

Učinkovitost umjetne inteligencije u razlikovanju malignih od benignih ultrazvučnih osobina tumora dojki

Marković, Filip

Master's thesis / Diplomski rad

2024

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **Josip Juraj Strossmayer University of Osijek, Faculty of Medicine Osijek / Sveučilište Josipa Jurja Strossmayera u Osijeku, Medicinski fakultet Osijek**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:152:377123>

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

Download date / Datum preuzimanja: **2025-01-22**



Repository / Repozitorij:

[Repository of the Faculty of Medicine Osijek](#)



**SVEUČILIŠTE JOSIPA JURJA STROSSMAYERA U OSIJEKU
MEDICINSKI FAKULTET OSIJEK**

**SVEUČILIŠNI INTEGRIRANI PRIJEDIPLOMSKI I
DIPLOMSKI STUDIJ MEDICINE**

Filip Marković

**UČINKOVITOST UMJETNE
INTELIGENCIJE U
RAZLIKOVANJU MALIGNIH OD
BENIGNIH ULTRAZVUČNIH OSOBINA
TUMORA DOJKI**

Diplomski rad

Osijek, 2024.

**SVEUČILIŠTE JOSIPA JURJA STROSSMAYERA U OSIJEKU
MEDICINSKI FAKULTET OSIJEK**

**SVEUČILIŠNI INTEGRIRANI PRIJEDIPLOMSKI I
DIPLOMSKI STUDIJ MEDICINE**

Filip Marković

**UČINKOVITOST UMJETNE
INTELIGENCIJE U
RAZLIKOVANJU MALIGNIH OD
BENIGNIH ULTRAZVUČNIH OSOBINA
TUMORA DOJKI**

Diplomski rad

Osijek, 2024.

Rad je ostvaren na Odjelu za radiologiju Nacionalne memorijalne bolnice „Dr. Juraj Njavro“, Vukovar.

Mentor rada: doc. dr. sc. Damir Štimac, dr. med., specijalist radiologije

Rad sadrži: 19 listova, 4 tablice i 3 slike.

PREDGOVOR

Velike zahvale mentoru, doc. dr. sc. Damiru Štimcu, dr. med., na predloženoj temi i stručnom vodstvu tijekom izrade ovog diplomskog rada.

Hvala mojim prijateljima koji su studiranje medicine učinili zanimljivijim i zabavnijim te hvala na podršci i lijepim uspomenama.

Hvala Teni što si bila uz mene sve ove godine i pomagala kad god je trebalo, koliko god teško bilo sa mnom. Studiranje je bilo lakše i ljepše uz tebe.

Najveće hvala mojim roditeljima, koji su gubili živce zbog mene, ali uvijek bili najveća podrška, gurali me naprijed i vjerovali u mene kada i sam nisam bio siguran kojim putem bih išao.

SADRŽAJ RADA

| | |
|---|----|
| 1. UVOD | 1 |
| 1.1. Građa dojke | 1 |
| 1.2. Tumori dojke | 1 |
| 1.3. Ultrazvuk dojke | 2 |
| 1.4. Core biopsija | 3 |
| 1.5. Umjetna inteligencija | 3 |
| 1.6. Neuronske mreže i <i>DenseNet</i> | 4 |
| 2. HIPOTEZA | 5 |
| 3. CILJ | 6 |
| 4. ISPITANICI I METODE | 7 |
| 4.1. Ustroj studije | 7 |
| 4.2. Ispitanici | 7 |
| 4.3. Metode | 7 |
| 4.4. UI Model | 7 |
| 4.5. Statističke metode | 8 |
| 5. REZULTATI | 9 |
| 5.1. Trening skup | 9 |
| 5.2. Validacijski skup | 9 |
| 5.3. Testni skup | 10 |
| 6. RASPRAVA | 12 |
| 7. ZAKLJUČAK | 15 |
| 9. SUMMARY | 17 |
| 10. LITERATURA | 18 |

POPIS KRATICA

2D – dvodimenzionalno

3D – trodimenzionalno

AUC – vrijednost površine ispod grafa

BI-RADS – sustav izvješćivanja i klasifikacije dojki, prema engl. *Breast imaging reporting and data system*

BIS – bolnički informacijski sustav

CAD – računalno potpomognuti dijagnostički sustavi, prema engl. *Computer-Aided Diagnostic systems*

CB – core biopsija

FN – lažno negativan, prema engl. *false negative*

FP – lažno pozitivan, prema engl. *false positive*

PTH – patohistologija

ROC – krivulja operativnih karakteristika prijemnika, prema engl. *Receiver Operating Characteristic*

TN – istinito negativno, prema engl. *true negative*

TP – istinito pozitivno, prema engl. *true positive*

UI – umjetna inteligencija

UZV – ultrazvuk

1. UVOD

1.1. Građa dojke

Dojka je modificirana žlijezda znojnica kože. Kod žena je to kompleksna funkcionalna jedinica, a u muškaraca rudimentarni organ. Građena je od kože s potkožnim masnim tkivom, žlijezda i vezivno-masne osnovice. Žljezdani ujedno je i funkcionalni dio dojke i građen je od 15 do 25 režnjeva (1). Svaki režanj ima duktuse (kanaliće) koji kreću od laktifernog duktusa te preko segmentalnih i subsegmentalnih kanalića završavaju u terminalnim duktulo-lobularnim jedinicama. Terminalne duktulo-lobularne jedinice građene su od ekstralobularnog duktusa i lobula, a sam lobul sadrži acinus (žlijezda) i intralobularni duktus okružen rahlom intalobularnom stromom. Uz samu bradavicu izvodni se kanalići šire u laktiferni sinus.

1.2. Tumori dojke

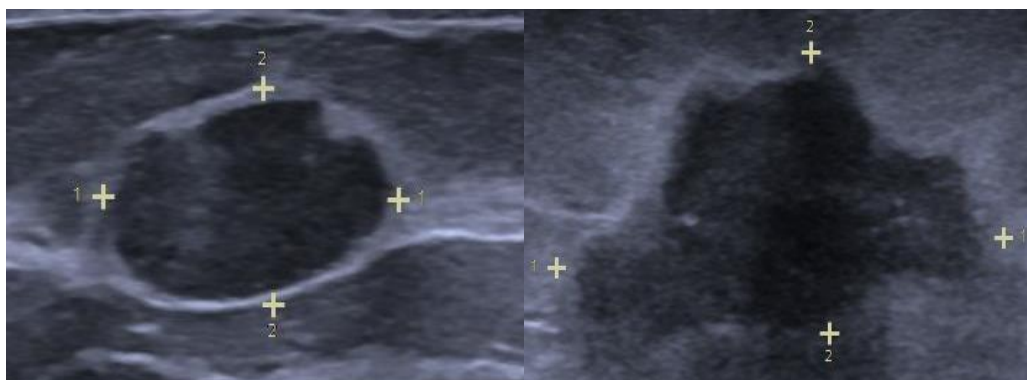
Tumori dojke mogu nastati iz epitela ili strome dojke te mogu biti benigni i maligni. Epitelni tumori mnogo su učestaliji od stromalnih tumora (1). Benigne bolesti dojke svrstavaju se u tri kliničke grupe: neproliferativne lezije i proliferativne lezije koje mogu biti bez atipije i s atipijom (2). Ova klasifikacija uključuje histološke kriterije koji su povezani s povećanim rizikom za razvoj raka dojke. Neproliferativne lezije dojke su: ciste, apokrina metaplazija, ektazija duktusa i periduktalni mastitis, blaga duktalna epitelna hiperplazija i pseudoangiomatozna stromalna hiperplazija i fibroadenom. Proliferativne lezije bez atipije uključuju: sklerozirajuća adenoza, radijalni ožiljak i kompleksne sklerozirajuće lezije, duktalna epitelna hiperplazija, intraduktalni papilom. Pod proliferativne lezije s atipijom spadaju: atipična lobularna hiperplazija, atipična duktalna hiperplazija, ravna epitelna atipija (2). Promjene kao ciste dojke, fibroadenome, intraduktalni papilom i sklerozirajuću adenozu mi možemo vizualizirati ultrazvukom (UZV), ali dijagnozu potvrđujemo biopsijom tkiva i patohistološkom (PTH) analizom. Među malignim tumorima najčešći su karcinomi dojke, koji su 100 puta učestaliji u žena nego u muškaraca. U ženskoj populaciji karcinom dojki je najčešća maligna bolest i mortalitet mu je 20 % zbog čega su veliki društveni i zdravstveni problem osobito u razvijenim zemljama (1). Karcinomi dojke mogu se podijeliti na neinvazivne (*in situ*) i invazivne karcinome dojke, prema histološkom podrijetlu i molekularnom tipu. Radi preciznog određivanja histološkog i molekularnog tipa karcinoma dojke, a radi uspješnog i individualiziranog liječenja, za pravilnu dijagnozu

karcinoma dojke potrebna je patohistološka analiza iz uzoraka dobivenih iglednom biopsijom (core ili vakuum asistirana biopsija).

1.3. Ultrazvuk dojke

Ultrazvuk je često korištena slikovna metoda u cijeloj medicini tako i u dijagnostici karcinoma dojke. Ultrazvuk može procijeniti morfologiju, orijentaciju, unutarnju strukturu i rubove lezija iz više kutova, pružajući visoku razlučivost i kod dojki s pretežno masnim tkivom, kao i kod gustih, žljezdanih struktura (3). Ultrazvuk je primarna metoda za razlikovanje cističnih od solidnih lezija otkrivenih mamografijom (4). U dijagnostici lezija dojke ultrazvukom, osim morfologije lezije, značajni čimbenici su okolno tkivo, oblik, kontura rubova, granica lezije i stražnje akustičke značajke (3). Maligni tumori obično imaju nepravilni oblik, nejasne i nepravilne rubove, značajno su hipoehogeni, mogu imati mikrokalcifikacije dok benigni tumori su ovalni ili okrugli, glatkih rubova, mogu biti hiperehogeni, izoehogeni ili blago hipoehogeni (Slika 1.) (3). Ultrazvučnim navođenjem najlakše je dovesti iglu biopsije do ciljanog tkiva za daljnju dijagnostiku. Sustav izvješćivanja i klasifikacije dojki (BI-RADS) koristi se u većini zemalja koje provode probir raka dojke. BI-RADS je osmišljen kako bi smanjio varijabilnosti među radiolozima prilikom pisanja nalaza za mamografiju, ultrazvuk ili magnetnu rezonanciju. BI-RADS je integriran izravno s mamografskim stanicama i implementiran u računalno potpomognute dijagnostičke sustave (CAD). Postoji više načina ultrazvučnog snimanja dojke: ultrasonografija dojke, ultrazvuk pojačan kontrastom, trodimenzionalni (3D) ultrazvuk, elastosonografija. 3D ultrazvuk pruža nove perspektive u području UZV dojki, gdje se čak i opremom za dvodimenzionalno (2D) snimanje može konstruirati 3D slika. Kontrastno pojačani ultrazvuk se koristi kako bi se vizualizirale vaskularne strukture i tkivo s različitom vaskularnošću, u kliničkim istraživanjima (3). Elastosonografija je postala rutinski alat u ultrazvučnoj dijagnostici koji može neinvazivno mjeriti konzistenciju ili tvrdoću tkiva kako bi se razlikovale benigne od malignih lezija dojke (3). Sva ta snimanja uvelike ovise o iskustvu radiologa te mogu imati veliku stopu varijacije među promatračima. Osim toga u kod korištenja 3D ultrazvuka postoji velika količina podataka koja se treba obraditi pa postoji rizik od predvidivih grešaka. Zato su u takve sustave integrirani CAD-ovi kako bi

pomogli radiolozima u otkrivanju i klasifikaciji raka dojke.



Slika 1. Fibroadenom (lijevo), Invazivni karcinom (desno), izvor: izradio autor rada.

1.4. Core biopsija

Core biopsija (CB) dojke je minimalno invazivna tehnika za dobivanje malih uzoraka tkiva radi histološke analize promjena u dojci. To je najdostupnija metoda za histološku dijagnozu tumora dojki. Radiolog pomoću neke od metoda oslikavanja (najčešće ultrazvučno) navodi iglu do lezije i opružnim mehanizmom uzima 3-7 cilindričnih uzoraka ciljanog tkiva. Uzorci se šalju na histološku analizu. Prijavljeni postotak lažno negativnih rezultata za malignitet kod core biopsije iznosi od 2 % do 6,7 %, s prosječnom stopom od 4,4 %. Oko 10% pokušaja biopsije bit će nerazjašnjeno (5).

1.5. Umjetna inteligencija

Umjetna inteligencija (UI) grana je računalne znanosti koja se bavi razvojem sustava za obavljanje zadataka koji inače zahtijevaju ljudsku inteligenciju. To uključuje procese poput prepoznavanja uzoraka, donošenja odluka, razumijevanja jezika, te učenja iz podataka. UI se oslanja na algoritme i matematičke modele koji omogućuju računalima da obrađuju velike količine informacija, identificiraju obrasce i izvode zaključke na temelju tih podataka. UI ima dva ključna načina učenja, a to su strojno učenje i duboko učenje. Strojno učenje podrazumijeva učenje iz podataka bez da su računala eksplicitno programirana za svaki pojedini zadatak. Duboko učenje je podskup strojnog učenja i koristi složene mreže inspirirane strukturom ljudskog mozga, poznate kao neuronske mreže, za analizu velikih skupova podataka i prepoznavanje složenih obrazaca. UI se danas koristi u raznim područjima života, pa je tako pronašla značajnu ulogu i u medicini, gdje je izrazito korisna za analizu medicinskih slika.

1.6. Neuronske mreže i *DenseNet*

Neuronske mreže ključni su koncept u umjetnoj inteligenciji, a sastoje se od slojeva čvorova koji su međusobno povezani te svojim funkcioniranjem simuliraju ljudske neurone. Svaki čvor prima podatke, obrađuje ih i šalje rezultate u sljedeći sloj, omogućujući mreži da prepoznaje složene uzorke u podacima. Sloj po sloj, neuronska mreža izvlači sve apstraktnije informacije iz podataka i zato je iznimno učinkovita u prepoznavanju slika i obradi jezika. U ovom diplomskom radu koristit će se neuronska mreža *DenseNet* (*Densely Connected Convolutional Networks*). Ova neuronska mreža ima specifičnu arhitekturu gdje svaki sloj ima pristup značajkama svih prethodnih slojeva (6). U tradicionalnim konvolucijskim mrežama informacije između slojeva se prenose samo između susjednih slojeva. Ova gusto povezana arhitektura pomaže smanjenju problema gubitka podataka ili degradacije informacija tijekom prolaska kroz mrežu, što rezultira učinkovitijom obradom. *DenseNet* pomaže u preciznom prepoznavanju karakteristika tumora i omogućuje dublje i preciznije učenje iz podataka, što je korisno u razlikovanju malignih od benignih lezija. Zbog takvog učenja smanjuje se potreba za ručnim prilagođavanjem značajki, što može ubrzati dijagnostički proces.

2. HIPOTEZA

Neuronska mreža *DenseNet* može učinkovito razlikovati maligne od benignih ultrazvučnih osobina tumora dojki, s visokom razinom točnosti koja je usporediva s dijagnozama dobivenim patohistološkom analizom iglene biopsije, kako bi se smanjio broj nepotrebnih invazivnih zahvata.

3. CILJ

Ciljevi istraživanja su:

1. Ispitati učinkovitost neuronske mreže *DenseNet* u analizi ultrazvučnih slika tumora dojki
2. Usporediti rezultate automatske klasifikacije tumora kao benignih ili malignih s dijagnozama dobivenim patohistološkom analizom iglene biopsije
3. Ocijeniti točnost (*accuracy*), osjetljivost (*sensitivity*), preciznost (*precision*) i *F1-score* UI modela u razlikovanju malignih od benignih tumora dojki

4. ISPITANICI I METODE

4.1. Ustroj studije

Retrospektivna studija s analizom sekundarnih podataka.

4.2. Ispitanici

Istraživanje se provodi na Odjelu za radiologiju Nacionalne memorijalne bolnice „Dr. Juraj Njavro“ Vukovar. U istraživanju se koristi 257 ultrazvučnih slika pacijentica s tvorbama u dojčkama koje su prošle ultrazvučni pregled i patohistološku dijagnostiku iglenom biopsijom u razdoblju od studenog 2021. do veljače 2024.

4.3. Metode

Rezultati patohistološke dijagnostike iglenom biopsijom i ultrazvučne slike bit će prikupljene preko bolničkog informacijskog sustava (BIS). Ultrazvučne slike bile su prilagođene za korištenje u UI modelu, uključujući označavanje relevantnih područja (tumora) i prilagođavanje na odgovarajuću veličinu piksela kako bi bile kompatibilne s UI modelom. Nakon prikupljanja podataka, implementirat će se neuronska mreža *DenseNet* za analizu ultrazvučnih slika. Model će se trenirati korištenjem prikupljenih podataka, pri čemu će se skup podataka podijeliti na trening, validacijski i testni skup. Rezultati modela bit će uspoređeni s dijagnozama dobivenim patohistološkom analizom iglene biopsije.

4.4. UI Model

UI model je korišten za analizu slike i klasifikaciju tumora dojke *DenseNet-121*. Ova mreža ima 121 sloj koji se koristi za analizu slike i klasifikaciju. Koristi se prethodno istrenirani model *DenseNet121* s unaprijed treniranim slikama s ImageNet-a gdje su zadnja 4 sloja otključana za fino podešavanje. Slike su podijeljene tako da 80 % slika čini trening skup, 10 % validacijski skup i 10 % testni skup. Trening skup je skup podataka na kojem se UI model uči. Model koristi te podatke da prepozna uzorke i odnose na podacima kako bi mogao prepoznati slike i donositi odluke. On je najveći od tri jer je modelu potrebno puno informacija za učenje. Validacijski skup je skup podataka koji se koristi tijekom treniranja za provjeru njegove izvedbe. Nakon što model nauči nešto iz trening skupa, validacijski skup pomaže provjeriti je li model zapravo naučio opće obrasce, a ne samo “naučio napamet“ podatke iz trening skupa. To je skup za fino podešavanje modela. Testni skup su podaci koji se koriste nakon što je model istreniran i

podešen te služi za konačnu provjeru izvedbe modela. To su potpuno novi podaci za model i oni testiraju kako se model ponaša u stvarnom kliničkom scenariju, kada dobije nove podatke. Slike tijekom treninga modela kroz model prolaze u epohama. Jedna epoha znači prolazak modela kroz cijeli skup podataka za treniranje jednom. U ovom modelu odabrano je da program odradi 15 epoha, ali je implementirano rano stajanje (*early stopping*) kako bi se spriječilo prekomjerno treniranje i poboljšala učinkovitost modela. *Early stopping* postavljen je s ciljem da prati performanse modela na validacijskom skupu i zaustavi treniranje ako se ne primijeti poboljšanje u određenom broju epoha. Kod prevelikog broja epoha postoji rizik da bi model mogao početi učiti napamet i izgubiti sposobnost generalizacije na nove i neviđene podatke pa broj epoha treba balansirati. Pratila se vrijednost *validation loss* i *validation accuracy* tijekom epoha kako bi se vidjelo uči li model napamet ili bi trebalo produžiti njegovo učenje. *Validation loss* predstavlja razliku između stvarnih vrijednosti u validacijskom setu i predviđenih vrijednosti koje daje model. *Validation accuracy* je preciznost modela na validacijskom skupu. Brzina učenja je isto jedan od parametara koji utječe koliko brzo model uči iz podataka. Ona označava koliko veliki korak model poduzima tijekom ažuriranja svojih parametara kako bi smanjio pogrešku na trening skupu. Svaki put kada model prođe kroz podatke, izračuna pogrešku i zatim prilagodi svoje parametre kako bi smanjio tu pogrešku. Ako je brzina učenja prevelika model bi mogao preskočiti optimalne postavke parametara, a s druge strane ako je premalena model bi učio presporo i možda bi mu trebalo puno više epoha da dosegne željeni rezultat. U ovom modelu je odabrana brzina 0,01 gdje model radi relativno umjerene korake.

4.5. Statističke metode

Koristit će se standardni dijagnostički testovi za evaluaciju točnosti, osjetljivosti = $TP/(TP + FN)$, preciznosti = $TP/(TP + FP)$, specifičnosti = $TN/(TN+FP)$, vrijednost površine ispod grafa (AUC) UI modela u razlikovanju benignih od malignih tumora dojki. $F1-score = 2 \times (Precision \times Sensitivity) / (Precision + Sensitivity)$.

5. REZULTATI

Istraživanje je provedeno sa 257 ultrazvučnih slika od kojih je 126 slika benignih tumora, a 131 slika malignih lezija. Program sam nasumično odabire 204 slika iz 257 za trening skup, 26 slika za validacijski skup i 26 slika za testni skup.

5.1. Trening skup

Nakon 11 epoha rezultati treninga su:

```
Epoch 1/15
7/7 [=====] - 24s 3s/step - loss: 4.6774 - accuracy: 0.5000 - val_loss: 0.6725 - val_accuracy: 0.5769
Epoch 2/15
7/7 [=====] - 14s 2s/step - loss: 0.9018 - accuracy: 0.5735 - val_loss: 0.6553 - val_accuracy: 0.5769
Epoch 3/15
7/7 [=====] - 15s 2s/step - loss: 0.6947 - accuracy: 0.5833 - val_loss: 0.7779 - val_accuracy: 0.5385
Epoch 4/15
7/7 [=====] - 15s 2s/step - loss: 0.5766 - accuracy: 0.6618 - val_loss: 0.5837 - val_accuracy: 0.7308
Epoch 5/15
7/7 [=====] - 15s 2s/step - loss: 0.5082 - accuracy: 0.7598 - val_loss: 0.6470 - val_accuracy: 0.6923
Epoch 6/15
7/7 [=====] - 15s 2s/step - loss: 0.4969 - accuracy: 0.7598 - val_loss: 0.5778 - val_accuracy: 0.8077
Epoch 7/15
7/7 [=====] - 15s 2s/step - loss: 0.4530 - accuracy: 0.7892 - val_loss: 0.5421 - val_accuracy: 0.8462
Epoch 8/15
7/7 [=====] - 16s 3s/step - loss: 0.4635 - accuracy: 0.7500 - val_loss: 0.5111 - val_accuracy: 0.8846
Epoch 9/15
7/7 [=====] - 16s 2s/step - loss: 0.3761 - accuracy: 0.8431 - val_loss: 0.5877 - val_accuracy: 0.8077
Epoch 10/15
7/7 [=====] - 16s 2s/step - loss: 0.3536 - accuracy: 0.8431 - val_loss: 0.5401 - val_accuracy: 0.6923
Epoch 11/15
7/7 [=====] - 15s 2s/step - loss: 0.3587 - accuracy: 0.8284 - val_loss: 0.5596 - val_accuracy: 0.7692
```

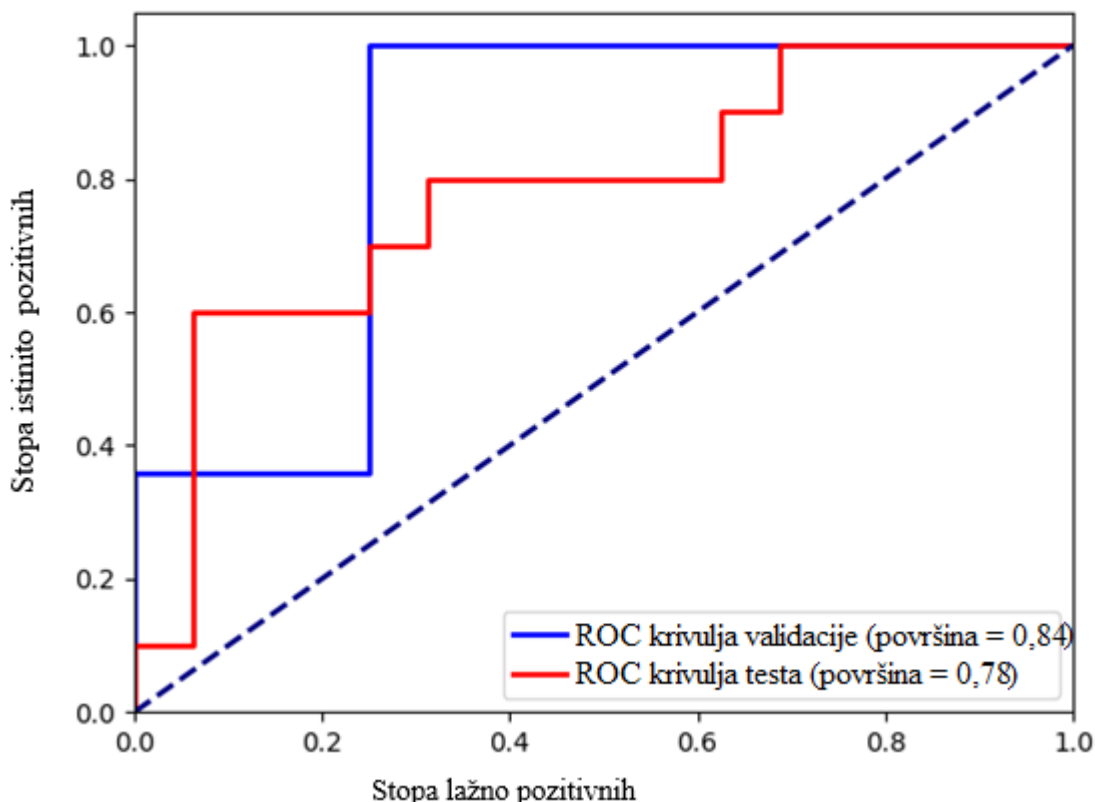
Slika 2. Trening UI modela

Trening je završen nakon 11 epoha. Točnost na trening skupu raste od 54,41 % do 93,14 %. *Validation loss* vrijednost konstantno pada i raste kroz sve epohe, a *validation accuracy* pada u zadnje 3 epohe. Rezultati validacijske točnosti mijenjaju se od 57,69 % do 88,46 % (Slika 2.).

5.2. Validacijski skup

Tablica 1. Matrica zabune validacijski skup

| | Benigno | Maligno |
|-----------------|---------|---------|
| Stvarno benigno | 9 | 3 |
| Stvarno maligno | 0 | 14 |



Slika 3. Krivulja operativnih karakteristika prijemnika (ROC)

Vrijednost površine ispod ROC grafa (AUC) za validacijski skup iznosi 0,84 (Slika 3.).

U validacijskom skupu bilo je ukupno 26 slika, od kojih je 12 bilo benignih, a 14 malignih. Osjetljivost ovog modela za maligne tvorbe dojke je 1, isto je izrazito visok i *F1-score* za maligne lezije sa 0,9. Točnost modela na validacijskom skupu iznosi 88 % (Tablica 2.).

Tablica 2. Tablica klasifikacije validacijski skup

| | Preciznost | Osjetljivost | F1- score | Broj slika |
|---------|------------|--------------|-----------|------------|
| Benigno | 1 | 0,75 | 0,86 | 12 |
| Maligno | 0,82 | 1 | 0,90 | 14 |
| Točnost | | | 0,88 | 26 |

5.3. Testni skup

U testnom skupu od 26 slika, 16 je slika benignih lezija, a 10 malignih. AUC testnog skupa je 0,78 (Slika 3.). Testirana osjetljivost ovog modela za maligne lezije je 80 %. *F1-score* je u ovom slučaju manji za maligne od benignih promjena. Točnost je pala na 73 %, a pale su i preciznost i osjetljivost za obje grupe lezija (Tablica 3.). Izračunata specifičnost za maligne lezije testnog skupa je 68,8 %.

Tablica 3. Tablica klasifikacija testni skup

| | Preciznost | Osjetljivost | F1- score | Broj slika |
|---------|------------|--------------|-----------|------------|
| Benigno | 0,85 | 0,69 | 0,76 | 16 |
| Maligno | 0,62 | 0,80 | 0,70 | 10 |
| Točnost | | | 0,73 | 26 |

U matrici zabune vidimo da je, za razliku od validacijskog skupa gdje nijednu malignu promjenu nije svrstao kao benignu, ovdje 2 maligne promjene klasificirao kao benigne (Tablica 1., Tablica 4.).

Tablica 4. Matrica zabune testni skup

| | Benigno | Maligno |
|-----------------|---------|---------|
| Stvarno benigno | 11 | 5 |
| Stvarno maligno | 2 | 8 |

6. RASPRAVA

U našem istraživanju bio je podjednak broj malignih i benignih tumora sa 126 benigna i 131 malignim tumorom, što je dobar preduvjet za kvalitetno treniranje modela zbog balansiranja traženih grupa. U ostalim istraživanjima raspored grupa znao je varirati pa u jednom istraživanju raspored je bio ujednačen sa 93 maligne i 102 benigne lezije (7) dok neka druga istraživanja imaju velike razlike između grupa kao npr. 575 benignih masa i 326 malignih promjena (8). To znači da se program može istrenirati i na nebalansiranim grupa, ali potrebno je obratiti pažnju na ostale hiperparametre treniranja i evaluacije. Jedna od čestih zabluda potreba za velikim brojem slika kako bi se pravilno i što bolje istrenirao UI model. Veći broj slika je prednost u treningu i provjeri učinkovitosti, ali iz našeg istraživanja vidimo da se i sa 200 slika može napraviti model i istrenirati da pokazuje dobre rezultate. Kvaliteta slika korištenih za treniranje UI modela igra ključnu ulogu u njegovoj preciznosti i sposobnosti razlikovanja različitih karakteristika. Model treniran na slikama visoke kvalitete, koje su jasno definirane i pravilno označene, ima veću šansu za postizanje boljih rezultata. Niske rezolucije, šum ili loša kontrastnost mogu otežati modelu prepoznavanje relevantnih obilježja, što direktno utječe na njegovu učinkovitost i pouzdanost u stvarnim scenarijima. U ovom istraživanju korištena je ultrazvučna slika u B-modu sa slikama različitih rezolucija i kontrasta zbog takve dostupnosti slika na odjelu. Postoji istraživanje koje je sve pacijente snimalo u istom položaju i isključivala nejasne slike UZV dojke kako bi osigurali dobar trening i učinkovitost UI modela (9). Istraživanja su isto koristila različite pristupe od običnih ultrazvučnih slika dojke do dodatnih mogućnosti ultrazvuka npr. *power Doppler* ultrazvuk snimke ili ultrazvuk s kontrastom za poboljšanje klasifikacije malignih promjena dojke (6, 9).

Trening skup imao je 126 slika koje je trebao provući kroz svoje čvorove u 15 epoha, ali zbog dodanog *early stopping* mehanizma stao je nakon 11. *Early stopping* mehanizam prati vrijednosti *validation loss* i *validation accuracy* i prati smanjuju li se one ili povećavaju te ima dozvoljene 3 epohe strpljenja. U slučaju povećanja *validation loss*-a ili smanjenja *validation accuracy* treba se zaustaviti treniranje. To se vidi na rezultatima gdje *validation loss* kroz svaku epohu varira tako da se vrijednost ili povećava ili smanjuje, ali *validation accuracy* tek krene padati nakon 8 epohe. Nakon 8 epohe imamo 3 epohe strpljenja i prekida se treniranje jer *validation loss* i dalje pada. Na validacijskom skupu vidimo da model odlično funkcionira sa poznatim podacima i dobro prepoznaje benigne tumore. To se vidi i po preciznosti za benigne tumore koja iznosi 100 %, a za maligne tumore je manje specifičan sa 82 % zbog svrstavanja 3 maligna tumora u benigne. Osjetljivost na maligne tumore je 100 % što znači da je svih 14 slika

malignih tumora tako i klasificirao, ali kod benignih je ta osjetljivost manja zbog 3 maligna tumora klasificirana kao benigna. *F1-score* u validacijskom skupu za benignu klasu iznosi 0.86 što pokazuje da model dobro balansira između preciznosti i osjetljivosti, ali postoji prostora za napredak. Za malignu klasu *F1-score* je još i bolji s 0,9 (idealno 1). $AUC = 0,84$ pokazuje da model ima veliku sposobnost razlikovanja između benignih i malignih tumora. Model dobro razlikuje između klasa i ima malu stopu pogreške prilikom predviđanja. U provjeri sa novim podacima iz testnog skupa došlo je do znatnog pada u učinkovitosti modela. AUC vrijednost testnog skupa (0,78) je dobra, ali ne izvrsna izvedba. Model ima dobru sposobnost razlikovanja između benignih i malignih tumora, iako je nešto niža u odnosu na AUC postignutu na validacijskom skupu. Ova razlika može ukazivati na varijacije između validacijskog i testnog skupa, što je normalno u praksi. Dobivena AUC vrijednost testnog skupa je konkurentna drugim istraživanjima, gdje kod korištenja multimodalnih ultrazvuka AUC je bio 0,85 ili kod korištenja 3D *power Doppler* ultrazvuka najmanja vrijednost je bila 0,782. U jednom istraživanju su koristili multimodalni UZV u poboljšanju točnosti procjene sentinelnih limfnih čvorova u slučajevima raka dojke. U tom istraživanju za dvodimenzionalni (2D) UZV, imao je vrijednost AUC 0,678 što je znatno manje od dobivene AUC u ovome istraživanju, ali bitno je naglasiti da UI modeli nisu klasificirali isto (10, 11). Točnost ovog modela na testnom skupu bila je 73 % što je vjerojatno zbog različitih kvaliteta slika. Točnost u istraživanjima znatno varira od nekih vrijednosti kao 70,5 % na standardnom ultrazvuku pa sve do točnosti od preko 90 % korištenjem multimodalnih ultrazvuka (9, 10, 12) i neki čak i 96 % korištenjem modulacije slike prije prolaska kroz model (13). Model iz ovog rada ima dosta prostora za napredak, ali iako 73 % ne izgleda kao velika točnost zapravo vidimo da postoje još uvijek modeli koji su na toj razini. Primjenom dodatnih tehnika ultrazvuku ili dodatnim procesiranjem slika točnost ovog modela bi sigurno narasla. Preciznost u ovom modelu bila je 62 %. U istraživanju preciznost na analizi 3D ultrazvučne slike iznosila je 84,7 % dok na običnoj ultrazvučnoj slici 75,9 % što govori da UI modeli još uvijek nisu dovoljno precizni u određivanju malignih lezija i to je nešto na čemu se treba raditi (9). Ovaj model pokazuje znatno nižu specifičnost (68,8 %) od modela koji su analizirali slike multimodalnih UZV gdje je specifičnost bila 92,4 %, nešto nižu specifičnost od 3D UZV analize (75,6 %), ali ima bolju specifičnost od modela koji je analizirao 2D UZV slike (64,9 %) (9, 10, 12). U usporedbi s modelom koji je analizirao iste slike ovaj model bolje prepoznaje benigne slučajeve. U slučaju osjetljivosti ovaj model je dosta bolji (80 %) i čak pokazuje bolju osjetljivost, nego model na osnovu multimodalnih UZV slika (77,4 %) (10). Model na osnovi 3D UZV slike pokazuje najveću osjetljivost 90,4 %, a model iz istraživanja koji je na temelju UZV slike ima nešto višu osjetljivost od 81,7 % (9, 12). Model

iz ovog istraživanja konkurentan je u usporedbi s drugim modelima, ali model na 3D UZV slikama pokazuje znatno bolju osjetljivost što znači da postoji dosta prostor za napredovanje. *F1-score* (0,70) je dosta niži u usporedbi s drugim istraživanjima, gdje čak istraživanje s nebalansiranim grupama ima *F1-score* od 0,87, a istraživanje sa približnom istom razlikom u broju malignih i benignih tumora ima *F1-score* 0,97 (8, 13).

7. ZAKLJUČAK

Na temelju provedenog istraživanja i dobivenih rezultata mogu se izvesti sljedeći zaključak:

- Model je učinkovit u razlikovanju benignih od malignih tumora
- Model ne može potpuno isključiti invazivne zahvate iz dijagnostike tumora dojki

8. SAŽETAK

CILJEVI ISTRAŽIVANJA: Ispitati učinkovitost neuronske mreže DenseNet u analizi ultrazvučnih slika tumora dojki.

NACRT STUDIJE: Retrospektivna studija s analizom sekundarnih podataka.

ISPITANICI I METODE: Istraživanje se provodi na Odjelu za radiologiju Nacionalne memorijalne bolnice „Dr. Juraj Njavro“ Vukovar. U istraživanju se koristi 257 ultrazvučnih slika pacijentica s tvorbama u dojkama koje su prošle ultrazvučni pregled i patohistološku dijagnostiku iglenom biopsijom u razdoblju od studenog 2021. do veljače 2024. Prikupljene slike su iz bolničke dokumentacije i prilagođene za korištenje u UI modelu. Implementira se neuronska mreža *DenseNet* za analizu ultrazvučnih slika. Rezultati modela uspoređeni su s dijagnozama dobivenim patohistološkom analizom iglene biopsije. Korišteni su standardni dijagnostički testovi za evaluaciju.

REZULTATI: Istraživanje je provedeno na ultrazvučnim slikama 126 benignih i 131 maligna tumora, što je omogućilo kvalitetno treniranje modela. Model je pokazao točnost od 73 %, AUC od 0,78, te specifičnost od 68,8% i osjetljivost od 80%, *F1-score* od 70 %.

ZAKLJUČAK: Istraživanje je pokazalo da model za klasifikaciju tumora dojke ima dobru sposobnost u razlikovanju malignih od benignih tumora te može smanjiti broj nepotrebnih invazivnih zahvata u dijagnostici tumora dojki.

KLJUČNE RIJEČI: tumor dojke; ultrazvuk dojke; umjetna inteligencija

9. SUMMARY

The Efficiency of Artificial Intelligence in Distinguishing Malignant from Benign Ultrasound Features of Breast Tumors.

OBJECTIVES: Assess the effectiveness of the DenseNet neural network in analyzing ultrasound images of breast tumors.

STUDY DESIGN: Retrospective study with analysis of secondary data.

PARTICIPANTS AND METHODS: The study is conducted at the Radiology Department of the National Memorial Hospital "Dr. Juraj Njavro" in Vukovar. It uses 257 ultrasound images of patients with breast masses who underwent ultrasound examination and pathohistological diagnosis through needle biopsy from November 2021 to February 2024. The images were collected from hospital records and adapted for use in the artificial intelligence model. The DenseNet neural network is implemented for ultrasound image analysis. The model's results will be compared with diagnoses obtained from pathohistological needle biopsy analysis. Standard diagnostic tests will be used for evaluation.

RESULTS: The study was conducted on ultrasound images of 126 benign and 131 malignant tumors, which enabled effective model training. The model demonstrated an accuracy of 73 %, an AUC of 0.78, specificity of 68.8 %, sensitivity of 80 %, and an F1-score of 70 %.

CONCLUSION: Research has shown that the breast tumor classification model has a good ability to distinguish between malignant and benign tumors, and there is significant room for improvement to reduce unnecessary invasive procedures.

KEYWORDS: artificial Intelligence; breast tumor; breast ultrasound

10. LITERATURA

1. Ivan Damjanov , Sven Seiwerth , Stanko Jukić , Marin Nola, Patologija, 5 izd. Zagreb: Medicinska naklada; 2018; 639, 645 – 652.
2. Kirby I. Bland; Edward M. Copeland, III; William J. Gradishar; V. Suzanne Klimberg, The Breast: Comprehensive Management of Benign and Malignant Diseases, 5. izd. Philadelphia: Elsevier; 2017; 87 – 92.
3. Guo R, Lu G, Qin B, Fei B. Ultrasound Imaging Technologies for Breast Cancer Detection and Management: A Review. Ultrasound Med Biol. 2018 Jan
4. Staren ED, O'Neill TP. Breast ultrasound. Surg Clin North Am. 1998 Apr..
5. Roger P. Smith MD, Netter's Obstetrics and Gynecology, 4. izd. Philadelphia: Elsevier, 2019; 382 – 396.
6. G. Huang, Z. Liu, L. Van Der Maaten and K. Q. Weinberger, "Densely Connected Convolutional Networks," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017.
7. Xu Z, Wang Y, Chen M, Zhang Q. Multi-region radiomics for artificially intelligent diagnosis of breast cancer using multimodal ultrasound. Comput Biol Med. 2022 Oct.
8. Wei Q, Yan YJ, Wu GG, Ye XR, Jiang F, Liu J, i sur. The diagnostic performance of ultrasound computer-UIded diagnosis system for distinguishing breast masses: a prospective multicenter study. Eur Radiol. 2022 Jun.
9. Xiao X, Gan F, Yu H. Tomographic Ultrasound Imaging in the Diagnosis of Breast Tumors under the Guidance of Deep Learning Algorithms. Comput Intell Neurosci. 2022 Feb 28.
10. Xu Z, Wang Y, Chen M, Zhang Q. Multi-region radiomics for artificially intelligent diagnosis of breast cancer using multimodal ultrasound. Comput Biol Med. 2022 Oct.
11. Su S, Ye J, Ke H, Zhong H, Lyu G, Xu Z. Multimodal ultrasound imaging: a method to improve the accuracy of sentinel lymph node diagnosis in breast cancer. Front Oncol. 2024 Mar 25.

12. Ma H, Tian R, Li H, Sun H, Lu G, Liu R, i sur. Fus2Net: a novel Convolutional Neural Network for classification of benign and malignant breast tumor in ultrasound images. Biomed Eng Online. 2021 Nov 18.
13. Chen H, Ma M, Liu G, Wang Y, Jin Z, Liu C. Breast Tumor Classification in Ultrasound Images by Fusion of Deep Convolutional Neural Network and Shallow LBP Feature. J Digit Imaging. 2023 Jun.