

Detekcija ozljeda abdomena iz CT slika

Ivana Krišto

Fakultet elektrotehnike i računarstva
Zagreb, Hrvatska
ivana.kristo@fer.hr

Majda Bakmaz

Fakultet elektrotehnike i računarstva
Zagreb, Hrvatska
majda.bakmaz@fer.hr

Alen Vodopija

Fakultet elektrotehnike i računarstva
Zagreb, Hrvatska
alen.vodopija@fer.hr

Zrinka Pećanić

Fakultet elektrotehnike i računarstva
Zagreb, Hrvatska
zrinka.pecanic@fer.hr

Bruno Maršić

Fakultet elektrotehnike i računarstva
Zagreb, Hrvatska
bruno.marsic@fer.hr

Sažetak — Ovaj rad obrađuje problematiku detekcije ozljeda abdomena iz CT slika, s fokusom na unutarnje krvarenje i ozljede sljedećih organa: crijeva (debelo i tanko), bubrezi, jetra, slezena. Pravovremena detekcija ozljede te njezina klasifikacija s obzirom na težinu iste ključ je za uspješno liječenje i (potencijalno) spašavanje života. Stoga, korištenje umjetne inteligencije može biti korisno za brži i kvalitetniji rad zdravstvenih djelatnika u tom području. U radu se istražuje korištenje modela konvolucijskih neuronskih mreža za detekciju ozljeda.

Ključne riječi—detekcija, abdomen, ozljeda, neuronske mreže, konvolucija, klasifikacija, umjetna inteligencija, analiza slike, CT snimanje, strojno učenje, medicinska dijagnostika, trauma

I. UVOD

Traumatske ozljede najčešći su uzrok smrti u prvih 40 godina života čovjeka i veliki zdravstveni problem diljem svijeta budući da uzrokuju više od 5 milijuna smrtnih slučajeva godišnje [1]. Brza i točna dijagnoza traumatskih ozljeda ključna je za pokretanje pravodobnih intervencija kako bi se osigurala što veća stopa preživljenja i kvalitetniji oporavak pacijenta. Računalna tomografija (CT) je zbog svoje mogućnosti snimanja detaljnih presjeka abdomena postala ključan (i za sada, nezamjenjiv) alat u dijagnozi pacijenata sa sumnjom na abdominalne ozljede. Međutim, tumačenje CT snimki može biti kompleksno i vremenski zahtjevno, pogotovo u slučajevima kada se radi o ozljedama više unutarnjih organa, uključujući jetru, slezenu, bubrege, debelo i tanko crijevo. Analizom CT slika korištenjem umjetne inteligencije i principa strojnog učenja može se smanjiti teret i ubrzati posao zdravstvenih djelatnika. Ovaj rad fokusira se na detekciju sljedećih vrsta ozljeda: unutarnje krvarenje, ozljeda crijeva, bubrega, jetre i/ili slezene uz primjenu modela neuronskih mreža.

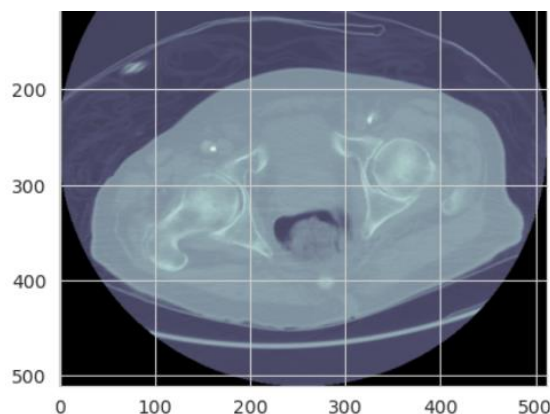
II. PREGLED POSTOJEĆIH RJEŠENJA

Osnova rješenja bilo je natjecanje *RSNA 2023 Abdominal Trauma Detection AI Challenge* koje je organizirala RSNA (Radiological Society of North America) u suradnji s ASER-om (American Society of Emergency Radiology) i SAR-om (Society for Abdominal Radiology).

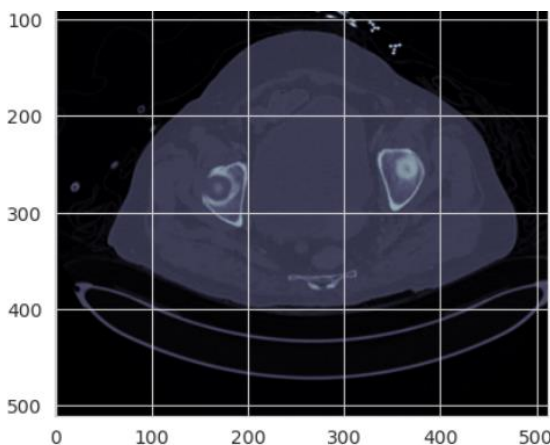
Slike su dostupne u DICOM formatu, što je često korišteni format medicinskih slikovnih podataka. Metapodaci se pružaju izvan DICOM formata, u CSV datotekama, kao i oznake skupa za treniranje. Podaci se sastoje od 4711 snimaka od 3147 pacijenata (svaki pacijent je sniman najmanje jednom, a najviše dva puta). Značajke u skupu za treniranje su ID pacijenta te sljedeće oznake:

- [bowel/extravasation]_[healthy/injury] – tipovi ozljeda s binarnim ciljnim oznakama (bez ozljede/ozljeda)
- [kidney/liver/spleen]_[healthy/low/high] – tipovi ozljeda s 3 ciljne oznake (bez ozljede/lakša ozljeda/teža ozljeda)
- any_injury – označava prisutnost bilo kakve ozljede (0 = bez ozljede, 1 = ozljeda)

Zadatak natjecanja bio je napraviti i istrenirati model na tim slikama te ga primijeniti na grupu slika za testiranje. Mnogo je timova sudjelovalo, i nisu sva rješenja javno dostupna, no među najboljima su korištene neuronske mreže, uglavnom CNN i RNN koje su autori posebno napravili za sudjelovanje u ovom natjecanju [2].



Slika II.1 CT snimka abdomena bez ozljede



Slika II.2 CT snimka abdomena s unutarnjim krvarenjem

III. OPIS RJEŠENJA PROJEKTOG TIMA

Kao osnova našeg rješenja poslužila je jedna od bilježnica na natjecanju koja se sastojala od koda za pretprocesiranje slika i treniranje te nije bila evaluirana od strane samog natjecanja budući da je bila nedovršena. Dodavanjem koda za testiranje modela te podešavanjem hiperparametara dobiveno je funkcionalno rješenje za dani problem. Za vizualizaciju ulaznih podataka i rezultata korišteni su Python i Tableau.

Budući da je javno dostupan skup podataka za testiranje bio premalen da bi testiranje na istom imalo smisla (svega 3 primjera), početni skup podataka podijeljen je na skup za treniranje i skup za testiranje u omjeru 4:1, te je skup za treniranje podijeljen na skup za treniranje i validaciju u istom tom omjeru. Slike u DICOM formatu se pretprocesiraju tako da s pretvaraju u crno-bijele i zatim skaliraju dijeljenjem s 255 u niz u kojem svaka vrijednost predstavlja vrijednost piksela slike u rasponu [0, 1] gdje veća vrijednost znači svjetliji piksel. Niz se zatim transformira (budući da imamo mnogo sličnih slika) transformacijama kao što su *RandomHorizontalFlip()*, *ColorJitter()*, *RandomAffine()* i *RandomErasing()*. Takav se niz zatim pretvara u tenzor da bi se s njim moglo raditi dalje u neuronskoj mreži.

Model koji koristimo je vrsta konvolucijske neuronske mreže, i to već istrenirani model EfficientNet-B0, točnije verziju IMAGENET1K_V1.

```
class CNNModel(nn.Module):
    def __init__(self):
        super().__init__()

        self.input = nn.Conv2d(4, 3, kernel_size = 3)
        model = models.efficientnet_b0(weights = 'IMAGENET1K_V1')

        self.features = model.features
        self.avgpool = model.avgpool

        self.bowel = nn.Linear(1280, 1)
        self.extravasation = nn.Linear(1280, 1)
        self.kidney = nn.Linear(1280, 3)
        self.liver = nn.Linear(1280, 3)
        self.spleen = nn.Linear(1280, 3)
```

Slika III.1 Arhitektura korištenog modela

Hiperparametri modela su: learning rate = 0.0004 uz Adam optimizator, batch size = 16, broj epoha = 12. Za gubitak se koristi gubitak unakrsne entropije. Uspješnost modela mjeri se metrikama točnosti (accuracy) i AUC (Area under the ROC Curve) te gubitkom. Na kraju svake epohe radi se validacija, te ako model ima gubitak koji je manji od zadnjeg najmanjeg gubitka, njegove težine se spremaju da bi se mogle dalje koristiti pri testiranju.

Za crijeva (*bowel*) i unutarnje krvarenje (*extravasation*) predviđa se jedna izlazna vrijednost (za svaku) u obliku vjerojatnosti da postoji ozljeda u rasponu [0, 1] gdje 0 znači da nema ozljede, a 1 da postoji ozljeda. Za svaki od ostalih unutarnjih organa (bubrezi (*kidney*), jetra (*liver*), slezena (*spleen*)) predviđaju se 3 izlazne vrijednosti u rasponu [0, 1]:

- [a, b, c]
 - a = vjerojatnost da organ nema ozljede
 - b = vjerojatnost da je organ lakše ozlijeđen
 - c = vjerojatnost da je organ teže ozlijeđen

Konačna klasifikacija između 3 vrijednosti jest ona čija je vjerojatnost najveća. U nastavku je dan primjer izlaznog skupa podataka za neku CT snimku:

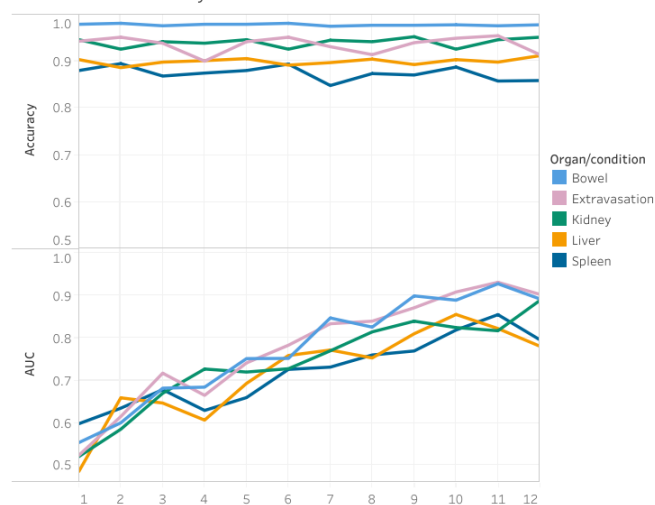
- *bowel* [0.01] - nema ozljede na crijevima
- *extravasation* [0.8] – postoji unutarnje krvarenje
- *kidney* [0.1, 0.8, 0.1] – bubreg ima lakšu ozljedu
- *liver* [0.9, 0.1, 0.001] – jetra nema ozljedu
- *spleen* [0.2, 0.1, 0.7] – slezena ima težu ozljedu.

IV. EKSPERIMENTALNI REZULTATI

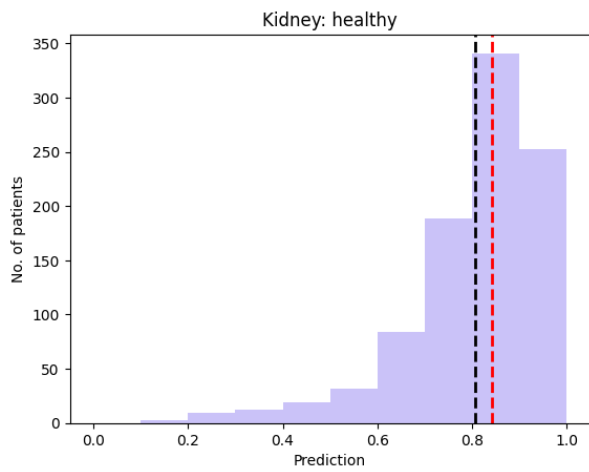
Podaci koje smo koristili nisu balansirani – samo približno četvrtina ulaznih podataka sadrži neki oblik ozljede. Rezultate uspoređujemo prema odgovarajućim kriterijima – prikazujući podatke u tablici predictions.csv, koju, zbog količine podataka, vizualiziramo histogramom, i ROC_AUC vrijednostima koje smo koristili za ocjenu kvalitete modela. Točnost (*accuracy*) nije primjerena u ovom kontekstu zbog neuravnoteženosti u skupu podataka za učenje, kao i zbog neprikladnosti za prirodu problema. Na primjer, ako imamo 100 pacijenata, od kojih 99 nema ozljede, a jedan ima, te naš model predviđa da su svi pacijenti bez ozljeda, tada bismo imali točnost predikcije od 99% (vrlo visoka i 'dobra'), no jedan pacijent bi imao nedetektiranu ozljedu abdomena, što može imati ozbiljne posljedice po njegovo zdravlje, ali i život. Stoga koristimo ROC i AUC kao mjere kvalitete predikcija. ROC krivulja (*Receiver Operating Characteristic curve*) je graf koji prikazuje dva parametra: frekvenciju točno klasificiranih slučajeva i frekvenciju lažno pozitivnih slučajeva. AUC (*Area Under the ROC Curve*) označava površinu ispod te krivulje i varira od 0 do 1, gdje su niže vrijednosti indikativne za lošu, a više za dobru klasifikaciju. Ovakav manje 'polarni' pristup bolje odgovara prirodi problema kojim se bavimo.

Slijede prikazi rezultata za svaki tip ozljede i metrike izračunate tijekom 12 epoha. Crvena linija na histogramima označava medijan, a crna srednju vrijednost.

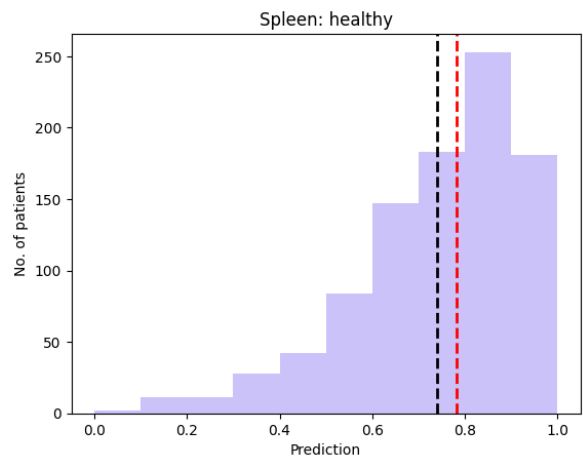
Validation accuracy and AUC



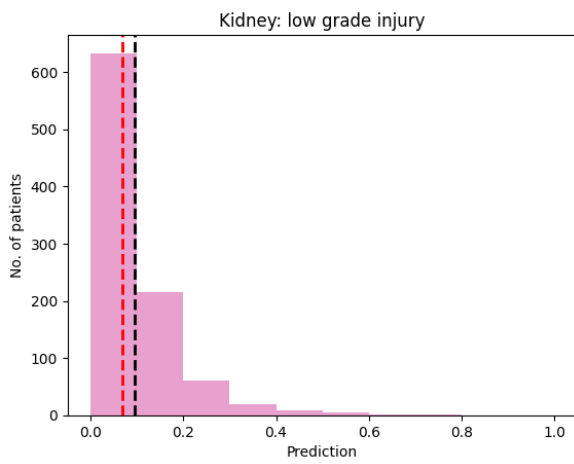
Slika IV.1 Kretanje vrijednosti točnosti i AUC na validacijskom skupu podataka tijekom 12 epoha



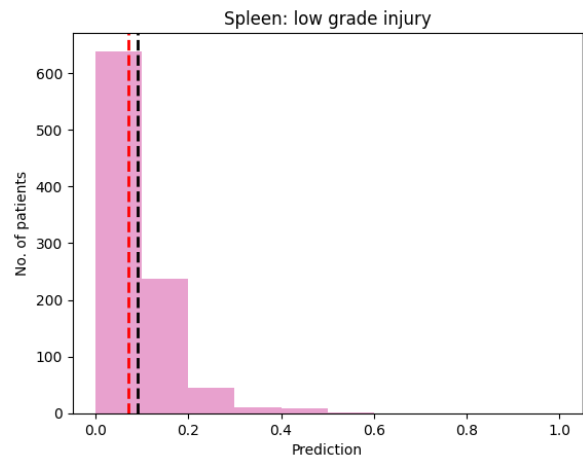
Slika IV.2 Predikcije: bubrezi bez ozljeda



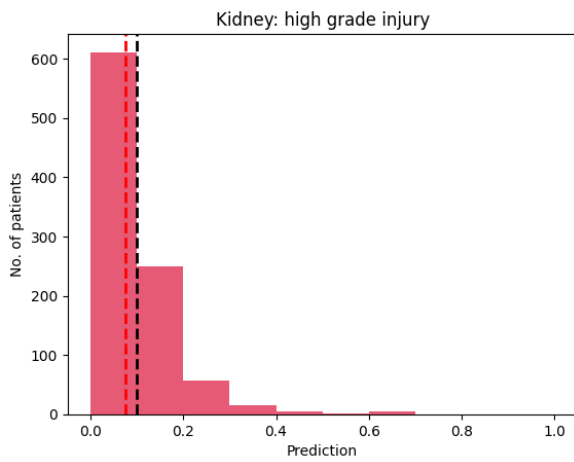
Slika IV.5 Predikcije: slezena bez ozljeda



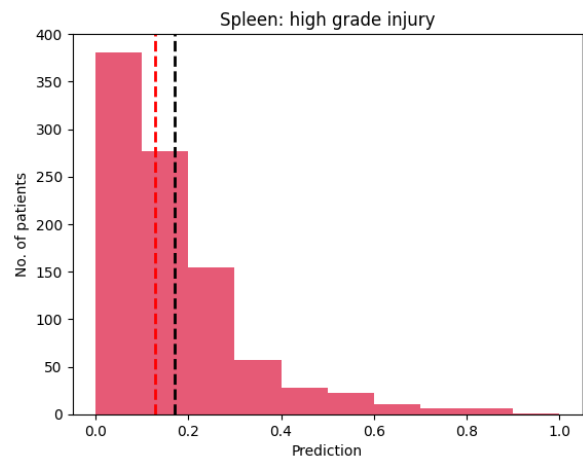
Slika IV.3 Predikcije: bubrezi s lakšom ozljedom



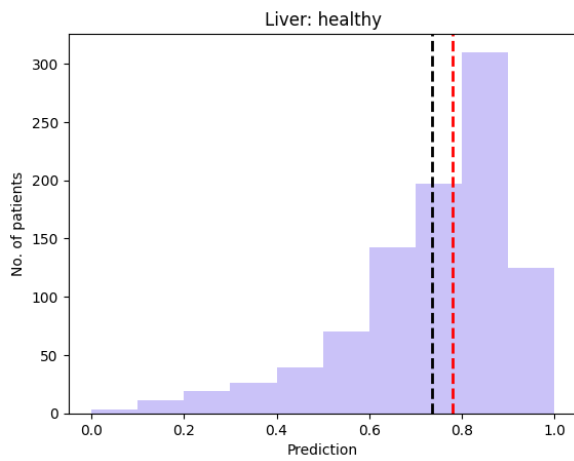
Slika IV.6 Predikcije: slezena s lakšom ozljedom



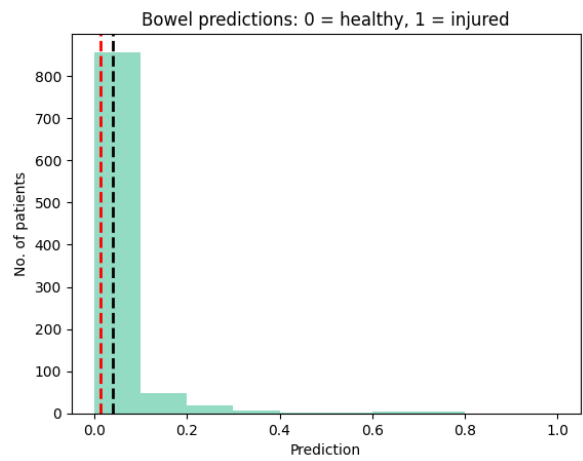
Slika IV.4 Predikcije: bubrezi s težom ozljedom



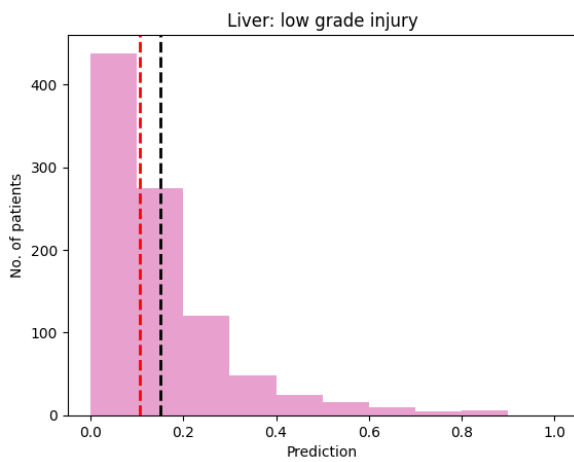
Slika IV.7 Predikcije: slezena s težom ozljedom



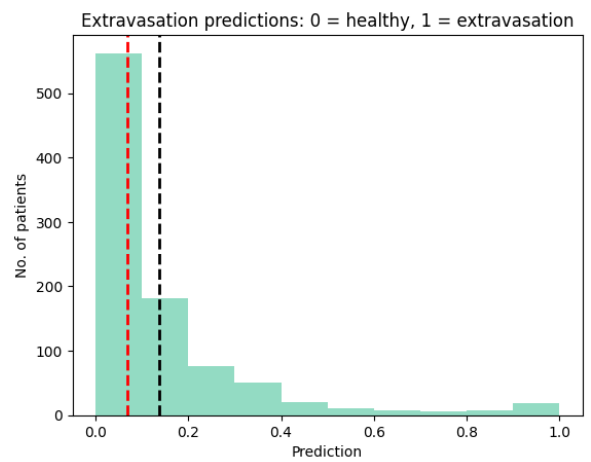
Slika IV.8 Predikcije: jetra bez ozljeda



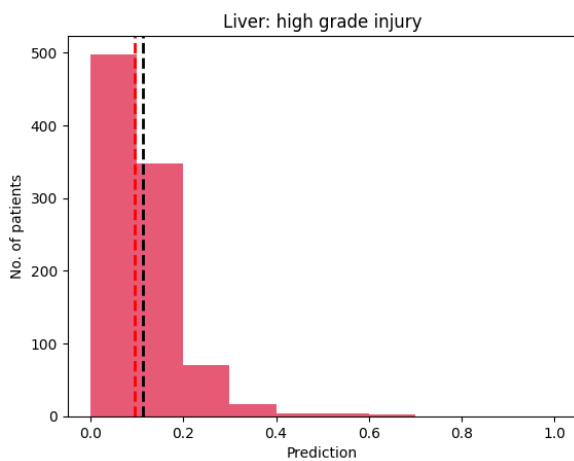
Slika IV.11 Predikcije: vjerojatnost ozljede crijeva



Slika IV.9 Predikcije: jetra s lakšom ozljedom

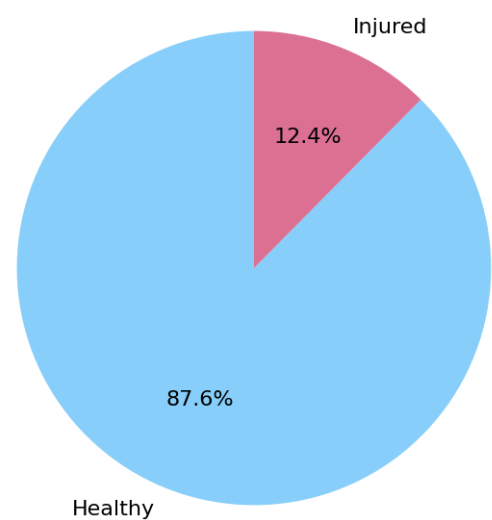


Slika IV.12 Predikcije: vjerojatnost prisutnosti unutarnjeg krvarenja



Slika IV.10 Predikcije: jetra s težom ozljedom

Abdominal injury detection (any injury)



Slika IV.13 Učestalost bilo kakve ozljede na testnom skupu

Injury type	accuracy	AUC
<i>Bowel</i>	0.975584	0.927473
<i>Extravasation</i>	0.954352	0.930937
<i>Liver</i>	0.898089	0.821857
<i>Kidney</i>	0.945860	0.816676
<i>Spleen</i>	0.857749	0.854666

Tablica IV.1 Prikaz najboljih vrijednosti dobivenih za točnost i AUC na validacijskom skupu

V. DISKUSIJA REZULTATA

Najbolje težine modela (najbolji AUC i najmanji gubitak) postiže se u 11. epohi. Gledajući samo metrike, očekuje se da će model najbolje raditi za detekciju unutarnjeg krvarenja i ozljede crijeva (što se može pretpostaviti i iz činjenice da za te ozljede radimo s jednostavnijim binarnim oznakama), dok će za jetru, bubrege i slezenu predikcije biti nešto manje kvalitetne (još uvijek zadovoljavajuće).

Idealno, na histogramima bi predikcije trebale biti nakupljene oko vrijednosti 0 i 1, budući da radimo s vrijednostima koje tumačimo kao vjerojatnost te stoga želimo da model daje što 'sigurnije' predikcije, primjerice – vrijednost 0.9 (kod crijeva i unutarnjeg krvarenja) tumačimo kao vrlo visoku vjerojatnost da se na slici nalazi ozljeda, no to ne možemo tvrditi s istom sigurnošću za vrijednost 0.55.

Najkvalitetnije predikcije (gledajući AUC metriku) dobivamo za ozljede crijeva (0.927473) i unutarnje krvarenje (0.930937), a to podupiru i njihovi histogrami na kojima se vidi da su predikcije lokalizirane oko 0 (i mali dio oko 1 za unutarnje krvarenje). Unutarnje krvarenje ima nešto više 'nejasnih' vrijednosti (oko 0.5) što znači da je za neke slike bilo teže odrediti je li prisutno unutarnje krvarenje ili ne.

Među preostala tri organa najkvalitetnije predikcije dobivamo za slezenu, čiji histogrami to podupiru. Vidljivo je

da kod slezene, jetre, i bubrega ima mnogo više 'nejasnih' klasifikacija i to se odražava i u AUC vrijednostima koje su sve ispod 0.9 (no još uvijek zadovoljavajuće). U svakom slučaju, na histogramima je vidljiva jasna lokalizacija vrijednosti oko 0 i 1. Također se može primijetiti da na većini slika nije detektirana ozljeda što je u skladu s ulaznim podacima i maloj zastupljenosti snimka s ozljedom.

VI. ZAKLJUČAK

Traumatske ozljede predstavljaju značajan zdravstveni izazov, često rezultirajući smrtnim ishodom, s više od 5 milijuna godišnjih smrtnih slučajeva. Brza i precizna dijagnoza traumatskih ozljeda ključna je za pravovremene intervencije.

Ovaj rad se fokusira na primjenu umjetne inteligencije, posebno neuronskih mreža, u detekciji različitih vrsta ozljeda unutarnjih organa. Korištenjem modela EfficientNet-B0 postignuti su obećavajući rezultati, posebno za ozljede crijeva i unutarnje krvarenje.

Analiza rezultata ukazuje na najbolje performanse modela u 11. epohi, s naglaskom na detekciju unutarnjeg krvarenja i ozljeda crijeva. Histogrami ukazuju na koncentraciju vrijednosti oko 0 i 1, što odražava želju za 'sigurnijim' predikcijama. Unatoč zadovoljavajućim rezultatima, predikcije za jetru, bubrege i slezenu pokazuju veći stupanj nejasnoća, što se odražava i u nižim AUC vrijednostima.

Nastavak istraživanja i izrada potpuno novog modela (umjesto korištenja već treniranog modela) prilagođenog ovoj problematici po uzoru na pobjednike natjecanja mogli bi dalje unaprijediti kvalitetu predikcija u prepoznavanju različitih ozljeda unutarnjih organa.

LITERATURA

- [1] <https://www.kaggle.com/competitions/rsna-2023-abdominal-trauma-detection/overview>, natjecanje korišteno za izbor modela i podatke
- [2] <https://www.kaggle.com/competitions/rsna-2023-abdominal-trauma-detection/leaderboard>, ljestvica najboljih rješenja natjecanja