obitelji na svemu što ste učinili za mene, bili ste mi oslonac tijekom cijelog školovanja, od najranijih dana pa sve do danas.

SADRŽAJ

1. Uvod	1
2. Pregled radova	2
3. Metoda	5
3.1. <i>NeRF - Neural Radiance Fields</i>	5
3.1.1. <i>Generative Radiance Fields</i>	7
3.2. <i>MedNeRF</i>	9
4. Generiranje skupa podataka	13
4.1. CT2XRAY	13
5. Treniranje	15
5.1. Dotreniranje i rekonstrukcija na dvije slike	18
6. Rezultati	21
7. Zaključak	32
Literatura	33

1. Uvod

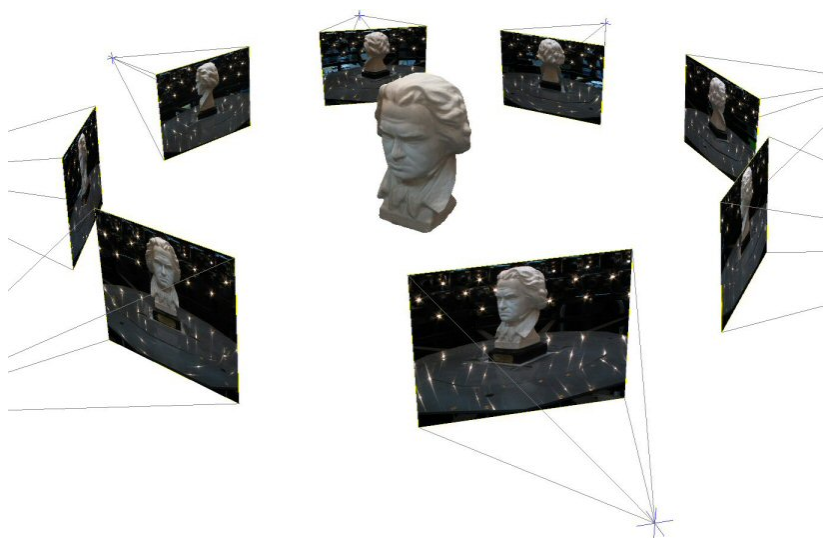
Koncept rekonstrukcije trodimenzionalnih modela iz dvodimenzionalnih slika predstavlja temeljno pitanje u području računarstva, čije istraživanje traje desetljećima, nastojeći istražiti i pronaći revolucionarne metode kako bi se premostio jaz između stvarnog i virtualnog svijeta. Trodimenzionalna rekonstrukcija ima iznimno široku primjenu i pokriva raznovrstan spektar područja korištenja. Primjeri uključuju industriju, medicinu, proizvodnju igara, robotiku i sl. Ovo multidisciplinarno područje služi kao most između više grana znanosti, pružajući konstantan napredak spajajući više područja primjene. U medicinskom sektoru, rekonstrukcija modela izrazito je važna jer omogućava neinvazivnu vizualizaciju i analizu kompleksnih struktura unutar pacijenta. Korištenjem ovih metoda liječnici mogu vizualizirati unutarnje organe te putem tih vizualizacija planirati i pratiti tijek liječenja. Istovremeno, ovaj pristup znatno smanjuje rizik za pacijente i ubrzava vrijeme potrebno za postavljanje dijagnoze.

U okviru ovog rada, istražuje se napredna metoda rekonstrukcije modela nazvana NeRF to jest *Neural Radiance Fields*, predstavljenu kao iznimno inovativan pristup rješenju ovog problema. Ova metoda koristi duboko učenje u kontekstu rekonstrukcije modela kako bi postigla iznimno visoku razlučivost i preciznost. U ovom radu posebni fokus je na primjenu NeRF-ova za rekonstrukciju oblika organa iz para rendgenskih slika, što ima izrazito velik potencijal za unaprjeđenje medicinske dijagnostike i olakšavanja tereta na medicinsko osoblje. Za učenje NeRF neuronske mreže potreban je skup podataka, u slučaju ovog rada odabrano je stvaranje sintetskog skupa tehnikom digitalno rekonstruiranog radiografa (DRR) koristeći označene CT slike. Ova tehnika koristi informacije iz volumetrijskih podataka kako bi se simulirale 2D rendgenske slike iz različitih kutova, omogućavajući zdravstvenim stručnjacima da vizualiziraju unutarnje strukture tijela. Kod izrade tih podataka potrebno je osigurati sličnost generiranih podataka sa stvarnim rendgenskim slikama. Nakon treniranja, provodi se testiranje mreže na DRR i stvarnim rendgenskim slikama kako bi se ocijenila njezina učinkovitost. Cilj rada je pridonijeti napretku u području rekonstrukcije oblika organa, te time unaprijediti medicinsku dijagnostiku i olakšati rad medicinskog osoblja.

2. Pregled radova

Rekonstrukcija trodimenzionalnih prikaza objekata i scena iz dvodimenzionalnih slika je dugoročni problem u računalnom vidu, suočen s nizom različitih pristupa. Tradicionalne metode kao SfM (*Structure from Motion*) [5, 21] i MVS (*Multi View Stereo vision*) [1] su već ukazale na dobre rezultate u stvaranju kvalitetnih modela iz velike količine detaljnih ulaznih podataka. Osim njih, obećavajući pristupi korišteni u industriji za modularnu i detaljnu rekonstrukciju su također fotogrametrija [12] i metode bazirane na dubokom učenju [7, 17]. Isto tako, proteklih godina svjedočimo iznimno brzom napretku u razvitku i korištenju implicitnih reprezentacija za zadatke 3D sinteze pogleda što je dovelo do stvaranja zanimljivih metoda rekonstrukcije scena. Inovativan i zanimljiv primjer takvih metoda bio bi NeRF (*Neural Radiance Fields*) [17] i njemu srodni pristupi kao što su *Mip-NeRF*, *Generative Radiance Fields* i *Few-View Object Reconstruction* [3, 22, 13] koji pokazuju izrazito visoku sposobnost modeliranja složene geometrije i izgleda traženih prikaza, dok *Instant-NGP* [18] pruža mogućnost iznimno brzog stvaranja visoko-kvalitetnih scena i modela od danih ulaznih podataka. Te metode iskorištavaju mrežu koja može naučiti kako svaki piksel na 2D slici korelira s prostornim pozicijama u 3D sceni, te iskoristiti to znanje da stvara reprezentacije scene ili objekta iz bilo koje pozicije. Budući da su metode poput NeRF-a relativno nove i obećavajuće, predviđa se mnoštvo inovacija, doprinosa i uzbudljivih prilika u daljnjem razvoju rekonstrukcije 3D modela iz 2D slika.

Medicina predstavlja znatno zanimljivo područje primjene za metode kao što su *NeRF* i slični pristupi, čime se otvaraju mogućnosti za olakšanje tereta na medicinsko osoblje, ubrzavanje dijagnostike ili poboljšanje interpretacije rezultata. Ovi napredni modeli mogu omogućiti precizniju analizu i optimizirane metode liječenja kroz korištenje sinteze pogleda dobivenog pomoću neuralnog polja zračenja tj. *NeRF-a*.

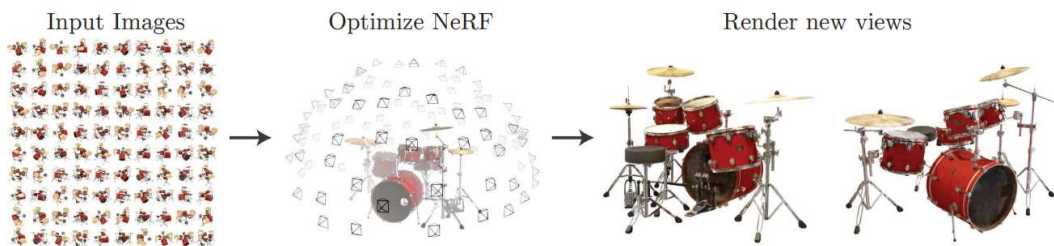


Slika 2.1: Pojednostavljena vizualizacija rekonstrukcije 3D modela iz više slika različitih pogleda s [16] web stranice

Tema ovog rada usredotočena je na rekonstrukciju oblika organa, pitanje koje je već razmatrano u drugim radovima kao što su [2] i [24]. Problem rekonstrukcije modela unutrašnjih struktura čovjeka jest to što osim činjenice da su svi ljudi različiti, također su i uređaji kojim se barata različitih kvaliteta, stoga prikazi neće biti ujednačeni kroz skupove podataka. Potencijalna rješenja tog problema razmatrana su u radovima poput [8] gdje autori generiraju sintetički skup podataka putem tehnike digitalno rekonstruiranog radiografa kao što je napomenuto u uvodu ovog rada. Ta tehnika smanjuje potrebu za iscrpnim dobavljanjem stvarnih skenova podešavajući parametre stvarnih sustava iskorištavanih u bolnicama ili istraživačkim centrima i dozvoljava dobivanje što veće sličnosti sa stvarnim rendgenskim slikama.

Autori [18] su uspješno istraživali i razvijali metode brze sinteze pogleda u sklopu NeRF-ova, no njihov rad nije direktno vezan za medicinsku primjenu, to pitanje brze sinteze ostaje neistraženo u trenutku pisanja ovog rada i ostavlja prostor za buduće istraživače. Čest problem u više sektora istraživanja je manjak podataka za obradu, potencijalno zbog nedostupnosti, financijskih razloga ili težine dobavljanja istih. Rješenje tog problema ispitali su autori metoda pod nazivima *DS-NeRF* [9], *PixelNeRF* [25] i *FORGE* [13] koristeći izrazito malu količinu ulaznih slika iz kojih su dobili izrazito dobre 3D prikaze njihovih ulaznih podataka. Ostvarili su to koristeći metodiku gdje iskorištavaju značajke slika u *feed-forward* algoritmu prije treniranja na tim istim slikama da pojačaju učenje neuronske mreže za dani problem. Fokus na potpunu rekonstrukciju modela i njegove površine koristeći NeRF metode rekonstrukcije imali su autori u *UNISURF* [19].

Najznačajniji i najsličniji pristup za ovaj rad po tematici predstavljen je u [8] pod nazivom MedNeRF, u kojem je korištena bazna metoda slična NeRF-u nazvana GRAF (*Generative Radiance Fields*) [22], gdje su također koristili vrlo malo pogleda za rekonstrukciju modela iz rendgenskih slika, odnosno samo jedan naspram naših dva. Osim njega, vrlo zanimljiv pristup u medicini je *CuNeRF* to jest Cube-based NeRF [6] gdje je cilj stvaranje volumetrijske reprezentacije CT i MRI skenova. Ovaj pristup provodi tehniku *super-resolution* proizvoljnog skaliranja, oboje koristeći uzorkovanje kockom naspram originalnog rada u kojem se koriste zrake.



Slika 2.2: NeRF pipeline

Osim razvoja raznih metodologija i varijanti NeRF-ova nedavno je predstavljen modularni programski okvir otvorenog koda pod nazivom *Nerfstudio* [23] koji pruža mogućnosti korištenja nekih od prijašnje spomenutih primjera s velikom jednostavnošću. Ovaj okvir omogućuje korisnicima korištenje nekih od prethodno spomenutih primjera s iznimnom jednostavnošću. Ističe se nekolicinom implementiranih prihvatljivih znanstvenih radova i metoda, čineći proces korištenja izuzetno lako prilagodljivim i pristupačnim.

3. Metoda

Neuronske mreže, kao bitan alat dubokog učenja predstavljaju izrazito moćan i fleksibilan instrument u razvoju metoda za volumnu rekonstrukciju oblika iz niže-dimenzijskih značajki. Tako sofisticirana tehnologija bazira se na automatskom učenju izraženih značajki iz naizgled jednostavnih ulaznih podataka. Njihova iznimna sposobnost ekstrakcije i povezivanja složenih uzoraka iz ograničenog skupa informacija omogućuje im duboko razumijevanje i reprezentacije struktura koje su inherentne podacima, a često izvan dosega ljudskog opažanja.

Fleksibilnost neuronskih mreža čini ih neophodnim alatom u mnogim područjima, od financijskih predikcija i analize tržišta do generativnih zadataka poput sinteze slika. Njihova sposobnost prilagodbe raznovrsnim podacima postavlja ih ključnim u kontekstu rekonstrukcije oblika, posebice iz medicinskih slika, gdje se često susrećemo s izazovima manjka informacija ili neujednačenosti odnosno nehomogenosti podataka. Stvaranje kontinuirane volumne reprezentacije iz niže-dimenzijskih značajki pomoću neuronskih mreža predstavlja ključni korak prema boljem razumijevanju i analizi podataka.

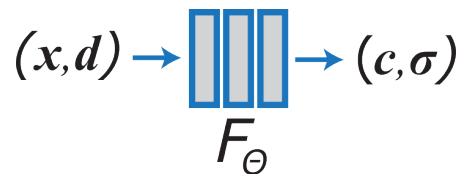
U svrhu ostvarivanja ovih ciljeva u kontekstu medicinske rekonstrukcije, primjenjujemo sofisticiranu arhitekturu poznatu kao MedNeRF [8]. Ovaj model predstavlja inovativnu kombinaciju različitih arhitektura uključujući varijaciju *Generative Adversarial Networks*, nazvanu *Generative Radiance Fields* [22] u kombinaciji s *Neural Radiance Fields* [17]. Takva arhitektura omogućava modelu da nauči kompleksne uzorke i disentanglira oblik i volumetrijsku dubinu iz ograničenog skupa ulaznih podataka.

3.1. *NeRF - Neural Radiance Fields*

Prije nego što se temeljito pristupi opisivanju MedNeRF mreže u potpunosti, potrebno je opisati njezine osnovne komponente, počevši s *Neural Radiance Fields*. *Neural Radiance Fields* (NeRF) predstavlja metodu dubokog učenja namijenjenu sintetiziranju kompleksnih trodimenzijskih prikaza scene iz dvodimenzijskih slika.

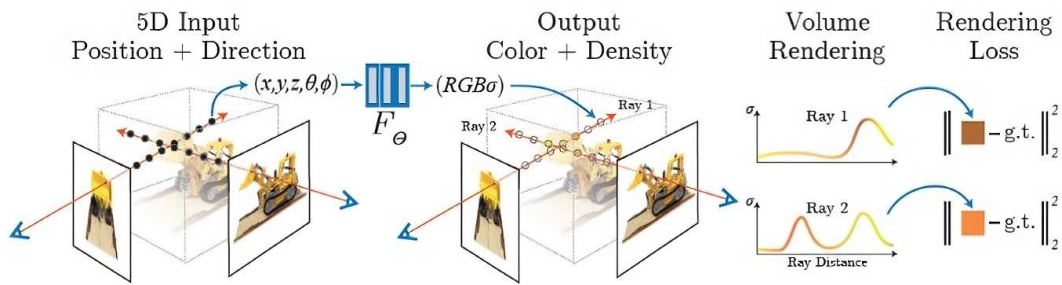
Model NeRF mreže je potpuno povezana ne-konvolucijska duboka neuronska mreža također poznata pod nazivom *Multi layer perceptron* (MLP). NeRF metodika razlikuje se od drugih metoda dubokog učenja, time što koristi niz slika za treniranje jedne potpuno povezane neuronske mreže, fokusirane na generiranje novih prikaza tog istog objekta.

Ta potpuno povezana mreža reprezentira scenu s pomoću *Radiance fields*, matematičkog modela koji mapira svaku točku u 3D prostoru na informacije o svjetlu koje prolazi kroz tu točku. Ovaj model radi s 5D vektorskim ulazom definiranim 3D lokacijom vektora $\mathbf{x} = (x, y, z)$ i 2D orijentacija $\mathbf{d} = (\theta, \phi)$, dok je izlaz 4D, to jest boja $\mathbf{c} = (r, g, b)$ i gustoća volumena σ , ilustracija te funkcije prikazana je na slici 3.1. Stvaranje scene ostvaruje se ispitivanjem 5D koordinata duž zrake kamere, nalik propuštanju laserskih zraka kroz kameru i uzorkovanju na specifičnim udaljenostima. Primjenom klasičnih tehnika volumetrijskog renderiranja, koje iskorištavaju akumulirane boje i gustoće iz različitih točaka prostora duž zrake kamere, postiže se kompleksno oblikovanje slike. Vizualni prikaz moguće je vidjeti na slici 3.2, gdje se kroz boje i intenzitete zraka dobiva precizna vrijednost u toj točki.



Slika 3.1: Funkcija volumetrijske scene MLP-a

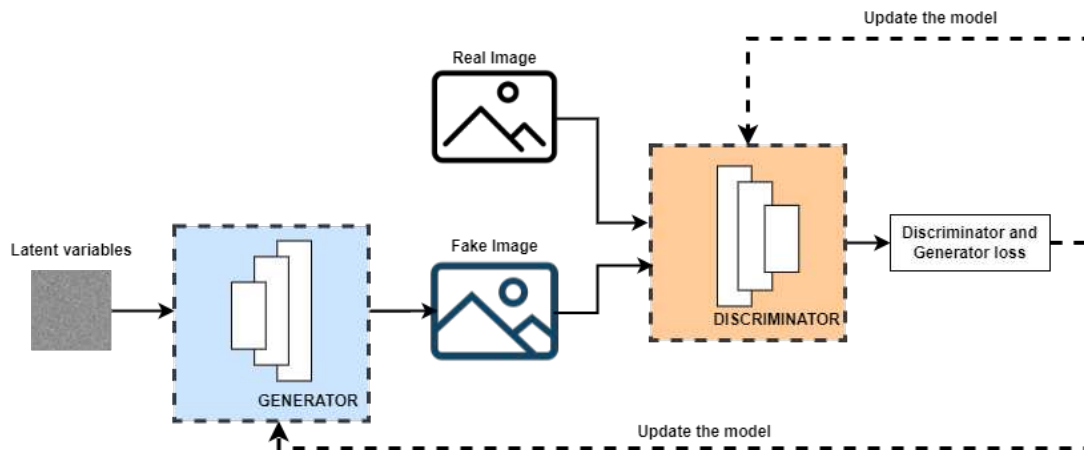
Treniranje koristi slike objekta ili scene snimljene iz različitih kuteva u kombinaciji s položajima istih. Proces treniranja usredotočuje se na prilagodbu težina mreže kako bi njezin izlaz bio usklađen s pikselima na slikama, optimizirajući tako rekonstrukciju trodimenzionalne scene. To se radi tako da se u svakoj iteraciji generirana slika uspoređuje sa stvarnom slikom u istoj poziciji, dobiva gubitak i tako podešavaju težine pri treniranju, također vidljivo na slici 3.2



Slika 3.2: Primjer funkcionalnosti NeRF-a

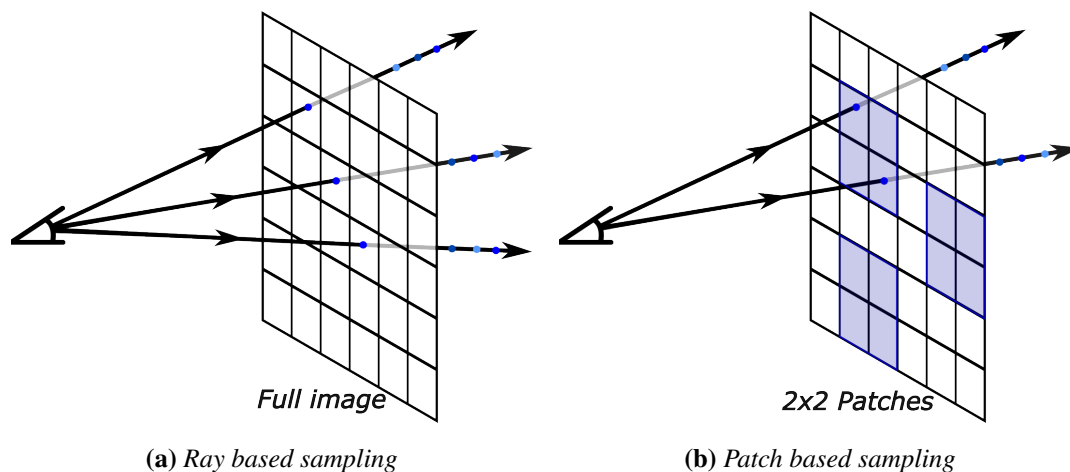
3.1.1. *Generative Radiance Fields*

Generative radiance fields (GRAF) predstavlja naprednu tehniku u dubokom učenju usredotočenu za područje sinteze slika s posebnim fokusom na duboko razumijevanje trodimenzionalnog prostora i procesa generiranja slika. GRAF tehnologija bazirana je na modelu *Generative adversarial network* [11], to jest GAN-ova. GAN je vrsta duboke neuronske mreže koja se, kao što naziv implicira, koristi za generiranje novih podataka, poput slika ili zvuka iz prijašnje naučenih uzoraka. Sastoji se od dva modela - generatora, koji stvara nove sintetske primjere i diskriminatora, koji ocjenjuje kvalitetu generiranih primjera na temelju usporedbe sa stvarnim. Ta dva modela su u međusobnom natjecanju, što potiče generator da stvara što uvjerljivije primjere, dok diskriminator razvija sposobnost razlikovanja stvarnih i sintetski generiranih podataka. Generator je neuronska mreža koja prima slučajne uzorke šuma, to jest latentne varijable i transformira ih u podatke s ciljem naučenosti distribucije stvarnih podataka kako bi stvorio uvjerljive i realistične uzorke i time prevario diskriminatora. Diskriminator je također neuronska mreža, ali ona ocjenjuje stvarne podatke i one koje generira prijašnje spomenuti generator s ciljem prilagođavanja kako bi postao što bolji u prepoznavanju razlika između stvarnih i sintetskih podataka. Dijagram te funkcionalnosti vidljiv je na slici 3.3.



Slika 3.3: Pojednostavljen način rada GAN-ova i GRAF-ova

Generative radiance fields proširuje osnovni koncept *Generative adversarial networks*-a primjenjujući ga na područje sinteze slika uz korištenje *radiance fields* za reprezentaciju scene u generatoru umjesto tradicionalnih GAN pristupa. Ovaj pristup omogućuje GRAF modelu bolje razumijevanje i reprodukciju trodimenzionalnog prostora, čime postiže precizniju i bržu kontrolu nad generiranim slikama objekata ili scena. Uz oslonac na *radiance fields* kao matematički model za reprezentaciju scene, GRAF unapređuje duboko razumijevanje trodimenzionalnog prostora korištenjem *Patch based sampling* umjesto standardnog *Ray based sampling* kao u običnom NeRF modelu. Taj način uzorkovanja omogućuje GRAF modelu da analizira scenu s pomoću određenih diskretnih područja unutar scene umjesto cjelokupnih slika odjednom, pružajući efikasnije prikupljanje informacija o trodimenzionalnom prostoru. U fazi treniranja, GRAF-ov generator ne sintetizira potpunu sliku, već generira tzv. *patch* koji se zatim uspoređuje od strane diskriminatora s odgovarajućim dijelom stvarne cjelokupne slike. Usporedba navedenog modaliteta uzorkovanja s *ray-based* pristupom jasno je prikazana na slici 3.4. Na slici 3.4a vrši se uzorkovanje cijelog skupa piksela, dok se na slici 3.4b uzorkovanje provodi na manjim podskupovima piksela, to jest *patch*-evima, u ovom slučaju veličine 2x2 piksela. Ovakav pristup ne samo da doprinosi smanjenju troškova i vremenske kompleksnosti treniranja, već poboljšava efikasnost modela za generaliziranje na različite trodimenzionalne scene. Uz to, smanjuje zahtjeve za resurse zbog manjih dimenzija ulaznih podataka, čineći GRAF model veoma praktičan za mnoštvo primjena.



Slika 3.4: Usporedba različitih metoda uzorkovanja

3.2. *MedNeRF*

Medicinska slikovna rekonstrukcija bitna je tematika u suvremenoj medicini pružajući neophodne alate za neinvazivnu dijagnostiku i praćenje bolesti. Različiti modaliteti slika postoje za slikovnu dijagnostiku, svaki sa svojim karakterističnim prednostima i ograničenjima. pri čemu se među najčešće korištene ubrajaju računalna tomografija (CT), rendgensko snimanje (X-Ray) i magnetska rezonanca (MRI). Klasični pristupi rekonstrukciji trodimenzionalnih modela anatomskih struktura pacijenata zahtijevaju prikupljanje velikog broja slika, što rezultira iznimno visokom dozom zračenja, čiju količinu je potrebno minimizirati ili u potpunosti izbjeći. Uzimajući ovaj nedostatak u obzir, predstavljen je inovativan koncept iskorištavanja neuronskih mreža za rekonstrukciju 3D modela iz ograničenog broja 2D slika nazvan MedNeRF.

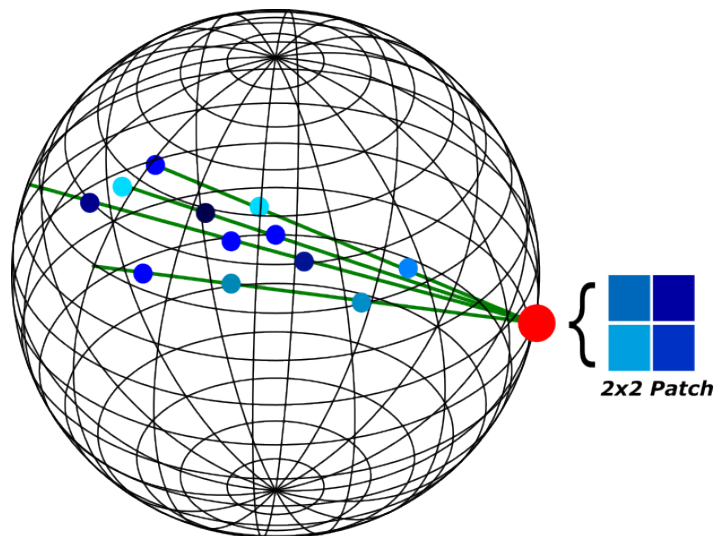
MedNeRF ili *Medical Neural Radiance Fields* predstavlja inovativan pristup u području rekonstrukcije trodimenzionalnih modela koristeći dvodimenzionalne slike za ulaz [8]. Ova tehnologija temeljena je na prijašnje spomenutom konceptu *Generative radiance fields*-a, koji se u svom osnovnom obliku koristi za generiranje novih podataka iz standardnih oblika RGB slika. Posebnost MedNeRF-a jest ta što preuzima taj koncept i nadograđuje ga za primjenu na područje medicinske slikovne rekonstrukcije gdje je preciznost i dobra naučenost od suštinskog značenja za postizanje visokokvalitetnih rezultata. Model GRAF, na kojem se temelji MedNeRF koristi *radiance fields* za reprezentaciju scene. Oni sami po sebi predstavljaju matematički model koji mapira svaku točku u 3D prostoru na informaciju o svjetlosti koja prolazi kroz tu istu točku.

Taj pristup po samoj definiciji čini se iznimno primamljiv za tematiku rekonstrukcije medicinskih slika zbog same funkcionalnosti uređaja koji se koriste u procesu akvizicije slika. Integracija ovog koncepta omogućuje MedNeRF-u da efikasno modelira i rekonstruira kompleksne medicinske strukture.

MedNeRF za svoj ulaz koristi digitalno rekonstruirani radiograf (DRR), simulaciju konvencionalnih 2D rendgenskih slika stvorenih iz snimki dobivenih putem računalne tomografije (CT). Upotreba DRR-ova omogućava dobivanje visokokvalitetnih skupova podataka bez dodatnog izlaganja pacijenta radijaciji. Ovaj pristup ne samo da omogućuje kontrolu parametara snimanja, uključujući kutove gledanja i rezoluciju slike, već također pruža mogućnost prilagodbe intenziteta piksela, čime se određuju varijacije unutrašnjih struktura, doprinoseći dubljoj raznolikosti i vjerodostojnosti generiranih rendgenskih slika. Odabrana tehnika za generaciju DRR slika u MedNeRF radu bila je korištenje softverskog alata *Plastimatch* [20] koji je softver otvorenog koda za visokoučinkovitu volumetrijsku registraciju, manipulaciju i stvaranje medicinskih slika. Integracija ovakve metode stvaranja sintetskih podataka iznimno olakšava proces generiranja skupova podataka prema specifičnim zahtjevima njihovih istraživanja ili kliničkih potreba.

Osim digitalno rekonstruiranih radiografa (DRR), kako u GRAF-u tako i u MedNeRF-u, koristi se konfiguracijska datoteka za ulaz. Ova konfiguracijska datoteka ima ključnu ulogu u definiranju parametara procesa treniranja u svrhu rekonstrukcije. Primjerice, "radius" je parametar koji utječe na veličinu volumena oko centralne točke u prostoru, odnosno, u kontekstu medicinskih slika, na udaljenost rendgenskog izvora od središta pacijenta. *Far* i *near* definiraju raspon unutar kojeg će MedNeRF analizirati strukture. *Field of view* (FOV) odnosi se na vidno polje, to jest kut koji obuhvaća projekcija. Također, postoje (**u_min, u_max**) i (**v_min, v_max**), parametri koji određuju minimalni i maksimalni azimutalni (u) i polarni (v) kut unutar kojeg se može generirati nasumična lokacija ξ . Ovi parametri igraju ključnu ulogu u preciznom podešavanju modela DRR-ova kako bi se postigla što veća usklađenost sa stvarnim rendgenskim skenovima. Kombinacija ovih parametara omogućuje preciznu kontrolu i fokusiranje na određeni prostor interesa, poboljšavajući učinkovitost rekonstrukcije.

U kontekstu MedNeRF-a, glavni cilj je dobivanje intenziteta zračenja na specifičnim lokacijama piksela. To se dobiva uzorkovanjem **KxK** patch-eva zraka na nasumičnoj lokaciji ξ generiranoj na sferi oko pacijenta. Svaka zraka uzorkuje **N** 3D točaka duž te rendgenske zrake koja potječe iz ξ . *Radius* određuje geometrijske karakteristike te sfere, kako je prikazano u primjeru na slici 3.5, dok *near*, *far* i područja azimutnih i polarnih kutova određuju iskoristivo područje točaka unutar te sfere.



Slika 3.5: Pojednostavljena vizualizacija uzorkovanja u MedNeRF-u gdje je $N = 3$, $K = 2$ i ξ je nasumična pozicija prikazana crvenom bojom

Inovativne nadogradnje nad GRAF arhitekturom za medicinsku domenu potrebne su zbog ograničenih veličina medicinskih skupova podataka. Prvotna nadogradnja jest unaprjeđenje to jest zamjena GRAF-ovog diskriminatora s inovativnim tehnikama samo-podržanog učenja. Diskriminator D_φ će iskoristavati tehniku samo-podržanog učenja zasnovanu na auto-encodiranju da uči globalne i lokalne značajke radiografskih slika. Radi to tako da izvodi analizu rekonstrukcije i ekstrakciju značajki, a zatim ih koristi za ocjenu i treniranje generatora G_θ . Za procjenu kvalitete strukture u dekodiranim značajkama iz D_φ koristili su *Learned Perceptual Image Patch Similarity* (LPIPS) [27] metriku za rekonstrukcijski gubitak, koja procjenjuje sličnost između rekonstruiranih volumetrijskih podataka i stvarnih podataka putem VGG16 mreže. Radi to na temelju usporedbe percepcijskih razlika između dviju slika, što ga čini iznimno korisnim alatom za procjenu kvalitete generiranih slika.

Osim što nadograđuje arhitekturu diskriminatora, MedNeRF dodatno uvodi i augmentaciju podataka za unaprjeđenje kvalitete učenja. To postižu uvođenjem transformacija kao što su horizontalno i vertikalno okretanje i rotacija T_k s pomoću više glava diskriminatora. Ove transformacije pridonose većoj robusnosti modela, čineći ga otpornijim na promjene u perspektivi i poziciji snimanja.

Nedvojbeno najbitnija stavka za potrebe zadatka ovog rada jest posljednja inovacija koju MedNeRF pridonosi nad ostalim modelima, a to je volumetrijsko renderiranje iz jedne slike digitalno rekonstruiranog radiografa. To je realizirano putem posebne python skripte koja obavlja to rekonstruiranje. Nakon provođenja treniranja modela,

rekonstruira se potpuna X-ray projekcija pacijenta u obliku videa koristeći samo jedan DRR. Funkcionira tako da se fine-tune-aju parametri generatora i diskriminatora nad tom singularnom slikom digitalno rekonstruiranog radiografa sa pozicijom kamere. To se radi minimizacijom generativne funkcije gubitka koja se sastoji od prijašnje spomenutog gubitka rekonstrukcije ($\mathcal{L}_r$), srednje kvadratne pogreške ($\mathcal{L}_{MSE}$) i negativnog logaritma izglednosti ($\mathcal{L}_{NLL}$).

Ukupni gubitak izgleda ovako:

$$\mathcal{L}_{\text{gen}} = \lambda_1 \mathcal{L}_r(\text{VGG16}) + \lambda_2 \mathcal{L}_{\text{MSE}}(G) + \lambda_3 \mathcal{L}_{\text{NLL}}(z) \quad (3.1)$$

Gdje λ_1 , λ_2 i λ_3 predstavljaju hiperparametre koji kontroliraju doprinose pojedinih dijelova jednadžbe, dok z su latentni vektori.

Ova funkcija gubitka bit će od velikog značaja u daljnjem predstavljanju prilagođene metode rekonstrukcije u ovom radu. Navedena struktura igra ključnu ulogu u pružanju izvrsne osnove za nadogradnju algoritma s obzirom na naš problem rekonstrukcije iz dva različita pogleda. Detaljnije pojašnjenje ideje i funkcionalnosti dostupno je kasnije u radu.

4. Generiranje skupa podataka

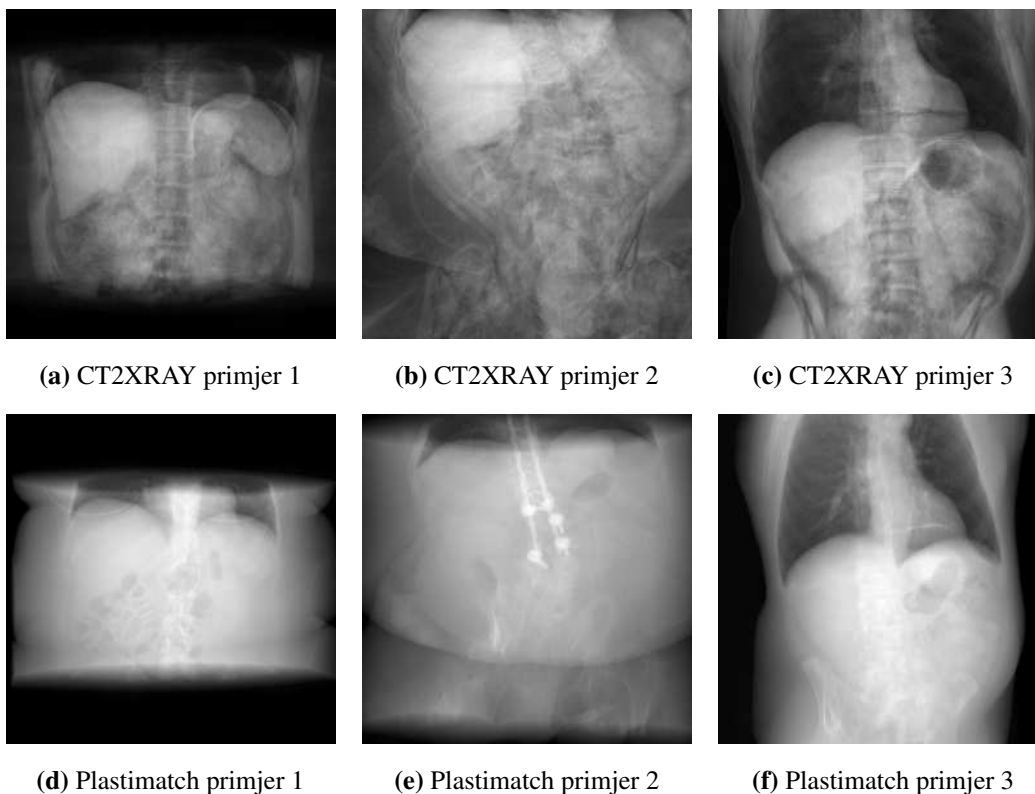
Skupovi podataka, bilo to sintetski ili stvarni predstavljaju ključnu ulogu u razvoju inovativnih tehnologija u području umjetne inteligencije. Kvalitetno prikupljeni i stvoreni skupovi podataka pružaju temelj za poučavanje dubokih modela, poput spomenutih digitalno rekonstruiranih radiografa (DRR) korištenih u MedNeRF modelu. Generacija DRR-ova u ovom kontekstu predstavlja olakšan pristup stvaranju izrazito kvalitetnih skupova podataka bez potrebe za dodatnim izlaganjem pacijenata zračenju. Kroz preciznu i raznovrsnu generaciju DRR-ova, omogućava se virtualna vizualizacija unutar tijela pacijenta. Korištenje DRR-ova kao skup podataka za treniranje neuronskih mreža pruža fleksibilnost u prilagodbi različitih parametara izgleda scene, što značajno unaprjeđuje funkcionalnost i modularnost sustava. Sve navedeno dodatno naglašava važnost skupova podataka u razvoju inovativnih tehnologija umjetne inteligencije gdje autorovo pažljivo promišljanje igra ključnu ulogu u odabiru optimalnog načina njihova stvaranja.

4.1. CT2XRAY

Prije nego što se definira algoritam generacije korišten u ovom radu, potrebno je pojasniti već često spomenute digitalno rekonstruirane radiografe (DRR). Digitalno rekonstruirani radiografi predstavljaju simulirane rendgenske slike koje se stvaraju s pomoću digitalne rekonstrukcije slika dobivenih iz medicinskih skeniranja, točnije putem računalne tomografije (CT) i magnetske rezonance (MRI). Kreirani su sumiranjem intenziteta po zraci svakog piksela do simuliranog izvora rendgenskog zračenja [14].

Za razliku od standardnog MedNeRF modela koji koristi Plastimatch za generaciju svojih DRR-ova, za potrebe ovog rada koristit će se LARICS-ov softver pod nazivom CT2XRAY [10]. Riječ je o softveru koji se koristi za generiranje DRR skupova podataka iz CT skenova.

Iz slike 4.1 jasno se uočavaju značajne razlike u razlučivosti unutrašnjih struktura

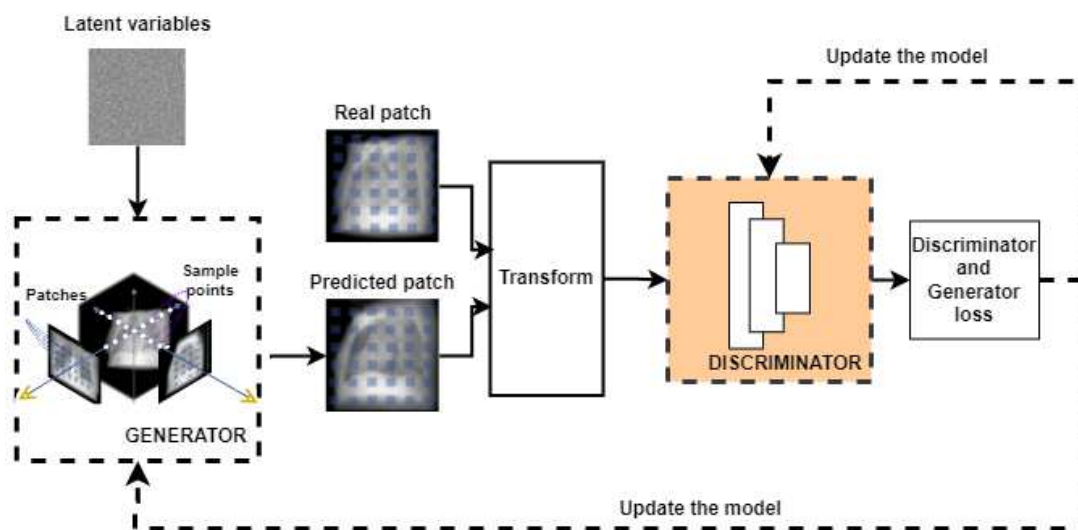


Slika 4.1: Usporedba CT2XRAY i defaultne Plastimatch generacije iz MedNeRF rada

pacijenata. Na gornje 3 slike, to jest na primjerima generiranim s CT2XRAY iznimno je lako razlučiti abdominalne strukture i organe, kao što bi bila jetra, bubrezi ili slezena. S druge strane na donjim slikama to nije toliko očito. Ovdje se te unutrašnje strukture dosta lako stapaju u pozadinski intenzitet abdomena i time ne pružaju dovoljno kvalitetne informacije mreži za efikasno treniranje. Iz ovih razloga odabrana metoda generacije je CT2XRAY s očekivanjem da će njezina primjena rezultirati poboljšanjem u razlučivosti i prepoznavanju unutrašnjih struktura, što će značajno doprinijeti učinkovitosti treniranja mreže za rekonstrukciju iz dva različita pogleda.

5. Treniranje

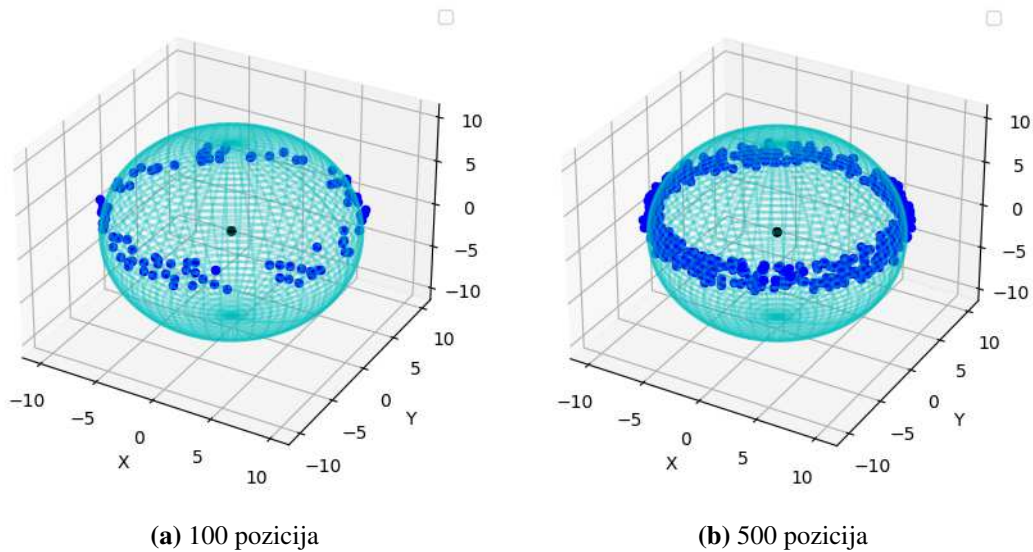
Nakon pojašnjenja procesa generiranja skupa podataka i pozadinskih funkcionalnosti MedNeRF modela, potrebno je skrenuti fokus na centralnu stavku, a to je treniranje neuronske mreže za rekonstrukciju trodimenzionalnih oblika organa. Ovaj segment rada ima ključnu ulogu u postizanju preciznih i relevantnih rezultata. U njemu će se analizirati detalji metode optimizacije i algoritma treniranja s ciljem postizanja visokokvalitetnih rezultata u rekonstrukciji oblika organa iz više različitih perspektiva. Skup CT podataka korišten za generaciju DRR-ova kojim će se trenirati mreža u ovom radu je AbdomenCT-1K [15], javno dostupan medicinski skup podataka već korišten za mnoštvo radova i projekata slične tematike zbog svoje opsežnosti i kvalitete.



Slika 5.1: Pojednostavljen MedNeRF algoritam jedne iteracije treniranja, modeliran prema GAN strukturi 3.3

Početni korak u procesu treniranja sastoji se od preuzimanja parametara iz konfiguracijske datoteke za stvaranje potrebnih distribucija i komponenata za rad. Ta datoteka sadrži važne informacije o arhitekturi mreže, hiperparametrima korištenima za optimizaciju iste i parametre za uzorkovanje pozicije kamere $\mathcal{E} = [\mathbf{R}|\mathbf{t}]$ pri treniranju.

Pozicija kamere, kako je prijašnje spomenuto dobiva se nasumično iz uniformne distribucije definirane azimutalnim (u) i polarnim (v) kutevima. Tim parametrima definira se prsten oko sfere na slici 3.5 unutar kojeg se uzima pozicija kao izvor zraka koje će uzorkovati intenzitete piksela, a time i scenu. Primjer nasumičnog stvaranja 500 pozicija kamere koristeći lateralne kuteve između ($u_{\min} = 0^\circ$, $u_{\max} = 360^\circ$), polarne kuteve između ($v_{\min} = 70^\circ$, $v_{\max} = 85^\circ$) i radius jednak 10 vidljiv je na slici 5.2. Te točke će u koraku treniranja služiti kao izvori zraka korištenih za proračun intenziteta piksela.



Slika 5.2: Nasumično generirane pozicije unutar MedNeRF-a

Nakon što je konfiguracijska datoteka iskorištena za potrebe inicijalizacije parametara i komponenata bitnih za učenje, kao što su *dataloader-i*, optimizatori i slično, proces treniranja započinje. Treniranje započinje s kreacijom latentnih varijabli z_{dist} uzorkovanih iz Gaussove distribucije, koje će biti korištene tijekom cijelog treniranja. Mreža koristi te varijable za pomoć pri učenju varijacija u obliku i izgledu anatomskih struktura tijekom procesa rekonstrukcije.

Oblik treniranja je iterativan, gdje mreža u svakoj iteraciji obrađuje *batch_size* digitalno rekonstruiranih radiografa na kojima će optimizirati parametre. Također, uzima se uzorak distribucije latentnih varijabli koji će pružiti dodatnu fleksibilnost generatoru, omogućavajući mu da se prilagodi različitim varijacijama u strukturi pacijentovih anatomskih dijelova. Prije samog ulaska u treniranje diskriminatora i generatora, stvarne slike pretvaraju se u *patch-eve* zbog arhitekture modela. Korak treniranja diskriminatora sastoji se od minimizacije regularizacijskog gubitka i gubitka diskriminatora 5.2 kroz što bolje prilagođavanje prepoznavanju razlika između stvarnih i gene-

riranih slika. Također, korak treniranja generatora sastoji se od minimizacije gubitka generatora 5.3 usklađivanjem generiranih slika sa stvarnima, pokušavajući prevariti diskriminator. Diskriminator korak sastoji se prvotno od testiranja diskriminatora na stvarnim podacima proračunavanjem njegovog gubitka sumacijom *hinge* gubitka ($\mathcal{L}_h$) i gubitka rekonstrukcije ($\mathcal{L}_r$) (5.1). Zatim, generiraju se lažni podaci nad kojima će se provoditi proračunavanje istih gubitaka. Gubitci na stvarnim i lažnim slikama se sumiraju te se dobiva ukupni gubitak zapisan u 5.2. Nakon toga diskriminator proračunava regularizacijski gubitak izračunavanjem kvadrata razlike između izlaza diskriminatora na stvarnim podacima i izlaza na generiranim podacima, uz primjenu hiperparametra za kontrolu jakosti regularizacije. Korak treniranja generatora sastoji se samo od stvaranja umjetnih podataka, proračunavanja gubitka generatora 5.3 i optimiziranja istog. Gubitak rekonstrukcije - *LPIPS* [27]:

$$\mathcal{L}_r = \mathbb{E} \left[\frac{1}{whd} \|\phi_i(G(f)) - \phi_i(T(p))\|_2 \right] \quad (5.1)$$

Gdje $\phi_i()$ označava izlaz unaprijed treniranog VGG16 modela, w , h i d predstavljaju širinu, visinu i dubinu prostora značajki. $G(f)$ predstavlja obradu na generiranim značajkama, a $T(p)$ obradu na stvarnim slikama u obliku *patch-eva*. Ovaj izraz predstavlja gubitak rekonstrukcije koji mjeri razliku između značajki generiranih slika i stvarnih slika u prostoru značajki VGG16 mreže. Gubitak diskriminatora nije ništa više nego suma ovog izraza iznad s gubitkom zgloba tj *hinge* gubitkom.

Ukupni gubitak diskriminatora:

$$\mathcal{L}_d(\theta, \phi_k) = \mathbb{E}[f(D_\phi(G_\theta(z, \xi, \nu)))] + \mathbb{E}[f(-D_\phi(I, \nu))] + \mathcal{L}_r \quad (5.2)$$

Prvi dio formule je gubitak na generiranim primjerima, to jest mjera koliko dobro generator reproducira stvarne podatke. Koristi se funkcija f (*hinge loss*) za procjenu izlaza diskriminatora na generiranim podacima. Drugi dio je gubitak na stvarnim primjerima, točnije mjera koliko dobro generator vara diskriminator. I u formuli predstavlja stvarne podatke, dok ν predstavlja *patch*. Ponovno se koristi funkcija f za procjenu izlaza diskriminatora na stvarnim podacima. Te na kraju, rekonstrukcijski gubitak, mjera koliko dobro generator reproducira stvarne podatke, ali procjenjuje uz korištenje percepcijskog gubitka (LPIPS).

Gubitak generatora:

$$L(\theta, \{\phi_k\}) = L(\theta, \phi_0) + \frac{\lambda}{n-1} \sum_{k=1}^n L(\theta, \phi_k) \quad (5.3)$$

Jednostavan gubitak koji samo sumira prvotno gubitak na slikama koje nisu transformirane i n transformiranih slika. n je odabran kao 4, dok je λ jednaka 0.2.

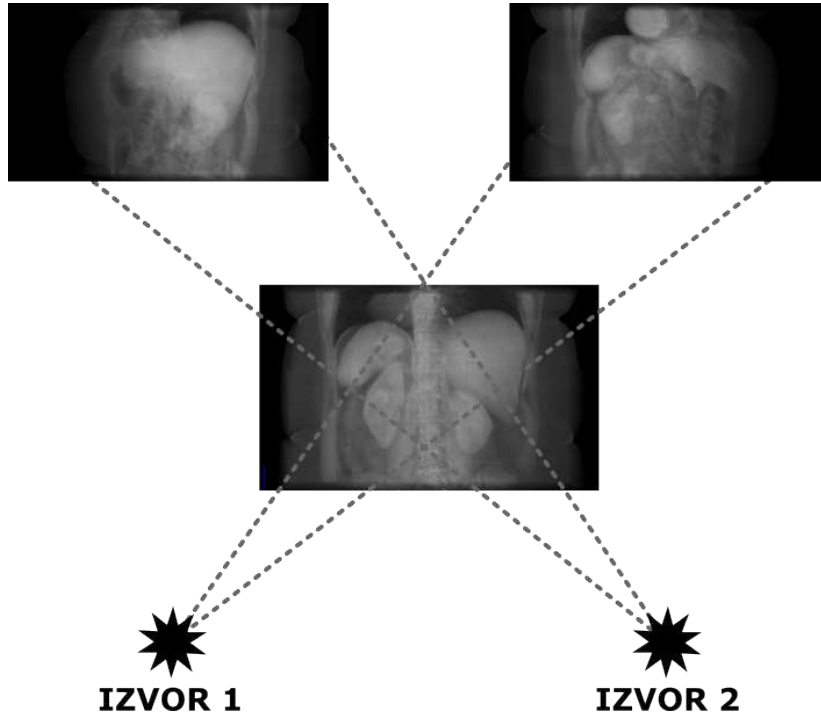
Nakon prijašnje spomenuta koraka diskriminatora i generatora, ažuriraju se težine i obavljaju se dodatne evaluacije kvalitete mreže. Tijekom procesa treniranja, model se redovito evaluira na posebnom generiranom skupu podataka kako bi se pratila kvaliteta generiranih slika. Tijekom procesa treniranja, koriste se objektivne evaluacijske mjere poput Fréchet inception distance (FID) i Kernel Inception Distance (KID). Ove mjere pružaju kvantitativnu analizu kvalitete generiranih slika, što omogućuje finu prilagodbu parametara modela kako bi se postigao optimalan rezultat. FID je metrika evaluacije generativnih modela koji pretpostavlja da se značajke iz realnih i generiranih slika ravnaju po Gaussovoj distribuciji. Niži FID obično ukazuje na bolju kvalitetu generiranih slika. KID, s druge strane, ne pretpostavlja Gaussovu distribuciju, već mjeri kvadratni Maksimalni Prostorni Disparitet (MMD) između reprezentacija stvarnih i generiranih uzoraka, koristeći polinomnu jezgru [4]. Ove mjere pružaju objektivne pokazatelje uspješnosti generativnih modela poput MedNeRF-a te pomažu u finom podešavanju mreže tijekom procesa, gdje se svakih N_{FID_KID} iteracija izračunava FID i KID, te se sprema najbolje postignuto stanje mreže do tog trenutka. Mreža je trenirana (XY) iteracija s parametrom N_{FID_KID} postavljenim na (XY).

5.1. Dotreniranje i rekonstrukcija na dvije slike

Za potrebe testiranja mreže, MedNeRF koristi prijašnje spomenutu skriptu stvorenu za dotreniranje i rekonstrukciju nad jednim digitalno rekonstruiranim radiografom. Skripta zahtjeva konfiguracijsku datoteku kojom se opisuje okruženje u kojem je trenirana mreža, spomenutu sliku, to jest radiograf i poziciju odakle je dobivena ta slika. S pomoću tih ulaznih komponenata, mreža se prilagođava i optimizira na jednu sliku za potrebe trodimenzionalne rekonstrukcije unutar pune vertikalne rotacije tog medicinskog primjera. Ovaj proces uključuje *fine-tune*-anje parametara generatora G s latentnim vektorima z . Njihova skripta minimizira gubitak koji je suma percepcijskog gubitka, to jest gubitka rekonstrukcije (5.1) sa gubitkom kvadratične pogreške (MSE) i negativnom logaritmu vjerojatnosti gubitka (NLLL) (5.4). Uz to, kvaliteta generiranja slike praćena je mjerama *Peak Signal to Noise Ratio* (PSNR) i *Structural Similarity Index Measure* (SSIM). PSNR uspoređuje originalnu i rekonstruiranu sliku mjereći omjer između maksimalne snage signala (najveća moguća razina signala) i srednje kvadratne pogreške (prosječna razlika između svakog piksela na originalnoj i rekonstruiranoj slici). SSIM procjenjuje sličnost između slika, uzimajući u obzir percepciju ljudskog oka mjereći tri ključna faktora: kontrast, strukturu i svjetlost.

$$\mathcal{L}_{\text{gen}} = \lambda_1 \mathcal{L}_r(\text{VGG16}) + \lambda_2 \mathcal{L}_{\text{MSE}}(G) + \lambda_3 \mathcal{L}_{\text{NLL}}(z) \quad (5.4)$$

Za potrebe rekonstrukcije koristeći dvije slike napisana je posebna skripta s kojom je cilj poboljšati rezultate rekonstrukcije i stvoriti vjerodostojan 3D model pacijenta. U suštini ta skripta se ne razlikuje značajno, ona također koristi konfiguracijsku datoteku kao ulaz, te uz nju umjesto jedne slike i pozicije kamere ona koristi dvije. Slike su dobivene postavljanjem kamere u generaciji DRR-ova na kuteve od 45° i 135°s prednje strane abdomena. Vizualni prikaz ovakvog koncepta prikazan je na slici 5.3 gdje zamišljamo pacijenta kako leži na leđima i ispod njega pod kutevima razmaknutima za 45° i 135°u negativnom smjeru su postavljeni su generatori rendgenskog zračenja. Ovaj pristup, koristeći dvije slike s različitih kutova, omogućuje MedNeRF modelu dublji uvid u unutarnju strukturu pacijenta, smanjujući potencijalne artefakte i poboljšavajući kvalitetu trodimenzionalne rekonstrukcije. Dotreniranje modela na ovaj način osigurava da se model prilagođava specifičnostima pojedinih pacijenata, pridonoseći preciznijim i vjerodostojnijim rezultatima.



Slika 5.3: Prikaz strukture stvaranja DRR-ova, slika dobivena s [26]

Dotreniranje modela provodit će se tako da se zatraži generator da izgenerira dvije umjetne slike iz istih pozicija kao što su i referentne, to jest stvarne slike uz korištenje latentnih vektora kao ulaz. Slike se zatim uspoređuju koristeći novi ukupni gubitak

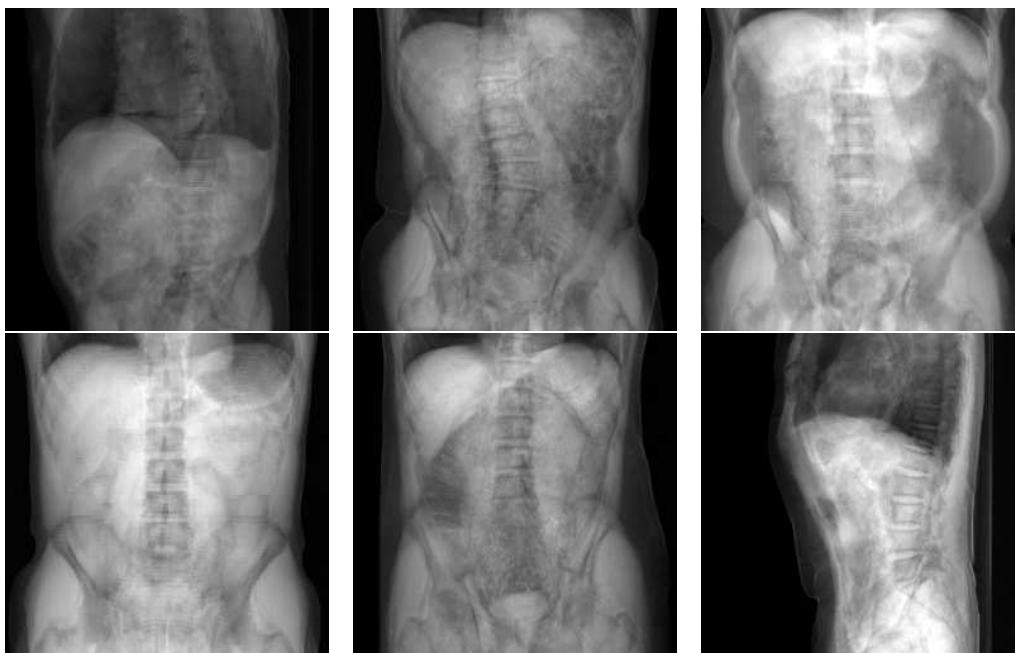
koji je osmišljen kao suma za obje slike prijašnje korištenog gubitka $\mathcal{L}_{\text{gen}}$ sa PSNR i SSIM gubitcima. PSNR gubitak je samo $PSNR_{ref} - PSNR$ to jest razlika između referentnog PSNR (vrijednost pri kojoj se dotreniranje zaustavlja) i izračunatog. Ovo osigurava da model konvergira prema željenom PSNR-u, što ukazuje na visoku kvalitetu rekonstrukcije. SSIM gubitak, definiran kao $1 - SSIM$, dodan je kako bi se osiguralo da struktura, kontrast i svjetlina rekonstruirane slike ostaju slični onima na stvarnim rendgenskim snimkama. Ova mjera doprinosi očuvanju detalja i teksture u konačnoj rekonstrukciji. Ukupni gubitak $\mathcal{L}_{\text{new}}$ zapisan je u jednadžbi 5.5. Intenziteti su podešeni da reguliraju utjecaj na ukupni gubitak, λ_4 i λ_5 postavljeni su na 0.5.

$$\mathcal{L}_{\text{new}} = \sum_{k=1}^2 (\mathcal{L}_{\text{gen}}(k) + \lambda_4(PSNR_{ref} - PSNR(k)) + \lambda_5(1 - SSIM(k))) \quad (5.5)$$

Skripta rutinski generira video rotacije oko vertikalne osi modela radiografa svakih N_{render} iteracija, pružajući vizualni uvid u postignutu kvalitetu rekonstrukcije. Zaustavljanje procesa do-treniranja i finalno generiranje trodimenzionalnog modela kontrolira se kroz hiperparametar $PSNR_{ref}$, kojim se određuje rubna vrijednost izračunatog omjera signala. Taj hiperparametar određuje prag PSNR-a, a njegovo pažljivo podešavanje ključno je za postizanje optimalne ravnoteže između kvalitete rekonstrukcije i trajanja procesa do-treniranja. Ovaj pristup osigurava da rezultati odražavaju vjerodostojnost rekonstruiranih modela pacijenata, prilagođenih specifičnostima njihovih anatomskih struktura.

6. Rezultati

U ovom poglavlju detaljno ćemo analizirati i predstaviti rezultate treniranja i rekonstrukcije modela pacijenta koristeći *MedNeRF* mrežu. Za potrebe analize rezultata trenirane su 3 *MedNeRF* mreže, svaka na posebnom skupu podataka. Podaci, točnije digitalno rekonstruirani radiografi generirani su iz podskupa prijašnje spomenutog AbdomenCT-1K skupa podataka. Taj podskup se sastoji od 53 pacijenta čiji skenovi su karakterizirani poprilično sličnim strukturama, osiguravajući ujednačenost treniranja. Prvi skup podataka sastoji se od 36 slika po pacijentu, s rotacijom svakih 10 stupnjeva oko pacijenta, drugi skup podataka sadrži 72 slike s rotacijom svakih 5 stupnjeva, dok je treći skup podataka sačinjen od 144 slike s razmakom od 2.5 stupnjeva između svake slike. Primjeri generiranih digitalno rekonstruiranih radiografa moguće je vidjeti u nastavku rada.



Slika 6.1: Primjer DRR-ova u različitim pozicijama (rezolucija - 256 x 256 px)

Mreže su trenirane 120 000 iteracija s osnovnim parametrima korištenim u *MedNeRF* radu, s proračunavanjem *Fréchet Inception Distance*-a i *Kernel Inception Distance*-a svakih 1000 iteracija za odabir najboljeg modela. To su mjere pojašnjene u prijašnjem poglavlju, te će poslužiti s procjenu najboljeg modela u treniranju. Nadalje, svi digitalno rekonstruirani rendgenski radiografi su skalirani na rezoluciju od (96 x 96) umjesto exportane (256 x 256) rezolucije zbog ograničenja hardvera potrebnog za rekonstrukciju.

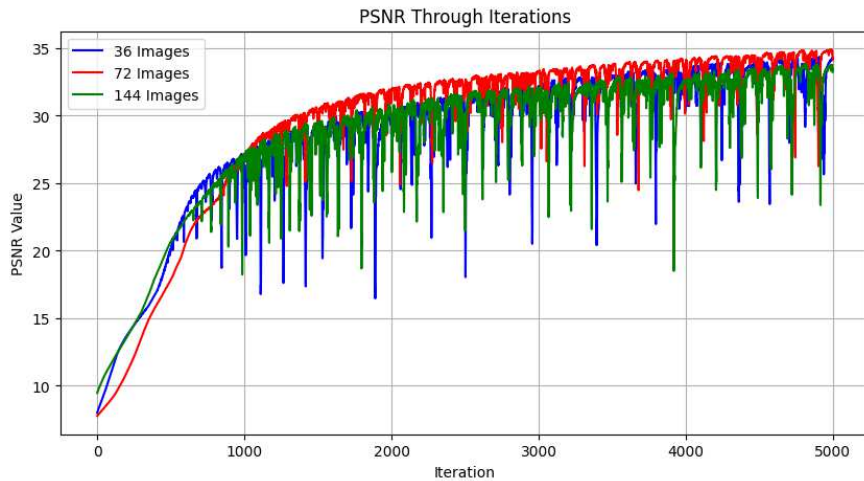
Broj slika	Fréchet Inc. Distance (FID)	Kernel Inc. Distance (KID)
36 slika	62.30	3.42
72 slike	53.81	2.66
144 slike	66.89	4.25

Tablica 6.1: FID i KID vrijednosti za najbolji model s obzirom na veličinu skupa podataka

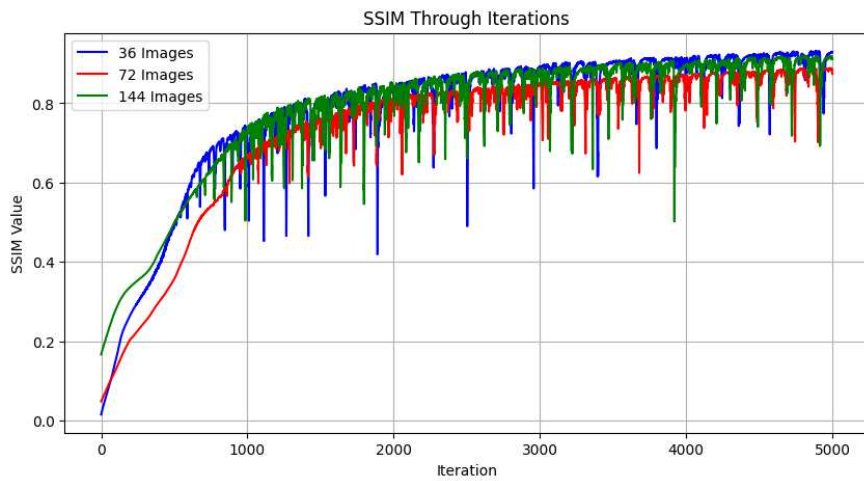
Dobivene procjene najboljih modela koristeći FID i KID vidljive su u tablici 6.1. U toj tablici vidimo da pri korištenju 72 slike za treniranje ovaj model ima najbolje rezultate zbog takve prilagodbe *MedNeRF* algoritma u kojem koristi 72 pogleda, a ostale generira. Zanimljiv rezultat imala je mreža trenirana na 144 slike, očekujući poboljšane rezultate zbog veće količine podataka, ona zapravo iskazuje gore vrijednosti od 3 uspoređivane mreže. Ali, unatoč tim višim FID i KID vrijednostima, ne mora se nužno pretpostavljati gora kvaliteta rekonstrukcije modela.

Na grafovima 6.4 i 6.5 vidljivi su rezultati rekonstrukcije to jest mjere PSNR i SSIM za 3 trenirane mreže kroz 5000 iteracija do-treniranja na jednom od primjera iz skupa podataka. Prva spomenuta slika 6.4 predstavlja standardnu *MedNeRF* rekonstrukciju iz jedne slike, to jest jednog pogleda, dok 6.5 predstavlja rekonstrukciju iz dvaju pogleda. Područje PSNR metrike jest od 0 do beskonačnosti, a SSIM između 0 i 1, gdje obje reprezentiraju bolje rezultate s visokim vrijednostima, a lošije sa nižim. Iz oba grafa moguće je primijetiti da po PSNR mjeri, mreža trenirana na 72 slike po pacijentu ima najbolje rezultate i najmanje oscilacije u vrijednosti, dok gledajući SSIM metriku rezultati su dosta isprepleteni i nije izražena dominacije nijedne mreže. Ovo opažanje je u skladu s rezultatima tablice 6.1 u kojoj također mreža trenirana na skupu u kojem svaki pacijent je predstavljen sa 72 slike ima najbolje rezultate. Ostala opažanja ispod koristit će tu mrežu za dobivanje numeričkih rezultata.

Rezultati tablice 6.2 dobiveni su koristeći najbolju mrežu, tj mrežu treniranu na pacijentima gdje svaki se sastoji od 72 slike. za potrebe dobivanja usporednih rezultata



(a) PSNR

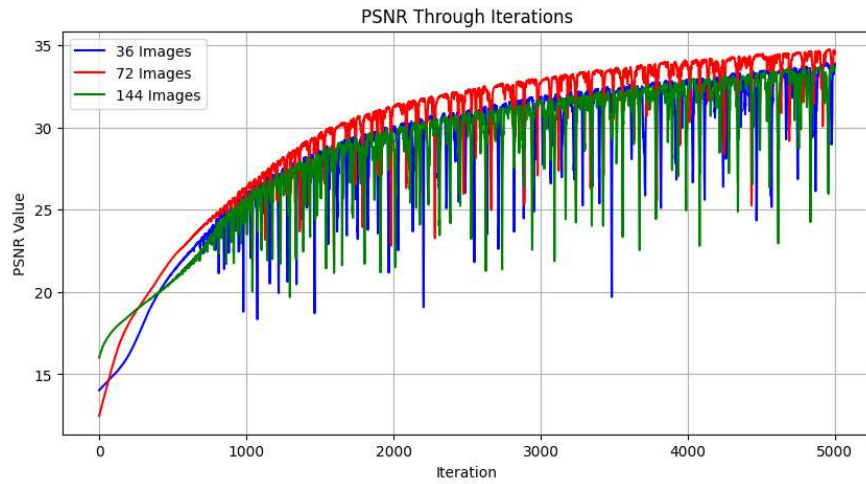


(b) SSIM

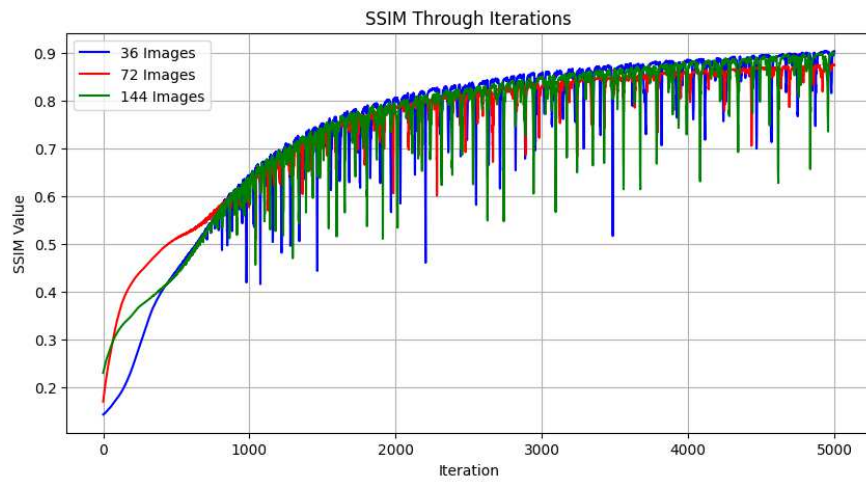
Slika 6.2: Primjer mjera PSNR i SSIM s obzirom na broj slika korištenih u treniranju koristeći *MedNeRF* algoritam za rekonstrukciju iz jedne slike

korištena su 3 pacijenta, vidljiva u prvom redu na skupovima slika 6.7, 6.8 i 6.9. Za te pacijente se izračunavala PSNR i SSIM vrijednost na cijeloj rotaciji oko vertikalne osi. Ideja je bila uzeti prosječnu vrijednost i proračunati joj devijaciju. Na sva tri pacijenta PSNR i SSIM vrijednosti su veće pri korištenju dvije slike te je iz tablice moguće zaključiti da korištenjem dvaju slika pri rekonstrukciji dobivamo znatno bolje rezultate.

Na grafovima 6.4, 6.5 i 6.6 prikazani su prijašnje spomenuti izračuni PSNR i SSIM kroz 360 stupnjeva oko vertikalne osi prikaza abdomena. Jasno je vidljivo da osim intenzivnih vrhova na dvije pozicije umjesto jedne, mreža rekonstrukcije s dvije slike je kroz skoro cijeli okret vidljivo bolja u obje mjere kvalitete.



(a) PSNR slika

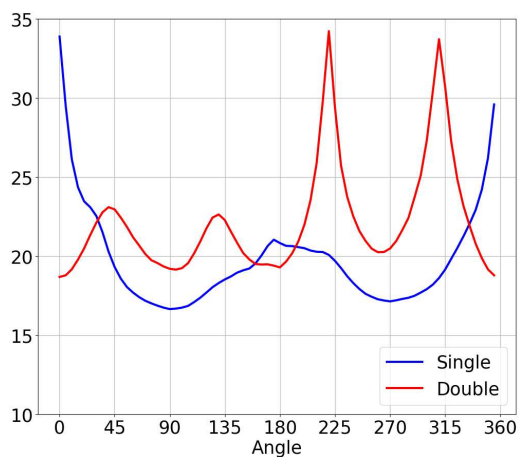


(b) SSIM slika

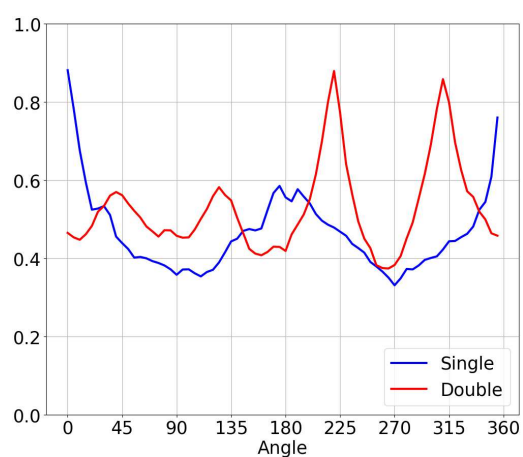
Slika 6.3: Primjer mjera PSNR i SSIM s obzirom na broj slika korištenih u treniranju koristeći naš algoritam za rekonstrukciju iz dvaju slika

	Jedna slika		Dvije slike	
	PSNR ($\mu \pm \sigma$)	SSIM ($\mu \pm \sigma$)	PSNR ($\mu \pm \sigma$)	SSIM ($\mu \pm \sigma$)
Pacijent 1	19.81 ± 3.26	0.46 ± 0.10	22.13 ± 3.46	0.53 ± 0.11
Pacijent 2	16.19 ± 3.51	0.36 ± 0.14	19.89 ± 3.48	0.52 ± 0.12
Pacijent 3	17.44 ± 3.39	0.40 ± 0.13	19.86 ± 3.66	0.49 ± 0.12

Tablica 6.2: Usporedba PSNR i SSIM rezultata na cjelokupnom volumenu koristeći mrežu treniranu na 72 pogleda oko pacijenta

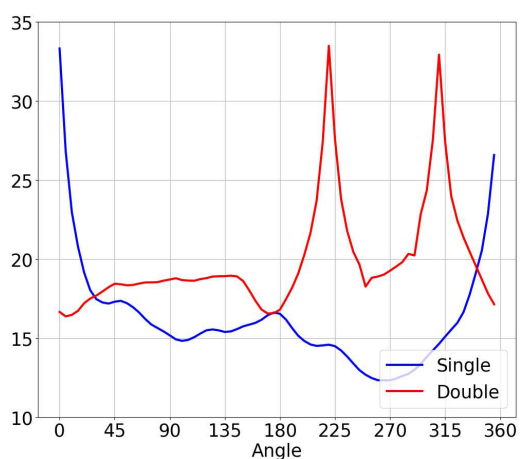


(a) PSNR

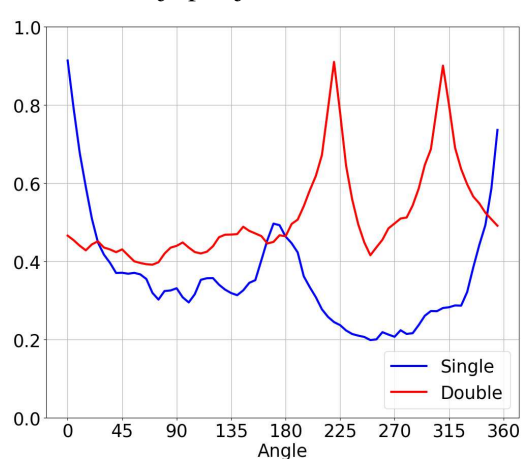


(b) SSIM

Slika 6.4: Rezultati PSNR i SSIM oko rotacije pacijentu 1

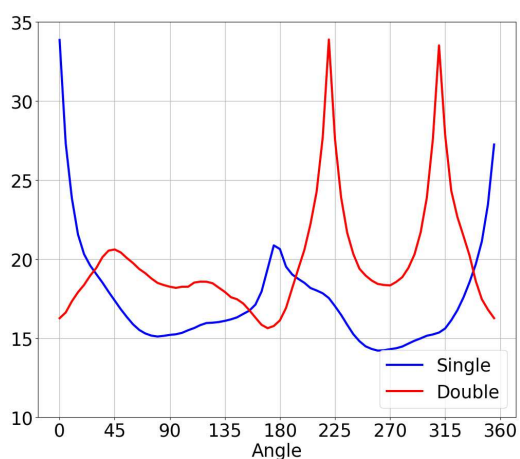


(a) PSNR

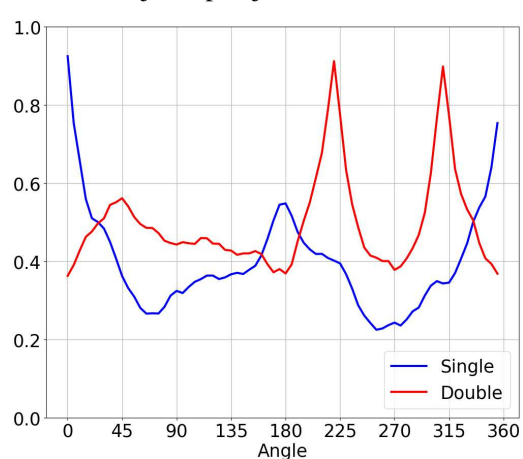


(b) SSIM

Slika 6.5: Rezultati PSNR i SSIM oko rotacije na pacijentu 2

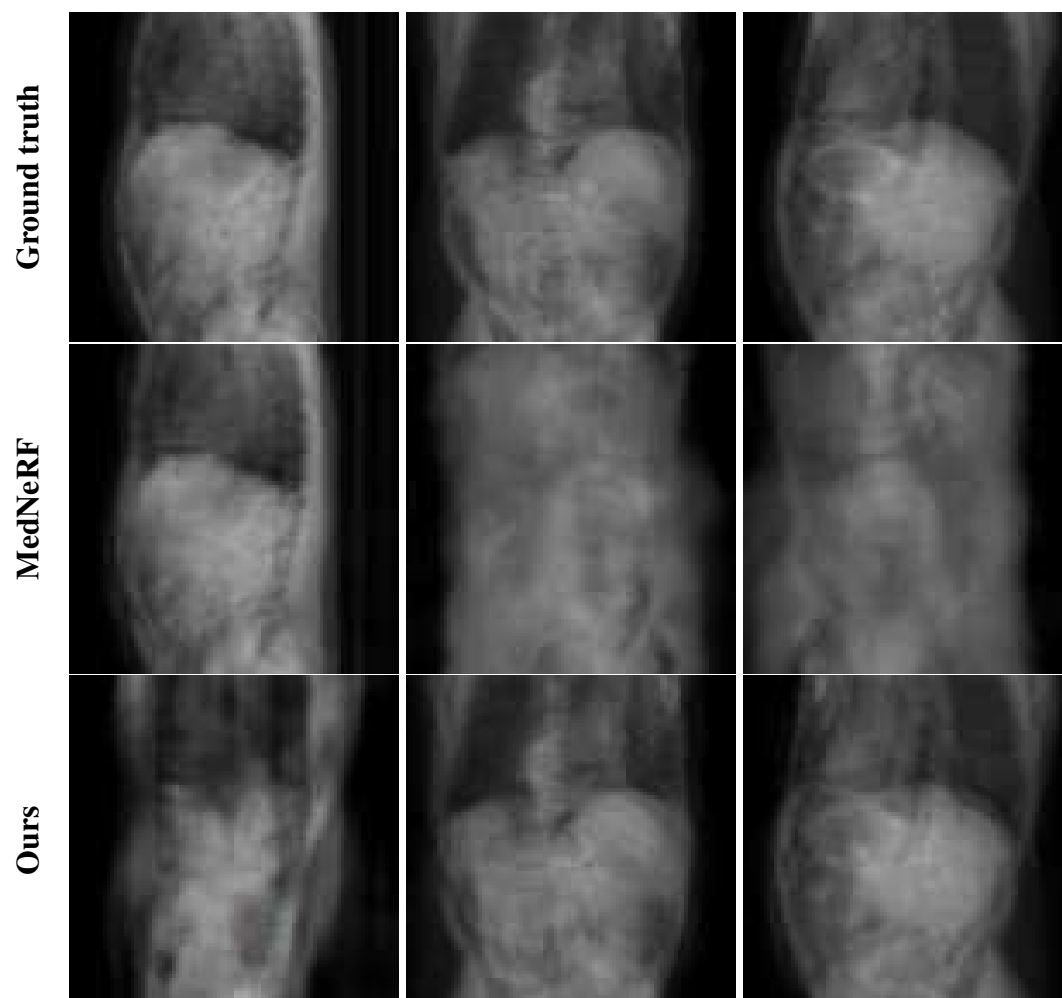


(a) PSNR

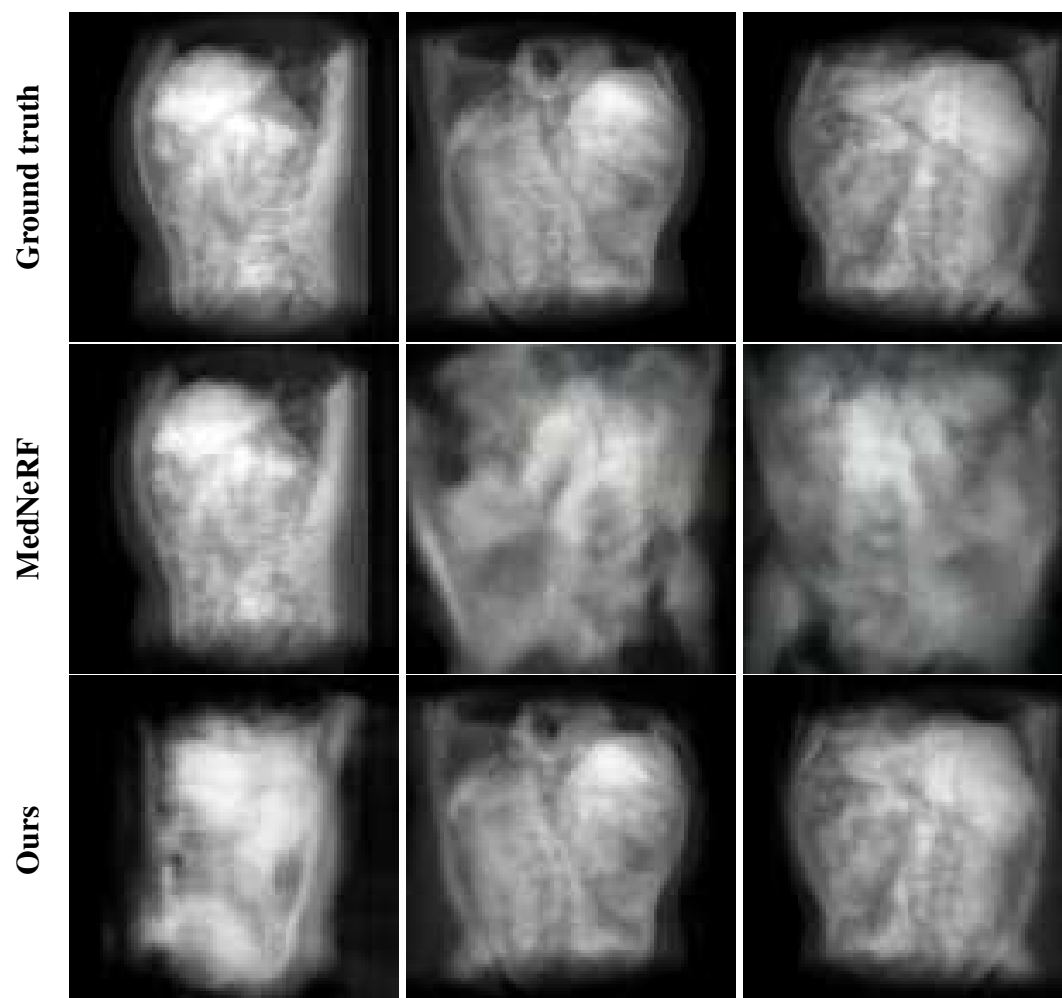


(b) SSIM

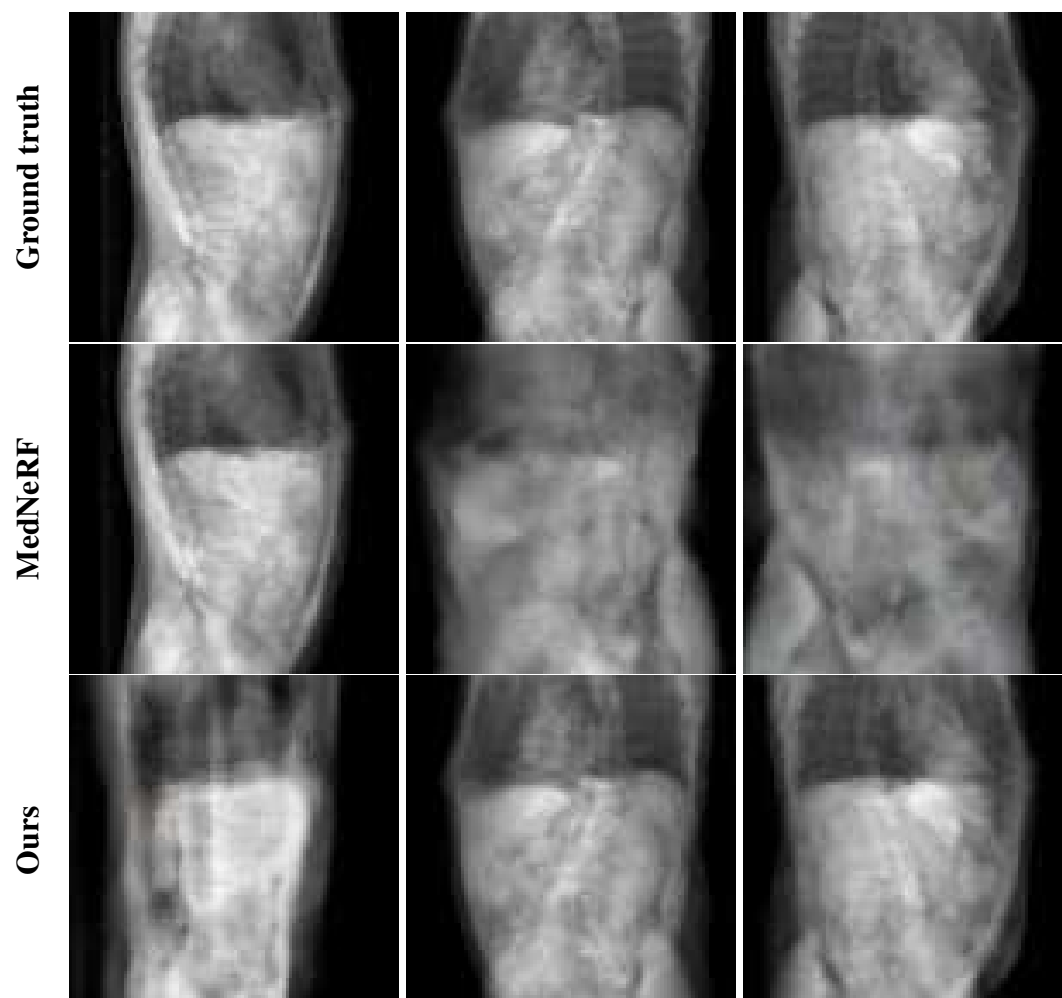
Slika 6.6: Rezultati PSNR i SSIM oko rotacije na pacijentu 3



Slika 6.7: Primjeri rekonstrukcije za prvog pacijenta koristeći metodu iz jedne slike i metodu iz dvije



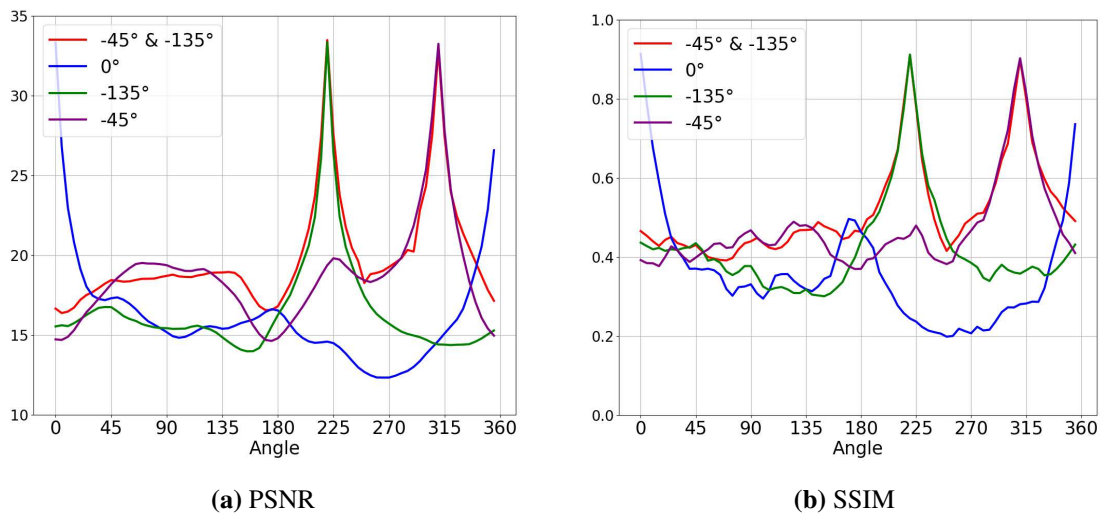
Slika 6.8: Primjeri rekonstrukcije za drugog pacijenta koristeći metodu iz jedne slike i metodu iz dvije



Slika 6.9: Primjeri rekonstrukcije za trećeg pacijenta koristeći *MedNeRF* metodu iz jedne slike i našu metodu iz dvije

Vizualne reprezentacije rekonstrukcija na prijašnje spomenuta 3 pacijenta vidljiva su na slikama 6.7, 6.8 i 6.9. Prvi red predstavlja takozvani *ground truth* to jest izgled kojem teži rekonstrukcija, drugi red predstavlja rekonstrukciju iz jedne slike koristeći njihovu funkciju gubitka, dok treći red predstavlja našu rekonstrukciju iz dvaju slika uz korištenje naše funkcije gubitka vidljive na 5.1.

Uz testiranja rekonstrukcije pri kojem se uspoređuje zadana struktura *MedNeRF* algoritma s našim, za potrebe analize provedeno je ispitivanje mijenjanja pozicije pri rekonstrukciji s jednom slikom s ciljem provjere postoji li ikakvo poboljšanje u kvaliteti ako se koristi drugačiji kut snimanja tj vizualizacije. Ti rezultati vidljivi su na slici 6.10 i ukazuju na jednaki problem koji je osnovni pogled (0°) imao pri rekonstrukciji iz jedne slike, a to je iznimno dobar rezultat točno u toj poziciji, dok lošiji u svim ostalima u usporedbi s rekonstrukcijom iz dva pogleda (-45° i -135°).

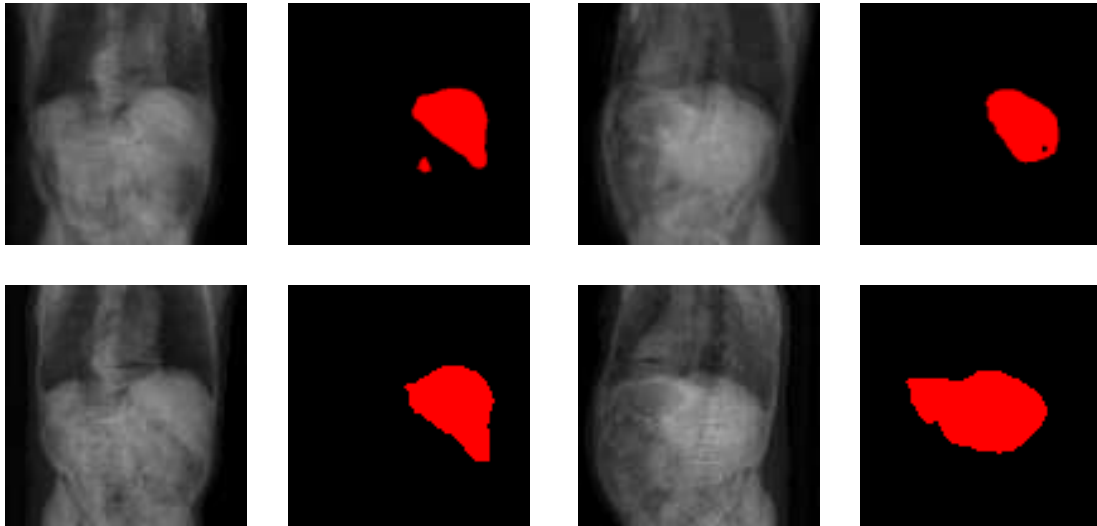


Slika 6.10: Rezultati PSNR i SSIM pri različitim kutevima rekonstrukcije

Posljednji eksperiment koji smo proveli za potrebe ovog rada bio je testirati detekciju oblika organa, preciznije, segmentaciju jetre korištenjem već utrenirane ENet mreže na parovima DRR-ova dobivenih putem CT2XRAY softvera. Spomenuti parovi DRR-ova generirani su na istim pozicijama kao prijašnje spomenuto, 45 i 135 stupnjeva negativne rotacije oko pacijenta. Ideja je testirati detekciju jetre na generiranim slikama i usporediti te rezultate sa *ground truth* segmentacijskim maskama tih slika. Eksperimenti bit će prikazani na jednom pacijentu, ali rezultati su ujednačeni kroz ostatak skupa podataka.

Prvi eksperiment sastoji se od korištenja mreže dotrenirane za rekonstrukciju iz -45 i -135 stupnjeva, izvlačenja reprezentativnih slika u tim pozicijama i uspoređivanju detekcije sa stvarnom maskom. Očekivani rezultati su da mreža iznimno dobro de-

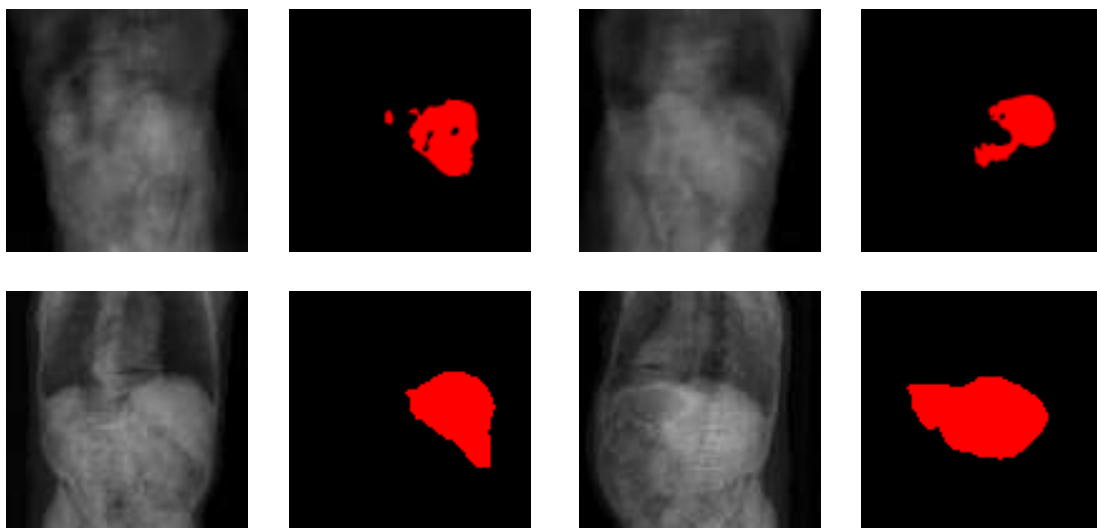
tektira jetru s obzirom na to da su detektor i *MedNeRF overfittane* na tim pogledima. Prikaz rezultata detekcije vidljiv je na 6.11, gdje u prvom redu vidimo *ground truth* sliku i njenu masku za kuteve -45 i -135, dok u drugom redu vidimo generiranu sliku i detektiranu masku na istim kutevima. Mreža ima poprilično dobre rezultate kao što je i očekivano.



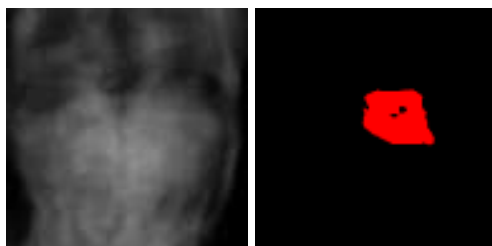
Slika 6.11: Prvi eksperiment - rekonstrukcija iz -45 i -135 i detekcija na istim kutevima

Drugi eksperiment zapravo dotrenira mrežu za potrebe rekonstrukcije na kutevima 90 i -90, ali koristi za detekciju generirane slike na istim kutevima kao prijašnje spomenuto s ciljem testiranja kvalitete generacije ne-viđenih pogleda i detekcije na istim. Usporedba rezultata vidljiva je na slici 6.12 gdje je raspored slika jednak kao u prošlom eksperimentu. Unatoč tome, iako je detekcija pokazala smanjenu preciznost zbog nedostatka obuke na tim specifičnim pogledima, mreža je ipak uspjela djelomično prepoznati jetru.

U posljednjem eksperimentu, također smo koristili dotreniranu mrežu kao i u prethodnom, ali umjesto generiranja slika pod ciljanim kutevima od 90 i -90 stupnjeva, generirane su slike u blizini tih kuteva. Cilj je bio usporediti kvalitetu detekcije na pogledima za koje mreža nije bila dotrenirana. Na prikazu 6.13 vidljiv je jedan takav pogled, točnije pogled pod kutem -55 stupnjeva, s rezultatom detekcije jetre na njemu. Kao što je očekivano, mreža je u ovom slučaju pokazala znatno lošije rezultate detekcije jetre, budući da rekonstrukcija nije bila izvršena iz naučenih pogleda i mreža nije bila trenirana za prepoznavanje jetre na -55 stupnjeva.



Slika 6.12: Drugi eksperiment - rekonstrukcija iz -90 i 90, ali detekcija na -45 i -135



Slika 6.13: Treći eksperiment - rekonstrukcija iz -90 i 90, ali detekcija na -55

Putem analize ovih rezultata očito je vidljivo da prvotno najbolja količina pogleda za treniranje *MedNeRF* mreže za potrebe rekonstrukcije oblika organa, u ovom slučaju abdomena jest 72 slike oko vertikalne osi pacijenta, gdje se svaka slika uzima s pomakom u rotaciji od 5 stupnjeva. Ovi nalazi ističu poboljšanje u kvaliteti rekonstrukcije kada se primjenjuje navedeni pristup. Međutim, kako bi se dublje razumjela dinamika ovog procesa, daljnja istraživanja mogu biti od koristi kako bi se detaljnije istražile različite konfiguracije i njihov utjecaj na rezultate.

7. Zaključak

Svrha ovog rada bila je istraživanje dostupnih rješenja za rekonstrukciju 3D modela iz dvodimenzionalnih slika korištenjem metoda koje ukomponiraju Neural Radiance Fields (NeRF). Ideja je korištenje NeRF-a ili njegovih varijacija za rekonstrukciju oblika organa iz para rendgenskih slika s motivacijom poboljšanja medicinske dijagnostike i olakšanja rada medicinskog osoblja.

Skup podataka sastojao se od sintetički generiranih digitalno rekonstruiranih radiografa dobivenih korištenjem CT2XRAY metode na AbdomenCT-1K skupu slika računalne tomografije. Ovaj skup podataka omogućio je analizu i evaluaciju pristupa rekonstrukciji, pružajući vrijedan uvid u performanse i mogućnosti različitih metoda.

MedNeRF metoda treniranja i rekonstrukcije pokazala se kao izrazito kvalitetan i efikasan pristup problemu rekonstrukcije trodimenzionalnog modela iz dvodimenzionalnih slika, posebice na medicinskim slikama. Naša prilagodba za navedeni problem sastojala se od nadogradnje koda za korištenje dvije slike odjednom, ali i nove funkcije gubitka. Evaluacija modela pri treniranju sastojala se od provjeravanja Fréchet Inception Distance i Kernel Inception Distance, dok se pri rekonstrukciji kvaliteta evaluacije provjeravala mjerama Peak Signal to Noise Ratio (PSNR) i Structural Similarity Index Measure (SSIM).

Analiza rezultata svih mjera dovela je do zaključaka o odabiru optimalnih količina slika po pacijentu, ali i razlikama u kvaliteti rekonstrukcije koristeći MedNeRF metodiku i našu nadograđenu metodiku iz dvije slike. S tim rezultatima dolazimo do zaključka da postoji značajno poboljšanje pri korištenju naše metodologije za rekonstrukciju oblika organa iz dvodimenzionalnih slika.

Za daljnji rad vrijedno bi bilo istražiti poboljšanje spomenute metode, primjerice putem analize učinka više rezolucije slika na kvalitetu rekonstrukcije. Također, razmatranje alternativnih funkcija gubitka ili eksperimentiranje s drugim varijacijama modela moglo bi pružiti bolje rezultate za probleme rekonstrukcije. Ovi pristupi mogu pomoći u daljnjem unaprjeđenju rekonstrukcije 3D modela iz dvodimenzionalnih slika i proširiti primjenu ove tehnike u medicinskoj dijagnostici.

LITERATURA

- [1] Henrik Aanæs, Rasmus Ramsbøl Jensen, George Vogiatzis, Engin Tola, i Anders Bjorholm Dahl. Large-scale data for multiple-view stereopsis. *Int. J. Comput. Vision*, 120(2):153–168, nov 2016. ISSN 0920-5691. doi: 10.1007/s11263-016-0902-9. URL <https://doi.org/10.1007/s11263-016-0902-9>.
- [2] Elena Balashova, Jiangping Wang, Vivek Singh, Bogdan Georgescu, Brian Teixeira, i Ankur Kapoor. 3d organ shape reconstruction from topogram images, 2019.
- [3] Jonathan T. Barron, Ben Mildenhall, Matthew Tancik, Peter Hedman, Ricardo Martin-Brualla, i Pratul P. Srinivasan. Mip-nerf: A multiscale representation for anti-aliasing neural radiance fields, 2021.
- [4] Eyal Betzalel, Coby Penso, Aviv Navon, i Ethan Fetaya. A study on the evaluation of generative models, 2022.
- [5] Simone Bianco, Gianluigi Ciocca, i Davide Marelli. Evaluating the performance of structure from motion pipelines. *Journal of Imaging*, 4:98, 08 2018. doi: 10.3390/jimaging4080098.
- [6] Zixuan Chen, Jian-Huang Lai, Lingxiao Yang, i Xiaohua Xie. Cunerf: Cube-based neural radiance field for zero-shot medical image arbitrary-scale super resolution, 2023.
- [7] Christopher B. Choy, Danfei Xu, JunYoung Gwak, Kevin Chen, i Silvio Savarese. 3d-r2n2: A unified approach for single and multi-view 3d object reconstruction, 2016.
- [8] Abril Corona-Figueroa, Jonathan Frawley, Sam Bond-Taylor, Sarath Bethapudi, Hubert P. H. Shum, i Chris G. Willcocks. Mednerf: Medical neural radiance fields for reconstructing 3d-aware ct-projections from a single x-ray, 2022.

- [9] Kangle Deng, Andrew Liu, Jun-Yan Zhu, i Deva Ramanan. Depth-supervised nerf: Fewer views and faster training for free, 2022.
- [10] Frano Petric, LARICS. Ct2xray. URL <https://larics.fer.hr/>. Posjećeno: 3.1.2024.
- [11] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, i Yoshua Bengio. Generative adversarial networks, 2014.
- [12] Matthew D Howland, Falko Kuester, i Thomas E Levy. Photogrammetry in the field: Documenting, recording, and presenting archaeology. *Mediterranean Archaeology and Archaeometry*, 14(4):101–108, 2014.
- [13] Hanwen Jiang, Zhenyu Jiang, Kristen Grauman, i Yuke Zhu. Few-view object reconstruction with unknown categories and camera poses, 2023.
- [14] LUMITOS AG. Digitally reconstructed radiograph. URL https://www.bionity.com/en/encyclopedia/Digitally_Reconstructed_Radiograph.html. Posjećeno: 3.1.2024.
- [15] Jun Ma, Yao Zhang, Song Gu, Cheng Zhu, Cheng Ge, Yichi Zhang, Xingle An, Congcong Wang, Qiyuan Wang, Xin Liu, Shucheng Cao, Qi Zhang, Shangqing Liu, Yunpeng Wang, Yuhui Li, Jian He, i Xiaoping Yang. Abdomenct-1k: Is abdominal organ segmentation a solved problem?, 2021.
- [16] Martin Oswald, Maria Klodt, Jörg Stücker, Prof. Dr. Daniel Cremers. Multi-view 3d reconstruction. URL https://cv.g.cit.tum.de/research/image-based_3d_reconstruction/multiviewreconstruction. Posjećeno: 12.10.2023.
- [17] Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, i Ren Ng. Nerf: Representing scenes as neural radiance fields for view synthesis, 2020.
- [18] Thomas Müller, Alex Evans, Christoph Schied, i Alexander Keller. Instant neural graphics primitives with a multiresolution hash encoding. *ACM Transactions on Graphics*, 41(4):1–15, jul 2022. doi: 10.1145/3528223.3530127. URL <https://doi.org/10.1145%2F3528223.3530127>.

- [19] Michael Oechsle, Songyou Peng, i Andreas Geiger. Unisurf: Unifying neural implicit surfaces and radiance fields for multi-view reconstruction, 2021.
- [20] Plastimatch development team. Plastimatch. URL <https://plastimatch.org/>. Posjećeno: 29.12.2023.
- [21] Johannes Lutz Schönberger i Jan-Michael Frahm. Structure-from-Motion Revisited. U *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
- [22] Katja Schwarz, Yiyi Liao, Michael Niemeyer, i Andreas Geiger. Graf: Generative radiance fields for 3d-aware image synthesis, 2021.
- [23] Matthew Tancik, Ethan Weber, Evonne Ng, Ruilong Li, Brent Yi, Terrance Wang, Alexander Kristoffersen, Jake Austin, Kamyar Salahi, Abhik Ahuja, David McAllister, Justin Kerr, i Angjoo Kanazawa. Nerfstudio: A modular framework for neural radiance field development. U *Special Interest Group on Computer Graphics and Interactive Techniques Conference Conference Proceedings*. ACM, jul 2023. doi: 10.1145/3588432.3591516. URL <https://doi.org/10.1145/3588432.3591516>.
- [24] Fei Tong, Megumi Nakao, Shuqiong Wu, Mitsuhiro Nakamura, i Tetsuya Matsuda. X-ray2shape: Reconstruction of 3d liver shape from a single 2d projection image. *svezak 2020*, stranice 1608–1611, 07 2020. doi: 10.1109/EMBC44109.2020.9176655.
- [25] Alex Yu, Vickie Ye, Matthew Tancik, i Angjoo Kanazawa. pixelnerf: Neural radiance fields from one or few images, 2021.
- [26] Chulong Zhang, Jingjing Dai, Tangsheng Wang, Xuan Liu, Yinping Chan, Liu Lin, Wenfeng He, Yaoqin Xie, i Xiaokun Liang. Xtransct: Ultra-fast volumetric ct reconstruction using two orthogonal x-ray projections via a transformer network, 05 2023.
- [27] Richard Zhang, Phillip Isola, Alexei A. Efros, Eli Shechtman, i Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. U *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, stranice 586–595, 2018. doi: 10.1109/CVPR.2018.00068.

Rekonstrukcija oblika organa iz para rendgenskih slika korištenjem neuronskih mreža

Sažetak

Ovaj rad istražuje primjenu Neural Radiance Fields (NeRF) za problem rekonstrukcije trodimenzionalnih modela iz dvodimenzionalnih rendgenskih slika. Korištenjem generiranih digitalno rekonstruiranih radiografa iz AbdomenCT-1K skupa podataka, analizirane su različite metode rekonstrukcije radi evaluacije performansi na rekonstrukciji modela. MedNeRF metoda, posebno prilagođena medicinskoj rekonstrukciji korištena je za potrebe izvedbe ovog rada. Evaluacija modela izvodila se tijekom treniranja putem FID i KID mjera kvalitete modela, dok pri rekonstrukciji koristile su se mjere PSNR i SSIM. Rezultate predstavljamo kao analizu optimalne količine slika po pacijentu, ali i usporedbu između standardne MedNeRF implementacije koristeći 1 sliku za rekonstrukciju, naspram naše gdje koristimo 2 slike i novu funkciju gubitka.

Ključne riječi: NeRF, medicina, 3D rekonstrukcija, 3D model, polje zračenja, DRR

Organ shape reconstruction from a pair of X-ray images using neural networks

Abstract

This paper investigates the application of Neural Radiance Fields (NeRF) to the problem of reconstructing three-dimensional models from two-dimensional X-ray images. Using generated digitally reconstructed radiographs from the AbdomenCT-1K dataset, different reconstruction methods were analyzed to evaluate the model's performance. The MedNeRF method, specially adapted to medical reconstruction, was used to conduct this work. Model evaluation was performed during training using FID and KID quality measures, while PSNR and SSIM measures were used during reconstruction. We present the results as an analysis of the optimal amount of images per patient, but also a comparison between the standard MedNeRF implementation using 1 image for reconstruction, versus ours where we use 2 images and a new loss function.

Keywords: NeRF, medicine, 3D reconstruction, 3D model, radiance fields, DRR