

SVEUČILIŠTE U SPLITU
PRIRODOSLOVNO MATEMATIČKI FAKULTET

DIPLOMSKI RAD

**PRIMJENA GENERATIVNIH SUPARNIČKIH
MREŽA U PROŠIRENJU MEDICINSKIH
SKUPOVA PODATAKA**

Lucija Dumančić

Split, rujan 2025.

Temeljna dokumentacijska kartica

Diplomski rad

Sveučilište u Splitu
Prirodoslovno-matematički fakultet
Odjel za informatiku
Studijski program: Podatkovna znanost i inženjerstvo

PRIMJENA GENERATIVNIH SUPARNIČKIH MREŽA U PROŠIRENJU MEDICINSKIH SKUPOVA PODATAKA

Lucija Dumančić

SAŽETAK

Ovaj rad istražuje primjenu generativnih suparničkih mreža za unaprjeđenje klasifikacije medicinskih slika u uvjetima ograničene količine podataka, s posebnim naglaskom na detekciju upale pluća na rendgenskim snimkama prsnog koša. Razvijena duboka konvolucijska generativna suparnička mreža korištena je za generiranje sintetičkih slika koje su uključene u proces obuke klasifikatora temeljenog na konvolucijskoj neuronskoj mreži. Eksperimenti provedeni na skupovima podataka različitih veličina pokazali su da proširenje podataka GAN-om značajno poboljšava uspješnost klasifikacije i konzistentnost obuke u uvjetima ograničene količine podataka, smanjujući broj lažno pozitivnih predikcija i stabilizirajući proces učenja. Prednosti se smanjuju s većim skupovima podataka, što naglašava potrebu za pažljivim podešavanjem strategija proširenja.

Ključne riječi: duboko učenje, konvolucijska neuronska mreža, rendgenske snimke prsnog koša, klasifikacija upale pluća, generiranje sintetičkih slika

Rad je pohranjen u repozitoriju Prirodoslovno-matematičkog fakulteta u Splitu, koji je dio nacionalnog repozitorijskog sustava Dabar.

Rad sadrži: 62 stranice, 32 grafička prikaza, 3 tablice i 27 literaturnih navoda. Izvornik je na hrvatskom jeziku.

Mentor: izv. prof. dr. sc. Goran Zaharija, izvanredni profesor Prirodoslovno-matematičkog fakulteta, Sveučilišta u Splitu

Ocjenjivači: izv. prof. dr. sc. Goran Zaharija, izvanredni profesor Prirodoslovno-matematičkog fakulteta, Sveučilišta u Splitu

prof. dr. sc. Saša Mladenović, redoviti profesor Prirodoslovno-matematičkog fakulteta, Sveučilišta u Splitu

doc. dr. sc. Monika Mladenović, docent Prirodoslovno-matematičkog fakulteta, Sveučilišta u Splitu

Rad prihvaćen: rujan 2025.

Basic documentation card

Thesis

University of Split
Faculty of Science
Department of Computer Science
Study programme: Data Science and Engineering

APPLICATION OF GENERATIVE ADVERSARIAL NETWORKS IN MEDICAL DATASET AUGMENTATION

Lucija Dumančić

ABSTRACT

This thesis investigates the application of generative adversarial networks for improving the classification of medical images in conditions of limited amount of data, with special emphasis on the detection of pneumonia on chest X-rays. The developed deep convolutional generative adversarial network was used to generate synthetic images that were included in the training process of a classifier based on a convolutional neural network. Experiments conducted on data sets of different sizes showed that data augmentation with GAN significantly improves classification performance and training consistency in conditions of limited amount of data, reducing the number of false positive predictions and stabilizing the learning process. The benefits diminish with larger datasets, highlighting the need for careful tuning of augmentation strategies.

Key words: deep learning, convolutional neural networks, chest x-ray imaging, pneumonia classification, synthetic data generation

The thesis is deposited in the repository of the Faculty of Science, University of Split, which is part of the national repository system, Dabar.

Thesis consists of: 62 pages, 32 figures, 3 tables and 27 references

Original language: Croatian

Mentor: **Goran Zaharija, Ph.D.** *Associate Professor of Faculty of Science, University of Split*

Reviewers: **Goran Zaharija, Ph.D.** *Associate Professor of Faculty of Science, University of Split*

Saša Mladenović, Ph.D. *Full Professor of Faculty of Science, University of Split*

Monika Mladenović, Ph.D. *Assistant Professor of Faculty of Science, University of Split*

Thesis accepted: **September 2025.**

IZJAVA

kojom izjavljujem s punom materijalnom i moralnom odgovornošću da sam diplomski rad s naslovom PRIMJENA GENERATIVNIH SUPARNIČKIH MREŽA U PROŠIRENJU MEDICINSKIH SKUPOVA PODATAKA izradila samostalno pod mentorstvom izv. prof. dr. sc. Gorana Zaharije. U radu sam primijenila metodologiju znanstvenoistraživačkog rada i koristila literaturu koja je navedena na kraju diplomskog rada. Tuđe spoznaje, stavove, zaključke, teorije i zakonitosti koje sam izravno ili parafrazirajući navela u diplomskom radu na uobičajen, standardan način citirala sam i povezala s fusnotama s korištenim bibliografskim jedinicama. Rad je pisan u duhu hrvatskog jezika.

Studentica

Lucija Dumančić

Sadržaj

Uvod	1
1. Pregled istraženosti područja	2
1.1. Strategija pretraživanja	2
1.2. Meta-analiza	4
1.2.1. Skup podataka	7
1.2.2. Pretprocesiranje podataka	8
1.2.3. Klasifikacijski model i evaluacijske metrike	9
1.2.4. Vrsta generativne suparničke mreže	10
1.2.5. Rezultati	11
2. Primjena u medicini	14
2.1. Identifikacija upale pluća	15
3. Konvolucijska neuronska mreža	17
4. Proširenje podataka	21
4.1. Tradicionalne tehnike proširenja podataka	21
4.1.1. Geometrijske transformacije	22
4.1.2. Fotometrijske transformacije	27
4.1.3. Konvolucijski filtri	29
4.1.4. Umetanje šuma	29
4.1.5. Nasumično brisanje	31
4.2. Tehnike proširenja podataka temeljene na dubokom učenju	32
4.2.1. Adverzalno učenje	33
4.2.2. Generativna suparnička mreža	34
4.2.3. Duboka konvolucijska generativna suparnička mreža	35
4.2.4. Neuronski prijenos stila	36
4.2.5. Meta-učenje strategija proširenja podataka	37

5.	Metodologija.....	39
5.1.	Skup podataka	39
5.2.	Podjela skupa podataka	40
5.3.	Pretprocesiranje podataka.....	40
5.4.	Klasifikacija slika	42
5.5.	Generativni model	42
5.6.	Eksperimentalni dizajn	46
6.	Rezultati.....	48
	Zaključak	58
	Literatura	59
	Skraćenice.....	62

Uvod

Ograničena dostupnost medicinskih skupova podataka i njihova neuravnoteženost znatno otežavaju razvoj pouzdanih modela dubokog učenja za analizu medicinskih slika, što otežava postizanje visoke razine performansi i sposobnosti generalizacije modela. Ručno označavanje slika zahtijeva mnogo vremena, visoku razinu stručnosti i često dovodi do nedosljednosti među različitim modalitetima snimanja, čime se dodatno usporava proces razvoja sustava. Budući da se patološki slučajevi u praksi pojavljuju rjeđe od zdravih, mnogi su skupovi podataka neuravnoteženi, što predstavlja dodatnu prepreku u izradi kvalitetnih modela. Kako bi se prevladali ovi izazovi, istraživači sve više primjenjuju napredne tehnike proširenja podataka, osobito generativne suparničke mreže (engl. *Generative Adversarial Networks*, skraćeno GAN), koje omogućuju stvaranje novih, raznolikih i realističnih uzoraka izvan okvira postojećih podataka.

U ovom radu istražena je primjena sintetičkih rendgenskih snimaka prsnog koša generiranih GAN mrežama za poboljšanje klasifikacije upale pluća pomoću konvolucijske neuronske mreže (engl. *Convolutional Neural Network*, skraćeno CNN). Razvijena je duboka konvolucijska generativna suparnička mreža (engl. *Deep Convolutional Generative Adversarial Network*, skraćeno DCGAN) za generiranje sintetičkih podataka za obuku, koji su zatim uključeni u proces obuke klasifikatora. Eksperimenti su provedeni na skupovima različitih veličina kako bi se sustavno analizirale prednosti i ograničenja proširenja podataka GAN-om pri različitim razinama dostupnosti podataka. Performanse modela vrednovane su pomoću više metrika, uključujući točnost, preciznost, odziv (engl. *recall*) i F1 mjeru (engl. *F1 score*), uz dodatne analize distribucije pogrešaka i stabilnosti obuke. Rezultati pokazuju da proširenje podataka GAN-om značajno poboljšava performanse i konzistentnost obuke kod manjih skupova podataka, dok se te prednosti smanjuju kako veličina skupa raste.

1. Pregled istraženosti područja

Sustavni pregled je strukturiran i sveobuhvatan pristup pregledu postojeće literature o određenom istraživačkom pitanju, s ciljem minimiziranja pristranosti korištenjem unaprijed definiranih kriterija uključivanja i isključivanja za odabir relevantnih studija. Razlikuje se od tradicionalnih pregleda po tome što sustavno identificira, procjenjuje i sintetizira sve dostupne radove kako bi pružio sveobuhvatan i nepristran sažetak postojeće literature. Meta – analiza je statistička metoda koja se često koristi unutar sustavnih pregleda za kvantitativno kombiniranje rezultata više studija (Nightingale, 2009).

1.1. Strategija pretraživanja

Kako bi se odabrali relevantni radovi za ovo istraživanje koristi se citatna baza *Web of Science* (skraćeno WoS). Navedena citatna baza ima i programsku podršku koja omogućuje jednostavno pretraživanje baze znanstvenih članaka prema različitim kriterijima. Kriteriji pretraživanja definirani su uz poštivanje specifičnih termina korištenih u području istraživanja. Tako je citatna baza pretraživana pod sljedećim upitom: („GAN“ OR „Generative Adversarial Network*“) AND („medical imag*“) AND („augmentation“) (Topic). S početnim upitom WoS baza kao rezultat prikazuje 220 radova u fazi identificiranja (engl. *identification phase*). Prelaskom u sljedeću fazu, fazu pregledavanja (engl. *screening phase*), pregledani su naslovi i sažeci svih 220 radova. Primijenjeni su kriteriji isključivanja koji su prikazani u Tablica 1.

Tablica 1 Kriteriji isključivanja

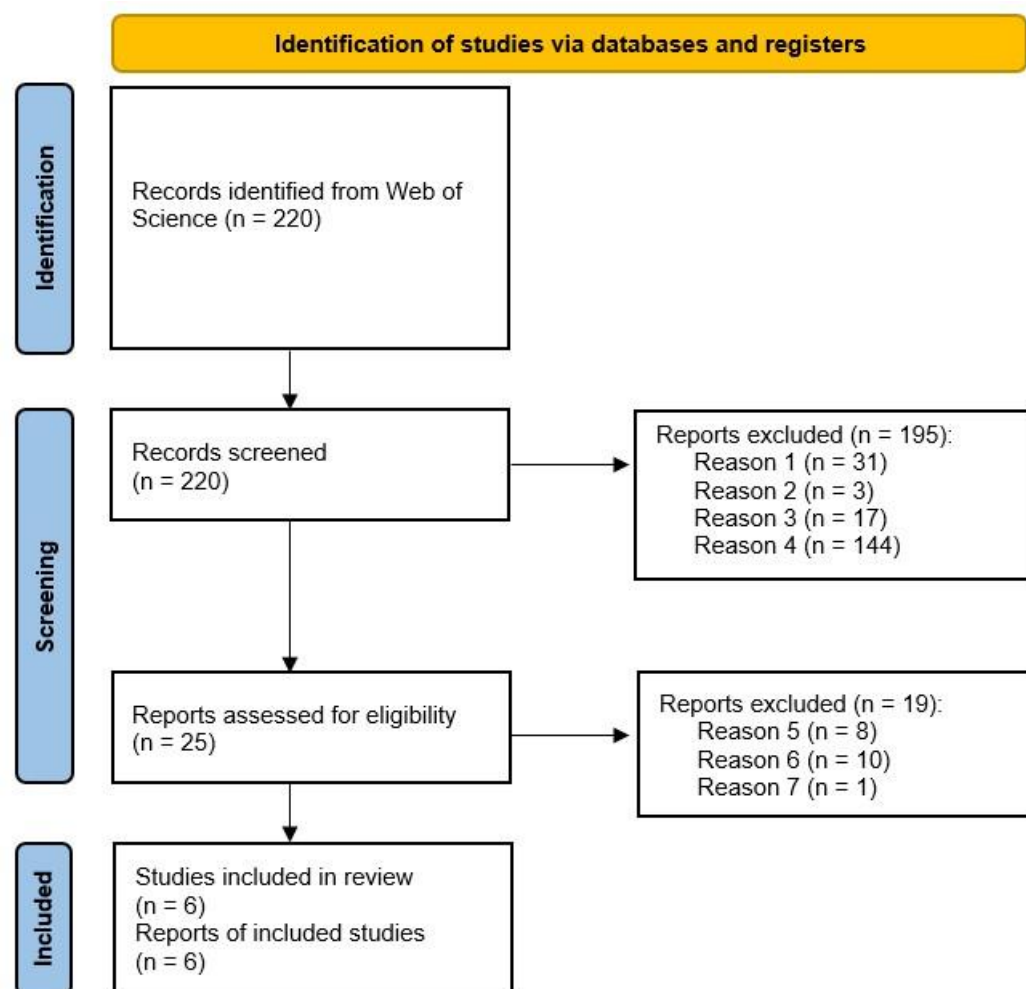
	<i>Kriterij isključivanja</i>
1	Članak je pregledni rad
2	Sadržaj nije vezan uz medicinu
3	Opće medicinske slike
4	Sadržaj nije povezan s rendgenom prsnog koša
5	Nema pristupa punom tekstu članka

6	Fokus nije na binarnoj klasifikaciji upale pluća i normalnih rendgenskih slika prsnog koša
7	Fokus nije na proširenju podataka

U ovoj fazi 31 rad je odbačen jer istraživanje predstavlja pregled postojećih radova, a ne primarnu studiju. Isključeni su radovi čija tema rada nije povezana s medicinskim istraživanjima ili primjenama, a takva su bila 3 rada. Radovi koji se bave medicinskim slikama općenito, a ne slikama prsnog koša, rezultirali su odbacivanjem 17 radova. Najviše radova, točnije 144 rada, odbačeno je iz razloga jer nisu upotrebljavali rendgenske slike prsnog koša. S time je u fazi pregledavanja isključeno ukupno 195 radova. Preostalih 25 radova je detaljno analizirano u fazi podobnosti (engl. *eligibility phase*). Uslijed nemogućnosti pristupa punom tekstu rada, odbačeno je 8 radova. U ovoj fazi najčešći razlog isključivanja radova je taj jer fokus rada nije na binarnoj klasifikaciji rendgenskih slika prsnog koša koje prikazuju upalu pluća i normalnih rendgenskih slika prsnog koša, pri čemu je odbačeno 10 takvih radova. Jedan rad je odbačen jer njegov fokus nije na proširenju podataka. Sustavni pregled literature rezultirao je identifikacijom 6 radova koji udovoljavaju definiranim kriterijima. Pregled tih kriterija dan je u Tablica 2. Cjelokupni proces odabira i filtriranja radova sistematično je prikazan PRISMA (engl. *Preferred Reporting Items for Systematic Reviews and Meta-Analyses*) dijagramom toka (Slika 1).

Tablica 2 Kriteriji uključivanja

	<i>Kriterij uključivanja</i>
1	Vrsta dokumenta je znanstveni rad ili konferencijski članak
2	Glavni fokus je na primjeni GAN-a za proširenje podataka rendgenskih slika prsnog koša za klasifikaciju upale pluća i normalnih slika



Slika 1 PRISMA dijagram toka

1.2. Meta-analiza

Identificirani radovi analizirani su kroz nekoliko ključnih aspekata. Usporedba je provedena prema korištenim skupovima podataka, postupcima pretprocesiranja, odabranim klasifikacijskim modelima i evaluacijskim metrikama, kao i prema primijenjenim GAN arhitekturama te ostvarenim rezultatima. Sažeti prikaz ovih elemenata prikazan je u Tablica 3.

Tablica 3 Meta-analiza

<i>Referenca</i>	<i>Godina</i>	<i>Skup podataka i veličina skupa</i>	<i>Tehnika proširenja podataka</i>	<i>GAN arhitektura</i>	<i>Točnost</i>
(Bhagat et al., 2019)	2019	Chest X-ray Images (Pneumonia) 5856 slika (1024×1024×3)	Tradicionalne tehnike GAN	ProGAN	91%
(Ganesan et al., 2019)	2019	RSNA Pneumonia Detection Challenge 20607 slika (512×512)	Bez proširenja podataka Tradicionalne tehnike GAN	PG-GAN	83,94%
(Poola et al., 2024)	2024	COVID-19 Radiography Database 13808 slika (229×229×1)	Bez proširenja podataka GAN	CE-GAN	97,6%
(Qin et al., 2022)	2022	Chest X-ray Images (Pneumonia) 5856 slika RSNA Pneumonia Detecetion Challenge 26684 slika (256×256×1)	Klasične geometrijske transformacije GAN	DCGAN PGGAN	91,7%

(Sashank et al., 2022)	2022	COVID-19 Radiography Database 2680 slika (128×128×1)	Bez proširenja podataka	Simple GAN	97,86%
			Tradicionalne tehnike	Transfer Learning	
			GAN	GAN	
(Wang et al., 2021)	2021	RSNA Pneumonia Detecetion Challenge	GAN	CycleGAN	86,7%
		25282 slika (224×224×1)			

1.2.1. Skup podataka

Ključna komponenta istraživanja analize medicinskih slika je izbor skupa podataka, jer izravno utječe na pouzdanost, generalizaciju i usporedivost razvijenih modela. U šest pregledanih radova pretežno su korištena tri javno dostupna skupa podataka rendgenskih snimaka prsnog koša:

1. Chest X-ray Images (Pneumonia) (Mooney, 2018)
2. RSNA Pneumonia Detection Challenge (RSNA Pneumonia Detection Challenge Dataset, 2018)
3. COVID-19 Radiography Database (Tawsifur et al., 2020)

Prvi skup podataka se široko koristi u zadacima klasifikacije upale pluća, izvorno je sastavljen od pedijatrijskih pacijenata u *Guangzhou Women and Children's Medical Center*. Sadrži rendgenske snimke kategorizirane u dvije klase (*normal*, *pneumonia*), s oznakama koje su provjerili stručni radiolozi. Sastoji se od 4273 slike upale pluća i 1583 normalne slike. (Bhagat et al., 2019) i (Qin et al., 2022) koriste ovaj skup podataka u svojim istraživanjima.

Drugi skup podataka je razvijen kao dio natjecanja koje su zajednički organizirali Radiološko društvo Sjeverne Amerike (engl. *Radiological Society of North America*) i Društvo torakalne radiologije (engl. *Society of Thoracic Radiology*). Predstavlja podskup kolekcije *NIH ChestX-ray14* i sastoji se od preko 25000 slika. Skup podataka je označen za prisutnost zamućenja pluća, često povezanih s upalom pluća, te je stoga prikladan i za klasifikaciju i za zadatke detekcije. (Ganesan et al., 2019) koristi 8954 normalne i 11653 abnormalne slike, kasnije podijeljene u skupove za obuku i testiranje. (Qin et al., 2022) također koristi ovaj skup podataka, restrukturirajući ga u binarne klasifikacijske kategorije (normalna, neprozirnost pluća) i odbacujući dvosmislene ili nekvalitetne uzorke. (Wang et al., 2021) je usvojio ovaj skup podataka, naglašavajući njegovu inherentnu neravnotežu, jer skup za obuku uključuje 5137 slučajeva pozitivnih na upalu pluća i 19905 slučajeva negativnih na upalu pluća.

Dva rada koriste treći skup podataka koji je dizajniran za otkrivanje upale pluća povezane s COVID-19. Ovaj skup podataka sastoji se od rendgenskih snimaka grupiranih u kategorije normalne i virusne upale pluća (uključujući COVID-19). U (Poola et al., 2024) radu, korišteno je 10192 normalnih slučajeva i 3616 slučajeva upale pluća, koji su kasnije

podijeljeni u skupove za obuku i testiranje. (Sashank et al., 2022) je, s druge strane, koristio ograničeni podskup, točnije 1340 normalnih slika i 1340 slika virusne upale pluća.

1.2.2. Pretprocesiranje podataka

Prije obuke modela, (Bhagat et al., 2019) je standardizirao sve rendgenske snimke prsnog koša promjenom veličine na ujednačenu rezoluciju od 1024×1024 piksela. Kako bi se poboljšala varijabilnost skupa podataka i ublažilo prekomjerno prilagođavanje, u radu su korištene tradicionalne tehnike proširenja podataka. To je uključivalo afine transformacije, poput rotacije slike, horizontalnog okretanja i dodavanja Gaussovog šuma, kao i geometrijske transformacije usmjerene na poboljšanje vizualne jasnoće, konkretno, primijenjeno je poboljšanje kontrasta, izoštravanje slike i izjednačavanje histograma kako bi se istaknule strukturne razlike unutar rendgenskih snimaka i povećao ukupni kontrast slike.

(Ganesan et al., 2019) pristup pretprocesiranju bio je usmjeren na izoliranje dijagnostički relevantnih područja rendgenskih snimaka prsnog koša. Polja pluća prvo su segmentirana pomoću *dropout UNET* modela kako bi se uklonile suvišne anatomske strukture. Nakon segmentacije, svaka je slika izrezana na granični okvir koji obuhvaća pluća, čime je analiza ograničena na područje od interesa. Na kraju su sve slike promijenjene na standardiziranu rezoluciju od 512×512 piksela, čime je osigurana ujednačenost unutar skupa podataka za obuku modela.

Kako bi se standardizirali ulazni podaci i smanjili računalni zahtjevi, u (Poola et al., 2024) je provedeno nekoliko koraka pretprocesiranja. Prvo su sve rendgenske snimke prsnog koša, koje su izvorno varirale u rezoluciji od 229×229 do 1024×1024 piksela, promijenjene na jedinstvenu dimenziju od $229 \times 229 \times 1$ piksela. Zatim su sve slike pretvorene u sive tonove (engl. *grayscale*) kako bi se pojednostavilo izdvajanje značajki. Na kraju su vrijednosti piksela normalizirane na ujednačen raspon, čime je osigurana stabilna i učinkovita obuka modela.

(Qin et al., 2022) je rješio problem heterogenosti dvaju korištenih skupova podataka tako što je prvo standardizirao i normalizirao vrijednosti piksela. Sve slike su promijenjene na rezoluciju od 256×256 piksela pomoću bilinearne interpolacije, uz dodatno centrirano izrezivanje radi naglašavanja područja pluća. Za arhitekture poput *Inception v3*, ulazne slike dodatno su promijenjene na $299 \times 299 \times 3$, dok su druge mreže dobivale slike veličine

256×256×1. Kako bi se ublažila neravnoteža među klasama u skupovima za obuku, primijenjene su geometrijske tehnike proširenja podataka, uključujući nasumičnu rotaciju, horizontalno zrcaljenje, horizontalno i vertikalno pomicanje na manjinskoj klasi.

(Sashank et al., 2022) je primijenio jednostavan korak pretprocesiranja usmjeren na smanjenje računalnih troškova. Sve rendgenske snimke prsnog koša, koje su izvorno bile veličine 1024×1024×1, smanjene su na 128×128×1 piksela, čime su ulazne dimenzije standardizirane uz istodobno skraćivanje vremena obuke i smanjenje potreba za resursima.

(Wang et al., 2021) je proveo višestupanjski postupak pretprocesiranja osmišljen za poboljšanje kvalitete ulaza i standardizaciju podataka. Prvo su slike razvrstane u kategorije visoke i niske kvalitete pomoću klasifikatora *ResNet-18*. Sve su slike zatim promijenjene na rezoluciju od 224×224 piksela u sivim tonovima radi dosljednosti i računalne učinkovitosti. Slike niske kvalitete dodatno su poboljšane putem *CycleGAN*-a, primijenjenog zasebno za klase pozitivne i negativne na upalu pluća, osiguravajući očuvanje dijagnostičkih značajki. Poboljšane slike zamijenile su izvorne uzorke niske kvalitete, čime se povećala kvaliteta skupa podataka i podržala pouzdanija obuka modela.

1.2.3. Klasifikacijski model i evaluacijske metrike

Pregledani radovi koristili su niz modela strojnog i dubokog učenja za klasifikaciju upale pluća i abnormalnosti, s odgovarajućim evaluacijskim metrikama koje odražavaju njihovu dijagnostičku učinkovitost.

(Bhagat et al., 2019) je koristio varijantu AlexNet duboke konvolucijske neuronske mreže (engl. *deep convolution neural network*, skraćeno DCNN) za binarnu klasifikaciju rendgenskih snimaka prsnog koša. Učinkovitost modela ocijenjena je korištenjem standardnih metrika, uključujući točnost, preciznost, odziv i F1 mjeru.

(Ganesan et al., 2019) je implementirao prilagođeni, prethodno obučeni VGG16 model za klasifikaciju abnormalnosti na rendgenskim snimkama prsnog koša. Evaluacija je provedena pomoću točnosti, F1 mjere, Matthewovog koeficijenta korelacije (engl. *Matthew's correlation coefficient*) i područja ispod krivulje (engl. *area under the curve*), pružajući sveobuhvatnu procjenu robusnosti modela i prediktivne moći.

(Poola et al., 2024) je primijenio komparativni pristup, uključujući stroj potpornih vektora (engl. *support vector machines*, skraćeno SVM), *k-nearest neighbors*, neuronske mreže, *Gaussian Naive Bayes* (skraćeno GNB) i logističku regresiju, na matrici proširenih značajki

dobivenoj iz sintetičkih i izvornih slika. Kod evaluacije modela korištena je točnost, preciznost, odziv, F1 mjera, područje ispod ROC krivulje (engl. *Area under the ROC Curve*, skraćeno AUC-ROC) uz mjerenje indeksa strukturne sličnosti (engl. *structural similarity index*) za usporedbu kvalitete sintetičkih i autentičnih slika.

(Qin et al., 2022) je istraživao opsežan skup arhitektura dubokog učenja, uključujući *multi-scale* konvolucijsku neuronsku mrežu, standardnu konvolucijsku mrežu, *ResNet*, *DenseNet*, *EfficientNet* i okvire prijenosnog učenja. Evaluacijske metrike uključivale su točnost, odziv, preciznost, F1 mjeru, specifičnost, *G-mean*, *AUC* i *Area under precision-recall curve*, odražavajući kako razinu performansi po klasama, tako i ukupnu prediktivnu učinkovitost.

(Sashank et al., 2022) je koristio standardnu konvolucijsku neuronsku mrežu za klasifikaciju, s evaluacijom ograničenom na točnost, naglašavajući ukupnu ispravnost predikcija.

(Wang et al., 2021) je koristio više *ResNet-18* mreža za dijagnozu, usklađenih s kategorijama kvalitete rendgenskih snimaka prsnog koša, temeljenim na prethodnoj obradi. Odvojene mreže su obučene na skupovima slika visoke kvalitete, niske kvalitete, poboljšane kvalitete i kombiniranih skupova slika kako bi se procijenio utjecaj poboljšanja slike na klasifikaciju. Performanse su evaluirane pomoću točnosti, specifičnosti, odziva, preciznosti, F1 mjere i AUC-ROC, omogućujući usporedbu osnovne i poboljšane kvalitete slike.

1.2.4. Vrsta generativne suparničke mreže

Pregledani radovi koristili su niz GAN arhitektura, što odražava razlike u istraživačkim ciljevima i karakteristikama skupova podataka. Zajednička tema im je oslanjanje na napredne GAN varijante za generiranje visokokvalitetnih sintetičkih rendgenskih snimaka prsnog koša ili za poboljšanje postojećih skupova podataka, pri čemu svaki model nudi različite prednosti.

Najčešće korišteni bili su *Progressive Growing GAN* (skraćeno ProGAN ili PGGAN), što se vidi u radovima (Bhagat et al., 2019), (Ganesan et al., 2019) i (Qin et al., 2022). Ovi modeli postupno povećavaju razlučivost slika tijekom obuke, čime se stabilizira proces učenja i ublažavaju problemi poput kolapsa modova (engl. *mode collapse*). (Bhagat et al., 2019) je primijenio ProGAN za sintezu visokorezolucijskih slika koje su podržavale naknadne klasifikacijske zadatke, dok su (Ganesan et al., 2019) i (Qin et al., 2022) posebno koristili PGGAN, pokazujući njegovu sposobnost generiranja detaljnih, dijagnostički značajnih rendgenskih snimaka pri razlučivostima do 512×512 piksela.

Nasuprot tome, (Poola et al., 2024) je predstavio *Classifier-Embedded GAN* (skraćeno CE-GAN), dizajniran za generiranje sintetičkih slika uvjetovanih oznakama klasa. Integriranjem klasifikacije u generativni okvir, CE-GAN nije samo riješio izazov ograničenih medicinskih skupova podataka, već je ima za cilj i smanjenje potencijalnih pristranosti prisutnih u stvarnim kolekcijama.

(Sashank et al., 2022) je koristio jednostavniju arhitekturu GAN-a (*Simple GAN*), obučavajući odvojene modele za normalne slike i slike s upalom pluća. Strategija prijenosnog učenja (engl. *transfer learning*) primijenjena je prethodnim obučavanjem GAN-a na velikom, nepovezanom skupu podataka (slike lica), koji je zatim fino podešen (engl. *fine-tuned*) na medicinskim podacima. Ovaj pristup ilustrira praktično rješenje za situacije u kojima dostupni medicinski skupovi nisu dovoljni za obučavanje GAN-a od početka.

(Wang et al., 2021) koristi *CycleGAN*, generativnu suparničku mrežu dizajniranu za nesporeno prevođenje slike u sliku, kako bi se poboljšale rendgenske snimke prsnog koša niske kvalitete pacijenata s upalom pluća. CycleGAN se sastoji od dviju GAN mreža koje provode gubitak konzistentnosti ciklusa (engl. *cycle-consistency loss*), osiguravajući da se slike prevedene u ciljnu domenu mogu mapirati natrag u njihovu izvornu domenu bez gubitka ključnih značajki. Dva odvojena CycleGAN modela obučena su za poboljšanje slika niske kvalitete iz klasa pozitivnih i negativnih na upalu pluća, čuvajući detalje povezane s bolešću i istovremeno poboljšavajući ukupnu kvalitetu slike. Ovaj pristup učinkovito poboljšava interpretativne značajke u rendgenskim snimkama niske kvalitete, značajno poboljšavajući točnost dijagnoze upale pluća.

Ova istraživanja ističu svestranost GAN-ova u istraživanju medicinskog snimanja. Dok su ProGAN/PGGAN-ovi bili preferirani za generiranje sintetičkih uzoraka visoke rezolucije, alternativni modeli poput CE-GAN-a, jednostavnih GAN-ova s prijenosnim učenjem i CycleGAN-ova bavili su se specifičnim izazovima, uključujući uvjetovanje klasa, nedostatak podataka i poboljšanje kvalitete slike.

1.2.5. Rezultati

(Bhagat et al., 2019) je pokazao da proširenje skupova podataka slikama generiranim GAN-om znatno poboljšava točnost klasifikacije. Obuka DCNN-a na proširenom skupu podataka rezultirala je s 91% točnosti, 89,4% preciznosti, 92,8% odziva i F1 mjerom od 91,06%, u usporedbi sa samo 79,5% točnosti na izvornom skupu podataka. To je potvrdilo da

integriranje sintetičkih slika sa stvarnim podacima značajno poboljšava performanse modela, iako je daljnje usavršavanje klasifikatora i dalje potrebno za kliničku pouzdanost.

(Ganesan et al., 2019) je utvrdio da proširenje podataka dosljedno poboljšava performanse klasifikatora u usporedbi s obukom bez proširenja. Performanse klasifikacijskog modela analizirane su u tri scenarija: bez proširenja podataka, tradicionalne tehnike proširenja podataka i GAN proširenje podataka. Transformacije primijenjene u sklopu tradicionalnog proširenja podataka uključivanje su Gaussovo zaglađivanje (engl. *Gaussian smoothing*), neizoštreno maskiranje (engl. *unsharp masking*) i minimalno filtriranje (engl. *minimum filtering*). U oba tipa proširenja podataka, slike su dodavane u postotcima od 25%, 50%, 75% i 100% ukupnog broja uzoraka u skupu za obuku. Rezultati pokazuju da su i tradicionalne tehnike proširenja podataka i GAN proširenje značajno poboljšale performanse modela u odnosu na scenarij bez proširenja podataka, čija točnost iznosi 79,32%. Zanimljivo je da je tradicionalno proširenje pokazalo bolje rezultate u usporedbi s GAN proširenjem, pri čemu je porast performansi bio izraženiji kako se povećavao udio sintetičkih slika. Na primjer, sa 75% dodanih sintetičkih slika, točnost s tradicionalnim proširenjem raste na 82,23%, dok s GAN proširenjem raste na 80,41%. Na 100% sintetičkih slika, točnost s tradicionalnim proširenjem iznosi 84,25%, dok s GAN proširenjem iznosi 83,94%. Autori su to pripisali GAN-ovima koji generiraju ograničenu varijancu, dok je tradicionalno proširenje uvelo biološki značajnu raznolikost.

(Poola et al., 2024) je predstavio okvir *SyntDiaNet*, kombinirajući generiranje sintetičkih podataka temeljenih na CE-GAN-u s naprednom ekstrakcijom značajki (npr. *Contour Detection*, *Gabor Texture Analysis*, *Blob Detection*). Kako bi se istražio učinak generiranih sintetičkih slika na klasifikaciju, analizirana su dva pristupa: sintetičke slike generirane CE-GAN-om i stvarni podaci bez proširenja. Rezultati pokazuju poboljšanja u performansama modela uz primjenu CE-GAN-a, pri čemu su postignute točnosti između 90,5% (GNB) i 97,6% (SVM), preciznost do 0,981, odziv do 0,987 te F1 mjeru do 0,980. Osim dijagnostičke točnosti, *SyntDiaNet* je istaknuto snažan potencijal za generalizaciju, interpretabilnost i primjenjivost na širi spektar zadataka medicinskog snimanja, iako i dalje postoje izazovi vezani uz kliničku validaciju i računalnu učinkovitost.

Kako bi se istražilo ima li kvaliteta slika generiranih GAN-ovima značajan utjecaj na klasifikacijske modele, (Qin et al., 2022) analizira tri scenarija proširenih podataka: sintetičke slike generirane PGGAN-om, slike generirane DCGAN-om te slike nastale primjenom klasičnih geometrijskih transformacija. Rezultati pokazuju da slike generirane

PGGAN-om uvijek imaju višu kvalitetu od onih generiranih DCGAN-om. Povedena ispitivanja na većem skupu podataka potvrđuju široku primjenjivost metode. Proširenje podataka pomoću GAN-ova poboljšalo je klasifikaciju za modele: CNN, *ResNet*, *DenseNet* i *EfficientNet*. Međutim, poboljšanje performansi *Transfer Learning* nije bilo značajno, s povećanom točnošću s 87,9% na 88,2%, pretpostavljajući zbog razlika između skupa podataka korištenog za unaprijed uvježbanu mrežu i ciljnog skupa podataka. Iako PGGAN zahtijeva više računalnih resursa i duže vrijeme obuke u usporedbi s metodama poput DCGAN-a i geometrijskih transformacija, rezultati pokazuju da slike više kvalitete obično doprinose boljim klasifikacijskim rezultatima. Tako je s DCGAN-om postignuta točnost od 91,2%, dok je s PGGAN-om postignuto 91,7%, a klasične geometrijske transformacije su postigle 90%.

(Sashank et al., 2022) analizira performanse klasifikacijskog modela uspoređujući rezultate na izvornom skupu podataka, skupu generiranom tradicionalnim tehnikama proširenja podataka, skupu koji kombinira izvorne podatke s podacima generiranim Simple GAN-om te skupu koji kombinira izvorne podatke s podacima generiranim *Transfer Learning GAN-om*. Pokazali su da je tradicionalno proširenje podataka imalo minimalne koristi, dok je proširenje podataka temeljeno na GAN-u donijelo značajna poboljšanja. Obuka CNN-a s podacima proširenim Simple GAN-om postigla je 97,7% točnost na skupu za obuku i 100% na validacijskom skupu, nadmašivši modele obučene izvornim (93,48% točnost na skupu za obuku i 92,59% točnost na validacijskom skupu) ili tradicionalno proširenim podacima (90,13% točnost na skupu za obuku i 92,31% točnost na validacijskom skupu). Transfer Learning GAN također je poboljšao rezultate sa 97,86% točnosti na skupu za obuku i 95,74% točnosti na validacijskom skupu, pokazujući znakove prekomjernog prilagođavanja (engl. *overfitting*).

(Wang et al., 2021) je pokazao da je poboljšanje slike temeljeno na CycleGan-u značajno unaprijedilo performanse modela. Mreže obučene na poboljšanim rendgenskim snimkama (*Res-He*, *Res-E*) nadmašile su one obučene na niskokvalitetnim ili čak izvornim visokokvalitetnim slikama, s dobitima od 15,8% u točnosti i 19,4% u F1 mjeri. Rezultati ističu sposobnost Cycle-GAN-a da poboljša kvalitetu slike uz očuvanje dijagnostičkih značajki, nudeći ovi pristup za unapređenje klasifikacije rendgenskih snimki prsnog koša bez potrebe za proširenjem skupova podataka.

2. Primjena u medicini

Tehnike medicinskog snimanja od presudne su važnosti za dijagnostiku bolesti jer omogućuju ne samo prepoznavanje poremećaja, već u pojedinim slučajevima i predviđanje njihova napredovanja. Modaliteti snimanja poput rendgenskih snimaka, CT-a, magnetne rezonance i ultrazvuka predstavljaju temelj zdravstvene skrbi te doprinose ranom otkrivanju bolesti, preciznoj dijagnozi, planiranju liječenja i kontinuiranom praćenju. Omogućuju zdravstvenim djelatnicima vizualizaciju unutarnjih struktura s izvanrednom jasnoćom, olakšavajući identifikaciju stanja poput raka, kardiovaskularnih bolesti, neuroloških poremećaja i ozljeda mišićno-koštanog sustava (Abhisheka et al., 2024).

Tradicionalno su stručnjaci ručno provodili klasifikaciju i segmentaciju slika. Međutim, s brзом evolucijom slikovnih tehnologija, danas se svakodnevno proizvodi golema količina medicinskih slika. Obrada tako velikih količina podataka nije samo vremenski zahtjevana, već i zahtjeva specijalizirano znanje. Posljedično, sve je izraženija potreba za automatizacijom u ovome području. Najdjelotvorniji pristup automatizaciji analize medicinskih slika jest primjena dubokih neuronskih mreža, razvijenih obukom modela dubokog učenja. Povećanje točnosti i robusnosti takvih modela uvelike ovisi o dostupnosti dovoljno velikih, raznolikih i precizno označenih skupova podataka. Kada je riječ o klasifikaciji bolesti, osobito je važno osigurati uravnoteženu zastupljenost klasa s ispravnim oznakama. Nedostatak raznolikosti ili neuravnoteženost u skupovima za učenje može dovesti do modela s lošim sposobnostima generalizacije.

Razvoj učinkovitih modela dubokog učenja za obradu medicinskih slika suočava se s brojnim izazovima. Primarni problem predstavlja dostupnost podataka, koja je ograničena potrebom razmjene medicinskih slika između institucija uz istodobno očuvanje privatnosti pacijenata. Drugi značajan izazov jest nedostatak označenih podataka. Ručno označavanje medicinskih slika izrazito je složeno i vremenski zahtjevno te često nedosljedno među različitim modalitetima snimanja. Dodatno, mnogi su skupovi podataka inherentno neuravnoteženi, budući da se patološki slučajevi javljaju znatno rjeđe od zdravih, što dodatno otežava razvoj pouzdanih dijagnostičkih sustava.

Kako bi se prevladala ova ograničenja, istraživači su usmjerili pozornost na unaprjeđenje medicinskih skupova podataka putem proširenja podataka. Postoje dva primarna pristupa:

1. Tradicionalne tehnike proširenja podataka, poput rotacije, zrcaljenja, rezanja i prilagodbe boja
2. Tehnike proširenja podataka temeljene na dubokom učenju, uključujući generativne suparničke mreže.

Dok su tradicionalne tehnike proširenja podataka ograničene u sposobnosti proširenja raznolikosti skupa podataka, budući da se oslanjaju isključivo na transformacije postojećih slika, generativne suparničke mreže mogu generirati nove, raznolike uzorke koji nisu uvjetovani izvornim skupom podataka (Makhlouf et al., 2023).

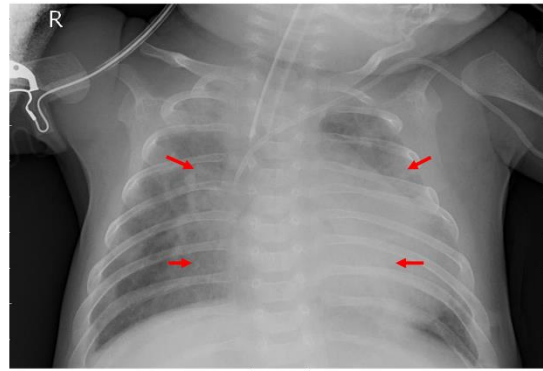
2.1. Identifikacija upale pluća

U ovom radu naglasak je na klasifikaciji upale pluća na temelju rendgenskih snimaka prsnog koša. Upala pluća, akutna respiratorna infekcija, uzrokuje 14% smrti među djecom mlađom od pet godina (Organization, 2022). Zbog visoke smrtnosti u ovoj ranjivoj populaciji, rana i točna dijagnoza od presudne je važnosti. Rendgensko snimanje prsnog koša predstavlja neinvazivno, cjenovno pristupačno i široko dostupno sredstvo za identifikaciju bolesti (Kundu et al., 2021). Učestalost i klinička ozbiljnost upale pluća potaknule su razvoj opsežnih skupova rendgenskih snimaka prsnog koša, označenih pouzdanim oznakama kojima se identificiraju slučajevi upale pluća. Takvi su skupovi osobito značajni za istraživanja jer omogućuju obuku modela dubokog učenja.

Slika 2 prikazuje primjer iz (Kundu et al., 2021), uspoređujući rendgensku snimku zdravih pluća sa snimkom na kojoj je prisutna upala pluća. Radiolozi obično postavljaju dijagnozu upale pluća prepoznavanjem bijelih infiltrata u plućima, označenih crvenim strelicama na drugoj slici, koji se smatraju karakterističnim pokazateljima bolesti. Važno je naglasiti da je takvo stručno tumačenje nužno u kliničkoj praksi, dok modeli dubokog učenja ne zahtijevaju specijalizirana stručna znanja. Konvolucijske neuronske mreže mogu učinkovito naučiti razlikovati zdrava od bolesnih pluća kada se obučavaju na dovoljno velikom broju označenih primjera.



a)



b)

Slika 2 Primjer dviju rendgenskih snimki pluća koje prikazuju: a) zdrava pluća, b) upala pluća (Kundu et al., 2021)

3. Konvolucijska neuronska mreža

Konvolucijske neuronske mreže predstavljaju temeljnu arhitekturu dubokog učenja, široko priznatu zbog iznimne sposobnosti identificiranja i izdvajanja relevantnih obilježja iz podataka, ponajprije slika, bez ljudskog nadzora (Alzubaidi et al., 2021). Razvoj CNN-ova inspiriran je strukturom životinjskog vidnog korteksa, gdje složeni niz neurona automatski obrađuje vizualne ulaze, omogućujući naprednu percepciju i kategorizaciju. Ova biološka analogija odražava se u načinu na koji CNN-ovi obrađuju ulazne podatke, oponašajući lokalnu i hijerarhijsku obradu prisutnu u prirodi. U svojoj osnovi, CNN-ovi organiziraju ulazne podatke, poput slika, u tri dimenzije: visinu, širinu i dubinu (broj kanala). Mreža se oslanja na konvolucijske slojeve koji koriste filtere, odnosno jezgre, za skeniranje ulaznih podataka i generiranje mapa značajki. Svaka jezgra sastoji se od mreže vrijednosti (težina) koje su u početku nasumično postavljene, a potom se iterativno ažuriraju tijekom učenja, čime mreža usvaja spoznaju o tome koje su značajke najrelevantnije za ciljani zadatak. Operacija konvolucije uključuje pomicanje svakog filtera preko ulaznih podataka i izračunavanje skalarnog produkta na svakoj poziciji, što rezultira mapom značajki koja ističe specifične obrasce unutar podataka.

Jedna od ključnih prednosti konvolucijskih slojeva je rijetka povezanost i dijeljenje težina. Za razliku od potpuno povezanih neuronskih mreža, u kojim je svaki neuron povezan sa svim neuronima u idućem sloju, CNN-ovi ograničavaju povezanost na lokalizirane regije. Ovo značajno smanjenje broja parametara koje je moguće uvježbavati pojednostavljuje proces učenja, povećava računalnu učinkovitost i čini mrežu robusnijom na prekomjerno prilagođavanje fokusiranjem na bitne značajke.

Nakon konvolucije, CNN arhitekture obično koriste slojeve za sažimanje (engl. *pooling layer*). Svrha ovog sloja je smanjenje uzorka mapa značajki, odnosno njihovo pojednostavljenje, uz očuvanje dominantnih informacija. Uobičajene metode, poput maksimalnog (engl. *max pooling*) ili prosječnog sažimanja (engl. *average pooling*), biraju najveću, odnosno prosječnu vrijednost u danom prostoru mape značajki. Glavni nedostatak ovog sloja očituje se u potencijalnom smanjenju ukupne učinkovitosti CNN-a. Iako omogućuje modelu prepoznavanje prisutnosti određenih značajki u ulaznoj slici, njegova je funkcionalnost primarno usmjerena na njihovu prostornu lokaciju, što može rezultirati gubitkom relevantnih informacija (Yamashita et al., 2018).

Kako bi se mreži omogućilo modeliranje složenih, nelinearnih odnosa, nakon konvolucije i sažimanja primjenjuju se aktivacijske funkcije. Iako postoji više vrsta, najčešće korištena u CNN-ovima je *ReLU* (engl. *Rectified Linear Unit*), koja transformira ulazne vrijednosti tako da negativne postavlja na nulu, dok pozitivne ostavlja nepromijenjenima. Ova jednostavna nelinearnost omogućuje CNN-ovima učenje vrlo sofisticiranih reprezentacija, jer omogućuje povratno širenje pogreške i učinkovito učenje. Varijacije *ReLU*-a, poput *Leaky ReLU* i *Parametric ReLU*, rješavaju specifične probleme poput *Dying ReLU*-a, gdje neuron može prestati s aktivacijom (Purwono et al., 2022).

U završnim fazama CNN-a, potpuno povezani slojevi (engl. *fully connected layers*) koriste se kao klasifikator koji agregira naučene značajke iz prethodnih slojeva i provodi konačnu klasifikaciju. Ovaj sloj funkcionira slično tradicionalnoj višeslojnoj perceptronskoj mreži, djelujući kao klasifikator izdvojenih značajki. Funkcija gubitka, poput unakrsne entropije kod klasifikacije, kvantificira razliku između predviđenih i stvarnih rezultata. Tijekom učenja parametri mreže se podešavaju optimizacijskim algoritmima, najčešće varijantama gradijentnog spusta, s ciljem minimizacije gubitka i poboljšanja prediktivne točnosti.

Tehnike regularizacije ključne su za suzbijanje prekomjernog prilagođavanja u CNN-ovima. Među njima se ističu *dropout* (slučajno onemogućavanje neurona tijekom učenja), *drop-weights* (slučajno uklanjanje veza između neurona), proširenje podataka (umjetno proširenje skupa podataka) i normalizacija po serijama (standardizacija izlaza svakog sloja radi stabilizacije učenja). Ove strategije osiguravaju da model dobro generalizira, odnosno da ostvaruje dobre rezultate i na podacima koji nisu viđeni tijekom učenja (Alzubaidi et al., 2021).

Kako bi se procijenila učinkovitost predloženog modela, izbor evaluacijskih metrika igra ključnu ulogu. Za probleme binarne klasifikacije, evaluacija se obično oslanja na matricu konfuzije, koja bilježi broj stvarno pozitivnih (engl. *true positive*, skraćeno TP), stvarno negativnih (engl. *true negative*, skraćeno TN), lažno pozitivnih (engl. *false positive*, skraćeno FP) i lažno negativnih (engl. *false negative*, skraćeno FN) rezultata. Iz ove matrice može se izvesti nekoliko važnih mjera koje obuhvaćaju različite aspekte učinkovitosti klasifikatora (Hossin & Sulaiman, 2015).

Najčešće korištena mjera je točnost, koja izražava ukupni udio ispravno klasificiranih instanci. Izračunava se na sljedeći način:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

Točnost je jednostavna za izračun i lako se interpretira, no može biti zavaravajuća kada je skup podataka neuravnotežen, budući da može biti pristrana prema većinskoj klasi. Kako bi se prevladalo ovo ograničenje, razmatrane su dodatne metrike.

Preciznost naglasak stavlja na pouzdanost pozitivnih predikcija, tj. koliko su predviđeni pozitivni slučajevi doista točni. Računa se sljedećom formulom:

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

Odziv, također poznat kao osjetljivost, mjeri udio stvarnih pozitivnih slučajeva koji su ispravno identificirani, na sljedeći način:

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

Preciznost i odziv su komplementarne mjere: preciznost naglašava izbjegavanje lažno pozitivnih rezultata, dok odziv naglašava izbjegavanje lažno negativnih rezultata. U praksi, poboljšanje jedne mjere često smanjuje drugu. Kako bi se oba aspekta sažela u jedinstvenu vrijednost, koristi se F1 mjera, definirana kao harmonijska sredina preciznosti i odziva:

$$F1\ score = 2 * \frac{precision*recall}{precision+recall} \quad (4)$$

Ova mjera uravnotežuje obje perspektive i osobito je korisna kada su klase neuravnotežene.

U kontekstu klasifikacije medicinskih slika, poput klasifikacije upale pluća na rendgenskim slikama prsnog koša, interpretacija ovih metrika postaje klinički značajna. Sama točnost može biti nedovoljna, jer broj zdravih slika često premašuje broj slika koje prikazuju upalu pluća, što može dovesti do zavaravajuće visoke točnosti ako model pretežno predviđa većinsku klasu. Preciznost je važna kako bi se osiguralo da predviđeni slučajevi upale pluća zaista budu točni, čime se smanjuju nepotrebne dijagnostičke procedure i dodatna opterećenja za pacijente uzrokovana lažno pozitivnim rezultatima. Odziv se, međutim, često smatra najkritičnijom metrikom u ovom kontekstu jer propuštanje slučajeva upale pluća (lažno negativni rezultati) može odgoditi liječenje i ugroziti pacijente. F1 mjera, balansirajući preciznost i odziv, pruža realističniji sažetak performansi modela od same točnosti, osobito kada i laži alarmi i propuštene detekcije imaju značajne posljedice.

Iako su i druge mjere, poput specifičnosti ili područja ispod ROC krivulje, relevantne u medicinskim istraživanjima, u ovom radu fokus je na točnosti, preciznosti, odzivu i F1 mjeri. Zajedno, ove četiri metrike pružaju sveobuhvatan i klinički relevantan okvir vrednovanja,

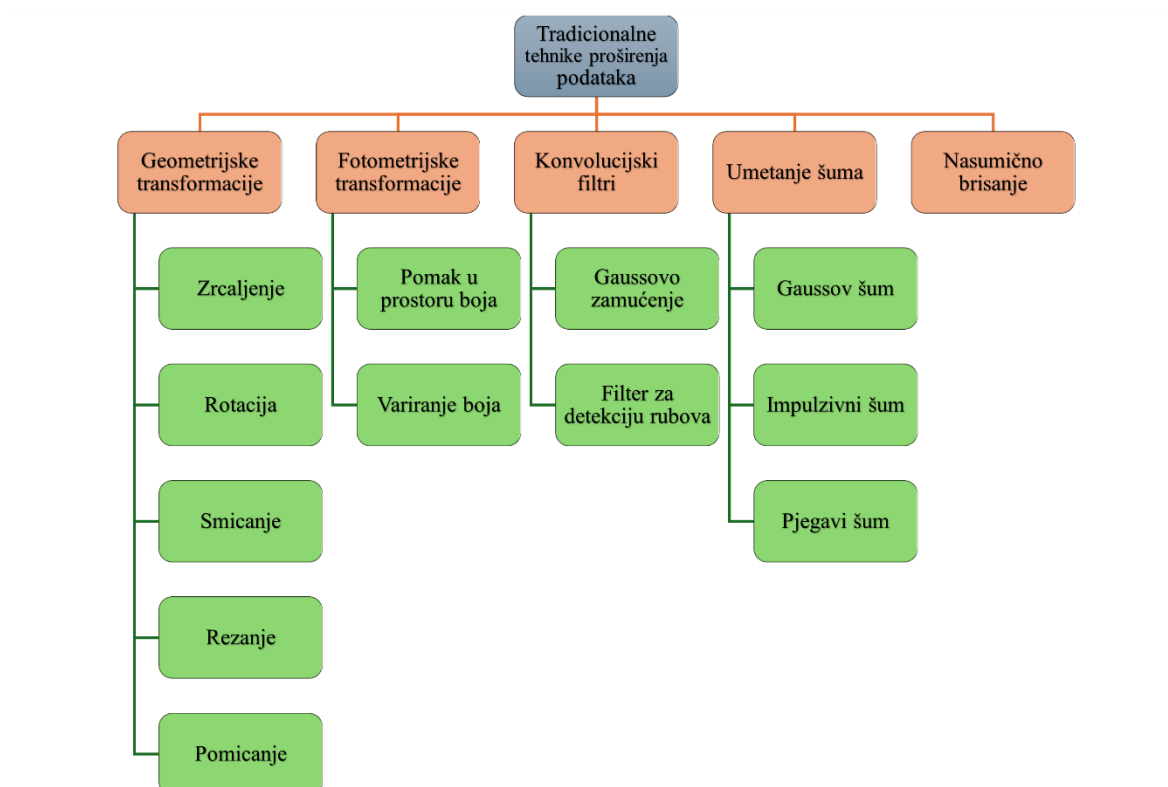
osiguravajući da se u procjeni modela za identifikaciju upale pluća uzimaju u obzir i ukupna ispravnost i klinički rizici povezani s pogrešnom klasifikacijom.

4. Proširenje podataka

Proširenje podataka je strategija koja se koristi u dubokom učenju, osobito u računalnom vidu, za umjetno proširenje i povećanje raznolikosti skupa podataka za obuku stvaranjem izmijenjenih kopija postojećih podataka. Time se pomaže u rješavanju izazova kao što su ograničena dostupnost podataka i prekomjerno prilagođavanje (Alomar et al., 2023). Potreba za proširenjem podataka proizlazi iz zahtjeva modernih algoritama strojnog učenja, posebice dubokog učenja, za velikim količinama podataka. Budući da je prikupljanje i označavanje takvih skupova često skupo i vremenski zahtjevno, proširenje podataka omogućuje generiranje dodatnih varijacija postojećih primjera. Tako se poboljšava točnost, generalizacija i robusnost modela, osobito u slučajevima ograničenih ili neuravnoteženih skupova (Zha et al., 2025).

4.1. Tradicionalne tehnike proširenja podataka

Tradicionalne tehnike proširenja podataka, u kontekstu obrade slike, mogu se smatrati osnovnim tehnikama proširenja podataka. Uključuju primitivne tehnike manipulacije podacima. Obuhvaćaju geometrijske transformacije, fotometrijske transformacije, konvolucijske filtre, umetanje šuma i nasumično brisanje. Slika 3 predstavlja taksonomiju tradicionalnih tehnika proširenja podataka.



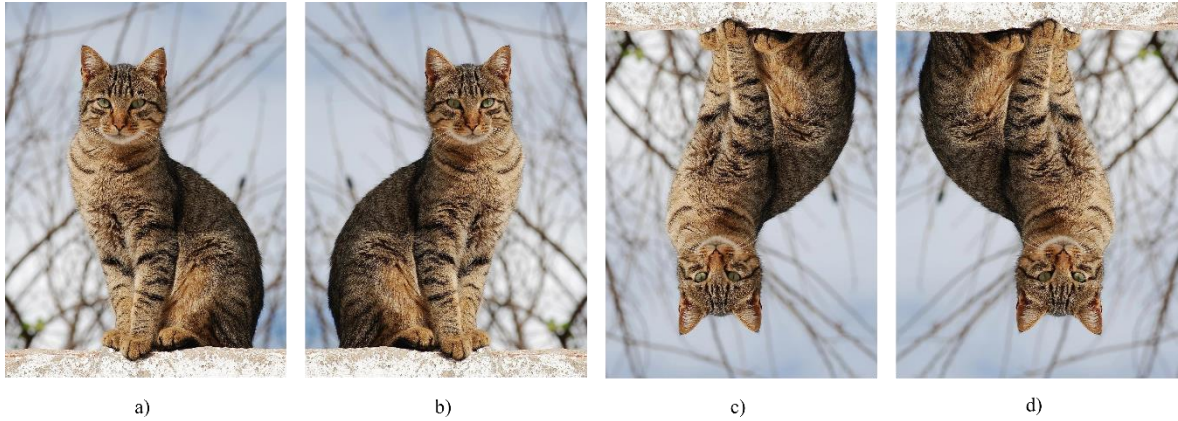
Slika 3 Taksonomija tradicionalnih tehnika proširenja podataka

4.1.1. Geometrijske transformacije

Geometrijske transformacije predstavljaju skup tehnika proširenja podataka kojima se mijenjaju prostorna svojstva slike, uključujući položaj, orijentaciju i oblik, radi stvaranja dodatnih varijacija unutar skupa podataka. Tehnike poput zrcaljenja, rotacije, smicanja, izrezivanja i pomicanja jednostavne su za implementaciju i često se koriste u računalnom vidu kako bi se poboljšala otpornost modela na različite rasporede i deformacije objekata. Prilikom primjene ovih transformacija važno je razmotriti njihovu sigurnost, odnosno utjecaj na očuvanje oznake slike. Dok su neke transformacije sigurne u većini konteksta, druge mogu izmijeniti semantičko značenje slike, što treba uzeti u obzir pri izgradnji učinkovitih strategija proširenja podataka. U nastavku su detaljno opisane najčešće geometrijske transformacije korištene u proširenju slikovnih podataka.

Zrcaljenje (engl. *flipping*) je tehnika proširenja podataka koja podrazumijeva horizontalno ili vertikalno okretanje slike kako bi se povećala raznolikost skupa podataka. Najčešće se koristi horizontalno zrcaljenje zbog veće realističnosti, osobito kada nedostaju prikazi objekata iz različitih smjerova. Iako učinkovita, ova metoda nije prikladna za smjerom

osjetljive ili asimetrične podatke, poput slova i brojki, jer može dovesti do pogrešnih oznaka (Alomar et al., 2023). Prikaz tehnike zrcaljenja nalazi se na Slika 4.



Slika 4 Tehnika zrcaljenja: a) izvorna slika, b) horizontalno zrcaljenje, c) vertikalno zrcaljenje, d) vertikalno i horizontalno zrcaljenje

Vertikalnim zrcaljenjem slika se preslikava tako da se zamijene gornji i donji dijelovi, pri čemu se os y projicira prema vrhu, a os x prema dnu slike. Vrijednosti f_x i f_y predstavljaju nove koordinate svakog piksela nakon zrcaljenja duž vertikalne osi, kako je prikazano u jednadžbi (5).

$$\begin{bmatrix} f_x \\ f_y \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} \quad (5)$$

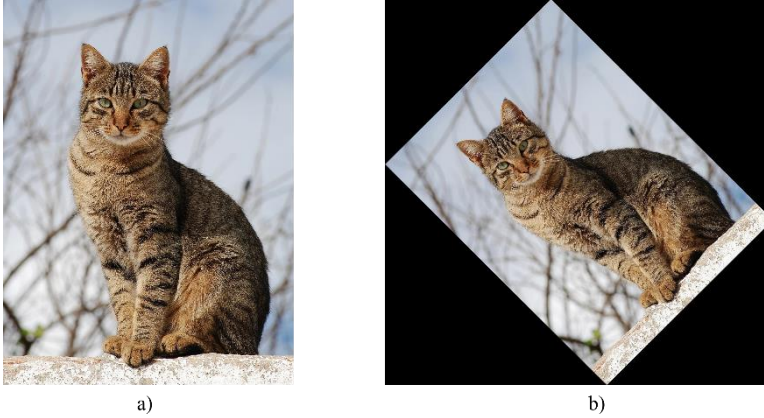
Horizontalnim zrcaljenjem dolazi do preslikavanja slike u vodoravnom smjeru, pri čemu se zamjenjuju lijeva i desna strana. Vrijednosti f_x i f_y predstavljaju nove koordinate svakog piksela nakon zrcaljenja duž horizontalne osi y , što je prikazano u jednadžbi (6).

$$\begin{bmatrix} f_x \\ f_y \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} \quad (6)$$

Kombiniranjem vertikalnog i horizontalnog zrcaljenja slika se istovremeno preslikava duž obje osi, pri čemu dolazi do zamjene gornjeg i donjeg, kao i lijevog i desnog dijela slike. Time se očuvaju rasporedi piksela duž vodoravnih i okomitih linija. Vrijednosti f_x i f_y predstavljaju nove koordinate svakog piksela nakon zrcaljenja u odnosu na obje osi, što je prikazano u jednadžbi (7).

$$\begin{bmatrix} f_x \\ f_y \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} \quad (7)$$

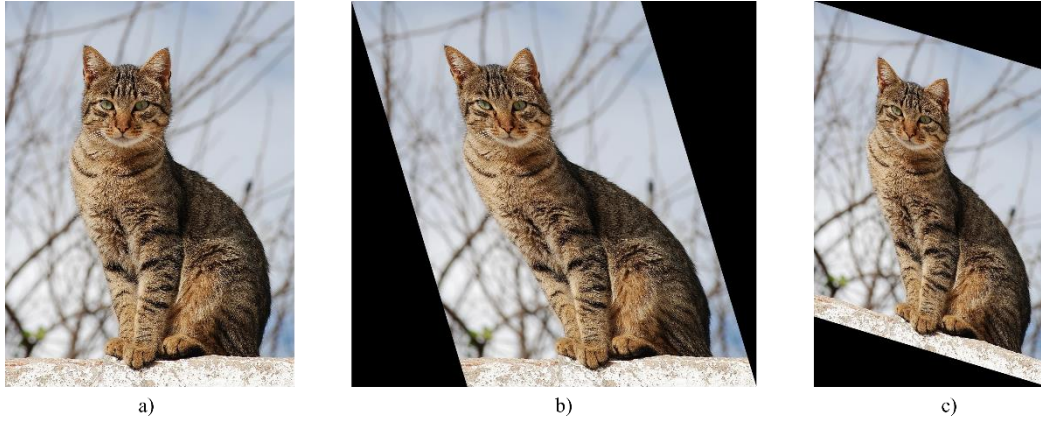
Rotacija (engl. *rotation*) predstavlja klasičnu geometrijsku tehniku proširenja podataka za slike gdje se slike rotiraju za određeni kut. Proces rotacije provodi se okretanjem slike oko osi, bilo u desnom ili lijevom smjeru, pod kutovima između 1° i 359° (Shorten & Khoshgoftaar, 2019). Jednadžba rotacije prikazana je u jednadžbi (8). Vrijednosti f_x i f_y predstavljaju nove koordinate svakog piksela nakon procesa rotacije, dok su x i y koordinate izvorne slike. Slika 5 prikazuje primjer slike s kutom rotacije od 45° .



Slika 5 Tehnika rotacije: a) izvorna slika, b) rotirana slika

$$\begin{bmatrix} f_x \\ f_y \end{bmatrix} = \begin{bmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} \quad (8)$$

Smicanje (engl. *shearing*) je tehnika proširenja podataka kojom se dijelovi slike pomiču u suprotnim smjerovima, čime se postiže transformacija geometrije slike (Kumar et al., 2024). Ova tehnika doprinosi većoj raznolikosti podataka i može poboljšati robusnost modela u zadacima računalnog vida. Međutim, ako se primijeni u prevelikoj mjeri, može uzrokovati ozbiljne deformacije slike koje otežavaju prepoznavanje objekata, čime se smanjuje učinkovitost modela. Stoga je važno pažljivo odabrati intenzitet smicanja kako bi se postigla ravnoteža između korisne raznolikosti i izbjegavanja uvođenja neželjenog šuma u podatke. Smicanje se može provesti na dva načina, duž osi x ili duž osi y . U prvom slučaju, kada se transformacija primjenjuje duž osi x , riječ je o horizontalnom smicanju, što je prikazano jednadžbom (9). U drugom slučaju, smicanje se provodi duž osi y , što rezultira vertikalnim pomakom piksela, što je prikazano jednadžbom (10). Vrijednosti f_x i f_y predstavljaju nove koordinate svakog piksela nakon primjene smicanja, dok su x i y izvorne koordinate prije transformacije. Primjeri obje vrste smicanja prikazani su na Slika 6.



Slika 6 Tehnika smicanja: a) izvorna slika, b) slika smicana u smjeru osi x , c) slika smicana u smjeru osi y

$$\begin{bmatrix} f_x \\ f_y \end{bmatrix} = \begin{bmatrix} 1 & shX \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} \quad (9)$$

$$\begin{bmatrix} f_x \\ f_y \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ shY & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} \quad (10)$$

Rezanje (engl. *cropping*) je osnovna tehnika proširenja podataka koja uključuje nasumično ili središnje izrezivanje dijela slike, nakon čega se izrezani segment skalira na izvornu veličinu, pri čemu se čuva izvorna oznaka slike. Međutim, ova tehnika može dovesti do problema s točnošću oznaka, osobito kada slike sadrže više objekata, a oznaka se odnosi na dominantni objekt. U takvim slučajevima, moguće je da rezanje izvuče područje koje sadrži detalje sporednog objekta, što može rezultirati netočnim označavanjem proširenih uzoraka. Neadekvatan izbor intenziteta rezanja može rezultirati gubitkom semantički značajnih dijelova slike, što negativno utječe na sposobnost modela da pravilno uči i klasificira uzorke (Alomar et al., 2023). Jednadžba (11) prikazuje jednadžbu skaliranja, vrijednosti f_x i f_y predstavljaju nove koordinate svakog piksela nakon skaliranja, a x i y predstavljaju koordinate izvornog položaja na slici. Slika 7 prikazuje primjer rezanja.



a)



b)

Slika 7 Tehnika izrezivanja: a) izvorna slika, b) izrezana slika

$$\begin{bmatrix} f_x \\ f_y \end{bmatrix} = \begin{bmatrix} X_{scale} & 0 \\ 0 & Y_{scale} \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} \quad (11)$$

Pomicanje (engl. *translation*) je tehnika proširenja podataka kojom se slika pomiče u horizontalnom ili vertikalnom smjeru, s ciljem smanjenja prostorne pristranosti unutar skupa podataka. Ova tehnika je osobito korisna u slučajevima kada su svi uzorci centrirani, kao što je često slučaj u skupovima za prepoznavanje lica, jer omogućuje modelima veću fleksibilnost prilikom obrade podataka koji nisu savršeno poravnati. Prilikom pomicanja, praznine nastale pomakom slike mogu se ispuniti konstantnim vrijednostima (npr. 0 ili 255), slučajnim vrijednostima ili šumom poput Gaussovog. Ovakvim popunjavanjem (engl. *padding*) moguće je sačuvati izvorne prostorne dimenzije slike i nakon primjene transformacije (Shorten & Khoshgoftaar, 2019). Jednadžba pomaka prikazana je jednadžbom (12). Vrijednosti f_x i f_y predstavljaju nove koordinate svakog piksela nakon pomicanja, dok su x i y izvorne koordinate prije transformacije. Slika 8 prikazuje primjer pomicanja.



a)



b)

Slika 8 Tehnika pomicanja: a) izvorna slika, b) pomaknuta slika

$$\begin{bmatrix} f_x \\ f_y \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} T_x \\ T_y \end{bmatrix} \quad (12)$$

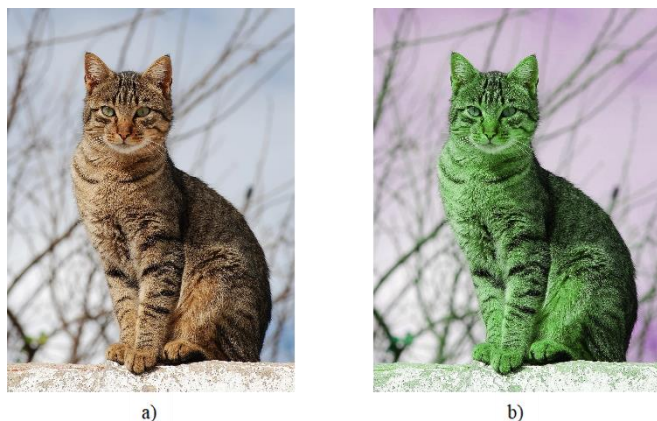
4.1.2. Fotometrijske transformacije

Fotometrijske transformacije predstavljaju skup tehnika proširenja slika koje djeluju tako da mijenjaju vrijednost piksela unutar slike, bez promjene njihove prostorne raspodjele. Cilj ovih transformacija je simulirati širok raspon uvjeta osvjetljenja i boja, čime se podržava razvoj modela računalnog vida otpornijih na promjene u osvjetljenju i varijacije u boji. Uobičajeno fotometrijsko proširenje podataka uključuje prilagodbe svjetline, kontrasta i boje, koje izravno utječu na način na koji se slika kodira i percipira. Digitalne slike obično se pohranjuju kao trodimenzionalni nizovi (visina \times širina \times kanali boje). Svaki kanal boje predstavlja vrijednost intenziteta određene boje. Manipulacijom tim kanalima moguće je primijeniti razne strategije proširenja podataka, od jednostavnog podešavanja svjetline i kontrasta do složenijih prilagodbi boja (Alomar et al., 2023; Shorten & Khoshgoftaar, 2019).

Prostor boja (engl. *color space*) je matematički okvir za prikazivanje boja u digitalnoj obradi slike. Izbor prostora boje uvelike utječe na učinkovitost fotometrijske transformacije. Najčešće korišteni prostor boja je RGB (engl. *Red, Green, Blue*, skraćeno RGB). Ipak, alternativni prostori poput HSL (engl. *Hue, Saturation, Lightness*) i HSV (engl. *Hue, Saturation, Value*, skraćeno HSV) često se koriste za proširenje podataka jer razdvajaju različite perceptivne komponente boje. Na primjer, nijansa predstavlja pigment, zasićenost određuje intenzitet, a svjetlina ili vrijednost označava koliko je boja svijetla ili tamna. Transformacija prostora boja uključuje pretvorbu slike iz jednog prostora boja u drugi, zatim primjenu tehnika proširenja podataka, a potom, po potrebi, povratak u RGB prostor. Manipulacija pojedinačnim komponentama omogućuje preciznije proširenje podataka, primjerice, pomicanje samo nijanse u HSV prostoru stvara raznolike varijante boja, a pritom čuva strukturalni sadržaj slike. Ovaj postupak ne samo da obogaćuje skup podataka za obuku, već i povećava robusnost modela na varijacije boje i osvjetljenja.

Pomak u prostoru boja (engl. *color space shifting*) odnosi se na operacije koje mijenjaju raspodjelu ili ravnotežu boja unutar slike, bilo izmjenom vrijednosti pojedinih kanala, bilo promjenom ukupnog sadržaja boje. To može biti jednostavno kao pojačavanje ili smanjivanje intenziteta jednog kanala (npr. pojačavanje crvenog kanala za topliji izgled slike), ili složenije poput permutacije kanala i izjednačavanje histograma. Ovi pomaci mogu simulirati različite scenarije, kao što su promjene u osvjetljenju ili dobu dana. U praksi,

pomak u prostoru boja izlaže modele strojnog učenja većem broju varijacija boja, osiguravajući da naučene značajke nisu previše ovisne o određenoj pristranosti boje iz izvornog skupa podataka, pretjerani ili neprimjereni pomaci mogu narušiti ključne značajke boje koje su važne za točnost modela, osobito u zadacima gdje je boja bitna za prepoznavanje. Primjer pomaka u prostoru boja nalazi se na Slika 9.



Slika 9 Pomak u prostoru boja: a) izvorna slika, b) slika pretvorena iz RGB u HSV i primijenjen pomak nijanse od 72°

Variranje boja (engl. *colour jittering*) je široko korištena tehnika fotometrijskog proširenja podataka koja uvodi kontrolirane, nasumične varijacije u boje slike, uključujući svjetlinu, kontrast, zasićenost i nijansu (Slika 10). Tijekom obuke, svaka slika se podvrgava tim prilagodbama s pomoću slučajno odabranih parametara unutar unaprijed definiranih raspona, čime se stvara niz vizualno raznolikih slika. Ova proces pomaže modelima da postanu neovisni o slučajnim promjenama boje ili osvjetljenja, poput onih uzrokovanih okolišnim faktorima ili postavkama kamere. Iako je variranje boja učinkovito za proširenje raznolikosti skupa podataka i poboljšava generalizaciju modela, mora se pažljivo primjenjivati. Pretjerano variranje može uništiti informacije o boji koje su važne za zadatke klasifikacije ili unijeti nerealne vizualne artefakte.

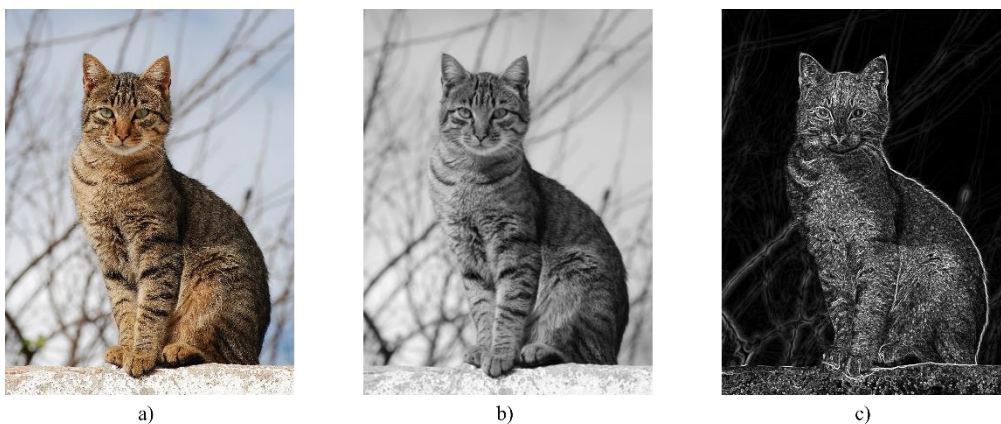


Slika 10 Variranje boja: a) izvorna slika, b) zasićenost, c) svjetlina, d) kontrast, e) nijansa

4.1.3. Konvolucijski filtri

Konvolucijsko filtriranje (engl. *kernel filter*) je tehnika proširenja podataka koja se koristi za izoštravanje ili zamućivanje slika. Postiže se primjenom kliznog prozora veličine $n \times n$ preko slike, koristeći Gaussovo zamućenje (engl. *Gaussian-blur filter*) ili filter za detekciju rubova (engl. *edge filter*). Gaussovo zamućenje služi za zaglađivanje slike smanjenjem visokofrekventnih komponenti, dok filter za rubove naglašava granice pojačavanjem gradijenata u horizontalnom ili vertikalnom smjeru, kako je prikazano na Slika 11.

Iako konvolucijsko filtriranje može poboljšati generalizaciju modela uvođenjem varijabilnosti u podatke za obuku, ono također ima određena ograničenja. Pretjerano zamućivanje može prikriti ključne strukturne detalje, čime se otežava sposobnost modela da nauči i prepozna važne značajke. S druge strane, pretjerano izoštravanje može pojačati šum i uvesti artefakte, što potencijalno može dovesti do pogrešne klasifikacije ili smanjenja performansi modela. Stoga je pažljivo podešavanje parametara konvolucijskog filtriranja ključno kako bi se postigla ravnoteža između koristi proširenja podataka i moguće degradacije kvalitete slike (Kumar et al., 2024).



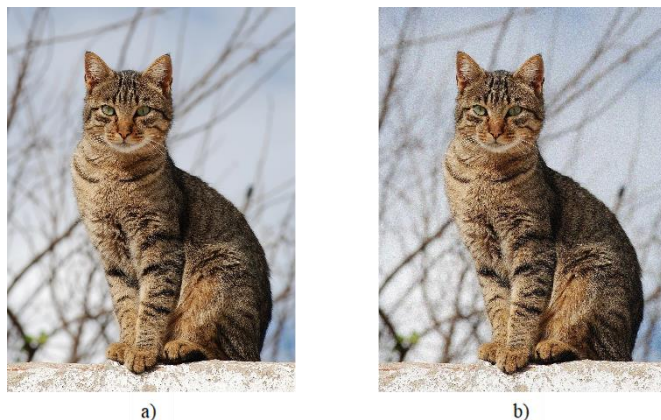
Slika 11 Konvolucijski filtri: a) izvorna slika, b) Gaussovo zamućenje, c) filter za detekciju rubova

4.1.4. Umetanje šuma

Šum na slikama odnosi se na nasumične promjene u svjetlini ili boji, koje su često posljedica ograničenja senzora ili nepovoljnih uvjeta u okruženju. Takvi šumovi su neizbježni u stvarnim uvjetima i predstavljaju izazov prilikom obrade i analize slika. Neuronske mreže su posebno osjetljive na prisutnost šuma, što može negativno utjecati na njihovu preciznost i sposobnost generalizacije tijekom testiranja na podacima koji nisu potpuno čisti. Kako bi

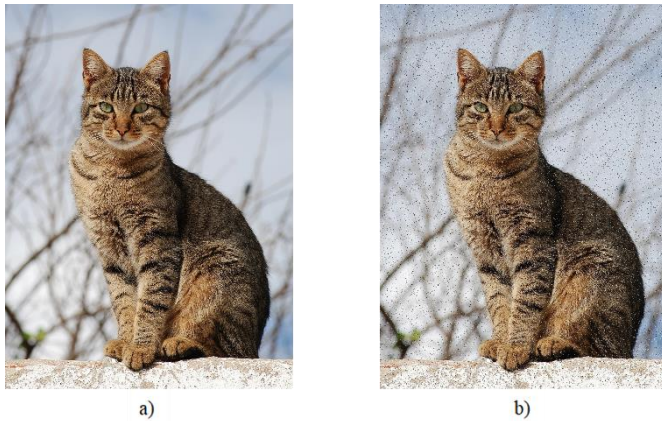
se povećala otpornost modela, skup podataka za obuku može se proširiti dodavanjem različitih vrsta šuma, poput Gaussovog, impulzivnog i pjegavog šuma.

Gaussov šum (engl. *Gaussian noise*) je vrsta statističkog šuma koju karakterizira funkcija gustoće vjerojatnosti koja slijedi normalnu raspodjelu. Ovaj šum je ravnomjerno raspoređen kroz cijelo područje signala. Kao aditivni oblik šuma, mijenja piksele slike dodavanjem nasumičnih vrijednosti izvučenih iz normalne raspodjele na izvorne jačine piksela. Statistički je neovisan za svaki piksel i ne ovisi o amplitudi signala (Alomar et al., 2023). Primjer Gaussovog šuma dan je na Slika 12.



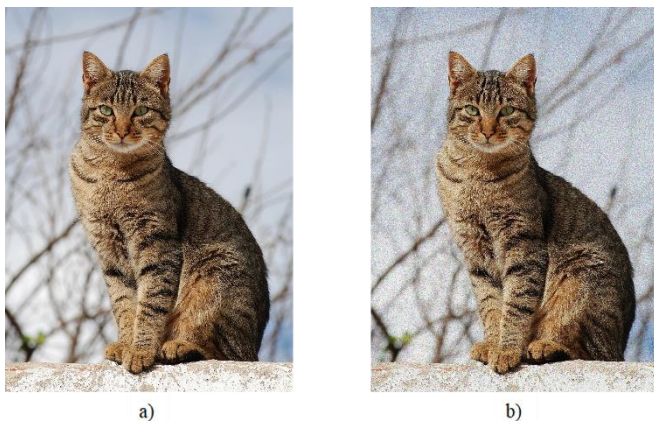
Slika 12 Umetanje šuma: a) izvorna slika, b) Gaussov šum

Impulzivni šum (engl. *impulsive noise*), također poznat kao šum soli i papra (engl. *salt and pepper noise*), čest je oblik izobličenja prisutan u digitalnim slikama. Očituje se kao nasumično raspoređeni crni i bijeli pikseli, koji se često pojavljuju kao svijetle točke u tamnim područjima slike ili tamne točke u svijetlim dijelovima. Različiti čimbenici mogu dovesti do pojave ove vrste šuma, uključujući mrtve piksele, pogreške tijekom pretvorbe iz analognog u digitalni signal ili oštećenja bitova tijekom prijenosa podataka (Vyas et al., 2018). Primjer impulzivnog šuma prikazan je na Slika 13.



Slika 13 Umetanje šuma: a) izvorna slika, b) impulzivni šum

Pjegavi šum (engl. *speckle noise*) je oblik multiplikativnog šuma koji se karakterizira granuliranim izgledom i narušava kvalitetu slika snimljenih aktivnim sustavima za snimanje, uključujući radar i radar sa sintetičkom aperturom (engl. *synthetic aperture radar*, skraćeno SAR). Ovaj šum nastaje zbog slučajnih varijacija u povratnom signalu. Pjegavi šum povećava lokalnu srednju razinu sive i otežava interpretaciju SAR slika, uglavnom zbog koherentne obrade signala koji se odbijaju od višestrukih raspršenih objekata (Vyas et al., 2018). Na Slika 14 prikazan je primjer pjegavog šuma.

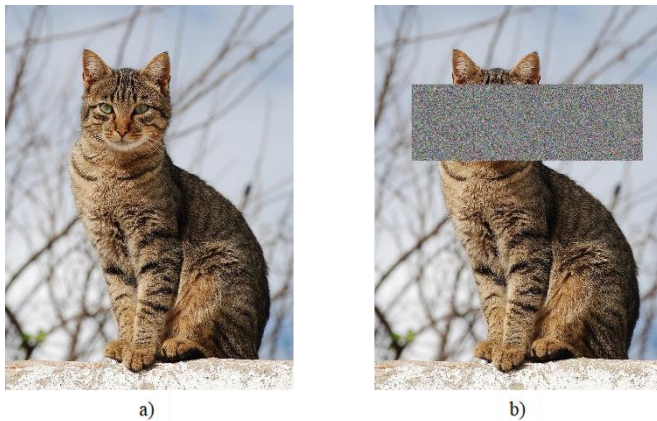


Slika 14 Umetanje šuma: a) izvorna slika, b) pjegavi šum

4.1.5. Nasumično brisanje

Nasumično brisanje (engl. *random erasing*) je tehnika proširenja podataka koja, umjesto mijenjanja pojedinačnih vrijednosti piksela, zamjenjuje sve vrijednosti piksela unutar nasumično odabranog pravokutnog područja slike slučajnim vrijednostima (Slika 15). Ova se tehnika može smatrati oblikom lokaliziranog umetanja šuma usmjerenog na poboljšanje robusnosti modela na djelomične okluzije (Alomar et al., 2023).

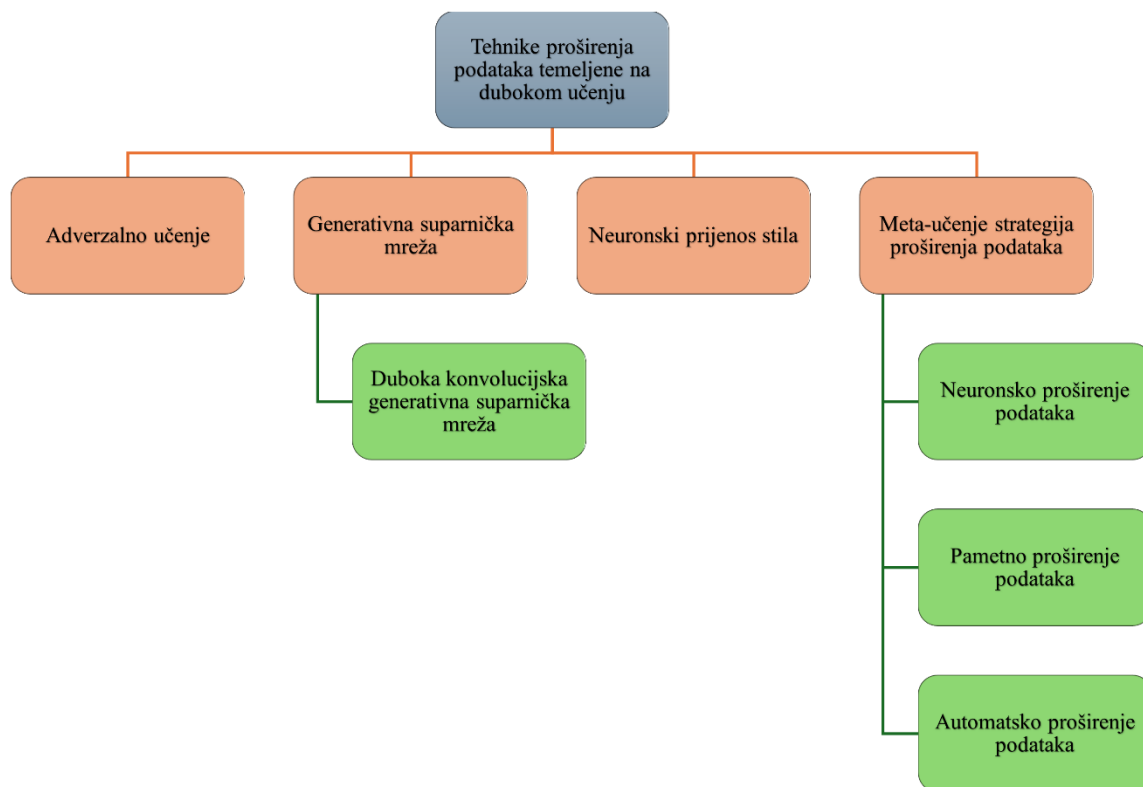
Uvođenjem varijabilnosti u lokalizirana područja, nasumično brisanje smanjuje prekomjerno prilagođavanje i poboljšava raznolikost podataka bez povećanja veličine skupa podataka. Međutim, budući da nasumično zaklanja dijelove slike, može prikriti ključne značajke nužne za klasifikaciju. Stoga ova tehnika možda nije prikladna za zadatke koji uključuju osjetljive ili precizne vizualne informacije, poput prepoznavanja znakova ili registarskih oznaka, gdje je potpuna vidljivost od presudne važnosti.



Slika 15 Nasumično brisanje: a) izvorna slika, b) slika nakon nasumičnog brisanja

4.2. Tehnike proširenja podataka temeljene na dubokom učenju

Tehnike proširenja podataka temeljene na dubokom učenju uključuju adverzalno učenje, generativne suparničke mreže, neuronski prijenos stila i meta-učenje strategija proširenja podataka. Za razliku od tradicionalnih tehnika, koje se uglavnom oslanjaju na jednostavne transformacije, tehnike temeljene na dubokom učenju koriste složene modele sposobne za automatsko učenje obrazaca i generiranje potpuno novih uzoraka. Taksonomija ovih tehnika prikazana je na Slika 16.

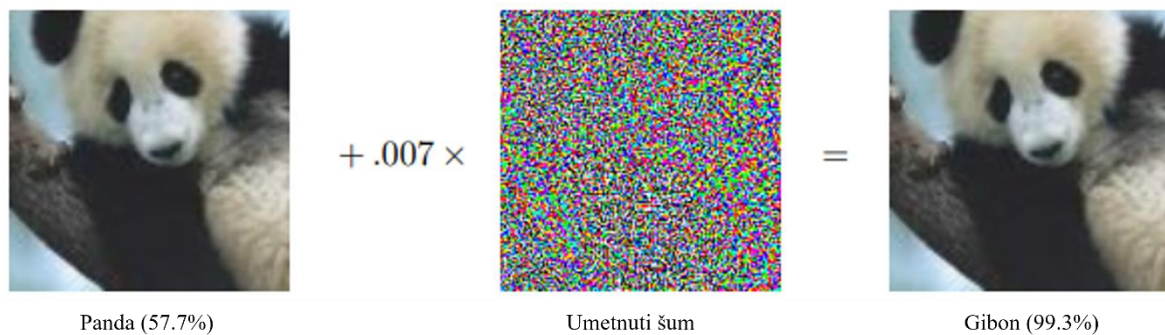


Slika 16 Taksonomija tehnika proširenja podataka temeljenih na dubokom učenju

4.2.1. Adverzalno učenje

Adverzalni primjeri (engl. *adversarial examples*), često nazivani „strojne iluzije“, predstavljaju ključno područje interesa unutar domene dubokog učenja. Ovi primjeri mogu se interpretirati kao specifičan oblik proširenja podataka temeljen na šumu, pri čemu se u ulazne slike uvode namjerne i suptilne perturbacije. Iako su te izmjene neprimjetne ljudskom oku, mogu rezultirati potpuno pogrešnim predikcijama dubokih neuronskih mreža. Ova pojava ukazuje na ranjivosti u ponašanju modela te ističe osjetljivost dubokih arhitektura na varijacije ulaza. Jedan od pristupa za povećanje robusnosti modela je adverzalno učenje (engl. *adversarial training*), koje podrazumijeva integraciju adverzalnih primjera tijekom procesa obuke, čime se modelu omogućuje bolje suočavanje s malicioznim ili neočekivanim ulazima. Shodno tome, adverzalno proširenje podataka ne samo da doprinosi većoj sigurnosti modela, već i služi kao vrijedan alat za identifikaciju i rješavanje strukturnih slabosti neuronskih modela (Alomar et al., 2023). Konkretni primjer adveralnog učenja nalazi se na Slika 17, gdje je istraživanje koje su proveli (Goodfellow et al., 2015) pokazalo da čak i iznimno točni modeli poput GoogLeNet-a mogu biti zavarani primjenom male perturbacije na ulaznu sliku. Korištenjem metode poznate kao metoda znaka brzog gradijenta (engl. *fast gradient sign method*), autori su dodali

minimalnu promjenu svakom pikselu u smjeru gradijenta funkcije gubitka. Iako je perturbacija bila vizualno neprimjetna, izmijenjena slika dovela je model do pogrešne klasifikacije uz visoku razinu pouzdanosti. Ovaj primjer ilustrira kako i najmanje manipulacije ulazom mogu otkriti ranjivosti u inače vrlo učinkovitim neuronskim mrežama te naglašava važnost adverzalnih primjera u otkrivanju osjetljivosti modela i povećanju njegove robusnosti kroz adverzorno učenje.



Slika 17 Adverzorno učenje (Goodfellow et al., 2015)

4.2.2. Generativna suparnička mreža

Generativne suparničke mreže predstavljaju klasu modela dubokog učenja osmišljenih za učenje statističke distribucije podataka iz skupa za obuku te generiranje novih uzoraka koji toj distribuciji nalikuju (Alqahtani et al., 2021; Creswell et al., 2018). Postale su značajan alat u polunadziranom i nenadziranom učenju, osobito kada je količina označenih podataka ograničena. GAN-ovi se obučavaju kroz međusobno nametanje dviju neuronskih mreža: jedna stvara sintetičke podatke, dok ih druga pokušava razlikovati od stvarnih.

Generator je odgovoran za stvaranje sintetičkih uzoraka podataka iz slučajnih ulaza izvučenih iz latentnog prostora. Putem niza nelinearnih transformacija, on te varijable pretvara u podatke koji oponašaju karakteristike skupa za obuku. Cilj generatora je proizvesti izlaze dovoljno uvjerljive da zavaraju diskriminator.

Diskriminator je neuronska mreža obučena za klasifikaciju ulaznih podataka kao stvarnih ili sintetičkih. Mreža mapira ulazne podatke u vjerojatnosni rezultat, pri čemu vrijednosti bliže jedinici označavaju stvarne uzorke, a vrijednosti bliže nuli sintetičke.

Kroz ovaj suparnički proces obje mreže poboljšavaju svoje performanse, što vodi prema generiranju sve realističnijih sintetičkih uzoraka. Ako je diskriminator previše točan prerano, generator može dobivati nedovoljno korisnih povratnih informacija, što otežava učenje. U

idealnom slučaju, obuka doseže točku u kojoj diskriminator ne može pouzdano razlikovati stvarne od sintetičkih podataka, predviđajući vjerojatnost blizu 0,5 za sve ulaze.

Osim generiranja sintetičkih podataka, GAN-ovi se mogu primijeniti u povezanim zadacima poput super rezolucije, prijenosa stila, prijenosa slike u sliku i prilagodbe domena. U području medicinskog snimanja to može značiti poboljšanje niskorezolucijskih snimki, stvaranje dodatnih podataka kada su označeni skupovi mali ili prilagodbu sintetičkih slika kako bi odgovarale karakteristikama stvarnih pacijenata. Kontrola nad latentnim prostorom i sposobnost modeliranja složenih, visokodimenzionalnih distribucija čini GAN-ove osobito pogodnima za ovakve zadatke.

S vremenom je razvijeno više varijanti izvornog GAN-a koje rješavaju izazove obuke i proširuju mogućnosti modela, među kojima su: osnovni GAN (engl. *Vanilla GAN*), duboki konvolucijski GAN, uvjetovani GAN (engl. *Conditional GAN*), suparnički autoenkoder (engl. *Adversarial Autoencoder*), dvosmjerni GAN (engl. *Bidirectional GAN*), suparnički naučena inferencija (engl. *Adversarially Learned Inference*), GAN za maksimizaciju informacija (engl. *Information Maximizing GAN*).

4.2.3. Duboka konvolucijska generativna suparnička mreža

Duboke konvolucijske generativne suparničke mreže predstavljaju značajan napredak u nenadziranom učenju primjenom konvolucijskih neuronskih mreža. Prvi put su ih predstavili (Radford et al., 2015), kako bi odgovorili na probleme nestabilnog učenja i ograničene kvalitete generiranih slika prisutne u ranijim GAN modelima.

DCGAN kombinira okvir suparničkog učenja GAN-ova s konvolucijskim arhitekturama te uvodi niz arhitektonskih smjernica koje poboljšavaju stabilnost obuke i učenje značajki:

- Uklanjanje potpuno povezanih skrivenih slojeva
- Zamjena slojeva sažimanja konvolucijama s korakom (engl. *strided convolutions*) u diskriminatoru te konvolucijama s razlomljenim korakom (engl. *fractional-strided convolutions*) u generatoru
- Primjena normalizacije po serijama u generatoru i diskriminatoru
- Korištenje *ReLU* aktivacijske funkcije u svakom sloju generativnog modela osim u posljednjem sloju te korištenje *LeakyReLU* aktivacijske funkcije u svim slojevima diskriminativnog modela

Ove inovacije omogućuju pouzdano i stabilno nenadzirano generiranje slika te učinkovito učenje njihovih reprezentacija. DCGAN-ovi mogu usvajati smislene, hijerarhijski organizirane značajke, što ih čini posebno pogodnima za zadatke koji zahtijevaju realistične sintetičke slike i bogatu ekstrakciju značajki.

4.2.4. Neuronski prijenos stila

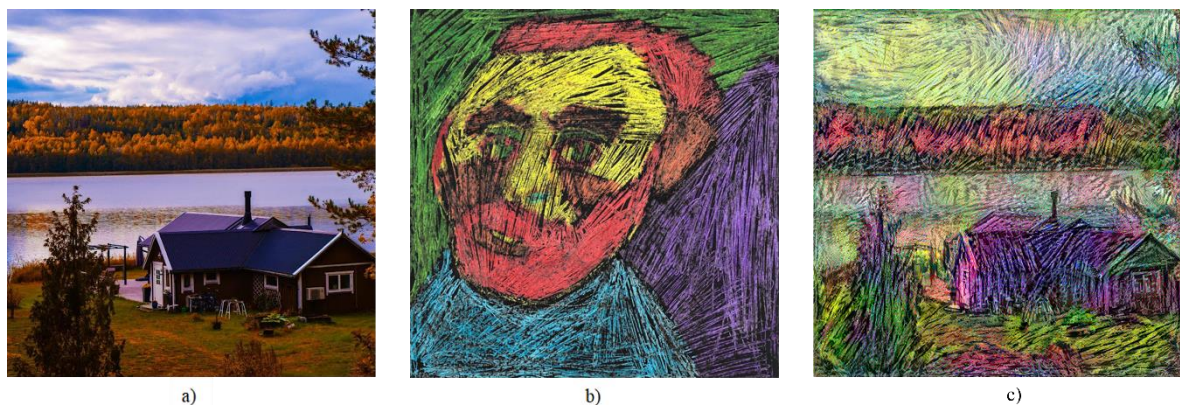
Neuronski prijenos stila (engl. *Neural Style Transfer*, skraćeno NST) predstavlja izvanredan primjer sposobnosti dubokog učenja u manipulaciji vizualnim reprezentacijama unutar konvolucijskih neuronskih mreža (Shorten & Khoshgoftaar, 2019). Iako je u početku populariziran zbog svojih kreativnih i umjetničkih primjena, NST se pokazao i kao učinkovita tehnika za proširenje podataka za obuku. Njegovo djelovanje temelji se na izmjeni reprezentacija značajki unutar neuronske mreže tako da se stilski elementi jedne slike prenesu na strukturu sadržaja druge (Slika 18). Time se omogućuje generiranje novih slika koje zadržavaju izvorni sadržaj, ali istovremeno odražavaju potpuno drugačiji vizualni stil. Ključno unaprjeđenje ovog pristupa predstavlja optimizacija algoritma u cilju postizanja brže obrade. Zamjenom tradicionalnih funkcija gubitka na razini piksela percepcijskim gubicima, koji se izводе iz reprezentacije viših razina apstrakcije, stilizacija se može izvoditi znatno učinkovitije. Nadalje, promjene u arhitekturi mreže, osobito uvođenje instance normalizacije (engl. *instance normalization*), dodatno su doprinijele poboljšanju kvalitete i brzine prijenosa stilova.

Primjena NST-a u kontekstu proširenja podataka može se promatrati kao napredni ekvivalent tradicionalnim transformacijama u prostoru boja, s time da se koncept proširuje na uključivanje umjetničkih tekstura i vizualnih stilova. Međutim, izazov ostaje u određivanju prikladnih stilova za generiranje novih primjera za obuku. U pojedinim kontekstima, poput prepoznavanja vizualnih podražaja pod različitim vremenskim ili svjetlosnim uvjetima, izbor stilova je intuitivan. S druge strane, u apstraktnijim domenama taj proces može biti znatno složeniji. Odabir ograničenog skupa stilova, koji se sustavno primjenjuje na cijeli skup podataka, predstavlja jednu od praktičnih strategija. Alternativno, stilovi se mogu nasumično uzorkovati iz šire distribucije kako bi se osigurala veća raznolikosti i umanjila mogućnost stilske pristranosti.

Tehnike prijenosa stilova pokazale su osobitu vrijednost u situacijama gdje je prikupljanje podataka iz stvarnog svijeta otežano ili rizično, primjerice pri obuci robotskih sustava u simuliranim okruženjima. Takvi modeli često pokazuju ograničenu sposobnost

generalizacije zbog razlika između sintetičkih i stvarnih podataka. Uključivanjem širokog spektra vizualnih stilova tijekom faze obuke, modeli postaju robusniji i prilagodljiviji, interpretirajući ulaze iz stvarnog svijeta kao još jednu varijaciju. U određenim slučajevima, pokazalo se da raznolikost stilova može imati veći pozitivan učinak od pokušaja da se sintetska okruženja učine fotorealističnim.

Uporaba sintetičkih podataka u zadacima računalnog vida dobiva sve veću pažnju, osobito s obzirom na visoke troškove i vremenske zahtjeve ručnog označavanja u stvarnim skupovima podataka. Generiranje stiliziranih podataka nudi učinkovit način za povećanje vizualne raznolikosti bez potrebe za opsežnim ljudskim angažmanom. Ipak, jedno od ograničenja proširenja podataka temeljenog na NST-u leži u računalnim i memorijskim zahtjevima, osobito pri primjeni velikog broja različitih stilova. Ako je skup stilova preusko definiran, može se nenamjerno uvesti pristranost u podatke. Osim toga, iako brže implementacije NST-a postoje, one obično ograničavaju korisnike na unaprijed definirane skupove stilova, što može umanjiti fleksibilnost cijelog procesa proširenja podataka.



Slika 18 Neuronski prijenos stila: a) izvorna slika b) referentni stil, c) generirani rezultat

4.2.5. Meta-učenje strategija proširenja podataka

Meta-učenje (engl. *meta-learning*) je napredni pristup u dubokom učenju koji uključuje optimizaciju neuronskih mreža uz pomoć drugih neuronskih mreža. Ovaj pristup dobiva sve više na značaju, osobito s razvojem *Neural Architecture Search* tehnike, koja primjenjuje potkrijepljeno učenje za automatsko pronalaženje učinkovitih arhitektura neuronskih mreža. Meta-učenje stoga označava pomak prema automatizaciji i unapređenju procesa učenja (Shorten & Khoshgoftaar, 2019).

U kontekstu proširenja podataka, razvijene su različite strategije meta-učenja koje koriste neuronske mreže za automatsko generiranje učinkovitih tehnika proširenja podataka. Među

najistaknutijima su neuronsko proširenje podataka, pametno proširenje podataka i automatsko proširenje podataka.

Neuronsko proširenje podataka (engl. *neural augmentation*) koristi meta-učenje za generiranje proširenih slika kombiniranjem značajki iz dviju ulaznih slika iste klase. Početne slike prolaze kroz konvolucijsku neuronsku mrežu kako bi se proizvela nova sintetička slika, koja se dodatno transformira tehnikom prijenosa stila. Dobiveni izlaz koristi se za obuku klasifikacijskog modela, a pogreška modela se propagira unatrag kako bi se ažurirala mreža za proširenje podataka. Zahvaljujući ovoj povratnoj vezi, sustav dinamički uči optimalna mapiranja koja očuvaju sadržaj i stil slike. Eksperimenti su provedeni na zadacima klasifikacije slika s ograničenim brojem primjera, ova metoda je pokazala bolje rezultate u usporedbi s tradicionalnim tehnikama poput geometrijskih transformacija ili fiksnog prijenosa stila. Najbolje performanse postignute su kombinacijom neuronskog i tradicionalnog proširenja podataka.

Pametno proširenje podataka (engl. *smart augmentation*) koristi sličan princip, ali uz dodatnu strukturu. Sastoji se od dvije mreže: mreže za proširenje podataka, koja generira sintetičke slike, i ciljne mreže, koja se obučava na tim podacima. Povratna informacija iz ciljne mreže, zajedno s dodatnim gubitkom koji osigurava dosljednost klase, koristi se za ažuriranje mreže za proširenje podataka. Za razliku od jednostavnih metoda koje koriste prosječno kombiniranje piksela, ovaj pristup razvija adaptivne kombinacije tijekom samog procesa učenja. Pokazao se vrlo učinkovit u zadacima poput prepoznavanja spola, nadmašujući tradicionalne tehnike proširenja podataka.

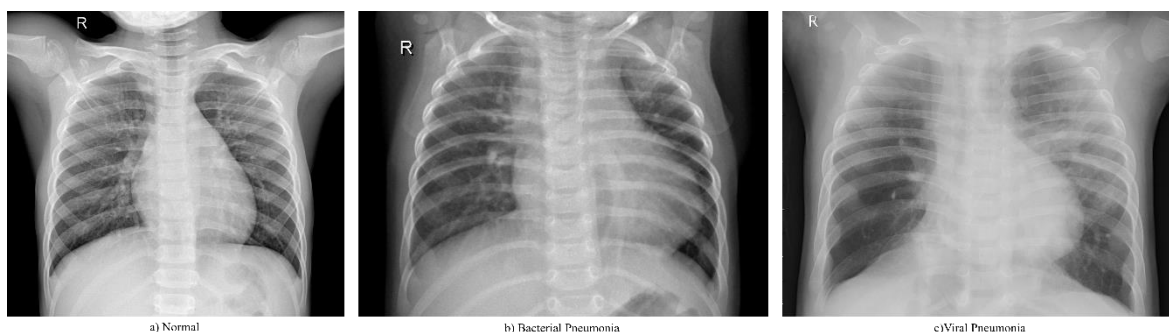
Automatsko proširenje podataka (engl. *auto augmentation*) primjenjuje potkrijepljeno učenje za odabir optimalnih politika proširenja podataka unutar unaprijed definiranog skupa geometrijskih transformacija. Svaka operacija i njezin intenzitet tretiraju se kao elementi diskretnog problema pretrage. Model uči koje kombinacije najbolje poboljšavaju performanse klasifikatora. Naučene politike pokazale su visoku prenosivost među različitim skupovima podataka, a dodatna istraživanja usmjerena su na kontinuirane prostore pretrage i proširenja podataka specifična za pojedine instance, što dodatno doprinosi povećanju performansi i robusnosti modela.

5. Metodologija

Praktični dio ovog rada implementiram je u programskom jeziku Python, uz korištenje biblioteka za duboko učenje za razvoj modela i provedbu eksperimenata. DCGAN model za generiranje sintetičkih rendgenskih snimaka prsnog koša obučen je lokalno na računalu s *NVIDIA Quadro RTX A4000* grafičkom karticom s 16 GB memorije, što je omogućilo učinkovitu obuku modela i generiranje slika. Eksperimentalna evaluacija CNN klasifikatora, uključujući obuku s različitim veličinama skupova podataka i analizu performansi, provedena je u okruženju *Google Colab Pro*, koje je pružilo prilagodljive računalne resurse te olakšalo izvođenje eksperimenata.

5.1. Skup podataka

Za izradu praktičnog, odnosno eksperimentalnog, dijela istraživanja odabran je skup podataka pod nazivom *Chest X-Ray Images (Pneumonia)* (Mooney, 2018). Ovaj skup podataka je odabran zbog slobodnog pristupa te jer se koristi i u drugim radovima ovog konteksta. Sadrži 5863 rendgenske slike u JPEG formatu, raspoređene u tri mape (*train*, *test*, *val*) s podmapama u dvije kategorije (*pneumonia*, *normal*). Slike su prikupljene iz retrospektivnih kohorti pedijatrijskih pacijenata (1-5 godina) iz *Guangzhou Women and Children's Medical Center*. Nakon inicijalne kontrole kvalitete, dijagnoze su potvrđene od strane dva stručna liječnika, dok je evaluacijski skup dodatno provjeren od strane trećeg stručnjaka. Iako u skupu podataka možemo razlikovati rendgenske slike prsnog koša s bakterijskom i virusnom upalom pluća (Slika 19b i Slika 19c), u ovom radu fokus je na prepoznavanju prisutnosti upale pluća, neovisno o njenom uzroku. Na Slika 19a prikaz je izvorne rendgenske slike prsnog koša bez znakova upale pluća.



Slika 19 Primjer rendgenskih snimaka prsnog koša

5.2. Podjela skupa podataka

Izvorna raspodjela skupa podataka je sljedeća:

1. skup za obuku: 3875 slika klase *pneumonia* + 1341 slika klase *normal*
2. skup za testiranje: 390 slika klase *pneumonia* + 234 slike klase *normal*
3. skup za validaciju: 8 slika klase *pneumonia* + 8 slika klase *normal*

Iako je skup podataka velik, na strategiju podjele utjecale su dvije praktične poteškoće. Prva se odnosi na neravnotežu klasa, budući da je u skupu za obuku broj slučajeva upale pluća gotovo tri puta veći od broja normalnih slučajeva. Druga poteškoća proizlazi iz ograničenih računalnih resursa, jer su eksperimenti provedeni na sustavu s jednom GPU jedinicom i ograničenom memorijom. Zbog toga obuka na cijelom skupu podataka, osobito kada se primjenjuje proširenje podataka i generiranje putem GAN-a, nije bilo izvedivo.

Kako bi se riješile ove poteškoće, skup podataka za obuku uravnotežen je smanjenjem broja slika s upalom pluća kako bi se uskladio s brojem dostupnih normalnih slučajeva (po 1341). Ovaj je korak osigurao jednaku zastupljenost obje klase uz smanjenje računalnog opterećenja.

Iz ovog uravnoteženog skupa od 2682 slike izdvojen je konstantni testni podskup od 500 slika, koji se sastoji od 250 slika klase *pneumonia* i 250 slika klase *normal*. Ove slike su potpuno isključene iz obuke i generativnog modeliranja, osiguravajući nepristranu evaluaciju u svim eksperimentima.

Preostale 2182 slike zatim su podijeljene u skupove za obuku i validaciju prema omjeru 90:10.

5.3. Pretprocesiranje podataka

Prije obuke konvolucijske neuronske mreže, izvorne rendgenske slike prsnog koša prošle su sveobuhvatan postupak pretprocesiranja. Time je osigurano da podaci budu u prikladnom formatu za model i pomoglo je poboljšati sposobnost modela za učenje i generalizaciju. Ovaj proces uključivao je nekoliko ključnih faza: formatiranje slika, kodiranje oznaka, integraciju sintetičkih podataka i augmentaciju podataka.

Sve slike u skupu podataka standardizirane su na jedinstvenu veličinu i format kako bi odgovarale ulaznim zahtjevima CNN-a. Svaka slika promijenjena je na dimenzije 160×160

piksela i pretvorena u RGB format. Ova standardizacija je ključna jer CNN-ovi zahtijevaju fiksnu veličinu ulaznog tenzora. Kategoričke oznake pretvorene su u numerički format prikladan za funkciju gubitka modela. Primijenjeno je *one-hot* kodiranje, pretvarajući svaku oznaku u binarni vektor. Tako je oznaka *normal* kodirana kao [1, 0], a oznaka *pneumonia* kao [0, 1]. Ovaj postupak osigurava da model ne pretpostavlja nikakav ordinalni odnos između klasa.

Ključna komponenta ovog istraživanja bila je analiza učinka sintetičkih podataka na performanse klasifikatora. U određenom skupu eksperimenata, skup za obuku nadopunjen je sintetički generiranim slikama dobivenim pomoću DCGAN-a. Implementacija je dizajnirana da bude fleksibilna, ako su dostupne sintetičke slike, one se pridodaju izvornim slikama za obuku. Povezane oznake su u skladu s tim proširene, stvarajući obogaćen skup za obuku koji je potom proslijeđen u završne korake proširenja podataka i obuke. U drugom skupu eksperimenata ovaj je korak izostavljen kako bi se dobila osnovna metrika performansi modela temeljena isključivo na izvornim podacima.

Postavljen je generator slika za obuku, validaciju i testiranje modela, koristeći *ImageDataGenerator* iz *Keras* biblioteke. Generatori omogućuju normalizaciju i proširenje podataka u stvarnom vremenu, kao i učitavanje podataka u grupama, što je ključno za rad s velikim skupovima podataka koji ne mogu stati u memoriju odjednom. Slike iz skupa za obuku su skalirane, centrirane i normalizirane, uz dodatne transformacije poput povećanja, pomaka po širini i pomaka po visini. Vrijednosti piksela slike su inicijalno u rasponu od 0 do 255. Skaliranjem se normaliziraju te vrijednosti u rasponu od 0 do 1, što omogućuje bržu i stabilniju konvergenciju tijekom obuke modela. Osim toga, standardizacija po uzorcima osigurava da svaka slika ima srednju vrijednost 0, a standardnu devijaciju 1, čineći podatke ujednačenijim. Radi povećanja otpornosti modela, primjenjuje se nasumično povećanje do 20%, kao i horizontalni i vertikalni pomaci do 10%. Ovo proširenje podataka je korisno isključivo kod obuke modela, jer se model obučava na varijantama slika, što poboljšava njegovu sposobnost generalizacije. Za validacijske i testne skupove koristi se samo normalizacija slika, bez dodatnih transformacija, kako bi se performanse modela ocijenile na stvarnim, neizmijenjenim podacima. Time se osigurava da rezultati validacije i testiranja točno odražavaju kako će model funkcionirati na izvornim, neviđenim podacima, bez uvođenja umjetnih varijacija. S obzirom na to da se u određenim slučajevima uz izvorne slike dodaju i sintetičke slike, tehnike proširenja podataka su primijenjene i na slike generirane DCGAN-om.

Na kraju, pretprocesirani podaci su u model prosljeđivani u grupama od 16 slika pomoću generatora podataka, što predstavlja učinkovit način obrade velikih skupova podataka jer slike učitava i obrađuje postupno, čime se smanjuje potrošnja memorije.

5.4. Klasifikacija slika

Arhitektura klasifikacijskog modela sastoji se od četiri konvolucijska bloka dizajnirana za postupno izdvajanje hijerarhijskih značajki iz rendgenskih snimki prsnog koša. Svaki blok sastoji se od dvodimenzionalnog konvolucijskog sloja, normalizacije po serijama za stabilizaciju i ubrzanje procesa učenja, *ReLU* aktivacijske funkcije za uvođenje nelinearnosti te sloja maksimalnog sažimanja dimenzija 2×2 za smanjenje dimenzija značajki. Konvolucijski slojevi koriste redom 32, 64, 128 i 128 filtera, svaki s jezgrom veličine 3×3 , što omogućava mreži prepoznavanje sve složenijih značajki slike. Mape značajki generirane u posljednjoj konvolucijskoj fazi ravnaju se u jednodimenzionalni vektor i prosljeđuju potpuno povezanom sloju s 512 neurona i *ReLU* aktivacijskom funkcijom, što omogućava modeliranje nelinearnih odnosa između izdvojenih značajki. Završni klasifikacijski sloj sadrži dva izlazna neurona s *sigmoid* aktivacijskom funkcijom, generirajući vjerojatnosne vrijednosti za dvije klase. Model je optimiziran pomoću stohastičkog gradijentnog spusta uz eksponencijalno smanjenje stope učenja. Za optimizaciju klasifikacijskih performansi koristi se binarna unakrsna entropija kao funkcija gubitka. Učenje se provodi kroz 20 epoha, pri čemu je broj koraka po epohi određen veličinom skupa za učenje i veličinom serije do 16. Nakon obuke, model se evaluira na testnom skupu, a performanse se mjere pomoću točnosti, preciznosti, odziva i F1 mjere. Ove metrike, zajedno s detaljima o sastavu skupa podataka i parametrima generativnog proširenja podataka, bilježe se kako bi se omogućio sveobuhvatan pregled ponašanja i rezultata modela.

5.5. Generativni model

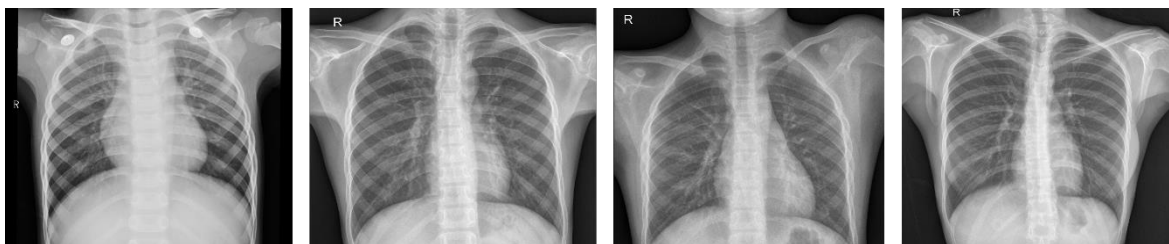
Sintetičke medicinske slike prsnog koša u ovom istraživanju generirane su korištenjem dosljedne arhitekture DCGAN-a. Mreža je prilagođena ulaznim i izlaznim dimenzijama slika korištenih u istraživanju. Obuka DCGAN-a provedena je kroz 1001 epohu.

Generativni model je dizajniran kao duboka konvolucijska neuronska mreža koja sintetizira realistične slike iz slučajnog šuma. Prati postupni pristup povećanja rezolucije (engl. *progressive upsampling*), počevši od male 10×10 latentne značajke i postupno povećavajući

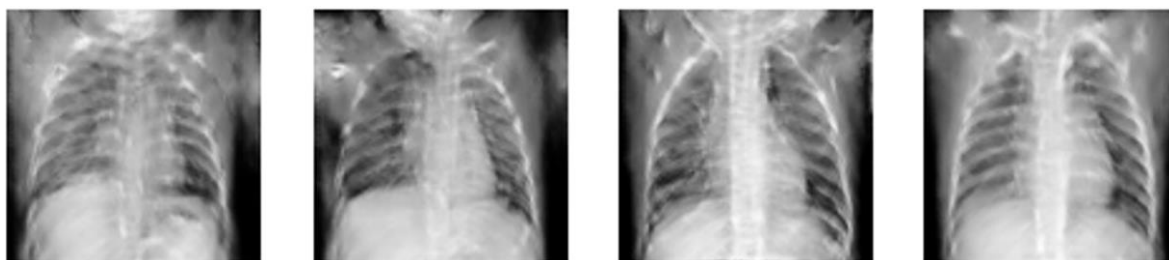
rezoluciju do željene veličine slike, u ovom slučaju 160×160 . Model prima 100-dimenzionalni latentni vektor kao ulaz, koji se preslikava u prostor više dimenzije pomoću potpuno povezanog sloja koji ga proširuje u $10 \times 10 \times 256$ tenzor. Kako bi se poboljšala stabilnost i konvergencija, primjenjuje se normalizacija po serijama, nakon čega slijedi *Leaky ReLU* aktivacijska funkcija. Glavni dio generatora sastoji se od četiri bloka koji povećavaju rezoluciju, implementiranih pomoću transponiranih konvolucijskih slojeva (*Conv2DTranspose*), koji postupno povećavaju prostorne dimenzije, dok se broj filtera smanjuje. Svaki blok primjenjuje 5×5 transponiranu konvoluciju, nakon čega slijedi normalizacija po serijama i *Leaky ReLU* aktivacijska funkcija. Veličina značajki se mijenja pri svakom koraku, prelazeći kroz $10 \times 10 \times 128$, $20 \times 20 \times 32$, $40 \times 40 \times 16$ i na kraju $80 \times 80 \times 8$. Završni sloj je transponirana konvolucija koja generira sliku ciljnih dimenzija koristeći *tanh* aktivacijsku funkciju, koja skalira vrijednosti piksela u raspon $[-1, 1]$. Ovaj postupni pristup povećanja rezolucije omogućuje generatoru postupnu sintezu sve detaljnijih slika.

S druge strane, diskriminator je također duboka konvolucijska neuronska mreža, ali s obrnutim pristupom. Umjesto povećanja rezolucije, on postupno smanjuje dimenzije ulazne slike (engl. *downsampling*), istovremeno povećavajući dubinu značajki kako bi izdvojio ključne uzorke koji razlikuju stvarne slike od generiranih. Model započinje konvolucijskim slojem s 64 filtera veličine 5×5 i korakom 2, čime se smanjuju dimenzije slike dok se zadržavaju prostorne značajke. Primjenjuje se *Leaky ReLU* aktivacijska funkcija i *Dropout* (0.3), kako bi se smanjilo prekomjerno prilagođavanje. Slijedi drugi konvolucijski sloj s 128 filtera, također 5×5 jezgrama i korakom 2, nakon kojega se ponovno primjenjuje *Leaky ReLU* i *Dropout* (0.3). Na kraju, izdvojene značajke se poravnavaju u jednodimenzionalni vektor, koji se prosljeđuje kroz potpuno povezani sloj s jednim neuronom, što daje skalarni izlaz koji predstavlja procjenu diskriminatora o autentičnosti slike.

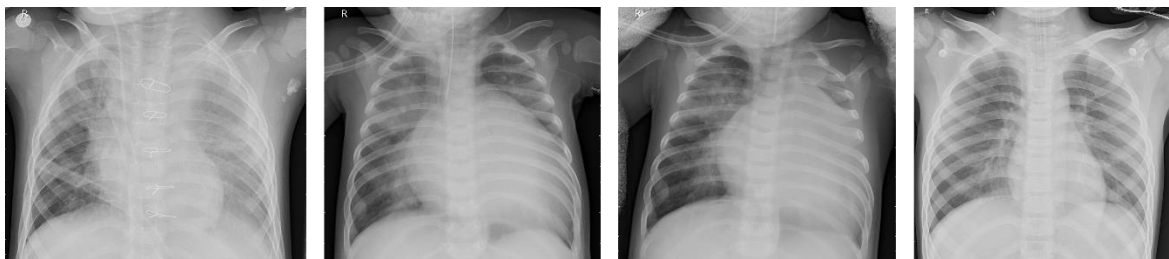
Slika 20 i Slika 21 prikazuju usporedbu izvornih i sintetičkih rendgenskih snimki zdravih pluća, dok Slika 22 i Slika 23 prikazuju usporedbu izvornih i sintetičkih snimki upale pluća. Slika 24 prikazuje proces učenja DCGAN modela tijekom sinteze rendgenskih snimaka. Svaki redak sadrži četiri generirane slike, pri čemu redci od vrha prema dnu odgovaraju sve većem broju epoha obuke. Na taj se način prikazuje postupna evolucija kvalitete i jasnoće generiranih slika, što upućuje na učinkovitost postupka obuke u poboljšanju reprezentativnosti rezultata.



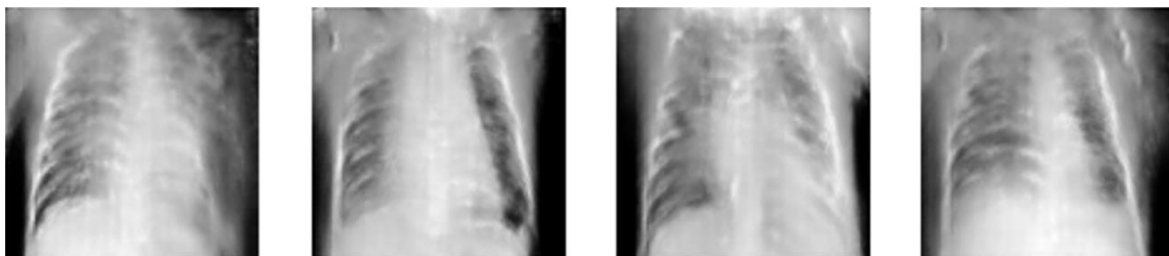
Slika 20 Primjer izvornih rendgenskih snimki zdravih pluća



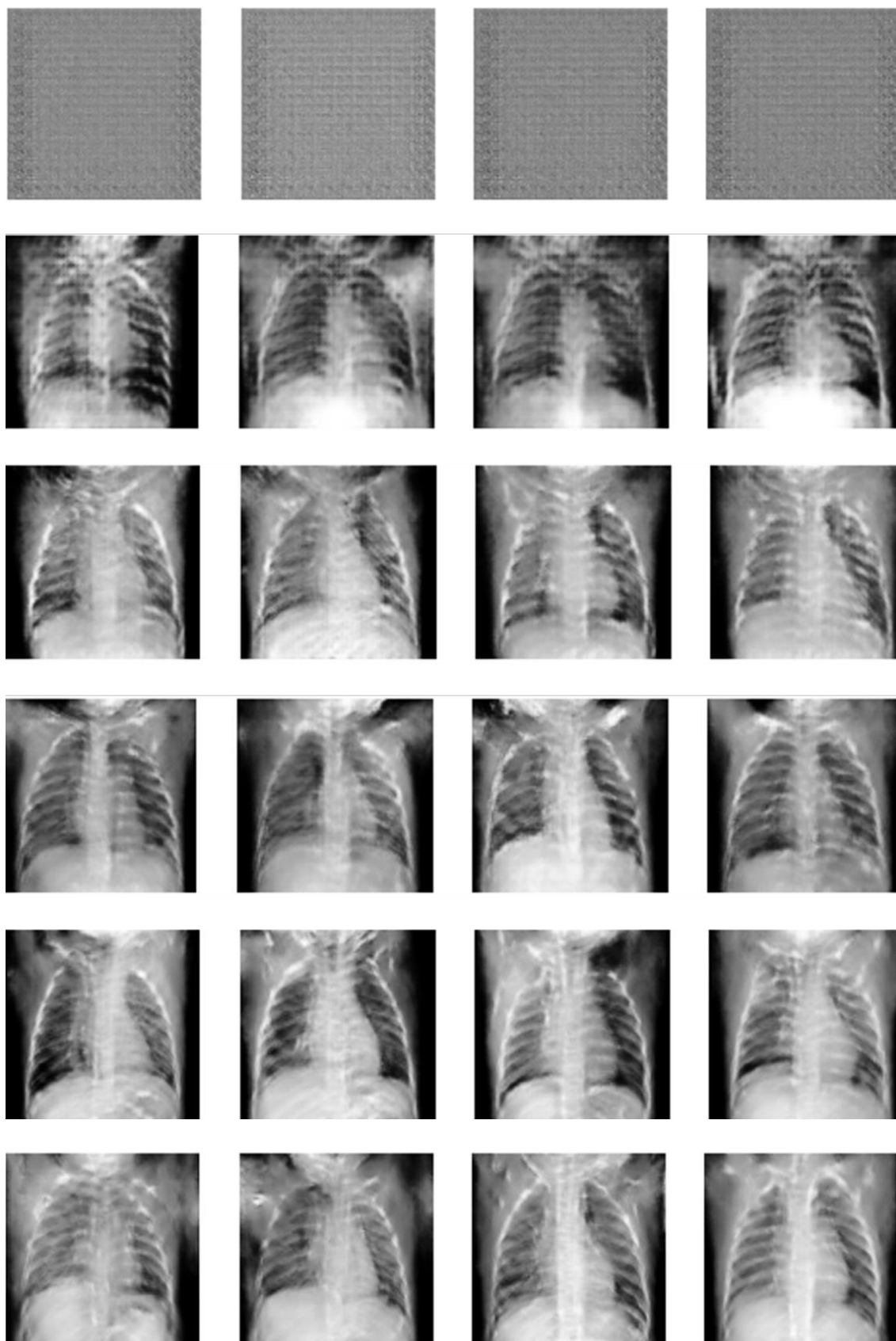
Slika 21 Primjer sintetičkih rendgenskih snimki zdravih pluća



Slika 22 Primjer izvornih rendgenskih snimki upale pluća



Slika 23 Primjer sintetičkih rendgenskih snimki upale pluća



Slika 24 Proces učenja DCGAN modela tijekom epoha

5.6. Eksperimentalni dizajn

Eksperimentalni dizajn ovog istraživanja strukturiran je kako bi se ispitao utjecaj dvaju primarnih čimbenika, veličine skupa za obuku i uključivanje sintetičkih slika, na performanse konvolucijske neuronske mreže zadužene za klasifikaciju upale pluća s rendgenskih snimaka prsnog koša. Glavni cilj bio je utvrditi može li proširenje malih skupova podataka sintetičkim uzorcima generiranim dubokom konvolucijskom generativnom suparničkom mrežom poboljšati generalizaciju modela i dijagnostičku točnost. Radi sustavnog i ponovljivog provođenja istraživanja, dizajn je oblikovan kroz nezavisne i zavisne varijable.

Primarna nezavisna varijabla istraživana u ovom radu bila je primjena generativnog proširenja podataka temeljenog na DCGAN-u. Cilj je bio ispitati može li proširenje skupa podataka za obuku sintetičkim slikama generiranim pomoću DCGAN modela poboljšati performanse klasifikatora temeljenog na konvolucijskoj neuronskoj mreži u razlikovanju rendgenskih snimaka prsnog koša oboljelih od upale pluća i onih bez oboljenja.

Za provedbu ovog postupka, prvo su obučena dva odvojena DCGAN modela: jedan na slikama zdravih pacijenata, a drugi na slikama pacijenata s upalom pluća. Svaki model bio je obučen samo jednom, a potom je korišten za generiranje velikog skupa sintetičkih slika, koje su zatim pohranjene i korištene u svim ponavljanjima eksperimenta.

Za klasifikacijski zadatak korištena je konvolucijska neuronska mreža, prethodno opisana u potpoglavlju „Klasifikacija slika“. Provedena su dva eksperimentalna uvjeta za svaku veličinu skupa podataka:

1. Osnovni uvjet: CNN je obučen isključivo na stvarnim rendgenskim snimkama.
2. Prošireni uvjet: CNN je obučen na istom skupu stvarnih slika, proširenom dodatnim sintetičkim slikama generiranim DCGAN-om, pri čemu je broj dodanih sintetičkih slika jednak broju stvarnih.

U oba uvjeta modeli su evaluirani na istom, unaprijed određenom testnom skupu stvarnih rendgenskih snimaka, čime je osigurana izravna usporedivost rezultata. Eksperimenti su višestruko ponavljani uz različite veličine skupova za obuku i različite slučajne inicijalizacije, kako bi se obuhvatila varijabilnost rezultata.

Druga nezavisna varijabla je veličina skupa za obuku. Broj stvarnih uzoraka za obuku varirao je tijekom eksperimenata. Točnije, podskupovi rastuće veličine izvučeni su iz

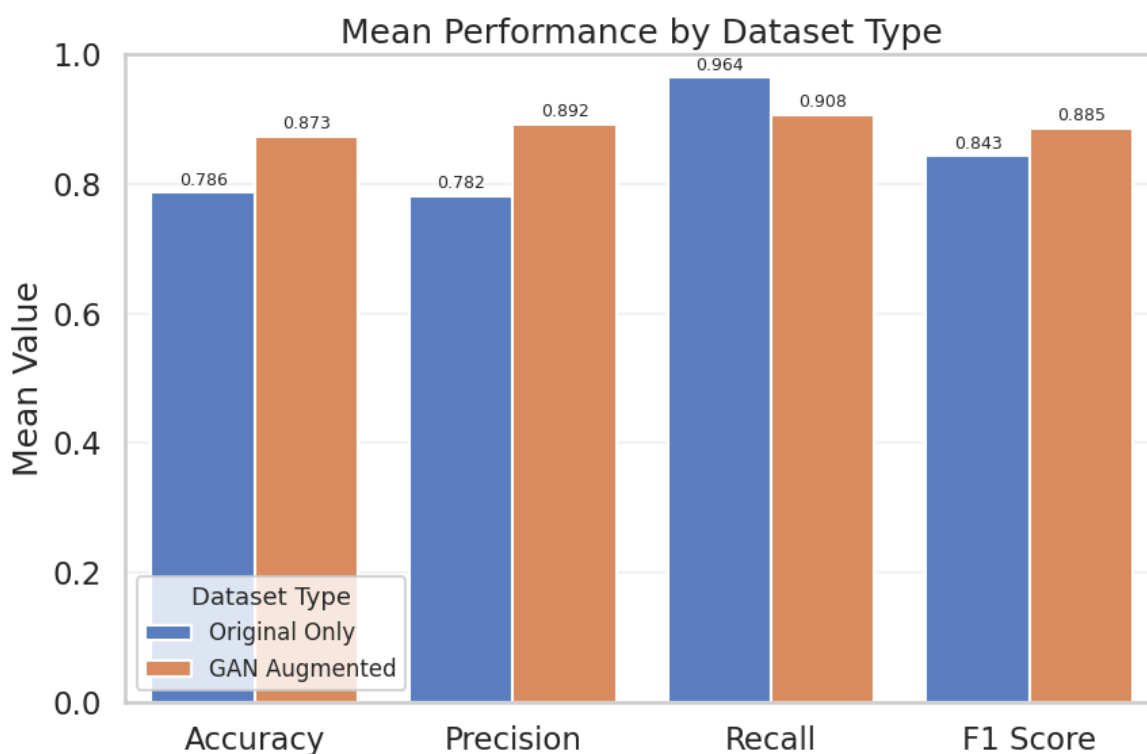
dostupnog skupa rendgenskih snimaka prsnog koša, počevši sa 100 slika i povećavajući se u koracima od 10 do 1000. Ova varijabla omogućila je izravnu procjenu kako količina podataka za obuku utječe na performanse modela, kako u osnovnim uvjetima tako i u kombinaciji sa sintetičkim proširenjem podataka.

Zavisne varijable sastoje se od metrika performansi klasifikatora, koje odražavaju utjecaj nezavisnih varijabli na učinkovitost modela. Glavni je fokus na procjeni kako generativno proširenje podataka utječe na sposobnost klasifikatora da točno razlikuje rendgenske snimke prsnog koša zdravih i oboljelih od upale pluća. U svakom eksperimentu obučavaju se dva klasifikatora, jedan na izvornom skupu podataka, a drugi na proširenom skupu podataka, te se oba potom evaluiraju na istom testnom skupu. Učinkovitost se kvantificira pomoću mjera poput točnosti, preciznosti, odziva i F1 mjere, koje su detaljno opisane u poglavlju „Konvolucijska neuronska mreža“.

6. Rezultati

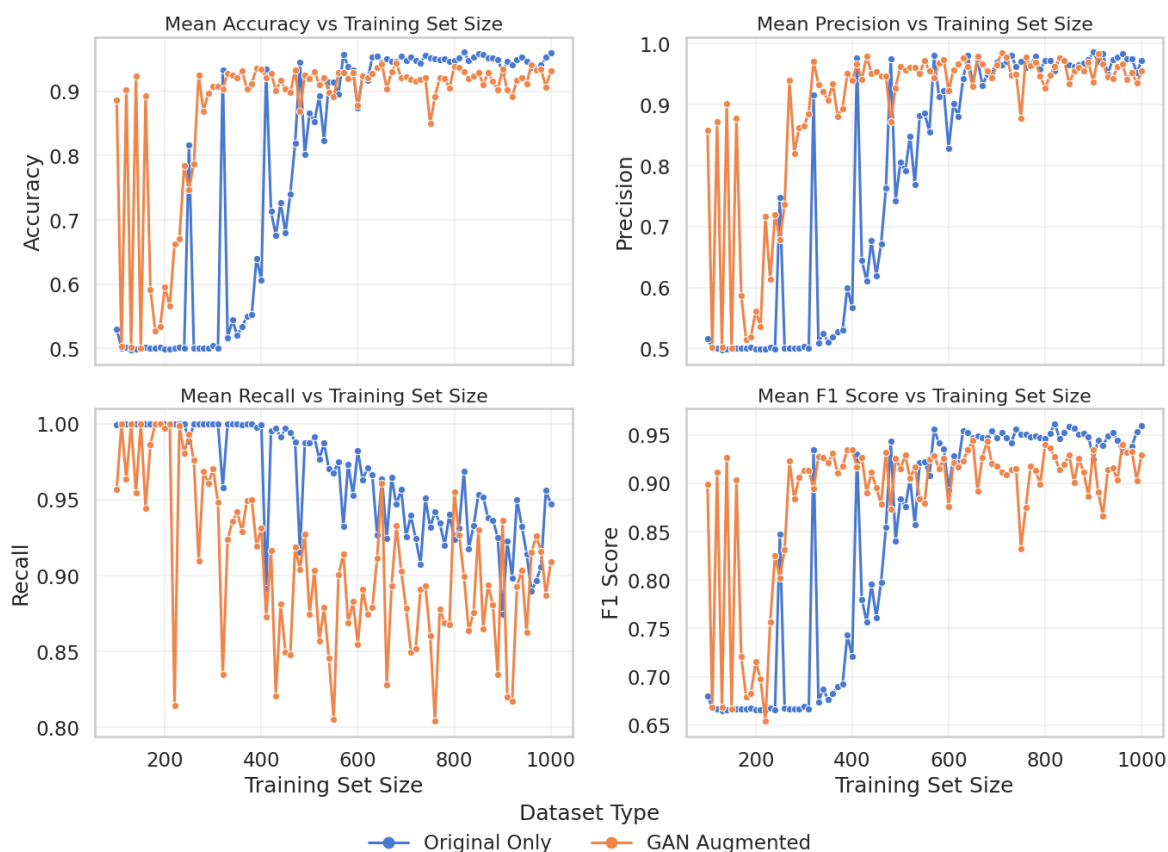
Ovo poglavlje predstavlja eksperimentalnu evaluaciju konvolucijske neuronske mreže razvijene za klasifikaciju upale pluća na temelju rendgenskih snimaka prsnog koša, s naglaskom na učinak proširenja podataka pomoću generativnih suparničkih mreža. Performanse su analizirane kroz različite veličine skupova za obuku kako bi se istaknuli učinci proširenja podataka u uvjetima ograničenih resursa. Ključne metrike, uključujući točnost, preciznost, odziv i F1 mjeru, razmatrane su zajedno s distribucijama pogrešaka, varijabilnošću između ponovljenih pokusa i dinamikom obuke, čime se pruža sveobuhvatan uvid u ponašanje modela i ulogu sintetičkih podataka u poboljšanju sposobnosti generalizacije.

Slika 25 prikazuje prosječne vrijednosti performansi u svim ponovljenim eksperimentima za dva načina obučavanja klasifikatora: *Original Only* (isključivo stvarni podaci) i *GAN Augmented* (kombinacija stvarnih i sintetičkih podataka). Analiza pokazuje da uključivanje uzoraka generiranih pomoću GAN modela povećava točnost s 0,786 na 0,873 te preciznost s 0,782 na 0,892, što jasno upućuje na smanjenje broja lažno pozitivnih klasifikacija i ukupno višu razinu ispravnosti. Ovaj porast popraćen je umjerenim smanjenjem odziva (s 0,964 na 0,908), što sugerira konzervativniji pristup klasifikaciji, s povećanim brojem lažno negativnih rezultata. Pod konzervativnijim pristupom klasifikaciji misli se da je model stroži u proglašavanju uzorka pozitivnim u odabranoj instanci. U praksi se to očituje tako da preciznost raste (manje lažno pozitivnih rezultata), dok odziv opada (više lažno negativnih rezultata). Ipak, ukupna F1 mjera bilježi porast s 0,843 na 0,885, što potvrđuje da su dobici u preciznosti značajniji od gubitaka u odzivu. Ukratko, GAN proširenje podataka ima neto pozitivan učinak na performanse modela, povećavajući selektivnost uz određeni kompromis u osjetljivosti. U narednim analizama detaljno se razmatraju navedeni prosjeci s obzirom na veličinu skupa za obuku, broj sintetički generiranih uzoraka i strukturu pogrešaka, kako bi se preciznije objasnilo gdje su poboljšanja najizraženija te što je uzrokovalo smanjenje odziva.



Slika 25 Prosječna točnost, preciznost, odziv i F1 mjera za modele obučene na izvornim podacima i na skupovima proširenima pomoću GAN-a

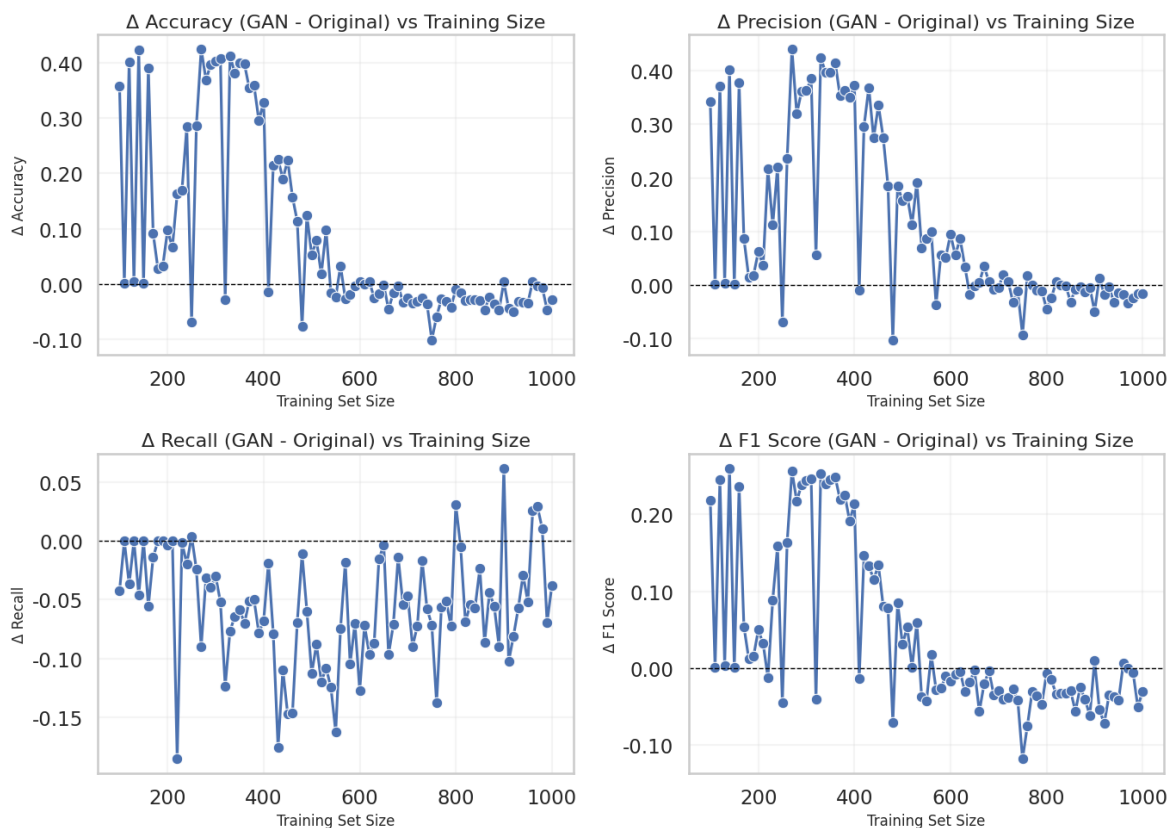
Grafovi na Slika 26 prikazuju prosječne vrijednosti točnosti, preciznosti, odziva i F1 mjere pri različitim veličinama skupa za obuku za dva načina obučavanja modela. Za manje skupove za obuku (ispod približno 300 slika), prošireni skup općenito poboljšava točnost, preciznost i F1 mjeru, što se vidi po višim i stabilnijim narančastim krivuljama na gornjem i donjem desnom grafikonu. To sugerira da proširenje podataka temeljeno na GAN-u pruža vrijednu raznolikost u scenarijima s ograničenim brojem podataka, poboljšavajući generalizaciju modela. Međutim, za veće veličine skupova za obuku, performanse oba skupa podataka konvergiraju, pri čemu model obučen samo na stvarnim slikama postiže nešto veću preciznost i F1 mjeru te pokazuje manju varijabilnost, posebno iznad 500 slika. Zanimljivo je da metrika odziva pokazuje obrnuti trend. Prošireni skup podataka dosljedno ima niže vrijednosti odziva u svim veličinama skupa, što implicira da, iako sintetičko proširenje podataka poboljšava preciznost, može smanjiti osjetljivost tako što model češće izbjegava pozitivna predviđanja. Za manje veličine skupova za obuku metrike snažno osciliraju zbog šuma uzorkovanja i nedostatka podataka, no ta se varijabilnost značajno smanjuje kako skup prelazi približno 400 uzoraka, kada se oba modela stabiliziraju s visokom točnošću i uravnoteženim kompromisom između preciznosti i odziva.



Slika 26 Prosječna točnost, preciznost, odziv i F1 mjera u odnosu na veličinu skupa za obuku za modele obučene na izvornim podacima i na skupovima proširenima pomoću GAN-a

Skup grafova na Slika 27 prikazuje razlike u metrikama između modela obučenog na proširenim podacima i modela obučenog samo na izvornim podacima, kojima se prethodne analize mogu potvrditi na još jedan način. Kod manjih skupova podataka, pozitivne delta vrijednosti za točnost, preciznost i F1 mjeru (do +0,4 za točnost i preciznost te +0,2 za F1 mjeru) naglašavaju značajan doprinos proširenja podataka performansama, dok negativne delta vrijednosti odziva potvrđuju blagi gubitak osjetljivosti. S povećanjem veličine skupa podataka, ove razlike naglo opadaju i stabiliziraju se blizu nule, što ukazuje na to da se marginalna korisnost sintetičkih podataka smanjuje kada je dostupan dovoljan broj stvarnih podataka. Varijabilnost u ovim grafovima također ilustrira nestabilnost uzrokovanu malim i potencijalno nereprezentativnim skupovima za obuku, koja se smanjuje rastom skupa podataka, što potvrđuje robusnost oba pristupa obučavanju pri većim veličinama skupa podataka. Rezultati pokazuju da proširenje podataka temeljeno na GAN-u predstavlja učinkovitu strategiju za poboljšanje performansi modela u uvjetima ograničenih skupova za obuku, osobito u pogledu povećanja preciznosti i ukupne pouzdanosti modela. Ipak, s rastom

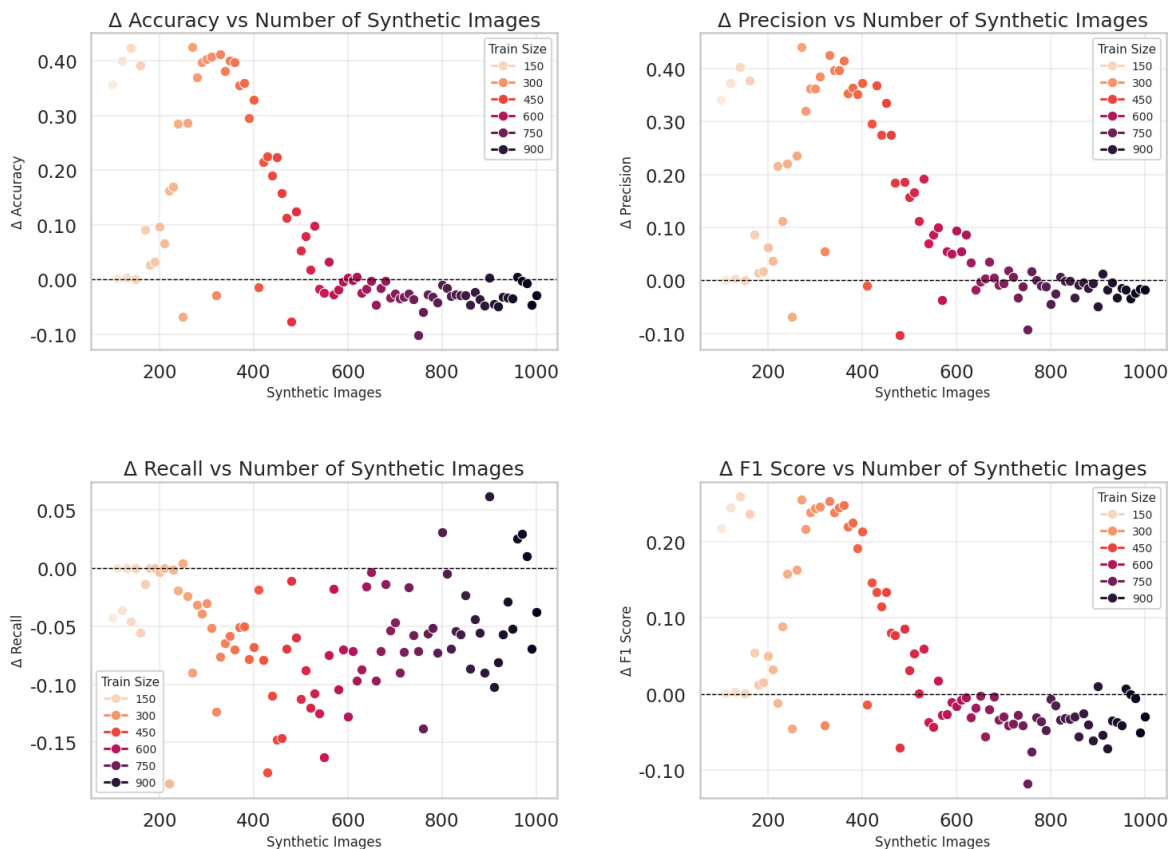
dostupnosti stvarnih podataka njezin se doprinos postupno smanjuje te, u pojedinim slučajevima, može imati blago nepovoljan učinak na odziv modela.



Slika 27 Razlika u performansama (*GAN Augmented – Original Only*) za točnost, preciznost, odziv i F1 mjeru u odnosu na veličinu skupa za obuku

Uz procjenu razlika u performansama u odnosu na ukupnu veličinu skupa za obuku, istražen je i odnos između broja sintetičkih slika i performansi modela, što je prikazano na Slika 28. Za razliku od prethodnih grafova, koji naglašavaju učinak proširenja podataka u kontekstu ukupne veličine skupa podataka, ovaj prikaz izolira izravan utjecaj količine sintetičkih slika. Svaka točka predstavlja razliku u performansama (*GAN Augmented – Original Only*) pri određenom broju sintetičkih uzoraka, dok boja označava ukupan broj stvarnih slika za obuku. Rezultati otkrivaju jasan trend: točnost, preciznost i F1 mjera dosežu vrhunac kada se koristi umjerena količina sintetičkih slika (otprilike 300-400), osobito za manje skupove za obuku (svjetlije točke), dok prekomjerno dodavanje sintetičkih podataka dovodi do slabljenja učinka, pa čak i negativnih rezultata. Ovaj učinak zasićenja sugerira da podaci generirani pomoću GAN-a mogu učinkovito nadopuniti stvarne uzorke do određene granice, no pretjerano oslanjanje na sintetičke podatke može uvesti šum ili smanjiti raznolikost skupa u odnosu na stvarne uzorke.

Metrika odziva dodatno naglašava ovaj kompromis, budući da veći broj sintetičkih slika sustavno smanjuje vrijednosti odziva za sve veličine skupova za obuku, što odražava sklonost proširenja podataka povećanju konzervativnosti modela, dajući prednost preciznosti u odnosu na osjetljivost. Uključivanje veličine skupa za obuku kao dimenzije boje također ističe da utjecaj sintetičkih podataka slabi s porastom broja stvarnih uzoraka, tamnije točke grupiraju se oko nultih delta vrijednosti, što pokazuje da proširenje podataka postaje manje važna za dobro dizajnirane skupove podataka. Zajedno s prethodnim grafovima razlika u performansama, koji su pokazali da proširenje podataka donosi najveće dobitke u uvjetima ograničenih podataka, ovi grafovi pojašnjavaju kako apsolutni broj sintetičkih uzoraka, a ne samo njihova prisutnost, utječe na klasifikacijske uzorke. Ova perspektiva posebno je vrijedna za oblikovanje strategija proširenja podataka, jer sugerira optimalnu razinu proširenja podataka u odnosu na veličinu skupa podataka, izbjegavajući pretjerano proširenje podataka koje bi moglo negativno utjecati na odziv ili sposobnost generalizacije.



Slika 28 Razlika u performansama (*GAN Augmented – Original Only*) za točnost, preciznost, odziv i F1 mjeru u odnosu na broj sintetičkih slika, pri čemu boja označava veličinu stvarnog skupa za obuku

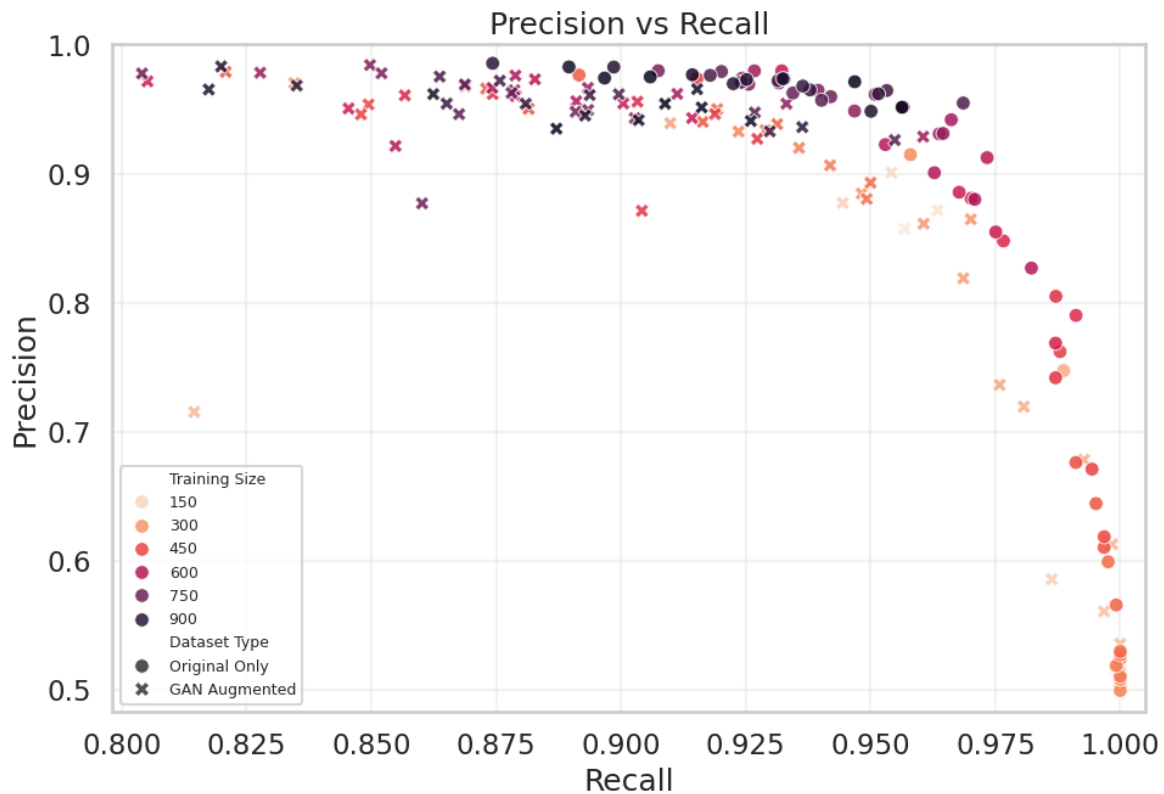
Slika 29 prikazuje prosječan broj lažno pozitivnih i lažno negativnih predikcija za modele obučene na izvornom skupu podataka u usporedbi s modelima obučenima na skupu podataka proširenom pomoću GAN-a, u ovisnosti o veličini skupa za obuku. Prikazani obrasci pružaju uvid u dinamiku odnosa između preciznosti i odziva koja je istaknuta u prethodnim grafovima. Lijevi graf pokazuje da proširenje podataka pomoću GAN-a značajno smanjuje broj lažno pozitivnih predikcija u gotovo svim veličinama skupova za obuku, osobito kada je dostupno manje od 400 stvarnih uzoraka. Nasuprot tome, model obučen samo na izvornim podacima pokazuje veći broj lažno pozitivnih predikcija pri manjim veličinama skupa, što je u skladu s nižim vrijednostima preciznosti u tom rasponu. Ovo smanjenje broja lažno pozitivnih predikcija potvrđuje ulogu proširenja podataka u poboljšanju stabilnosti granice odlučivanja, što dovodi do konzervativnijih predikcija i manjeg broja pogrešnih pozitivnih klasifikacija.

S druge strane, desni graf otkriva komplementaran učinak. Broj lažno negativnih predikcija raste u modelima obučenima s proširenim podacima, osobito kada veličina skupa prelazi približno 500 uzoraka. Ovo opažanje u skladu je s uočenim padom odziva kod obuke s proširenim podacima, potvrđujući da, iako sintetički podaci povećavaju sigurnost modela u predikcijama, to čine na račun osjetljivosti, potencijalno propuštajući više slučajeva upale pluća. Oba grafa također naglašavaju visoku varijabilnost broja lažno pozitivnih i lažno negativnih predikcija kod malih skupova za obuku, što se može pripisati nestabilnosti obuke s ograničenim brojem uzoraka. Varijabilnost se značajno smanjuje kada skup prelazi 500 uzoraka, pri čemu se performanse obje metode stabiliziraju i konvergiraju prema nižim ukupnim stopama lažno pozitivnih i lažno negativnih predikcija. Rezultati ističu da proširenje podataka pomoću GAN-a pomiče model prema točki s višom preciznošću i nižim odzivom, što može biti korisno za smanjenje broja lažnih alarma, ali zahtijeva pažljivo razmatranje u kliničkim kontekstima u kojima je osjetljivost od ključne važnosti.



Slika 29 Prosječan broj lažno pozitivnih predikcija (lijevo) i lažno negativnih predikcija (desno) u odnosu na veličinu skupa za obuku za modele obučene na izvornim podacima i na skupovima proširenima pomoću GAN-a

Slika 30 prikazuje kompromis između preciznosti i odziva za CNN modele obučene na izvornim podacima i na podacima proširenima pomoću GAN-a, pri čemu boja točaka označava veličinu skupa za obuku. Modeli obučeni s proširenjem podataka označeni su križićima, dok su modeli obučeni samo na izvornim podacima prikazani krugovima. Graf jasno pokazuje da proširenje podataka prvenstveno usmjerava model prema višoj preciznosti uz blago smanjen odziv, što je u skladu s ranijim analizama distribucija lažno pozitivnih i lažno negativnih predikcija. Ovaj je učinak najizraženiji kod manjih skupova za obuku (svjetlije točke), gdje proširenje omogućuje modelu postizanje vrijednosti blizu ili iznad 0,95, dok se odziv umjereno smanjuje, što odražava konzervativnije ponašanje u klasifikaciji.

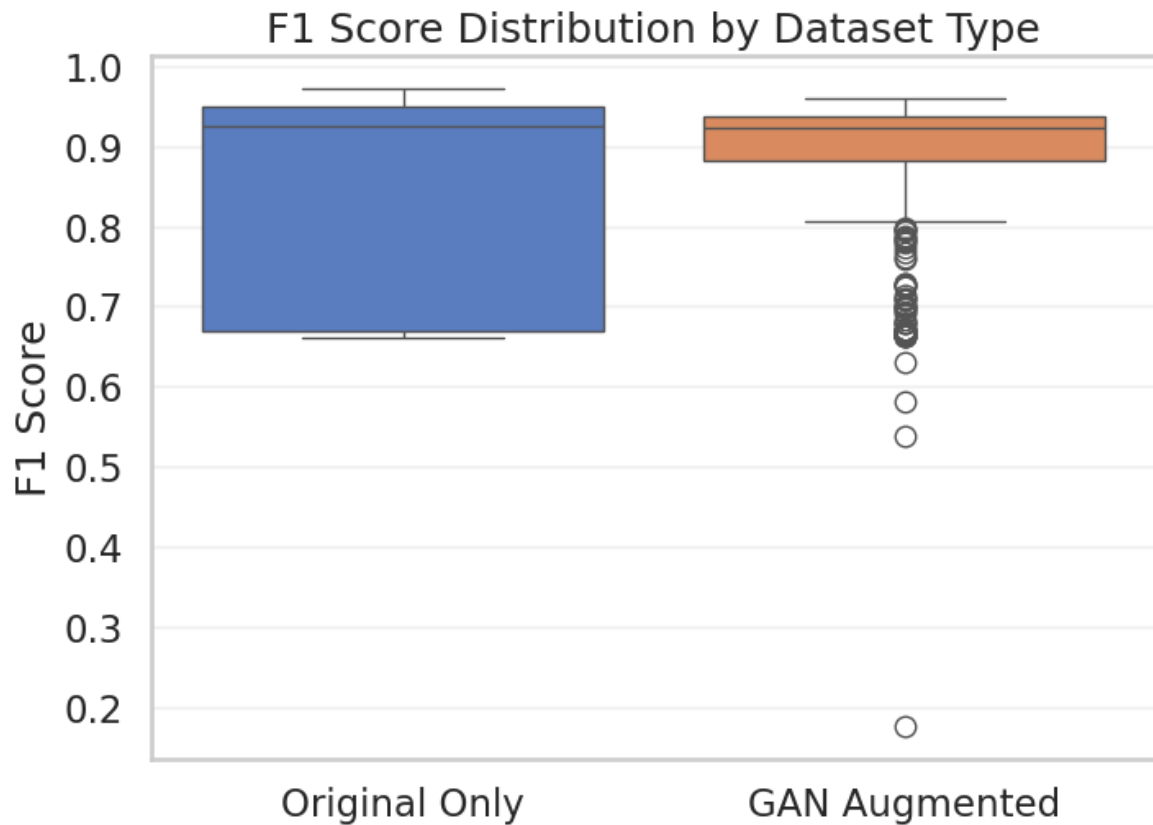


Slika 30 Odnos između preciznosti i odziva za modele obučene na izvornim podacima i na skupovima proširenima pomoću GAN-a, s bojama koje označavaju veličinu skupa za obuku

Slika 31 uspoređuje distribuciju F1 mjera za modele obučene isključivo na izvornom skupu podataka i one obučene s proširenim skupom podataka generiranim pomoću GAN-a. Dijagrami kutije i brkova pokazuju da oba pristupa postižu usporedive medijane F1 mjere, što sugerira da proširenje podataka ne mijenja drastično ukupne performanse modela kada se promatraju u cjelini. Međutim, uočljive su razlike u rasponu podataka i prisutnosti odstupajućih vrijednosti. Distribucija F1 mjera za model s proširenjem podataka pokazuje užu interkvartilni raspon, što ukazuje na dosljednije performanse modela kroz više eksperimenata, iako uključuje veći broj modela s lošijim rezultatima. To sugerira da proširenje podataka općenito stabilizira performanse, ali povremeno može dovesti do slabijih modela, potencijalno zbog šuma uzrokovanog manje realističnim sintetičkim uzorcima. Suprotno tome, model obučen samo na izvornim podacima pokazuje veću varijabilnost u središnjoj distribuciji, što odražava veću osjetljivost na slučajnost manjih skupova za obuku, ali s manje ekstremno slabih rezultata.

Ova vizualizacija potvrđuje zaključke iz prethodnih analiza. Proširenje podataka pomoću GAN-a najučinkovitije je u smanjenju varijabilnosti performansi kada je skup podataka ograničen, no njegove prednosti opadaju povećanjem veličine skupa i može dovesti do

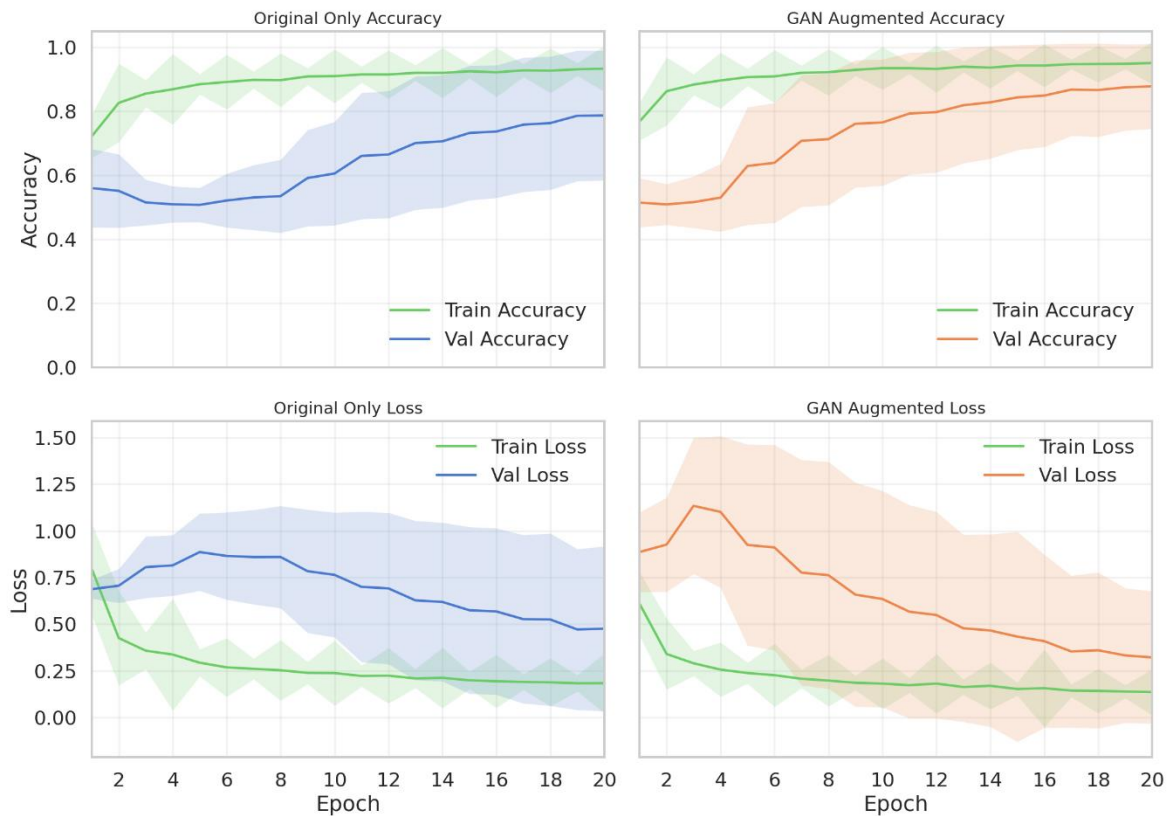
rijetkih slučajeva degradiranih performansi. U praksi, ovi rezultati naglašavaju važnost pažljivog odabira i izrade sintetičkih podataka kako bi se maksimizirala njihova vrijednost, a istovremeno smanjili rizici povezani s nestabilnošću modela.



Slika 31 Distribucija F1 mjera za modele obučene na izvornim podacima i na skupovima podataka proširenim pomoću GAN-a

Slika 32 pruža sveobuhvatan prikaz dinamike obuke i validacije za CNN modele obučene na isključivo izvornom skupu podataka i na skupu proširenom pomoću GAN-a, kombinirajući krivulje točnosti i gubitka kroz 20 epoha sa zasjenjenim područjima koja predstavljaju srednju vrijednost \pm standardnu devijaciju kroz ponovljene pokuse. Oba modela pokazuju postojan napredak u točnosti na skupu za obuku i dosljedno smanjenje gubitka, što potvrđuje učinkovito učenje u oba pristupa. Ipak, model s proširenjem podataka postiže višu točnost na validacijskom skupu i niži gubitak tijekom cijelog procesa obuke, uz sve užu varijabilnost kako epohe napreduju. Ovi trendovi potvrđuju da dodavanje sintetičkih slika ne samo da ubrzava konvergenciju, već i stabilizira rezultate obuke među ponovljenim pokusima. Nasuprot tome, model obučen isključivo na izvornim podacima pokazuje veći jaz između performansi na skupu za obuku i validaciju te veću varijabilnost, osobito u ranim epohama, što odražava njegovu osjetljivost na ograničenu količinu podataka i veću sklonost

prekomjernom prilagođavanju. Proširenjem skupa sintetičkim uzorcima, model dobiva veću raznolikost u podacima za učenje, što poboljšava njegovu sposobnost generalizacije čak i pri manjem broju stvarnih slika. Ovi grafovi sažimaju ključan zaključak da proširenje podataka pomoću GAN-a poboljšava konvergenciju, konzistentnost i sposobnost generalizacije, pružajući značajnu prednost u scenarijima s ograničenim količinama podataka.



Slika 32 Točnost i gubitak na skupu za obuku i validaciju (srednja vrijednost \pm standardna devijacija) kroz 20 epoha za modele obučene na izvornim podacima i na skupu podataka proširenom pomoću GAN-a

Zaključak

U ovom radu je istražena primjena generativnih suparničkih mreža, konkretno arhitekture DCGAN, za generiranje sintetičkih rendgenskih snimaka prsnog koša u svrhu klasifikacije upale pluća pomoću konvolucijske neuronske mreže. Primarni cilj bio je procijeniti može li proširenje podataka temeljeno na GAN-u poboljšati performanse modela, osobito u scenarijima s ograničenom dostupnošću stvarnih podataka.

Provedeni eksperimenti pokazali su da proširenje podataka slikama generiranim GAN-om donosi jasne prednosti u uvjetima ograničenih podataka, poboljšavajući točnost, preciznost i F1 mjeru uz smanjenje broja lažno pozitivnih predikcija. Međutim, ta su poboljšanja popraćenom kompromisom u odzivu, budući da modeli obučeni s proširenjem podataka teže konzervativnijim predikcijama koje daju prednost preciznosti nad osjetljivošću. S povećanjem veličine skupa za obuku razlike u performansama između modela obučenih s i bez proširenja podataka smanjile su se, što ukazuje na to da su sintetički podaci najdjelotvorniji kada stvarnih podataka nedostaje. Analiza dinamike obuke dodatno je otkrila da proširenje podataka stabilizira proces konvergencije, smanjuje varijabilnost među pokusima i omogućuje brže učenje.

Rezultati istraživanja naglašavaju važnost proširenja temeljenog na GAN-u u analizi medicinskih slika, posebno za rješavanje problema nedostatka podataka i poboljšanje robusnosti klasifikatora. Istovremeno, pad odziva ukazuje na potrebu uravnoteženih strategija proširenja i pažljivog odabira sintetičkih podataka kako bi se izbjegle nenamjerne pristranosti u performansama.

Buduća istraživanja pružaju mogućnost primjene naprednijih GAN arhitektura te proširenja eksperimenata na veće i raznovrsnije skupove podataka, što bi dodatno učvrstilo zaključke i omogućilo širu kliničku primjenjivost.

Literatura

- [1] Abhisheka, B., Biswas, S. K., Purkayastha, B., Das, D., & Escargueil, A. (2024). Recent trend in medical imaging modalities and their applications in disease diagnosis: a review. *Multimedia Tools and Applications*, 83(14), 43035-43070.
- [2] Alomar, K., Aysel, H. I., & Cai, X. (2023). Data augmentation in classification and segmentation: A survey and new strategies. *Journal of Imaging*, 9(2), 46.
- [3] Alqahtani, H., Kavakli-Thorne, M., & Kumar, G. (2021). Applications of Generative Adversarial Networks (GANs): An Updated Review: H. Alqahtani et al. *Archives of Computational Methods in Engineering*, 28(2), 525-552.
- [4] Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., . . . Farhan, L. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *Journal of big Data*, 8(1), 53.
- [5] Bhagat, V., & Bhaumik, S. (2019). Data augmentation using generative adversarial networks for pneumonia classification in chest xrays. *2019 Fifth International Conference on Image Information Processing (ICIIP)* (str. 574-579). IEEE.
- [6] Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., & Bharath, A. A. (2018). Generative adversarial networks: An overview. *IEEE signal processing magazine*, 35(1), 53-65.
- [7] Ganesan, P., Rajaraman, S., Long, R., Ghoraani, B., & Antani, S. (2019). Assessment of data augmentation strategies toward performance improvement of abnormality classification in chest radiographs. *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (str. 841-844). IEEE.
- [8] Goodfellow, I. J., Shlens, J., & Szegedy, C. (2015). Explaining and harnessing adversarial examples. *International Conference on Learning Representations*.
- [9] Hossin, M., & Sulaiman, M. N. (2015). A review on evaluation metrics for data classification evaluations. *International journal of data mining & knowledge management process*, 5(2), 1.

- [10] Kumar, T., Brennan, R., Mileo, A., & Bendeche, M. (2024). Image data augmentation approaches: A comprehensive survey and future directions. *IEEE Access*.
- [11] Kundu, R., Das, R., Geem, Z. W., Han, G.-T., & Sarkar, R. (2021). Pneumonia detection in chest X-ray images using an ensemble of deep learning models. *PloS one*, 16(9), e0256630.
- [12] Makhoulf, A., Maayah, M., Abughanam, N., & Catal, C. (2023). The use of generative adversarial networks in medical image augmentation. *Neural Computing and Applications*, 35(34), 24055-24068.
- [13] Mooney, P. T. (2018). *Chest X-Ray Images (Pneumonia)*. Preuzeto 06. 02 2025 iz Kaggle: <https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia>
- [14] Nightingale, A. (2009). A guide to systematic literature reviews. *Surgery (Oxford)*, 27(9), 381-384.
- [15] Organization, W. H. (11. 11 2022). *Pneumonia in children*. Preuzeto 19. 8 2025 iz World Health Organization: <https://www.who.int/news-room/fact-sheets/detail/pneumonia>
- [16] Poola, R. G., Pl, L., & Yellampalli, S. S. (2024). SyntDiaNet: Integrating feature extraction, transfer learning and classifier-embedded generative adversarial network for advanced pneumonia diagnosis. *Multimedia Tools and Applications*, 1-39.
- [17] Purwono, P., Ma'arif, A., Rahmانيar, W., Fathurrahman, H. I., Frisky, A. Z., & ul Haq, Q. M. (2022). Understanding of convolutional neural network (cnn): A review. *International Journal of Robotics and Control Systems*, 2(4), 739-748.
- [18] Qin, X., Bui, F. M., Nguyen, H. H., & Han, Z. (2022). Learning from limited and imbalanced medical images with finer synthetic images from gans. *Ieee Access*, 10, 91663-91677.
- [19] Radford, A., Metz, L., & Chintala, S. (2015). Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*.
- [20] *RSNA Pneumonia Detection Challenge Dataset*. (2018). Preuzeto 06. 02 2025 iz Kaggle: <https://www.kaggle.com/c/rsna-pneumonia-detection-challenge/data>

- [21] Sashank, M. S., Maddila, V. S., Boddu, V., & Radhika, Y. (2022). Efficient deep learning based data augmentation techniques for enhanced learning on inadequate medical imaging data. *ACTA IMEKO*, 11(1), 6.
- [22] Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. *Journal of big data*, 6(1), 1-48.
- [23] Tawsifur, R., Muhammad, C., & Amith, K. (2020). *COVID-19 Radiography Database*. Preuzeto 06. 02 2025 iz Kaggle: <https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database>
- [24] Vyas, A., Yu, S., & Paik, J. (2018). *Multiscale transforms with application to image processing* (Svez. 1). Springer.
- [25] Wang, Z., Hall, J., & Haddad, R. J. (2021). Improving pneumonia diagnosis accuracy via systematic convolutional neural network-based image enhancement. *SoutheastCon 2021* (str. 1-6). IEEE.
- [26] Yamashita, R., Nishio, M., Do, R. K., & Togashi, K. (2018). Convolutional neural networks: an overview and application in radiology. *Insights into imaging*, 9(4), 611-629.
- [27] Zha, D., Bhat, Z. P., Lai, K. H., Yang, F., Jiang, Z., Zhong, S., & Hu, X. (2025). Data-centric artificial intelligence: A survey. *ACM Computing Surveys*, 57(5), 1-42.

Skraćenice

AUC-ROC	<i>Area under the ROC Curve</i>	područje ispod ROC krivulje
CE-GAN	<i>Classifier-Embedded GAN</i>	
CNN mreža	<i>Convolutional Neural Network</i>	konvolucijska neuronska
DCGAN generativna suparnička mreža	<i>Deep Convolutional GAN</i>	duboka konvolucijska
DCNN neuronska mreža	<i>Deep Convolutional Neural Network</i>	duboka konvolucijska
FN	<i>False Negative</i>	lažno negativno
FP	<i>False Positive</i>	lažno pozitivno
GAN	<i>Generative Adversarial Network</i>	generativna suparnička mreža
GNB	<i>Gaussian Naive Bayes</i>	
HSV	<i>Hue, Saturation, Value</i>	
NST	<i>Neural Style Transfer</i>	neuronski prijenos stila
ProGAN / PGGAN	<i>Progressive Growing GAN</i>	
RGB	<i>Red, Green, Blue</i>	
SAR	<i>Synthetic Aperture Radar</i>	radar sa sintetičkom aperturom
SVM	<i>Support Vector Machine</i>	stroj potpornih vektora
TN	<i>True Negative</i>	stvarno negativno
TP	<i>True Positive</i>	stvarno pozitivno
WOS	<i>Web of Science</i>	