

**SVEUČILIŠTE U SPLITU
FAKULTET ELEKTROTEHNIKE, STROJARSTVA I
BRODOGRADNJE**

DIPLOMSKI RAD

**UTJECAJ GENERATIVNIH MODELA NA
UČINKOVITOST KLASIFIKACIJE
MEDICINSKIH SLIKA**

Ante Šerić

Split, rujan 2025.



Sveučilišni diplomski studij: **Elektronika i računalno inženjerstvo**

Smjer/Usmjerenje: **Računalno inženjerstvo**

Oznaka programa: 222

Akademска godina: 2024./2025.

Ime i prezime: **Ante Šerić**

JMBAG: 0023143414

ZADATAK DIPLOMSKOG RADA

Naslov: **UTJECAJ GENERATIVNIH MODELA NA UČINKOVITOST KLASIFIKACIJE MEDICINSKIH SLIKA**

Zadatak: Istražiti i predstaviti generativne modele za izradu sintetičkih slika. Napraviti klasifikaciju medicinskih slika korištenjem dubokih neuronskih mreža. Dobivene rezultate koristiti kao polaznu točku za analizu generativnih modela. Implementirati odabrani generativni model i izraditi sintetičke slike. Temeljem standardnih mjera procijeniti učinkovitost generativnog modela u odnosu na duboku neuronsku mrežu.

Rad predan:

Predsjednik
Odbora za diplomski rad:

prof. dr. sc. Josip Musić

Mentor:

izv. prof. dr. sc. Duje Čoko

IZJAVA

Ovom izjavom potvrđujem da sam diplomski rad s naslovom *Utjecaj generativnih modela na učinkovitost klasifikacije medicinskih slika* pod mentorstvom izv. prof. dr. sc. Duje Čoke pisao samostalno, primjenivši znanja i vještine stečene tijekom studiranja na Fakultetu elektrotehnike, strojarstva i brodogradnje, kao i metodologiju znanstveno-istraživačkog rada, te uz korištenje literature koja je navedena u radu. Spoznaje, stavove, zaključke, teorije i zakonitosti drugih autora koje sam izravno ili parafrazirajući naveo u diplomskom radu citirao sam i povezao s korištenim bibliografskim jedinicama.

Student

Ante Šerić

SADRŽAJ

1. UVOD	1
2. PREGLED LITERATURE	2
3. DUBOKE NEURONSKE MREŽE (DNN)	9
3.1. Osnovne komponente i primjeri DNN mreža.....	9
3.2. Metrike za analizu performansi modela	14
3.3. Prijenos znanja.....	18
4. GENERATIVNE KONTRADIKTORNE MREŽE (GAN).....	20
4.1. Razvoj GAN-a i ključne inovacije.....	21
4.2. Svrha i upotreba GAN-a u medicini	23
4.3. Metrike evaluacije za generativne kontradiktorske mreže.....	25
4.4. Izazovi i ograničenja GAN modela	29
5. POSTAVKE EKSPERIMENTA	31
5.1. Opis ulaznih podataka	31
5.2. Dizajn generatora i diskriminadora u odabranoj GAN mreži	34
5.3. Odabir postavki za treniranje GAN mreže	35
6. REZULTATI	38
6.1. Bazni rezultati korištenjem DNN mreže	38
6.2. Rezultati treniranja DCGAN-a	42
6.3. Rezultati DNN klasifikacije s dodanim umjetno generiranim slikama	47
7. ZAKLJUČAK.....	54
8. LITERATURA	55
POPIS OZNAKA I KRATICA	60
SAŽETAK.....	62
SUMMARY.....	63
DODATAK A – popis slika i tablica	64
Popis slika	64
Popis tablica	65

1. UVOD

Primjena umjetne inteligencije u domeni medicinske dijagnostike doživjela je veliki napredak posljednjih godina, posebno u dijagnostici različitih stanja te klasifikaciji bolesti. Međutim, razvoj se robusnih, klinički primjenjivih modela strojnog učenja i dalje suočava s temeljnim izazovom – nedostatkom podataka. To je ograničenje posebno izraženo u primjenama medicinskog snimanja, gdje dobivanje velikih, raznolikih i visokokvalitetnih skupova podataka postaje skupo te dugotrajno. Ovu prepreku dodatno pogoršavaju strogi propisi o privatnosti te etička razmatranja koja ograničavaju dijeljenje podataka pacijenata u svrhu različitih istraživanja.

U ovom radu, navedeni će se problemi i prepreke istražiti te pokušati riješiti u kontekstu klasifikacije stadija demencije na temelju slika magnetske rezonancije (MR) mozga. Nažalost, demencija pogađa milijune ljudi diljem svijeta, a rano otkrivanje ključno je za učinkovitu intervenciju te planiranje liječenja. Generativne kontradiktorne mreže (*engl. Generative adversarial network - GAN*) postavile su se kao obećavajuće rješenje u izazovima nedostatka vrijednih podataka. Učenjem kako generirati sintetičke slike koje blisko oponašaju stvarne medicinske podatke, uz očuvanje privatnosti, GAN modeli nude potencijal proširenja postojećih malih ili neuravnoteženih skupova podataka te poboljšanje generalizacije modela dubokog učenja. U radu će se posebno baviti uvjetnim dubokim konvolucijskim GAN mrežama (*engl. conditional Deep Convolutional GAN - cDCGAN*), vrstom generativnih modela koja je iznimno vrijedna zbog mogućnosti generiranja podataka specifičnih za određenu klasu, što je bitno u medicinskim primjenama s detekcijom različitih zdravstvenih stanja.

Ovaj se rad bavi ključnim pitanjem, može li umjetno proširenje podataka korištenjem cDCGAN-a poboljšati učinkovitost klasifikacije za dijagnosticiranje znakova demencije na MR slikama mozga. Sljedeća će poglavljia sustavno predstaviti sveobuhvatno istraživanje provedeno u ovome projektu, od teorijskih osnova i metodologija kroz provedbu eksperimentalnog dijela i detaljnog prikaza rezultata, do komparativne analize i procjene performansi predložene arhitekture generativnog modela. Na kraju će se procijeniti učinkovitost korištenja GAN modela kao alata za rješavanje problema nedostatka vrijednih medicinskih podataka na temelju promjene točnosti klasifikacije demencijskih stanja te dati prostor mogućim budućim poboljšanjima.

2. PREGLED LITERATURE

Razni istraživački naporci koristili su generativne kontradiktorne mreže za rješavanje izazova u medicinskom snimanju, često se fokusirajući na zadatke proširenja setova podataka, sinteze slika i klasifikacije. U nastavku se donosi kratak pregled odabralih radova i njihovih postignuća.

- **Dijagnoza poremećaja iz spektra autizma (engl. Autism Spectrum Disorder - ASD)** – Y. Zhou i suradnici predložili su novi pristup nazvan *GARN* za dijagnozu ASD-a korištenjem podataka neuroloških snimanja. Inovativno integrira generativne kontradiktorne mreže i duboko Q-učenje (engl. *Deep Q-Learning*). Koristili su nekoliko GAN modela (posebno InfoGAN, DCGAN i WGAN) kako bi proširili ograničene medicinske skupove podataka ABIDE I i II (*Autistic Brain Imaging Data Exchange*). Kombinacija InfoGAN modela s dubokim Q-učenjem donijela je najznačajnije poboljšanje, postižući točnost od 0,873 i osjetljivost od 0,916 na skupu podataka ABIDE II. [1]
- **Generiranje i validacija sintetičkih MR snimki mozga** – Koshino je, sa suradnicima, koristio DCGAN (engl. *Deep Convolutional GAN - DCGAN*) za sintezu umjetnih MR slika mozga iz T1-ponderiranih slika zdravih ispitanika i pacijenata s cerebrovaskularnim oštećenjem.

Kontrola kvalitete slika učinjena je uz pomoć nekoliko radiologa (5 ispitanika). Važno je napomenuti kako su neuroradiolozi (2 ispitanika) ocijenili 45% i 71% generiranih slika kao stvarne, što ukazuje da sintetičke slike mogu biti teško razlučive od stvarnih u nekim situacijama. Ostala tri radiologa drugih specijalizacija postigli su veću točnost raspoznavanja (64% - 83%). Velike razlike u rezultatima sugeriraju varljivu prirodu generiranih slika te da su takve slike sposobne ispravno oponašati stvarnu ljudsku anatomiju. [2]

Slično tome, Marión je razvio arhitekture DCGAN i WGAN-GP (engl. *Wasserstein GAN with Gradient Penalty*) modela za generiranje dvodimenzionalnih MR snimki mozga. U tu svrhu koristio je *Amsterdam Open MRI Collection* set podataka. Njegov WGAN-GP uspješno je generirao visokokvalitetne slike koje su stručnjaci mogli razlikovati od stvarnih s točnošću od 61,25%, pokazujući ostvarivu vjerodostojnost slika. [3]

- **Poboljšanje kvalitete slike i klasifikacija stupnja Alzheimerove bolesti** – studija koja je uključivala MR snimke iz nekoliko medicinskih setova podataka: *Alzheimer's Disease Neuroimaging Initiative* (ADNI), *Australian Imaging, Biomarker & Lifestyle Flagship Study of Ageing* (AIBL) te *National Alzheimer's Coordinating Center* (NACC). Razvila je GAN model koji je treniran na snimkama dobivenih uređajima jakosti 1,5 Tesla (1,5 T) i 3 Tesla (3 T). Navedeni je model generirao nove slike, označene kao 1,5 T*, koje su pokazale poboljšanu kvalitetu u usporedbi s originalnim snimkama (1,5 T). Poboljšanje je primjetno i u srednjoj kvaliteti slike (8,3% na testnim podatcima ADNI-a, 10,0% na AIBL i 9,0% na podatcima NACC). Klasifikacijski modeli za određivanje stadija Alzheimerove bolesti (*engl. Alzheimer's disease - AD*) razvijeni korištenjem ovako generiranih slika pokazali su veću učinkovitost klasifikacije od onih koji su koristili originalne snimke (1,5 T). [4]
- **Nadzirana segmentacija medicinskih slika s generiranjem maski** – Neff i suradnici predložili su novu varijantu GAN-a koja generira ne samo sintetičke medicinske slike, već i odgovarajuće segmentacijske maske. Ovaj je pristup testiran na zadatku segmentacije pluća korištenjem rendgenskih snimaka prsnog koša. Pokazao je kako treniranje modela isključivo na generiranim uzorcima postiže rezultate usporedive s modelom treniranim na stvarnim slikama, za vrlo male skupove podataka. Iako same generirane slike nisu dovoljne za zamjenu čitavog skupa podataka stvarnih slika, itekako su vrijedan alat u ublažavanju nedostatka podataka u zadatcima strojnog učenja. [5]
- **Detekcija unutarnjeg krvarenja** – Jeong je, sa svojim suradnicima, osmislio vrstu cDCGAN-a kao alat za proširenje korištenih uzoraka. Navedenim se koristio u zadatku detekcije unutarnjeg krvarenja koristeći CT snimke iz *RSNA 2019 Brain CT Hemorrhage Challenge* skupa podataka. Za manjinske klase, poput epiduralnih krvarenja, korištenje generativnog modela dovelo je do najmanje dvostrukog poboljšanja performansi u usporedbi s tradicionalnim metodama proširenja skupa uzoraka. Uočeno je poboljšanje i u klasifikaciji korištenih neuronskih mreža (veća točnost kod manjinskih klasa u višeklasnoj klasifikaciji). [6]
- **Sinteza i segmentacija medicinskih slika** – Skandarani i suradnici proveli su empirijsku studiju o različitim GAN arhitekturama (DCGAN, LSGAN, WGAN, HingeGAN i

StyleGAN) na različitim modalitetima medicinskog snimanja, kao što su MR srca, CT jetre i RGB slike mrežnice. Otkrili su da su GAN-ovi s najboljim performansama sposobni generirati vizualno realistične medicinske slike koje bi mogle zavarati obučene stručnjake u vizualnom Turingovom testu. Međutim, njihovi rezultati segmentacije pokazali su da nijedan GAN model nije mogao u potpunosti replicirati bogatstvo stvarnih medicinskih skupova podataka. Jednostavniji GAN modeli nisu ostvarili dobre rezultate, dok je sofisticiraniji StyleGAN, unatoč mogućnosti generiranja realističnih slika, zahtjevao značajne računalne resurse za treniranje (otprilike 30 dana za treniranje nad podatcima iz ACDC skupa). [7]

- **Sinteza visokorezolucijskih medicinskih slika zdjelice** – Fenty i suradnici istraživali su potencijal StyleGAN modela za sintezu medicinskih slika visoke rezolucije, s posebnim naglaskom na razumijevanje i manipuliranje njegovim latentnim prostorom. Cilj je bio generirati prilagođene sintetičke slike pacijenata koje bi mogле biti korisne za treniranje algoritama dubokog učenja. Za treniranje modela koristili su slike kompjutorizirane tomografije (*engl. Computed Tomography - CT*) i T2-ponderirane MR snimke zdjelične regije, prikupljene od 117 pacijenata.

Istraživanje je uspješno prezentiralo sposobnost StyleGAN modela za manipuliranje specifičnim značajkama slike poput modaliteta snimke (MR snimke u CT te obrnuto), spola i položaja uzdužnog presjeka podešavanjem latentnih vektora. Model je postigao FID (*engl. Fréchet Inception Distance*) rezultat od približno 12,3. [8]

- **Klasna neravnoteža u segmentaciji MR snimki** – istraživački tim, pod vodstvom prof. Cuija, razvio je novu arhitekturu dubokog učenja pod nazivom uvjetna kontrastna generativna kontradiktorna mreža (*engl. Conditional Contrastive Generative Adversarial Network - cCGAN*) kako bi riješili problem neravnoteže klase u medicinskim skupovima podataka za zadatke segmentacije. Navedeni je model bazirao svoj izlaz s obzirom na klasu ulaznih podataka, a sadržavao je modul za uravnoteženje klase te nadziranu neuronsku mrežu temeljenu na kontrastima (*engl. Supervised Contrastive-based Learning Network - SCoLN*). Model je postigao vrhunske performanse s visokim rezultatima koeficijenata sličnosti na različitim vrlo neuravnoteženim skupovima podataka za segmentaciju

medicinskih snimki (korišteni skupovi: BUS2017, DDTI, LiTS MICCAI 2017, ATLAS i BRATS 2015). [9]

- **Klasifikacija vrsta tumora mozga s pomoću Pix2Pix GAN-a** – Istraživački laboratorij, na čelu s Onakpojeruo-m, postao je pionir u korištenju Pix2Pix GAN modela za stvaranje sintetičkih skupova podataka za klasifikaciju tumora mozga. Fokusirali su se na četiri klase: gliom, meningeom, tumor hipofize te dobroćudni tumor.

Za potrebe treniranja i testiranja modela korišten je set podataka pronađen na mrežnoj stranici Kaggle. Sastojao se od 3264 stvarnih MR slika koji su istraživači proširili s 10000 sintetičkih slika.

Predloženi je klasifikacijski model (uvjetna konvolucijska neuronska mreža – cDCNN) postigao točnost od 0,86 koristeći sintetičke slike.

Studija je otkrila kako korištenje setova podataka koji su prošireni generiranim slikama (načinjene Pix2Pix modelom) može značajno poboljšati točnost klasifikacije tumora na mozgu u usporedbi s modelima treniranim na samo izvornim podatcima (točnost od samo 0,72 s izvornim podatcima). [10]

- **Proširenje podataka za segmentaciju duboke sive tvari mozga** – Bowles i suradnici koristili su skup podataka OASIS-1 za svoju studiju gdje su obrađivali problem s više klase. Kao označeni dio seta podataka koristili su 35 ručno označenih slika (nastale trudom medicinskih stručnjaka) iz OASIS-1 seta. Cijeli navedeni skup uzoraka (436 slika), koji uključuje ispitanike širokog raspona dobi i različitih stadija demencije, poslužio je kao neoznačeni skup podataka koji će se koristiti za umjetno proširenje uzoraka s pomoću GAN modela.

Izradili su generativni model koji je uspješno povećao anatomsku varijabilnost testiranih medicinskih slika, posebno u skupini starijih pacijenata te onih s obilježjima Alzheimerove bolesti. Navedeno je dovelo do poboljšanja točnosti segmentacije za strukture poput hipokampusa i amigdale. [11]

- **Generirani podatci za otkrivanje lezija uzrokovanih multiplom sklerozom (engl. Multiple Sclerosis - MS)** – Brugnara i suradnici koristili su T1-ponderirane snimke iz

OASIS-3 skupa podatka (735 MR snimki, 535 pacijenata) za generiranje sintetičkih FLAIR slika s MS lezijama. Slike su dobili korištenjem vlastitog uvjetnog GAN-a. U tu svrhu, kombinirali su T1-ponderirane sekvence iz OASIS-3 skupa sa FLAIR sekvencama (iz vlastitog, internog skupa podataka za treniranje).

Heterogenost OASIS-3 skupa uzoraka pomogla je mreži da uči iz raznolikijeg seta stanja, značajno poboljšavajući generalizaciju i performanse klasifikacijskih modela na novim, neviđenim MR podatcima (površina ispod krivulje – *engl. Area Under Curve* - AUC: 0,836 s originalnim podatcima, 0,933 s proširenjem sintetičkim podatcima). [12]

- **Treniranja 3D-StyleGAN modela** – studija (Hong i suradnici) je koristila veliku kolekciju uzoraka od približno 12 000 T1 MR snimki mozga iz više javno dostupnih skupova podataka (OASIS, ADNI, ABIDE, ADHD2000, MCIC, PPMI, HABS i Harvard GSP) za treniranje vlastitog trodimenzionalnog StyleGAN modela. Od ukupnog broja slika, 7392 uzorka je korišteno za treniranje modela.

Rad je pokazao mogućnost sinteze realističnih 3D (trodimenzionalnih) slika mozga u različitim rezolucijama (rezolucija 2 milimetra, što odgovara 80 x 96 x 112 prostornih jedinica - voksela). Također je pokazano kako se mogućnost miješanja stilova može koristiti za kreiranje anatomskeh varijacija između slika. [13]

Tablica 2.1. Pregled korištenih metoda i rezultata iz srodrne literature

Autori studije	Korišteni GAN model(i) i ostale DNN mreže	Zadatak	Ključni rezultati / performanse
Neff i suradnici (2017.)	nova GAN varijanta (generira slike + maske za segmentaciju); U-Net za segmentaciju	nadzirana segmentacija medicinskih slika (rendgen pluća)	treniranje modela na isključivo generiranim slikama postiglo je rezultate usporedive s treniranjem na stvarnim slikama za vrlo male skupove podataka
Koshino i suradnici (2018.)	DCGAN	generiranje sintetičkih MR slika mozga i validacija kvalitete s pomoću stručnjaka	u testiranju je sudjelovalo 5 radiologa: 2 neuroradiologa ocijenili su 45% i 71% slika stvorenih DCGAN-om kao stvarne; ostala 3 (drugih specijalizacija) – 64% do 80% točnosti
Bowles i suradnici (2018.)	GAN s progresivnim rastom (<i>engl. Progressive Growing GAN - PGGAN</i>)	segmentacija duboke sive tvari (više klase) na T1-ponderiranim MR slikama unutar skupa podataka o AD	poboljšana je segmentacija struktura za koje se zna da su pogodjene AD-om, poput hipokampa i amigdale.; pokazano je da 6 označenih slika proširenih sintetičkim podacima može postići rezultate ekvivalentne korištenju 12 ili 24 označene slike bez proširenja.
Fetty i suradnici (2020.)	StyleGAN	sinteza medicinskih slika visoke rezolucije (CT zdjelice i T2-ponderirane MR snimke); manipulacija latentnim prostorom za kontrolu atributa slike	demonstrirana manipulacija atributima slike putem latentnih vektora; postignut FID rezultat od 12,3
Studija o klasifikaciji stadija AD bolesti (2020.)	modificirani GAN (3D konvolucijski blokovi u generatoru, 5 konvolucijskih blokova u diskriminatoru); FCN klasifikator	poboljšava učinkovitost klasifikacije AD-a i kvalitetu MR slika	poboljšana srednja kvaliteta slike (8,3% na testnim podatcima ADNI-a, 10,0% na AIBL i 9,0% na podatcima NACC)
Skandarani i suradnici (2021.)	DCGAN, LSGAN, WGAN, HingeGAN, StyleGAN, SPADE GAN; U-Net za segmentaciju	procjena korisnosti medicinskih podataka generiranih GAN-om za sintezu i segmentaciju slike	nijedan GAN nije u potpunosti replicirao bogatstvo stvarnih medicinskih skupova podataka; obuka StyleGAN-a zahtjevala je značajne računalne resurse (oko 30 dana za ACDC skup podataka)

Hong i dr. (2021.)	3D-StyleGAN (proširenje StyleGAN2)	generativno modeliranje 3D medicinskih slika, posebno sinteza visokokvalitetnih 3D MR snimki mozga	demonstrirana je mogućnost sinteze realističnih 3D MR snimki mozga u rezolucijama poput 2 mm; miješanje stilova može se koristiti za "prijenos" anatomske varijacije između slika
Jeong i suradnici (2022.)	cDCGAN; klasifikatori: DenseNet-169, ResNet-152, ResNeXt-101, SE-ResNeXt-101	detekcija i klasifikacija unutarnjeg krvarenja s pomoću CT slika (umjetno proširenje skupa uzoraka)	proširenje cDCGAN-om dovelo je do najmanje 2x poboljšanja performansi za super-manjinske klase (npr. epiduralna krvarenja) u usporedbi s tradicionalnim metodama proširenja
Marión (2023.)	DCGAN, WGAN-GP	generiranje sintetičkih 2D i 3D MR snimki mozga kao rješenje nedostatka podataka	WGAN-GP je proizveo visokokvalitetne slike koje su radiološki stručnjaci razlikovali od stvarnih s točnošću od samo 61,25%, pokazujući vjerodostojnost uzoraka
Cui i dr. (2024.)	cCGAN	rješavanje neravnoteže klase u skupovima podataka MR segmentacije	postignuti su najsuvremeniji DSC (<i>engl. Dice Similarity Coefficient</i>) rezultati za vrlo neuravnotežene skupove podataka: BUS2017 ($0,965 \pm 0,012$), DDTI ($0,896 \pm 0,091$), LiTS MICCAI 2017 ($0,786 \pm 0,046$), ATLAS ($0,712 \pm 1,5$), BRATS 2015 ($0,877 \pm 1,2$)
Onakpojeruo i dr. (2024.)	Pix2Pix GAN; cDCNN	klasifikacija tumora mozga (gliom, meningeom, hipofiza, benigni) korištenjem sintetičkih skupova podataka	klasifikacijski je model postigao točnost od 86% sa sintetičkim slikama (u odnosu na 72% s originalnim setom uzoraka)
Brugnara i suradnici (2024.)	cGAN	poboljšati generalizaciju i performanse modela te ublažiti pomak domene i neravnotežu klase (klasifikacija lezija uzrokovanih MS-om)	pokazalo je značajno poboljšanje performansi kada se primijeni na testne podatke, s AUC-om koji se povećao s 83,6% bez sintetičkih podataka na 93,3% s proširenjem sintetičkim podatcima
Y. Zhou i suradnici (2024.)	GARL (InfoGAN, DCGAN, WGAN + duboko Q učenje)	dijagnoza ASD-a korištenjem podataka neuroloških pretraga (tehnike proširenja uzoraka)	GARL je, s InfoGAN-om i DQN-om, dao točnost od 0,873 i osjetljivost od 0,916 na skupu podataka ABIDE II; proširenje podataka s pomoću GAN-a općenito je poboljšalo performanse modela klasifikacije

3. DUBOKE NEURONSKE MREŽE (DNN)

Duboke neuronske mreže (DNN) računalni su sustavi načinjeni od mnogo slojeva te dizajnirani za učenje složenih uzoraka s ulaznih podataka. **Konvolucijske neuronske mreže** (CNN) posebna su vrsta DNN mreža koje su izrazito dobre u pronalaženju značajki na slikama kako bi se detektirali sami objekti ili kategorije objekata. [14]

DNN arhitekture posebno su dizajnirane za prepoznavanje i izdvajanje uzoraka iz slika. Njihovu srž čine konvolucijski slojevi, vješto upareni za izdvajanje značajki ili uzoraka te slojevi udruživanja (*engl. pooling layers*) koji čine takvu detekciju neosjetljivom na male varijacije ili pomake značajki unutar ulazne slike. U višeslojnim DNN-ovima, plitki slojevi služe izdvajajuju lokalnih značajki, dok oni dublji mogu razaznati i veće, globalne uzorke. [14]

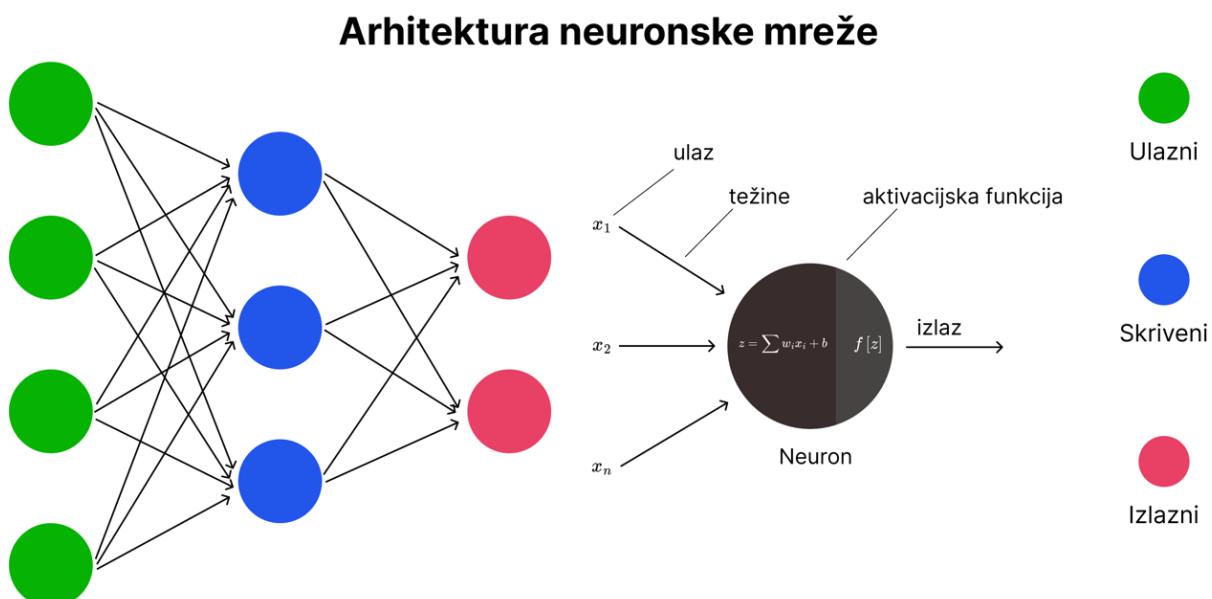
Učinkovitost i mogućnosti generalizacije modela strojnog učenja, uključujući DNN-ove, uvelike ovisi o kvaliteti, količini i raznolikosti podataka korištenih u treniranju modela. [15] Međutim, prikupljanje velikih, raznovrsnih i adekvatno označenih skupova podataka često predstavlja značajan izazov, posebno u specijaliziranim slučajevima kao što je medicinska dijagnostika. Ove prepreke proizlaze iz problema poput nedostatka podataka zbog zabrinutosti za privatnost pacijenta ili etičkih ograničenja, visokih troškova i dugotrajnog vremena potrebnog za precizno i stručno ručno označavanje. [5] Nedostatak podataka dovodi do pojave neuravnoteženosti korištenih klasa u skupovima podataka koji za posljedicu može imati smanjenje točnosti klasifikacije. Navedena situacija dovodi i do povećanog rizika od prekomjernog prilagođavanja modela (*engl. overfitting*) zastupljenijim klasama tijekom faze treninga. [7]

3.1. Osnovne komponente i primjeri DNN mreža

Osnovne građevne slojeve DNN mreže čine:

- **konvolucijski slojevi** – primjenjuju skup filtera koji se tijekom treniranja prilagođavaju i postupno mijenjaju na setu ulaznih slika, detektirajući različite značajke poput rubova, tekstura ili uzoraka.
- **slojevi udruživanja** (npr. MaxPooling, Avarage Pooling) – ravnomjerno raspoređeni u strukturi mreže, ovi slojevi smanjuju prostorne dimenzije mapa značajki, čineći mrežu robusnijom na male varijacije ulazne slike te smanjujući računalno opterećenje.

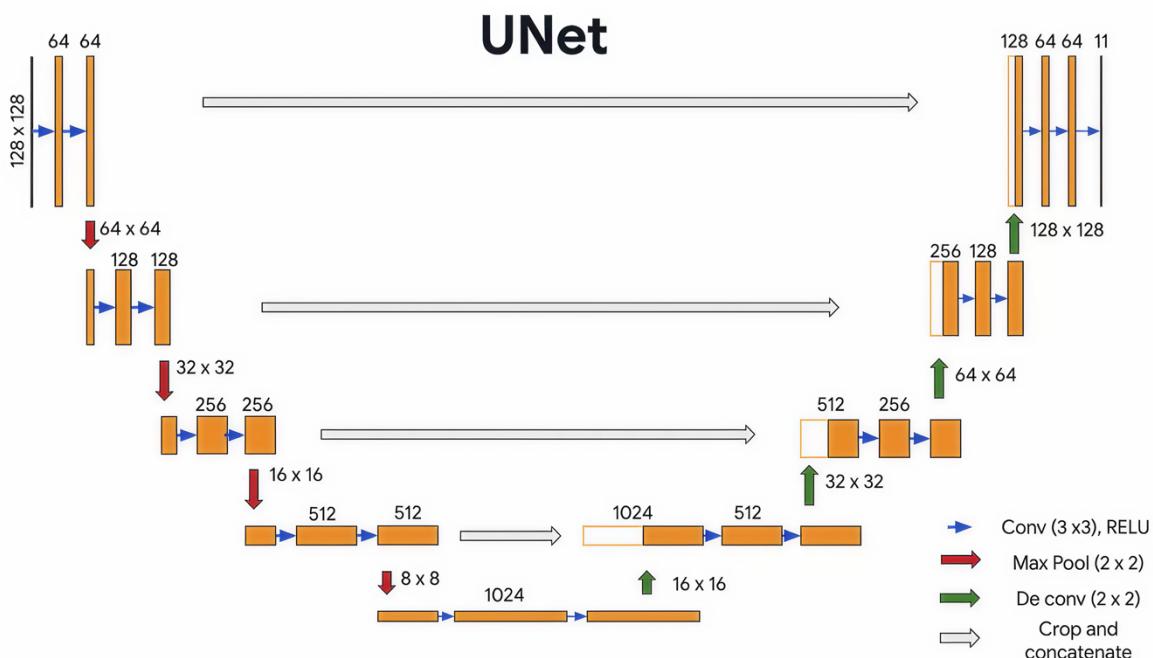
- **aktivacijske funkcije** – nelinearne funkcije koje se primjenjuju nakon konvolucijskih slojeva kako bi se uvela nelinearnost u model, omogućujući učenje složenih uzoraka. Uobičajeni korišteni primjeri uključuju ReLU (*engl. Rectified Linear Unit*) i LeakyReLU [52].
- **normalizacija po grupi** (*engl. Batch Normalization*) – koristi se u postupku normalizacije ulaznih podataka u slojeve, poboljšavajući optimizaciju modela te stabilnost učenja rješavanjem unutarnjeg pomaka kovarijance.
- **potpuno povezani slojevi** (*engl. Fully Connected Layers*) – ovi slojevi povezuju svaki neuron u jednom sloju sa svakim neuronom u sljedećem susjednom. Obično se koriste pri kraju mreže za proces klasifikacije, na temelju značajki koje su izdvojili raniji slojevi.
- **izlazni sloj** – završni sloj koji daje predviđanje modela. Kod klasifikacije s više klasa, može koristiti sigmoidne neurone (koriste sigmoidnu aktivacijsku funkciju) za nezavisna predviđanja između klasa.



Slika 3.1. Arhitektura neuronske mreže s označenim glavnim dijelovima

Neki od primjera DNN arhitektura koje se često koriste za analizu slika:

- **U-Net** – potpuno konvolucijska mreža koja je široko prihvaćena. Posebno je uspješna u segmentaciji biomedicinskih slika. Njena arhitektura značajna je po zamjeni procesa udruživanja utemeljenog na funkciji maksimuma (*engl. max pooling*) s varijantom utemeljenom na funkciji prosjeka (*engl. average pooling*) za smanjenje dimenzija mapa značajki. Sastoji se od kodera i dekodera čiji su neki slojevi vezani direktnim vezama za prijenos informacija. Ove veze dekoderu omogućuju kombiniranje lokalnih i globalnih značajki za precizniju lokalizaciju te poboljšanu točnost segmentacije. U-Net omogućava treniranje i na manjem setu podataka korištenjem jakih tehniki umjetnog proširenja podataka (*engl. data augmentation techniques*).



Slika 3.2. Prikaz primjera U-Net arhitekture [16]

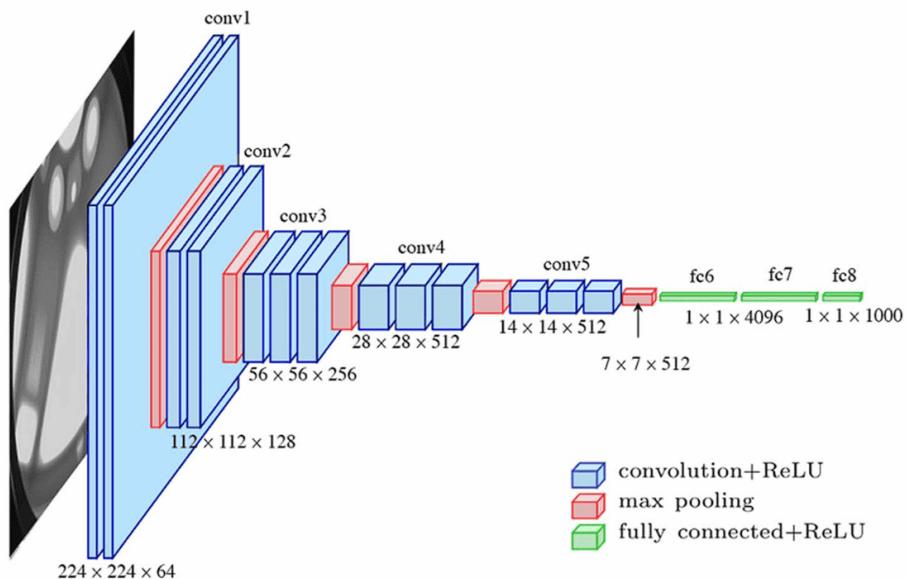
Postoji više standardnih klasifikacijskih DNN arhitektura, a za analizu odabrana su tri najsvremenija modela dubokih neuronskih mreža: VGG19, MobileNetV2 i EfficientNetB0.

- **VGG19** – duboki je CNN model objavljen 2014. godine. Poznat je po svojoj dubini te korištenju jednostavnih, ponavljajućih građevnih blokova. Ukupno ima 19 slojeva, uključujući konvolucijske (za pronađak značajki), potpuno povezane slojeve (za kombiniranje informacija) te MaxPooling slojeve udruživanja.

Ova arhitektura zamjenjuje velike konvolucijske filtere s nekoliko sekvenčijalnih veličine 3×3 , čime je povećana učinkovitost detekcije složenih značajki. Svi skriveni slojevi koriste istu aktivacijsku funkciju, ReLU. Završni slojevi su potpuno povezani,

posljednji ima dimenzije 1000, što odgovara 1000 klasa u skupu podataka ImageNet na kojima je model izvorno treniran.

VGG19 se široko koristi u istraživanjima medicinskog snimanja. Često se odabirao za klasifikaciju slika u ovoj domeni zbog svojih dobrih performansi pri testiranju na ImageNet-u (točnost od 71,3%). [17] Međutim, posjeduje relativno velik broj parametara (22,9 milijuna) u usporedbi s drugim suvremenim modelima. No, njegovi rezultati za klasifikaciju slika na medicinskim skupovima podataka usporedivi su s drugim istraživanjima.

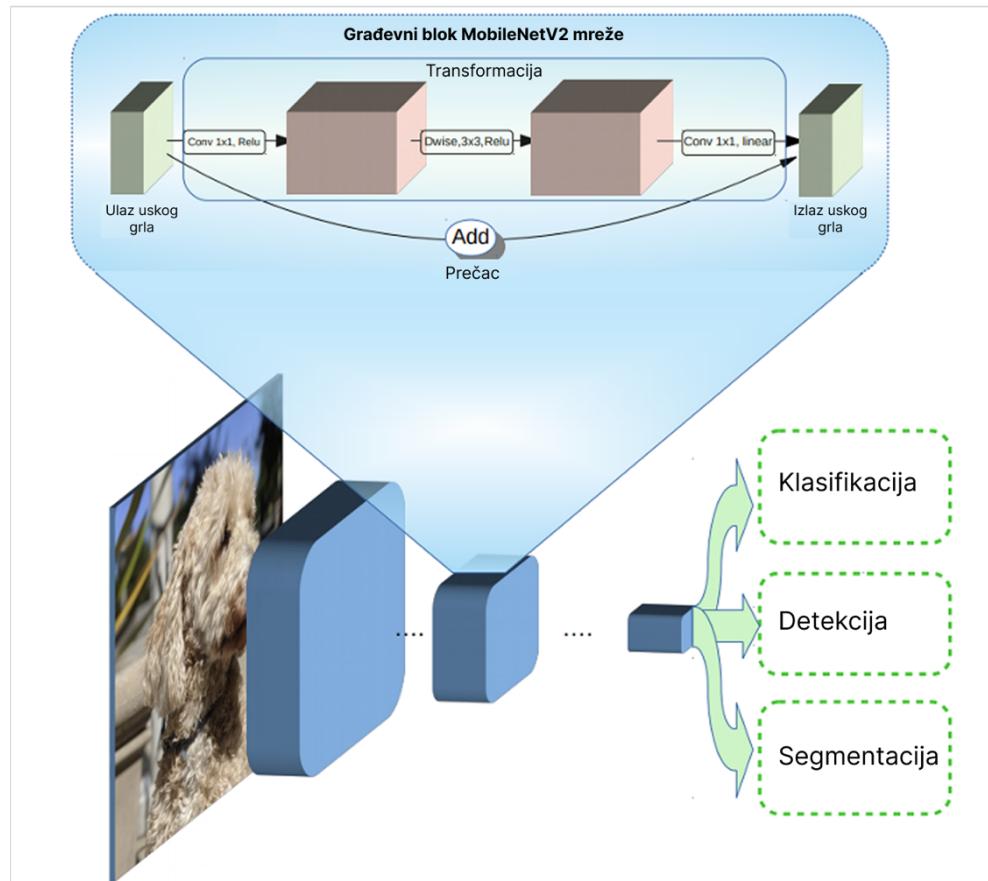


Slika 3.3. Arhitektura VGG19 modela [18]

- **MobileNetV2** – razvijen je 2018. godine u Googleovom istraživačkom timu. Njegovo ime ističe njegovu namjenu: biti dovoljno učinkovit i lagan za pokretanje na mobilnim uređajima za zadatke poput klasifikacije i detekcije objekata na slikama. Koristi konvolucijske slojeve koji su odvojivi po dubini, a koji su učinkoviti građevni blokovi.

Novost koji model uvodi su linearne uske grle između slojeva te prečaci koji povezuju ta uske grle. Navedene značajke pomažu modelu u razumijevanju slika od osnovnih piksela do koncepata više razine, a prečaci ubrzavaju fazu treniranja te poboljšavaju točnost modela, slično rezidualnim vezama u drugim mrežama.

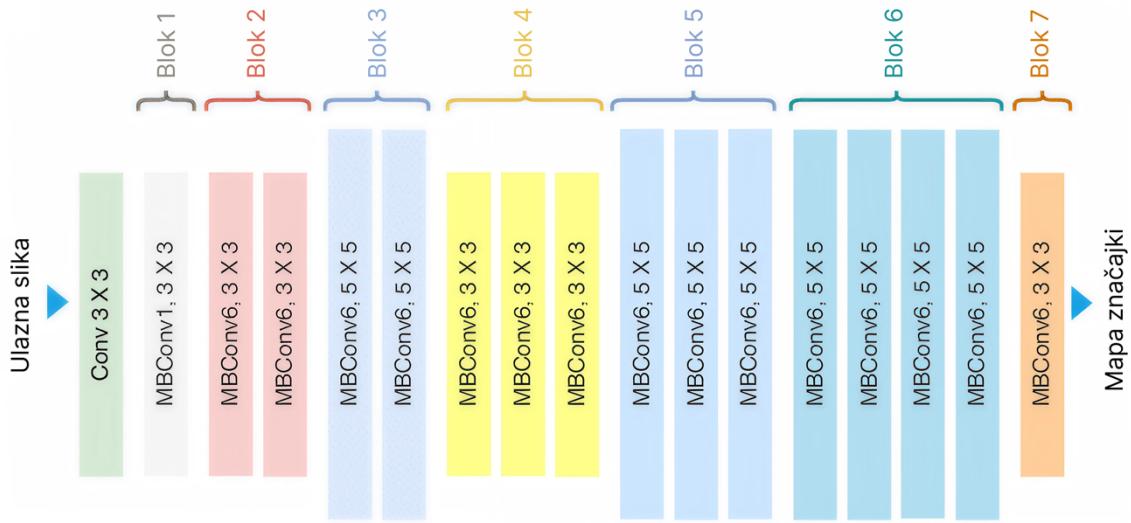
MobileNetV2 cijenjen je u području medicinske dijagnostike zbog svoje učinkovitosti. Postigao je istu točnost nad ImageNet setom podataka kao i VGG19 (71,3%), ali sa znatno manje parametara (svega 3,5 milijuna). To ga čini vrlo prikladnim za hardversku implementaciju zbog manje veličine. U studijama se često pokazao kao najbolji ili drugi najbolji model, a istovremeno je nudio značajne uštede vremena tijekom obuke (npr. 44 minute manje od VGG19 i 19 minuta manje od EfficientNetB0 u jednom slučaju). [17] Smatra se snažnim kompromisom između točnosti i složenosti.



Slika 3.4. Arhitektura MobileNetV2 modela [19]

- **EfficientNetB0** (EfficientNetV2) – predstavljen 2021. godine, pripada grupi CNN modela usmjerenih na brže treniranje i učinkovitije korištenje parametara. Dizajniran je s pomoću naprednog procesa pretraživanja koji je uključivao nove operacije (koristenje MBConv i Fused-MBConv u ranijim slojevima). Sposoban je mnogo brže učiti te je znatno manji (do 6,8 puta) od prethodnih vodećih modela. Preferira korištenje manjih veličina filtera (3×3) zbog postizanja konzistentnosti među slojevima.

Jedan je od najčešćih suvremenih odabira za klasifikaciju slika medicinske naravi zbog visoke točnosti na testnom setu ImageNet-a (78,7%) te učinkovitog broja parametara (7,2 milijuna). [17] Unatoč činjenici da ima više parametara od prethodno spomenutog MobileNetV2 modela, općenito pokazuje najbolje performanse među testiranim modelima.



Slika 3.5. Arhitektura EfficientNetB0 modela [20]

3.2. Metrike za analizu performansi modela

Pri procjeni performansi DNN-ova, posebno u zadatcima klasifikacije slika, koriste se različite metrike kako bi se pružilo sveobuhvatno razumijevanje učinkovitosti modela. Metrike predstavljaju ključ u usmjeravanju napretka prema maksimalnoj iskoristivosti modela i podataka. U primjenama klasifikacije s više klasa, posebno u radu s neuravnoteženim skupovima podataka, primjena je metrika ključna za izbjegavanje donošenja pogrešnih zaključaka.

Neke od najčešće korištenih metrika za analizu performansi DNN modela, koje se obično primjenjuju u zadatcima medicinske naravi, opisane su u nastavku.

3.2.1. Matrica zbumjenosti i pogreške predviđanja modela

Matrica zbumjenosti (*engl. Confusion Matrix*) dijagnostički je alat koji prikazuje broj točnih i netočnih predviđanja za svaku klasu kod primjena klasifikacije. Nudi detaljan uvid u greške koje je učinio model. Kod klasifikacije s više klasa, obično se sastoji od četiri ključne komponente:

- **Stvarno pozitivni** (*engl. True Positives - TP*) – broj pozitivnih uzoraka koje je model ispravno predvidio kao pozitivne
- **Stvarno negativni** (*engl. True Negatives - TN*) – broj negativnih uzoraka koje je model ispravno predvidio kao negativne
- **Lažno pozitivni** (*engl. False Positives - FP*) – broj negativnih uzoraka koje je model netočno predvidio kao pozitivne; poznate i kao pogreška tipa I
- **Lažno negativni** (*engl. False Negatives - FN*) – broj pozitivnih primjera koje je model netočno predvidio kao negativne; poznate i kao pogreške tipa II

Ćelije na dijagonali matrice zbnjenosti predstavljaju točna predviđanja modela (TP ili TN), dok preostale ćelije predstavljaju pogreške (FP ili FN).

		Predviđanje modela	
		Pozitivni	Negativni
Stvarne vrijednosti	Pozitivni	Stvarno pozitivni (TP)	Lažno pozitivni (FP)
	Negativni	Lažno negativni (FN)	Stvarno negativni (TN)

Slika 3.6. Struktura matrice zbnjenosti

3.2.2. Točnost, preciznost, osjetljivost i F1 rezultat

Točnost (*engl. accuracy*) je metrika koja izražava točnost pozitivnih predviđanja, pokazujući, od svih pozitivnih predviđanja, koliko ih je zaista bilo pozitivnih. Ključna je stavka ako je prioritet smanjenje lažno pozitivnih predviđanja (FP). Točnost se računa kao:

$$Točnost = \frac{TP}{TP + FP} \quad (3.1)$$

gdje su parametri TP i FP već ranije objašnjeni.

U primjerima s više klase, točnost je moguće usrednjiti korištenjem mikro, makro ili ponderiranih tehnika.

Preciznost (*engl. precision*) je, s druge pak strane, specifična za određenu klasu uzorka. Određuje koliko je uzoraka koji su predviđeni kao dio neke klase zaista dio te klase. Fokusira se na kvalitetu pozitivnih predviđanja za određenu klasu – to je lokalna, klasno-ovisna mjeru. Vrijednost se također kreće od 0 do 1 (ili od 0% do 100%), gdje više vrijednosti označavaju bolje performanse modela.

Recall faktor (poznat kao **osjetljivost** ili stopa stvarno pozitivnih) je metrika koja pokazuje potpunost pozitivnih predviđanja. Izražava odnos: od svih stvarno pozitivnih uzoraka, koliko je ispravno predviđeno kao pozitivnih. Važna je ako je cilj minimizirati lažno negativne (FN) uzorce. Osjetljivost se kao takva računa:

$$\text{Recall (osjetljivost)} = \frac{TP}{TP + FN} \quad (3.2)$$

gdje su parametri TP i FN već ranije objašnjeni.

U zadatcima s više klase, vrijednosti osjetljivosti mogu se usrednjiti.

F1 rezultat je mjeru koja pruža harmonijsku sredinu preciznosti i osjetljivosti, sažimajući matricu zbnjenosti u jedan broj. Predstavlja uravnoteženu mjeru u slučaju neravnomjerne raspodjele klasa. F1 se rezultat računa kao:

$$F1 \text{ rezultat} = \frac{2 * \text{preciznost} * \text{osjetljivost}}{\text{preciznost} + \text{osjetljivost}} = \frac{2 * TP}{2 * TP + FP + FN} \quad (3.3)$$

gdje su parametri TP, FP i FN već ranije objašnjeni.

Poput točnosti i osjetljivosti, može biti mikro, makro ili ponderirano usrednjena u scenarijima s više klasa.

3.2.3. ROC krivulja i površina ispod krivulje (AUC)

ROC krivulja (*engl. Receiver Operating Characteristic*) je vrsta grafa koja vizualizira stopu stvarno pozitivnih rezultata (*engl. True Positive Rate - TPR*) na y-osi u odnosu na stupu lažno pozitivnih rezultata (*engl. False Positive Rate - FPR*) na x-osi.

- **Stopa stvarno pozitivnih rezultata (TPR)** ekvivalentna je parametru osjetljivosti (recall) te se računa kao

$$TPR = \frac{TP}{TP + FN} \quad (3.4)$$

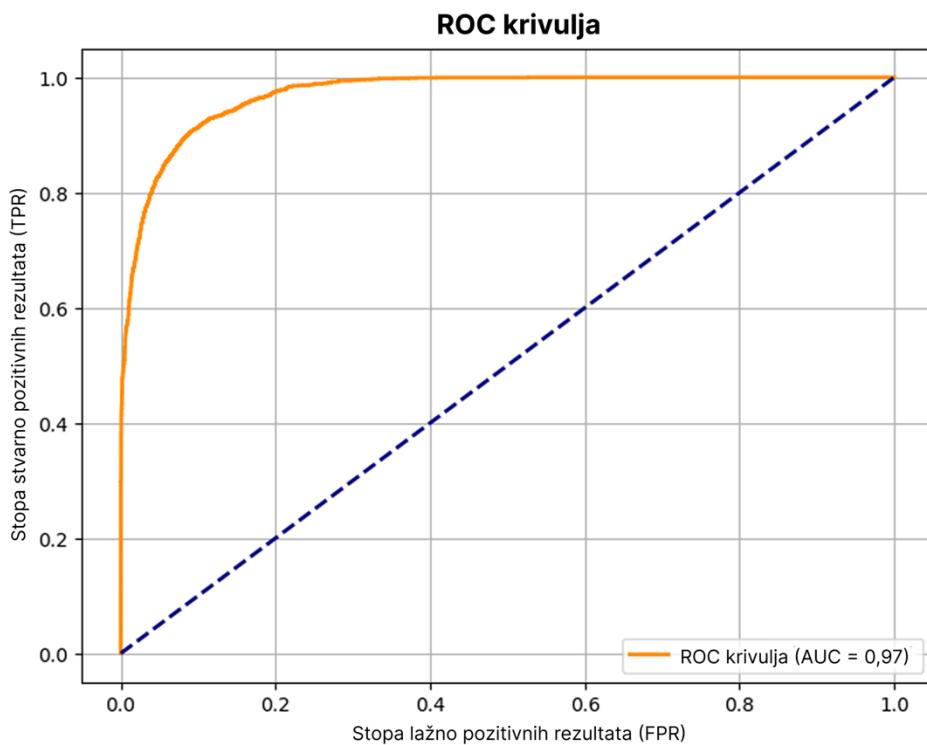
gdje su parametri TP i FN već ranije objašnjeni.

- **Stopa lažno pozitivnih rezultata (FPR)** pokazuje koliko je negativnih uzoraka model netočno klasificirao kao pozitivne, izračunat kao

$$FPR = \frac{FP}{FP + TN} \quad (3.5)$$

gdje su parametri FP i TN već ranije objašnjeni.

Površina ispod krivulje (AUC) kvantificira ukupne performanse predstavljene ROC krivuljom. Viša vrijednost AUC-a, gdje se krivulja približava gornjem lijevom kutu grafa, ukazuje na bolje performanse klasifikacije. Kada se izračun AUC-a te crtanjem ROC krivulje izolira posebno za svaku klasu, ove mjere postaju korisne u procjeni performansi modela u neuravnoteženim skupovima podataka.



Slika 3.7. Primjer ROC krivulje i površine ispod krivulje (AUC)

Izbor metrika ovisi o specifičnoj primjeni i zahtjevima koje je potrebno ispuniti. Iako jednostavne metrike, poput točnosti klasifikacije, osjetljivosti ili F1 rezultata pružaju brz

pregled u ponašanje promatranog modela, mogu biti zavaravajuće, posebno s neuravnoteženim skupovima podataka. Navedeno dovodi do potrebe uključivanja i drugih metrika kako bi se dobio širi pogled na performanse modela.

3.3. Prijenos znanja

Prijenos znanja (*engl. Transfer Learning - TL*) široko je prihvaćen pristup u području dubokog učenja, posebno pri radu s ograničenim skupovima podataka, što je uobičajeno u medicinskoj dijagnostici. Umjesto treniranja modela „od nule“, standardni pristup utedeljen na prijenosu znanja uključuje korištenje modela koji su već prethodno trenirani na masivnim skupovima podataka opće namjene, kao što je ImageNet. Ideja koja stoji iza principa je da su mnoge značajke niske razine (lokalne značajke kao rubovi, oblici, boje), pa čak i neki uzorci visoke razine naučeni iz velikog izvornog seta podataka, prenosivi i mogu biti korisni za sličan zadatak.

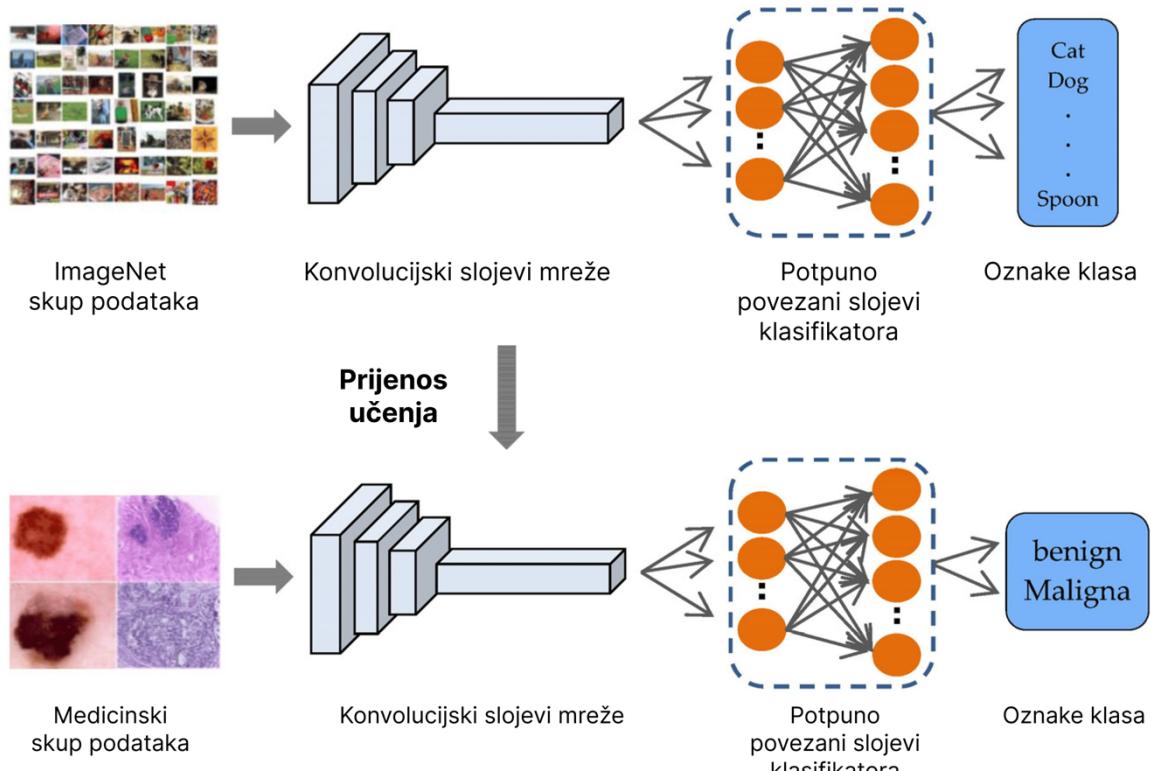
Prednosti prijenosa znanja uključuju:

- **smanjeno vrijeme treniranja modela** – ponovna uporaba prethodno obučenih modela značajno štedi vrijeme i računalne resurse, budući da je početno učenje značajki već odrađeno
- **poboljšana kvaliteta naučenih značajki** – prethodno obučeni modeli naučili su robusne i generalizirajuće značajke koje se mogu učinkovito primijeniti i na nove zadatke, čak i s relativno manjim setovima ulaznih podataka

U kontekstu klasifikacije slika medicinske dijagnostike, modeli VGG19, MobileNetV2 i EfficientNetB0 posebno su odabrani zbog dostupnosti njihovih prethodno treniranih verzija u Keras aplikacijama [21]. Važni kriteriji pri odabiru korištenih modela bili su i njihova točnost na skupu podataka za validaciju ImageNet-a, broj parametara (važno za ograničene računalne resurse) te godina izlaska. Pokazano je kako su VGG19 i MobileNetV2 ostvarili slične rezultate točnosti, premda se ta dva modela izrazito razlikuju u broj parametara.

Međutim, postupak prijenosa znanja u području medicinskog snimanja može predstavljati izazov zbog značajnih razlika između prirodnih slika (korištenih za početno učenje modela i medicinskih slika (ciljna domena). Naučene značajke i uzorci iz velikih setova prirodnih slika ne moraju uvijek biti optimalno prikladni za različite karakteristike medicinskih slika. Unatoč tim izazovima, prijenos znanja ostaje učinkovit pristup analizi medicinskih slika.

Tipičan proces uključuje fino podešavanje određenih ili svih slojeva prethodno trenirane mreže kako bi bila prilagođena ciljanom skupu podataka.



Slika 3.8. Osnovna ideja prijenosa učenja [22]

4. GENERATIVNE KONTRADIKTORNE MREŽE (GAN)

Generativne kontradiktorne mreže klase su modela dubokog učenja koja je privukla značajnu pozornost zbog svoje sposobnosti stvaranja visokokvalitetnih umjetnih podataka. Početke ove tehnologije obilježila je grupa istraživača pod vodstvom prof. Goodfellowa (2014. godine). GAN modeli rade na principu suparničkog učenja, gdje se dvije neuronske mreže, generator (G) i diskriminatator (D), natječe jedna protiv druge u minmax igri. [15]

Generator (G) ima zadatak proizvesti nove, sintetičke uzorke podataka koji su što realniji, s ciljem oponašanja karakteristika stvarne distribucije podataka korištenih za treniranje. Obično uzima vektor slučajnog šuma (*engl. noise vector*) kao ulaz i transformira ga u sintetičku podatkovnu točku (sintetički uzorak) u latentnom prostoru.

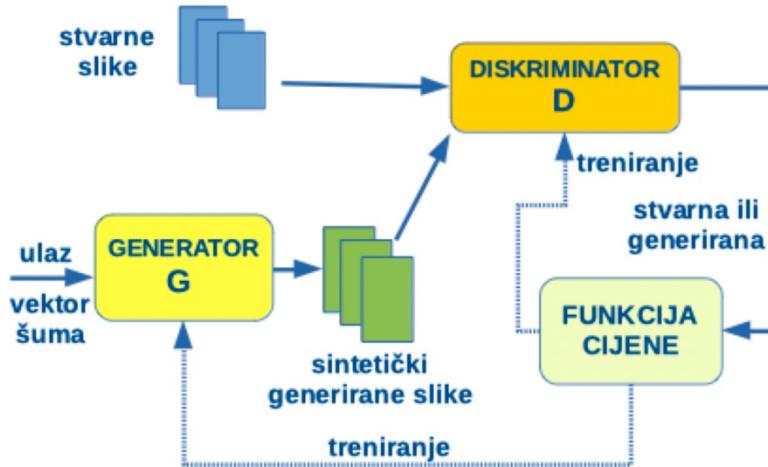
Diskriminatator (D) djeluje kao klasifikator, pokušavajući razlikovati stvarne uzorke podataka iz skupa za treniranje od lažnih uzoraka podataka koje generira generator. Njegov je cilj ispravno označiti stvarne slike kao „stvarne“ (izlaz blizu 1), a sintetičke slike kao lažne (izlaz blizu 0).

Tijekom procesa treniranja, ove se dvije mreže iterativno ažuriraju:

- Diskriminatator se trenira kako bi poboljšao svoju sposobnost razlikovanja stvarnih i sintetičkih podataka.
- Istovremeno, generator uči kako stvoriti što realističnije sintetičke podatke kako bi „prevario“ diskriminatator.

Ovaj se suparnički proces nastavlja sve dok generator ne stvori podatke toliko uvjerljive da ih diskriminatator više ne može pouzdano razlikovati od stvarnih. U tom trenutku, distribucija generiranih uzoraka je slična onoj kod stvarnih, ulaznih podataka.

Primarna je svrha GAN-ova stvaranje realističnih umjetnih podataka, što je izrazito vrijedno u prevladavanju izazova povezanih s ograničenim i neuravnootženim skupovima podataka u raznim primjenama strojnog učenja.



Slika 4.1. Proces suparničkog treniranja u GAN mreži [23]

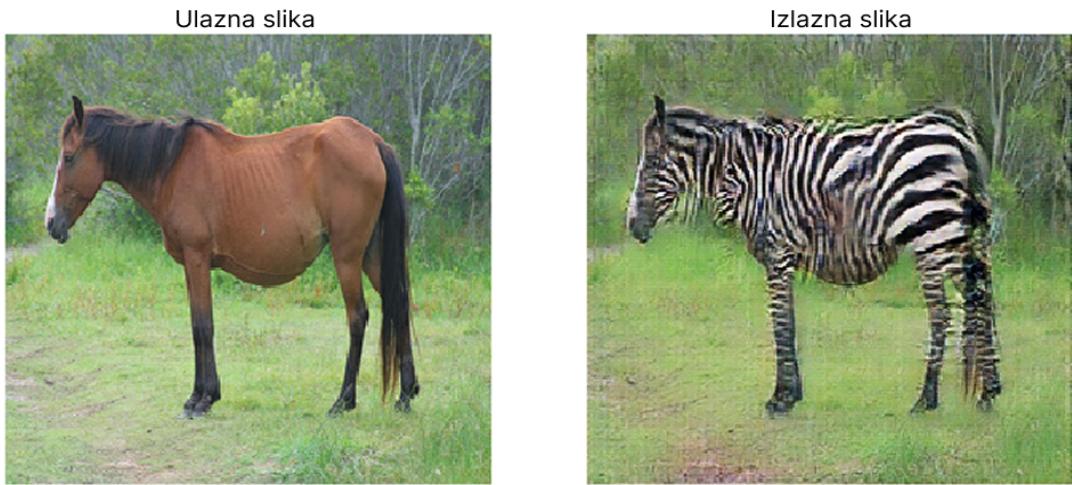
4.1. Razvoj GAN-a i ključne inovacije

Od prve pojave, GAN-ovi su se brzo razvijali, s brojnim proširenjima i poboljšanjima izvorne „vanilla“ GAN arhitekture. Ključne točke u razvoju ove tehnologije uključuju:

- **Početna verzija (2014. godine)** – izvorni GAN Goodfellowa i suradnika koristio je potpuno povezane slojeve. Iako revolucionarna, rana se GAN verzija suočavala s izazovima poput nedostatka funkcije gubitka (*engl. loss function*) povezane s kvalitetom slike i potrebe za uravnoteženjem ciklusa treniranja između dviju mreža. [11]
- **DCGAN** – uvela ga je grupa prof. Radforda i suradnika (2015. godine). Prvi su integrirali konvolucijske slojeve u GAN, značajno poboljšavši kvalitetu generiranih slika te stabilnost učenja. DCGAN se često smatra osnovnim GAN modelom i široko se primjenjuje. [7]
- **cGAN** – predložili su ga Mirza i Osindero 2014. godine. cGAN je poboljšao upravlјivost modela dodavanjem uvjetnog ulaza (npr. oznaka klase podatka ili referentnog uzorka) i generatoru i diskriminatoru. Ta novost omogućuje GAN-u generiranje slika sa specifičnim karakteristikama, što ih čini popularnim izborom za proširenje analize medicinskih slika (rješavanje neuravnoteženosti u klasama skupa podataka). [26]
- **Modeli prevodenja slike u sliku:**
 - **pix2pix** – predstavljen 2017. godine (Isola i suradnici), pix2pix je revolucionirao prevodenje uparenih slika iz jedne domene u drugu (*engl. image-to-image*).

to-image translation) primajući cijele slike kao ulaz u generator, umjesto slučajnog šuma. [25]

- **CycleGAN** – predstavljen iste godine kao i pix2pix (Zhu i suradnici), CycleGAN je otišao i korak dalje. Omogućio je prevođenje nesparenih slika između različitih domena koristeći koncepte kao što su konzistentnost ciklusa (*engl. cycle consistency*) i gubitke identiteta (*engl. identity losses*). Ova arhitektura je vrlo svestrana i poznati je izbor za mnoge primjene, posebno u rukovanju s multimodalnim podatcima. [27]



Slika 4.2. Primjer rezultata dobivenog korištenjem CycleGAN modela [24]

- **WGAN** – uvela ga je grupa pod vodstvom Arjovskyja 2017. godine. WGAN i njegova poboljšana varijanta, WGAN-GP (Gulrajani i suradnici, 2017. godine), usredotočili su se na stabilizaciju treniranja GAN modela korištenjem Wasserstein-ove udaljenosti kao funkcije gubitka, rješavajući probleme nestabilnosti i eventualne pojave moda kolapsa (*engl. mode collapse*). [25]
- **PGGAN** – predstavljen iste godine kao i WGAN varijanta (Karras i suradnici), PGGAN je omogućio generiranje visokokvalitetnih slika progresivnim povećavanjem rezolucije te dodavanjem slojeva generatoru i diskriminatoru tijekom treniranja modela. [26]
- **StyleGAN** (Karras i suradnici, 2019. godine) - nadovezujući se na PGGAN, StyleGAN uvodi inovacije u arhitekturi modela za najsuvremeniju sintezu slika. Posebno je dobar u generiranju realističnih ljudskih lica jer uči veze između različitih značajki. Ta sposobnost omogućuje korisniku mijenjanje željenih značajki na slici unoseći sitne

promjene u dijelove vektora šuma. zbog toga je poboljšao istraživanje i razumijevanje latentnog prostora. [25]



Slika 4.3. Primjer promjene značajki na slikama ljudskog lica pomoću StyleGAN-a [28]

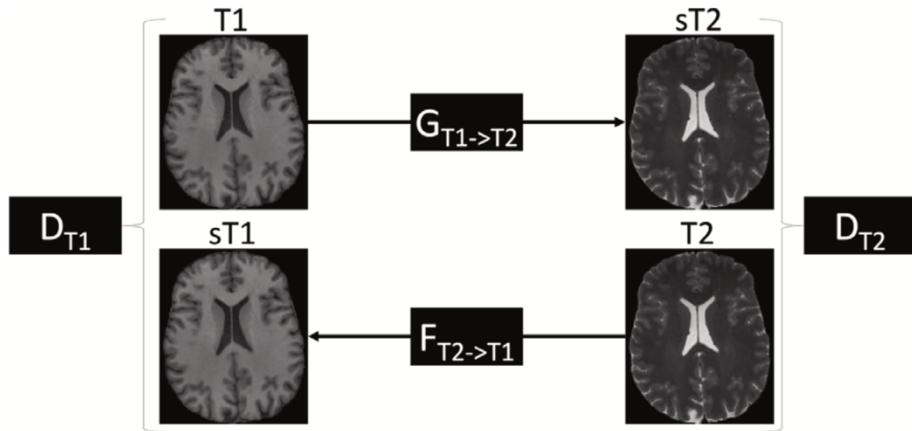
4.2. Svrha i upotreba GAN-a u medicini

GAN su modeli privukli značajnu pozornost u području medicinskog snimanja. Omogućili su prevladavanje uobičajenih izazova poput ograničenih i neuravnoteženih skupova podataka, koji su česti u kontekstu medicine. Njihove primarne primjene u medicini uključuju:

- **Umjetno povećavanje broja podataka** (*engl. Data Augmentation*) – ovo je jedna od najčešćih upotreba GAN-ova u medicini. Generiranjem više slika, GAN-ovi mogu:
 - povećati veličinu skupova podataka, što je ključno za modele dubokog učenja koji zahtijevaju veći broj podatka za treniranje [27]
 - donijeti ravnotežu kod neuravnoteženih skupova podataka, posebno onih s manjim ili rijetkim klasama. Neke od primjena gdje se navedeno uspješno koristilo: klasifikacija različitih vrsta tumora i njihovih stanja (npr. otkrivanje metastatskog raka jetre, klasifikacija raka kože), segmentacija područja od interesa i anomalija (npr. segmentacija plućnih čvorova, segmentacija srčanog ultrazvuka, segmentacija melanoma) [26]
 - riješiti probleme neslaganja podataka za treniranje s distribucijom podataka s kojima se model susreće u stvarnim primjenama. Uklonivši moguću pristranost u ulaznim podatcima te mane zbog načina uzorkovanja i odabira podataka,

omogućuje modelima dubokog učenja bolju generalizaciju na novom setu uzoraka iz ciljane domene.

- očuvati privatnost pacijenta i pružiti anonimnost procesu korištenjem sintetičkih slika, a ne stvarnih slika pacijenata ili ispitanika
- **Prijevod slike u sliku** – osim mogućnosti generiranja, GAN modeli mogu transformirati slike iz jednog modaliteta, domene ili stila u drugi. To uključuje:
 - pretvaranje MR snimaka u CT snimke te obrnuto,
 - prijevod MR slike iz jednog modaliteta u drugi (T1, T2, FLAIR),
 - stvaranje segmentacijskih slika iz neobrađenih snimki (*engl. raw scans*) [27]



Slika 4.4. Ilustracija prijevoda između modaliteta T1 i T2 MR snimki mozga [27]

- **Povećanje rezolucije slika** (*engl. Super-Resolution*) – sposobnost generiranja medicinskih slika visoke rezolucije iz manje kvalitetnih snimki (niska rezolucija, stariji primjeri snimki itd.). Navedena korist je izrazito vrijedna s obzirom na složenost te visoku cijenu prikupljanja kvalitetnih medicinskih podataka.



Slika 4.5. Primjer povećanja rezolucije MR snimke mozga [27]

- **Uklanjanje šuma i artefakata** – primjena kod uklanjanja šuma i ispravljanja smetnji kod medicinskih snimaka, nastalih zbog mnogih vanjskih faktora (greške prilikom snimanja, korištenja starijih uređaja manje kvalitete, pomaka pacijenta, prisustva stranog tijela itd.)
- **Planiranje i praćenje liječenja** – GAN modeli mogu pomoći i u automatizaciji procesa planiranja terapije kod oboljelih pacijenata (kirurški zahvati, zračenje, eksperimentalni lijekovi itd.). Jedna od primjena je simuliranje scenarija napredovanja bolesti, popust rasta glioma, vrste tumora u mozgu. [25]

4.3. Metrike evaluacije za generativne kontradiktorne mreže

Evaluacija performansi GAN modela predstavlja složen i neriješen izazov u području generativnog modeliranja. Za razliku od tradicionalnih modela dubokog učenja koji nastoje postići konvergenciju s minimalnom vrijednošću funkcije gubitka, GAN-ovi uključuju složen proces suparničkog treniranja između generatorske i diskriminatorske mreže. Navedeni princip rada zahtijeva posebne metrike za procjenu kvalitete i raznolikosti generiranih uzoraka, ključnih aspekata za određivanja performansi GAN mreže.

4.3.1. Fréchet Inception udaljenost

Fréchet Inception udaljenost pojavila se kao jedna od najčešće korištenih metrika za procjenu GAN-ova. Bazira se na određivanju sličnosti između distribucija generiranih te stvarnih slika uspoređujući prikazane značajke. Ekstrakciju značajki sa slika vrši prethodno trenirana neuronska mreža (*engl. pre-trained neural network*). Posebno je vrijedna metrika jer se dobro slaže s ljudskom percepcijom kvalitete slike, a istovremeno je osjetljiva i na probleme kolapsa načina rada te raznolikosti uzoraka. [29]

Rad FID metrike odvija se u 3 koraka:

1. izdvajanje značajki iz stvarnih i generiranih uzoraka s pomoću prethodno istrenirane InceptionV3 mreže (obično se odvija na pretposljednjem sloju neuronske mreže)
2. modeliranje značajki kao normalne (Gaussove) distribucije sa srednjom vrijednošću μ i kovarijancom σ za stvarne (r) i generirane slike (g) [30]
3. izračun Fréchet-ove udaljenosti između promatranih distribucija s pomoću formule:

$$FID = \|\mu_r - \mu_g\|^2 - Tr\left(\Sigma_r + \Sigma_g - 2\sqrt{\Sigma_r \Sigma_g}\right) \quad (4.1)$$

gdje je:

- μ_r i μ_g – srednja vrijednost distribucije značajki sa stvarnih, odnosno generiranih slika
- Σ_r i Σ_g – matrica kovarijance distribucije značajki sa stvarnih, odnosno generiranih slika
- $Tr()$ – trag matrice (*engl. trace operation*) – operacija koja zbraja elemente duž glavne dijagonale matrice [30]

Niže vrijednosti FID-a ukazuju na povećanu sličnost distribucija generiranih i stvarnih uzoraka, što sugerira bolje GAN performanse.

Unatoč svojoj popularnosti, FID ima nekoliko ograničenja: pretpostavlja da stvarne značajke slijede normalnu distribuciju (što ne mora uvijek biti slučaj), za stabilnu procjenu potreban je veliki broj uzoraka (barem 10 000) te ovisi o prethodno treniranoj mreži (obično istreniranoj na ImageNet setu podataka) koja možda nije optimalna za sve domene.

4.3.2. Maksimalno odstupanje srednjih vrijednosti

Maksimalno odstupanje srednjih vrijednosti (*engl. Maximum Mean Discrepancy - MMD*) – predstavlja još jedan način izračuna udaljenosti između distribucija. Za razliku od FID vrijednosti, MMD ne pretpostavlja temeljni oblik distribucije, nudeći fleksibilniju metriku evaluacije. Istiće se u otkrivanju suptilnih razlika između distribucija, a osjetljivost pokazuje u scenarijima kolapsa načina rada. [31]

MMD funkcioniра na sljedeći način:

- mapira uzorke iz obje distribucije u Hilbertov prostor reproducirajućih jezgri (*engl. Reproducing Kernel Hilbert Space - RKHS*) korištenjem jezgrene funkcije (obično Gaussova ili linearna jezgra)
- mjeri udaljenost između srednjih vrijednosti distribucija u visokodimenzionalnom prostoru značajki [32]

- MMD rezultat računa se kao:

$$MMD^2 = \mathbb{E}[k(x, x')] + \mathbb{E}[k(y, y')] - 2\mathbb{E}[k(x, y)] \quad (4.2)$$

gdje je:

- k – jezgrena funkcija
- x i x' – uzorci iz stvarne distribucije
- y i y' – uzorci iz distribucije generiranih slika
- $\mathbb{E}[]$ – očekivanje, očekivana vrijednost

Kao i kod FID-a, niža vrijednost MMD vrijednosti ukazuje na zanemarive razlike u distribucijama te ispravan rad generativnog modela. [32]

4.3.3. Test 1-najbližeg susjeda

Test 1-najbližeg susjeda (*engl. 1-Nearest Neighbor - 1-NN*) nudi jednostavan, ali snažan pristup procjeni rada GAN modela. Izravno procjenjuje jesu li stvarni te generirani uzorci različiti. Istiće se u otkrivanju prekomjerne prilagodbe modela (kada model pamti uzorce korištene za treniranje) te pruža intuitivnu mjeru preklapanja razmatranih distribucija. Test 1-NN je posebno vrijedan alat jer daje minimalne prepostavke o distribuciji podataka te primjenu pronalazi u mnogim domenama. [31]

Ukratko, princip testa 1-NN bazira se na:

- kombiniranju uzoraka iz stvarne i generirane distribucije u jedan set podataka s odgovarajućim oznakama
- za svaki uzorak pronalazi se njegov najbliži susjed i provjerava se dolazi li iz iste distribucije
- računa se točnost putem unakrsne provjere „izdvoji jednog“ (*engl. Leave-one-out Cross Validation - LOOCV*) na novom setu podataka [31]

U idealnom scenariju gdje se distribucije savršeno preklapaju, LOOCV točnost trebala bi biti približno 0,5, što ukazuje da se stvarni i generirani uzorci ne razlikuju. Točnost znatno iznad 0,5 sugerira veliku razliku u distribucijama (loše generiranje uzorka), dok točnost ispod 0,5 može ukazati na prekomjerno prilagođavanje modela ulaznim podatcima. [31]

4.3.4. Indeks kreativnosti - nasljeđivanja - raznolikosti

Indeks kreativnosti - nasljeđivanja - raznolikosti (*engl. Creativity - Inheritance - Diversity Index – CID*) predstavlja nešto drugčiji pristup GAN evaluaciji izravnom analizom generiranih slika prema tri temeljna aspekta idealnog procesa generiranja uzorka. Za razliku od metrika koje se oslanjanju na prethodno istrenirane klasifikatore, CID pruža eksplicitna mjerena o tome jesu li generirani uzorci novi (nisu duplicitne inačice podataka za treniranje), održavaju li stil i značajke ciljane distribucije te pokazuju li dovoljno međusobnu raznolikost. [33]

CID se indeks sastoji od tri komponente:

- **Kreativnost** – mjeri minimalnu udaljenost između generiranih i stvarnih uzorka, kako bi se osiguralo da GAN ne pamti samo podatke za treniranje. Viši rezultat kreativnosti ukazuje na originalnost generiranih slika.
- **Nasljeđivanje** – procjenjuje koliko dobro generirani uzorci održavaju ključne karakteristike ciljane distribucije, obično koristeći metrike sličnosti poput kosinusne ili strukturne sličnosti (*engl. Structural similarity index measure - SSIM*). Više vrijednosti nasljeđivanja ukazuju bolje očuvanje značajki koje su specifične za domenu.
- **Raznolikost** – procjenjuje varijacije među generiranim uzorcima mjerenoj udaljenosti unutar generiranog skupa. Viši rezultati ukazuju na manje ponavljanja i smanjenu vjerojatnost za kolaps moda. [33]

Ukupni CID indeks kombinira sve tri komponente, pružajući sveobuhvatnu mjeru performansi GAN-a koja uravnotežuje originalnost, vjerodostojnost i raznolikost generiranih uzorka. [33]

4.3.5. Preciznost i osjetljivost

Preciznost i osjetljivost za distribucije - već otprije spomenuti, u kontekstu modela dubokog učenja, ove poznate i jednostavne metrike primjenu pronalaze i u domeni generativnog modeliranja. nude intuitivan okvir za razdvajanje aspekata kvalitete i raznolikosti uzorka.

Preciznost određuje koliko je generirane distribucije sadržano unutar stvarne (kvaliteta), dok **osjetljivost** mjeri koliko je stvarne distribucije obuhvaćeno generiranom (pokrivenost, raznolikost). [34]

Moderne metrike preciznosti i osjetljivosti GAN modela obično:

- izdvajaju značajke iz stvarnih i generiranih uzoraka s pomoću prethodno istrenirane mreže (klasifikatora)
- procjenjuje razinu preklapanja dviju distribucija korištenjem metode najbližeg susjeda
- računa preciznost kao udio generiranih uzoraka koji bi se mogli smatrati dijelom stvarne distribucije
- računa osjetljivost kao udio stvarnih uzoraka koji je pokriven distribucijom generiranih inačica

Ovaj pristup pruža odvojene rezultate omogućujući finiju procjenu performansi te jasniju identifikaciju problema poput kolapsa načina rada (niska osjetljivost) ili loše kvalitete uzoraka (niska preciznost). [34]

4.3.6. Vizualni Turingov test

Vizualni Turingov test (*engl. Visual Turing Test - VTT*) ili **vizualna procjena** – smatra se najpouzdanijom metodom za određivanje vjerodostojnosti generiranih podataka. Klinički stručnjaci zamole se da klasificiraju slike kao stvarne ili lažne. Rezultat blizu 50% točnih odgovora sugerira da stručnjaci ne mogu razlikovati stvarne i sintetičke podatke. Iako je proces vrlo učinkovit, dugotrajan je, skup i podložan varijacijama između promatrača. U jednoj studiji, medicinski su stručnjaci postigli svega 60% prosječne točnosti u klasifikaciji MR snimki generiranih StyleGAN modelom kao stvarnih ili lažnih, što ukazuje na vizualnu točnost generiranih slika. [7]

4.4. Izazovi i ograničenja GAN modela

Unatoč svim svojim mogućnostima, GAN se modeli suočavaju s nekoliko izazova i ograničenja, posebno u zahtjevnoj domeni medicinskog snimanja:

- **problem nestabilnosti i konvergencije pri treniranju** – modeli znaju postati nestabilni tijekom treniranja, varirati u točnosti iz epohe u epohu te propasti u procesu konvergencije. Proces učenja u suparničkom sustavu, koji uključuje dvije konkurentske mreže, generator i diskriminatory, može biti teško uravnotežiti.

- **kolaps načina rada** – pojava koja dovodi do toga da generator nauči proizvoditi samo uzorke ograničene raznolikosti (istraži i poboljša samo mali dio latentnog prostora) koji mogu prevariti diskriminator. Navedeno dovodi do gubitka raznolikosti generiranih podataka (*engl. diversity, variety*), čime se dobiju slike visoke kvalitete, no uvjek istih značajki. To može ozbiljno našteti performansama modela dubokog učenja koji bi se trenirali s ovim podatcima.
- **zahtjevi za računalnim resursima** – generiranje visokorezolucijskih ili trodimenzionalnih medicinskih podataka, kao i treniranje takvih modela, zahtijeva značajnu računalnu snagu te korištenje hardverske i softverske podrške najnovije generacije.
- **vjerodostojnost generiranih slika** – GAN modeli mogu stvoriti vizualno realistične slike koje bi, u nekim slučajevima, mogle prevariti i stručnjake u vizualnom Turingovom testu. No, studije pokazuju da nisu uvjek sposobni reproducirati puninu medicinskih skupova podataka ili prenijeti suptilne informacije potrebne za ispravnu dijagnostiku. Vjerodostojnost i klinička upotrebljivost generiranih slika ostaju primarna briga.
- **halucinacije značajki** – uvjetni GAN-ovi mogu nenamjerno prikazati nepostojeće ili nerealne artefakte na slikama pacijenata, što predstavlja značajan rizik ako se koristi u kliničkim uvjetima. Potrebna su daljnja istraživanja kako bi se ove halucinacije pouzdano izbjegle te otkrile.
- **nedostatak opće prihvaćenih metrika za evaluaciju** – vizualni je pregled stručnjaka još uvjek često primarni alat za evaluaciju. Pored toga, koriste se metrike poput FID-a te IS-a (*engl. Inception Score*), no njihova primjenjivost u medicinskoj domeni zahtjeva daljnje istraživanje.
- **etička i sigurnosna pitanja** – sposobnost GAN-a da stvori vrlo realistične sintetičke slike izaziva zabrinutost zbog mogućnosti prijevara ili zlonamjernog mijenjanja medicinskih slika (dodavanje ili uklanjanje anomalija na slikama, narušavanje kvalitete, lažiranje podataka u istraživanjima itd.). Postoji i zabrinutost oko privatnosti zbog postojanja mogućnosti rekonstrukcije stvarnih podataka ispitanika (korištenih za treniranje modela) iz generiranih.

5. POSTAVKE EKSPERIMENTA

Ovo poglavlje detaljno opisuje eksperimentalne postavke za ostvarenu sintezu medicinskih slika, opisujući korišteni skup podataka, arhitekturu korištene generativne kontradiktorne mreže, ključne parametre treniranja i druge značajne funkcije bitne za razumijevanje metodologije zadatka.

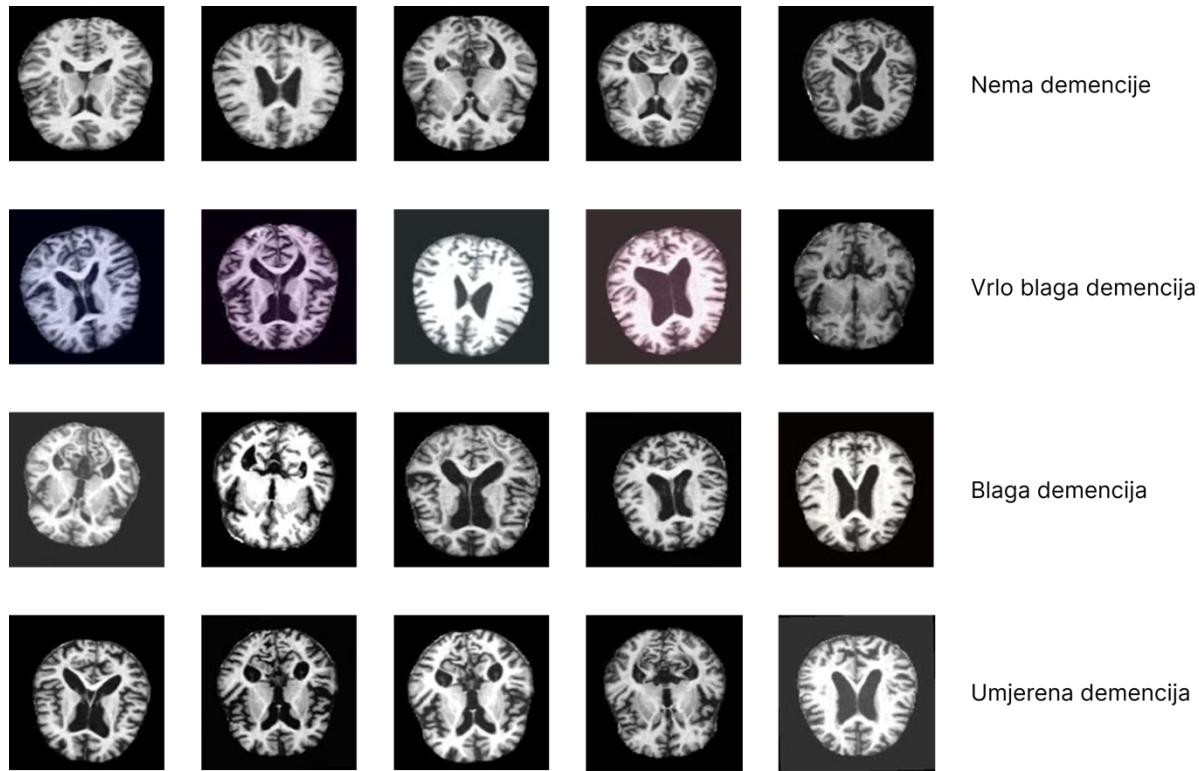
5.1. Opis ulaznih podataka

Skup podataka OASIS (*engl. Open Access Series of Imaging Studies*) značajan je i javno dostupan resurs u istraživanju neurološkog zdravlja, prvenstveno usmjeren na učinke zdravog starenja i Alzheimerove bolesti (AD). U poleđini svega nalazi se program koji je olakšao globalnu suradnju među brojnim istraživačkim centrima te dao značajan doprinos razumijevanju različitih neurorazvojnih stanja. Ranije inačice, *OASIS-Cross-sectional* i *OASIS-Longitudinal*, široko su citirane i korištene za razvoj neuroloških alata poput atlasa mozga i algoritama za segmentaciju tkiva, kao i za stvaranje modela zdravog starenja i progresije AD bolesti. [35]

Skup podataka OASIS-3 predstavlja najveće proširenje, prikupljajući MR i PET (*engl. positron emission tomography*) snimke, s povezanim kliničkim informacijama, od 1098 ispitanika. Ova se raznolika skupina sastoji od pacijenata između 42 i 95 godina starosti te uključuje i kognitivno zdrave osobe kao i one u različitim fazama kognitivnog opadanja. Ovaj set uzorka nudi preko 2000 MR snimanja, od kojih se svako sastoji od više modaliteta (T1, T2, FLAIR itd.), pružajući istraživačima mogućnost ispitivanja longitudinalnih promjena u različitim tipovima tkiva. Uključeno je i gotovo 1500 PET metaboličkih snimki s pripadnim kliničkim procjenama. MR snimke su prikupljene na različitim modelima medicinskih uređaja, uključujući one jakosti 1,5 T i 3 T. [36]

Za potrebe ovoga rada, korišteni set podataka dohvaćen je s Kaggle repozitorija [37]. Radi se o prilagođenom skupu uzorka koji sadrži samo dio opisanog OASIS seta, i to približno 44 000 MR slika koje su podijeljene u četiri različite klase na temelju stadija demencije. Navedeni je repozitorij namijenjen korištenju u treniranju i testiranju modela dubokog učenja.

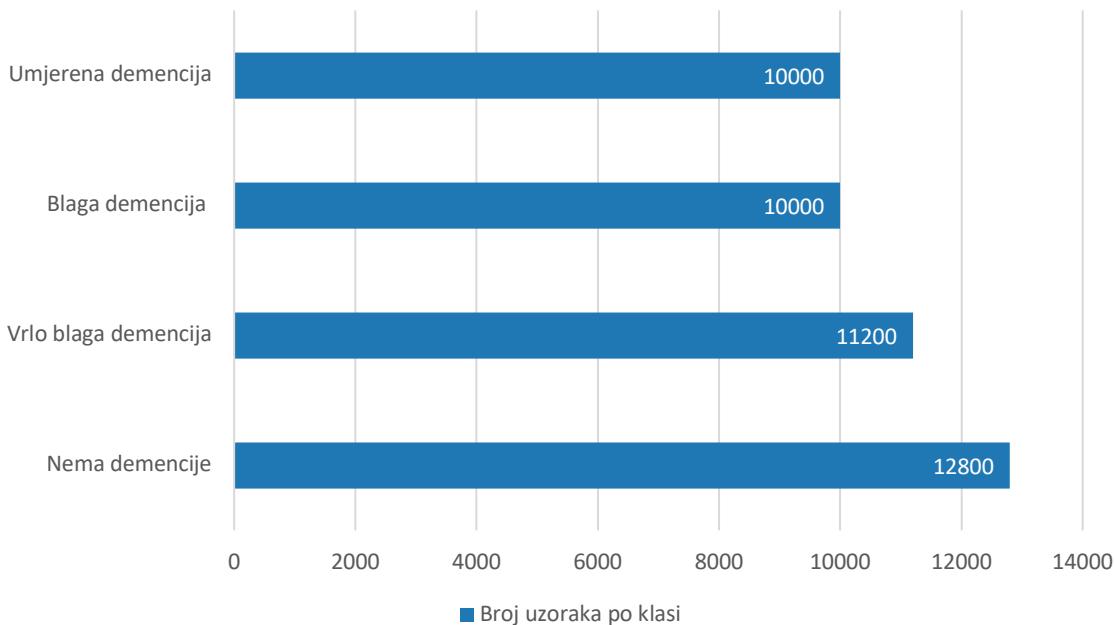
Sve slike su obrađene, sve smetnje uklonjene te je prikazano samo moždano tkivo (sa snimki je uklonjen prikaz lubanje te okolnog tkiva).



Slika 5.1. Primjeri uzoraka iz korištenog skupa podataka po klasama [37]

Skup podataka organiziran je u četiri direktorija, od kojih svaki predstavlja različitu klasu stanja:

- **Nema demencije** (*engl. No Dementia*) – sadrži 12 800 MR snimki ispitanika bez znakova demencije
- **Vrlo blaga demencija** (*engl. Very Mild Dementia*) – snimke mozga s ranim znakovima AD bolesti, vrlo blagi simptomi (11 200 snimki)
- **Blaga demencija** (*engl. Mild Dementia*) – 10 000 MR slika s jasnim znakovima demencije
- **Umjerena demencija** (*engl. Moderate Dementia*) – prikazuje uznapredovali stadij AD bolesti, izraženiji simptomi (10 000 slika)



Slika 5.2. Prikaz distribucije korištenog skupa podataka

Sve slike (44 000 uzoraka) sadržane su u JPG (*engl. Joint Photographic Experts Group*) formatu. Inače se snimke magnetske rezonancije spremaju u DICOM (sadrži i slike i metapodatke, za uporabu u bolnicama i centrima za snimanje) ili NIfTI formatu (uobičajeno korišteno u istraživanjima, nije standardno za kliničku uporabu) koje omogućuju rekonstrukciju snimki u trodimenzionalnom obliku te precizno lociranje i proučavanje određenih metaboličkih stanja.

Za potrebe ovoga rada te fokusiranja na jedan djelić problema u medicinskoj dijagnostici, navedene prednosti izvornih formata nisu bitne te je iz tog razloga odabran navedeni skup podataka, koji slike sprema u kompresijskom JPG formatu. Odabrani je format slika pogodan za osnovne klasifikacijske zadatke te znatno olakšava izvođenje programa na ograničenim računalnim resursima. Potrebno je naglasiti kako je u procesu dodatnog rasterećenja dostupnih resursa te eksperimentiranja u izradi i izvođenju generativnih modela, ukupan broj korištenih slika smanjen na 12 000. Tako se svaka od četiri klase sastoji od 3000 slika, što se pokazalo kao optimalna brojka za postizanje zadovoljavajuće kvalitete rezultata, uz razumne vremenske i računalne troškove.

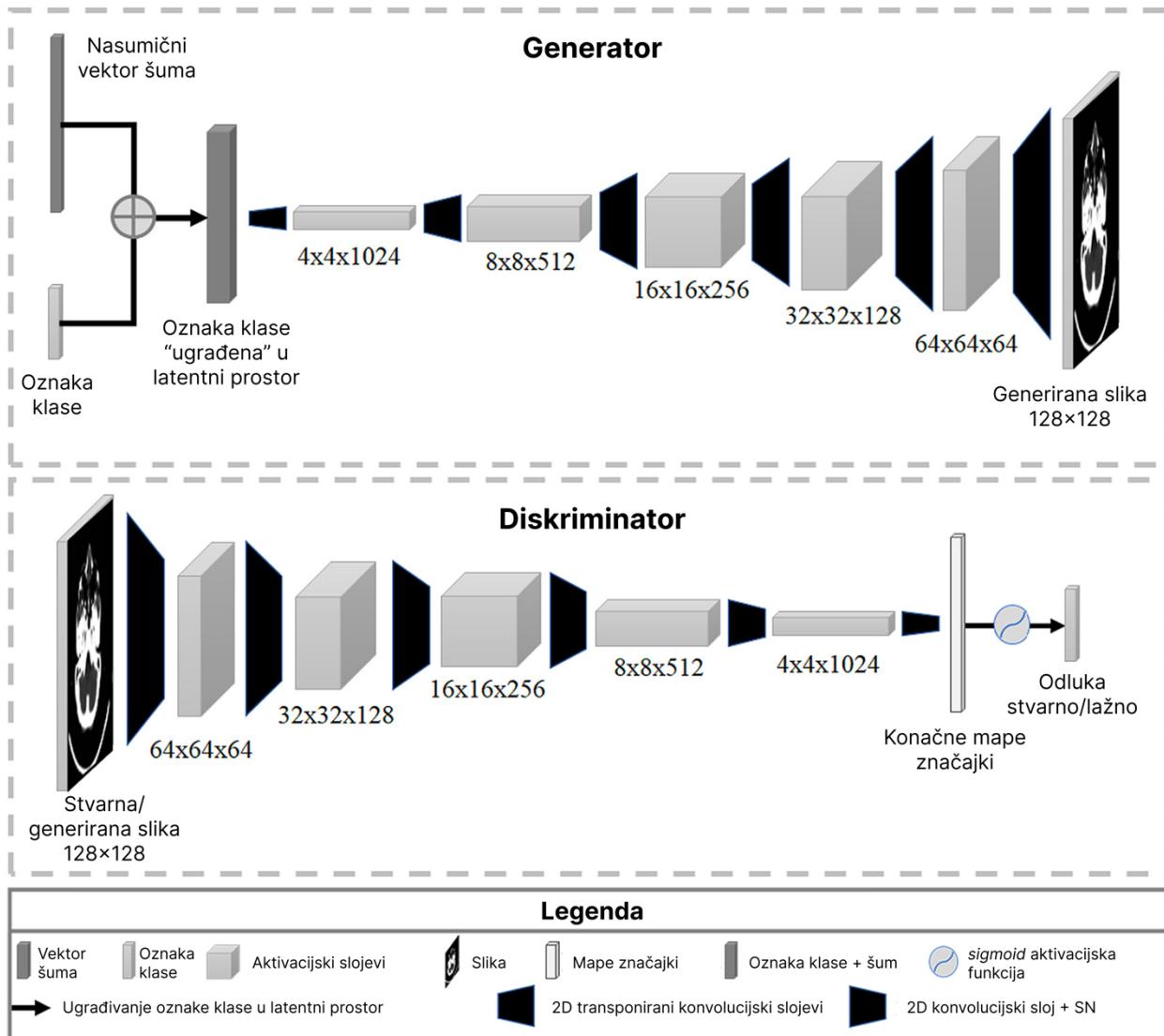
5.2. Dizajn generatora i diskriminatora u odabranoj GAN mreži

Generator u predloženoj arhitekturi dizajniran je za sintezu realističnih RGB slika veličine 128 x 128 piksela koje su uvjetovane oznakama klase. Takav princip slijedi paradigmu rada cDCGAN-a. Na samom početku prihvata dva parametra: vektor nasumičnog šuma, dimenzije 256 elemenata, te oznake klase. Oznaka se prvo „ugrađuje“ u gusti vektorski prostor koji odgovara dimenziji vektora šuma, a zatim se množi s, već spomenutim, vektorom šuma, učinkovito spajajući informaciju o klasi uzorka u latentni prostor. Ovaj će se spojeni prikaz projicirati putem potpuno povezanog sloja i preoblikovati u prostornu mapu niže rezolucije.

Generator koristi niz transponiranih konvolucijskih slojeva (također poznatih kao dekonvolucije), nakon čega slijede slojevi normalizacije i *ReLU* aktivacijska funkcija s propuštanjem (*Leaky ReLU*), kako bi se postupno povećalo uzorkovanje te poboljšale mape značajki (*engl. feature maps*). Mreža proširuje prostorne dimenzije s 16 x 16 piksela na 128 x 128, istovremeno smanjujući broj kanala. Mreža završava konačnim konvolucijskim slojem s *tanh* aktivacijskom funkcijom koja daje sliku s tri kanala (RGB). Ovakva arhitektura osigurava da i šum i informacije o klasi vode proces sinteze slika. Na taj je način osigurana raznolikost i vjernost u generiranim uzorcima.

Diskriminator je izrađen za robusno razlikovanje stvarnih i generiranih slika, također uvjetovanih oznakama klase. Kao ulaze prima sliku te oznaku klase. Oznaka se „ugrađuje“ te preoblikuje kako bi odgovarala dimenzijama slike, a zatim se spaja sa slikom duž osi kanala. Tako omogućuje mreži da od samog početka zajednički obrađuje slike i informacije o klasi.

Diskriminator koristi niz konvolucijskih slojeva, gdje je svaki „obavijen“ spektralnom normalizacijom kako bi se stabilizirao proces treniranja, što je ključno za konvergenciju modela. Svaki konvolucijski sloj prati aktivacijska funkcija *ReLU* s propuštanjem i sloj s ispuštanjem (*engl. dropout layer*), što poboljšava sposobnost generalizacije modela te odupiranje od prekomjernog prilagođavanja. Nakon nekoliko koraka smanjenja dimenzija uzorka, mape značajki se „izravnjavaju“ i prolaze kroz sloj sa *sigmoid* aktivacijom, kako bi se dobila skalarna vrijednost. Dobiveni broj ukazuje na vjerojatnost da je ulazna slika stvarna u kontekstu promatrane klase. Ovaj dizajn osigurava da diskriminator procjenjuje realnost slika, ali i konzistentnost s obzirom na značajke klase, pružajući tako snažan nadzor generatoru u procesu treniranja.



Slika 5.3. Prikaz arhitektura generatora i diskriminatora u korištenom rješenju [6]

5.3. Odabir postavki za treniranje GAN mreže

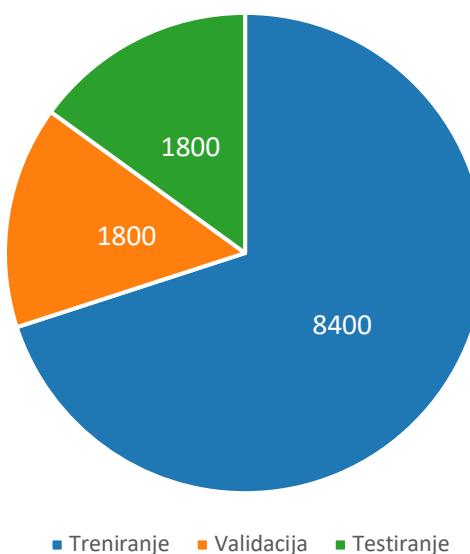
Odabir hiperparametara u ovom projektu vođen je empirijski najboljim praksama za cGAN-ove te specifičnim zahtjevima generiranja 128 x 128 RGB slika. Testirane su različite kombinacije hiperparametara iz odabrane literature i radova kako bi se pronašao najbolji pristup za razmatranu problematiku. Dimenzija latentnog vektora postavljena je na 256, što generatoru osigurava dovoljno prostora za otkrivanje različitih značajki slike. Model je treniran s veličinom grupe (*engl. batch size*) od 32, čime se postiže ravnoteža između stabilnosti treniranja i učinkovitog korištenja GPU memorije. Stopa je učenja (*engl. learning rate*) bila fiksirana na $1e^{-4}$ i za generator i diskriminatore, s Adam optimizatorom konfiguriranim s $\beta_1 = 0,5$ i malim padom težine od 1×10^{-6} kako bi se potaknula generalizacija. Broj je epoha bio postavljen na 50

za početno eksperimentiranje, a u konačnom izvođenju programa korišteno ih je 500. Mnogi izvori sugeriraju kako bi optimalna konvergencija modela mogla zahtijevati 500 – 1000 epoha, ovisno o složenosti skupa podataka i traženoj vjernosti slika.

Tablica 5.1. Pregled korištenih hiperparametara modela

Hiperparametar	Vrijednost
Veličina grupe	32
Optimizator	Adam ($\beta_1 = 0,5$ i pad težine od 1×10^{-6})
Stopa učenja	1×10^{-4}
Broj epoha	500
Dimenzija latentnog vektora	256

Predobrada ulaznih podataka provedena je postupkom normalizacije. Svim je slikama promijenjena veličina na 128 x 128 piksela, a vrijednosti su im skalirane na raspon [-1, 1] s pomoću *Rescaling* sloja, što je standardni pristup za GAN-ove koji koriste *tanh* aktivacijsku funkciju na izlazu generatora. Skup podataka učitan je iz lokalnog direktorija, automatski podijeljen na dijelove za treniranje, validaciju i testiranje (u omjerima 70%, 15%, 15%, redom). Oznake su klase kodirane kao cijeli brojevi za uvjetno generiranje. Ovakav princip osigurava da i generator i diskriminatore primaju konzistentno skalirane ulaze, što je ključno za stabilnu dinamiku treniranja.



Slika 5.4. Prikaz podjele korištenog skupa podataka (ukupno 12 000 slika)

Treniranje je modela provedeno u potpunosti lokalno na radnoj stanici opremljenoj Nvidia RTX 4070 Ti (12 GB VRAM-a) grafičkom karticom, 64 GB RAM-a i Intel Core i7 CPU-om 13. generacije. Cijelo je okruženje postavljeno s pomoću *Windows Subsystem for Linux* (WSL) alata [38]. Ovakvo je računalno okruženje pružilo dovoljno memorije i protoka za manje veličine grupe te slike visoke rezolucije. Prije toga, početni eksperimenti provedeni su na besplatnoj verziji Google Colaba [39] te sustavu za računarstvo u oblaku Vrančić [40]. Međutim, prelazak na lokalni hardver omogućio je dulje treniranje modela i veću fleksibilnost u dodjeli računalnih resursa.

Programsko je okruženje ponovo konfiguirirano kako bi se osigurala međusobna kompatibilnost i tražene performanse [41]. Korišteni su sljedeći alati te njihove verzije [42]:

- CUDA Toolkit: 12.5
- cuDNN SDK: 9.3
- Tensorflow: 2.19
- Python: 3.9

Svi su potrebni alati i biblioteke instalirani unutar WSL okruženja kako bi se iskoristila izvorna CUDA podrška te maksimiziralo GPU ubrzanje.

Pažljivi odabir hiperparametara, predobrada, odabrani hardver te softver pružili su robusno okruženje za treniranje uvjetnog GAN-a na podatcima slika visoke rezolucije.

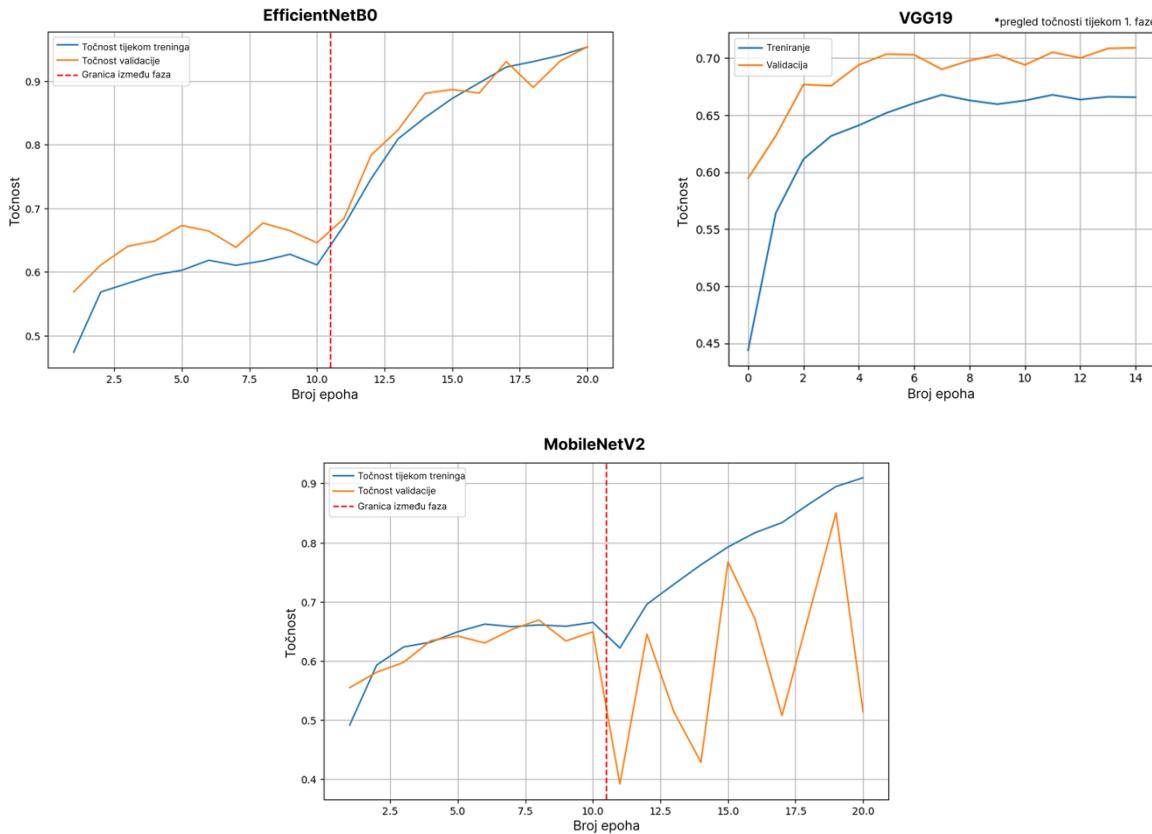
6. REZULTATI

Nakon pregleda teorijske podloge koja se krije iza ovoga projekta, u ovome će se poglavlju predstaviti rezultati provedenog eksperimenta, koji za cilj ima istražiti utjecaj korištenja GAN modela kao tehniku proširenja podataka na učinkovitost klasifikacije. U početku, cilj je bio dobiti osnovne, bazne vrijednosti, procjenjujući točnost CNN-a koji je treniran isključivo na izvornom skupu podataka. Potom će se prezentirati rezultati treniranja cDCGAN modela, koji je korišten kao tehnika proširenja podataka za generiranje dodatnih, sintetičkih slika. Konačno, procijenit će se kako uključivanje generiranih slika utječe na rezultate klasifikacije, pružajući komparativnu analizu učinkovitosti odabranog načina povećanja broja i raznolikosti slika.

6.1. Bazni rezultati korištenjem DNN mreže

U svrhu provedbe osnovne klasifikacije korišten je tzv. prijenos znanja s prethodno istreniranim konvolucijskim dubokim neuronskim mrežama. Koristili su se modeli koji su prethodni trenirani na velikim skupovima podataka, poput ImageNet-a. Ovaj pristup omoguće modelima korištenje robustnih mapa značajki naučenih na milijunima slika, koje se zatim mogu fino podesiti za specifičan zadatak, poput klasifikacije stupnja demencije iz MR snimki mozga.

Tijekom početne faze istraživanja, procijenjene su tri različite prethodno istrenirane arhitekture: EfficientNetB0, VGG19 i MobileNetV2. Svaki je model procijenjen na temelju točnosti klasifikacije, učinkovitosti učenja i računalnih zahtjeva. EfficientNetB0 se pokazao kao optimalan izbor zbog svojih superiornih performansi u svim metrikama evaluacije. Iako je VGG19 pokazao konkurentne rezultate, njegovo znatno dulje vrijeme učenja (otprilike 40 minuta po epohi) učinilo ga je nepraktičnim za iterativno eksperimentiranje. MobileNetV2, unatoč svojoj računalnoj učinkovitosti, postigao je nižu točnost u usporedbi s EfficientNetB0 te je pokazao naznake prekomjernog prilagođavanja, već nakon manjeg broja epoha. Posljedično, EfficientNetB0 odabran je kao primarni model za sljedeće eksperimente.

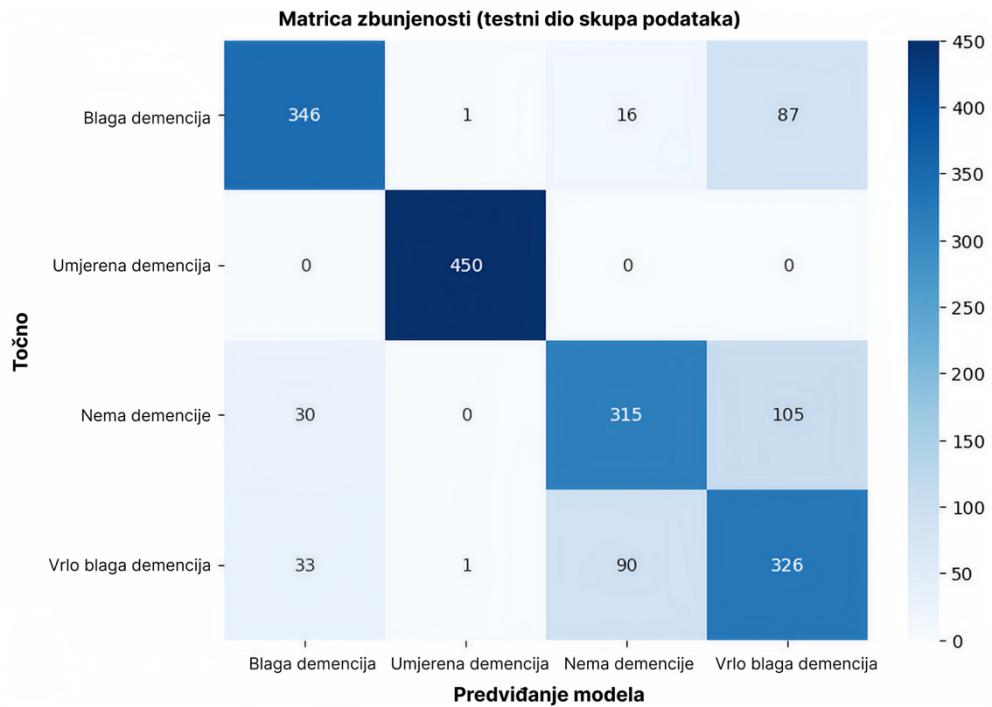


Slika 6.1. Rezultati točnosti promatranih DNN modela tijekom treninga i validacije

Proces je prijenosa znanja slijedio postupak treniranja u dvije faze: **osnovno učenje**, a zatim **fino podešavanje** (engl. *fine tuning*). U prvoj fazi, prethodno je trenirani model EfficientNetB0 prilagođen zadatku klasifikacije stadija demencije, sa zamrznutim svim konvolucijskim slojevima, omogućujući treniranje samo novododanog klasifikacijskog sloja. Početna je faza omogućila prilagođavanje gotove strukture modela specifičnim zahtjevima projekta, a istovremeno je sačuvala značajke niže razine koje je model prethodno naučio iz ImageNet-a. Druga je faza uključivala proces finog podešavanja, gdje se eksperimentiralo s različitim brojevima odmrznutih slojeva kako bi se omogućila prilagodba značajki više razine ciljanoj medicinskoj domeni.

Početno je testirano odmrzavanje samo posljednjih 5 do 10 slojeva tijekom faze finog podešavanja, na temelju uobičajene prakse postupnog prilagođavanja dubljih slojeva, uz očuvanje slojeva za izdvajanje lokalnih značajki. Međutim, ovaj je pristup dao prosječne rezultate, postižući samo 80% točnosti s relativno slabijim performansama na zahtjevnijim klasa poput *Vrlo blaga demencija* (F1-rezultat: 0,67) i *Nema demencije* (F1-rezultat: 0,72). Pristup s ograničenim odmrzavanjem slojeva zahtjevalo je znatno duže vrijeme treniranja

(preko 20 minuta) te je pokazalo znakove problema s konvergencijom (zaustavljen na 33. epohi, od planiranih 40).

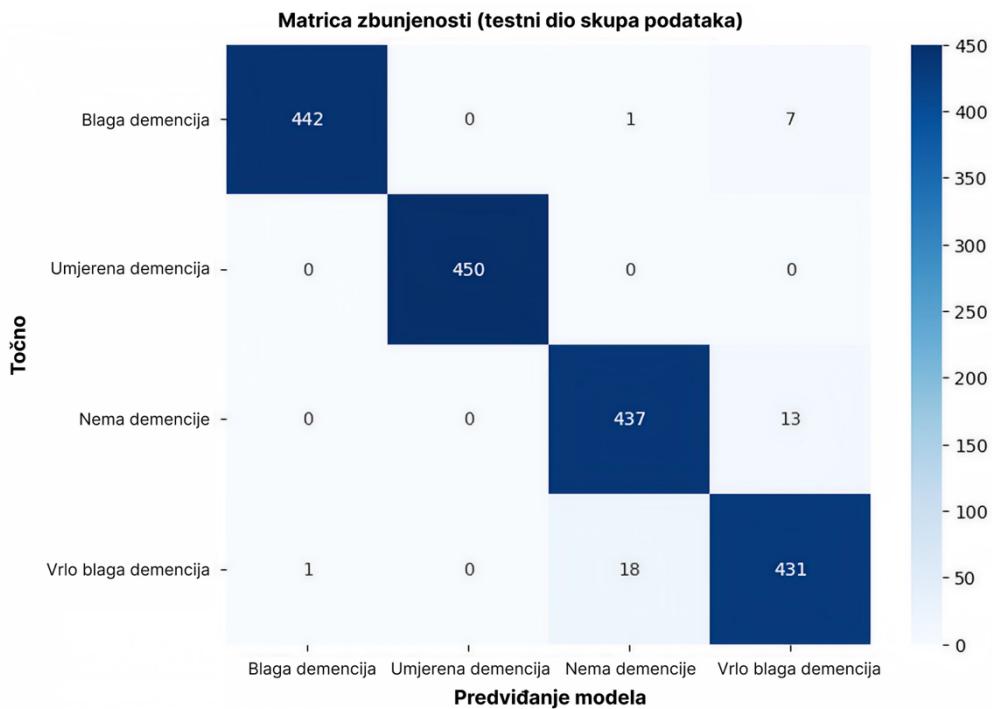


Slika 6.2. Prikaz matrice zbunjenosti EfficientNetB0 modela (odmrzavanje 10 posljednjih slojeva mreže u fazi finog podešavanja)

Nasuprot tome, odmrzavanje svih slojeva tijekom faze finog podešavanja pokazalo je znatno bolje rezultate. Ovaj je pristup postigao točnost od 97% u znatno kraćem vremenu treniranja (ukupno približno 10 minuta), sa značajnim poboljšanjima u svim klasama demencije. Tako je, klasa *Vrlo blaga demencija*, pokazala najveće poboljšanje (skok F1-rezultata s 0,67 na 0,95), a slično je primjetno i kod klase *Nema demencije* (novi F1-rezultat: 0,96). Ovo otkriće sugerira kako specifična priroda MR analize mozga za klasifikaciju stadija demencijskog poremećaja ima koristi od sveobuhvatne prilagodbe cijele mreže, odbacujući tako ustaljene pristupe kod finog podešavanja.

Tablica 6.1. Izvješće klasifikacije EfficientNetB0 modela (odmrzavanje svih slojeva mreže u fazi finog podešavanja)

Ime klase	Preciznost	Osjetljivost	F1-rezultat	Površina ispod krivulje (AUC)
<i>Nema demencije</i>	0,95833	0,97111	0,96468	0,99278
<i>Vrlo blaga demencija</i>	0,95565	0,95778	0,95671	0,99321
<i>Blaga demencija</i>	0,99774	0,98222	0,98992	0,99562
<i>Umjerena demencija</i>	1,0	1,0	1,0	1,0
Točnost		0,97793		



Slika 6.3. Prikaz matrice zbunjenosti EfficientNetB0 modela (odmrzavanje svih slojeva mreže u fazi finog podešavanja)

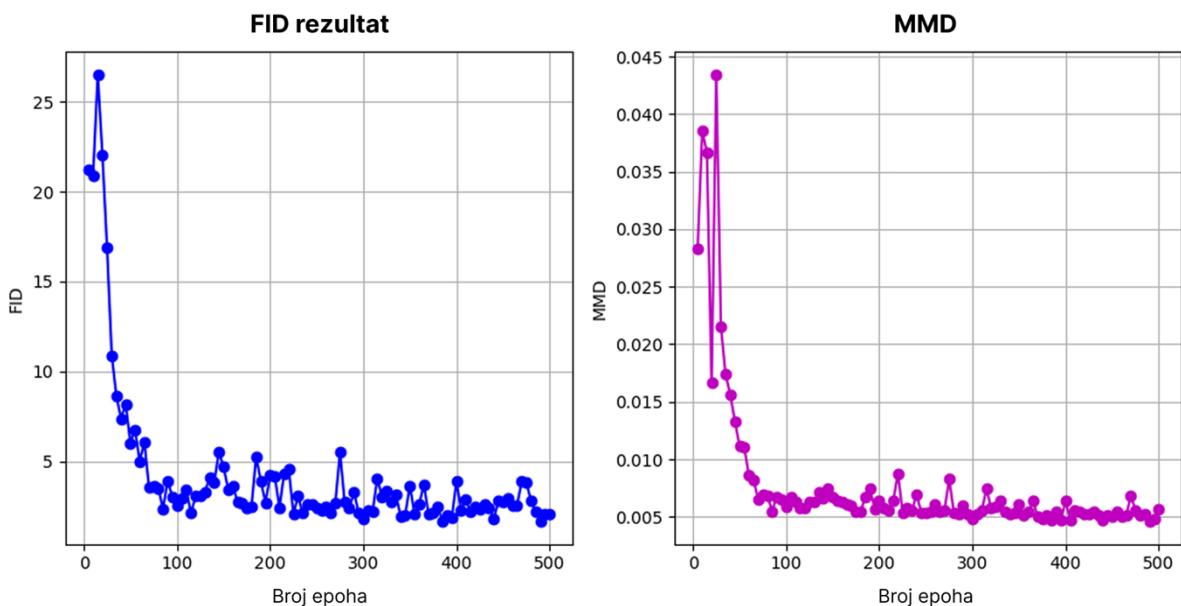
Predstavljene osnovne performanse modela predstavljaju mjerilo u odnosu na koje će se procijeniti učinkovitost proširenja skupa podataka s pomoću GAN-a, kako bi se vidjelo može li se generiranjem sintetičkih slika dodatno poboljšati točnost klasifikacije.

6.2. Rezultati treniranja DCGAN-a

Treniranje uvjetne duboke konvolucijske GAN mreže provedeno je korištenjem dvije različite rezolucije slika: 64×64 i 128×128 piksela. Obje su arhitekture trenirane tijekom 500 epoha korištenjem identičnih hiperparametara, uključujući latentnu dimenziju od 256, veličinu grupe od 32 te Adam optimizator s $\beta_1 = 0,5$. Iako su obje verzije prikazale uspješnu konvergenciju, analiza otkriva razlike u performansama koje zaslužuju detaljno ispitivanje, s posebnim naglaskom na verziju veće rezolucije (128×128 piksela) kao primarni fokus ovog projekta.

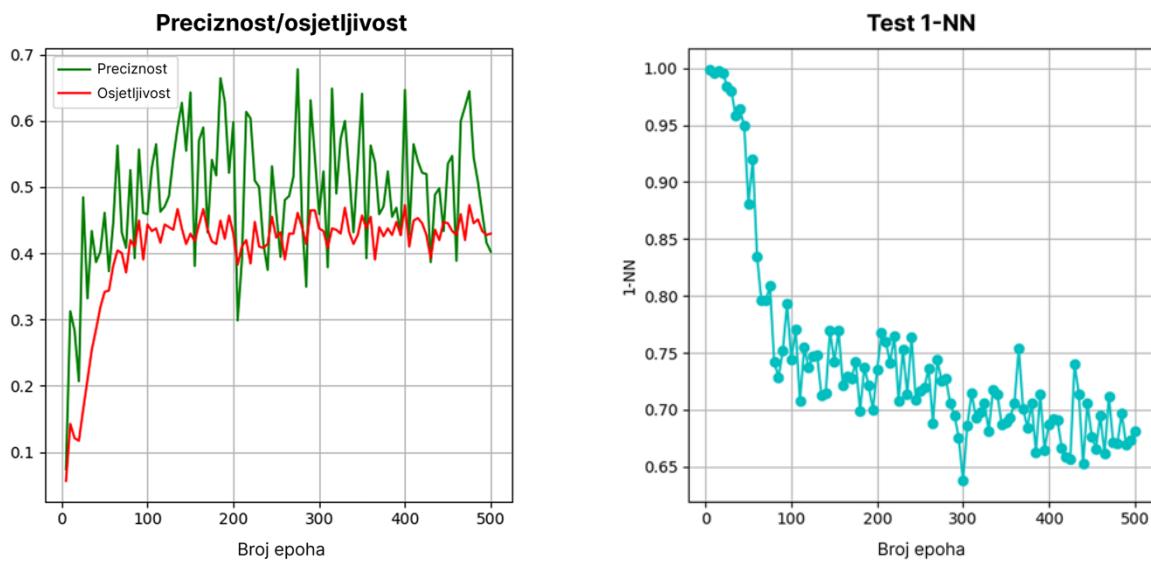
6.2.1. Analiza performansi verzije GAN-a veće rezolucije (128×128 piksela)

Ova je verzija GAN-a pokazala robusnu dinamiku treniranja u svim procijenjenim metrikama, s nekim očekivanim izazovima povezanim s generiranjem slika veće rezolucije. FID rezultat, koji mjeri kvalitetu i raznolikost generiranih slika, pokazao je izvrstan primjer konvergencije, počevši od približno 25 i brzo se smanjujući, da bi se stabilizirao oko 2 - 4 do epohe 100. Konačni FID raspon od 2 - 3 govori da su generirane slike visoke kvalitete jer niži FID rezultat ukazuje na bolju sličnost između generirane i stvarne distribucije slika. MMD metrika potvrdila je prijašnje navode, smanjujući vrijednost s početnih 0,045 na približno 0,005 - 0,010, što ukazuje na visoko podudaranje distribucija sintetičkih i stvarnih uzoraka magnetske rezonancije mozga.



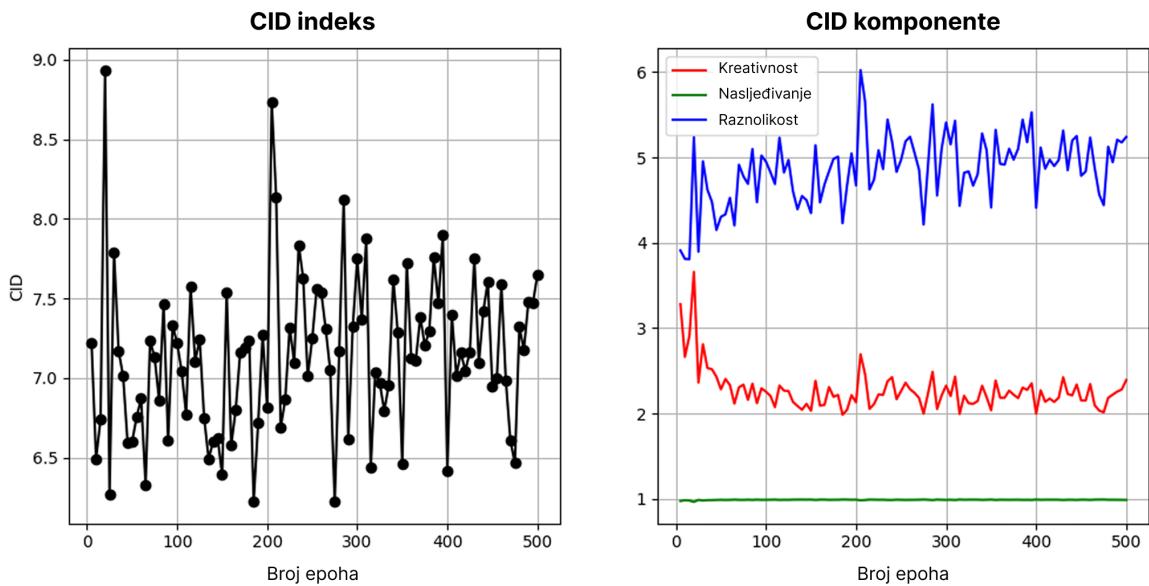
Slika 6.4. Prikaz rezultata FID i MMD metrika za odabrani GAN model

Metrike preciznosti i osjetljivosti potvrdile su sposobnost generatora da proizvede i visokokvalitetne i raznolike slike. Preciznost se stalno poboljšavala od 0,1 do približno 0,5 – 0,6, dok se osjetljivost mijenjala od 0,1 do 0,4 – 0,5, što pokazuje uspješnu ravnotežu između kvalitete i raznolikosti uzorka. Test 1-NN dosegao je raspon vrijednosti 0,65 - 0,75, gdje vrijednosti bliže 0,5 ukazuju na optimalne performanse. Premda ne postiže optimalnu vrijednost 1-NN testa, dobiveni raspon vrijednosti sugerira da diskriminator može razlikovati stvarne i generirane slike s umjerenom pouzdanošću. Navedeno ukazuje na prostor za daljnja poboljšanja, no pokazuje i značajan napredak u suparničkom treniranju dviju mreža.

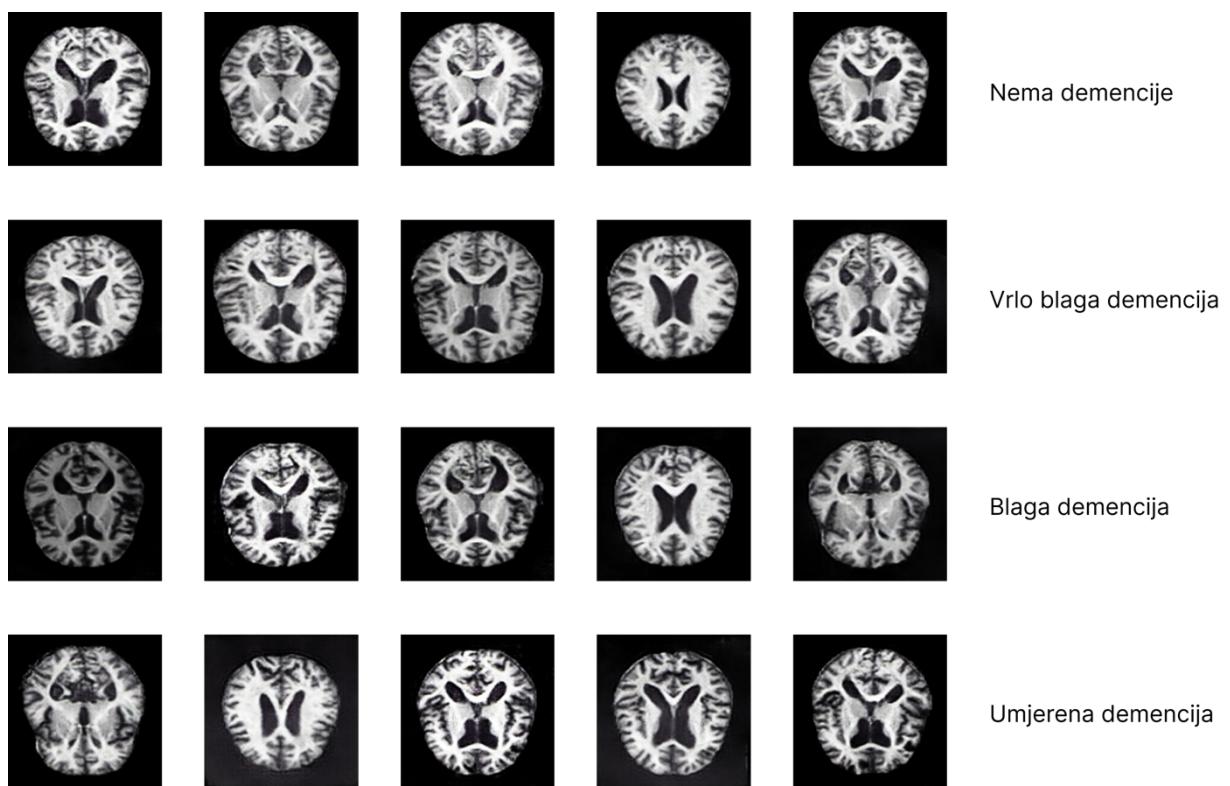


Slika 6.5. Prikaz rezultata preciznosti, osjetljivosti i 1-NN testa za odabrani GAN model

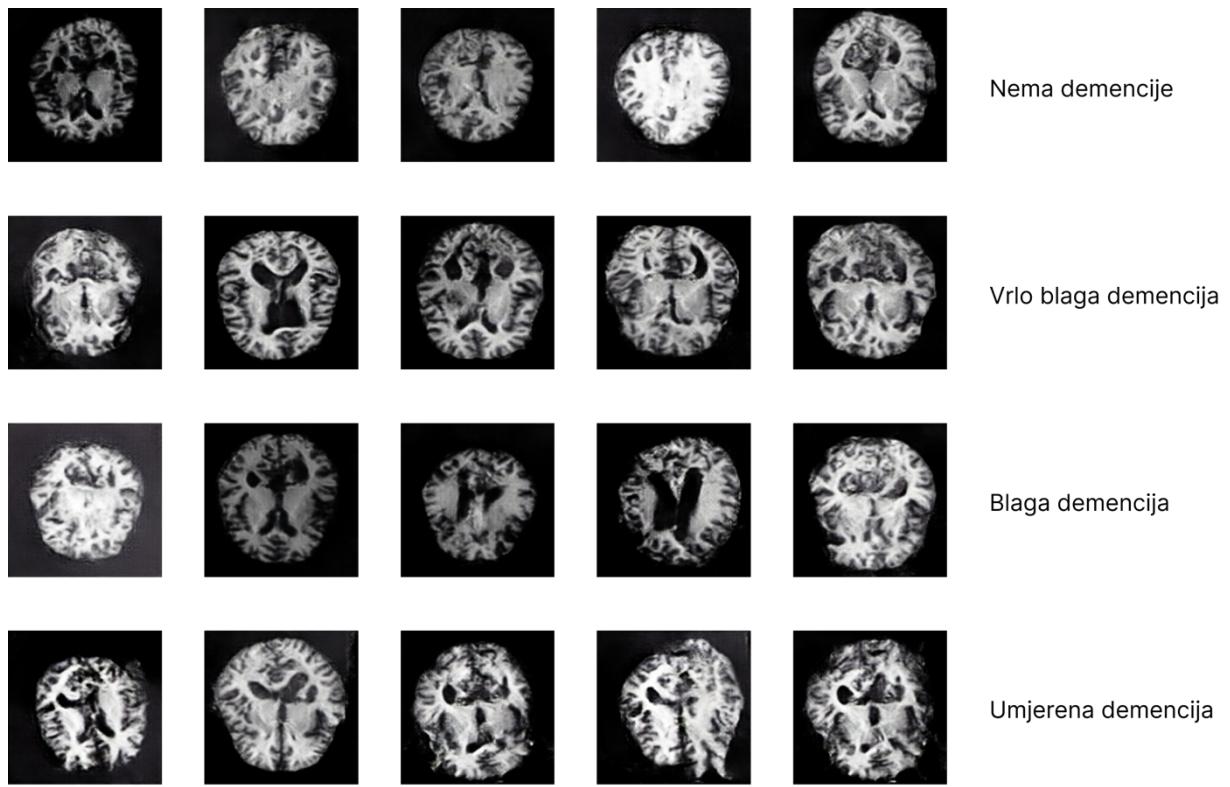
CID indeks je pružio uvid u kreativne sposobnosti generatora. Komponenta kreativnosti se smanjila s približno 3,5 na vrijednost 2,3 te stabilizirala, što sugerira da su generirane slike postale sličnije stvarnim podatcima tijekom treninga. Nasljeđivanje je ostalo relativno stabilno (oko 1) što ukazuje na očuvanje bitnih značajki iz distribucije podataka za treniranje. Vrijednost raznolikosti je, pak, fluktuirala između 4 i 6, približavajući se konačnoj vrijednosti 5, što sugerira dosljednu varijabilnost u generiranim uzorcima. Objedinjenu vrijednost CID indeksa je teško odrediti iz grafa, no prilikom evaluacije modela na testnom dijelu seta podataka pokazalo se da iznosi 7,45. Sve navedeno ukazuje kako je DCGAN model uspješno naučio generirati raznolike slike primjerene klasi, istovremeno održavajući varijabilnost unutar svake od četiri klase.



Slika 6.6. Prikaz rezultata CID metrike i njenih komponenti za odabrani GAN model



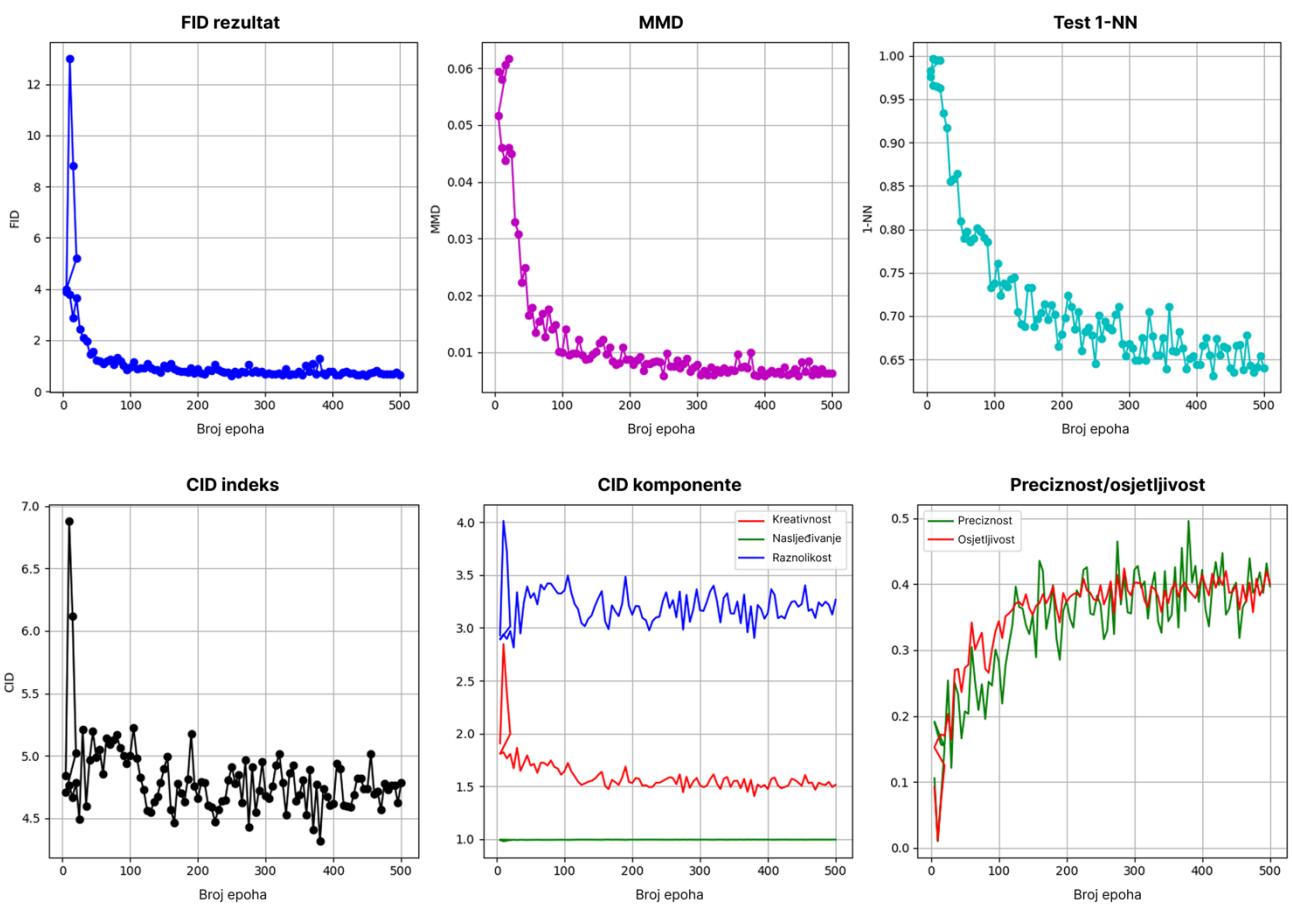
Slika 6.7. Primjeri slika izvrsne kvalitete generiranih predloženom DCGAN arhitekturom



Slika 6.8. Primjeri slika loše kvalitete generiranih predloženom DCGAN arhitekturom

6.2.2. Komparativna analiza s verzijom GAN-a niže rezolucije (64×64 piksela)

Druga verzija GAN-a, ona niže rezolucije, pokazala je superiorne performanse na više metrika, ističući neizbjegne kompromise između rezolucije i kvalitete generiranih slika. Ovaj je model postigao bolji FID rezultat (1 – 2) u usporedbi s glavnom verzijom (128 x 128 piksela). Pokazao je i brže te stabilnije obrasce konvergencije. Test 1-NN postigao je vrijednost oko 0,65, približivši se dodatno optimalnom rasponu od 0,5. Navedeno ukazuje na uspješnije zavaravanje diskriminatora. Također, metrike preciznosti i osjetljivost postigle su dobre rezultate, s istim konačnim vrijednostima (obje oko 0,4) te konzistentnim putanjama konvergencije.



Slika 6.9. Prikaz rezultata performansi GAN modela s rezolucijom 64x64 piksela

Za ovakve razlike u performansama moguće je pripisati nekoliko čimbenika. Arhitektura 64 x 64 piksela ima koristi od dublje mreže diskriminatora, s dodatnim konvolucijskim slojem (512 filtera), što pruža mogućnost robusnijeg izdvajanja značajki. Nadalje, manja računalna složenost omogućuje stabilniji put ka konvergenciji te smanjenje vjerojatnosti za pojavu kolapsa moda. Manja veličina ulaznih slika smanjuje složenost zadatka generiranja, omogućujući modelu da usmjeri računalne resurse na poboljšanje kvalitete uzorka.

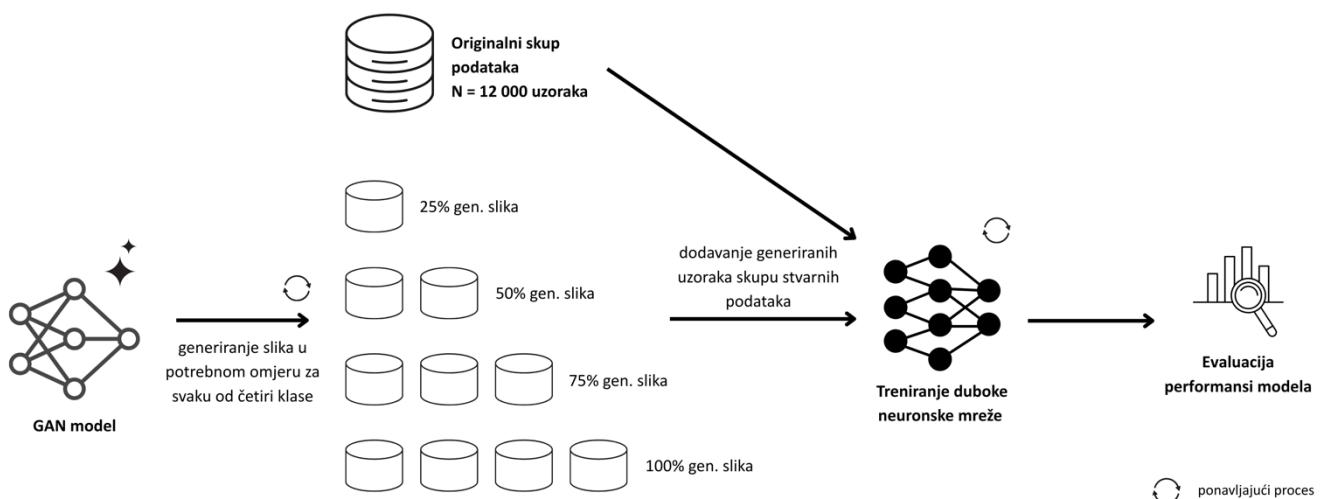
Unatoč svim izazovima, GAN model od 128 x 128 piksela postigao je odličan uspjeh u generiranju MR slika mozga visoke rezolucije, koje su pogodne za proširenje postojećeg skupa podataka. Metrike tijekom treniranja pokazuju da je, bez obzira na to što je modelu bilo potrebno više epoha za postizanje stabilnosti u usporedbi s manjim modelom, u konačnici proizveo vizualno uvjerljive rezultate s odgovarajućom kvalitetom i raznolikošću. Nešto više vrijednosti FID rezultata te preciznosti i osjetljivosti predstavljaju prihvatljiv kompromis za značajno povećanje rezolucije slike i očuvanje anatomskeih značajki potrebnih za ispravnu

klasifikaciju demencijskih promjena. Stoga je, za daljnje eksperimente proširenja seta podataka, odabrana navedena verzija GAN modela.

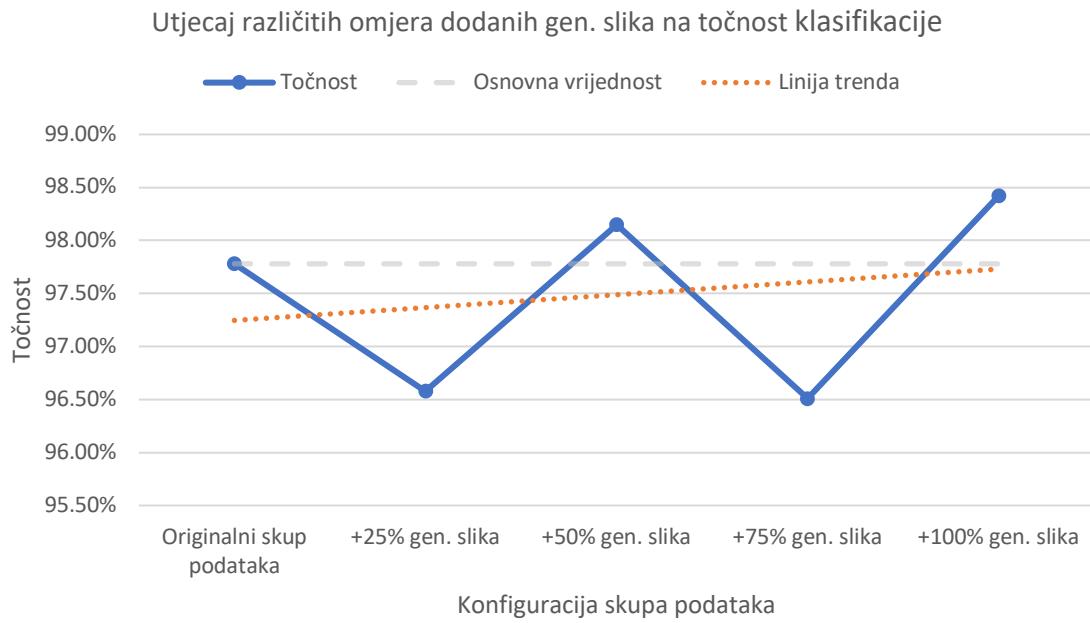
6.3. Rezultati DNN klasifikacije s dodanim umjetno generiranim slikama

Evaluacija učinkovitosti klasifikacije nakon integracije sintetičkih slika dobivenih GAN modelom pokazala je ne baš očite obrasce u tome kako različiti omjeri proširenja utječu na točnost modela te performanse specifične za pojedinu klasu.

Eksperimentalni je dio uključivao sustavno dodavanje različitih udjela sintetičkih slika (25%, 50%, 75% i 100%) izvornom uravnoteženom skupu podataka od 12 000 MR slika mozga. Za svaki omjer proširenja podataka, generiran je potreban broj sintetičkih uzoraka za svaku od četiri klase demencije pozivom specijalizirane funkcije.



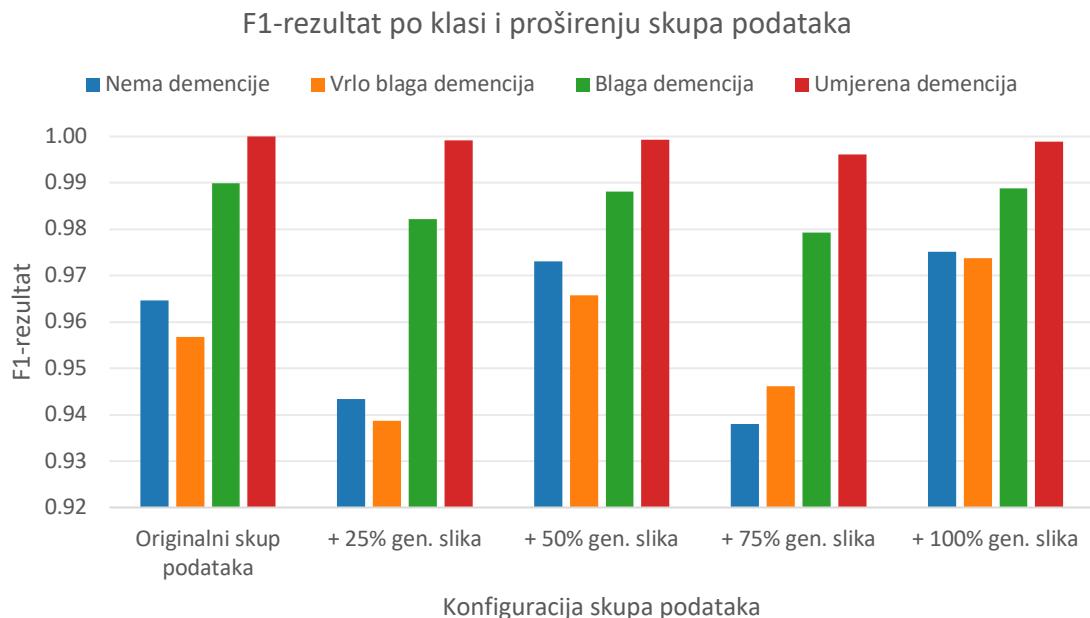
Slika 6.10. Prikaz izvođenja eksperimentalnog dijela projekta



Slika 6.11. Ukupna točnost klasifikacije s različitim omjerima umjetnog proširenja podataka

6.3.1. Analiza obrazaca i promjena performansi

Ukupna je točnost klasifikacije pokazala nelinearni odnos s različitim omjerima generiranih podataka, dovodeći u pitanje pretpostavku kako više sintetičkih podataka može dovesti do boljih performansi modela. Polazna točnost od 97,78% utvrđena izvornim skupom podataka poslužila je kao robusna referentna vrijednost, a rezultati proširenja pokazali su poboljšanja, ali i degradaciju, ovisno o primjenjenom omjeru. Stopostotno proširenje generiranim uzorcima postiglo je najveću točnost od 98,42%, što predstavlja poboljšanje od 0,66% u odnosu na početnu vrijednost. Nasuprot tome, proširenje sa 75% sintetičkih podataka pokazalo je najslabije performanse (točnosti 96,51%), što upućuje na degradaciju od 1,27%. Zanimljivo je da je omjer proširenja od 50% također premašio osnovne performanse (98,15%), što sugerira da umjerene razine proširenja mogu biti korisne. Omjeri od 25% i 75% rezultirali su smanjenjem performansi u ovome eksperimentu.



Slika 6.12. Performanse F1-rezultata po klasama uz različite omjere umjetnog proširenja podataka

6.3.2. Analiza performansi po klasama

Utjecaj se proširenja skupa podataka sintetičkim uzorcima značajno razlikovao među različitim klasama demencije, otkrivajući osjetljivosti klase na kvalitetu i količinu generiranih podataka.

Klase *Umjerena demencija* dosljedno je održavala iznimne performanse u svim omjerima proširenja, s F1-rezultatom koji je ostao iznad 99,6%, što pokazuje robusnost ove klase na varijacije broja sintetičkih uzoraka.

Klase *Blaga demencija* pokazala je relativno stabilne performanse, s F1-rezultatom koji je minimalno varirao između 97,9% i 99,0%, što ukazuje na dobru kvalitetu generiranih slika za ovu klasu.

Istiće se klasa *Nema demencije*, s najvećim varijacijama performansi s obzirom na omjer proširenja. Za ovu klasu, F1-rezultat se mijenjao u rasponu od 93,8% (75%-tno proširenje) do 97,5% (100%-tno proširenje), što dovodi do zaključka da ova klasa ima najviše koristi od visokokvalitetnih sintetičkih podataka. Kada su sintetički uzorci neoptimalni, tada su i vrijednosti performansi niže.

Klasa *Vrlo blaga demencija* pokazala je značajno poboljšanje pri 100%-tnom proširenju izvornog seta uzoraka (F1-rezultat: 97,4%) u usporedbi s početnom vrijednošću (95,7%). Navedeno ukazuje da ova klasa ima koristi od povećanja raznolikosti podataka koji se koriste za treniranje modela.

6.3.3. Analiza obrazaca u pogreškama predviđanja modela

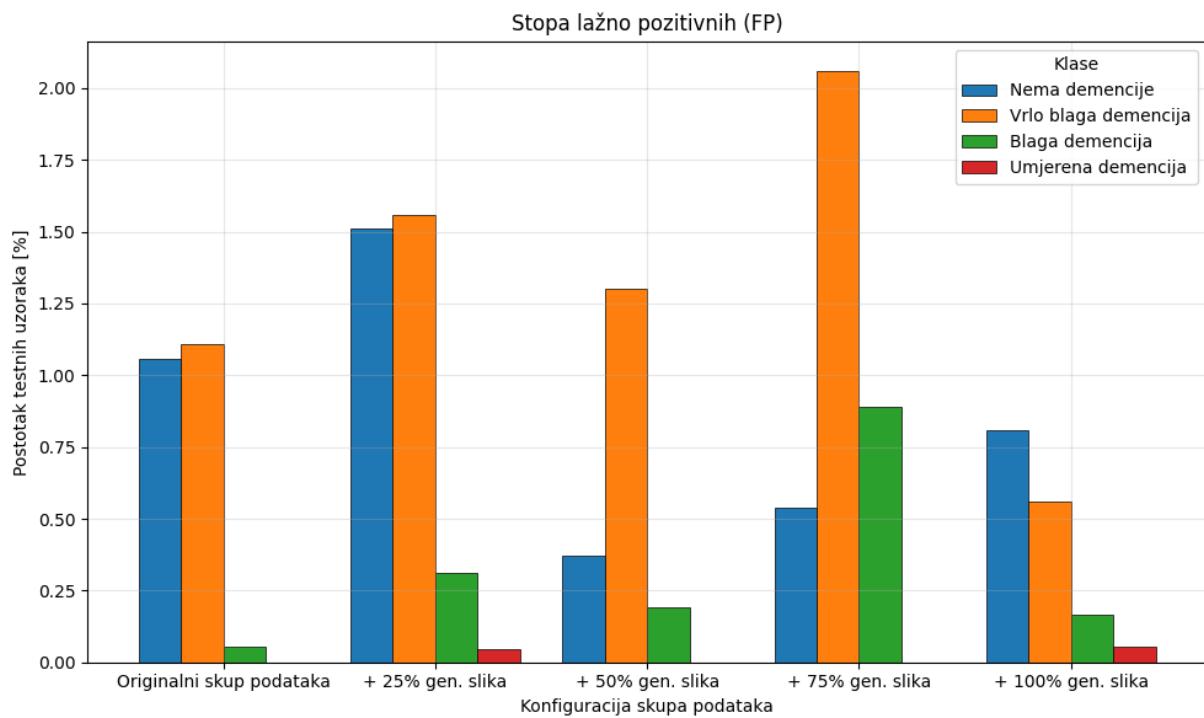
Detaljnom analizom postotaka TP, TN, FP i FN može se otkriti kako ovakav princip sintetičkog proširenja podataka utječe na pogrešna predviđanja klasifikacijskog modela.

Konfiguracija proširenja od 25% pokazala je povećane stope pogrešaka u većini klase, s povećanjem i lažno pozitivnih i lažno negativnih uzoraka, posebno za klasu *Nema demencije* (FN se povećao za 0,61%).

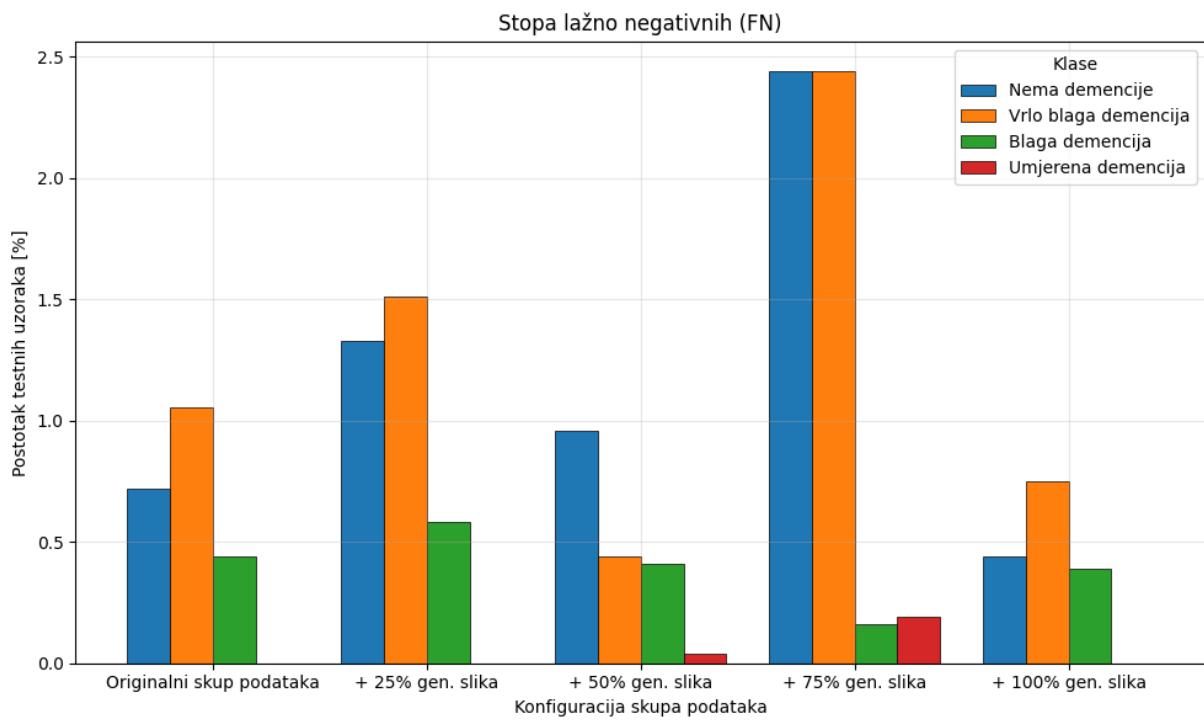
Proširenje podataka od 50% pokazalo je poboljšanu ravnotežu pogrešaka, sa smanjenim stopama lažno pozitivnih za klasu *Nema demencije* (FP se smanjio za 0,69%).

75%-tno proširenje otkrilo je problematične obrasce, posebno za klasu *Nema demencije*, gdje se stopa lažno negativnih povećala (FN se povećao za 1,72%). Ovo sugerira moguće pogoršanje kvalitete sintetičkih podataka pri ovome omjeru.

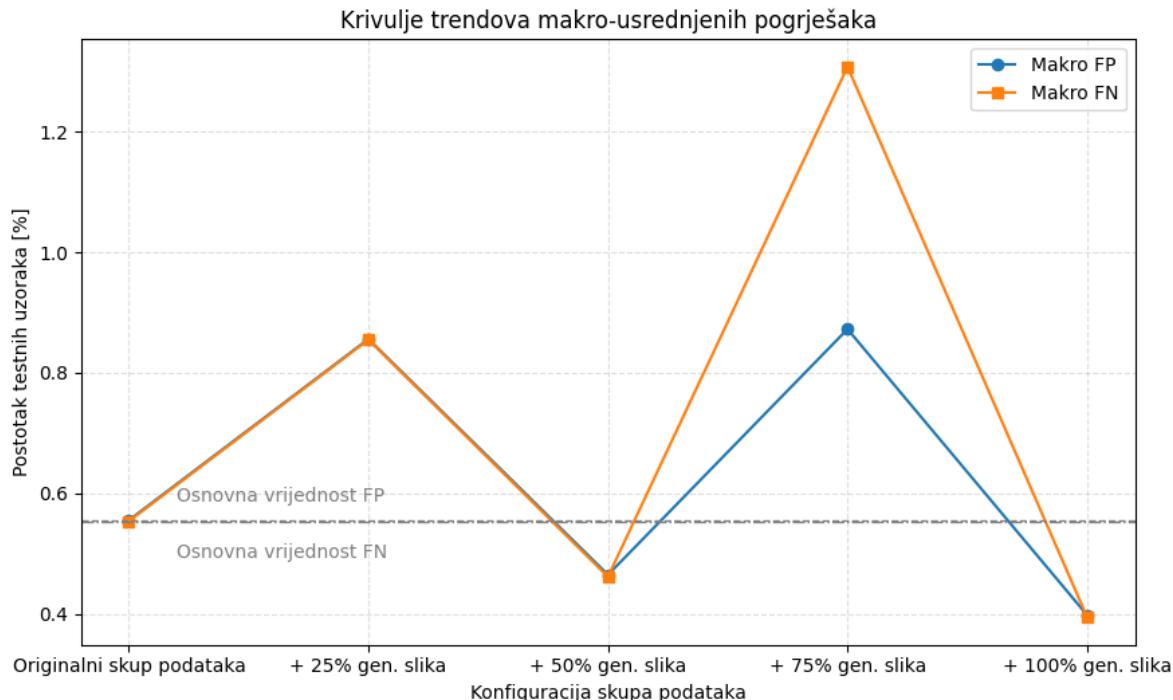
Povećanje u omjeru od 100% podataka postiglo je najbolju ukupnu ravnotežu pogrešaka, sa značajnim poboljšanjima u stopama lažno pozitivnih za klasu *Nema demencije* (FP smanjen za 0,25%) i *Vrlo blaga demencija* (FP smanjen za 0,55%), uz održavanje stabilnih stopa stvarno pozitivnih uzoraka u svim klasama.



Slika 6.13. Utjecaj različitih omjera umjetnog proširenja podataka na postotak lažno pozitivnih uzoraka po klasama



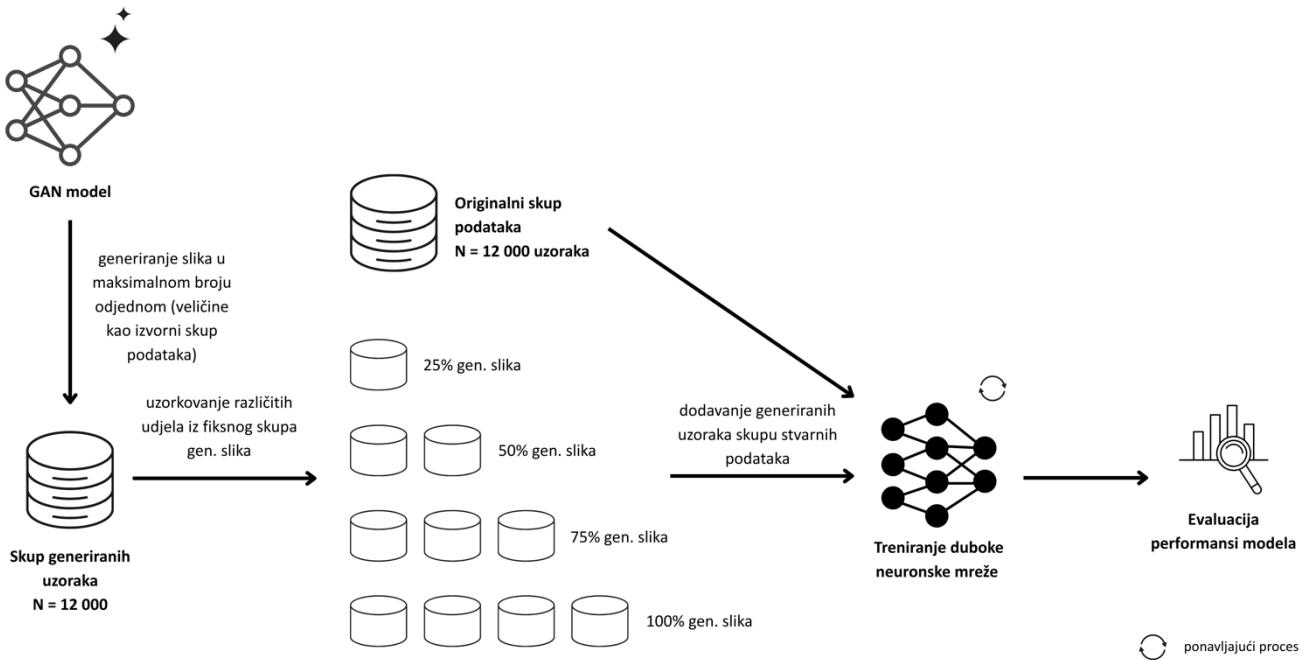
Slika 6.14. Utjecaj različitih omjera umjetnog proširenja podataka na postotak lažno negativnih uzoraka po klasama



Slika 6.15. Prikaz promjene makro-usrednjeni pogrešaka klasifikacije prema omjerima proširenja podataka

6.3.4. Razmatranje alternativne metodologije

Alternativni eksperimentalni pristup koji bi mogao pružiti veću kontrolu nad evaluacijom sintetičkih podataka uključivao bi generiranje velike serije sintetičkih slika (100% podataka – 3 000 slika po klasi) korištenjem specijalizirane funkcije samo jednom. Zatim bi se provodilo sustavno uzorkovanje različitih udjela (25%, 50%, 75%, 100%) iz ovog fiksnog skupa za svaki eksperiment. Ova metodologija bi potencijalno mogla eliminirati varijabilnost zbog generiranja različitih sintetičkih uzoraka za svaki udio proširenja, omogućujući precizniju procjenu kako dodavanje sintetičkih podataka utječe na performanse klasifikacije. Takav bi pristup mogao omogućiti izravnu usporedbu istih sintetičkih uzoraka na različitim razinama proširenja, pružajući jasniji uvid u optimalne omjere stvarnih i generiranih podataka, uz kontrolu varijacije kvalitete sintetičkih podataka.



Slika 6.16. Prikaz alternativnog pristupa izvođenja eksperimenta

Sveobuhvatna analiza otkriva kako umjetno proširenje podataka korištenjem cDCGAN modela može učinkovito poboljšati performanse klasifikacije. U ovom projektu, optimalni rezultati postignuti su pri 100%-tnom proširenju (udvostručenje veličine skupa podataka). Međutim, uspjeh kritično ovisi o količini i kvaliteti umjetno stvorenih uzoraka, pri čemu svaki omjer (udio) povećanja podataka potencijalno uvodi neoptimalne sintetičke uzorke koji mogu negativno utjecati na performanse. Otkriveni obrasci specifični za klasu sugeriraju da neke klase demencije imaju više koristi od umjetnog proširenja od drugih, što naglašava važnost prilagođenih strategija proširenja seta podataka u zadatcima klasifikacije medicinskih slika.

7. ZAKLJUČAK

Ovaj je projekt uspješno pokazao učinkovitost korištenja cDCGAN mreža kao održive strategije proširenja podataka za klasifikaciju demencije iz MR slika mozga. Osnovni su rezultati dobiveni koristeći metodu prijenosa znanja s EfficientNetB0 modelom, postižući točnost klasifikacije od 97,78%. Potom je razvijen i treniran kvalitetni cGAN model koji je sposoban generirati realistične sintetičke MR slike mozga rezolucije 128 x 128 piksela (za četiri klase demencije). Sustavna evaluacija umjetnog proširenja skupa podataka otkrila je da je stopostotno proširenje (udvostručenje veličine skupa podataka) postiglo optimalne rezultate s točnošću od 98,42%, predstavljajući poboljšanje 0,65% u odnosu na već izvrsne, početne rezultate.

U radu se bavilo ključnim izazovima koje donose primjene umjetne inteligencije u medicini. Pokazano je kako korištenje generativnih mreža može učinkovito doprinijeti rješavanju problema malih ili neuravnoteženih skupova medicinskih podataka, uz očuvanje privatnosti ispitanika. Sveobuhvatne metode evaluacije, koje uključuju tradicionalne metrike performansi klasifikacije te napredne mjere specifične za GAN (FID, MMD, CID, preciznost, osjetljivost), postavljaju vrijedne referentne vrijednosti za buduća istraživanja u području sinteze dijagnostičkih slika. Rezultati otkrivaju kako utjecaj proširenja skupa generiranjem novih uzoraka značajno varira među različitim klasama demencije, s optimalnim rezultatima postignutim pažljivom ravnotežom između količine i kvalitete generiranih podataka.

Budući bi se rad trebao usredotočiti na unaprjeđenje arhitekture generativne mreže, uključujući sintezu veće rezolucije (256 x 256 ili 512 x 512 piksela), implementaciju najsuvremenijih GAN varijanti poput StyleGAN-a ili difuzijskih modela. Nadalje bi se moglo razmisiliti i o uključivanju mehanizama pažnje (*engl. attention mechanisms*) za poboljšano očuvanje anatomskih detalja. Razvoj sustava multimodalnog generiranja podataka koji kombinira MR snimke s kliničkim metapodatcima mogao bi pridonijeti stvaranju sveobuhvatnijih dijagnostičkih setova podataka. Ostala proširenja koja bi rad mogao doživjeti uključuju: istraživanje optimalnih omjera stvarnih i sintetičkih uzoraka putem sustavnih metodologija uzorkovanja, integracija naprednih tehnika treniranja modela (progresivni rast, spektralna normalizacija) te istraživanje međumodalne sinteze između različitih sekvenci MR snimki (T1, T2, FLAIR). Sve navedeno predstavlja obećavajuće smjerove za poboljšanje kvalitete sinteze medicinskih slika i kliničke primjenjivosti.

8. LITERATURA

- [1] Zhou Y. i dr.: “Advancing ASD identification with neuroimaging: a novel GARN methodology integrating Deep Q-Learning and generative adversarial networks”, *BMC Medical Imaging*, Vol. 24, No. 1, 2024.
- [2] Kazuhiro, K. i dr.: “Generative adversarial networks for the creation of realistic artificial brain magnetic resonance images”, *Tomography*, Vol. 4, No. 4, pp. 159–163, 2018.
- [3] Oriol, M.: “Generative Adversarial Networks (GAN) to generate synthetic MRI brain images”, *Dipòsit Digital De Documents De La UAB*, 2024.
- [4] Zhou, X. i dr.: “Enhancing magnetic resonance imaging-driven Alzheimer’s disease classification performance using generative adversarial learning”, *Alzheimer S Research & Therapy*, Vol. 13, No. 1, 2021.
- [5] Neff, T. i dr.: „Generative Adversarial Network based Synthesis for Supervised Medical Image Segmentation“, *OAGM&ARW Joint Workshop 2017*, pp. 140–145, 2017.
- [6] Jeong, J.; Patel, B.; Banerjee, I.: “GAN augmentation for multiclass image classification using hemorrhage detection as a case-study”, *Journal of Medical Imaging*, Vol. 9, No. 3, 2022.
- [7] Skandarani, Y.; Jodoin, P.-M.; Lalande, A.: “GANS for Medical Image Synthesis: An Empirical study”, *Journal of Imaging*, Vol. 9, No. 3, 2023.
- [8] Fetty L. i dr.: “Latent space manipulation for high-resolution medical image synthesis via the StyleGAN”, *Zeitschrift Für Medizinische Physik*, Vol. 30, No. 4, pp. 305–314, 2020.
- [9] Cui, L. i dr.: “Towards reliable healthcare Imaging: conditional contrastive generative adversarial network for handling class imbalancing in MR Images”, *PeerJ Computer Science*, Vol. 10, 2024.
- [10] Onakpojeruo, E. P. i dr.: “Enhanced MRI-based brain tumor classification with a novel Pix2pix generative adversarial network augmentation framework”, *Brain Communications*, Vol. 6, No. 6, 2024.
- [11] Bowles, C. i dr.: “GANsfer Learning: Combining labelled and unlabelled data for GAN based data augmentation”, *arXiv (Cornell University)*, 2018.

- [12] Brugnara, G. i dr.: “Addressing the Generalizability of AI in Radiology Using a Novel Data Augmentation Framework with Synthetic Patient Image Data: Proof-of-Concept and External Validation Classification Tasks in Multiple Sclerosis”, *Radiology Artificial Intelligence*, Vol. 6, No. 6, 2024.
- [13] Hong, S. i dr.: “3D-StyleGAN: a Style-Based Generative Adversarial network for generative Modeling of Three-Dimensional Medical Images”, *Lecture notes in computer science*, pp. 24–34, 2021.
- [14] Koshino, K. i dr.: “Narrative review of generative adversarial networks in medical and molecular imaging”, *Annals of Translational Medicine*, Vol. 9, No. 9, 2021.
- [15] Devi, S. i dr.: “Data Augmentation using GANs for Image Classifications”, Global Scientific Journals, Vol. 13, No. 5, pp. 522-526, 2025.
- [16] Öztemiz, M.: “Building a U-Net Architecture for Image Segmentation with Python and Keras”, s Interneta, <https://medium.com/@mustii27/building-a-u-net-architecture-for-image-segmentation-with-python-and-keras-30a381af7731>, 12. srpnja 2025.
- [17] Durdov, B.: “Classification performance analysis of medical histopathology images using deep neural networks”, Diplomski rad, Sveučilište u Splitu, Fakultet elektrotehnike, strojarstva i brodogradnje, Split, 2024.
- [18] Boesch, G.: “Very Deep Convolutional Networks (VGG) Essential Guide”, s Interneta, <https://viso.ai/deep-learning/vgg-very-deep-convolutional-networks/>, 11. srpnja 2025.
- [19] Sandler, M.; Howard, A.: “MobileNetV2: The Next Generation of On-Device Computer Vision Networks”, s Interneta, <https://research.google/blog/mobilenetv2-the-next-generation-of-on-device-computer-vision-networks/>, 12. srpnja 2025.
- [20] Ahmed, T.; Sabab, N. H. N.: “Classification and understanding of cloud structures via satellite images with EfficientUNet”, *arXiv (Cornell University)*, 2020.
- [21] Keras.io: “Keras Applications”, s Interneta, <https://keras.io/api/applications/>, 11. srpnja 2025.
- [22] Mukhlif, A. A.; Al-Khateeb, B.; Mohammed, M. A.: “Incorporating a novel dual transfer learning approach for medical images”, *Sensors*, Vol. 23, No. 2, 2023.
- [23] Stipanićev, D.; Šerić, Lj.; Braović, M.: “Uvod u umjetnu inteligenciju”, Fakultet elektrotehnike, strojarstva i brodogradnje, Split, 2021.

- [24] TensorFlow: "CycleGAN", <https://www.tensorflow.org/tutorials/generative/cyclegan>, 11. srpnja 2025.
- [25] Osuala, R. i dr.: "Data synthesis and adversarial networks: A review and meta-analysis in cancer imaging", *Medical Image Analysis*, Vol. 84, 2022.
- [26] Jeong, J. J. i dr.: "Systematic Review of Generative Adversarial Networks (GANs) for medical image classification and Segmentation", *Journal of Digital Imaging*, Vol. 35, No. 2, pp. 137–152, 2022.
- [27] Ferreira, A. i dr.: "GAN-based generation of realistic 3D volumetric data: A systematic review and taxonomy", *Medical Image Analysis*, Vol. 93, 2024.
- [28] Khan, J.: „StyleGAN: Use machine learning to generate and customize realistic images“, s Interneta, <https://www.comet.com/site/blog/stylegan-use-machine-learning-to-generate-and-customize-realistic-images/>, 12. srpnja 2025.
- [29] Lawton, G.: „What is Fréchet inception distance (FID)?“, s Interneta, <https://www.techtarget.com/searchenterpriseai/definition/Frechet-inception-distance-FID>, 12. srpnja 2025.
- [30] Thakur, A.: „How to Evaluate GANs using Frechet Inception Distance (FID)“, s Interneta, <https://wandb.ai/ayush-thakur/gan-evaluation/reports/How-to-Evaluate-GANs-using-Frechet-Inception-Distance-FID---Vmlldzo0MTAxOTI>, 11. srpnja 2025.
- [31] Xu, Q. i dr.: "An empirical study on evaluation metrics of generative adversarial networks", *arXiv (Cornell University)*, 2018.
- [32] Zhou, Z. i dr.: "DC-MMD-GAN: A new maximum mean discrepancy generative adversarial network using divide and conquer", *Applied Sciences*, Vol. 10, No. 18, 2020.
- [33] Guan, S.; Loew, M.: "Evaluation of Generative Adversarial Network Performance Based on Direct Analysis of Generated Images", *2019 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, SAD, pp. 1-5, 2019.
- [34] Sajjadi, M. S. M. i dr.: „Assessing generative models via precision and recall“, *32nd International Conference on Neural Information Processing Systems (NIPS'18)*, pp. 5234–5243, 2018.
- [35] Washington University: "Open Access Series of Imaging Studies (OASIS)", s Interneta, <https://sites.wustl.edu/oasisbrains/>, 11. srpnja 2025.

- [36] LaMontagne P. J. i dr.: "OASIS-3: Longitudinal Neuroimaging, Clinical, and Cognitive Dataset for Normal Aging and Alzheimer Disease", *medRxiv (Cold Spring Harbor Laboratory)*, 2019.
- [37] Kaggle: "Alzheimer's Disease Multiclass Images Dataset", s Interneta, <https://www.kaggle.com/datasets/aryansinghal10/alzheimers-multiclass-dataset-equal-and-augmented>, 11. srpnja 2025.
- [38] Microsoft Learn: "Install WSL", s Interneta, <https://learn.microsoft.com/en-us/windows/wsl/install>, 12. srpnja 2025.
- [39] Google Colab: „Google Colaboratory“, s Interneta, <https://colab.google/>, 12. srpnja 2025.
- [40] Sveučilišni računski centar (Srce): "Napredno računanje", s Interneta, <https://www.srce.unizg.hr/napredno-racunanje>, 12. srpnja 2025.
- [41] TensorFlow: "Install TensorFlow with pip", s Interneta, <https://www.tensorflow.org/install/pip#windows-wsl2>, 12. srpnja 2025.
- [42] TensorFlow: "Build from source", s Interneta, <https://www.tensorflow.org/install/source#gpu>, 12. srpnja 2025.
- [43] Goodfellow, I. J.: "NIPS 2016 Tutorial: Generative Adversarial Networks", *arXiv (Cornell University)*, 2017.
- [44] Antoniou, A.; Storkey, A. J.; Edwards, H.: "Data augmentation generative adversarial networks", *arXiv (Cornell University)*, 2017.
- [45] Brock, A.; Donahue, J.; Simonyan, K.: "Large scale GAN training for high fidelity natural image synthesis", *arXiv (Cornell University)*, 2018.
- [46] Shrivastava, A. i dr.: "Learning from Simulated and Unsupervised Images through Adversarial Training", *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2242-2251, 2017.
- [47] Ahmad, W. i dr.: "A new generative adversarial network for medical images super resolution", *Scientific Reports*, Vol. 12, No. 1, 2022.
- [48] Kynkänniemi, T. i dr.: „Improved precision and recall metric for assessing generative models“, *33rd International Conference on Neural Information Processing Systems*, Article 353, pp. 3927–3936, 2019.

- [49] Dayarathna, S. i dr.: “Deep learning based synthesis of MRI, CT and PET: Review and analysis”, *Medical Image Analysis*, Vol. 92, 2023.
- [50] Conte G. M. i dr.: “Generative adversarial networks to synthesize missing T1 and FLAIR MRI sequences for use in a multisequence brain tumor segmentation model”, *Radiology*, Vol. 299, No. 2, pp. 313–323, 2021.
- [51] Zhang, H. i dr.: „Self-Attention Generative Adversarial Networks“, *36th International Conference on Machine Learning*, pp. 7354-7363, 2019.
- [52] Amini, A.: „Introduction to Deep Learning“, s Interneta, https://introtodeeplearning.com/slides/6S191_MIT_DeepLearning_L1.pdf, 12. srpnja 2025.

POPIS OZNAKA I KRATICA

MR	Magnetska rezonancija
GAN	<i>Generative adversarial network</i> (generativna kontradiktorna mreža)
cDCGAN	<i>conditional Deep Convolutional GAN</i> (uvjetni duboki konvolucijski GAN)
ASD	<i>Autism Spectrum Disorder</i> (poremećaj iz spektra autizma)
GARL	<i>Generative Adversarial Reinforcement Learning</i> (pojačano generativno kontradiktorno učenje)
WGAN	Wasserstein GAN
WGAN-GP	<i>Wasserstein GAN with Gradient Penalty</i> (Wassersteinov GAN s gradijentnom kaznom)
LSGAN	<i>Least Squares GAN</i> (GAN najmanjih kvadrata)
CT	<i>Computed Tomography</i> (kompjutorizirana tomografija)
RGB	<i>red – green – blue</i> (model boja crvena – zelena – plava)
ACDC	<i>Automated Cardiac Diagnosis Challenge</i> (izazov automatizirane srčane dijagnoze)
FID	<i>Fréchet Inception Distance</i> (Fréchet Inception udaljenost)
cCGAN	<i>Conditional Contrastive GAN</i> (uvjetna kontrastna generativna kontradiktorna mreža)
SCoLN	<i>Supervised Contrastive-based Learning Network</i> (nadzirana neuronska mreža temeljena na kontrastima)
OASIS	<i>Open Access Series of Imaging Studies</i> (Serija dijagnostičkih studija otvorenog pristupa)
MS	<i>Multiple sclerosis</i> (Multipla skleroza)

ROC	<i>Receiver Operating Characteristic</i> (Radna karakteristika prijemnika)
AUC	<i>Area Under Curve</i> (Površina ispod krivulje)
3D	trodimenzionalni
AD	<i>Alzheimer's disease</i> (Alzheimerova bolest)
PGGAN	<i>Progressive Growing GAN</i> (Progresivno rastući GAN)
DNN	<i>deep neural network</i> (duboka neuronska mreža)
CNN	<i>convolutional neural network</i> (konvolucijska neuronska mreža)
ReLU	<i>Rectified Linear Unit</i> (Ispravljena linearna jedinica)
IS	<i>Inception Score</i> (Inception rezultat)
PET	<i>positron emission tomography</i> (pozitronska emisijska tomografija)
JPG (JPEG)	<i>Joint Photographic Experts Group</i> (Skupina stručnjaka za fotografsku obradu)
DICOM	<i>Digital Imaging and Communications in Medicine</i> (Standard za digitalnu dijagnostiku i komunikaciju u medicini)
NIfTI	<i>Neuroimaging Informatics Technology Initiative</i> (Inicijativa za informatičku tehnologiju u neurološkoj dijagnostici)

NASLOV

UTJECAJ GENERATIVNIH MODELA NA UČINKOVITOST KLASIFIKACIJE MEDICINSKIH SLIKA

SAŽETAK

Ovaj rad istražuje upotrebu uvjetnih dubokih konvolucijskih generativnih kontradiktornih mreža (cDCGAN) za proširenje skupa slika magnetske rezonancije mozga koji se koristi za klasifikaciju stadija demencije. Prvo se dohvataju bazni rezultati procesom prijenosa znanja s modernim konvolucijskim neuronskim mrežama, a potom se razvija klasno uvjetovani GAN model za generiranje MR slika visoke kvalitete. Sveobuhvatni proces evaluacije, koji kombinira metrike za procjenu kvalitete i raznolikosti, potvrđuje realističnost generiranih uzoraka. Sustavno se istražuje utjecaj dodavanja različitih udjela sintetičkih podataka stvarnim slikama na performanse klasifikacije. Pokazujući kako pažljivo dodano umjetno proširenje može nadoknaditi nedostatak podataka te učiniti klasifikacijske modele robusnijima, bez ugrožavanja privatnosti pacijenata, ovaj rad postavlja temelje za daljnji napredak u sintezi medicinskih slika, uključujući generiranje slika veće rezolucije te razvijanje snažnijih GAN arhitektura.

KLJUČNE RIJEČI

proširenje skupa podataka, klasifikacija stadija demencije, uvjetni DCGAN, sintetičke MR slike, prijenos znanja

TITLE

THE IMPACT OF GENERATIVE MODELS ON THE EFFECTIVENESS OF MEDICAL IMAGE CLASSIFICATION

SUMMARY

This thesis investigates the use of conditional Deep Convolutional GANs (cDCGANs) to augment brain MRI dataset for dementia classification. It achieves this by first establishing a transfer-learning baseline with modern convolutional neural networks, then developing class-conditional GAN to produce high-resolution MRI images. A comprehensive evaluation framework - combining generative quality and diversity metrics - confirms the realism of the generated images. The work systematically explores the impact of blending various proportions of synthesized data with actual images on classification performance. By demonstrating that carefully blended synthetic augmentation can offset data scarcity and render classification models more robust without compromising patient privacy, this work lays the groundwork for continued advances in medical image synthesis, including higher-resolution generation and more powerful GAN architectures.

KEYWORDS

data augmentation, dementia classification, cDCGAN, synthetic MRI images, transfer learning

DODATAK A – POPIS SLIKA I TABLICA

Popis slika

<i>Slika 3.1. Arhitektura neuronske mreže s označenim glavnim dijelovima</i>	10
<i>Slika 3.2. Prikaz primjera U-Net arhitekture [16]</i>	11
<i>Slika 3.3. Arhitektura VGG19 modela [18]</i>	12
<i>Slika 3.4. Arhitektura MobileNetV2 modela [19]</i>	13
<i>Slika 3.5. Arhitektura EfficientNetB0 modela [20]</i>	14
<i>Slika 3.6. Struktura matrice zbirjenosti</i>	15
<i>Slika 3.7. Primjer ROC krivulje i površine ispod krivulje (AUC)</i>	17
<i>Slika 3.8. Osnovna ideja prijenosa učenja [22]</i>	19
<i>Slika 4.1. Proces suparničkog treniranja u GAN mreži [23]</i>	21
<i>Slika 4.2. Primjer rezultata dobivenog korištenjem CycleGAN modela [24]</i>	22
<i>Slika 4.3. Primjer promjene značajki na slikama ljudskog lica pomoću StyleGAN-a [28]</i>	23
<i>Slika 4.4. Ilustracija prijevoda između modaliteta T1 i T2 MR snimki mozga [27]</i>	24
<i>Slika 4.5. Primjer povećanja rezolucije MR snimke mozga [27]</i>	24
<i>Slika 5.1. Primjeri uzoraka iz korištenog skupa podataka po klasama [37]</i>	32
<i>Slika 5.2. Prikaz distribucije korištenog skupa podataka</i>	33
<i>Slika 5.3. Prikaz arhitektura generatora i diskriminadora u korištenom rješenju [6]</i>	35
<i>Slika 5.4. Prikaz podjele korištenog skupa podataka (ukupno 12 000 slika)</i>	36
<i>Slika 6.1. Rezultati točnosti promatranih DNN modela tijekom treninga i validacije</i>	39
<i>Slika 6.2. Prikaz matrice zbirjenosti EfficientNetB0 modela (odmrzavanje 10 posljednjih slojeva mreže u fazi finog podešavanja)</i>	40
<i>Slika 6.3. Prikaz matrice zbirjenosti EfficientNetB0 modela (odmrzavanje svih slojeva mreže u fazi finog podešavanja)</i>	41
<i>Slika 6.4. Prikaz rezultata FID i MMD metrika za odabrani GAN model</i>	42
<i>Slika 6.5. Prikaz rezultata preciznosti, osjetljivosti i 1-NN testa za odabrani GAN model</i>	43
<i>Slika 6.6. Prikaz rezultata CID metrike i njenih komponenti za odabrani GAN model</i>	44
<i>Slika 6.7. Primjeri slika izvrsne kvalitete generiranih predloženom DCGAN arhitekturom ..</i>	44
<i>Slika 6.8. Primjeri slika loše kvalitete generiranih predloženom DCGAN arhitekturom</i>	45
<i>Slika 6.9. Prikaz rezultata performansi GAN modela s rezolucijom 64x64 piksela</i>	46
<i>Slika 6.10. Prikaz izvođenja eksperimentalnog dijela projekta</i>	47

<i>Slika 6.11. Ukupna točnost klasifikacije s različitim omjerima umjetnog proširenja podataka</i>	48
<i>Slika 6.12. Performanse F1-rezultata po klasama uz različite omjere umjetnog proširenja podataka</i>	49
<i>Slika 6.13. Utjecaj različitih omjera umjetnog proširenja podataka na postotak lažno pozitivnih uzoraka po klasama</i>	51
<i>Slika 6.14. Utjecaj različitih omjera umjetnog proširenja podataka na postotak lažno negativnih uzoraka po klasama</i>	51
<i>Slika 6.15. Prikaz promjene makro-usrednjениh pogrješaka klasifikacije prema omjerima proširenja podataka</i>	52
<i>Slika 6.16. Prikaz alternativnog pristupa izvođenja eksperimenta</i>	53

Popis tablica

<i>Tablica 2.1. Pregled korištenih metoda i rezultata iz srodne literature</i>	7
<i>Tablica 5.1. Pregled korištenih hiperparametara modela</i>	36
<i>Tablica 6.1. Izvješće klasifikacije EfficientNetB0 modela (odmrzavanje svih slojeva mreže u fazi finog podešavanja)</i>	41