

# Segmentacija koljena difuzijskim modelima iz 3D slika magnetske rezonance

---

**Perić, Juraj**

**Master's thesis / Diplomski rad**

**2024**

*Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj:* **Josip Juraj Strossmayer University of Osijek, Faculty of Electrical Engineering, Computer Science and Information Technology Osijek / Sveučilište Josipa Jurja Strossmayera u Osijeku, Fakultet elektrotehnike, računarstva i informacijskih tehnologija Osijek**

*Permanent link / Trajna poveznica:* <https://um.nsk.hr/um:nbn:hr:200:562980>

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

*Download date / Datum preuzimanja:* **2024-11-04**

*Repository / Repozitorij:*

[Faculty of Electrical Engineering, Computer Science and Information Technology Osijek](#)



SVEUČILIŠTE JOSIPA JURJA STROSSMAYERA U OSIJEKU

FAKULTET ELEKTROTEHNIKE, RAČUNARSTVA I  
INFORMACIJSKIH TEHNOLOGIJA

Sveučilišni diplomski studij Računarstvo

SEGMENTACIJA KOLJENA  
DIFUZIJSKIM MODELIMA IZ 3D SLIKA  
MAGNETNE REZONANCE

Diplomski rad

Juraj Perić

Osijek, 2024.

**FERIT**FAKULTET ELEKTROTEHNIKE, RAČUNARSTVA  
I INFORMACIJSKIH TEHNOLOGIJA OSIJEK**Obrazac D1: Obrazac za ocjenu diplomskog rada na sveučilišnom diplomskom studiju****Ocjena diplomskog rada na sveučilišnom diplomskom studiju**

<b>Ime i prezime pristupnika:</b>	Juraj Perić
<b>Studij, smjer:</b>	Sveučilišni diplomski studij Računarstvo
<b>Mat. br. pristupnika, god.</b>	D1316R, 07.10.2022.
<b>JMBAG:</b>	0165082365
<b>Mentor:</b>	prof. dr. sc. Irena Galić
<b>Sumentor:</b>	dr. sc. Marija Habijan
<b>Sumentor iz tvrtke:</b>	
<b>Predsjednik Povjerenstva:</b>	doc. dr. sc. Hrvoje Leventić
<b>Član Povjerenstva 1:</b>	prof. dr. sc. Irena Galić
<b>Član Povjerenstva 2:</b>	doc. dr. sc. Krešimir Romić
<b>Naslov diplomskog rada:</b>	Segmentacija koljena difuzijskim modelima iz 3D slika magnetske rezonance
<b>Znanstvena grana diplomskog rada:</b>	<b>Umjetna inteligencija (zn. polje računarstvo)</b>
<b>Zadatak diplomskog rada:</b>	Segmentacija koljena iz slika magnetne rezonance važna je za praćenje progresije bolesti poput osteoartritisa i planiranje odgovarajuće terapije. U ovom radu predstavljen je model za automatsku segmentaciju koljena korištenjem difuzijskih modela, metode koja je u posljednje vrijeme postala popularna u području generiranja slika. Model koristi U-Net arhitekturu s modifikacijama prilagođenim za difuzijske procese, uključujući dinamičko uvjetno kodiranje i FF-pretvarač koji su predloženi u radu <i>Medical Image Segmentation with Diffusion Probabilistic Model</i> .
<b>Datum ocjene pismenog dijela diplomskog rada od strane mentora:</b>	09.09.2024.
<b>Ocjena pismenog dijela diplomskog rada od strane mentora:</b>	Izvrstan (5)
<b>Datum obrane diplomskog rada:</b>	13. 9.2024.
<b>Ocjena usmenog dijela diplomskog rada (obrane):</b>	Izvrstan (5)
<b>Ukupna ocjena diplomskog rada:</b>	Izvrstan (5)
<b>Datum potvrde mentora o predaji konačne verzije diplomskog rada čime je pristupnik završio sveučilišni diplomski studij:</b>	13.09.2024.



**FERIT**

FAKULTET ELEKTROTEHNIKE, RAČUNARSTVA  
I INFORMACIJSKIH TEHNOLOGIJA **OSIJEK**

## IZJAVA O IZVORNOSTI RADA

Osijek, 17.09.2024.

**Ime i prezime Pristupnika:**

Juraj Perić

**Studij:**

Sveučilišni diplomski studij Računarstvo

**Mat. br. Pristupnika, godina upisa:**

D1316R, 07.10.2022.

**Turnitin podudaranje [%]:**

10

Ovom izjavom izjavljujem da je rad pod nazivom: **Segmentacija koljena difuzijskim modelima iz 3D slika magnetske rezonance**

izrađen pod vodstvom mentora prof. dr. sc. Irena Galić

i sumentora dr. sc. Marija Habijan

moj vlastiti rad i prema mom najboljem znanju ne sadrži prethodno objavljene ili neobjavljene pisane materijale drugih osoba, osim onih koji su izričito priznati navođenjem literature i drugih izvora informacija.

Izjavljujem da je intelektualni sadržaj navedenog rada proizvod mog vlastitog rada, osim u onom dijelu za koji mi je bila potrebna pomoć mentora, sumentora i drugih osoba, a što je izričito navedeno u radu.

Potpis pristupnika:

# SADRŽAJ

<b>1. Uvod . . . . .</b>	<b>1</b>
<b>2. Pregled područja i problematike . . . . .</b>	<b>2</b>
2.1. U-Net . . . . .	2
2.2. DeepMedic . . . . .	3
2.3. Medical Image Segmentation with Diffusion Model . . . . .	3
2.4. Ostale metode . . . . .	4
<b>3. Klinička pozadina . . . . .</b>	<b>5</b>
<b>4. Prikupljanje i opis korištenih podataka . . . . .</b>	<b>7</b>
4.1. Magnetska rezonanca(MRI) . . . . .	7
4.2. Opis korištenog skupa podataka . . . . .	9
4.3. NIfTI format . . . . .	11
<b>5. Metodologija . . . . .</b>	<b>13</b>
5.1. Difuzijski modeli . . . . .	13
5.1.1. Proces difuzije prema naprijed . . . . .	13
5.1.2. Proces obrnute difuzije . . . . .	14
5.1.3. Treniranje difuzijskih modela . . . . .	15
5.1.4. Parametrizacija $\beta_t$ . . . . .	17
5.2. Konvolucijske neuronske mreže . . . . .	18
5.3. U-Net arhitektura . . . . .	21
5.4. Gaussovo zamućenje . . . . .	23
<b>6. Izrada modela . . . . .</b>	<b>24</b>
6.1. Podjela skupa podataka i predobrada . . . . .	24
6.2. Arhitektura modela . . . . .	25
6.2.1. Dinamičko uvjetno kodiranje . . . . .	25
6.2.2. FF-Pretvarač . . . . .	27
6.2.3. ResNet blok . . . . .	28
6.2.4. Blokovi pažnje . . . . .	28
6.2.5. Sinusoidalne pozicije . . . . .	28
6.2.6. Povezivanje komponenti u arhitekturi mreže . . . . .	29
6.3. Funkcija gubitka . . . . .	30

6.4. Metrike . . . . .	31
6.5. Parametri treninga . . . . .	31
<b>7. Analiza i naknadna obrada rezultata . . . . .</b>	<b>33</b>
7.1. Rezultati treninga. . . . .	33
7.2. Evaluacija i naknadna obrada. . . . .	35
7.3. Usporedba s ostalim modelima . . . . .	38
<b>8. Zaključak. . . . .</b>	<b>40</b>

# 1. UVOD

Magnetna rezonanca postala je jedan od najvažnijih alata u suvremenoj medicinskoj dijagnostici, omogućujući detaljan prikaz unutarnjih struktura tijela i analizu različitih patoloških stanja, uključujući ozljede i degenerativne promjene koljena. Unatoč visokoj kvaliteti prikaza, slike dobivene magnetnom rezonancom često sadrže šum, što može otežati njihovu preciznu interpretaciju. Kako bi se olakšala analiza i interpretacija ovih slika, predlažu se razne metode za automatsku segmentaciju koljena koja znatno olakšava razdvajanje područja interesa od okolnih struktura. Segmentacija koljena je posebno važna u praćenju osteoartritisa, degenerativne bolesti koja zahvaća kosti i zglobove. Kod osteoartritisa, posebno je bitno pratiti površinu hrskavice kako bi se pravodobno reagiralo na razvoj bolesti i pružila adekvatna terapija, što može biti otežavajuće s obzirom na to da je hrskavicu teško odvojiti od susjednih koštanih struktura zbog sličnog intenziteta signala na MRI slikama [1].

Segmentacija medicinskih slika jedan od najvažnijih koraka u suvremenoj dijagnostici, a razvijeni modeli koji koriste U-Net arhitekturu i njezine varijacije pokazali su se izuzetno učinkovitima u raznim zadacima segmentacije. U ovom radu analizirat će se performanse difuzijskog modela u usporedbi s klasičnim metodama.

Cilj ovog diplomskog rada je razviti i evaluirati metodu za automatsku segmentaciju MRI slika koljena difuzijskim modelima, koji su u zadnje vrijeme postali popularni u generiranju i segmentaciji slika zbog svoje sposobnosti stvaranja visokokvalitetnih sintetičkih prikaza. Korišteni skup podataka uključuje slike koljena snimljene MRI-om, a u radu će biti detaljno opisani postupci obrade i segmentacije tih slika.

Rad je strukturiran na sljedeći način. U drugom poglavlju bit će opisane trenutno aktualne metode segmentacije medicinskih slika. Sljedeće poglavlje detaljnije će se opisati kliničku pozadinu osteoartritisa i bitnost segmentacije za taj problem. Četvrto poglavlje opisuje korišteni skup podataka, kako je spremljen i opisuje strukture bitne za segmentaciju. Peto poglavlje nudi teorijski opis svih potrebnih tehnologija korištenih za izradu modela, uključujući detaljan opis rada difuzijskih modela, konvolucijskih neuronskih mreža, U-Net arhitekturu te Gaussovo zamućenje. U sljedećem poglavlju detaljno je opisana izrada modela te način i parametri za treniranje. U sedmom poglavlju vrši se analiza rezultata treninga, naknadna obrada podataka te usporedba s ostalim modelima.

## 2. PREGLED PODRUČJA I PROBLEMATIKE

Segmentacija koljena iz 3D slika magnetne rezonance je važan zadatak u medicinskoj dijagnostici koji se često koristi za analizu anatomskih struktura koljena, dijagnosticiranje ozljeda, praćenje razvoja bolesti i planiranje operativnih zahvata. U posljednjih nekoliko godina, razvijene su različite metode i arhitekture neuronskih mreža za segmentaciju koljena. Osim segmentacije koljena, slični problemi u medicinskoj dijagnostici koji su rješavani s pomoću MRI slika i neuronskih mreža uključuju segmentaciju drugih anatomskih struktura kao što su mozak, srce, pluća, jetra, bubreg, dojke i kralježnica. Osim same segmentacije anatomskih struktura, neuronske mreže se mogu koristiti i za detekciju i dijagnostiku bolesti kao što su tumori, multipla skleroza, artritis, artroza i ozljede ligamenata.

Također, neuronske mreže se koriste i za druga medicinska pitanja, kao što su predviđanje ishoda terapije, predviđanje vjerojatnosti pojave bolesti i procjena rizika na temelju snimke magnetne rezonance ili nekog drugog skupa podataka.

### 2.1. U-Net

Jedna od najčešće korištenih arhitektura za segmentaciju koljena iz 3D MRI slika je 3D U-Net. 3D U-Net arhitektura je proširenje 2D U-Net arhitekture koja se koristi za segmentaciju dvodimenzionalnih slika. Ova arhitektura ima nekoliko slojeva konvolucije, slojeva sažimanja (engl. *downsampling*) i slojeva za proširivanje dimenzija (engl. *upsampling*) koji omogućuju preciznu segmentaciju anatomskih struktura koljena iz 3D MRI slika. 3D U-Net arhitektura također koristi blokove konvolucije i sažimanja (engl. *poolinga*) koji pomažu u izdvajanju značajki iz slika i smanjuju broj parametara modela [3].

Kako je U-Net dugo vremena pokazivao dobre rezultate, istraživanja u tom području su se nastavila i razvijena je nnU-Net metoda koja se automatski prilagođava na zadani skup podataka i pokazuje odlične rezultate u tom području. Metoda analizira cijeli skup podataka, bio on dvodimenzionalan ili trodimenzionalan te prilagođava arhitekturu U-Net mreže i cijeli proces treniranja od predobrade do naknadne obrade podataka. Osim što se metoda prilagođava skupu podataka, nnU-Net će prilagoditi parametre treninga računalu na kojemu će se trening odvijati ovisno o njegovoj jačini. Ova metoda je bila revolucionarna s obzirom na to da nije potrebna visoka razina razumijevanja koncepata strojnog učenja kako bi se trening modela pokrenuo [4].



U-Net arhitektura će biti detaljnije opisana u poglavlju 5.3 jer će biti nužna za realizaciju metode ovoga rada.

## 2.2. DeepMedic

Još jedna metoda koja je pružala dobre rezultate u segmentaciji medicinskih slika je DeepMedic. Ova arhitektura specifično je dizajnirana za segmentaciju 3D volumetrijskih podataka pomoću konvolucijskih neuronskih mreža. Metoda koristi paralelne putanje s različitim razlučivostima gdje istovremeno analizira lokalne i globalne informacije iz volumena. Jedna putanja obrađuje podatke na visokoj rezoluciji kako bi uhvatila lokalne detalje, dok druga putanja obrađuje podatke na nižoj rezoluciji kako bi uhvatila globalni kontekst. Na kraju ovih paralelnih putanja, informacije se spajaju kako bi se postigla konačna segmentacija. DeepMedic koristi i posebno dizajnirane mehanizme za balansiranje između različitih klasa. Ovaj problem je vrlo čest u medicinskim podacima s obzirom na to da pozadina čini većinu segmentacije i da su manje strukture poprilično česte. Arhitektura je pokazala izvrsne rezultate u segmentaciji različitih anatomskih struktura u medicinskim slikama [5].

## 2.3. Medical Image Segmentation with Diffusion Model

U posljednjih nekoliko godina pojavila se nova metoda generiranja slika difuzijskim modelima. Modeli koriste proces difuzije za generiranje visokokvalitetnih umjetnih slika pokazujući odlične rezultate u tom području [9]. Samim time, pozitivno se odražavaju na problem segmentacije jer mogu generirati realističnije segmentacije.

*Medical Image Segmentation with Diffusion Model* je naziv projekta inspiriran segmentacijom koristeći difuzijske modele i transformere. Projekt ima dvije metode segmentacije te su obje metode dizajnirane samo za segmentaciju dvodimenzionalnih slika. Prva metoda koristi samo difuzijski model dok se u drugoj pojavljuje i transformer arhitektura. Druga metoda je nadmašila nekoliko prijašnjih metoda za segmentaciju dvodimenzionalnih slika mozga [7] [8]. Arhitektura prve metode je iskorištena u ovome radu i bit će dodatno pojašnjena u poglavlju 6.2.

## 2.4. Ostale metode

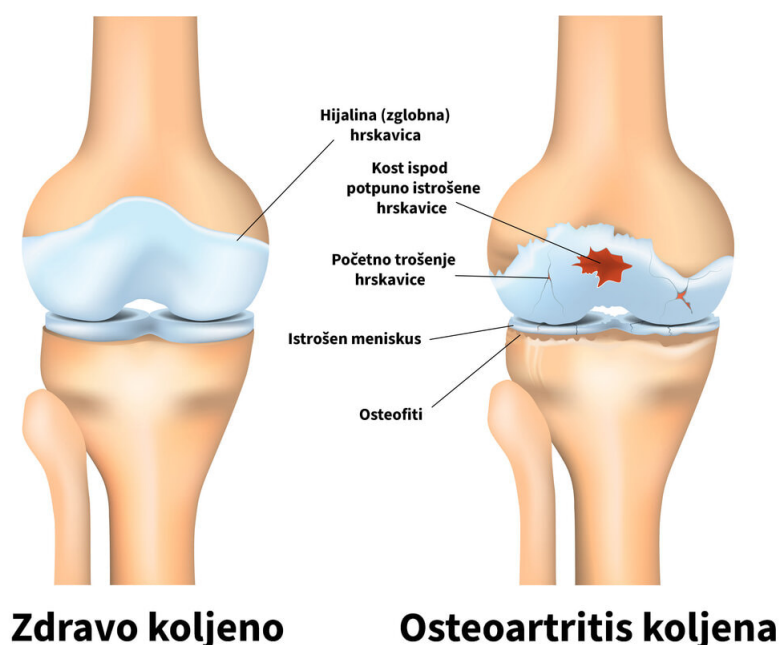
Uz spomenute, koriste se i druge metode za segmentaciju koljena iz 3D MRI slika, uključujući i metode bazirane na regiji interesa (engl. *region of interest*) i metode bazirane na prijenosnom učenju (engl. *transfer learning*). Metode bazirane na regiji interesa koriste dodatne informacije o regijama interesa na slici kako bi se poboljšala segmentacija. Metode bazirane na prijenosnom učenju koriste prethodno naučene modele kako bi se prenijela znanja iz drugih sličnih zadataka i poboljšala segmentacija koljena.

U zadnje vrijeme, uz korištenje neuronskih mreža za segmentaciju koljena iz 3D MRI slika, često se koristi i kombinacija različitih metoda kako bi se poboljšala kvaliteta i preciznost segmentacije.

### 3. KLINIČKA POZADINA

Osteoartritis je degenerativni poremećaj zglobova koji pogađa sve dobne skupine, a posebno one starije dobi. Glavni simptomi uključuju bol, oticanje i ograničenje pokreta u zahvaćenom zglobu, što rezultira postupnim propadanjem zglobne hrskavice. Zbog složenosti i progresivnog karaktera osteoartritisa, rano otkrivanje i praćenje promjena u zglobovima važno je za učinkovito liječenje i upravljanje ovom bolesti. Osteoartritis koljena jedna je od najčešćih vrsta osteoartritisa.

## Osteoartritis koljena



**Slika. 3.1:** Prikaz zdravog koljena i koljena zahvaćenog osteoartritisom [2]

Bitnost segmentacije dijelova koljena predstavlja identifikaciju i kvantifikaciju promjena u zglobu koje se javljaju tijekom osteoartritisa. Slike snimljene magnetnom rezonancom posebno su korisne za otkrivanje ranih promjena u zglobovima prije nego što postanu vidljive na rendgenskim snimkama. Segmentacija omogućuje razdvajanje dijelova koljena kao što su hrskavice, meniskusi, ligamenti i koštani dijelovi, što omogućuje precizniju analizu promjena u tim strukturama. Tijekom razvoja osteoartritisa bitno je pratiti kosti jer su podložne degeneraciji tijekom razvoja osteoartritisa, dok se hrskavice najčešće oštećuju tijekom bolesti. Hrkavica ima važnu ulogu u održavanju zdravlja zgloba jer služi kao glatka površina koja

omogućuje amortizaciju udaraca, klizanje kostiju tijekom pokreta i raspoređuje opterećenje. Ako se hrskavica istroši, potkoljениčna kost preuzima opterećenje i postepeno se stanjuje, što dodatno pogoršava stanje [1] [2]. Na slici 3.1 se nalazi primjer zdravog koljena i koljena zahvaćenim osteoartritisom. Slika jasno prikazuje promjene koje se mogu dogoditi na hrskavici, meniskusu i kosti tijekom razvoja bolesti.

Segmentacija dijelova koljena može biti korisna i za razvoj novih dijagnostičkih alata i terapija. Precizno mjerenje volumena i gustoće hrskavice može pomoći u procjeni učinkovitosti terapija za osteoartritis koljena. Također, segmentacija može biti korisna za praćenje napretka bolesti tijekom vremena i identificiranje novih dijagnostičkih značajki koje mogu biti korisne u identificiranju ranih stadija osteoartritisa.

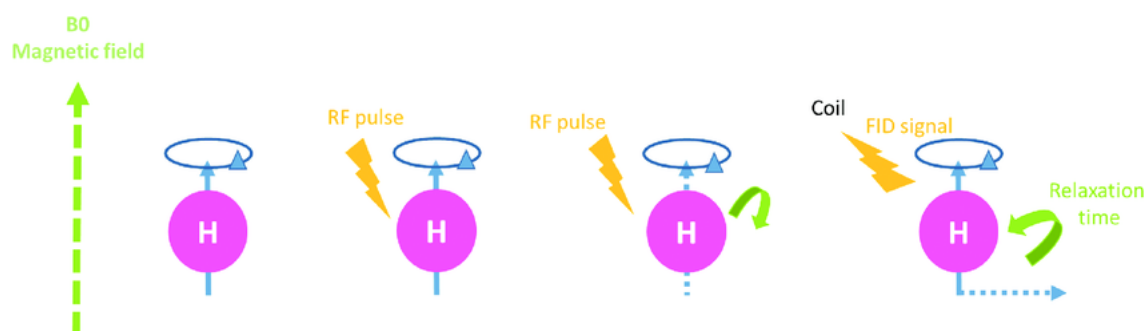
## 4. PRIKUPLJANJE I OPIS KORIŠTENIH PODATAKA

U ovom poglavlju opisan je skup podataka korišten za rješavanje problema, način na koji je prikupljan te format u kojemu je zapisan.

### 4.1. Magnetska rezonanca(MRI)

Magnetska rezonanca (MRI) je dijagnostička tehnologija koja koristi jak magnetni pol i radio valove za stvaranje detaljnih slika unutar tijela. MRI se često koristi u medicini za dijagnostiku i praćenje različitih zdravstvenih stanja i bolesti. Primjerice, MRI se može koristiti za analizu tumora, krvnih ugrušaka, upala, ozljeda i drugih abnormalnosti u tijelu.

MRI slike se dobivaju pomoću posebnog stroja, MRI skenera. Pacijent leži na pokretnom stolu koji se pomiče u tunelu MRI skenera ovisno o tome koji se dio tijela snima. Uređaj stvara jako homogeno magnetsko polje koje utječe na protone vodika, najzastupljenijeg elementa u ljudskom tijelu. Kada se tijelo pacijenta stavi u magnetsko polje MRI stroja, protoni vodika u tkivima se usmjere u smjeru polja. U tom stanju protoni vodika proizvode maleni magnetski dipol koji reagira na magnetsko polje MRI stroja. Tijekom snimanja, MRI stroj koristi radiofrekventne (RF) impulse koji stvaraju promjenjivo magnetno polje i uzrokuju ekscitaciju protona vodika, mijenjajući njihov položaj i rotaciju unutar tkiva. Kada RF puls prestane, protoni se vraćaju u prvobitno stanje, oslobađajući elektromagnetske signale koje antene u MRI stroju detektiraju. Slika 4.2 prikazuje proces kako atomi vodika reagiraju na radiofrekventne impulse i magnetsko polje MRI uređaja.



**Slika. 4.2:** Prikaz interakcije vodikovih atoma s magnetnim poljem i radio frekventnim impulsima [13]

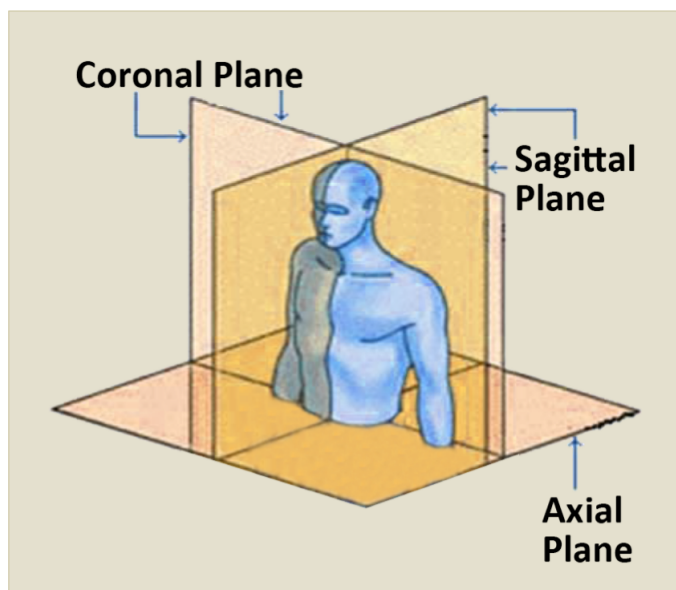
Signali koje hvataju antene su u osnovi analogni, kontinuirani signali koji se moraju digitalizirati prije nego što se mogu koristiti za stvaranje slike. Digitalizacija se provodi s pomoću analogno-digitalnih pretvarača koji pretvaraju kontinuirane signale u diskretne vrijednosti.

Nakon što se signali pretvore u diskretne vrijednosti, oni se organiziraju u matricu k-prostora, gdje svaka točka u matrici predstavlja jedan diskretni uzorak signala. Dimenzije matrice k-prostora ovise o veličini i razlučivosti slike koju se želi dobiti. Što je veća matrica k-prostora, to će slika biti detaljnija, ali će i vrijeme snimanja biti duže. Korištenjem matrice k-prostora moguće je primijeniti različite matematičke algoritme i transformacije na signale kako bi se dobile konačne slike. Jedan od najčešće korištenih algoritama je Fourierova transformacija koja omogućuje da se signali prikažu u frekvencijskoj domeni. Ovim postupkom dobivaju se podaci o različitim frekvencijama signala koje se zatim mogu koristiti za stvaranje slike u prostornoj domeni [12].

Tijekom magnetne rezonance mogu se dobiti slike tijela pacijenta u različitim ravninama koje se određuju položajem i orijentacijom magnetnog polja MRI stroja u odnosu na pacijenta. Tri ravnine u kojima se dobivaju MRI slike su:

1. Aksijalna ravnina (transverzalna ravnina) - ravnina koja dijeli tijelo pacijenta na gornju i donju polovicu. Kada se dobivaju MRI snimke u aksijalnoj ravnini, slike se prikazuju u horizontalnoj perspektivi, gdje se gornja strana tijela pacijenta prikazuje na gornjoj strani slike, a donja strana na donjoj strani slike. Ova ravnina je često korištena za dijagnostiku mozga.
2. Sagitalna ravnina - ravnina koja dijeli tijelo pacijenta na desnu i lijevu polovicu. Kada se dobivaju MRI snimke u sagitalnoj ravnini, slike se prikazuju u bočnom pogledu, gdje je lijeva strana tijela pacijenta prikazana na desnoj strani slike, a desna strana na lijevoj strani slike. Ova ravnina se često koristi za pregled mozga, kralježnice i hrskavica koljena.
3. Koronalna ravnina - ravnina koja prolazi kroz tijelo pacijenta i dijeli ga na prednju (ventralnu) i stražnju (dorzalnu) polovicu. Ova ravnina se također naziva frontalnom ravninom, budući da je u većini slučajeva prikazana u frontalnom ili prednjem pogledu. Kada se dobivaju MRI snimke u koronalnoj ravnini, slike se prikazuju u vertikalnom pogledu, gdje je prednja strana tijela pacijenta prikazana na stražnjoj strani slike, a stražnja strana na prednjoj strani slike. Često se koristi u ortopediji za pregled zglobova i kostiju, uključujući koljeno, stopalo, gležanj i kuk.

Slika 4.3 prikazuje spomenute ravnine u prostoru [14].



**Slika. 4.3:** Prikaz koronalne, sagitalne i aksijalne ravnine [14]

MRI ima nekoliko prednosti u odnosu na druge dijagnostičke tehnologije kao što su rendgenske snimke i CT (računalna tomografija). Jedna od najvećih prednosti je što MRI ne koristi ionizirajuće zračenje, što znači da je sigurniji za pacijente. Osim toga, MRI može pružiti detaljniji i precizniji prikaz mekih tkiva i organa u usporedbi s drugim dijagnostičkim tehnikama što je posebno bitno za analizu hrskavica i ligamenata.

## 4.2. Opis korištenog skupa podataka

Skup podataka korišten u ovome radu je *Osteoarthritis Initiative ZIB Dataset* [15]. Osteoarthritis Initiative (OAI) ZIB Dataset je veliki skup podataka koji se koristi u istraživanju osteoartritisa.

OAI ZIB Dataset je skup podataka koji su prikupljeni u sklopu inicijative OAI, koju su financirali Nacionalni institut za artritis i koštana oboljenja, Nacionalni institut za zdravstveno istraživanje i privatni donatori. Podaci su prikupljeni tijekom dugotrajnog istraživanja koje je započelo 2004. godine i još uvijek traje. Podaci uključuju kliničke, radiološke i biomehaničke podatke, kao i podatke o kvaliteti života, funkcionalnosti i drugim faktorima koji utječu na razvoj osteoartritisa. Podaci se prikupljaju od sudionika koji su u riziku od razvoja osteoartritisa ili koji već imaju dijagnozu osteoartritisa. Skup podataka je dostupan javnosti za korištenje u znanstvene svrhe, a njegova primjena je vrlo široka. Podaci se mogu koristiti za razumijevanje patofiziologije osteoartritisa, identifikaciju faktora rizika, razvoj novih terapija i procjenu učinkovitosti postojećih terapija. Također se mogu koristiti za razvoj novih

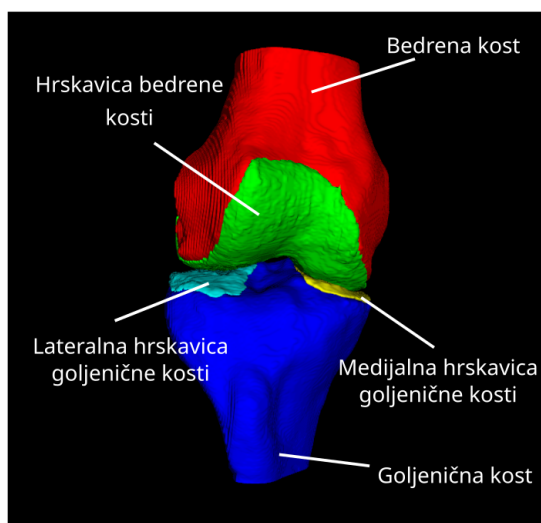
dijagnostičkih alata i za procjenu učinka intervencija kod osteoartritisa [15].

Radiološke slike u ovom skupu podataka sastoje se od slike magnetne rezonance, rendgenskih zraka ili računalne tomografije (CT). Za potrebe ovoga rada korištene su slike dobivene magnetnom rezonancom i njihove oznake koje će biti potrebne za segmentaciju. Postoji 5 označenih struktura koljena koje se mogu vidjeti na slikama 4.4 i 4.5. Slika 4.4 prikazuje primjer koljena iz skupa podataka za sve tri ravnine prikaza i trodimenzionalno, slika 4.5a 3D prikaz prednje strane koljena s oznakama i slika 4.5b 3D prikaz stražnje strane koljena s oznakama. Crvenom bojom označena je bedrena kost (lat. *Femur*), plavom bojom goljenična kost (lat. *Tibia*), zelenom bojom medijalna i lateralna hrskavica bedrene kosti (lat. *Condylus medialis et condylus lateralis Femori*), žutom bojom medijalna hrskavica goljenične kosti (lat. *Condylus medialis Tibiae*) te svijetlo plavom bojom lateralna hrskavica goljenične kosti (lat. *Condylus lateralis Tibiae*).

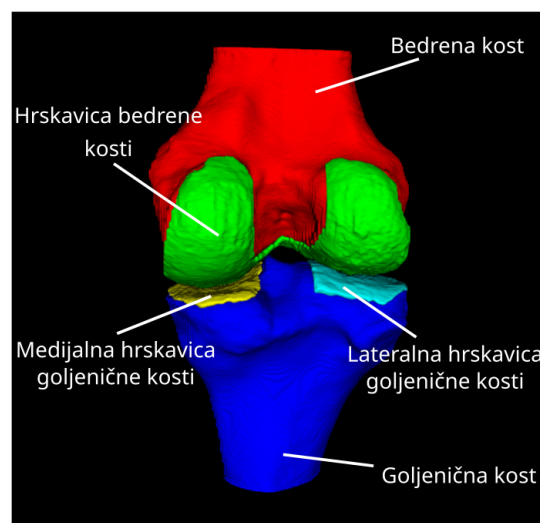


Slika. 4.4: Aksijalni, sagitalni, 3D i koronalni prikaz oznaka koljena





(a) 3D prikaz prednje strane koljena s oznakama



(b) 3D prikaz stražnje strane koljena s oznakama

Slika. 4.5: 3D prikazi koljena s oznakama

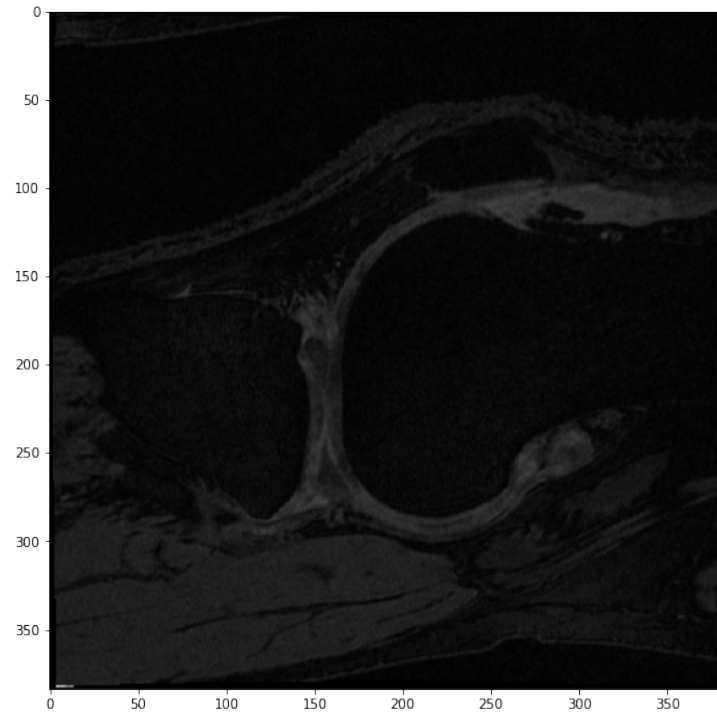
### 4.3. NIfTI format

NIfTI (engl. *Neuroimaging Informatics Technology Initiative*) format je standardni format za pohranu medicinskih slika u području neuroznanosti, uključujući MRI, CT i PET (pozitronska emisijska tomografija) slike.

NIfTI format se sastoji od dva dijela: zaglavlje i blokovi podataka. Zaglavlje sadrži informacije o slici, kao što su dimenzije, broj vremenskih točaka, orijentacija, vrsta slike itd. Blokovi podataka sadrže stvarne slike. Format koristi dvije datoteke za pohranu podataka: ".nii" datoteku koja sadrži zaglavlje i blokove podataka u jednoj datoteci, ili ".hdr" i ".img" datoteke koje sadrže zaglavlje i blokove podataka odvojeno. NIfTI format je fleksibilan i može se koristiti za pohranu slika s različitim dimenzijama i orijentacijama. Podržava i više dimenzija slika, što omogućuje pohranu više slika istovremeno. Također podržava pohranu različitih vrsta podataka, uključujući vrijednosti piksela, vektori piksela, tenzorski podaci itd. NIfTI format je vrlo popularan u neuroznanosti jer omogućuje lako dijeljenje i razmjenu podataka između različitih istraživačkih skupina [16].

Slike iz prethodno opisanog skupa podataka spremljene su u NIfTI formatu te su proučavane koristeći softverski alat ITK-SNAP, slika 4.4 prikazuje snimku zaslona u ITK-SNAP alatu [6].

Za korištenje NIFTI formata u Pythonu potrebno je uvesti biblioteku *nibabel* te korištenjem funkcije *load* iz biblioteke učitati sliku. Slika 4.6 prikazuje primjer sagitalne ravnine volumena koljena dobivenog magnetnom rezonancom. Ako uzimamo po drugoj dimenziji tenzora dobivamo sliku koronalne ravnine, a po trećoj dimenziji sliku aksijalne ravnine.



**Slika. 4.6:** *Sagitalni prikaz dvodimenzionalne slike koljena*

## 5. METODOLOGIJA

U ovom poglavlju opisane su teorijske osnove načina rada difuzijskih modela, konvolucijskih neuronskih, opis U-Net arhitekture neuronske mreže te Gaussovo zamućenje.

### 5.1. Difuzijski modeli

Difuzijski modeli su generativni modeli koji su posljednjih godina stekli značajnu popularnost. Nekoliko važnih radova objavljenih 2020-ih godina pokazalo je što difuzijski modeli mogu postići, poput nadmašivanja GAN-ova (eng. *Generative Adversarial Networks*) u sintetičkom generiranju slika.

U svojoj osnovi, difuzijski modeli djeluju tako da uništavaju podatke za treniranje postupnim dodavanjem Gaussovog šuma (5-1) [22], a zatim uče kako oporaviti podatke obrnutim procesom uklanjanja šuma. Nakon treniranja difuzijski model se može koristiti za generiranje podataka jednostavnim propuštanjem nasumično uzorkovanog šuma kroz naučeni proces uklanjanja šuma. Konkretno, difuzijski model je model latentnih varijabli koji mapira u latentni prostor koristeći fiksni Markovljev lanac. Ovaj lanac postupno dodaje šum podacima kako bi se dobila približna posteriorna distribucija, pri čemu su latentne varijable iste dimenzionalnosti kao i originalni podaci. Na kraju, slika se asimptotski transformira u čisti Gaussov šum. Cilj treniranja difuzijskog modela je naučiti obrnuti proces, odnosno trenirati model da generira podatke kretanjem unazad duž lanca, odnosno iz šuma kreirati sliku visoke kvalitete.

$$\mathcal{N}(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) \quad (5-1)$$

#### 5.1.1 Proces difuzije prema naprijed

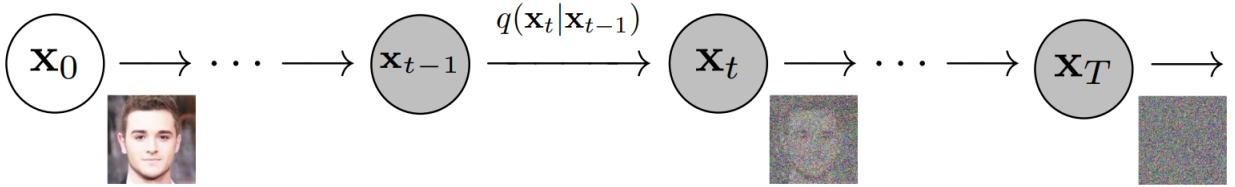
Za zadanu podatkovnu točku iz prave distribucije podataka  $\mathbf{x}_0 \sim q(\mathbf{x})$ , definiramo proces prednje difuzije koji dodaje Gaussov šum kroz  $T$  koraka, proizvodeći niz pošumljenih podataka  $\mathbf{x}_1, \dots, \mathbf{x}_T$ . Proces je upravljan rasporedom varijance  $\{\beta_t \in (0, 1)\}_{t=1}^T$ , ovaj raspored određuje koliki je raspon dodanog šuma u svakom koraku a kasnije će se u potpoglavlju 5.1.4 objasniti različite vrste rasporeda. Jednadžbe 5-2 prikazuju računanje posteriorne vrijednosti za korak  $t$  i generalnu reprezentaciju posteriorne vrijednosti za  $T$  koraka Markovljevog lanca.

$$\begin{aligned}
q(\mathbf{x}_t|\mathbf{x}_{t-1}) &= \mathcal{N}(\mathbf{x}_t; \sqrt{1 - \beta_t}\mathbf{x}_{t-1}, \beta_t\mathbf{I}) \\
q(\mathbf{x}_{1:T}|\mathbf{x}_0) &= \prod_{t=1}^T q(\mathbf{x}_t|\mathbf{x}_{t-1})
\end{aligned} \tag{5-2}$$

Korištenjem reparametrizacijskog trika Gaussove razdiobe,  $\mathbf{x}_t$  se može uzorkovati u bilo kojem proizvoljnom vremenskom koraku  $t$  u zatvorenom obliku kako je prikazano u 5-3. Što znatno komputacijski olakšava proces računanja šuma u trenutku  $t$ .

$$\begin{aligned}
q(\mathbf{x}_t|\mathbf{x}_0) &= \mathcal{N}(\mathbf{x}_t; \sqrt{\bar{\alpha}_t}\mathbf{x}_0, (1 - \bar{\alpha}_t)\mathbf{I}), \\
\alpha_t &= 1 - \beta_t, \\
\bar{\alpha}_t &= \prod_{i=1}^t \alpha_i
\end{aligned} \tag{5-3}$$

Podatak  $\mathbf{x}_0$  postepeno gubi svoje prepoznatljive značajke kako korak  $t$  postaje veći. Nakon nekog vremena kada  $T \rightarrow \infty$ ,  $\mathbf{x}_T$  postaje jednak izotropnoj Gaussovoj razdiobi, odnosno slika puna šuma. Slika 5.7 pokazuje proces difuzije prema naprijed postepenim dodavanjem šuma u Markovljevom lancu.



**Slika. 5.7:** Proces difuzije prema naprijed postepenim dodavanjem šuma u Markovljevom lancu [9]

### 5.1.2 Proces obrnute difuzije

Proces obrnute difuzije omogućava generiranje stvarnih uzoraka koji bi se mogli uzorkovati  $q(\mathbf{x}_{t-1}|\mathbf{x}_t)$  iz Gaussovog šuma  $\mathbf{x}_T \sim \mathcal{N}(0, \mathbf{I})$ .

Ako je dodani šum u koraku ( $\beta_t$ ) dovoljno mali,  $q(\mathbf{x}_{t-1}|\mathbf{x}_t)$  će također biti iz Gaussove distribucije. Budući da je teško procijeniti  $q(\mathbf{x}_{t-1}|\mathbf{x}_t)$  jer zahtijeva korištenje cijelog skupa podataka i prethodnih koraka, potrebno je naučiti model  $p_\theta$  koji aproksimira te uvjetne vjerojatnosti kako bi se mogao obaviti obrnuti proces difuzije 5-4. Odnosno, da bi se znalo iz kojeg koraka je nastao podatak  $\mathbf{x}_t$  trebao bi se imati pristup svim koracima pošumljavanja svih podataka kako bi se mogla obaviti usporedba, što je poprilično neefikasno i zahtjevno.

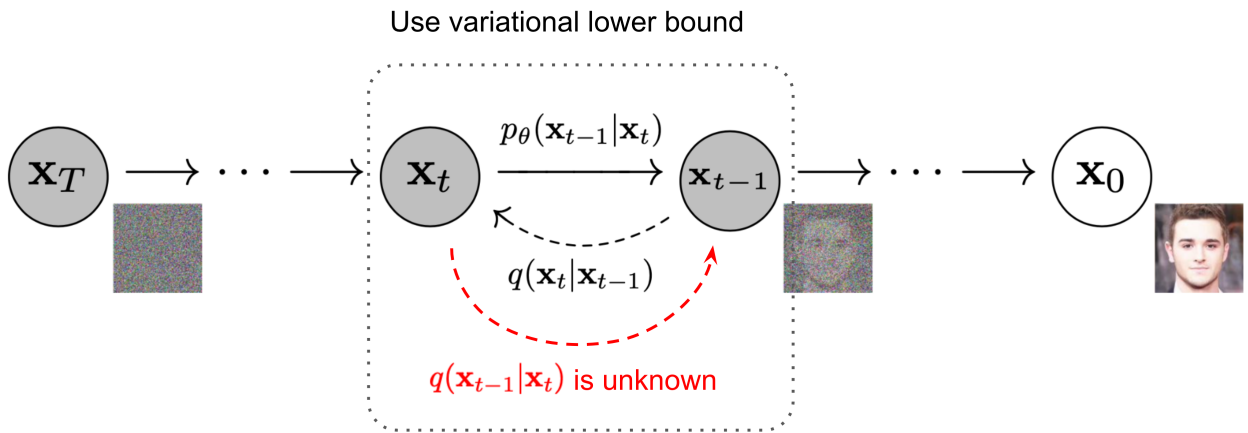
$$p_\theta(\mathbf{x}_{t-1} | \mathbf{x}_t) = \mathcal{N}(\mathbf{x}_{t-1}; \boldsymbol{\mu}_\theta(\mathbf{x}_t, t), \boldsymbol{\Sigma}_\theta(\mathbf{x}_t, t))$$

$$p_\theta(\mathbf{x}_{0:T}) = p(\mathbf{x}_T) \prod_{t=1}^T p_\theta(\mathbf{x}_{t-1} | \mathbf{x}_t) \quad (5-4)$$

Bitno je napomenuti da se obrnuta uvjetna vjerojatnost može pratiti kada je uvjetovana na  $\mathbf{x}_0$  što predstavlja jednadžba 5-5. Ovo je korisno jer tijekom treniranja imamo pristup originalnim slikama  $\mathbf{x}_0$ . Model uči kako ukloniti šum i obnoviti originalnu sliku putem reverznog procesa difuzije. Ova vjerojatnost omogućuje modelu da nauči koliko je šuma dodano u koraku  $t$ , odnosno  $\mathbf{x}_t$  i kako ukloniti taj šum tijekom generiranja novih slika.

$$q(\mathbf{x}_{t-1} | \mathbf{x}_t, \mathbf{x}_0) = \mathcal{N}(\mathbf{x}_{t-1}; \tilde{\boldsymbol{\mu}}(\mathbf{x}_t, \mathbf{x}_0), \tilde{\boldsymbol{\beta}}_t \mathbf{I}) \quad (5-5)$$

Postepenim prolascima kompletno pošumljene slike kroz naučeni model, iz slike se uklanja šum u svakom koraku  $t$  te tako nakon konačnog broja koraka  $T$  model iz kompletno pošumljene slike šumom iz Gaussove distribucije uspijeva dobiti realističnu sliku. Ovaj proces prikazan je na slici 5.8.



**Slika. 5.8:** Proces obrnute difuzije postepenim uklanjanjem šuma modelom u Markovljevom lancu [11]

### 5.1.3 Treniranje difuzijskih modela

Kako bi se neuronsku mrežu naučilo da aproksimira uvjetne vjerojatnosti u obrnutom procesu difuzije,  $p_\theta(\mathbf{x}_{t-1} | \mathbf{x}_t) = \mathcal{N}(\mathbf{x}_{t-1}; \boldsymbol{\mu}_\theta(\mathbf{x}_t, t), \boldsymbol{\Sigma}_\theta(\mathbf{x}_t, t))$ , potrebno je trenirati  $\boldsymbol{\mu}_\theta$  da bude što bliže stvarnom  $\tilde{\boldsymbol{\mu}}_t$ . Gdje su  $\boldsymbol{\mu}_\theta$  i  $\tilde{\boldsymbol{\mu}}_t$  koristeći reparametrizacijski trik definirani na način prikazan u 5-6,  $\epsilon_\theta$  predstavlja procijenjeni šum koji se koristi za uklanjanje šuma iz podataka,  $\epsilon_t$

predstavlja stvarni šum u koraku  $t$ .

$$\begin{aligned}\tilde{\mu}_t &= \frac{1}{\sqrt{\alpha_t}} \left( \mathbf{x}_t - \frac{1 - \alpha_t}{\sqrt{1 - \bar{\alpha}_t}} \epsilon_t \right) \\ \mu_\theta(\mathbf{x}_t, t) &= \frac{1}{\sqrt{\alpha_t}} \left( \mathbf{x}_t - \frac{1 - \alpha_t}{\sqrt{1 - \bar{\alpha}_t}} \epsilon_\theta(\mathbf{x}_t, t) \right)\end{aligned}\tag{5-6}$$

Funkcija gubitka je parametrizirana kako bi se minimizirala razlika od  $\tilde{\mu}$ :

$$L_t = \mathbb{E}_{\mathbf{x}_0, \epsilon} \left[ \frac{1}{2 \|\Sigma_\theta(\mathbf{x}_t, t)\|^2} \|\tilde{\mu}(\mathbf{x}_t, \mathbf{x}_0) - \mu_\theta(\mathbf{x}_t, t)\|^2 \right]$$

U radu [9], autori su empirijski utvrdili da treniranje difuzijskog modela bolje funkcionira ako se u funkciji gubitka zanemare težinski faktori pa je konačna funkcija gubitka:

$$\begin{aligned}L_t^{simple} &= \mathbb{E}_{t \sim [1, T], \mathbf{x}_0, \epsilon_t} [\|\epsilon_t - \epsilon_\theta(\mathbf{x}_t, t)\|^2] \\ L_{simple} &= L_t^{simple} + C\end{aligned}\tag{5-7}$$

gdje je  $C$  konstanta koja ne ovisi o parametrima modela  $\theta$  i može se zanemariti. U ovome radu fokus je stavljen na opis i funkcionalnost difuzijskih modela, a cijeli izvod funkcije gubitka može se pronaći u izvorima [9][10][11].

Nakon što je definirana funkcija gubitka model se može trenirati kako bi generirao nove podatke. Algoritam 1 pokazuje pseudokod za treniranje modela, a algoritam 2 pokazuje kako pomoću istreniranog modela u  $N$  koraka uzorkovati nove podatke. Difuzijski modeli se mogu učiti kako predvidjeti šum koji je dodan na originalnu sliku i kasnije ga otkloniti ili kako predvidjeti originalnu sliku iz pošumljene.

---

#### Algoritam 1 Treniranje

---

- 1: **repeat**
  - 2:    $\mathbf{x}_0 \sim q(\mathbf{x}_0)$
  - 3:    $t \sim \text{Uniform}(\{1, \dots, T\})$
  - 4:    $\epsilon \sim \mathcal{N}(0, \mathbf{I})$
  - 5:   uzmi korak gradijentnog spusta na  $\nabla_\theta \|\epsilon - \epsilon_\theta(\sqrt{\alpha_t} \mathbf{x}_0 + \sqrt{1 - \bar{\alpha}_t} \epsilon, t)\|^2$
  - 6: **until** konvergirano
-

---

**Algoritam 2** Uzorkovanje

---

```
1:  $\mathbf{x}_T \sim \mathcal{N}(0, \mathbf{I})$ 
2: for  $t = T, \dots, 1$  do
3:   if  $t > 1$  then
4:      $z \sim \mathcal{N}(0, \mathbf{I})$ 
5:   else
6:      $z = 0$ 
7:   end if
8:    $\mathbf{x}_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left( \mathbf{x}_t - \frac{1-\alpha_t}{\sqrt{1-\alpha_t}} \epsilon_\theta(\mathbf{x}_t, t) \right) + \sigma_t z$ 
9: end for
10: vrati  $\mathbf{x}_0$ 
```

---

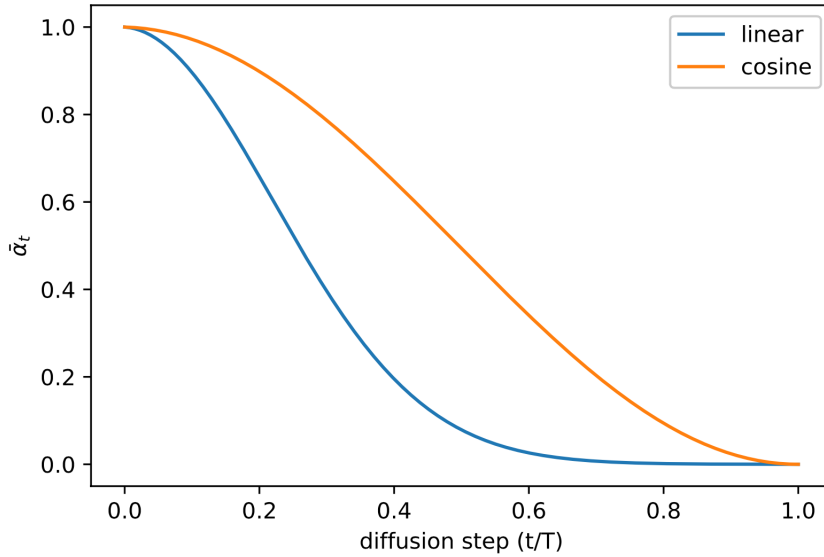
### 5.1.4 Parametrizacija $\beta_t$

Kod difuzijskih modela, raspored varijance  $\beta_t$  igra važnu ulogu u kontroliranju jačine šuma koji se dodaje u svakom koraku difuzije. Postoji nekoliko popularnih rasporeda varijance koji se koriste u praksi.

Varijance su unaprijed postavljene da budu niz linearno rastućih konstanti od  $\beta_1 = 10^{-4}$  do  $\beta_T = 0.02$ . One su relativno male u usporedbi s normaliziranim pikselima slike s vrijednostima između  $[-1, 1]$ . Difuzijski modeli u njihovim eksperimentima pokazali su uzorke visoke kvalitete, ali nisu uspjeli postići konkurentnu logaritamsku vjerojatnost modela kao drugi generativni modeli.

Nichol i Dhariwal [10] predložili su nekoliko tehnika poboljšanja kako bi pomogli difuzijskim modelima da postignu niže vrijednosti logaritamske vjerojatnosti. Jedno od poboljšanja je korištenje rasporeda varijance temeljenog na kosinusu. Prednost je što on osigurava gotovo linearni pad u sredini procesa treniranja i suptilne promjene oko  $t = 0$  i  $t = T$ . Slika 5.9 prikazuje razlike između linearne i kosinusne varijante rasporeda varijance.

Osim linearnog i kosinusnog rasporeda varijance, koriste se još i kvadratni, eksponencijalni i sigmoidni raspored varijance.



**Slika. 5.9:** Usporedba linearne i kosinusne varijante rasporeda  $\beta_t$  tijekom treniranja

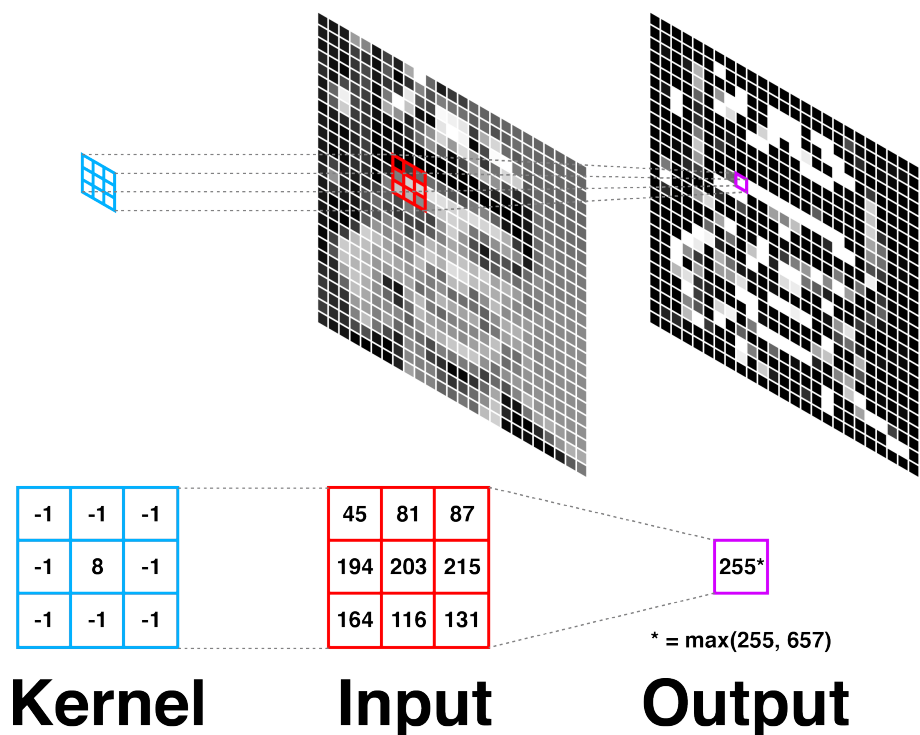
## 5.2. Konvolucijske neuronske mreže

Konvolucijske neuronske mreže (engl. *Convolutional Neural Networks - CNN*) su vrsta dubokih neuronskih mreža koje se koriste za obradu slika, govora i drugih vrsta podataka. One su inspirirane strukturom vizualnog korteksa u ljudskom mozgu, a sastoje se od slojeva konvolucijskih filtara, slojeva sažimanja (engl. *pooling*) i potpuno povezanih slojeva.

Prvi sloj konvolucijske neuronske mreže sastoji se od konvolucijskih filtara koji se primjenjuju na ulazne podatke kako bi se naučile značajke koje su bitne za klasifikaciju ili predviđanje. Konvolucijski filtri su matrice težina koje se mogu trenirati kako bi se prepoznale različite značajke. Ovi filtri se primjenjuju na ulazni podatak s pomoću konvolucije koja se može predstaviti kao množenje vektora težina filtera s podskupom ulaznih podataka, odnosno prelaskom jezgre (engl. *kernel*) konvolucije nad ulaznim podacima. Kada se primjeni na čitavu sliku, konvolucijski sloj će proizvesti skup značajki za daljnju obradu. Kernel prelazi preko cijele slike tako što se svaki piksel kernela i slike na mjestu poklapanja množi, zatim se rezultati množenja zbroje i dobiveni broj predstavlja značajku tog dijela slike. Na slici 5.10 je prikazan primjer konvolucije. Može se zaključiti da obavljanjem konvolucije slika gubi na dimenzionalnosti, ako to nije željeno ponašanje na sliku se može dodati ispunjavanje (engl. *padding*) kako bi se slika popunila na rubovima i zadržala svoje prvotne dimenzije.

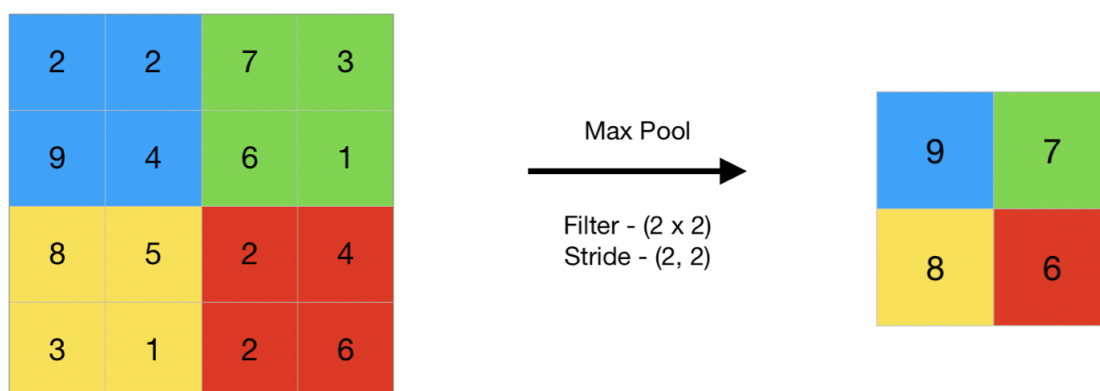
Nakon konvolucijskog sloja, slijedi sloj sažimanja (engl. *max pooling*). Slojevi sažimanja sma-





Slika. 5.10: Operacija konvolucije [19]

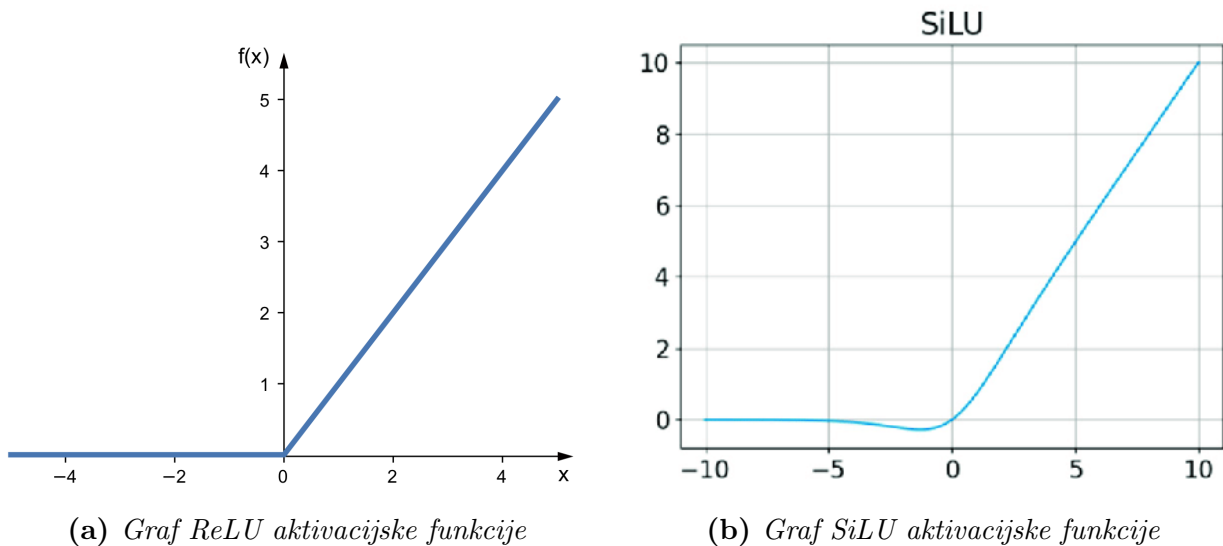
njuju dimenzije ulaznih podataka tako da se gubi na prostornoj razlučivosti, ali se zadržava informacija o značajkama. Najčešće se koristi maksimalno sažimanje (engl. *max-pooling*), gdje se uzima maksimalna vrijednost iz manjeg područja u ulaznom podatku. Na taj način, ovaj sloj će smanjiti broj težina koje se obrađuju u sljedećem sloju konvolucije, a da se puno ne izgubi na značajkama. Primjer maksimalnog sažimanja 4x4 matrice prikazan je na slici 5.11.



Slika. 5.11: Operacija maksimalnog sažimanja [20]

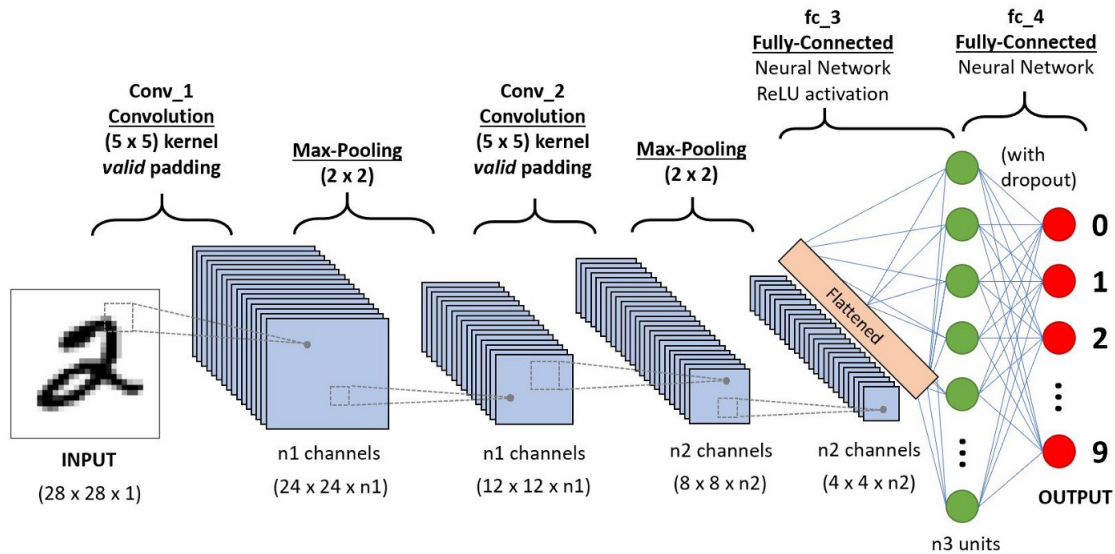
Nakon slojeva konvolucije i sažimanja, dolaze slojevi potpuno povezanih neurona koji vrše klasifikaciju (ako je to zadani problem) ili predviđanje na temelju naučenih značajki. Pot-

puno povezani slojevi imaju sve neurone povezane s neuronom prethodnog sloja. U ovim slojevima, svaki neuron iz prethodnog sloja je povezan sa svakim neuronima u trenutnom sloju, a izlazni vektori su obrađeni kako bi se donijela konačna klasifikacijska odluka. Svaki od neurona u ovom sloju sadrži svoju aktivacijsku funkciju. Aktivacijske funkcije se koriste za uvjetovanje aktivnosti neurona u neuronskoj mreži. One su ključne za učenje jer omogućuju neuronskoj mreži da nauči nelinearne funkcije koje su prisutne u mnogim problemima. Najčešće korištene aktivacijske funkcije su *ReLU* (engl. *rectified linear unit*) i *SiLU* (engl. *sigmoid linear unit*) korištene u skrivenim slojevima mreže, *sigmoid* kao funkcija odluke za binarnu klasifikaciju te *softmax* kao funkcija odluke za višeklasne odluke. Graf ReLU i SiLU funkcije su prikazani na slici 5.12.



**Slika. 5.12:** Usporedba ReLU i SiLU aktivacijskih funkcija

Sve zajedno povezane komponente čine konvolucijsku neuronsku mrežu [17]. Konvolucijske neuronske mreže su bitne za proces učenja difuzijskih modela jer one služe kao alat za izvlačenje značajki iz pošumljene slike kako bi se šum mogao ukloniti. Prikaz konvolucijske neuronske mreže za problem klasifikacije sa svim ranije spomenutim komponentama prikazan je na slici 5.13.



Slika. 5.13: Prikaz konvolucijske neuronske mreže [18]

### 5.3. U-Net arhitektura

U-Net arhitektura je konvolucijska neuronska mreža koja je prvenstveno razvijena za segmentaciju medicinskih slika. Naziv U-Net dolazi od oblika arhitekture, koji ima oblik slova „U”. Glavna ideja U-Net arhitekture je korištenje velikih brojeva filtriranja u konvolucijskim slojevima za učenje velikog broja značajki iz slika, uz istovremeno održavanje prostornih informacija. U-Net arhitektura sastoji se od tri dijela: enkodera, dekodera i mosta koji ih povezuje. Slika 5.14 prikazuje arhitekturu U-Net mreže, s lijeva na desno su prikazani enkoder, most, dekodera.

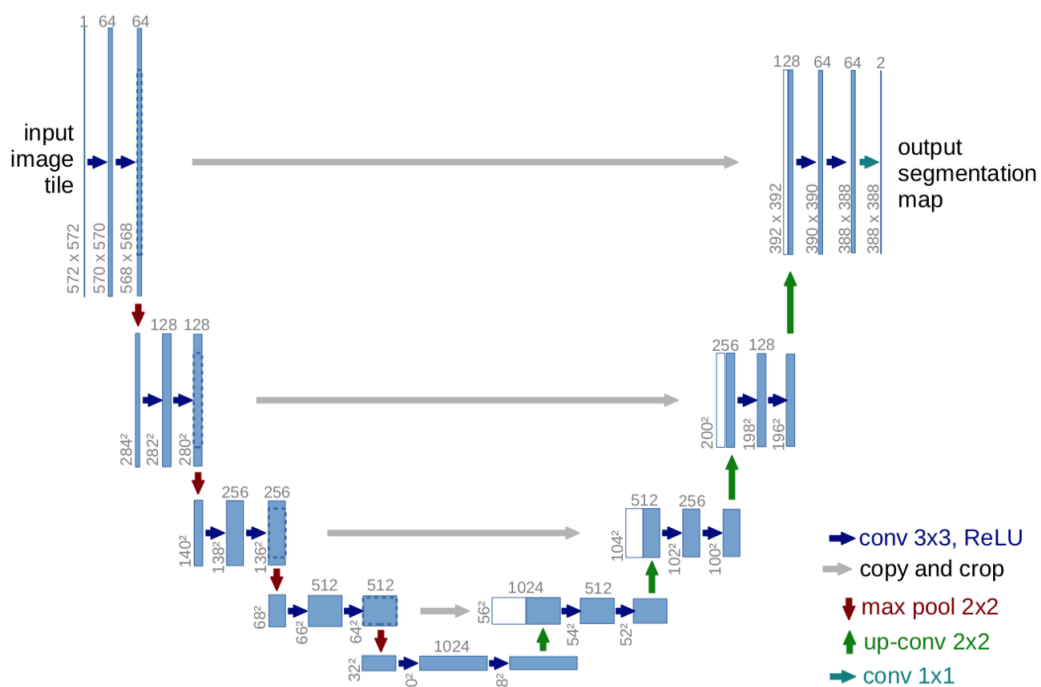
Enkoder je dio mreže koji izvlači značajke iz slike, smanjujući dimenziju slike pomoću slojeva sažimanja (engl. *max-pooling*) nakon svakog sloja konvolucije. Ovaj dio mreže nalikuje arhitekturi koja se koristi za klasifikaciju slika, a služi za ekstrakciju značajki iz slike. Enkoder je sastavljen od ponavljajućeg sekvencijalnog bloka koji se sastoji od dva konvolucijska sloja s aktivacijskom funkcijom *ReLU* i slojem za sažimanje. Tako se u svakoj sekvenci dimenzionalnost slike prepolovi.

Nakon što se smanji prostorna dimenzionalnost u enkoderu, dodaje se sloj koji povezuje enkoder s dekoderom. Ovaj sloj se sastoji od konvolucijskog sloja koji se primjenjuje na izlazu enkodera, a koji se proširuje prema gore (engl. *up-sampling*) kako bi se povećala prostorna dimenzionalnost. Ovaj sloj se zove most (engl. *bridge*) jer povezuje enkoder i dekodera te omogućuje prenošenje značajki iz enkodera u dekoder.

Dekoder je dio mreže koji proizvodi segmentiranu sliku na temelju značajki koje su naučene u enkoderu. Dekoder se sastoji od ponavljajućeg sekvencijalnog bloka koji se sastoji od sloja za proširivanje dimenzija (*up-sampling*) i konvolucijskog sloja s aktivacijskom funkcijom *ReLU*. Sloj za *up-sampling* koristi se umjesto sloja za sažimanje kako bi se povećala prostorna razlučivost ulaznih podataka [3].

U-Net arhitektura se pokazala izuzetno korisnom u učenju difuzijskih modela zbog svoje sposobnosti očuvanja prostornih informacija kroz slojeve enkodera i dekodera. Difuzijski modeli koji generiraju slike dodavanjem i uklanjanjem šuma kroz više koraka mogu iskoristiti U-Netovu strukturu za učinkovit prijenos i rekonstrukciju značajki. Kombinacija U-Net arhitekture s difuzijskim modelima omogućuje stvaranje detaljnih i visokokvalitetnih segmenata, čime se značajno poboljšava preciznost i učinkovitost u zadacima poput segmentacije medicinskih slika.

Kod difuzijskih modela, uz sliku se također enkodira trenutni trenutak  $t$ , odnosno korak unutar difuzijskog procesa. Ova vremenska komponenta obično se dodaje pomoću sinusoidalnog pozicijskog enkodiranja koji se kombiniraju sa slikovnim značajkama na različitim razinama mreže. Time se modelu omogućuje da nauči kako se šum mijenja kroz različite korake procesa difuzije i kako rekonstruirati sliku iz šuma u različitim trenutcima difuzijskog procesa.



Slika. 5.14: Arhitektura U-Net mreže [3]

## 5.4. Gaussovo zamućenje

Gaussovo zamućenje je tehnika zamućenja slike korištenjem Gaussove funkcije. Ova metoda se često koristi u grafičkim softverima za smanjenje šuma i detalja na slici.

U računalnom vidu, Gaussovo zamućenje često se koristi kao korak predobrade ili naknadne obrade za isticanje struktura slike na različitim skalama što je korisno za daljnje analize. Primjena Gaussovog zamućenja na sliku ima efekt smanjenja visokofrekventnih komponenti slike, zbog čega je Gaussovo zamućenje ujedno i niskopropusni filter [21].

Matematički gledano, primjenjivanje Gaussovog zamućenja na sliku je isto što i primjenjivanje konvolucije Gaussove funkcije na sliku. Gaussova funkcija zamućenja za dvodimenzionalnu sliku prikazana je jednadžbom 5-8, a za trodimenzionalni prostor jednadžbom 5-9.

$$G(x, y, z) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (5-8)$$

$$G(x, y, z) = \frac{1}{(2\pi\sigma^2)^{3/2}} \exp\left(-\frac{x^2 + y^2 + z^2}{2\sigma^2}\right) \quad (5-9)$$

Kako bi se kreirao kernel,  $x$ ,  $y$  i  $z$  predstavljaju udaljenost od centra kernela, a  $\sigma$  je raspršenje oko centra kernela. Kreiranje dvodimenzionalnog 3x3 kernela s vrijednosti  $\sigma \approx 0.8$  daje rezultat:

$$\frac{1}{16} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix}$$

U ovome radu ova tehnika je korisna za naknadnu obradu trodimenzionalnog volumena kako bi se dobili finiji prijelazi između svake ravnine što će biti naknadno objašnjeno.

## 6. IZRADA MODELA

U ovome poglavlju opisana je arhitektura i korišteni postupci pri treniranju modela za zadani problem segmentacije 3D slika magnetne rezonance koljena difuzijskim modelima, uključujući predobradu podataka, funkciju gubitka i korištene metrike za evaluaciju.

### 6.1. Podjela skupa podataka i predobrada

U skupu podataka nalazi se 507 slika magnetne rezonance koljena koje će biti ulaz u model i njihove maske koje treba predvidjeti. Podatci su podijeljeni u direktorije train i test koji sadrže po 253 i 254 primjerka. Kako se ovaj omjer nije pokazao najbolji pri treniranju neuronske mreže jer mreža nije dobro generalizirala i jer difuzijski modeli zahtijevaju puno veći broj podataka, podatci su podijeljeni na trening skup od 405 slika i testni skup od 102 slike. Odnosno 80 % podataka predodređeno je za trening, a 20 % za test. Od toga je za vrijeme treninga 40 % testnog skupa, odnosno 40 slika odvojeno za validaciju.

Pošto je problem trodimenzionalne segmentacije računalno zahtjvan, podatke je prije treninga potrebno još nekako podijeliti kako bi stali u RAM memoriju grafičke kartice zajedno s modelom. Slike koje se nalaze u skupu podataka su dimenzija 160 x 384 x 384 i teško ih je, uz kompleksan model koji obavljanja složene matematičke operacije za vrijeme treninga, cijele obrađivati na grafičkoj kartici računala. Ideja je slati jednu po jednu dvodimenzionalnu sliku u model. Time se model svodi na dvodimenzionalan problem i štede se resursi računala koji bi bili zauzimani za procese trodimenzionalne konvolucije i čuvanja dimenzija. Ovakvom podjelom trening skup sadrži 64 800, testni skup 16 320, a validacijski 6400 dvodimenzionalnih slika.

Osim podjele podataka potrebno je izvršiti dodatnu predobradu poput normalizacije piksela maski i ulaznih slika. Ulazne slike su normalizirane koristeći sloj za predobradu *Normalize* iz *torchvision* biblioteke. Ovaj sloj će skalirati ulaze ovisno o srednjoj vrijednosti i standardnoj devijaciji podataka na način prikazan formulom 6-10 [23]. Srednja vrijednost i standardna devijacija izračunate su prije treniranja nad trening skupom ulaznih slika te su dobivene vrijednosti  $\mu = 64.05$  i  $\sigma = 50.98$ .

$$x'_i = \frac{x_i - \mu}{\sigma} \quad (6-10)$$

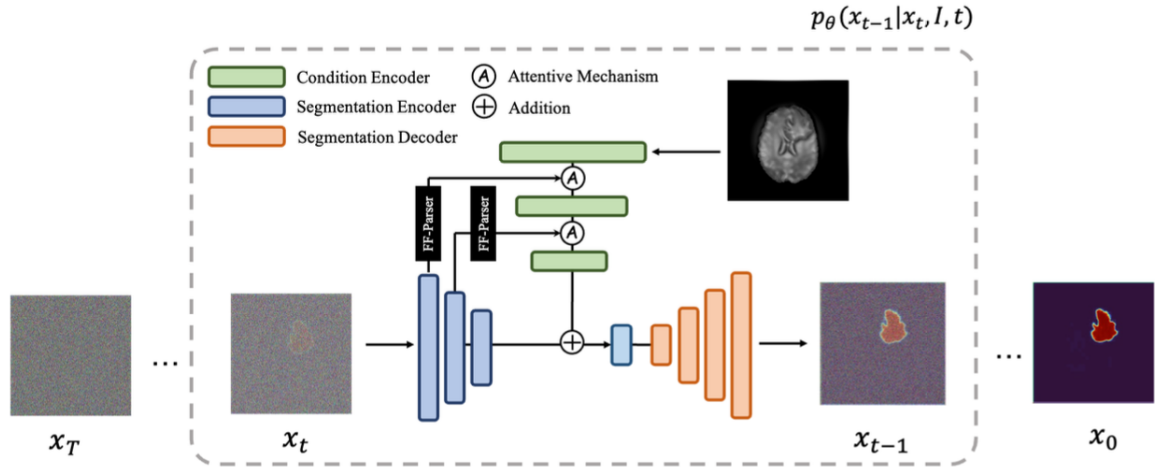
Svaka struktura maske predstavlja izlazni kanal modela. Tako će izlaz iz modela imati 6 kanala, 5 za strukture kosti i hrskavica, i 1 za pozadinu. Vrijednost kanala će imati vrijednost 1 ako piksel predstavlja tu strukturu i 0 ako ne. Svaki piksel može imati upaljen samo jedan kanal, odnosno može biti samo jedna struktura. Maske su normalizirane na vrijednosti u rasponu  $[-1, 1]$  jer je važno osigurati da ulazni podaci imaju sličan raspon vrijednosti kao i vrijednosti šuma za treniranje difuzijskih modela. Time se poboljšava stabilnost i učinkovitost optimizacije, smanjujući potencijalne numeričke probleme te omogućujući modelu da brže i preciznije uči. Normalizacija također olakšava usporedbu različitih maski i rezultata tijekom treniranja difuzijskih modela. Nakon normalizacije ako je kanal maske aktivan imat će vrijednost 1, a ako nije imat će vrijednost -1.

## 6.2. Arhitektura modela

Kako je već spomenuto, za treniranje difuzijskih modela koristi se U-Net arhitektura neuronske mreže. Kako je tema ovog rada segmentacija, arhitektura mreže je nešto drugačija od klasičnog U-Neta jer se slika ne generira samo iz šuma već je slika MRI jedan od uvjeta za generiranje. Također, u slike je potrebno enkodirati i trenutak  $t$  kako je već spomenuto u poglavlju 5.3. Ova mreža stoga ima 2 ulaza, odnosno 2 enkodera, most i dekodera. Ulazi u dekodere su šum s vremenskim trenutkom  $t$  i slika MRI kao uvjet po kojemu se segmentira s vremenskim trenutkom  $t$ , iz njih će se izvlačiti značajke i međusobno kombinirati kako bi se dobila segmentacija. Za spomenuti postupak koristit će se tehnika predložena u radu [7] koja koristi pretvarač brze Foureirove transformacije i svoju mapu značajki. Arhitektura modela iz spomenutog rada prikazana je na slici 6.15 gdje su zbog jasnoće izostavljeni podatci o enkodiranju vremenskog koraka  $t$ . U sljedećim potpoglavljima bit će opisani svi blokovi potrebni za sastavljanje ove arhitekture.

### 6.2.1 Dinamičko uvjetno kodiranje

U većini uvjetovanih difuzijskih modela, uvjetovana prethodna vjerojatnost zavisi od specifičnih informacija koje se koriste kao uvjet. To znači da će model, kada dobije neki uvjet, generirati skup slika koje su usklađene s tim uvjetom. Međutim, segmentacija medicinskih slika je poznata po svojim dvosmislenim objektima. Lezije ili tkiva često je teško razlikovati od pozadine. Modaliteti slika niskog kontrasta, poput MRI ili ultrazvučnih slika, dodatno pogoršavaju situaciju. Ako je dana samo statična slika  $I$  kao uvjet za svaki korak, učenje će



**Slika. 6.15:** *Arhitektura modela za segmentaciju difuzijskim modelima predloženog u Medical Image Segmentation with Diffusion Probabilistic Model [7]*

biti otežano. Kako bi se riješio ovaj problem, autori rada su predložili sljedeći princip rada.

S jedne strane uvjetna slika  $I$  sadrži informacije o cilju segmentacije, ali je teško razlikovati slične strukture, dok s druge strane segmentacijska mapa trenutnog koraka  $x_t$  sadrži poboljšane ciljne regije, ali nije točna i još uvijek sadrži šum kojega treba otkloniti. Trenutna segmentacijska informacija  $x_t$  integrirana je u kodiranje uvjetovane slike  $I$  za međusobno nadopunjavanje. Konkretno, u enkoderu uvjetovane slike  $I$  značajke srednjih slojeva se poboljšavaju sa značajkama srednjih slojeva trenutne segmentacijske informacije  $x_t$ . Svaki korak uvjetovane mape značajki  $m_I^k$  spaja se s mapama značajki  $m_x^k$  iste veličine, gdje je  $k$  indeks sloja enkodera. Spajanje se provodi mehanizmom  $\mathcal{A}$  prikazanim formulom 6-11, gdje  $\otimes$  označava množenje po elementima, a  $LN$  označava normalizaciju sloja. Na dvije mape značajki se prvo primjenjuje normalizacijski sloj, a zatim se množe zajedno da bi se dobila mapa afiniteta. Zatim se mapa afiniteta množi s uvjetovanim značajkama kodiranja slike  $I$  kako bi se poboljšala regija interesa [7].

$$\mathcal{A}(m_I^k, m_x^k) = (LN(m_I^k) \otimes LN(m_x^k)) \otimes m_I^k, \quad (6-11)$$

Operacija se primjenjuje na srednje slojeve enkodera što je vidljivo na slici 6.15. Iako je strategija učinkovita, specifičan problem je što integracija značajki iz slojeva ulaza  $x_t$  inducira dodatni visokofrekventni šum. Da bi se riješio ovaj problem, autori predlažu FF-pretvarač koji će ograničiti visokofrekventne komponente u značajkama.



### 6.2.2 FF-Pretvarač

FF-pretvarač je primjenjen na izlaze srednjih slojeva  $x_t$  enkodera, odnosno na  $m_x^k$ , prije primjene dinamičkog uvjetnog kodiranja kako je prikazano slikom 6.15. Njegova funkcija je ograničiti visokofrekventne komponente povezane sa šumom u  $x_t$  značajkama. Glavna ideja je naučiti parametriziranu mapu težina koja se primjenjuje na značajke Fourierovog prostora. S obzirom na mapu značajki dekodera  $m \in \mathbb{R}^{C \times H \times W}$ , prvo se provodi 2D FFT (engl. *Fast Fourier Transformation*) duž prostornih dimenzija, što je prikazano formulom 6-12, gdje  $\mathcal{F}[\cdot]$  označava 2D FFT.

$$M = \mathcal{F}[m] \in \mathbb{C}^{H \times W \times C}, \quad (6-12)$$

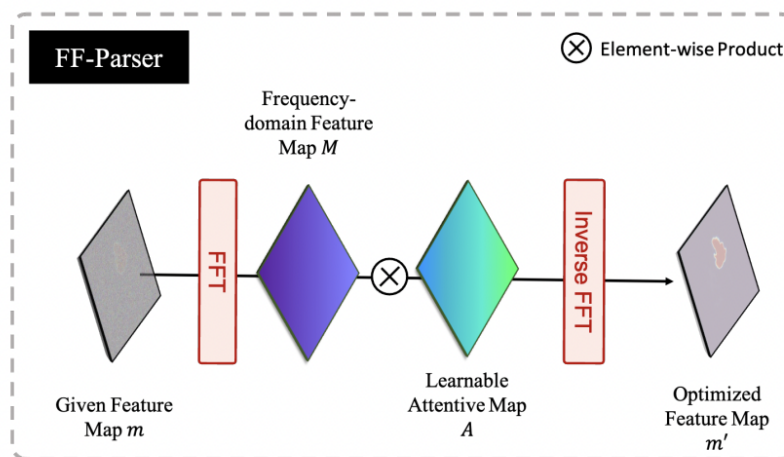
Zatim se modulira spektar od  $m$  množenjem parametrizirane mape težina  $A \in \mathbb{C}^{H \times W \times C}$ , koja će se naučiti treniranjem, s  $M$  prikazano u 6-13, gdje  $\otimes$  označava množenje po elementima.

$$M' = A \otimes M, \quad (6-13)$$

Na kraju,  $M'$  se ponovno vraća u prostornu domenu koristeći inverznu FFT (6-14).

$$m' = \mathcal{F}^{-1}[M']. \quad (6-14)$$

FF-pretvarač se smatra verzijom frekvencijskih filtera koji se može naučiti, a koji se široko primjenjuju u digitalnoj obradi slike. Za razliku od prostorne pažnje, globalno podešava komponente specifičnih frekvencija te na taj način može naučiti ograničiti visokofrekventne komponente [7]. Ilustracija FF-pretvarača prikazana je na slici 6.16.



Slika. 6.16: Ilustracija FF-pretvarača [7]

### 6.2.3 ResNet blok

ResNet blok predstavlja jednu od jedinica U-Neta koji služi za izvlačenje značajki. Svaki ResNet blok sastoji se od dva konvolucijska sloja s aktivacijskom funkcijom *SiLU* (Sigmoid Linear Unit), koja se pokazala učinkovitom u brojnim zadacima dubokog učenja. Osim toga, blokovi sadrže slojeve za normalizaciju značajki kako bi se osigurala stabilnost tijekom treniranja. Tijekom prolaza kroz mrežu, ulazni podaci prolaze kroz prvi konvolucijski sloj gdje se primjenjuje konvolucija, normalizacija i aktivacija, ako su prisutne značajke o vremenskom trenutku  $t$  difuzijskog procesa, one su također nadodane. Nakon toga, podaci prolaze kroz drugi konvolucijski sloj, nakon čega se izlaz kombinira s originalnim ulazom u ResNet blok. Rezultirajući izlaz je zbroj izlaza drugog konvolucijskog sloja i originalnog ulaza, čime se omogućuje bolje treniranje i učinkovitija ekstrakcija značajki. Ova struktura omogućuje stabilniji i brži proces treniranja, posebno u dubokim mrežama koje se koriste za generiranje visokokvalitetnih segmenata u difuzijskim modelima [26].

### 6.2.4 Blokovi pažnje

Blokovi pažnje(engl. *Attention*) su važni elementi u neuronskim mrežama koji omogućuju mreži da se usredotoči na bitnije dijelove ulaznih podataka. Svrha ovog bloka je da omogući mreži da dinamički prilagodi svoju pozornost na različite dijelove ulazne slike. Time se postiže bolje razumijevanje i rekonstrukcija kompleksnih značajki, što je od velike važnosti u zadacima segmentacije medicinskih slika.

Blok pažnje prima kao ulaz tensor značajki slike koji dolazi iz prethodnog sloja mreže. Ovi ulazi se zatim obrađuju kako bi se generirali upiti(engl. *queries*), ključevi(engl. *keys*) i vrijednosti(engl. *values*) koji se koriste za računanje pažnje. Upiti, ključevi i vrijednosti su također treningom naučene značajke. Korištenjem skalarnog produkta upita i ključeva, računa se matrica pažnje koja identificira relevantne dijelove ulazne slike na temelju sličnosti između upita i ključeva. Matrica se zatim koristi za ponderiranje vrijednosti, omogućujući mreži da se fokusira na relevantnije dijelove slike [24].

### 6.2.5 Sinusoidalne pozicije

Sloj za računanje sinusoidalne pozicije(engl. *Sinusoidal Position Embedding*) služi za kodiranje vremenskih informacija  $t$  u difuzijskim modelima. Ovaj sloj transformira skalarnu vrijednost vremena difuzijskog koraka  $t$  u višedimenzionalni vektor pomoću sinusoidalnih i

kosinusoidalnih funkcija. Ulazna vrijednost se skalira i kombinira sa sinusoidalnim i kosinusoidalnim funkcijama različitih frekvencija. Generirana reprezentacija vremenskog koraka  $t$  je učinkovitija od korištenja samog skalara jer svojom vrijednošću neće staviti veću važnost na vremenski korak  $t$  umjesto na ulaznu sliku, ali će i dalje izmijeniti ulazne vektore tako da modelu da do znanja o kojem se vremenskom trenutku  $t$  radi [25].

### 6.2.6 Povezivanje komponenti u arhitekturi mreže

Na početku se ulazne slike normaliziraju kako bi se postigao raspon vrijednosti prikladan za daljnju obradu. Prvi ulaz je pošumljena slika  $x_t$  koja prolazi kroz inicijalni konvolucijski sloj s parametrima za pošumljene slike, dok drugi ulaz, uvjetne slike  $I$ , prolaze kroz zaseban inicijalni konvolucijski sloj. Vrijeme  $t$  se također enkodira s pomoću sloja za računanje sinusoidalnih pozicija.

Tijekom prolaza kroz enkoder, značajke iz šuma  $x_t$  prolaze kroz 2 ResNet bloka, blok pažnje i zatim prolaze operaciju sažimanja (engl. *downsampling*). Niz ovih blokova zajedno predstavlja jedan blok enkodera  $x_t$ , odnosno šuma. Prije operacije sažimanja dobivena vrijednost se sprema jer će biti korištena u enkoderu uvjetne slike  $I$  za kreiranje spomenute uvjetovane mape značajki.

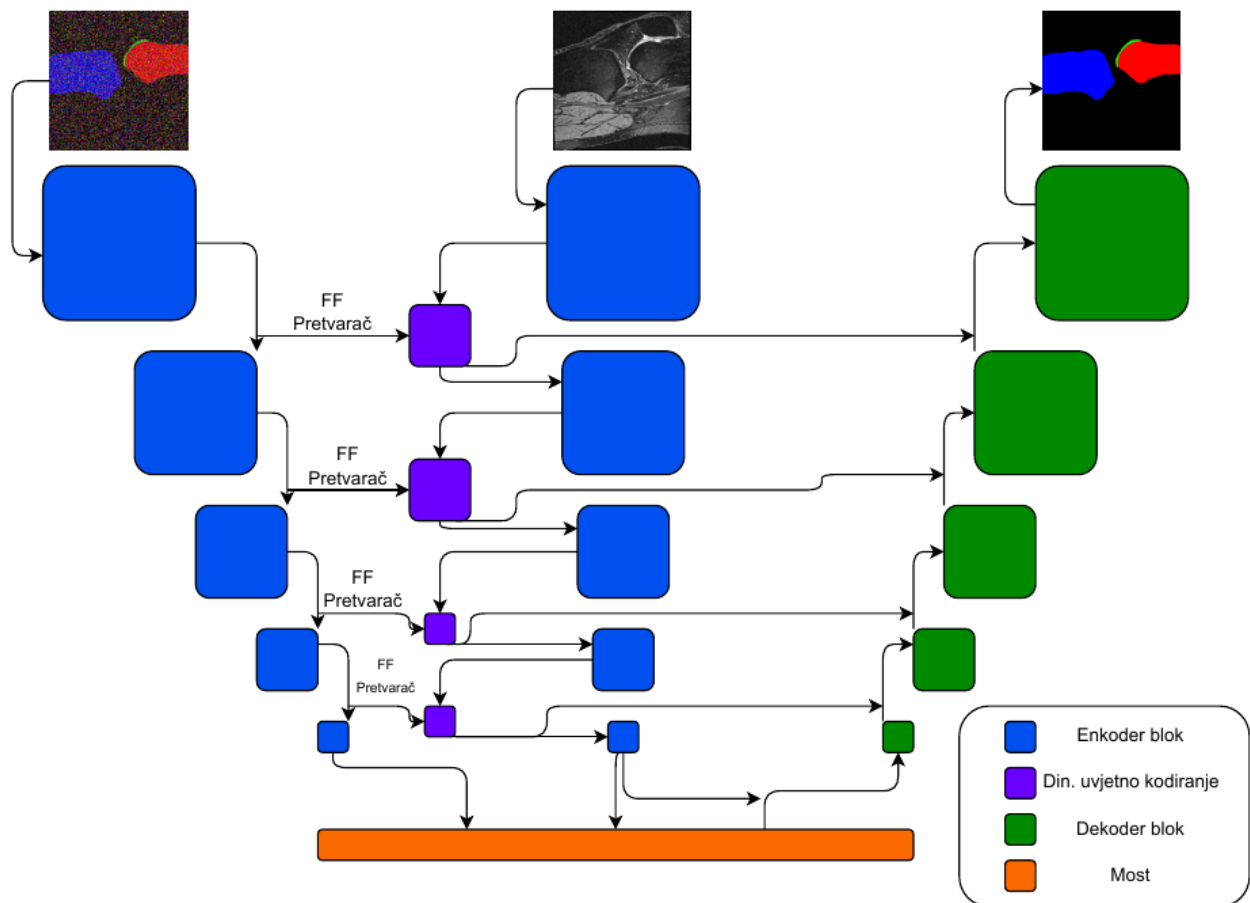
Prolazak kroz enkoder uvjetnih slika  $I$  sličan je enkoderu za  $x_t$ . Jedina razlika je što se prije sloja sažimanja računa uvjetovana mapa značajki spomenuta u 6.2.1 koristeći spremljenu vrijednost istih dimenzija iz segmentacijskog enkodera. I u ovom sloju se prije sažimanja spremaju vrijednosti koje će biti potrebne u dekoderu. Nakon svake operacije sažimanja u oba enkodera broj konvolucijskih filtara se poduplava prije ulaska u sljedeći blok enkodera, a dimenzije slike se smanjuju za duplo. Konkretno vrijednosti idu od 32 filtra do 512 filtara.

Nakon što se prođe kroz zadani broj enkoder blokova, kojih je u ovom slučaju 5, rezultati iz oba enkodera se kombiniraju i prolaze kroz ResNet blok, blok pažnje, ResNet blok te konvolucijski sloj u mostu U-Net mreže.

Tijekom dekodiranja, na značajke se prvo nadodaju spremljene vrijednosti iz enkodera uvjetne slike  $I$  za svaki sloj. Zatim prolaze kroz sličan niz blokova kao u enkoderu za  $x_t$  samo što se umjesto sloja sažimanja ovdje koristi sloj za proširivanje dimenzija. U dekoderu dimenzije se proširuju i konvolucijski filtri smanjuju kako se ide duž dekodera prema gore.

Na kraju, konačne značajke prolaze kroz završni ResNet blok i konvolucijski sloj te se primjenjuje softmax funkcija na 6 kanala kako bi se dobile konačne vrijednosti segmenta-

cijske mape. Konačne segmentacijske mape se zatim normaliziraju kako bi se dobio raspon vrijednosti  $[-1, 1]$  jer je to raspon normalne Gaussove razdiobe koja se koristi u difuzijskim modelima za treniranje. Vizualni prikaz opisane arhitekture nalazi se na slici 6.17.



Slika. 6.17: Arhitektura modela za problem segmentacije koljena difuzijskim modelima

### 6.3. Funkcija gubitka

U poglavlju 5.1.3 opisana je funkcija gubitka za treniranje difuzijskih modela. Za ovaj problem koristit će se slična funkcija s malim modifikacijama.

U spomenutom skupu podataka nalaze se maske koljena koje predstavljaju 6 klasa, pozadinu, bedrenu kost i njezinu hrskavicu i goljeničnu kost i njezine dvije hrskavice. Pošto pozadina čini najveći dio maske, a hrskavice najmanji, koristeći standardnu funkciju gubitka moguće je da se gubitak smanjuje a generirane slike ostanu neprecizne ili uopće nemaju smisla zbog nebalansiranog skupa podataka. Bolja opcija predstavljena je funkcijom 6-16. U ovoj funkciji težine pridodane svakoj klasi koriste se za ponderiranje gubitaka, pri čemu se gubitak za

svaku klasu množi odgovarajućom težinom. Na kraju se uzima prosječna vrijednost ponderirane funkcije gubitka za svaku klasu. Težine klasa su izračunate prije treniranja za svaku klasu formulom 6-15.

$$w_c = \frac{1}{\text{broj piksela klase}_c / \text{broj svih piksela}} \quad (6-15)$$

$$L_{\text{weighted}} = \mathbb{E}_{t \sim [1, T], x_0, \epsilon_t} \left[ \sum_{c=1}^C w_c \cdot \|\epsilon_{t,c} - \epsilon_{\theta}(\mathbf{x}_t, t)_c\|^2 \right] \quad (6-16)$$

## 6.4. Metrike

Predstavljena funkcija nije nužan pokazatelj dobre segmentacije modela, stoga je potrebno koristiti neke bolje metrike za procjenu segmentacije.

*Dice coefficient* ili *F1-Score* je metrika koja se često koristi u zadacima segmentacije slike. Mjeri sličnost između dvaju skupova, s posebnim naglaskom na zajedničke elemente te se koristi za procjenu preklapanja između tih skupova. Pri izračunu koristi vrijednosti preciznosti (engl. *precision*) i odziva (engl. *recall*). Izračunava ga se formulom 6-17 [27].

$$F1 = \frac{2 \cdot \text{točno pozitivni}}{2 \cdot \text{točno pozitivni} + \text{lažno pozitivni} + \text{lažno negativni}} \quad (6-17)$$

Uz *Dice coefficient* koristi se još i *Jaccard index*, odnosno *Intersection over Union* prikazan formulom 6-18. *Jaccard index* mjeri sličnost i raznolikost između dvaju skupova, računajući omjer veličine presjeka skupova prema veličini njihove unije [27].

$$IoU = \frac{\text{točno pozitivni}}{\text{točno pozitivni} + \text{lažno pozitivni} + \text{lažno negativni}} \quad (6-18)$$

## 6.5. Parametri treninga

Jedno od bitnih svojstava kod treniranja difuzijskih modela je to da funkcija gubitka prestaje napredovati, ali generirane slike i dalje postaju sve bolje. Razlog leži u tome što funkcija gubitka mjeri razliku između difuzijskih koraka, odnosno koliko je model dobro uklonio šum sa slike u koraku  $t$ . Tako su slike s većim  $t$  brže naučene i ranije spuste vrijednost funkcije

gubitka, dok vrijednost funkcije gubitka za predviđanje šuma u ranijim koracima  $t$  postaje minimalno i napreduje puno slabije, ali model i dalje generira sve kvalitetnije slike. Iz tog razloga, za napredovanje modela kroz epohe odabrana su 4 volumena MRI slika koljena iz validacijskog skupa nad kojima su se računali *Dice coefficient* i *Jaccard index* kako bi se pratilo napredovanje modela. Međutim, difuzijskim modelima treba neko vrijeme da počnu generirati smislene slike, tako je računanje ove dvije metrike odgođeno za prvih 20 epoha dok model ne nauči generirati smislene segmentacije.

Model je učio rekonstruirati slike u 100 difuzijskih koraka,  $t = 100$  i linearnim rasporedom varijance  $\beta_t$ . Veličina grupe(engl. *batch size*) podataka je 8 i akumulacija gradijenta provedena je svaka 2 trening koraka, odnosno gradijent se računao i propagirao u nazad svakih 16 dvodimenzionalnih slika.

Korištena je Adam optimizacijska funkcija sa stopom učenja  $5 \times 10^{-5}$  i eksponencijalnom stopom pada 0.9 za prvi i 0.99 za drugi moment. Broj epoha postavljen je na 250, ali moguće je ranije zaustavljanje ako kombinacija metrika *Dice coefficient* i *Jaccard index* ne napreduje nakon 6 epoha. Također, stopa učenja se također smanjivala 0.1 puta ako spomenute metrike ne napreduju nakon. Mogućnosti ranijeg zaustavljanja i smanjivanja stope učenja uključene su tek nakon prvih 20 epoha. Cijeli model ima 85 793 126 parametara koji se mogu trenirati.

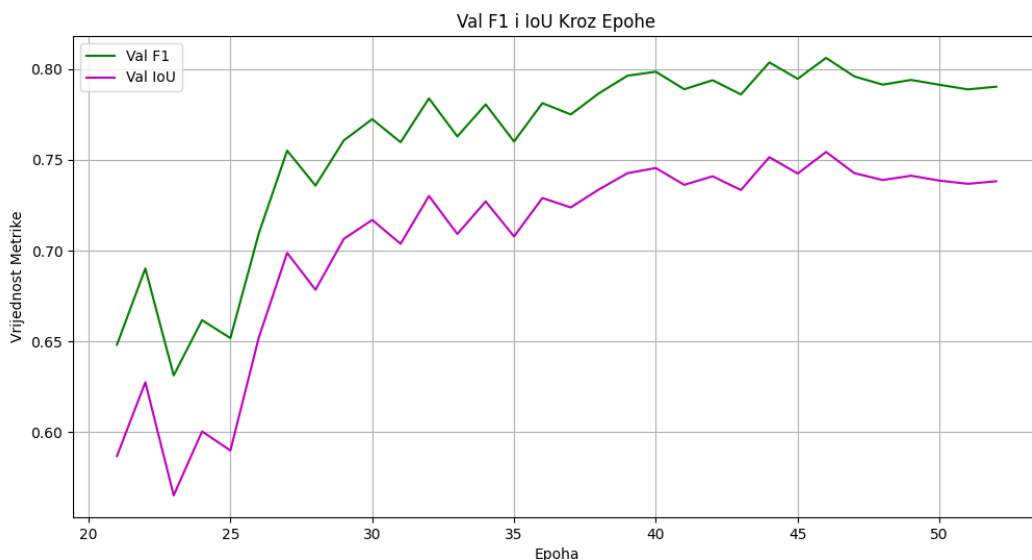
## 7. ANALIZA I NAKNADNA OBRADA REZULTATA

U ovom poglavlju prikazani su i analizirani rezultati treninga modela, provedena evaluacija i naknadna obrada podataka te usporedba s ostalim modelima.

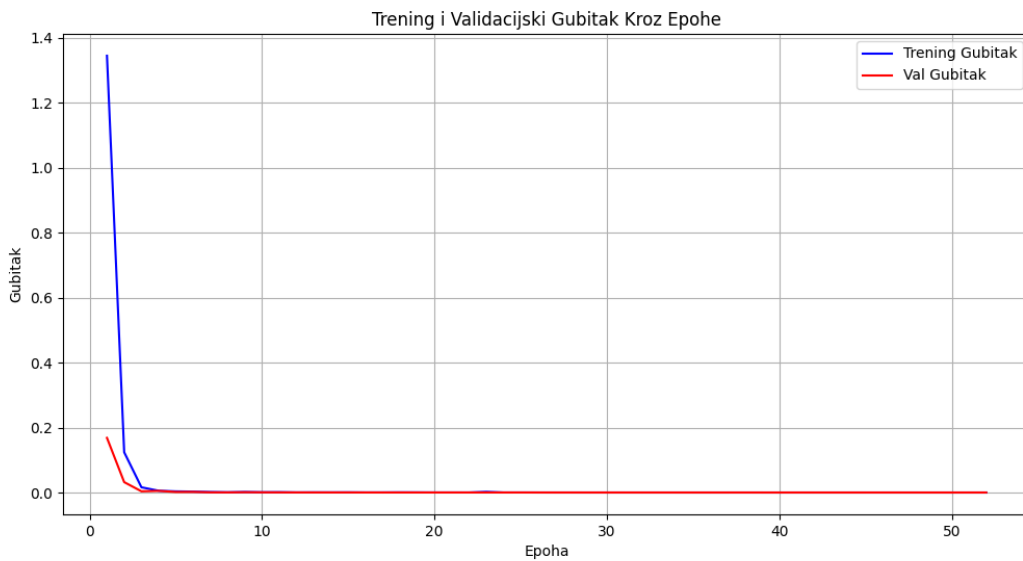
### 7.1. Rezultati treninga

Trening je završen nakon 52 epohe s vrijednosti gubitka  $3.64 \times 10^{-4}$ , a validacijskim skupom  $3.84 \times 10^{-4}$ . *IoU* je imao vrijednost od 73,82 %, a *F1-Score* 79,04 % no treba napomenuti da su se ove dvije metrike računale kao srednja vrijednost nad svim dvodimenzionalnim slikama umjesto nad cijelim volumenom. Stopa učenja spustila se do  $5 \times 10^{-10}$ .

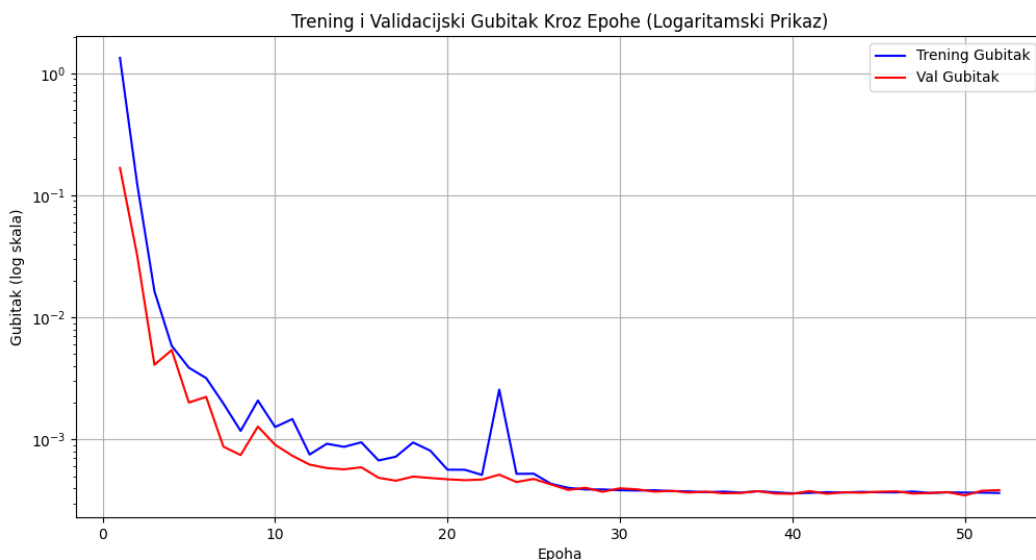
Grafovi pod 7.19 prikazuju kako je gubitak napredovao nad trening i testnim skupom kroz epohe. Napredak je bolje vidljiv na grafu 7.19b jer je gubitak prikazan logaritamskom skalom pa se bolje vidi napredak kroz veći broj epoha. Graf 7.18 prikazuje napredak *F1-Score* i *IoU* kroz epohe. Iz grafova se vidi kako su spomenute metrike i dalje rasle kroz epohe unatoč znatno sporijem padu gubitka nakon 15. epohe. Najbolji rezultat metrika nad validacijskim skupom model je postigao u 46. epohi.



Slika. 7.18: Graf napretka *F1* i *IoU* kroz epohe



(a) Graf napretka trening i validacijskog gubitka na linearnoj skali

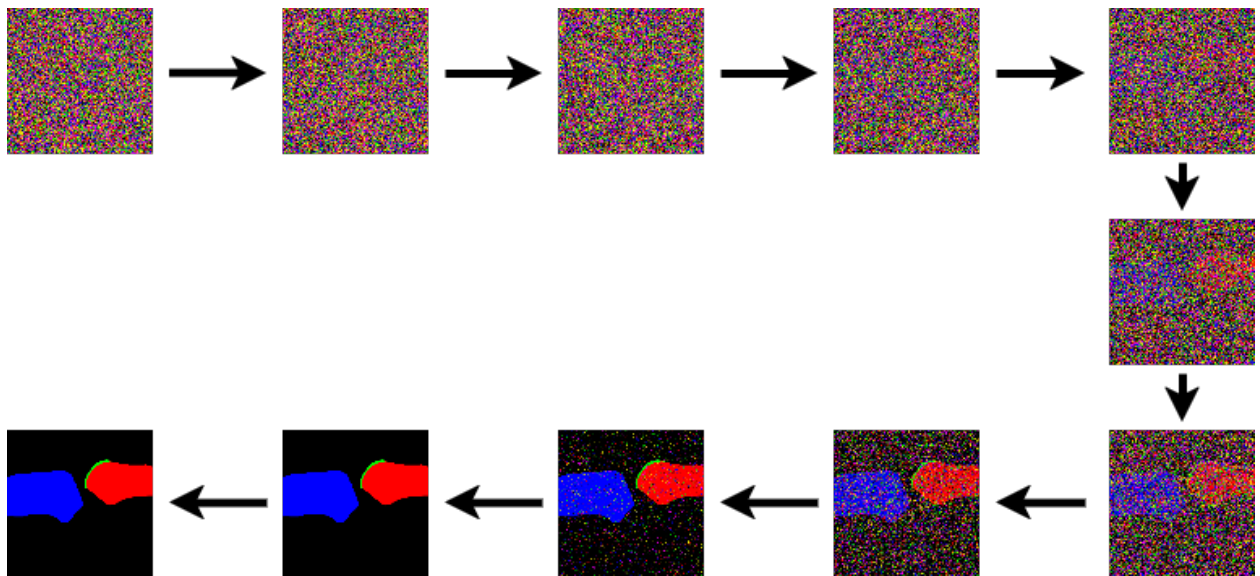


(b) Graf napretka trening i validacijskog gubitka na logaritamskoj skali

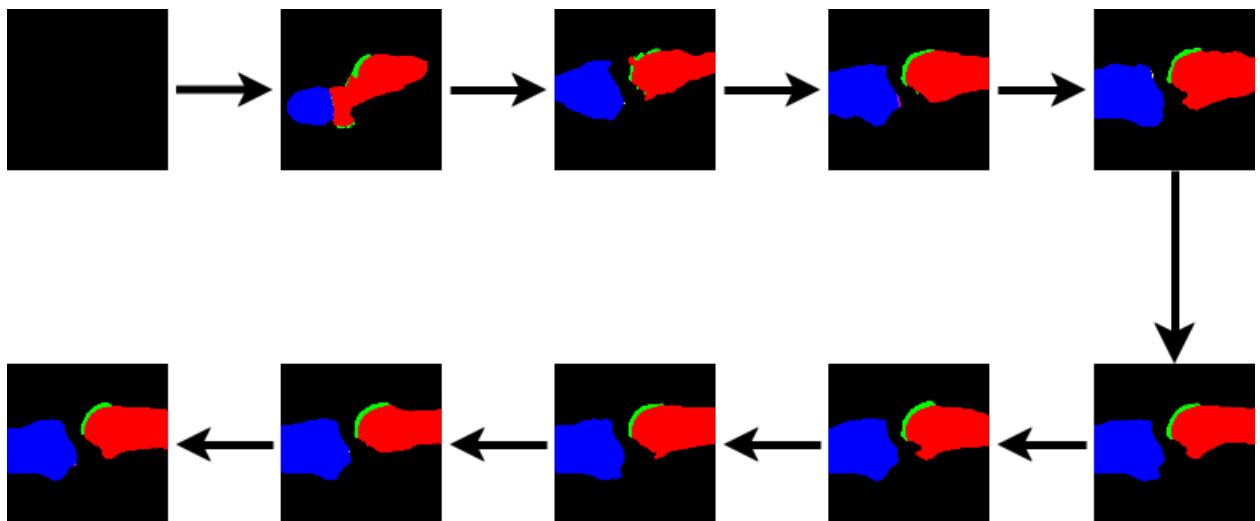
**Slika. 7.19:** Grafovi napretka trening i validacijskog gubitka kroz epohe

Za trening modela korišteno je 100 difuzijskih koraka, što znači da model radi 100 predviđanja iz šuma i u svakom predviđanju ga postepeno uklanja. Proces difuzije, odnosno postepenog uklanjanja šuma, za jednu dvodimenzionalnu sliku prikazan je slikom 7.20. Slika 7.21 prikazuje kako je generiranje slika modela napredovalo kroz epohe. Slika se radi praćenja generirala svakih 5 epoha.





Slika. 7.20: Proces difuzije za dvodimenzionalni presjek koljena



Slika. 7.21: Napredak generiranja slike kroz epohe

## 7.2. Evaluacija i naknadna obrada

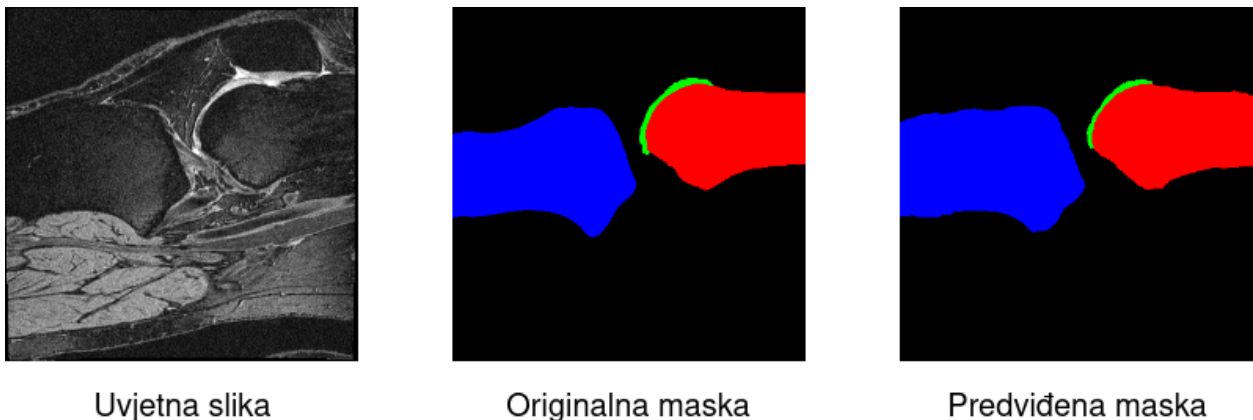
Evaluacija modela provedena je nad spomenutim skupom od 102 testna volumena te su postignuti rezultati  $F1\text{-Score} = 87,34\%$  i  $Intersection\ over\ Union = 79,24\%$  što se smatra dobrim rezultatima. Za usporedbu, čovjek segmentira slike s  $F1\text{-Score}$  preciznošću od  $85\%$  do  $95\%$  i  $IoU$  od  $75\%$  do  $85\%$ .

Usporedba uvjetne slike, originalne maske i maske predviđene modelom za dvodimenzionalni presjek prikazana je na slici 7.22.

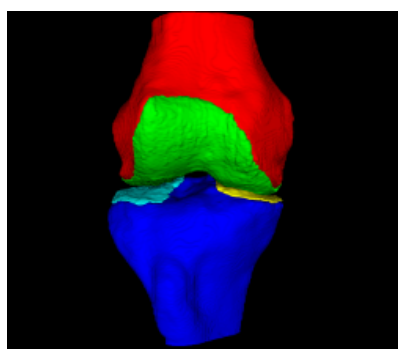
Kako je problem segmentacije koljena trodimenzionalan, a model predviđa segmentaciju dvodimenzionalnih slika, mogući su oštri prijelazi između presjeka dimenzije po kojoj

je segmentirano. Kako bi se riješio ovaj problem, na cijeli volumen je primijenjeno Gaussovo zamućenje. Parametri zamućenja su kernel veličine  $7 \times 7 \times 7$  sa standardnom devijacijom  $\sigma = 2.0$ . Optimalne vrijednosti ovih parametara dobivene su empirijski. Nakon primjene Gaussovog zamućenja na volumene testnog skupa podataka, vrijednost metrika se dodatno poboljšala. Dobiveni su rezultati  $F1\text{-Score} = 89,21 \%$  i  $Intersection\ over\ Union = 82,1 \%$ . Na slici 7.23 prikazani su originalni 3D volumen, 3D volumen predviđen modelom prije zamućenja i 3D volumen predviđen modelom nakon zamućenja. Iz slike se može zaključiti kako je ova naknadna obrada ublažila nagle prijelaze između presjeka i obradila volumen da izgleda realističnije.

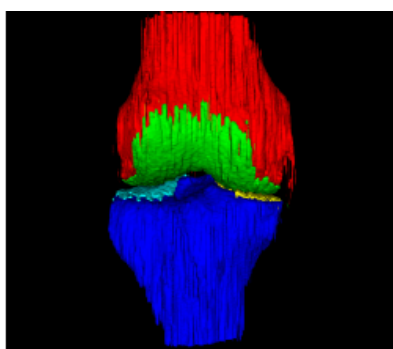
Na slici 7.24 prikazana je matrica konfuzije za nasumično izabran podatak iz testnog skupa. Matrica je izračunata na volumenu koji je naknadno obrađen te je normalizirana iz razloga nebalansiranosti skupa podataka. Iz matrice se vidi kako model dobro segmentira obje kosti i pozadinu, ali ima nešto slabiji performans nad hrskavicama. Hrskavice su najčešće zamijenjene s pozadinom ili susjednom kosti, što je i očekivano s obzirom na to da hrskavice zapravo leže na kostima i teško ih je odvojiti zbog njihovog prikaza na MRI.



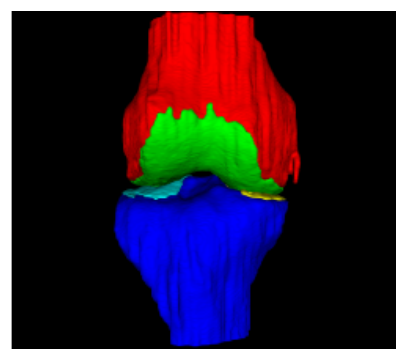
**Slika. 7.22:** *Usporedba uvjetne slike, originalne maske i maske predviđene modelom*



Originalna maska



Predviđena maska

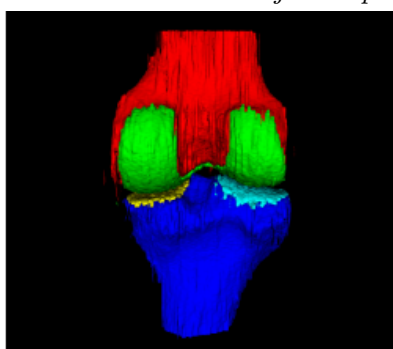


Izgladena predviđena maska

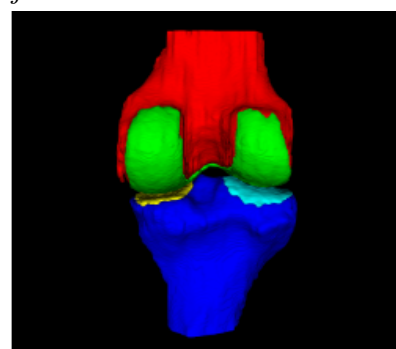
(a) *Usporedba 3D volumena koljena s prednje strane*



Originalna maska

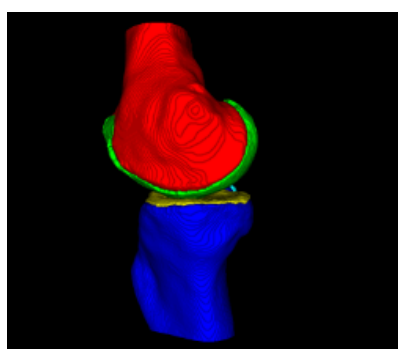


Predviđena maska

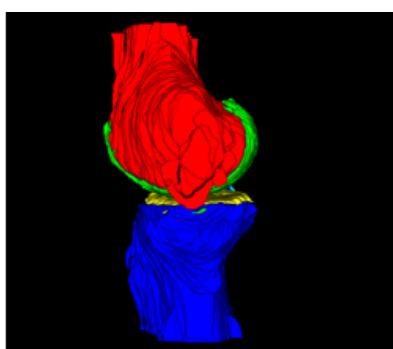


Izgladena predviđena maska

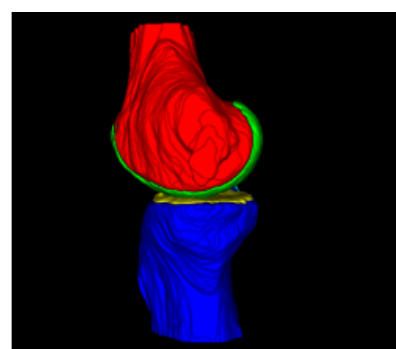
(b) *Usporedba 3D volumena koljena sa stražnje strane*



Originalna maska



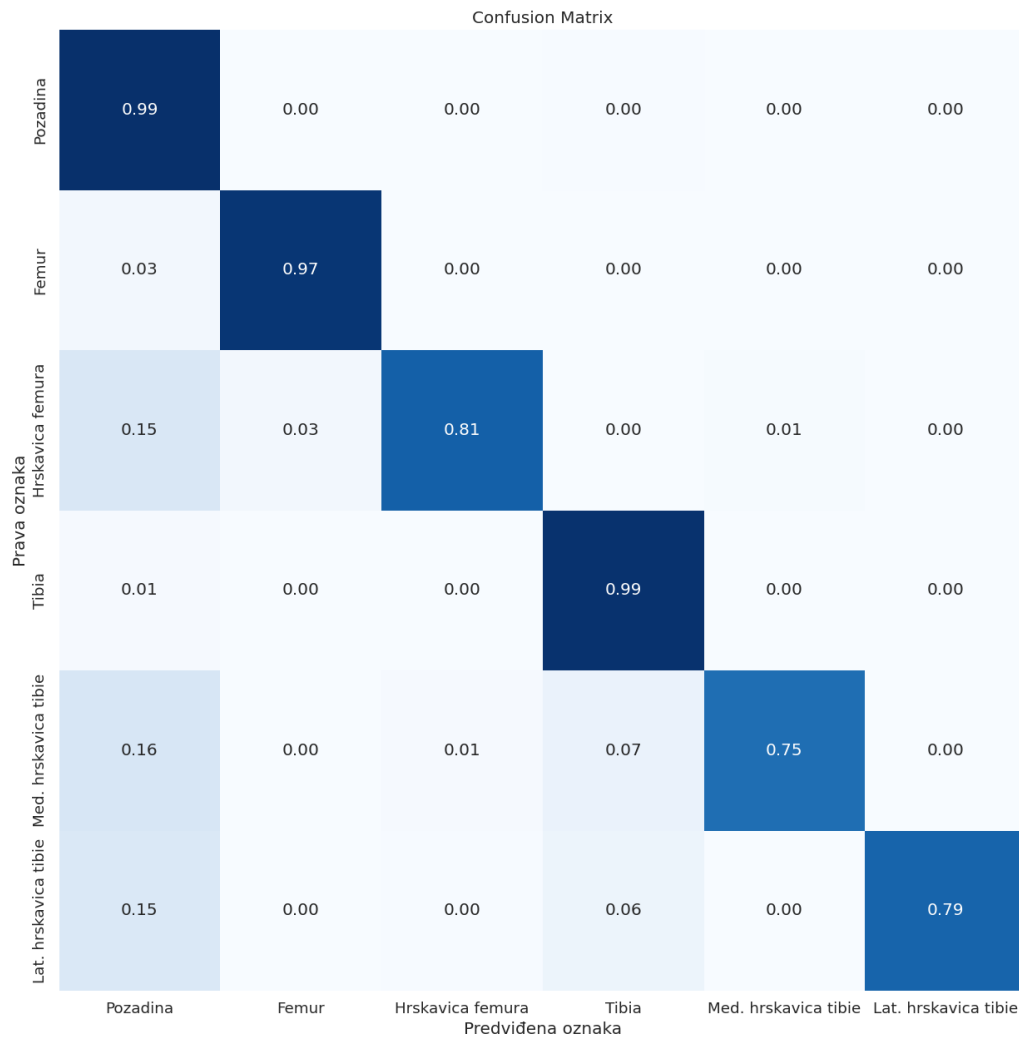
Predviđena maska



Izgladena predviđena maska

(c) *Usporedba 3D volumena koljena s boka*

**Slika. 7.23:** *Usporedba originalnih volumena segmentacije koljena s generiranim segmentacijama za 3 prikaza*



**Slika. 7.24:** *Matrica konfuzije za jedan primjerak segmentacije iz testnog skupa podataka*

### 7.3. Usporedba s ostalim modelima

Za problem segmentacije medicinskih slika postoji mnogo rješenja. Neka od njih su spomenuta u uvodu ovoga rada, a usporedba metrika modela ovoga rada s nekim od spomenutih nalazi se u tablici 7.1. Model segmentacije difuzijskim modelima preciznošću puno ne zaostaje za ostalim modelima. Također, performanse modela mogu se poboljšati dodavanjem novih slojeva ili dodatnih ulaznih podataka, poput segmentacije prethodnog presjeka, kako bi model stekao bolju svijest o okolini i time povećao svoju preciznost. Za usporedbu, klasični

U-Net model ima oko 240 000 000 parametara, a difuzijski 85 793 126 što je znatno manje. No za segmentaciju ovim difuzijskim modelom je potrebno model pozvati 100 puta kako bi uklonio šum, a ostale samo jednom, što znatno usporava sam proces segmentacije za razliku od običnog U-Neta.

Metoda	F1 Score	IoU Score
Difuzijski model	0,8734	0,7924
Difuzijski model sa izgladivanjem	0,8921	0,8210
nnU-Net	0,8947	0,8196
U-Net dizajniran za specifičan problem	0,9219	0,8643

**Tablica. 7.1:** *Usporedba segmentacije koljena difuzijskim modelima s ostalim metodama*

## 8. ZAKLJUČAK

Uspješna dijagnostika osteoartritisa zahtjeva preciznu segmentaciju zahvaćenih područja na slikama magnetne rezonance kako bi se mogao pratiti razvitak bolesti i uspješnost terapije. Iako su slike dobivene magnetnom rezonancom puno "čišće" od onih dobivenih računalnom tomografijom ili rendgenom, one i dalje sadrže dozu šuma i nije ih uvijek jednostavno čitati. Automatska segmentacija nudi puno brže i efikasnije segmentiranje dijelova koljena i može ukazati liječnicima na neke propuste ako nisu dovoljno dobro analizirali sliku.

U ovome radu opisan je način dobivanja slika magnetnom rezonancom kao i tri pogleda, aksijalni, sagitalni i koronalni koji nastaju pri stvaranju iste slike. Opisan je NIfTI format u koji se slike dobivene magnetnom rezonancom spremaju u računalu. Također, opisana je klinička pozadina i važnost same segmentacije koljena za osteoartritis.

Kako su difuzijski modeli u zadnje vrijeme postali aktualna metoda za generiranje novih slika, u ovome radu predložena je i opisana metoda koja segmentira strukture koljena iz 3D volumena magnetne rezonance koristeći difuzijske modele. Također, radi lakšeg razumijevanja i praćenja rada, opisani su načini rada difuzijskih modela, konvolucijskih neuronskih mreža, U-Net arhitekture i Gaussovog zamućenja koji su služili kao nužni alati za izradu ove metode. Detaljno je opisana metoda koja koristi sve ove alate kao i parametri treninga ovog modela.

Nakon treniranja modela, provedena je dodatna obrada generiranih slika Gausovim zamućenjem te su analizirani dobiveni rezultati. Daljnjom obradom rezultata, primjenom Gaussovog zamućenja na generirane volumene, postignuto je dodatno poboljšanje preciznosti segmentacije, što ukazuje na važnost dodatnog procesiranja u zadacima segmentacije. Rezultati  $F1-Score = 89,21 \%$  i  $Intersection\ over\ Union = 82,1 \%$  ukazuju na dobru preciznost segmentacije volumena za zadani problem. Uz to, performanse modela mogu se dodatno poboljšati uključivanjem dodatnih ulaznih informacija, poput segmentacije prethodnih presjeka, što bi omogućilo modelu bolje razumijevanje okolnih struktura i povećalo točnost segmentacije. Ovaj pristup pokazuje potencijal u razvoju naprednih metoda za segmentaciju medicinskih slika, posebno u kontekstu složenih anatomskih struktura kao što je koljeno.

Velika mana kreiranog modela je brzina njegovog generiranja. Iako je model manji od klasičnog U-Neta, potrebno mu je puno više vremena za segmentaciju zato što proces di-

fuzije iziskuje više prolazaka kroz model kako bi uklanjao šum. Stoga, iako je model konkurentan postojećim rješenjima i ima potencijala za dodatna poboljšanja, možda nije optimalno rješenje za zadani problem s obzirom na to da postojeće metode pružaju visoku razinu preciznosti uz znatno kraće vrijeme segmentacije. Također, treba uzeti u obzir i da povećanjem kompleksnosti difuzijskog modela, vrijeme segmentacije raste puno više u odnosu na povećanje kompleksnosti ostalih modela zbog potrebnog broja koraka za generiranje slike.

Kako se tehnologija nastavlja razvijati, MRI i dalje dobiva na značaju kao alat za praćenje i dijagnostiku bolesti u medicini kojega se često koristi u integraciji s neuronskim mrežama. Jedna od glavnih prednosti korištenja dubokih neuronskih mreža u dijagnostici jest njihova sposobnost obrade velike količine podataka u kraćem vremenu, čime se ubrzava proces dijagnostike i smanjuju se moguće greške ljudskog faktora. Također, neuronske mreže su sposobne prepoznati obrasce koji su teško vidljivi ljudskom oku, što može dovesti do ranije dijagnostike bolesti.

## LITERATURA

- [1] “Osteoarthritis - Symptoms and causes,” Mayo Clinic, 16. lipnja 2021. [Dostupno]: <https://www.mayoclinic.org/diseases-conditions/osteoarthritis/symptoms-causes/syc-20351925>. Pristupljeno: 15. svibnja 2024.
- [2] U. Blašković, “Osteoarthritis (artroza) koljena - I što da radite s njim,” Rekreativa Medical. [Dostupno]: <https://www.rekreativa-medical.com/osteoarthritis-koljena.htm>. Pristupljeno: 15. svibnja 2024.
- [3] N. Tomar, “What is UNET? - Analytics Vidhya - Medium,” Analytics Vidhya, 25. siječnja 2021. [Dostupno]: <https://medium.com/analytics-vidhya/what-is-unet-157314c87634>. Pristupljeno: 15. svibnja 2024.
- [4] F. Isensee, P. F. Jaeger, S. A. Kohl, J. Petersen, and K. H. Maier-Hein, “nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation,” ”Nature Methods”, vol. 18, no. 2, pp. 203-211, 2021.
- [5] K. Kamnitsas, C. Ledig, V. F. J. Newcombe, J. P. Simpson, A. D. Kane, D. K. Menon, D. Rueckert, and B. Glocker, “Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation,” ”Medical Image Analysis”, vol. 36, pp. 61-78, 2017, doi: 10.1016/j.media.2016.10.004.
- [6] P. A. Yushkevich, J. Piven, C. Hazlett, H. G. Smith, R. S. Ho, J. C. Gee, and G. Gerig, “User-Guided 3D Active Contour Segmentation of Anatomical Structures: Significantly Improved Efficiency and Reliability,” ”Neuroimage”, vol. 31, no. 3, pp. 1116-1128, 2006. [Dostupno]: [www.itksnap.org](http://www.itksnap.org). Pristupljeno: 15. svibnja 2024.
- [7] J. Wu, R. Fu, H. Fang, Y. Zhang, Y. Yang, H. Xiong, H. Liu, and Y. Xu, “MedSegDiff: Medical Image Segmentation with Diffusion Probabilistic Model”, in ”Proceedings of the Medical Imaging with Deep Learning”, 2023.
- [8] J. Wu, W. Ji, H. Fu, M. Xu, Y. Jin, and Y. Xu, “MedSegDiff-V2: Diffusion-Based Medical Image Segmentation with Transformer,” in ”Proceedings of the AAAI Conference on Artificial Intelligence”, vol. 38, no. 6, pp. 6030–6038, ožujka 2024, doi: 10.1609/aaai.v38i6.28418.



- [9] J. Ho, A. Jain, and P. Abbeel, "Denoising Diffusion Probabilistic Models," "Advances in Neural Information Processing Systems", vol. 33, pp. 6840-6851, 2020.
- [10] A. Nichol and P. Dhariwal, "Diffusion Models Beat GANs on Image Synthesis," "Advances in Neural Information Processing Systems", vol. 34, pp. 8780-8794, 2021.
- [11] L. Weng, "What are Diffusion Models?," Lil'Log, 11. srpnja 2021. [Dostupno]: <https://lilianweng.github.io/posts/2021-07-11-diffusion-models/>. Pristupljeno: 25. svibnja 2024.
- [12] Contributors to Wikimedia projects, "Magnetic resonance imaging," Wikipedia, 4. ožujka 2023. [Dostupno]: [https://en.wikipedia.org/wiki/Magnetic\\_resonance\\_imaging](https://en.wikipedia.org/wiki/Magnetic_resonance_imaging). Pristupljeno: 15. svibnja 2024.
- [13] E. Arsène, D. Barrière, H. Adriaensen, D. Palmer, T. Melzer, N. Mitchell, and M. Keller, "The development of brain magnetic resonance approaches in large animal models for preclinical research," "Animal", vol. 9, pp. 44-51, 2019, doi: 10.1093/af/vfz024.
- [14] "Axial, Coronal, & Sagittal Planes – IPF Radiology Rounds." [Dostupno]: <https://www.ipfradiologyrounds.com/hrct-primer/image-reconstruction/>. Pristupljeno: 15. svibnja 2024.
- [15] "Osteoarthritis Initiative (OAI)," National Institute on Aging. [Dostupno]: <https://www.nia.nih.gov/research/resource/osteoarthritis-initiative-oai>. Pristupljeno: 15. svibnja 2024.
- [16] A. M. Winkler, "The NIFTI file format," Brainder., 23. rujna 2012. [Dostupno]: <https://brainder.org/2012/09/23/the-nifti-file-format/>. Pristupljeno: 15. svibnja 2024.
- [17] Prabhu, "Understanding of Convolutional Neural Network (CNN) — Deep Learning," Medium, 21. studenoga 2019. [Dostupno]: <https://medium.com/@RaghavPrabhu/understanding-of-convolutional-neural-network-cnn-deep-learning-99760835f148>. Pristupljeno: 15. svibnja 2024.
- [18] Sanghviraajit, "On Vectorization of Convolution Layer in Convolution Neural Networks (CNNs)," Analytics Vidhya, 10. veljače 2021. [Dostupno]: <https://medium.com/analytics-vidhya/>

[on-vectorization-of-convolution-layer-in-convolution-neural-networks-cnns-1775cb4](#)

Pristupljeno: 15. svibnja 2024.

- [19] G. Gundersen, “From Convolution to Neural Network.” [Dostupno]: <https://gregorygundersen.com/blog/2017/02/24/cnns/>. Pristupljeno: 15. svibnja 2024.
- [20] GeeksforGeeks, “CNN,” GeeksforGeeks, 5. kolovoza 2019. [Dostupno]: <https://www.geeksforgeeks.org/cnn-introduction-to-pooling-layer/>. Pristupljeno: 15. svibnja 2024.
- [21] Contributors to Wikimedia projects, “Gaussian blur,” Wikipedia, 22. lipnja 2024. [Dostupno]: [https://en.wikipedia.org/wiki/Gaussian\\_blur](https://en.wikipedia.org/wiki/Gaussian_blur). Pristupljeno: 24. kolovoza 2024.
- [22] Contributors to Wikimedia projects, “Gaussian function,” Wikipedia, 16. lipnja 2024. [Dostupno]: [https://en.wikipedia.org/wiki/Gaussian\\_function](https://en.wikipedia.org/wiki/Gaussian_function). Pristupljeno: 24. kolovoza 2024.
- [23] “Normalize — Torchvision main documentation,” PyTorch. [Dostupno]: <https://pytorch.org/vision/main/generated/torchvision.transforms.Normalize.html>. Pristupljeno: 30. kolovoza 2024.
- [24] S. Cristina, “The Attention Mechanism from Scratch,” MachineLearningMastery.com, 26. kolovoza 2022. [Dostupno]: <https://machinelearningmastery.com/the-attention-mechanism-from-scratch/>. Pristupljeno: 25. svibnja 2024.
- [25] Rand, “Creating Sinusoidal Positional Embedding from Scratch in PyTorch,” AI Mind, 29. lipnja 2023. [Dostupno]: <https://pub.aimind.so/creating-sinusoidal-positional-embedding-from-scratch-in-pytorch-98c49e153d6>. Pristupljeno: 25. svibnja 2024.
- [26] A. Gokulaprasad, “Implementing ResNet from Scratch - Anirudh Gokulaprasad,” Medium, 26. lipnja 2023. [Dostupno]: <https://medium.com/@g.anirudh15/implementing-resnet-from-scratch-a1bb437b5f05>. Pristupljeno: 25. svibnja 2024.
- [27] S. Tas, “How To Evaluate Image Segmentation Models?,” Towards Data Science, 17. listopada 2020. [Dostupno]: <https://towardsdatascience.com/>

[how-accurate-is-image-segmentation-dd448f896388](#). Pristupljeno: 25. svibnja 2024.

## SAŽETAK

Segmentacija koljena iz slika magnetne rezonance važna je za praćenje progresije bolesti poput osteoartritisa i planiranje odgovarajuće terapije. U ovom radu predstavljen je model za automatsku segmentaciju koljena korištenjem difuzijskih modela, metode koja je u posljednje vrijeme postala popularna u području generiranja slika. Model koristi U-Net arhitekturu s modifikacijama prilagođenim za difuzijske procese, uključujući dinamičko uvjetno kodiranje i FF-pretvarač koji su predloženi u radu *MedSegDiff: Medical Image Segmentation with Diffusion Probabilistic Model*. Trening modela proveden je na skupu podataka *Osteoarthritis Initiative* (OAI), koji sadrži 3D MRI slike koljena s pripadajućim oznakama kostiju i hrskavica. Model je evaluiran korištenjem metrika *F1-Score* i *Intersection over Union* (IoU), pri čemu su postignuti rezultati od 89,21 % i 82,1 % nakon naknadne primjene Gaussovog zamućenja. Ovi rezultati su usporedivi s postojećim metodama segmentacije koljena. Iako model pokazuje visoku razinu preciznosti i otvorenost za moguću nadogradnju, njegova primjena u praksi može biti ograničena zbog dugog vremena segmentacije.

**Ključne riječi:** MRI, Segmentacija, Difuzijski modeli, Osteoartritis, Medicinske slike, U-Net

## ABSTRACT

Knee segmentation from magnetic resonance imaging is crucial for monitoring the progression of diseases like osteoarthritis and planning appropriate therapy. This paper presents a model for automatic knee segmentation using diffusion models, a method that has recently gained popularity in the field of image generation. The model employs a U-Net architecture with modifications tailored for diffusion processes, including dynamic conditional encoding and the FF-Parser proposed in the paper *MedSegDiff: Medical Image Segmentation with Diffusion Probabilistic Model*. The model was trained on the *Osteoarthritis Initiative* (OAI) dataset, which contains 3D MRI knee images with corresponding annotations of bones and cartilage. The model was evaluated using *F1-Score* and *Intersection over Union* (IoU) metrics, achieving results of 89.21% and 82.1% after applying post-processing with Gaussian blurring. These results are comparable with existing knee segmentation methods. While the model demonstrates a high level of accuracy and potential for further improvement, its practical application may be limited due to the long segmentation time.

**Keywords:** MRI, Segmentation, Diffusion models, Osteoarthritis, Medical imaging, U-Net

## ŽIVOTOPIS

Juraj Perić rođen je 13. listopada 2000. u Osijeku. Pohađao je Osnovnu školu Frana Krste Frankopana u Osijeku te se nakon završetka 2015. godine upisuje III. Gimnaziju Osijek koju završava 2019. godine. Nakon završene srednje škole upisuje se na preddiplomski sveučilišni studij Računarstvo na Fakultetu elektrotehnike, računarstva i informacijskih tehnologija u Osijeku koji uspješno završava 2022. godine. Iste godine upisuje se na diplomski sveučilišni studij Računarstvo, smjer Informacijske i podatkovne znanosti na Fakultetu elektrotehnike, računarstva i informacijskih tehnologija u Osijeku. Za vrijeme diplomskog studija, dobitnik je Rektorove nagrade i priznanja za uspješnost u studiranju na diplomskom studiju. Od 2022. godine radi kao student u tvrtki Barrage kao backend developer.