

Otkrivanje i klasifikacija lezija dojki na mamografskim snimkama pomoću dubokog učenja

Pintarić, Karlo

Master's thesis / Diplomski rad

2023

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, School of Medicine / Sveučilište u Zagrebu, Medicinski fakultet**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:105:029747>

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

Download date / Datum preuzimanja: **2025-02-23**



Repository / Repozitorij:

[Dr Med - University of Zagreb School of Medicine Digital Repository](#)



SVEUČILIŠTE U ZAGREBU
MEDICINSKI FAKULTET

Karlo Pintarić

**Otkrivanje i klasifikacija lezija dojki na
mamografskim snimkama pomoću
dubokog učenja**

DIPLOMSKI RAD



Zagreb, 2023.

Ovaj diplomski rad izrađen je na Katedri za medicinsku statistiku, epidemiologiju i medicinsku informatiku pod vodstvom doc.dr.sc Andrija Štajduhar i predan je na ocjenu u akademskoj godini 2022./2023.

KAZALO POJMOVA

AI (engl. *artificial intelligence*), umjetna inteligencija.

BI-RADS Breast Imaging-Reporting and Data System.

CAD (engl. *computer-aided diagnosis*), računalno potpomognuta dijagnostika.

CBIS-DDSM Curated Breast Imaging Subset of the Digital Database for Screening Mammography.

CNN (engl. *convolutional neural networks*), konvolucijske neuronske mreže.

DNN (engl. *deep neural networks*), duboke neuronske mreže.

MLP (engl. *multilayer perceptron*), višeslojni perceptron.

NLP (engl. *natural language processing*), obrada prirodnog jezika.

ViT Vision Transformer.

SADRŽAJ

1	Uvod	1
2	Mamografija	2
2.1	Radiološko oslikavanje dojki	2
2.2	Lezije dojki	3
2.3	Značaj ranog otkrivanja raka dojke	6
2.4	CAD sistemi i AI kod mamografije	7
3	Duboko učenje za računalni vid	9
3.1	Od ideje to primjene	9
3.2	Principi dubokog učenja	11
3.2.1	Model	11
3.2.2	Optimizacija	13
3.2.3	Prenaučenost i podnaučenost	16
3.3	Digitalni prikaz slika	17
4	Izrada modela za segmentaciju lezija dojki	20
4.1	Cilj	20
4.2	Korištene tehnologije	20
4.3	Model	21
4.4	Skup podataka	22
4.5	Mjere evaluacije modela	23
4.6	Razvoj modela	25
4.6.1	Obrada slika	25
4.6.2	Treniranje modela	25
4.6.3	Validacija modela	27
4.7	Rezultati	28
4.8	Rasprava	31
5	Zaključak	33
	Literatura	34
	Životopis	41

SAŽETAK

Otkrivanje i klasifikacija lezija dojki na mamografskim snimkama pomoću dubokog učenja

Karlo Pintarić

Rak dojke, kao najčešće dijagnosticirana vrsta raka, predstavlja veliki globalni zdravstveni izazov. Rano otkrivanje, u kojem mamografija igra ključnu ulogu, ključno je za smanjenje stope smrtnosti. No, razlikovanje između benignih i malignih lezija, koje se vizualno razlikuju po konturama na mamografskim snimkama, predstavlja značajan izazov. Da bi se ovaj proces olakšao i standardizirao, uveden je BI-RADS sustav kategorizacije lezija. Unatoč dugotrajnom korištenju računalom potpomognute dijagnostike (CAD) u analizi mamografskih snimaka, postignuti rezultati još uvijek nisu na očekivanoj razini. Upravo se tu otvara prostor za primjenu dubokog učenja, segmenta strojnog učenja temeljenog na umjetnim neuronskim mrežama. Ova tehnologija ima potencijal značajno olakšati i poboljšati rad radiologa. U sklopu ovog rada, razvijen je model dubokog učenja namijenjen za segmentaciju lezija dojki. Ovaj model analizira svaki pojedinačni piksel na slici te ga svrstava u jednu od tri kategorije: pozadina, benigna ili maligna lezija. Iako je model treniran na relativno ograničenom skupu podataka, postignuti rezultati su obećavajući - uspješno segmentira lezije bez obzira na njihovu veličinu i postiže značajan rezultat prema Dice koeficijentu. Suvremene tehnologije, kao što je duboko učenje, otvaraju nova vrata u području mamografije. No, potrebno je nastaviti s daljnjim kliničkim istraživanjima kako bi se potpuno otkrio njihov potencijal. Ovaj rad doprinosi tom cilju, istražujući mogućnosti koje pruža spoj zdravstva i tehnologije.

Ključne riječi: rak dojke, mamografija, duboko učenje, segmentacija slika

SUMMARY

Detection and classification of breast lesions on mammography images using deep learning

Karlo Pintarić

Breast cancer, as the most commonly diagnosed type of cancer, poses a significant global health challenge. Early detection, in which mammography plays a crucial role, is essential for reducing mortality rates. However, distinguishing between benign and malignant lesions, which visually differ in contours on mammographic images, presents a substantial challenge. To standardize this process, the BI-RADS system for categorizing lesions was introduced. Despite the longstanding use of computer-aided diagnostics (CAD) in the analysis of mammographic images, the results achieved are still not at the expected level. This is where the application of deep learning, a segment of machine learning based on artificial neural networks, comes into play. This technology has the potential to significantly ease and enhance the work of radiologists. In the scope of this thesis, a deep learning model intended for the segmentation of breast lesions was developed. This model analyzes each individual pixel in the image and classifies it into one of three categories: background, benign lesion, or malignant lesion. Although the model was trained on a relatively limited data set, the results achieved are promising - it successfully segments lesions regardless of their size and achieves a significant result according to the Dice coefficient. Modern technologies, such as deep learning, are opening new doors in the field of mammography. However, it is necessary to continue with further clinical research to fully uncover their potential. This thesis contributes to this goal, exploring the possibilities offered by the intersection of healthcare and technology.

Keywords: breast cancer, mammography, deep learning, image segmentation

1. Uvod

Rak dojke predstavlja globalni zdravstveni izazov s porastom incidencije diljem svijeta. Prema podacima Svjetske zdravstvene organizacije (WHO), u 2020. godini premašio je rak pluća kao najčešće dijagnosticiranu vrstu raka te rezultira najvećim brojem izgubljenih godina života prilagođenih na nesposobnost (DALYs) među ženama (1). U Republici Hrvatskoj, situacija je slična globalnim trendovima. Rak dojke je najčešći maligni tumor u žena, s 2869 zabilježenih slučajeva u 2020. godini (2). Ovi podaci ističu važnost ranog otkrivanja ove bolesti pomoću dijagnostičkih alata kao što su mamografija ili magnetska rezonanca. U tom kontekstu, dijagnostički kapaciteti postaju od velike važnosti. Međutim, suočavamo se s izazovom nedostatka kvalificiranog medicinskog osoblja, prvenstveno liječnika. Ovaj deficit može rezultirati dužim vremenima čekanja na pregled, rezultirajući ozbiljnim posljedicama za pacijente otkrivanjem bolesti u podmaklom stadiju. Tehnološki napredak olakšava i ubrzava procese u mnogim područjima, a uloga tehnologije i njene primjene u zdravstvu postaju sve važnije. Konkretno, umjetna inteligencija i njene grane poput strojnog učenja poprimaju sve značajniju ulogu u medicini, dijagnostici i podršci radu radiologa.

Algoritmi obrade i analize slika već se duže vrijeme primjenjuju na medicinskim podacima. Značajan uspjeh postigli su algoritmi strojnog učenja, matematički modeli koji koriste podatke za donošenje odluka bez eksplicitnog programiranja (3). No, za rad klasičnih algoritama strojnog učenja često je potrebna predobrada slika, ručno dizajniranje značajki i drugi koraci pripreme "sirovih" podataka prije samog modeliranja, što može biti vremenski zahtjevno i podložno greškama, posebno kada se radi o kompleksnim podacima. S druge strane, duboko učenje, podgrana strojnog učenja koja doživljava značajan napredak u posljednjih nekoliko godina, pruža rješenje za ovaj problem. U podlozi dubokog učenja su umjetne neuronske mreže, koje omogućuju automatsko stvaranje korisnih značajki slika (4, 5).

Integriranjem dubokog učenja u procese mamografije, radiolozima bi se mogao značajno olakšati i ubrzati rad, omogućujući im da se usmjere na kompleksnije slučajeve, doprinoseći većoj učinkovitosti zdravstvenog sustava. U okviru ovog diplomskog rada, istražuje se primjena dubokog učenja na radiološkim snimkama te daje pregled relevantnih studija i primjena. Konačno, kao primjer korištenja umjetne inteligencije u specifičnom zadatku, razvijen je i vlastiti model za detekciju i klasifikaciju lezija na mamografskim snimkama dojki.

2. Mamografija

2.1. Radiološko oslikavanje dojki

Mamografija je neinvazivna radiološka metoda koja koristi niskodozno ionizirajuće zračenje za vizualizaciju unutrašnjosti dojki. Prva primjena mamografije u medicinske svrhe datira iz 1913. godine, kada je kirurg Albert Salomon koristio rendgenske snimke za analizu izvađenih dojki. Međutim, moderna mamografija razvila se tek sredinom 20. stoljeća kada su se počeli koristiti posebno konstruirani mamografi. U tom razdoblju, Jacob Gershon-Cohen, pionir u području radiologije, postigao je značajne napretke u mamografskoj tehnici i dijagnostici. Dodatno, tijekom 1960-ih i 1970-ih, Robert L. Egan unaprijedio je mamografsku tehniku uvođenjem specifičnih protokola za snimanje, interpretaciju slika i standarde za kvalitetu slika i dozama zračenja, čime je pridonio poboljšanju sigurnosti i efikasnosti ove dijagnostičke metode (6, 7).

Dojka se sastoji od različitih vrsta tkiva koje djeluju zajedno kako bi ostvarile funkciju proizvodnje i transporta mlijeka do bradavice. Žljezdano tkivo je podijeljeno na 15-20 režnjeva (lobula), koji su građeni od granajućeg sustava kanalića (duktusa). Većina lezija dojki proizlazi iz terminalne duktulo-lobularne jedinice, koja čini osnovnu jedinicu građe dojke (8).

Na mamografskim snimkama, različite strukture dojki prikazuju se u različitim nijansama sive. Koštano tkivo i kalcifikacije unutar dojke prikazat će se kao bijele ili svijetle točke, masno tkivo je obično tamno, dok su normalno žljezdano tkivo i tumori srednje sive boje. Veličina, oblik, rubovi i gustoća tumora su ključni faktori koji pomažu u razlikovanju benignih i malignih lezija. Na primjer, tumori s nepravilnim oblicima ili grubim, spikuliranim rubovima često su povezani s malignim lezijama. S druge strane, tumori s glatkim rubovima obično su benigni (9). Na snimkama mogu biti vidljive i različite kalcifikacije.

Za standardizaciju izvješćivanja mamografskog, ultrazvučnog i MR nalaza dojki, Američka udruga radiologa (ACR) razvila je Breast Imaging-Reporting and Data System (BI-RADS) klasifikaciju. Klasifikacija uvodi standardne pojmove za opisivanje i procjenu rizika masa, kalcifikacija i drugih lezija. Prema klasifikaciji, rezultati snimanja raspoređuju se u sedam skupina, označenih kao BI-RADS 0 do 6. BI-RADS kategorija 1 upućuje na to da su nalazi bez ikakvih lezija, a kategorije 2 i 3 označavaju benigni, odnosno vjerojatno benigni nalaz. Sumnja na malignitet raste s kategorijama BI-RADS 4 i 5, dok kategorija BI-RADS 6 predstavlja malignitet

potvrđen biopsijom. BI-RADS 0 zahtijeva dodatne slikovne pretrage (Tablica 1). Klasifikacija također uvodi kategorizaciju gustoće dojki u četiri tipa prema količini žljezdanog i masnog tkiva (7, 10).

Tablica 1. BI-RADS kategorije (7)

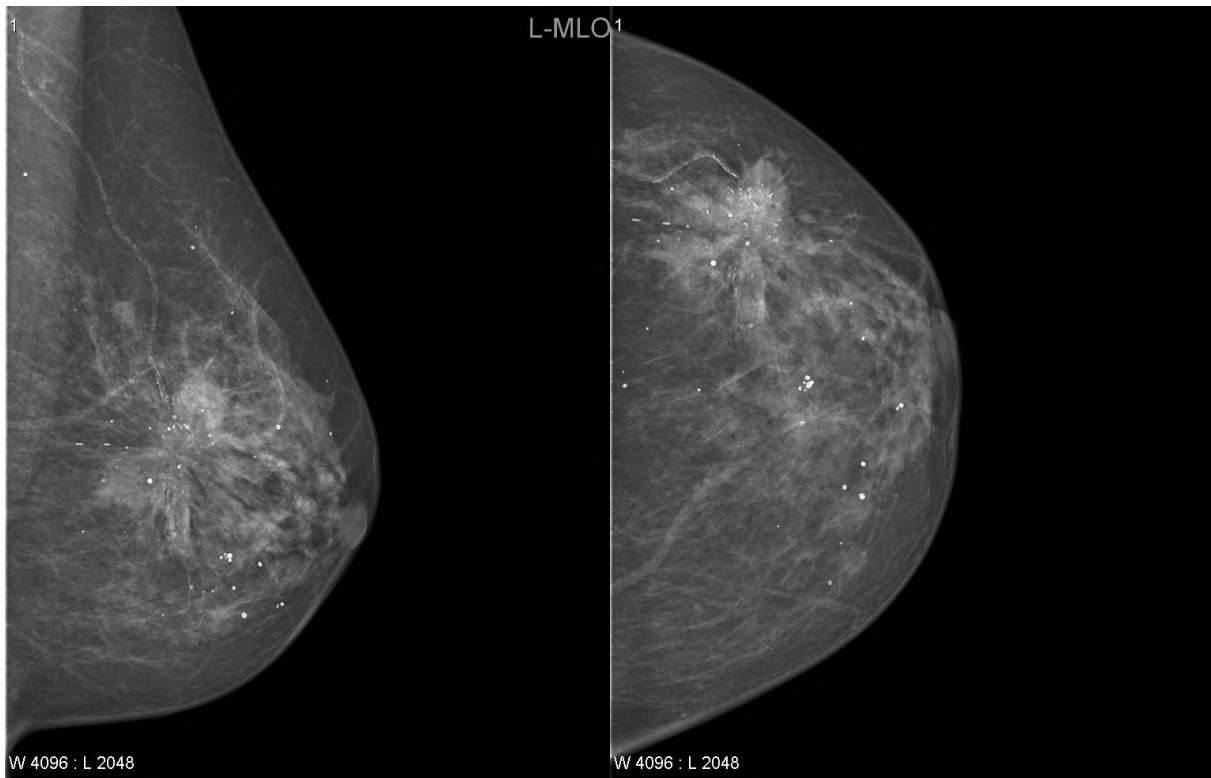
Kategorija	Daljnja procedura	Vjerojatnost karcinoma
0: Neodređeni nalaz	Potrebne su dodatne pretrage	-
1: Uredan nalaz	Rutinski probir	$\approx 0\%$
2: Benigni nalaz	Rutinski probir	$\approx 0\%$
3: Vjerojatno benigni nalaz	Kontrolni pregled u roku 6 mj.	$\leq 2\%$
4: Suspekti nalaz	Biopsija	2 – 94%
5: Vrlo suspekti nalaz	Biopsija	$\geq 95\%$
6: Malignitet potvrđen biopsijom	Potrebno je poduzeti odgovarajuće korake	-

Tijekom uobičajenog postupka snimanja, pacijenti se postavljaju u poziciju za snimanje dviju mamografskih projekcija: kraniokaudalnu (CC) i mediolateralnu kosu (MLO). Ove dvije projekcije su standardne projekcije korištene prilikom provođenja probira za rak dojke. CC pogled je vertikalna projekcija, dok MLO pogled prikazuje dojku pod kutom od 45 stupnjeva (Slika 1). MLO projekcija je izuzetno važna jer pruža najviše informacija o lateralnom dijelu dojke, posebno gornjem lateralnom kvadrantu, koji obuhvaća najveći volumen žljezdanog tkiva. Upravo u tom dijelu najčešće se javljaju maligne promjene. Oba pogleda pružaju različite perspektive unutrašnjosti i obično se koriste zajedno za sveobuhvatnu evaluaciju (11, 12).

2.2. Lezije dojki

Većina lezija dojki su benigne prirode, češće kod žena u reproduktivnoj dobi radi cikličkih hormonskih promjena (14). Radiološki ih karakteriziraju jasni znakovi poput oštro omeđenih kontura, okrugli oblik ponekad s makrokalcifikacijama. Oštra ocrtanost kontura je najvažniji pokazatelj benignosti jer ovakve mase imaju manje od 4% šanse da budu maligne (15). U tipične benigne mase (BI-RADS 2) pripadaju lezije kao što su fibroadenom, kalcificirane ciste te lezije koje sadrže mast poput uljnih cisti i lipoma (9).

Fibroadenom je najčešći benigni tumor, s vrhuncem incidencije u dobi između 15 i 35 godina (14). Radi se fibroepitelnoj proliferaciji tkiva dojke (8). Ova benigna tvorba može postići veličinu do 4 cm te se ponekad otkrije palpacijom. Ako je asimptomatska, vidljiva je na mamografiji kao ovalna tvorba, jasno definiranih rubova. Patognomoničan nalaz za fibroadenom

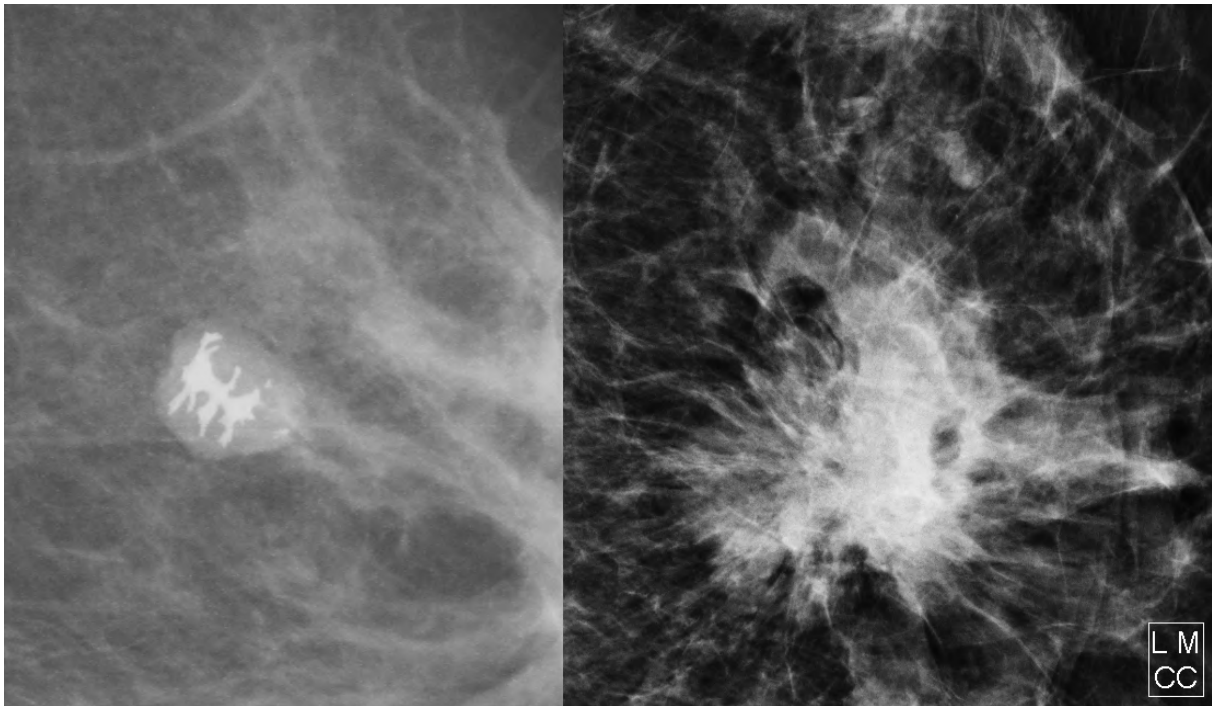


Slika 1. Standardne mamografske projekcije: MLO (lijevo), CC (desno). Autor: Frank Gaillard, Radiopaedia.org, rID: 12608 (13)

su tzv. "popcorn" kalcifikacije (9). Uljne ciste nastaju kao posttraumatska promjena pri kojoj dolazi do oštećenja masnih stanica. Sadrže fibroznu kapsulu i kalcifikacije kao posljedica saponifikacije masnih kiselina. Precipitacijski kalcij na kapsuli mogu poprimiti izgled poput "ljuske jajeta", što je patognomičan znak. Intramamarni limfni čvorovi mogu biti vidljivi na mamografiji kao okrugle ili ovalne tvorbe na kojima se jasno razlikuju hilum i korteks (9). Spikulirane benigne lezije su rijetkost, no zahtijevaju dodatnu obradu radi isključivanja maligne promjene. U podlozi ovakve lezije mogu biti razni ožiljci ili sklerozirajuća adenozna. Ova promjena karakterizirana je intralobularnom fibrozom i proliferacijom te se može zamijeniti za karcinom dojke (8, 9). Filodes tumori su rijetki bifazični tumori s epitelnom i vezivnom komponentom. Javljaju se kasnije od fibroadenoma te ih je mamografski i histološki teško razlikovati. Klinički se prikazuju kao palpabilne tvorbe, ili dovode do deformacija tkiva dojke ako su veliki. Važnost razlikovanja fibroadenoma i filodes tumora leži u činjenici da 15% posto svih filodes tumora je maligno (8).

Karcinomi čine veliku većinu malignih lezija dojke. Radi se o epitelnim tumorima koji nastaju proliferacijom stanica duktulo-lobularne jedinice. Ako stanice karcinoma probiju bazalnu membranu, takav karcinom nazivamo invazivnim, dok se pojam *in situ* koristi za karcinome s očuvanom bazalnom membranom. Najčešća vrsta neinvazivnog karcinoma je duktalni karcinom *in situ*. Ovaj tumor ne pokazuje nikakve simptome, no zbog sklonosti stvaranja mikrokalci-

fikacija, jasno je vidljiv na mamografiji (8). Invazivne karcinome dojke ovisno o histolškim karakteristikama dijelimo na: invazivni karcinom dojke ne-specijalnog tipa (engl. *no special type*, NST), lobularni invazivni i ostali rijedi oblici invazivnog karcinoma. Invazivni karcinomi karakterizirani su sporim rastom, što dovodi do opsežne dezmodoplastične reakcije. Radi toga će takvi tumori imati zvjezdasti oblik, nejasnih rubova s središnjim prolongacijama koje nazivamo spikulima (Slika 2). Mogući je i nalaz mikrokalcifikata. Takve tvorbe čini glavninu leziju klasificiranih kao BI-RADS kategorija 5 (9).



Slika 2. Benigna (lijevo) i maligna (desno) lezija. Autori: Mohammadtaghi Niknejad, rID: 163936; Giorgio M Baratelli, rID: 29466, Radiopaedia.org, (16, 17)

Trećina svih lobularnih invazivnih karcinoma vidljiva je na mamografiji kao spikulirana lezija. Ovaj tumor izaziva slabiju dezmodoplastičnu reakciju te je moguća situacija da uopće ne bude makroskopski vidljiv. Razlog tome je mogućnost rasta između normalnog žljezdanog tkiva dojke bez narušavanja vidljive arhitekture. Najčešće je kod takvog oblika rasta vidljiva asimetrija tkiva (9). Tubularni karcinom je rijedak oblik invazivnog karcinoma kojeg karakterizira izvrsna prognoza (8). Na mamografiji je vidljiv kao mala, zvjezdasta spikulirana lezija s mogućim kalcifikatima (9). Ukoliko invazivni tumori ubrzano rastu mogući je potpuni izostanak dezmodoplastične reakcije. Radi toga takve lezije ne će imati spikulirane rubove na mamografiji, već će se prikazati kao okrugle, ovalne ili lubolirane tvorbe s jasno ograničenom konturom. Obični invazivni karcinom može pokazivati ovakve karakteristike, zajedno sa rjeđim tipovima poput medularnog, mucinoznog i papilarnog karcinoma (9).

2.3. Značaj ranog otkrivanja raka dojke

Rak dojke često se razvija bez vidljivih simptoma, posebice u ranoj fazi. Stoga je ključno otkriti maligne promjene što je ranije moguće, prije regionalnog širenja ili metastaziranja raka. Rano otkrivanje raka dojke može se postići različitim pristupima. Samopregled dojki jest jedna od metoda koja je dostupna svim ženama i koja je korisna u otkrivanju neobičnih promjena. Međutim, iako je koristan, samopregled dojki nije dovoljno pouzdan za rano otkrivanje, osobito u voluminoznijim dojkama. Klinički pregled dojki od strane stručnjaka pruža nešto bolju točnost u otkrivanju abnormalnosti. Ipak, iako je klinički pregled važan, on također ima svoje ograničenja, poput niske osjetljivosti (18). Razvitkom i unaprijeđenjem tehnika oslikavanja dojki, samopregled i klinički pregled izgubili su na vrijednosti.

Mamografija se pokazala jedinom metodom ranog otkrivanja raka dojke koja dovodi do smanjenja smrtnosti (18). Nekoliko je istraživanja pokazalo da mamografija može smanjiti smrtnost od raka dojke za oko 20% (19, 20). Meta-analiza iz 2020. godine, koja je obuhvatila 27 međunarodnih kohortnih istraživanja, pokazala je da poziv na mamografiju u sklopu programa ranog otkrivanja raka dojke, može dovesti do relativnog smanjenja mortaliteta kod žena za 22% (RR 0.78, 95% CI 0.75–0.82). Za žene koje se odazovu pozivu, redukcija mortaliteta iznosi 33% (RR 0.67, 95% CI 0.61–0.75) (21).

U mnogim zemljama, uključujući Hrvatsku, provode se organizirani nacionalni programi probira za rak dojke pomoću mamografije (2). Ovi programi su od vitalnog značaja u borbi protiv ovog raka, jer omogućuju otkrivanje u ranoj fazi, kada su mogućnosti liječenja i šanse za preživljavanje najbolje. Osim boljeg preživljenja, potrebno je sagledati i ekonomsku stranu liječenja. Sustavni pregled iz 2019. godine, koji je obuhvatio istraživanja o troškovima liječenja raka dojke u 10 država s različitim stupnjevima socio-ekonomskog razvoja, procijenio je da prosječni trošak liječenja za prvi stadij bolesti iznosi \$29,724, uzimajući u obzir vrijednost dolara iz 2015. godine. Za bolesti u naprednijim stadijima, II, III i IV, troškovi liječenja su bili veći za 32%, 95% i 109% slijedom, u usporedbi s troškovima liječenja prvog stadija bolesti (22).

Pitanje kada započeti s mamografijom i koliko često ju ponavljati je predmet intenzivnih rasprava i razlikuje se od države do države. U većini razvijenih država, uključujući Hrvatsku, provodi se dvogodišnji probir mamografijom kod žena u dobi između 50 i 69 godina. Neke države proširuju ovaj interval do 75. godine. U SAD-u postoje različite preporuke za početak i učestalost probira, pa tako USPSTF (US Preventive Services Task Force) preporučuje dvogodišnji probir za žene između 50 i 74 godina, dok ACR (American College of Radiology) zagovara strože preporuke koje uključuju godišnji probir već od 40. godine (18, 23). Prema najnovijim europskim smjernicama za probir i dijagnostiku raka dojke, postoje snažni dokazi koji zagovaraju dvogodišnji probir kod žena između 50. i 69. godine. Te smjernice ne preporučuju probir

za žene mlađe od 45 godina, dok za žene između 45 i 49 godina daju uvjetnu preporuku za dvogodišnji ili trogodišnji probir. Isto tako, daju uvjetnu preporuku za trogodišnji probir u dobi od 70. do 74. godine. Uvjetna preporuka bazira se na slabijim dokazima, stoga kliničar treba individualizirati pristup ovisno o pacijentici, uzimajući u obzir individualne faktore rizika, kao što su obiteljska povijest raka dojke i drugi relevantni zdravstveni faktori (24).

Kao što je već navedeno, mamografija ima ključnu ulogu u ranom otkrivanju raka dojke, ali, postoji i nekoliko rizika i ograničenja povezanih s njome. Mamografija može dovesti do lažno pozitivnih rezultata, što znači da se u nekim slučajevima može ukazivati na rizičnu leziju iako nije prisutna. Analiza ukazuje na to da kumulativni rizik lažno pozitivnih nalaza tijekom desetogodišnjeg perioda iznosi 6.4% (95% CI, 5.6%-7.2%) za probir koji se provodi svake dvije godine kod žena koje su započele s pregledima u 50. godini (19). Kod takvih nalaza potrebno je provesti dodatna testiranja poput drugih slikovnih pretraga ili biopsije, kako bi se isključila vjerojatnost maligne lezije, što dovodi do nepotrebne anksioznosti i stresa kod žena (25).

Također, postoji i rizik od prekomjernog otkrivanja (engl. *overdiagnosis*), što znači da se mogu otkriti vrlo sporo rastući tumori koji možda nikada ne bi izazvali probleme tijekom života žene. Ovo može dovesti do prekomjernog liječenja, uključujući kirurške zahvate, radioterapiju i kemoterapiju, koje sa sobom nose rizike i nuspojave (18, 25).

2.4. CAD sistemi i AI kod mamografije

Računalno potpomognuta dijagnostika (engl. *computer-aided diagnosis*, CAD) služi kao alat koji podržava radiologe u procesu dijagnosticiranja bolesti. CAD sustavi su dizajnirani da identificiraju karakteristike koje mogu ukazivati na bolest na medicinskim slikama. Ovi sustavi koriste se za sve veći broj primjena, uključujući mamografiju. U mamografiji, CAD sustavi koriste se za detekciju i karakterizaciju abnormalnosti, poput kalcifikacija i tumora. Sustavi analiziraju digitalne slike mamograma, tražeći anomalije koje se mogu lako previdjeti ljudskim okom. CAD sustavi pomažu radiolozima da izdvoje područja na mamogramima koja bi mogla zahtijevati dodatnu pažnju. FDA (Food and Drug Administration), američka uprava za hranu i lijekove, odobrila je CAD sustave za mamografiju 1998. godine (26).

Postoje dvije glavne vrste CAD sustava u mamografiji: CADe (CAD za detekciju) i CADx (CAD za dijagnozu). CADe sustavi služe kao pomoć u otkrivanju potencijalno abnormalnih područja na mamogramima. Najčešće se primjenjuju kao sekundarni pregledači, nakon što radiolog pregleda mamogram. Ako CAD sustav identificira moguće područje interesa, to se područje može označiti za daljnje ispitivanje. S druge strane, CADx sustavi se upotrebljavaju za karakterizaciju i ocjenu identificiranih područja interesa. Oni mogu pomoći u određivanju je li identificirana lezija vjerojatno benigna ili maligna, pružajući podršku u procesu odlučivanja o daljnjem liječenju (12, 27). Važno je napomenuti da su značajke lezija koje koriste ovi sustavi

ručno dizajnirane od strane stručnjaka. Na temelju tih značajki, algoritmi klasičnog strojnog učenja, poput metode potpornih vektora (engl. *Support Vector Machine* - SVM), primjenjuju se za klasifikaciju lezija kao benignih ili malignih (12, 28).

Iako CAD sustavi nude brojne prednosti, istraživanja su pokazala nejasne rezultate u pogledu njihove učinkovitosti u povećanju preciznosti dijagnoze raka dojke. Neki radovi sugeriraju da CAD sustavi mogu povećati broj lažno pozitivnih nalaza, što može dovesti do nepotrebnih biopsija i dodatnih troškova (27, 28). Studija iz 2015. godine (26) pokazuje da je osjetljivost mamografskog probira korištenjem CAD-a iznosila 85.3% (95% CI, 83.6%-86.9%), a bez korištenja CAD-a 87.3% (95% CI, 84.5%-89.7%). Značajna razlika nije pronađena ni u specifičnosti (91.6% i 91.4%). Kod radiologa koji povremeno koriste CAD, pronađena je smanjena osjetljivost kada su ga koristili (omjer izglednosti, 0.53; 95% CI, 0.29-0.97), što bi se moglo pripisati lažnom osjećaju sigurnosti kod neiskusnih radiologa (27).

Razvojem tehnologija poput dubokog učenja i umjetne inteligencije, CAD sustavi u mamografiji postaju sve precizniji i efikasniji. CAD sustavi koji koriste duboko učenje mogu naučiti prepoznati složene obrasce bez potrebe za ručnim dizajniranjem radiomičkih značajki. Augmentacija podataka i prijenosno učenje (engl. *transfer learning*) mogu pomoći pri nedostatku kvalitetnih snimaka, dok multimodalno učenje omogućuje sintezu kliničkih i radioloških nalaza za postavljanje dijagnoze. Evaluacija na javno dostupnim skupovima podataka pokazuje dobre rezultate (4, 12, 27, 28).

Retrospektivno istraživanje (29) koje je evaluiralo performanse i generalizaciju AI sustava za probir raka dojke, pokazalo je da takav sustav postiže bolje rezultate u odnosu na jednog radiologa. Osjetljivost i specifičnost bile su jednake ako se AI koristi kao drugi pregledač umjesto radiologa. Također se pokazalo da sustav ima sposobnost međunarodne generalizacije, tako da je za treniranje modela korišten skup mamografskih snimaka iz Ujedinjenog Kraljevstva, dok je evaluacija provedena na skupu podataka iz SAD-a. Jedno je istraživanje pokazalo da nema statistički značajne razlike između performansi AI-CAD sustava i prosječnih performansi 101 radiologa. Naprotiv, AUC rezultat koji je AI sustav postigao bio je bolji od rezultata polovice radiologa (30). AI sustavi mogli bi u budućnosti čak i zamijeniti radiologe u trijaži jednostavnih negativnih slučajeva (31). Iako su rezultati obećavajući, potrebno je provesti dodatna istraživanja koja obuhvaćaju kliničku primjenu razvijenih modela.

3. Duboko učenje za računalni vid

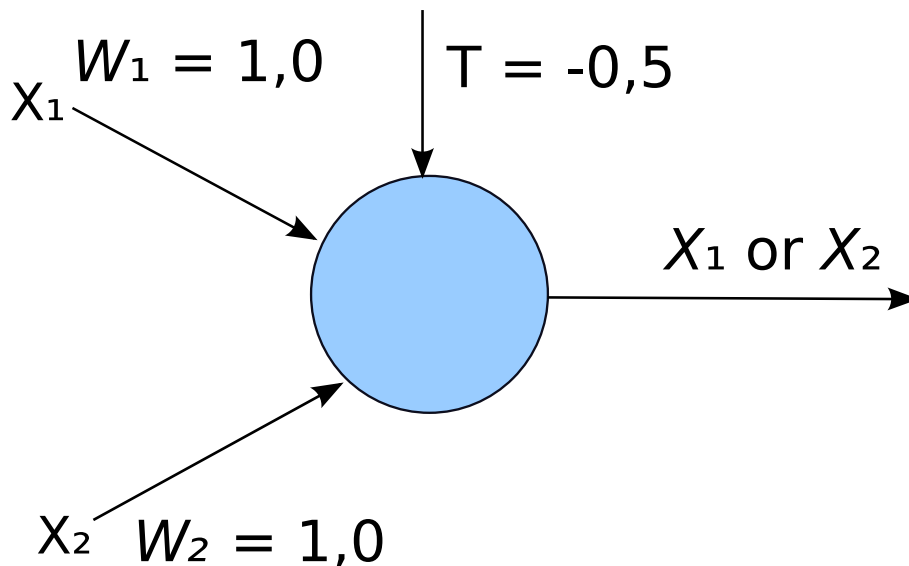
3.1. Od ideje to primjene

Povijesno gledajući, umjetna inteligencija (engl. *artificial intelligence*, AI) doživjela je nekoliko razdoblja uspona i padova. Postoje razdoblja u kojima je došlo do naglog razvoja ove tehnologije, no zbog raznih okolnosti, primjena u svakodnevnom životu nikada nije ostvarena. Tek je nedavno započela AI revolucija, koja se temelji na razvoju hardverskih performansi računala, dostupnosti velike količine podataka te napretku u metodama dubokog učenja (4, 32).

Već nakon izrade prvih računala, primjetni su pokušaji imitacije ljudskog mozga i razmišljanja pomoću logičkih sklopova. McCulloch i Pitts (33) navode da se ponašanje biološkog neurona, koji radi na principu "sve ili ništa", može modelirati pomoću umjetnog neurona koji koristi logičke operacije poput "i", "ili", "ne" iz Booleove algebre. Izlaz umjetnog neurona ovisi o zbroju svih primljenih ulaza te o pragu podražaja neurona, koji je definiran inhibicijskim signalima (Slika 3). Ovaj rad predstavlja ključni doprinos u razvoju umjetne inteligencije jer je dao ideju o modeliranju bioloških neuronskih mreža korištenjem matematičkih i logičkih modela koji omogućavaju rješavanje kompleksnih problema. Međutim, ograničenja ovog pristupa su da umjetni neuron nema sposobnost "učenja" te da može generirati samo dvije vrijednosti, nula ili jedan, zbog principa "sve ili ništa" (34).

Napredak u razvoju umjetnih neuronskih mreža dogodio se s Rosenblattovim radom iz 1958. godine (36), u kojem je predstavio perceptron - umjetnu neuronsku mrežu jednog sloja koja je bila sposobna generirati kontinuirane izlazne vrijednosti između nula i jedan. Izlaz modela temeljio se na linearnoj kombinaciji njegovih ulaza i usporedbi dobivenih vrijednosti s pragom. Primjenjujući ideju o plastičnosti veza između neurona, Rosenblatt je razvio algoritam učenja perceptrona koji mu je omogućio prilagodbu i poboljšanje izlaznih rezultata na temelju ulaznih podataka i očekivanih izlaza. Ovaj rad postavio je temelje dubokog učenja, jer je perceptron bio sposoban prepoznati uzorke u podacima i naučiti klasificirati obrasce. Kritike zbog njegove primjenjivosti samo na linearne klasifikacijske zadatke dovele su do smanjenja popularnosti i stagnacije u razvoju neuronskih mreža, iako se pokazalo da se dodavanjem samo jednog sloja može riješiti spomenuti problem (37).

Fokus na neuronske mreže vratio se krajem osamdesetih godina, s posebnim zanimanjem



Slika 3. Shematski prikaz McCullochova umjetnog neurona. (35)

usmjerenim na višeslojne mreže. Međutim, problem je bio nedostatak efikasnih algoritama za učenje dubokih neuronskih mreža (DNN) s više skrivenih slojeva. Rumelhart, Hinton i Williams (38) su pokazali da se algoritam propagacije greške prema unatrag (engl. *backpropagation*), može koristiti za efikasno učenje dubokih mreža. Ideja je bila da se prolaskom kroz mrežu unaprijed (engl. *forward pass*) računaju izlazi modela i funkcija gubitka između izlaza i stvarnih vrijednosti, dok se u prolasku unatrag računaju gradijenti funkcije gubitka u odnosu na parametre modela. Gradijenti se dobivaju primjenom pravila lanca na kompozitne funkcije. Tako se neuronska mreža uči prilagoditi ulaznim podacima i generirati izlaz koji odgovara očekivanom rezultatu. Ovaj algoritam je temelj učenja neuronskih mreža i koristi se i danas, uz nekoliko poboljšanja.

Revolucionarna primjena neuronskih mreža za prepoznavanje slika postignuta je modelom LeNet-5, kojeg je razvio LeCun 1989. godine. Ovaj model predstavljao prvu uspješnu primjenu konvolucijskih neuronskih mreža (CNN) za prepoznavanje rukom pisanih poštanskih brojeva (39). LeNet-5 se sastojao od nekoliko slojeva, uključujući slojeve konvolucije i agregacije, koji su trenirani algoritmom propagacije greške unatrag. Konvolucijski slojevi imaju zadatak detektiranja značajki u slikama, dok agregacijski slojevi reduciraju dimenzionalnost izlaza konvolucijskog sloja. Kombiniranjem tih slojeva, model je mogao sam naučiti razne značajke koje su bile korisne za prepoznavanje brojeva, bez ručnog dizajniranja.

Tijekom mnogih godina, CNN-ovi su predstavljali vodeću tehnologiju (engl. *state-of-the-art*, SOTA) u različitim zadacima koji zahtijevaju klasifikaciju slika ili detekciju objekata, kao što su prepoznavanje lica, autonomna vožnja vozila ili detekcija grešaka na proizvodima u tvornici (5). Međutim, nedavno je otkriveno da nova vrsta arhitekture nazvana Vision Transformer (ViT) (40) postiže bolje rezultate u većini zadataka gdje su CNN-ovi dominirali, poput pre-

poznavanja slika i semantičke segmentacije (41). Nova generacija modela baziranih na ViT-u donosi značajne napretke u području vizualne obrade i analize slika.

3.2. Principi dubokog učenja

Temeljni dijelovi svakog algoritma strojnog učenja su model, funkcija gubitka i algoritam optimizacije (42). U nadziranom dubokom učenju, model je skup funkcija f kojima želimo što bolje aproksimirati funkciju f^* , koja opisuje prave odnose između nezavisnih varijabli x_i i zavisne varijable y_i . Tijekom učenja, model pokušava identificirati veze između varijabli u skupu podataka, koji se sastoji od n uzoraka ili jedinki, označenih kao $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$. Na temelju tih uzoraka, model formira hipotezu $y_i = f(x_i, \theta)$, pri čemu θ označava parametre tog modela. Algoritam optimizacije zatim prilagođava parametre modela kako bi minimizirao gubitak, koji kvantificira odstupanja izlaza modela od pravih vrijednosti.

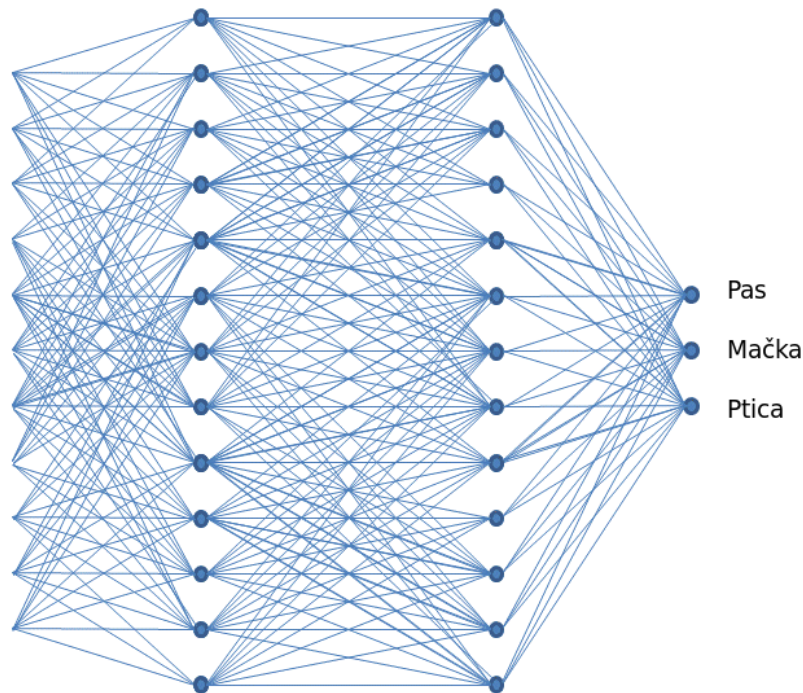
3.2.1. Model

S obzirom na protok informacija, modele dubokog učenja možemo podijeliti na unaprijedne i povratne modele (4). Povratni modeli (engl. *Recurrent Neural Networks* - RNNs) koriste se za obradu i analizu sekvenci podataka koji se mijenjaju tijekom vremena. U ovom modelu, svaki neuron, osim ulaznih podataka $x^{(t)}$, prima i skriveno stanje h^{t-1} , koje se sastoji od informacija o prethodnim ulaznim podacima (43). Radi toga su RNN-i dobri pri zadacima gdje je potrebno zadržati kontekst, kao što su govorne poruke, tekstovi ili senzorski podaci (5). Postoji nekoliko vrsta povratnih modela poput dvosmjernih RNN, LSTM (engl. *Long short-term memory*) i GRU (engl. *Gated recurrent unit*).

Unaprijedni modeli nemaju povratni protok informacija, već svaki sloj prima aktivacije neurona iz prethodnog sloja. Najjednostavniji primjer unaprijednih modela su unaprijedni potpuno povezani modeli, gdje je svaki neuron povezan sa svim neuronima iz prethodnog sloja. Potpuno povezani modeli se još nazivaju i višeslojnim perceptronima (MLP).

Glavne komponente MLP-a su ulazni sloj, skriveni slojevi i izlazni sloj. Ulazni sloj sadrži broj neurona jednak dimenzionalnosti ulaznog prostora. Broj neurona u izlaznom sloju odgovara broju klasa u višeklasnoj klasifikaciji, dok je za regresiju i binarnu klasifikaciju obično jednak jedan. Broj neurona u skrivenim slojevima je proizvoljan, ali može značajno utjecati na performanse modela. Skriveni slojevi koji su bliži ulaznom sloju nazivaju se nižim slojevima, dok se slojevi koji su bliži izlazu nazivaju višim (Slika 4). Postoji općenito prihvaćen pogled da se mreže smatraju dubokima ako imaju više od dva skrivena sloja, iako ne postoji precizna definicija (32).

Izlaz neurona u MLP-u dobiva se linearnom kombinacijom ulaznih značajki, što se može



Slika 4. Shematski prikaz unaprijedne potpuno povezane neuronske mreže. Mreža sadrži dva skrivena sloja i tri izlazna neurona za klasifikaciju. (44)

prikazati sljedećim izrazom:

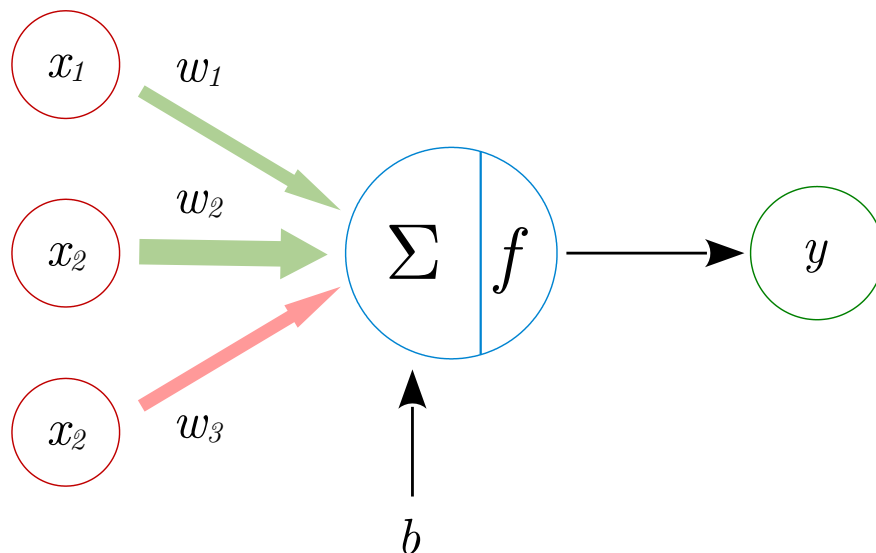
$$y = \sum_{i=1}^n w_i x_i + b.$$

U ovom izrazu, x predstavlja vektor ulaznih značajki (x_1, x_2, \dots, x_n) za jedan primjer iz skupa podataka ako se radi o prvom skrivenom sloju, dok konstanta b označava pomak (engl. *bias*) neurona. Za dublje slojeve, ulazni vektor se sastoji od aktivacija neurona iz prethodnog sloja (a_1, a_2, \dots, a_n) . Dimenzije vektora težina (w_1, w_2, \dots, w_n) moraju odgovarati dimenzijama vektora ulaznih značajki. Izlaz neurona zatim prolazi kroz nelinearnu aktivacijsku funkciju, koja omogućava modeliranje nelinearnih odnosa između značajki (Slika 5). Bez aktivacijske funkcije, cjelokupni model neuronske mreže bio bi ograničen na linearno modeliranje (4). S obzirom na veliki broj neurona u jednom sloju, korisno je koristiti matematičke operacije s matricama. Konačni matematički zapis transformacije u jednom sloju MLP-a pomoću matrica je sljedeći (45):

$$Z^{[l]} = W^{[l]} A^{[l-1]} + b^{[l]}$$

$$A^{[l]} = g(Z^{[l]}),$$

gdje je $W^{[l]}$ matrica težina, $A^{[l-1]}$ matrica aktivacija izlaza prethodnog sloja, $b^{[l]}$ vektor pomaka, a g predstavlja aktivacijsku funkciju.



Slika 5. Shematski prikaz umjetnog neurona u MLP-u. Izlaz neurona čini linearna kombinacija težina na koju je pridodana vrijednost pomaka. Nelinearnost izlaza se postiže primjenom aktivacijske funkcije. (46)

U skrivenim slojevima neuronskih mreža, vrlo često korištena aktivacijska funkcija je zglobnica (engl. *rectified linear unit*, ReLU) (47). Postoje i varijacije ReLU-a koje definiraju gradijente i za vrijednost $x < 0$, kao što su Leaky ReLU, GeLU, ELU (48). ReLU se pokazao kao odličan izbor za skrivene slojeve dubokih neuronskih mreža zbog svoje matematičke jednostavnosti i sposobnosti razrjeđivanja mreže, postavljanjem izlaza nekih neurona na nulu (49). Neuron koji koristi ReLU kao aktivacijsku funkciju je aktivan kada je linearna kombinacija ulaza veća od nule:

$$ReLU = \max(0, x).$$

Takav neuron proizvodi izlaz u rasponu $[0, \text{inf}]$.

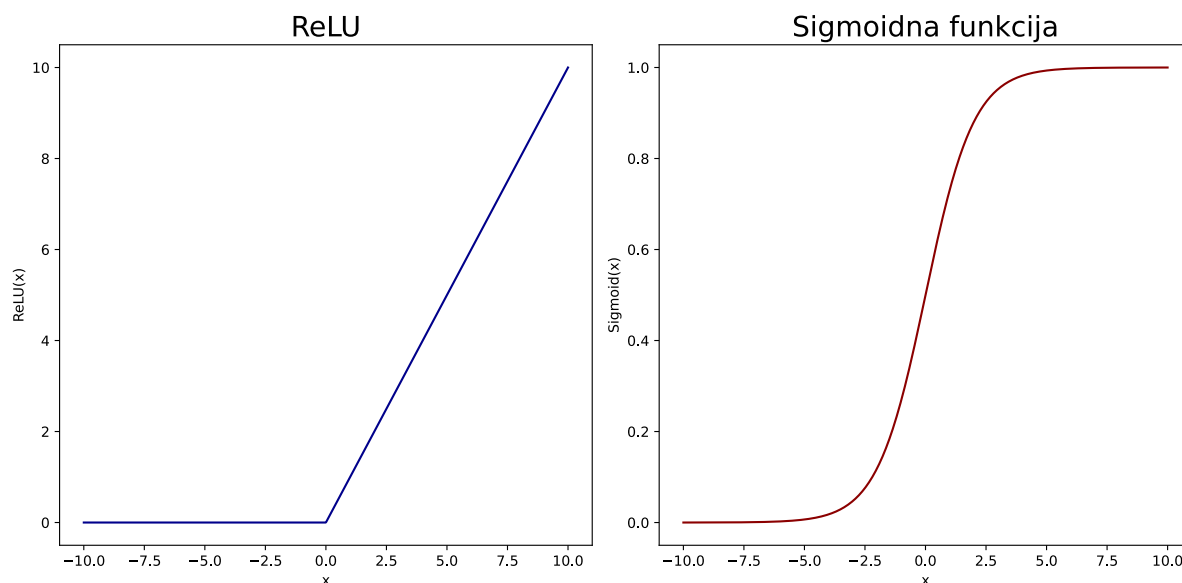
Na izlaznom sloju mreže, odabir aktivacijske funkcije ovisi o vrsti zadatka (47). Za binarne klasifikacijske probleme često se koristi sigmoidna funkcija:

$$\sigma(z) = \frac{1}{1 + e^{-z}}.$$

Izlaz te funkcije je realni broj u rasponu $[0, 1]$ što se tumači kao vjerojatnost da uzorak pripada pozitivnoj klasi (Slika 6). Kada se sigmoidna funkcija proširi na višeklasne probleme, dobiva se *softmax* funkcija (4). Izlaz te funkcije su vjerojatnosti za svaku klasu, pri čemu je zbroj tih vjerojatnosti jednak jedan. Klasa s najvećom vjerojatnošću se uzima kao predikcija mreže.

3.2.2. Optimizacija

Na početku procesa dubokog učenja, model s nasumično inicijaliziranim parametrima generira prediktivne vrijednosti, \hat{y}_i , temeljene na dostavljenim ulaznim podacima. Ove predviđene



Slika 6. ReLU (lijevo) i sigmoidna (desno) aktivacijska funkcija (Autorski rad)

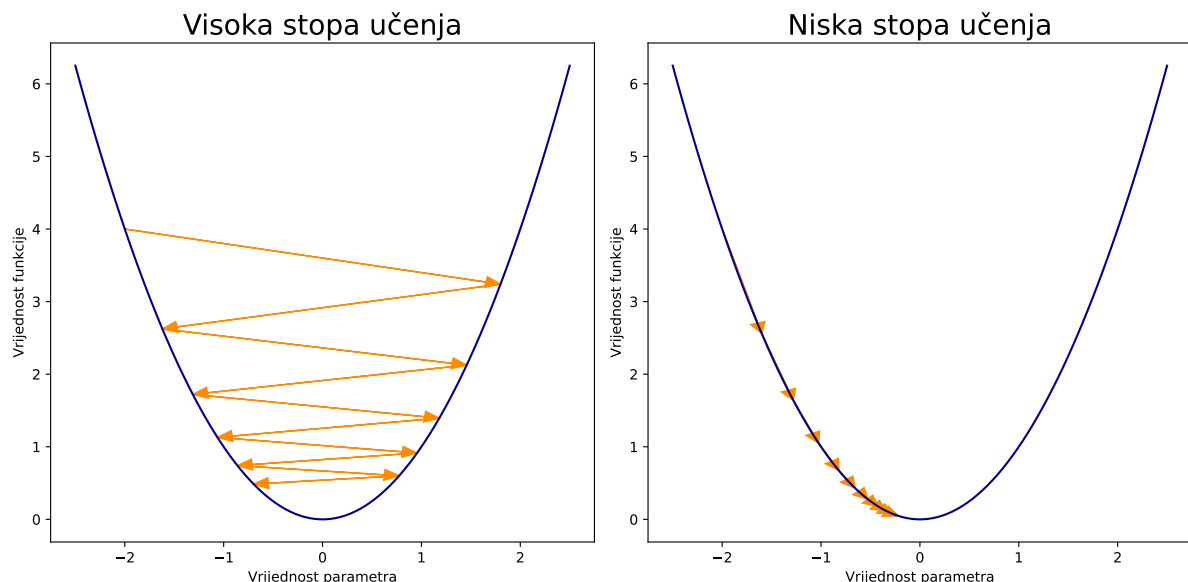
vrijednosti mogu znatno odstupati od stvarnih ciljanih vrijednosti, y_i , koje dolaze iz izvorne distribucije $p(x_i, y_i)$. Izvorna distribucija predstavlja teoretsku raspodjelu svih mogućih ulazno-izlaznih parova za određeni problem. No, u stvarnosti, ova distribucija je nepoznata i nepristupačna. Umjesto toga, model radi s ograničenim skupom podataka, poznatim kao empirijska distribucija, $\hat{p}(x_i, y_i)$. Ova empirijska distribucija predstavlja stvarno dostupne podatke iz trenutnog skupa za treniranje. Optimiziranjem modela prema empirijskoj distribuciji, nadamo se da ćemo smanjiti grešku i na izvornoj distribuciji. Ova pretpostavka počiva na načelu induktivnog zaključivanja, ali nije uvijek ispunjena, osobito u slučajevima kada skup podataka za treniranje nije reprezentativan za izvornu distribuciju (50).

Cilj u praksi nije nužno minimizacija empirijske greške, već minimizacija generalizacijske greške, što se postiže posredno preko funkcije gubitka. Generalizacijska greška odnosi se na razliku između performansi modela na skupu za treniranje i njegovih performansi na novim, neviđenim podacima, tj. ona mjeri sposobnost modela da se generalizira na nepoznate podatke. Kao mjeru generalizacijske greške često se koristi poseban skup podataka, testni set, kojem model nema pristup tijekom treniranja (4). Funkcija gubitka služi za kvantificiranje koliko predikcija modela odstupa od stvarne ciljane vrijednosti, pružajući numeričku mjeru ove pogreške koju možemo optimizirati. Ovaj proces optimizacije koristi derivacije funkcije gubitka u algoritmu gradijentnog spusta, čija je svrha pronalazak globalnog (ili lokalnog) minimuma funkcije, čime se indirektno smanjuje i generalizacijska greška (4, 51).

Pri odabiru funkcije gubitka, važno je odabrati derivabilnu funkciju, jer nam to omogućuje izračunavanje smjera i veličine promjene pogreške u odnosu na parametre modela. Izbor funkcije gubitka ovisi o vrsti zadatka i vrsti podataka. Na primjer, za regresijske probleme često se koristi srednja kvadratna pogreška (engl. *mean squared error*), dok se za klasifikacijske

probleme koristi kategorička pogreška unakrsne entropije (engl. *categorical cross-entropy*) ili binarna pogreška unakrsne entropije (engl. *binary cross-entropy*), ovisno o broju klasa u problemu (32, 37, 47).

U optimizaciji parametara modela, ključnu ulogu igra algoritam gradijentnog spusta koji iterativno mijenja parametre modela tako da se gubitak smanjuje. Parametri modela se ažuriraju u smjeru negativnog gradijenta funkcije gubitka. Veličina ažuriranja ovisi o stopi učenja, koja određuje koliko velik korak model napravi prilikom optimizacije parametara (4, 51). To je važno jer utječe na brzinu prilagodbe parametara na temelju gradijenta funkcije gubitka. Visoka stopa učenja može uzrokovati skakanje preko globalnog minimuma ili divergencije, dok niska stopa može dovesti do spore konvergencije ili zaglavljivanja u lokalnom minimumu (Slika 7). Kako bi se umanjili ovi problemi, razvijene su varijante gradijentnog spusta, kao što su stohastički gradijentni spust (SGD), RMSprop, Adagrad, Adam, koje koriste različite tehnike, poput momentuma, adaptivne stope učenja, i druge, kako bi se poboljšao proces optimizacije (51).



Slika 7. Primjer visoke stope učenja (lijevo) i niske stope učenja (desno). Visoka može uzrokovati oscilacije gradijentnog spusta bez postizanja minimuma, dok niska može dovesti do spore konvergencije i potencijalno preranog završetka treninga prije postizanja minimuma (Autorski rad)

Ovisno o veličini skupa podataka koji se koristi u jednoj iteraciji ažuriranja parametara, razlikujemo stohastički, mini-grupni (engl. *mini-batch*) i grupni (engl. *batch*) gradijentni spust. Stohastički i grupni gradijentni spust su dvije krajnosti algoritma, gdje se kod prvog koristi samo jedan primjer iz skupa podataka, dok se kod grupnog gradijentnog spusta koristi cijeli skup. Radi toga, stohastički dovodi do većih fluktuacija gubitka, no bržeg ažuriranja parametara, a grupni je puni sporiji i računalno zahtjevniji, no stabilniji. Ipak, danas je najčešće korišten mini-grupni gradijentni spust koji se nalazi na zlatnoj sredini što se tiče brzine i stabilnosti. U jednoj iteraciji mini-grupnog gradijentnog spusta koristi se određeni broj primjera, n , za optimiziranje

parametara θ (51):

$$\theta = \theta - \eta \cdot \Delta_{\theta} J(\theta; x^{(i:i+n)}; y^{(i:i+n)}),$$

pri čemu je η stopa učenja, a $\Delta_{\theta} J$ gradijent gubitka u odnosu na parametre modela. Pojmom epoha opisujemo jedan prolaz gradijentnog spusta kroz cijeli skup podataka za ažuriranje svakog parametra. Ukupni broj iteracija, odnosno ažuriranja parametara, stoga ovisi o broju epoha te o veličini grupe (4, 32, 47).

3.2.3. Prenaučenost i podnaučenost

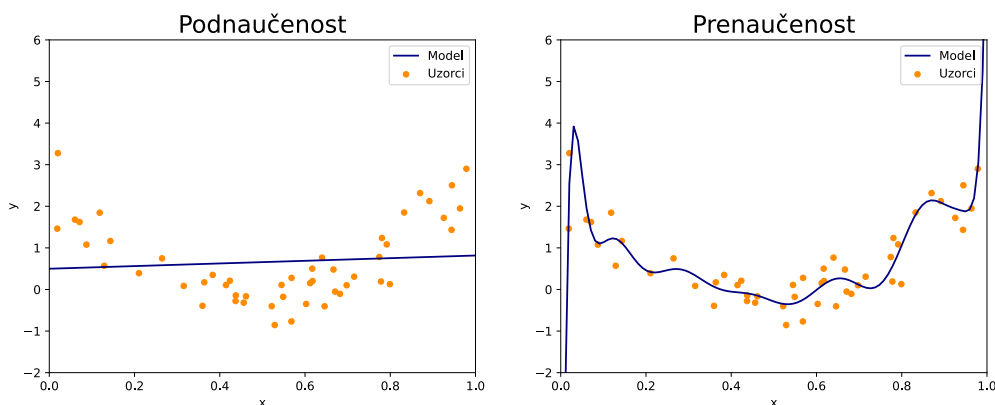
Prenaučenost i podnaučenost su dva ključna problema koja se javljaju tijekom treniranja modela dubokog učenja. Prenaučenost se javlja kada model previše dobro nauči značajke trening skupa podataka, do te mjere da postaje loš u generalizaciji na novim, neviđenim podacima. S druge strane, podnaučenost se javlja kada model nije dovoljno dobro naučio značajke trening skupa podataka, što rezultira lošim performansama na trening skupu, ali i na test skupu podataka (4, 47).

Koncept koji opisuje izazov povezan s problemima podnaučenosti i prenaučnosti naziva se pristranost-varijanca kompromis (engl. *bias-variance tradeoff*). Model visoke varijance izrazito je osjetljiv na fluktuacije u trening setu, do mjere da bi identificirao različit minimum ako bi se uzorkovanje podataka za treniranje promijenilo. Ovakav model pati od nedostatka generalizacijske sposobnosti. Povećanje složenosti modela do određenog stupnja rezultira smanjenjem greške kako na trening setu tako i na testnom setu. Međutim, daljnje povećanje složenosti dovodi do povećanja generalizacijske greške, budući da model postaje pretjerano prilagođen trening podacima (Slika 8). Stoga, cilj je identificirati model koji je dovoljno složen da učinkovito uči iz podataka, ali ne toliko složen da riskira prenaučnost (4, 47).

Jedan od načina za borbu protiv prenaučnosti je korištenje regularizacijskih tehnika. Regularizacija je metoda koja dodaje dodatni termin u funkciju gubitka kako bi se ograničila složenost modela i spriječila prenaučnost. L1 i L2 regularizacija su dvije najčešće korištene metode regularizacije u strojnom učenju. L1 regularizacija dodaje apsolutnu vrijednost težina u funkciju gubitka, što potiče model da ima manje težine, dok L2 regularizacija dodaje kvadrat težina u funkciju gubitka, što potiče model da ima manje i ravnomjernije raspodijeljene težine (4, 37, 47).

Dropout je popularna tehnika za borbu protiv prenaučnosti u dubokom učenju. Tijekom treniranja, dropout nasumično "isključuje" određeni postotak neurona u sloju, što znači da se njihovi izlazi postavljaju na nulu. Time se model ne oslanja previše na bilo koji pojedini neuron, što pomaže u sprečavanju prenaučnosti (4, 37).

Rano zaustavljanje (engl. *early stopping*) je još jedna tehnika koja se često koristi za smanjenje prenaučnosti. Ova tehnika uključuje praćenje performansi modela na validacijskom



Slika 8. Podnaučenost (lijevo) i prenaučnost (desno). Korištenje prejednostavnog modela za složen zadatak (u ovom primjeru, linearni model za aproksimiranje kvadratne funkcije) rezultira podnaučenošću. S druge strane, upotreba pretjerano složenog modela (poput polinoma višeg stupnja) vodi prema prenaučnosti i gubitku sposobnosti generalizacije. (Autorski rad)

skupu podataka tijekom treniranja. Kada se performanse na validacijskom skupu prestanu poboljšavati, treniranje se zaustavlja. Ovo pomaže modelu jer sprječava da previše dugo trenira na trening skupu podataka (4, 37, 47).

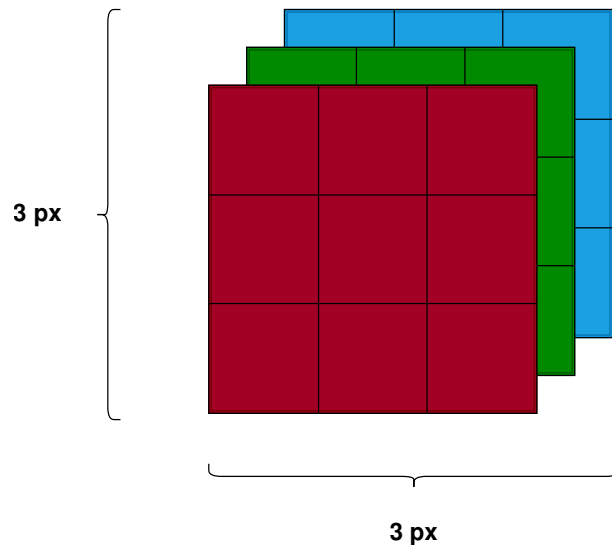
S druge strane, podnaučenost se može riješiti povećanjem složenosti modela, dodavanjem više podataka ili korištenjem tehnika augmentacije podataka. Augmentacija podataka uključuje stvaranje novih trening primjera putem malih modifikacija originalnih podataka, što kod slika može uključivati rotacije, translacije, zumiranje i druge tehnike. To omogućava modelu da nauči više značajki i poboljša svoju sposobnost generalizacije (4, 37, 47, 52).

3.3. Digitalni prikaz slika

Digitalne slike su diskretne reprezentacije analognih vizualnih informacija. Svaka slika se sastoji od osnovnih elemenata koje nazivamo pikselima. Kod skalarnih slika, svaki piksel je određen realnim brojem, čija vrijednost ovisi o dubini bita slike. Takve se slike često prikazuju u sivim tonovima, kao što je slučaj s rendgenskim snimkama. Često korišteni broj bitova za takve slike su 8 i 16. Kod 8-bitnih slika svaki piksel može imati maksimalnu razinu intenziteta od 255, a najmanju od 0. Generalno, formula koja povezuje dubinu bita slike i razine intenziteta piksela je: 2^k , gdje je k dubina bita slike (53). Viša dubina bita omogućuje veću preciznost u prikazu sivih tonova, ali također povećava veličinu datoteke slike. Današnji standard za slike u nijansama sive je 16 bita. (54).

Slike u boji se obično prikazuju koristeći RGB (Red-Green-Blue) model, koji je aditivni model boja. Takve slike su predstavljene vektorom piksela, gdje svaki piksel sadrži tri vrijednosti - jednu za crvenu, zelenu i plavu komponentu. Cijela slika je stoga trodimenzionalna matrica,

gdje treća dimenzija ovisi o broju kanala, odnosno komponentama kojima se prikazuje boja (Slika 9). Dubina bita u kontekstu RGB slika predstavlja broj mogućih nijansi crvene, zelene i plave, što za 8-bitne slike rezultira s više od 16 milijuna boja. (53, 54).



Slika 9. Prikaz teoretske slike u boji veličine 3x3 piksela (Autorski rad)

Rezolucija slike odnosi se na količinu detalja koje slika može prikazati. Prostorna rezolucija se mjeri u pikselima po inču (PPI) ili pikselima po centimetru. Dubina bita se smatra intenzitetnom rezolucijom slike. Viša rezolucija omogućuje prikaz finijih detalja, ali povećava veličinu datoteke i potrebnu računalnu snagu za obradu slike. Promjena rezolucije se često koristi za smanjenje veličine datoteke ili za prilagodbu slike određenom formatu prikaza. Međutim, promjena rezolucije može rezultirati gubitkom informacija ili stvaranjem artefakata (53).

Prilikom mijenjanja rezolucije koriste se interpolacijske metode koje izračunavaju vrijednosti novih piksela na temelju originalnih piksela. Najjednostavnija metoda, najbližeg susjeda, dodjeljuje vrijednost najbližeg originalnog piksela novom pikselu. Bilinearna interpolacija izračunava vrijednost novog piksela kao težinsku prosječnu vrijednost četiri najbliža piksela. Bikubična interpolacija koristi 16 najbližih piksela da izračuna novu vrijednost, što rezultira glađom slikom (53). Odabir interpolacijske metode prilikom smanjivanja rezolucije slike za potrebe dubokog učenja, može utjecati na performanse modela (55).

Kada se slike koriste za treniranje modela dubokog učenja, uobičajeno je koristiti tehniku poznatu kao augmentacija podataka. Augmentacija podataka obuhvaća niz tehnika koje stvaraju promjene u originalnim slikama, uključujući rotacije, translacije, zumiranje, izrezivanje, preokretanje i promjenu boje slika (Slika 10). Augmentacija se koristi u dubokom učenju kako bi se modeli bolje generalizirali na nove podatke, pružajući im više varijacija na trening podacima (52).

U medicinskim primjenama, augmentacija podataka može biti posebno korisna za povećanje veličine skupa podataka. Medicinski skupovi podataka slika često su manji u usporedbi s



Slika 10. Prikaz često korištenih augmentacija slika (Autorski rad)

konvencionalnim skupovima podataka u drugim poljima, primarno zbog poštivanja principa zaštite privatnosti pacijenata i stroge regulative vezane za osobne zdravstvene podatke (52). Isto tako, zbog inherentne prirode medicinskih podataka, često se susrećemo s neuravnoteženošću u distribuciji klasa. To jest, patološke slike su značajno rjeđe od slika koje prikazuju normalne zdravstvene uvjete. Augmentacija pomaže modelima da se nose s različitim varijacijama patoloških slika (56).

4. Izrada modela za segmentaciju lezija dojki

4.1. Cilj

Cilj ovog rada jest razviti model sposoban za detekciju i klasifikaciju benignih i malignih lezija dojke. Za izradu ovog modela korišten je CBIS-DDSM skup podataka, koji uključuje mamografske snimke benignih i malignih kalcifikacija te tumora dojke. Pored samih snimaka, u skupu podataka nalaze se i takozvane maske, binarne slike na kojima svjetlija područja označavaju lezije. Stoga, za svaku leziju unutar ovog skupa, određeno je li benigna ili maligna, uz identifikaciju točnih piksela koji pripadaju toj leziji.

Razvijen model ima mogućnost klasifikacije svakog pojedinačnog piksela u jednu od tri kategorije: benigni, maligni ili pozadinski piksel. Ovaj pristup klasifikaciji, gdje se svaki piksel pojedinačno klasificira umjesto da se klasificira cijela slika, poznat je kao semantička segmentacija. Ova metoda omogućuje lokalizaciju detektiranih lezija, što bi moglo pomoći radiologu ubrzati dijagnostiku mamografskih snimaka.

4.2. Korištene tehnologije

U sklopu ovog rada, za procesiranje slika i konstrukciju modela, primijenjen je programski jezik Python, specifično njegova 3.10 verzija. Python je stekao reputaciju najpopularnijeg izbora u polju dubokog učenja, zahvaljujući svojoj jednostavnosti korištenja i širokom spektru biblioteka koje implementiraju različite alate neophodne za strojno učenje. Biblioteke koje su se koristile u ovom istraživanju obuhvaćaju PyTorch, Hugging Face Transformers, Scikit-Learn, OpenCV.

Izrada i implementacija koda odvijala se putem razvojne platforme Microsoft Visual Studio Code te Jupyter bilježnica, gdje je preferirani stil programiranja bio objektno orijentirani. Najveći dio modela obučeni je koristeći Bioimage ANalysis Desktop (BAND) te Kaggle platforme, koje nude besplatno korištenje grafičkih kartica za duboko učenje.

4.3. Model

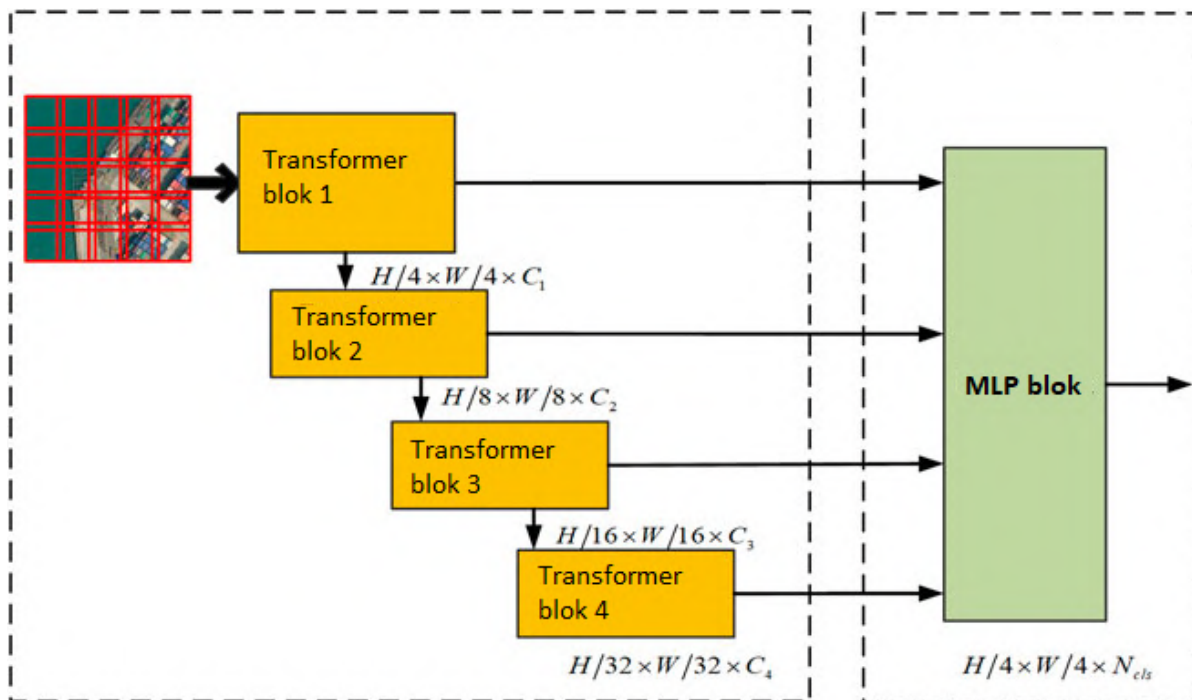
Prije razvoja metode semantičke segmentacije, proces detekcije i klasifikacije obavljao se u dva odvojena koraka. Razvoj dubokog učenja i hardvera za treniranje modela doveli su do stvaranja prvih modela koji ovaj proces mogu obaviti u jednom koraku. Jedan od prvih takvih modela razvijen je 2015. godine (57), a njegovom modifikacijom dobiven je U-Net, koji je primjenjen za biomedicinsku segmentaciju te je postigao obećavajuće rezultate. Posebnost ovog modela je njegova simetrična arhitektura koja nalikuje slovu U. Model se sastoji od enkoder dijela, koji sažima slika te ekstrahira značajke iz njih. Nakon toga slijedi dekodeo dio koji obavlja povećanje rezolucije kako bi se obnovila originalna veličina slike i dobila preciznija segmentacija. U-Net također koristi preskočne (engl. *skip*) veze, koje povezuju odgovarajuće slojeve između enkodera i dekodera. Te veze omogućuju prijenos informacija iz nižih razina slike, čime se očuvaju fini detalji, istovremeno kombinirajući kontekst iz viših razina (58).

Za moj zadatak odabran je SegFormer model, novi model nedavno predstavljen u domeni segmentacije slike. Njegov je enkoder zasnovan na Transformer arhitekturi, za koju je zanimljivo napomenuti da se prvobitno počela koristiti u polju obrade prirodnog jezika (engl. *natural language processing*, NLP). Najistaknutiji modeli koji koriste Transformer arhitekturu u NLP-u su Generative Pretrained Transformer (GPT) modeli. Ovi modeli su postali temelj za razvoj popularnog chatbota poznatog kao ChatGPT, koji ima mogućnost generiranja uvjerljivih i koherentnih odgovora (32). Primjena u domeni klasifikacija slika zaživjela je ViT modelom (40).

Transformer arhitektura sastoji se od blokova koji se temelje na mehanizmu pozornosti (engl. *attention*) i omogućuju modelu da modelira globalne odnose i kontekst između elemenata slike na različitim razinama rezolucije. Mehanizam pozornosti dopušta modelu da procijeni važnost svakog segmenta slike na temelju njegovih interakcija s ostalim segmentima (40, 43). Dodatno, mehanizam pozornosti pomaže vizualizaciju dijelova slike na koje se model koncentrira prilikom donošenja predikcija, poboljšavajući interpretabilnost neuronskih mreža, koje su se tradicionalno smatrale "crnim kutijama" (32).

U SegFormer arhitekturi, dekodeo se sastoji od jednostavnog MLP-a, što doprinosi brzini mreže. Konačna veličina segmentacijska maske koju model producira iznosi $\frac{H}{4} \times \frac{W}{4} \times N_{cls}$, gdje su H i W originalna visina i širina segmentacijske maske, a N_{cls} broj kategorija (Slika 11). Zbog toga je bilo potrebno provesti dodatnu operaciju povećanja rezolucije kako bi se dobila maska originalne veličine. Također, prije izračunavanja funkcije gubitka, provedena je *softmax* operaciju za izračunavanje vjerojatnosti za svaku klasu.

Autori SegFormer trenirali su model na poznatim skupovima slika za segmentaciju poput ADE20K, koji sadrži preko 20,000 slika koje pokrivaju širok spektar scena iz svakodnevnog života. Iako se navedene slike značajno razlikuju od mamografskih snimaka, pokazalo se da pret hodno treniranje modela na srodnim zadacima može poboljšati performanse. Naime, neuron-



Slika 11. Pojednostavljeni prikaz SegFormer arhitekture. Enkoderski dio modela se sastoji od niza Transformer blokova, dok enkoderski dio čini jednostavni MLP blok. Potrebno je povećati rezoluciju završnih segmentacijskih maski kako bi se dobila maska originalne veličine (59)

ske mreže tijekom procesa treniranja razvijaju vlastite značajke iz slika na kojima se provodi treniranje. Analizom tih značajki utvrđeno je da modeli na taj način usvajaju prepoznavanje različitih vizualnih elemenata, od jednostavnih oblika poput ravne linije različitih orijentacija i boja, preko složenijih struktura poput krugova, pa sve do vrlo kompleksnih entiteta poput očiju, usta i udova (60). Jednostavnije značajke, poput linija, detektiraju se u početnim slojevima modela, dok se složenije značajke, koje se sastoje od kombinacija tih jednostavnijih, predstavljaju u dubljim slojevima. Treniranjem na drugim podacima, tzv. fino podešavanje (engl. *fine-tuning*), model može iskoristiti prethodno naučene značajke nižeg reda i primijeniti ih na drugačiji zadatak od onoga na kojem je prvotno treniran. Ovaj pristup, poznat kao prijenosno učenje, omogućava bržu konvergenciju i bolju generalizaciju modela na novim zadacima, posebno kada je dostupan ograničen broj primjera za trening, kao što je slučaj za biomedicinske slike (37, 61).

4.4. Skup podataka

CBIS-DDSM (Curated Breast Imaging Subset of the Digital Database for Screening Mammography) (62) je skup podataka koji je razvijen za potrebe znanstvenih istraživanja u području dijagnostike dojke. Dostupan je na repozitoriju "The Cancer Imaging Archive" (63). Ovaj skup podataka je nastao kao selektivna podskupina Digitalne baze za probirnu mamografiju (Digital

Database for Screening Mammography - DDSM). CBIS-DDSM je dobio svoje ime zbog svoje kurirane prirode, što znači da su slike pažljivo odabrane i pregledane kako bi se osigurala njihova kvaliteta i relevantnost za istraživanje. Skup podataka se sastoji od 2620 digitaliziranih snimki film mamografije, koje su snimljene na 1566 pacijenata. Ove slike obuhvaćaju patološki verifikirana benigna i maligna stanja te uključuju dvije standardne mamografske projekcije: MLO i CC. Navedeni skup podataka ima važnu ulogu u podržavanju istraživanja u području računalnog vida, strojnog učenja i razvoja novih algoritama za dijagnostiku raka dojke. Zahvaljujući velikom broju raznolikih slika i označenim regijama interesa (ROI), omogućuje se razvoj i evaluacija novih algoritama za detekciju, klasifikaciju i segmentaciju lezija na mamografskim slikama.

Kako bi se omogućila relevantna evaluacija performansi razvijenih modela, CBIS-DDSM je podijeljen na dva podskupa: skup za treniranje i skup za testiranje. Na ovaj način, istraživači mogu koristiti jedan dio podataka za treniranje i optimizaciju svojih modela, dok se drugi dio koristi za nepristrano testiranje i validaciju rezultata. (64). Mamografske snimke u skupu su pohranjene u DICOM formatu (Digital Imaging and Communications in Medicine). DICOM je industrijski standard za pohranu, prikaz i razmjenu medicinskih slika. Ovaj format omogućuje standardiziranu reprezentaciju medicinskih slika, uključujući mamografske snimke, i sadrži metapodatke poput informacija o pacijentu, postupku snimanja i tehničkim parametrima. DICOM format osigurava interoperabilnost između različitih medicinskih uređaja i softverskih platformi (65).

4.5. Mjere evaluacije modela

Za vrednovanje segmentacijskog modela korištene su sljedeće mjere evaluacije: točnost, Dice mjera i Jaccard mjera. Točnost nam govori o udjelu točno klasificiranih piksela u skupu svih piksela:

$$\frac{TP}{TP + TN + FP + FN}$$

Ova mjera nije idealna kada se radi o izrazito neuravnoteženim podacima kao što je slučaj u ovom skupu podataka. Naime, ako skup ima veliki broj primjera koji pripadaju negativnoj klasi, visoka točnost može biti postignuta jednostavnim ignoriranjem ili pogreškom klasifikacije onih primjera koji pripadaju pozitivnoj klasi. Tako bi model mogao imati visoku točnost, ali bi bio neefikasan u klasificiranju piksela koji su zapravo relevantni, poput tumorskih.

Dice mjera, jednaka F1-mjeri za segmentaciju, pruža informacije o harmonijskoj sredini preciznosti i odziva, odnosno mjeri koliko se predviđena segmentacija poklapa s pravom segmentacijom te je jedna od najčešće korištenih mjera za segmentaciju. Preciznost nam govori koliko je od svih primjera koje smo klasificirali kao pozitivne, stvarno pozitivno. Naziva se još i pozitivnom prediktivnom vrijednošću. Odziv, s druge strane, pokazuje koliko smo od svih

stvarnih pozitivnih primjera uspjeti ispravno klasificirati, odnosno otkriti. Pokazuje koliko je klasifikator "osjetljiv" na pozitivne primjere, stoga se još naziva i osjetljivošću (engl. *sensitivity*). No, iako su preciznost i odziv korisne mjere, same za sebe ne daju cjelovitu sliku. Naime, moguće je da model ima visoku preciznost, ali niski odziv (ili obrnuto), što bi značilo da ima mnoštvo lažno negativnih nalaza te je time neefikasna. Radi toga se izračunava njihova harmonijska sredina kako bi dobili uvid u obje mjere. Formula za izračun Dice mjere je sljedeća:

$$\frac{2TP}{2TP + FP + FN}$$

U kontekstu segmentacije, korisno je Dice mjeru prikazati pomoću površina preklapanja između predviđenog i stvarnog segmenta. U tom slučaju, Dice mjera se izračunava kao dvostruko preklapanje između predviđenog i stvarnog segmenta podijeljeno sa zbrojem površina oba segmenta. Formula za Dice mjeru na temelju toga je:

$$Dice = \frac{2 \cdot |P \cap G|}{|P| + |G|}$$

gdje su P i G skupovi piksela predviđenog i stvarnog segmenta, $|P \cap G|$ predstavlja broj piksela koji se nalaze i u predviđenom i u stvarnom segmentu, a $|P| + |G|$ je ukupni broj piksela u oba segmenta.

Jaccardova mjera, koja se ponekad naziva i Jaccardov indeks ili Jaccardova sličnost, još je jedna često korištena mjera za ocjenjivanje performansi segmentacije. Jaccardov indeks je definiran kao površina preklapanja između predviđenog i stvarnog segmenta podijeljena s površinom unije tih dvaju segmenta, odnosno kao omjer presjeka i unije segmenta (engl. *intersection over union*, IoU). Izračunava se pomoću formule:

$$Jaccard = \frac{TP}{TP + FP + FN} = \frac{|P \cap G|}{|P \cup G|}$$

Obje ove mjere su visoko korelirane te daju podjednake rezultate između 0 i 1, pri čemu 1 označava potpunu podudarnost između predviđene i stvarne segmentacije, a 0 znači da nema preklapanja (66).

S obzirom na to da se ovdje radi o višeklasnoj segmentaciji, za potpunu analizu bilo je nužno izračunati ove mjere za svaku klasu posebno - pozadinsku, benignu i malignu. Nakon toga, korišteno je uprosječavanje tih mjera - takozvano makro uprosječavanje - kako bih se dobila jedinstvena mjera performansi. Makro uprosječavanje podrazumijeva izračunavanje metrike za svaku klasu neovisno, a zatim uzimanje aritmetičke sredine tih mjera. Ova metoda osigurava da svaka klasa ima jednaku težinu u konačnoj procjeni performansi modela, čime se neutralizira učinak neuravnoteženosti klasa. Primjenom ovakvog pristupa u evaluaciji, omogućeno je precizno vrednovanje modela u detekciji svih vrsta lezija, bez obzira na njihovu zastupljenost u skupu podataka, što je od ključne važnosti za kvalitetnu primjenu modela u stvarnom svijetu.

4.6. Razvoj modela

4.6.1. Obrada slika

Prva faza pristupa bila je temeljita obrada slika, što je ključna etapa za stvaranje preciznih modela strojnog učenja. Ovaj postupak je uključivao eliminaciju bijelih artefakata nastalih tijekom digitalizacije, jer oni mogu narušiti točnost modela. Također su uklonjene oznake sa slika kako bi se osiguralo da model nije obučavan na temelju nepotrebnih informacija. Izvršena je i rotacija svih slika u jedinstvenu orijentaciju kako bi se osigurala konzistentnost u podacima. Na kraju, korištena je tehnika histogram ekvalizacije da bih se pojačao kontrast na slikama, čime se poboljšava vidljivost detalja (53).

U idućem su koraku u fokusu bile segmentacijske maske. Svaka lezija unutar CBIS-DDSM skupa podataka dolazi sa svojom segmentacijskom maskom. Da bi se proces treniranja pojednostavio, sve su maske spojene u jednu za svaku sliku. Korištenjem jedinstvenih oznaka pacijenata, koje su bile pohranjene u DICOM formatu, svaka je maska povezana s odgovarajućom mamografskom snimkom. Također je svaka slika preimenovana da uključuje broj pacijenta i vrstu projekcije, olakšavajući daljnje manipulacije i analize.

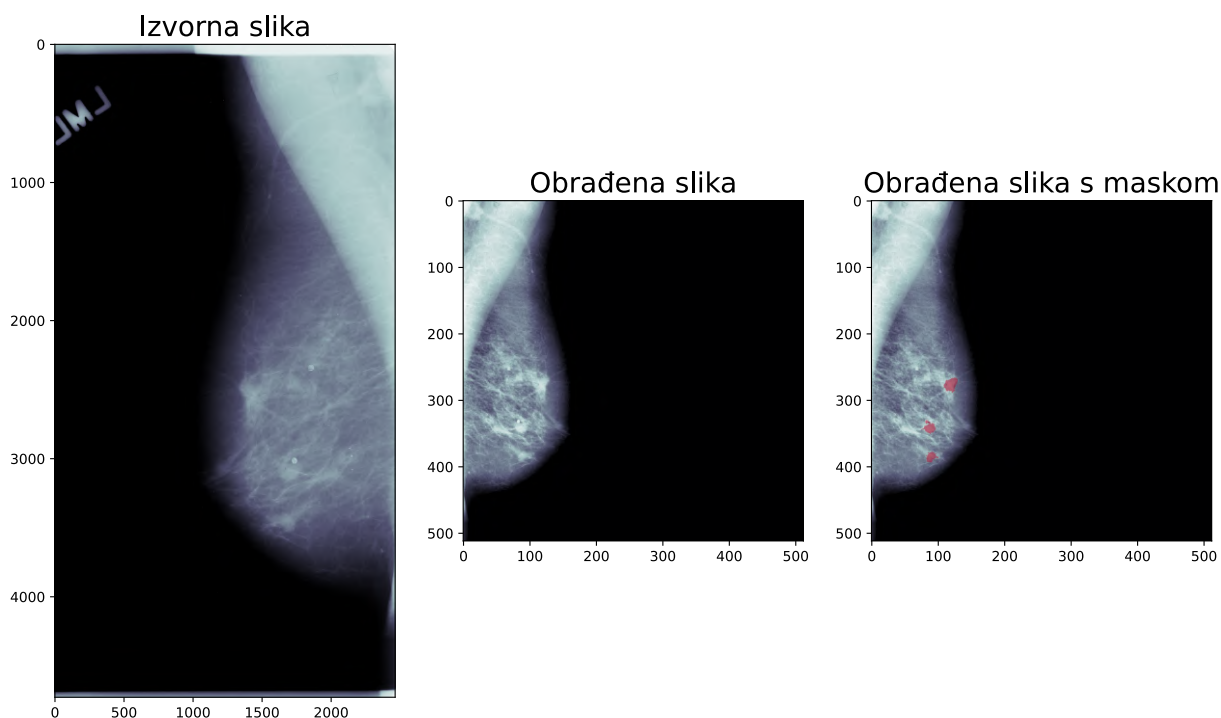
U sljedećem je koraku izvršena prilagodba veličine slika na dimenzije 256x256 i 512x512 piksela. Tijekom ovog procesa korištena je bilinearna interpolacija jer ne dovodi do stvaranja artefakata. Ova prilagodba veličine provedena je kako bi se osigurala kompatibilnost s arhitekturama dubokog učenja koje zahtijevaju kvadratne ulazne slike. Za slike koje nisu originalno bile kvadratne, korištene su tehnike proširivanja nulama kako bi se postigla kvadratna dimenzija (Slika 12).

Završna faza uključivala je konverziju svih slika u 8-bitni PNG format. Ovo je izvedeno kako bi se osiguralo očuvanje visoke kvalitete slika, ali i kako bi se smanjila njihova veličina, što je olakšalo njihovu pohranu i manipulaciju u kasnijim koracima istraživanja.

4.6.2. Treniranje modela

Kao ključni dio svog pristupa, implementirana je tehnika finog podešavanja modela SegFormer na skupu podataka CBIS-DDSM. Da bi se model prilagodio specifičnostima ovog skupa podataka, provedena je serija eksperimenata kako bi se odredili najprikladniji parametri za treniranje.

Funkcije gubitka su kritične za vođenje procesa učenja i pronalaska minimuma, te je odabir prave funkcije mogao značajno utjecati na točnost modela. Procjenjen je i utjecaj različitih vrsta histogram ekvalizacije. Histogram ekvalizacija utječe na kontrast slika, što zauzvrat može utjecati na sposobnost detekcije lezija. Korištena je "obična" normalizacija histograma te adaptivna (CLAHE) koja dovodi do manjeg povećanja šuma na slici (67). Primijenjena je i tehnika



Slika 12. Izvorna slika u DICOM formatu (lijevo) i procesirana slika s maskom i bez nje. Uočljivo je povećanje detalja i bolja vidljivost lezije zbog histogram ekvalizacije

izravnavanja oznaka (engl. *label smoothing*) koja dodaje šum oznakama prilikom izračunavanja gubitka (61). Nadalje, ispitan je utjecaj različitih veličina modela SegFormer arhitekture. Ovo je bitno jer modeli s više parametara imaju veći kapacitet za učenje složenih značajki iz podataka. Isto tako, isprobane su različite rezolucije slika, jer različite rezolucije mogu utjecati na performanse modela i brzinu treniranja.

Prilikom treniranja modela, početno je korištena stopa učenja od 1×10^{-3} , bez L2 regularizacije, i veličina grupe od 32. Primijenjeno je i rano zaustavljanje kao bi se spriječilo predugo treniranje. Kada je bilo nemoguće trenirati model koristeći predviđenu veličinu grupe zbog ograničenih resursa grafičke memorije, primijenjena je akumulaciju gradijentata. Akumulacija gradijenta je tehnika koja omogućava efikasno treniranje s većim veličinama grupe, čak i kada memorija ne dozvoljava direktan upis. To uključuje akumulaciju gradijenta preko više mini-grupa i ažuriranje težina samo nakon što su svi gradijenti akumulirani (37).

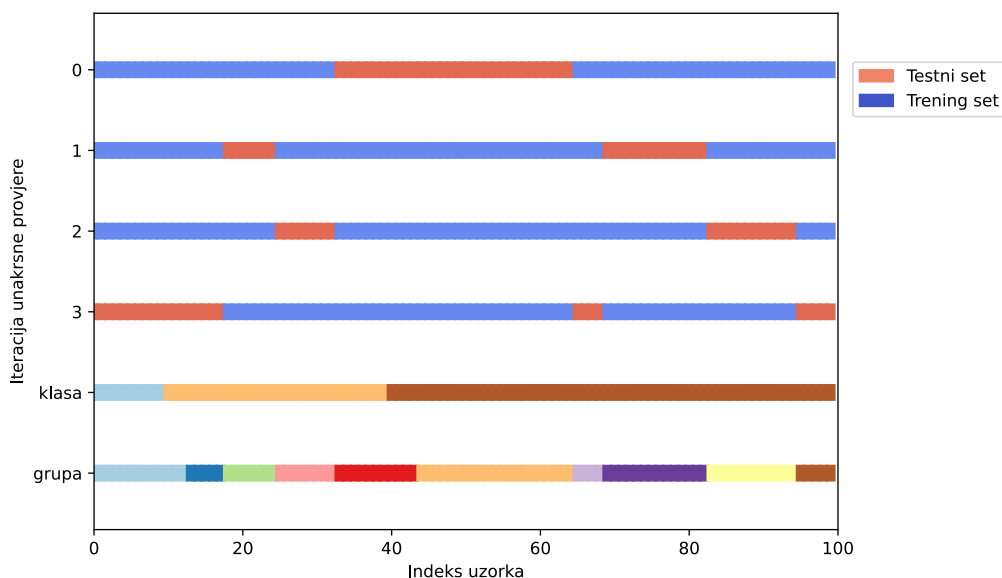
Nakon što je odabran model na temelju ovih eksperimenata, provedena je optimizacija hiperparametara. Koristeći ovaj postupak, odabrane su najbolje vrijednosti stope učenja i regularizacije, usmjeravajući se prema optimalnoj vrijednosti Dice mjere. Implementirane su različite stope učenja za enkoderski i dekoderski dio modela. Razlog tome je što se parametri dekoderskog dijela nasumično inicijaliziraju prilikom treniranja na novom skupu podataka kako bi imali odgovarajuće izlazne dimenzije, dok se parametri enkoderskog dijela očuvavaju jer sadrže generalne značajke primjenjive za više zadataka (60). Zbog toga je početnom dijelu modela potrebna manja stopa učenja, a završnom veća (37).

Tijekom treniranja, korištene su augmentacije slika kako bih povećao robusnost modela. U početku sam samo koristio horizontalno zrcaljenje. Kasnije su dodane nasumične promjene u kontrastu i svjetlini do maksimalno 5 posto i vertikalno zrcaljenje. Svi rezultati ovih eksperimenata detaljno su prikazani u tablici.

4.6.3. Validacija modela

Validacija modela izvedena je kroz postupak poznat kao stratificirana grupna k-struka unakrsna provjera (engl. *Stratified Group K-Fold Cross Validation*). Ova metoda se često koristi u situacijama gdje je važno očuvati distribuciju klasa među setovima za treniranje i validaciju.

Uobičajena k-struka unakrsna provjera dijeli skup podataka na k podskupova (engl. *folds*), gdje se jedan podskup koristi za validaciju modela, a ostali za treniranje. Ovaj postupak se ponavlja k puta, tako da svaki podskup jednom služi kao set za validaciju. Stratificirana k-struka unakrsna provjera ide korak dalje i osigurava da svaki podskup ima otprilike isti omjer klasa kao i cjelokupni skup podataka. Korisna je primjena stratifikacije kada imamo neravnomjeran broj primjera svake klase u skupu podataka. Dodavanjem koncepta grupa u ovaj postupak, osigurava se da podaci iz iste grupe (u ovom slučaju, podaci od istog pacijenta) nisu raspršeni između setova za treniranje i validaciju. Ovo je važno kako bi se izbjeglo "curenje informacija" između setova, što može dovesti do precijenjenih procjena performansi modela (Slika 13). U ovom radu, korištena je četverostruka unakrsna provjera, što znači da su podaci podijeljeni četiri puta na podskupove.



Slika 13. Vizualizacija stratificirane grupne k-struke unakrsne provjere. Generirano kodom dostupnim u scikit-learn biblioteci (68).

Konačno, za evaluaciju modela na testnom skupu, korišten je pristup poznat kao ansambl

modela, što podrazumijeva kombiniranje modela koji su trenirani na različitim podskupovima. Kombiniranje predikcija ovih modela često rezultira boljim performansama, jer se slabosti pojedinačnih modela međusobno balansiraju, dok se njihove snage pojačavaju (37).

4.7. Rezultati

Dice gubitak se pokazao najboljom funkcijom gubitka za ovaj zadatak (Tablica 2). Korištenjem ove funkcije postignut je najbolji rezultat Dice mjere, najvjerojatnije radi činjenice jer ovaj gubitak direktno optimizira istoimenu mjeru.

Tablica 2. Utjecaj različitih funkcija gubitka na validacijske rezultate

Funkcija gubitka	Dice mjera	Jaccard mjera	Točnost
Pogreška unakrsne entropija	0.478	0.425	0.998
Fokalni gubitak	0.493	0.434	0.998
Dice gubitak	0.503	0.434	0.997
Jaccard gubitak	0.499	0.436	0.997
Tversky gubitak	0.501	0.436	0.997

Napomena: Korišten je najmanji model SegFormer b0, stopa učenja 1×10^{-3} , bez augmentacija i L2 regularizacije

Primjenom histogram ekvalizacije dobiveni su iznenađujući rezultati. Ni obična ekvalizacija histograma ni CLAHE nisu donijele očekivane rezultate kada su primijenjene na sve kanale slike. Međutim, kada sam različite metode ekvalizacije primjenio na različite kanale - bez ekvalizacije na prvom kanalu, obična ekvalizacija na drugom kanalu, a CLAHE na trećem - postignuto je značajno poboljšanje (Tablica 3).

Tablica 3. Utjecaj histogram ekvalizacije na validacijske rezultate

Tip histogram ekvalizacije	Dice mjera	Jaccard mjera	Točnost
Obična histogram ekvalizacija	0.504	0.433	0.997
CLAHE	0.500	0.430	0.997
Kombinirano	0.514	0.441	0.997

Napomena: Korišten je Dice gubitak

Primjenom izravnavanja oznaka i augmentacija također su postignuti bolji rezultati (Tablica 4). Model s većim broje parametara primijenjen na većoj rezoluciji slika postigao je bolje rezultate u odnosu na najmanji model (Tablica 5). Nakon postupka optimizacije hiperparametara dobiven je konačni model te je evaluiran na testnom skupu podataka (Tablica 6, Slika 14)

Tablica 4. Utjecaj izravnavanja oznaka i augmentacija na validacijske rezultate

Mjera	<i>Label smoothing</i>	Augmentacije ¹
Točnost	0.997	0.997
Dice mjera	0.521	0.531
Jaccard mjera	0.446	0.455

¹ Horizontalno i vertikalno zrcaljenje, promjena kontrasta i svjetljine do 5%

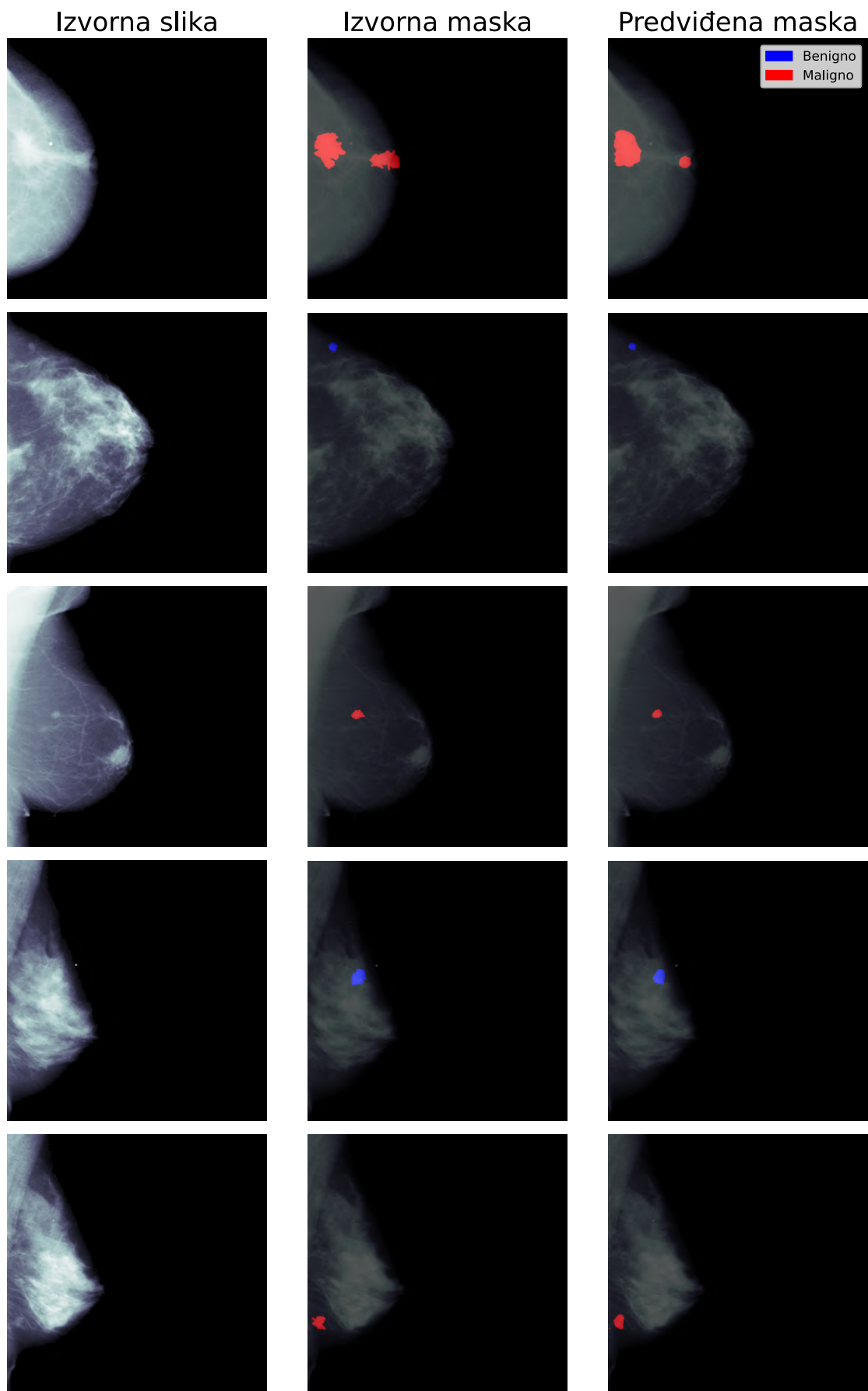
Tablica 5. Usporedba utjecaja veličine modela i rezolucije slika na validacijske rezultate (u zagradama je označen broj parametara modela)

Model	Mjera	256x256	512x512
SegFormer b0 (3.7M)	Dice mjera	0.544	0.578
	Jaccard mjera	0.463	0.489
	Točnost	0.998	0.998
Segformer b1 (13.7M)	Dice mjera	0.564	0.610
	Jaccard mjera	0.477	0.513
	Točnost	0.998	0.998

Napomena: Modeli su trenirani stopom učenja od 1×10^{-3} za MLP blok i 1×10^{-5} za enkoder dio

Tablica 6. Rezultati konačnog modela nakon postupka optimizacije hiperparametara

Skup	Dice mjera	Jaccard mjera	Točnost
Validacijski	0.642	0.535	0.998
Testni	0.629	0.527	0.998



Slika 14. Predviđanja modela na uzorku iz testnog skupa CBIS-DDSM

Konačni model baziran je na SegFormer b1 arhitekturi zbog optimalnog rezultata i brzine treniranja. Stopa učenja je postavljena na 1×10^{-2} za dekoderski dio, a za enkoderski dio na 1×10^{-4} . Tijekom evaluacije, model nije imao problema prepoznati i razlikovati lezije dojki bez obzira na veličinu. Rezultat na testnom setu ne razlikuje se značajno u odnosu na rezultat unakrsne provjere, što potvrđuje sposobnost generalizacije modela. Niži je rezultat postignut zbog toga što previđene konture lezija ne odgovaraju u potpunosti stvarnima u skupu podataka.

4.8. Rasprava

Tijekom treniranja modela suočavao sam se s nekoliko limitacija, od kojih su neke inherentne samom skupu podataka, dok su druge bile vezane za tehničke i resursne prepreke. Prva, i možda najbitnija limitacija, povezana je s prirodom CBIS-DDSM skupa podataka koji je korišten u ovom radu. Ovaj skup podataka sadrži digitalizirane filmske mamograme, koje su po definiciji manje precizne od *full-field* digitalnih mamograma. Usprkos tome što su ove slike bile neophodne za treniranje modela, suboptimalna kvaliteta podataka mogla je utjecati na rezultate dobivene tijekom procesa dubokog učenja i smanjiti sposobnost generalizacije modela na novim snimkama.

Druga značajna prepreka odnosila se na broj primjera dostupnih u skupu podataka. Kako transformerski modeli, poput SegFormera, zahtijevaju velike količine podataka za postizanje optimalnih performansi, ovaj nedostatak mogao je potencijalno smanjiti kvalitetu konačnih rezultata. Osim toga, nepravilnosti i nesavršenosti segmentacijskih maski unutar samog skupa podataka mogle su također utjecati na točnost modela.

Kada je riječ o izazovima u modeliranju, neke od potencijalnih optimizacija su ostale neiskorištene. U idealnim uvjetima, provela bi se opsežnija optimizacija hiperparametara modela, međutim, zbog ograničenih računalnih resursa, to nije bilo moguće. Nadalje, rezolucija slika i veličina modela također su bili ograničeni hardverskim kapacitetima.

Usporedbom ovoga rada s dostupnom literaturom, vidljivo je da predstavlja prvi takav pristup. Nijedan drugi rad koji koristi SegFormer arhitekturu za ovaj problem nije pronađen. Također, većina radova koji provode segmentaciju na CBIS-DDSM skupu ne razlikuje benigne od malignih masa, a često koriste vlastitu podjelu na skupove za trening i testiranje, što otežava usporedbu. Kao posljedica toga, rezultati prikazani u literaturi često djeluju viši u odnosu na model u ovom radu. Postoje i radovi koji provode segmentaciju masa bez razlikovanja benignih i malignih. Na primjer, YOLO-LOGO model, koji koristi dijelove transformerske arhitekture i primijenjen je za segmentaciju masa na cijeloj mamografskoj snimci, postiže Dice rezultat od 74.5%, a Jaccard rezultat od 64.0% (69). Najčešće korišteni modeli u literaturi su modifikacije klasične U-net arhitekture. Povezivanjem dva takva modela s opsežnom augmentacijom, neki autori su postigli Dice rezultat od 89.52% na CBIS-DDSM skupu podataka (70). Važno je na-

pomenuti da u tom radu nisu korištene cijele mamografske snimke, već samo njihovi dijelovi koji sadrže leziju. S druge strane, korištenjem cijelih mamografskih snimki, ali sa selekcijom manjeg broja za evaluaciju, drugi autori su postigli rezultat od 82.9% koristeći varijaciju U-net arhitekture (71). Pronađen je i rad u kojem autori navode da njihov model postiže Dice rezultat od 63.56% na cijelom testnom skupu podataka (72). Trend nižih rezultata na cijelom testnom skupu evidentan je i u radu gdje se pokušava standardizirati performanse različitih modela za segmentaciju masa (73). Autori su uspoređivali svoj model s često korištenim modelima u literaturi, utvrdivši znatno niže rezultate kada se koristi cijeli testni skup za evaluaciju. Važno je napomenuti da su korišteni isti parametri treniranja za sve modele, što može biti jedan od razloga nižih rezultata, ali ne može se isključiti mogućnost da rezultati prikazani u literaturi ne odražavaju stvarne performanse razvijenih modela.

Na temelju navedenog, ovaj rad predstavlja alternativni pristup detekciji i klasifikaciji lezija dojke putem višeklasne segmentacije u jednom koraku. Iako su rezultati obećavajući, usporedba s drugim radovima može biti otežana s obzirom na to da ovakav pristup nije često zastupljen u literaturi. Nadalje, testiranje modela na samo jednom skupu podataka možda ne prikazuje njegove stvarne performanse. Za sveobuhvatniju evaluaciju, primjena na snimkama iz različitih skupova podataka ili, idealno, u kliničkim scenarijima, bila bi potrebna.

5. Zaključak

U ovom radu istražio sam značaj dubokog učenja kao naprednog alata u ranom otkrivanju raka dojke na mamografskim snimkama. Usprkos ograničenim resursima i početničkom iskustvu u primjeni, razvijeni model pokazao je značajnu sposobnost segmentacije lezija dojki.

Primjenjujući temeljne principe dubokog učenja, optimizacije i strategije za izbjegavanje prenaučivosti, oblikovao sam model koji efikasno koristi dostupne podatke i pruža pouzdane rezultate. Kroz upotrebu metoda semantičke segmentacije, model je uspješno identificirao i segmentirao lezije dojki na mamografskim snimkama.

Iako su rezultati mog rada ohrabrujući, oni također ističu potrebu za daljnjim istraživanjima i razvojem ovakvih modela. Ključno je provesti dodatna ispitivanja u kliničkom okruženju, koristeći podatke iz stvarnog svijeta, kako bi se bolje razumjela korisnost i primjenjivost modela dubokog učenja u praksi. Moj rad služi kao poticaj za daljnje razvijanje i testiranje ovakvih modela, s krajnjim ciljem poboljšanja dijagnostike i liječenja raka dojke putem dubokog učenja.

LITERATURA

- [1] World Health Organization. Breast Cancer [internet]; 2021 [pristupljeno 24.6.2023.]. Dostupno na: <https://www.who.int/news-room/fact-sheets/detail/breast-cancer>.
- [2] Hrvatski zavod za javno zdravstvo. Odjel za programe probira raka dojke [internet]; 2023 [pristupljeno 24.6.2023.]. Dostupno na: <https://www.hzjz.hr/sluzba-epidemiologija-prevencija-nezaraznih-bolesti/odjel-za-programe-probira-raka-dojke>.
- [3] Samuel AL. Some Studies in Machine Learning Using the Game of Checkers. IBM Journal of Research and Development. 1959 Jul;3(3):210-29. doi: 10.1147/rd.33.0210.
- [4] Goodfellow I, Bengio Y, Courville A. Deep Learning [internet]. MIT Press; 2016 [pristupljeno: 15.5.2023.]. Dostupno na: <http://www.deeplearningbook.org>.
- [5] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015 May;521(7553):436-44. doi: 10.1038/nature14539.
- [6] Gold RH, Bassett LW, Widoff BE. Highlights from the history of mammography. Radiographics. 1990 Nov;10(6):1111-31. doi: 10.1148/radiographics.10.6.2259767.
- [7] D'Orsi CJ, Sickles EA, Mendelson EB, Morris EA. ACR BI-RADS Atlas: Breast Imaging Reporting and Data System ; Mammography, Ultrasound, Magnetic Resonance Imaging, Follow-up and Outcome Monitoring, Data Dictionary. 5th ed. ACR, American College of Radiology; 2013.
- [8] Damjanov I, Seiwerth S, Jukić S, Nola M. Patologija. 5th ed. Zagreb: Medicinska naklada; 2018.
- [9] Berment H, Becette V, Mohallem M, Ferreira F, Chérel P. Masses in mammography: What are the underlying anatomopathological lesions? Diagnostic and Interventional Imaging. 2014 Feb;95(2):124-33. doi: 10.1016/j.diii.2013.12.010.

- [10] Pesce K, Orruma MB, Hadad C, Cano YB, Secco R, Cernadas A. BI-RADS Terminology for Mammography Reports: What Residents Need to Know. *RadioGraphics*. 2019 Mar;39(2):319-20. doi: 10.1148/rg.2019180068.
- [11] Murphy A, Radswiki T. Mammography views [internet]. *Radiopaedia.org*; 2011 [pristupljeno 1.6.2023.]. Dostupno na: <https://doi.org/10.53347/rid-15023>.
- [12] Sechopoulos I, Teuwen J, Mann R. Artificial intelligence for breast cancer detection in mammography and digital breast tomosynthesis: State of the art. *Seminars in Cancer Biology*. 2021 Jul;72:214-25. doi: 10.1016/j.semcancer.2020.06.002.
- [13] Gaillard F. Spiculated breast cancer. Case study. *Radiopaedia.org*; 2010 [pristupljeno 25.6.2023.]. Dostupno na: <https://doi.org/10.53347/rID-12608>.
- [14] Stachs A, Stubert J, Reimer T, Hartmann S. Benign Breast Disease in Women. *Deutsches Ärzteblatt international*. 2019 Aug. doi: 10.3238/arztebl.2019.0565.
- [15] Burrell HC, Pinder SE, Wilson ARM, Evans AJ, Yeoman LJ, Elston CW, et al. The positive predictive value of mammographic signs: A review of 425 non-palpable breast lesions. *Clinical Radiology*. 1996 Apr;51(4):277-81. doi: 10.1016/s0009-9260(96)80346-1.
- [16] Niknejad M. Breast mass (BI-RADS 5). Case study. *Radiopaedia.org*; 2023 [pristupljeno 28.6.2023.]. Dostupno na: <https://doi.org/10.53347/rID-163936>.
- [17] Baratelli G. Fibroadenoma: involuting with macrocalcification. Case study. *Radiopaedia.org*; 2014 [pristupljeno 28.6.2023.]. Dostupno na: <https://doi.org/10.53347/rID-29466>.
- [18] Jordan V, Khan M, Prill D. Breast Cancer Screening: Why Can't Everyone Agree? *Primary Care: Clinics in Office Practice*. 2019 Mar;46(1):97-115. doi: 10.1016/j.pop.2018.10.010.
- [19] Myers ER, Moorman P, Gierisch JM, Havrilesky LJ, Grimm LJ, Ghatge S, et al. Benefits and Harms of Breast Cancer Screening. *JAMA*. 2015 Oct;314(15):1615. doi: 10.1001/jama.2015.13183.
- [20] Marmot MG, Altman DG, Cameron DA, Dewar JA, Thompson SG, et al. The benefits and harms of breast cancer screening: an independent review. *British Journal of Cancer*. 2013 Jun;108(11):2205-40. doi: 10.1038/bjc.2013.177.

- [21] Dibden A, Offman J, Duffy SW, Gabe R. Worldwide Review and Meta-Analysis of Cohort Studies Measuring the Effect of Mammography Screening Programmes on Incidence-Based Breast Cancer Mortality. *Cancers*. 2020 Apr;12(4):976. doi: 10.3390/cancers12040976.
- [22] Sun L, Legood R, dos Santos-Silva I, Gaiha SM, Sadique Z. Global treatment costs of breast cancer by stage: A systematic review. *PLOS ONE*. 2018 Nov;13(11):e0207993. doi: 10.1371/journal.pone.0207993.
- [23] Ebell MH, Thai TN, Royalty KJ. Cancer screening recommendations: an international comparison of high income countries. *Public Health Reviews*. 2018 Mar;39(1). doi: 10.1186/s40985-018-0080-0.
- [24] European Commission. European guidelines on breast cancer screening and diagnosis; 2023 pristupljeno 24.6.2023. Dostupno na: <https://healthcare-quality.jrc.ec.europa.eu/ecibc/european-breast-cancer-guidelines>.
- [25] Seely JM, Alhassan T. Screening for Breast Cancer in 2018—What Should We be Doing Today? *Current Oncology*. 2018 Jun;25(11):115-24. doi: 10.3747/co.25.3770.
- [26] Lehman CD, Wellman RD, Buist DSM, Kerlikowske K, Tosteson ANA, Miglioretti DL. Diagnostic Accuracy of Digital Screening Mammography With and Without Computer-Aided Detection. *JAMA Internal Medicine*. 2015 Nov;175(11):1828. doi: 10.1001/jama-intermed.2015.5231.
- [27] Kohli A, Jha S. Why CAD Failed in Mammography. *Journal of the American College of Radiology*. 2018 Mar;15(3):535-7. doi: 10.1016/j.jacr.2017.12.029.
- [28] Chan HP, Samala RK, Hadjiiski LM. CAD and AI for breast cancer—recent development and challenges. *The British Journal of Radiology*. 2020 Apr;93(1108):20190580. doi: 10.1259/bjr.20190580.
- [29] McKinney SM, Sieniek M, Godbole V, Godwin J, Antropova N, Ashrafian H, et al. International evaluation of an AI system for breast cancer screening. *Nature*. 2020 Jan;577(7788):89-94. doi: 10.1038/s41586-019-1799-6.
- [30] Rodriguez-Ruiz A, Lång K, Gubern-Merida A, Broeders M, Gennaro G, Clauser P, et al. Stand-Alone Artificial Intelligence for Breast Cancer Detection in Mammography: Comparison With 101 Radiologists. *JNCI: Journal of the National Cancer Institute*. 2019 Mar;111(9):916-22. doi: 10.1093/jnci/djy222.

- [31] Dembrower K, Wåhlin E, Liu Y, Salim M, Smith K, Lindholm P, et al. Effect of artificial intelligence-based triaging of breast cancer screening mammograms on cancer detection and radiologist workload: a retrospective simulation study. *The Lancet Digital Health*. 2020 Sep;2(9):e468-74. doi: 10.1016/s2589-7500(20)30185-0.
- [32] Géron A. *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly Media, Incorporated; 2019.
- [33] McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*. 1943 Dec;5(4):115-33. doi: 10.1007/bf02478259.
- [34] Muthukrishnan N, Maleki F, Ovens K, Reinhold C, Forghani B, Forghani R. Brief History of Artificial Intelligence. *Neuroimaging Clinics of North America*. 2020 Nov;30(4):393-9. doi: 10.1016/j.nic.2020.07.004.
- [35] Superm401, temeljeno na originalu korisnika User:Lightning13. Datoteka:Artificial neuron or.svg [slika s interneta]; 2007 [pristupljeno 25.6.2023.]. Javna domena, putem Wikimedia Commonsa. Dostupno na: https://commons.wikimedia.org/wiki/File:Artificial_neuron_or.svg.
- [36] Rosenblatt F. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*. 1958;65(6):386-408. doi: 10.1037/h0042519.
- [37] Howard J, Gugger S. *Deep Learning for Coders with Fastai and Pytorch: AI Applications Without a PhD*. O'Reilly Media, Incorporated; 2020.
- [38] Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. *Nature*. 1986 Oct;323(6088):533-6. doi: 10.1038/323533a0.
- [39] LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, et al. Back-propagation Applied to Handwritten Zip Code Recognition. *Neural Computation*. 1989 Dec;1(4):541-51. doi: 10.1162/neco.1989.1.4.541.
- [40] Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:201011929*. 2020.
- [41] Xie E, Wang W, Yu Z, Anandkumar A, Alvarez JM, Luo P. SegFormer: Simple and efficient design for semantic segmentation with transformers. *Advances in Neural Information Processing Systems*. 2021;34:12077-90.

- [42] Alpaydin E. Introduction to Machine Learning, fourth edition. Adaptive Computation and Machine Learning series. MIT Press; 2020.
- [43] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is all you need. Advances in neural information processing systems. 2017;30.
- [44] Cecbur. Datoteka:Artificial Neural Network.gif [slika s interneta]; 2019 [pristupljeno: 24.6.2023.]. CC BY-SA 4.0, preko Wikimedia Commonsa. Dostupno na: https://commons.wikimedia.org/wiki/File:Artificial_Neural_Network.gif.
- [45] Jordan J. Neural networks: representation [internet]; 2017 [pristupljeno 16.5.2023.]. Dostupno na: <https://www.jeremyjordan.me/intro-to-neural-networks>.
- [46] Gerazov B. Datoteka:Artificial neuron.svg [slika s interneta]; 2023 [pristupljeno: 24.6.2023.]. CC BY-SA 4.0, preko Wikimedia Commonsa. Dostupno na: https://commons.wikimedia.org/wiki/File:Artificial_neuron.svg.
- [47] Burkov A. The Hundred-Page Machine Learning Book. Andriy Burkov; 2019.
- [48] Hendrycks D, Gimpel K. Gaussian error linear units (gelus). arXiv preprint arXiv:160608415. 2016.
- [49] Glorot X, Bordes A, Bengio Y. Deep sparse rectifier neural networks. In: Proceedings of the fourteenth international conference on artificial intelligence and statistics. JMLR Workshop and Conference Proceedings; 2011. p. 315-23.
- [50] Vapnik VN. An overview of statistical learning theory. IEEE Transactions on Neural Networks. 1999;10(5):988-99. doi: 10.1109/72.788640.
- [51] Ruder S. An overview of gradient descent optimization algorithms. arXiv preprint arXiv:160904747. 2016.
- [52] Perez L, Wang J. The effectiveness of data augmentation in image classification using deep learning. arXiv preprint arXiv:171204621. 2017.
- [53] Gonzalez RC, Woods RE. Digital Image Processing [internet]. Pearson; 2018 [pristupljeno 12.6.2023.]. Dostupno na: <https://dl.icdst.org/pdfs/files4/01c56e081202b62bd7d3b4f8545775fb.pdf>.
- [54] Klette R. Concise Computer Vision: An Introduction into Theory and Algorithms [internet]. Undergraduate Topics in Computer Science. Springer London; 2014 [pristupljeno 11.6.2023.]. Dostupno na: https://doc.lagout.org/science/0_Computer%20Science/2_Algorithms/Concise%20Computer%20Vision_%20An%20Introduction%20into%20Theory%20and%20Algorithms%20%5BKlette%202014-01-20%5D.pdf.

- [55] Hirahara D, Takaya E, Kadowaki M, Kobayashi Y, Ueda T. Effect of the Pixel Interpolation Method for Downsampling Medical Images on Deep Learning Accuracy. *Journal of Computer and Communications*. 2021;09(11):150-6. doi: 10.4236/jcc.2021.911010.
- [56] Vasconcelos CN, Vasconcelos BN. Increasing Deep Learning Melanoma Classification by Classical And Expert Knowledge Based Image Transforms. *ArXiv*. 2017;abs/1702.07025.
- [57] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE; 2015. doi: 10.1109/cvpr.2015.7298965.
- [58] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: *Lecture Notes in Computer Science*. Springer International Publishing; 2015. p. 234-41.
- [59] Yang L, Wang X, Zhai J. Waterline Extraction for Artificial Coast With Vision Transformers. *Frontiers in Environmental Science*. 2022 Feb;10. doi: 10.3389/fenvs.2022.799250.
- [60] Zeiler MD, Fergus R. Visualizing and understanding convolutional networks. In: *Computer Vision—ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part I 13*. Springer; 2014. p. 818-33.
- [61] He T, Zhang Z, Zhang H, Zhang Z, Xie J, Li M. Bag of tricks for image classification with convolutional neural networks. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*; 2019. p. 558-67.
- [62] Sawyer-Lee R, Gimenez F, Hoogi A, Rubin D. Curated Breast Imaging Subset of Digital Database for Screening Mammography (CBIS-DDSM) [skup podataka]. *The Cancer Imaging Archive*; 2016 [pristupljeno 1.5.2023.]. Dostupno na: <https://doi.org/10.7937/K9/TCIA.2016.7O02S9CY>.
- [63] Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, et al. The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository. *Journal of Digital Imaging*. 2013 Jul;26(6):1045-57. doi: 10.1007/s10278-013-9622-7.
- [64] Lee RS, Gimenez F, Hoogi A, Miyake KK, Gorovoy M, Rubin DL. A curated mammography data set for use in computer-aided detection and diagnosis research. *Scientific Data*. 2017 Dec;4(1). doi: 10.1038/sdata.2017.177.
- [65] Bidgood WD, Horii SC, Prior FW, Syckle DEV. Understanding and Using DICOM, the Data Interchange Standard for Biomedical Imaging. *Journal of the American Medical Informatics Association*. 1997 May;4(3):199-212. doi: 10.1136/jamia.1997.0040199.

- [66] Taha AA, Hanbury A. Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. *BMC Medical Imaging*. 2015 Aug;15(1). doi: 10.1186/s12880-015-0068-x.
- [67] OpenCV. Histograms - 2: Histogram Equalization; 2023 [pristupljeno 28.6.2023.]. Dostupno na: https://docs.opencv.org/4.x/d5/daf/tutorial_py_histogram_equalization.html.
- [68] Scikit-learn developers. 3.1. Cross-validation: evaluating estimator performance; 2023 [pristupljeno 28.6.20.23.]. Dostupno na: https://scikit-learn.org/stable/modules/cross_validation.html.
- [69] Su Y, Liu Q, Xie W, Hu P. YOLO-LOGO: A transformer-based YOLO segmentation model for breast mass detection and segmentation in digital mammograms. *Computer Methods and Programs in Biomedicine*. 2022;221:106903. doi:10.1016/j.cmpb.2022.106903.
- [70] Baccouche A, Garcia-Zapirain B, Castillo Olea C, Elmaghraby AS. Connected-UNets: a deep learning architecture for breast mass segmentation. *NPJ Breast Cancer*. 2021;7(1):151. doi: 10.1038/s41523-021-00358-x.
- [71] Rajalakshmi NR, Vidhyapriya R, Elango N, Ramesh N. Deeply supervised u-net for mass segmentation in digital mammograms. *International Journal of Imaging Systems and Technology*. 2021;31(1):59-71. doi: 10.1002/ima.22516.
- [72] Tiryaki VM. Mass segmentation and classification from film mammograms using cascaded deep transfer learning. *Biomedical Signal Processing and Control*. 2023;84:104819. doi: 10.1016/j.bspc.2023.104819.
- [73] Sarker P, Sarker S, Bebis G, Tavakkoli A. ConnectedUNets++: Mass Segmentation from Whole Mammographic Images. In: *International Symposium on Visual Computing*. Springer; 2022. p. 419-30.

ŽIVOTOPIS

Karlo Pintarić rođen je 29. studenog 1998. godine u Čakovcu. Svoje srednjoškolsko obrazovanje stekao je u Prvoj gimnaziji Varaždin, gdje je 2016. godine ostvario iznimno postignuće postavši državni prvak iz biologije. Nakon srednje škole, 2017. godine, upisao se na Medicinski fakultet Sveučilišta u Zagrebu. Već na prvoj godini studija, njegov uspjeh prepoznat je dodjelom Dekanove nagrade. Karlo je od rane dobi pokazivao interes za računalnu tehnologiju, a od 2021. godine aktivno se posvećuje učenju programiranja u Pythonu. Poseban interes pokazuje za područje umjetne inteligencije i dubokog učenja. Sudjelovao je na nekoliko natjecanja iz područja strojnog učenja, dodatno potvrđujući svoju strast prema ovoj disciplini.