

Univerza v Ljubljani
Fakulteta *za matematiko in fiziko*



Binarna klasifikacija resnosti COVID-19 okužbe s pomočjo CT slik prsnega koša

Avtor: Simon Perovnik

Predavatelj: doc. dr. Simon Čopar

Asistent: Žan Klaneček

Peta naloga pri Praktikum strojnega učenja v fiziki

Ljubljana, maj 2024

1 Naloga

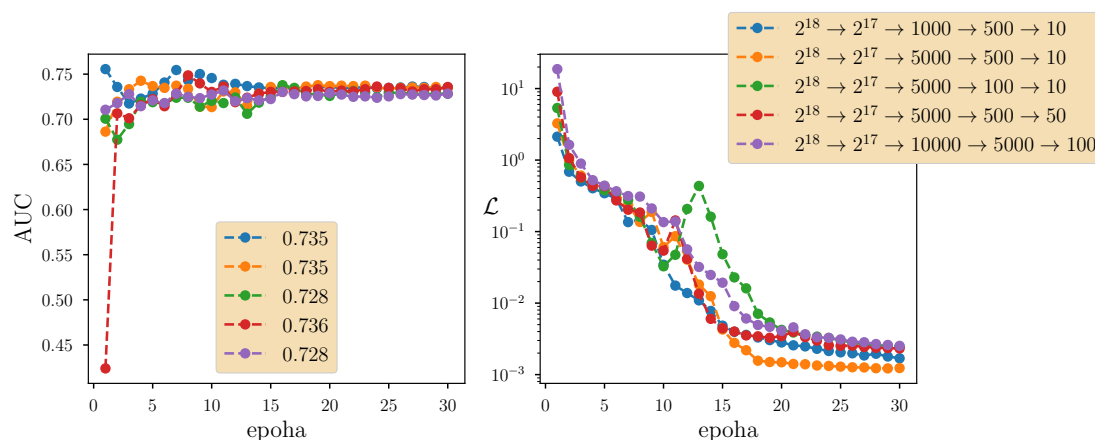
V peti nalogi obravnavamo CT slike prsnega koša pri obolelih s COVID-19. V podatkovnem setu imamo na voljo slike tako blago kot hudo bolnih ljudi, naša naloga pa bo uporabiti različne nevronske mreže za binarno klasifikacijo obeh razredov. V ta namen bomo osnovali gosto povezano mrežo ter konvolucijsko mrežo in uporabili splošno uveljavljeno mrežo za klasifikacijo slik imenovano **ResNET**. V prvem delu naloge bomo iskali hiperparametre mrež, s katerimi dosežemo najboljšo napoved, nato pa bomo vse modele med seboj primerjali in ovrednotili. V drugem delu naloge si bomo bolj natančno ogledali obe konvolucijski mreži - našo lastno in **ResNET**. Prikazali bomo kakšne filtre vsebujeta, kaj se dogaja z vhodno sliko na različnih plasteh mreže in kako lahko z mapo izpostavljenosti preverimo, kateri piksli najbolj prispevajo h končni napovedi.

2 Rezultati

2.1 Iskanje optimalnih hiperparametrov

2.1.1 Gosto povezana mreža

Za gosto povezano mrežo sicer vemo, da ni najbolj primerna za klasifikacijo slik, a lahko vseeno poiščemo kakšna arhitektura take mreže da najboljše rezultate. Vhodne podatke tako transformiramo v vektor. Spremljali bomo dve metriki - t. i. površino pod krivuljo (*Area Under the ROC Curve*, *AUC*) in funkcijo izgube. Preverimo nekaj različnih arhitektur, kjer pa je vsem skupno, da najprej začnemo s razpolovitvijo vhodnih vrednosti preko plasti **max_pooling**, nato pa si sledijo tri plasti nevronov, katere velikosti spreminjamo in se končajo v enem samem nevronu. Po pregledu različnih modelov za optimalne hiperparametre vzamemo $lr = 0.01$ in $mp = 0.9$. Rezultati se nahajajo na Sliki 1.



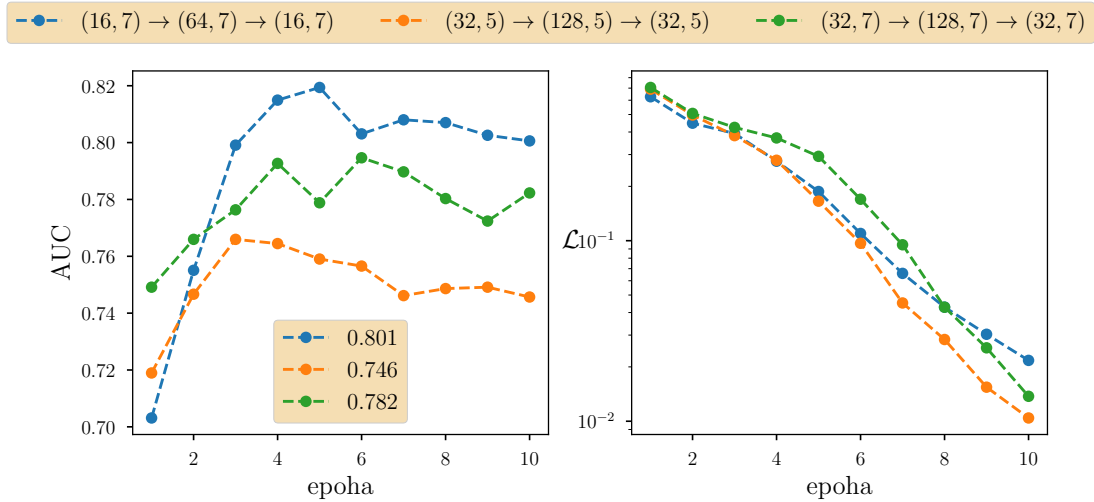
Slika 1: Sledenje AUC in funkciji izgube po epohah za različne arhitekture gosto povezane mreže. Legenda na levem grafu prikazuje končne vrednosti AUC.

Vsi modeli so približno enako uspešni, kar nakazujeta oba grafa s slike. Najvišjo AUC zabeležimo pri arhitekturi s 5000, 500 in 50 nevroni v skritih plasteh, ko le-ta znaša 0.736.

2.1.2 Konvolucijska mreža

Obetamo si, da bomo s konvolucijsko mrežo dosegli boljše rezultate. Z grid-search iskanjem določimo optimalne hiperparametre: $lr = 0.001$, $wd = 0.001$, $mp = 0.9$. Spet bomo preverili tri različne arhitekture. Vsaka ima tri s konvolucijske plasti, vsaki plasti pa sledijo še

BatchNorm2d, ReLU aktivacija in `max_pooling(x, (2, 2))`. Variiramo število filtrov in velikost konvolucijskega jedra. Z $(32, 7) \rightarrow (128, 7) \rightarrow (32, 7)$ označimo “osnoven” model z 32, 128 in 32 filtri in konstantno velikostjo konvolucijskega jedra (7×7) . Testiramo še dve arhitekturi z manjšim številom filtrov in manjšim konvolucijskim jedrom. Rezultati so prikazani na Sliki 2.

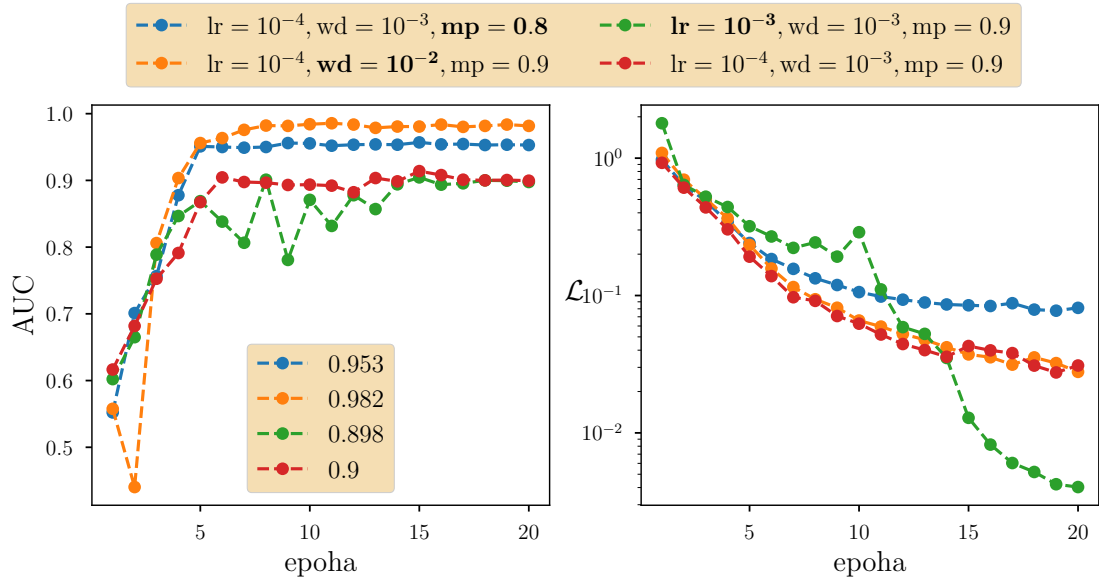


Slika 2: Sledenje AUC in funkciji izgube po epohah za različne arhitekture konvolucijske nevronske mreže. Legenda na levem grafu prikazuje končne vrednosti AUC.

Najvišji AUC dobimo pri mreži z manj filtri, ki ob koncu na validacijski množici znaša 0.8. Vidimo lahko, da se vse funkcije izgube tudi ob zadnji epohi še vedno znatno nižajo, zaradi česar bo verjetno smiselno izbrani konvolucijski model trenirat dlje časa.

2.1.3 ResNET

ResNET konvolucijska mreža je močan model za klasifikacijo slik. Ker je v osnovi narejen za “multi-case” klasifikacijo barvnih slik za naše potrebe spremenimo vhodno plast (spremenimo v `Conv2d(1, 64, k=(7, 7), s=(2, 2), p=(3,3))` in dodamo na koncu še `Conv2d(512, 1, k=(1, 1), s=(1, 1))` in `AdaptiveMaxPool2d((1, 1))`. S tem smo določili arhitekturo mreže. Dodatno lahko raziščemo kakšni so optimalni hiperparametri, kar je prikazano na Sliki 3, kjer smo variirali *learning rate* (*lr*), *weight decay* (*wd*) in multiplikator spreminjanja uteži (*mp*).



Slika 3: Sledenje AUC in funkciji izgube po epohah za različne hiperparametre ResNET mreže. Legenda na levem grafu prikazuje končne vrednosti AUC.

Vsi modeli dosegajo razmeroma dober končen AUC, najvišjega pa zabeležimo pri $lr = 10^{-4}, wd = 10^{-2}, mp = 0.9$, kjer znaša 0.98. V nadaljevanju bomo primerjali tudi model ResNET mreže, kjer ne bomo uporabili že natreniranih uteži.

2.2 Primerjava modelov

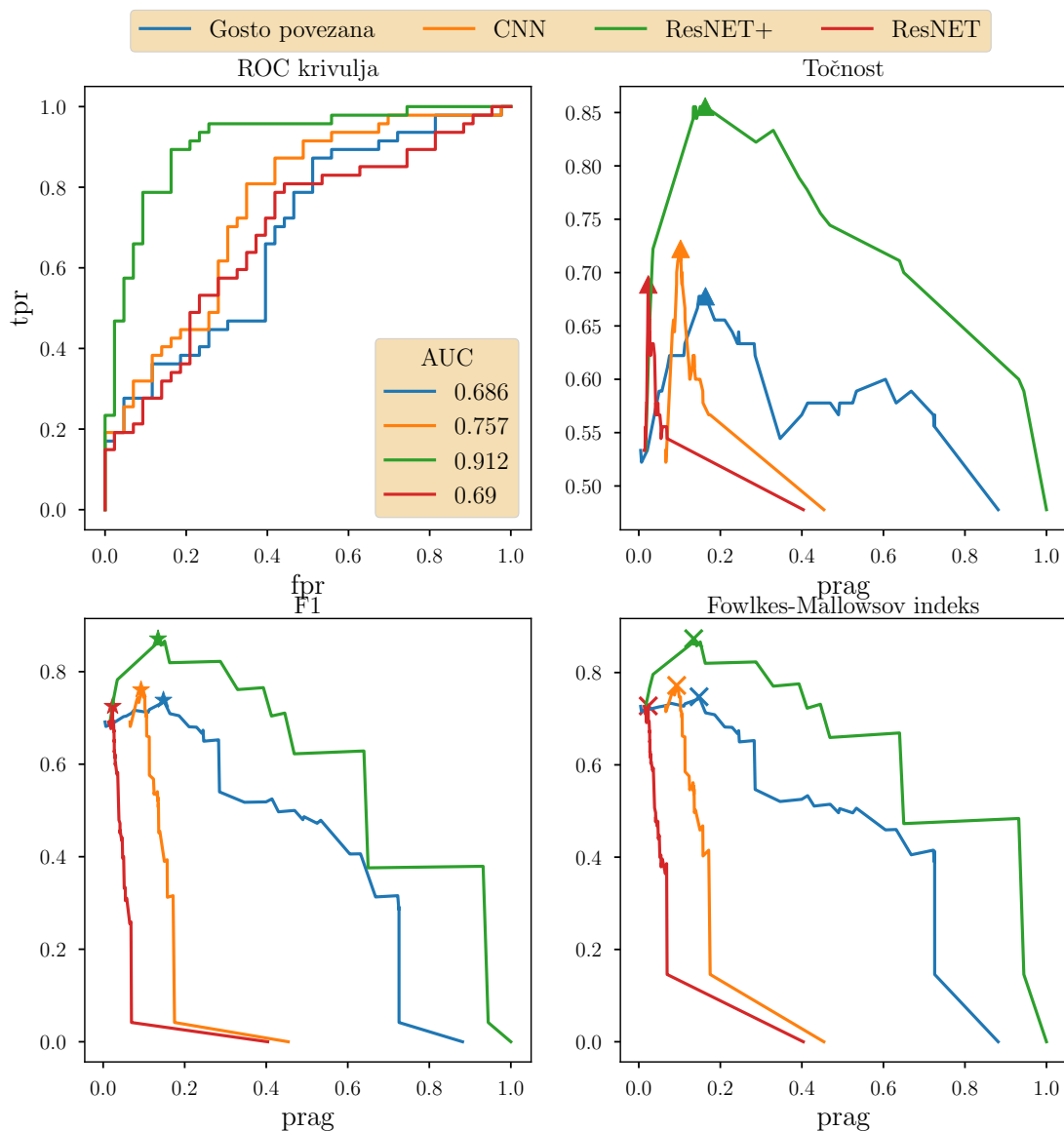
Vse predstavljene modele lahko med seboj primerjamo. Za primerjavo uporabimo najboljšega kandidata iz vsake skupine. To so:

- Gosto povezana mreža z arhitekturo $2^{18} \rightarrow 2^{17} \rightarrow 5000 \rightarrow 500 \rightarrow 50$,
- konvolucijska mreža z 32, 128 in 32 filtri ter konstantno velikostjo konvolucijskega jedra (7×7),
- prilagojena ResNET konvolucijska mreža z že natreniranimi utežmi (označimo z ResNET+ in
- prilagojena ResNET konvolucijska mreža brez inicializiranih uteži (označimo z ResNET.

V vseh primerih treniramo mreže 40 epoh s hiperparametri $lr = 10^{-4}, wd = 10^{-2}$ in $mp = 0.9$. Primerjamo 4 metrike modelov: AUC, točnost (delež pravih napovedi), F1 metriko in Fowlkes-Mallowsov indeks, slednja dva sta definirana z

$$F1 = 2 \frac{TPR \times PPV}{TPR + PPV}, \quad PPV = \frac{TP}{TP + FP}, \quad FM = \sqrt{TPR \times PPV}, \quad (1)$$

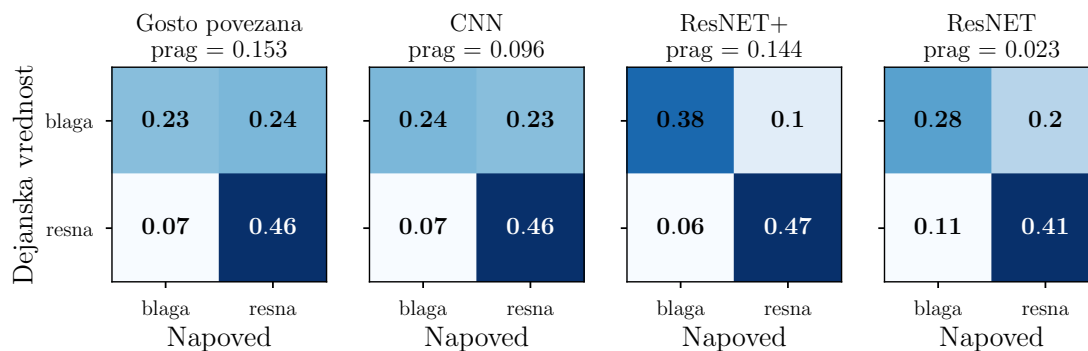
kjer TPR označuje občutljivost ($TP/(TP + FN)$). Uporabljamo oznake resnično negativen (TN), resnično pozitiven (TP), lažno negativen (FN) in lažno pozitiven (FP). Primerjave za določitev primerne praga so prikazane na Sliki 4.



Slika 4: Primerjava metrik AUC, točnosti, F1 metrike in Fowlkes-Mallowsov indeksa za 4 različne modele nevronske mreže. Maksimalne vrednosti, s katerimi določimo prag so prikazane z različnimi simboli.

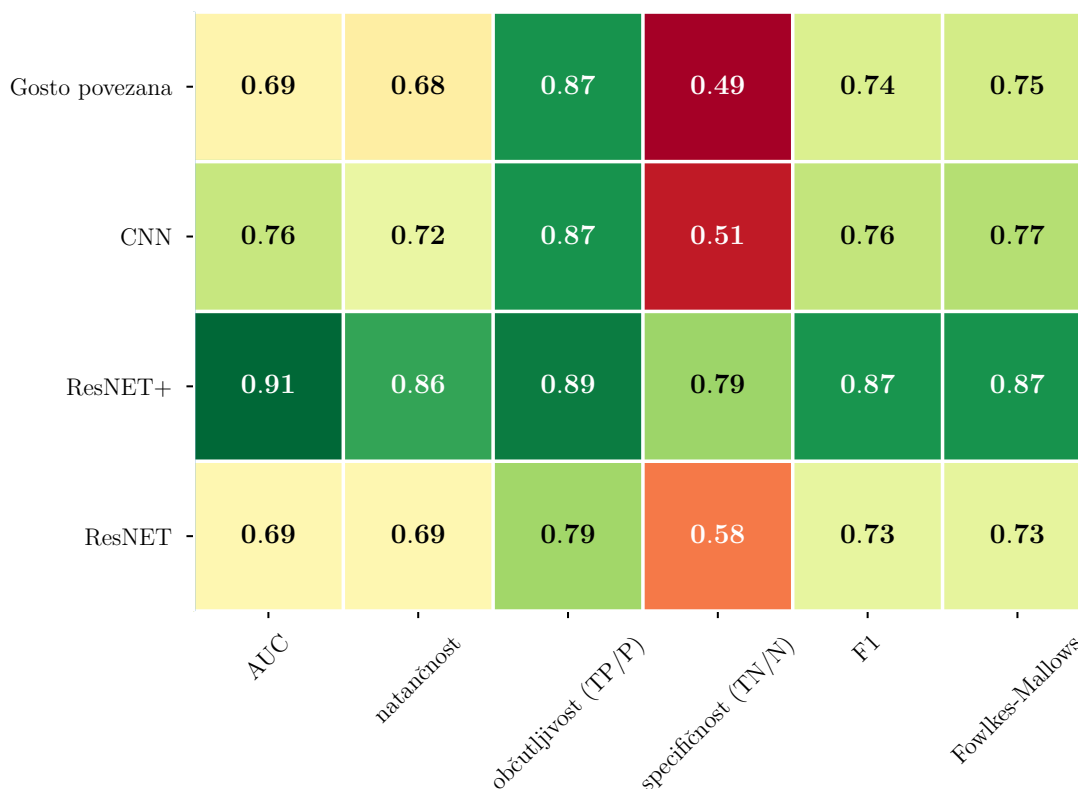
V vseh metrikah se kot najboljši izkaže model **ResNET+**, torej model z že inicializiranimi utežmi. V vseh primerih optimalni prag najdemo pri približno enakih vrednostih. Opazimo lahko, da že majhna sprememba vrednosti praga močno vpliva na večino metrik (ostri vrhovi). Ker izbira praga tako predstavlja odločitev, katero metriko želimo maksimizirati, se odločimo, da vzamemo prag, ki ga prepoznamo z metriko F1, saj se nam zdi ta najbolj primerna, glede na naravo klasifikacije.

Pri izbranih pragih si lahko pogledamo še uspešnost modelov, ki jo prikazujejo konfuzijske matrice na Sliki 5.



Slika 5: Konfuzijske matrike posameznih modelov. V celicah so podani deleži posameznih razredov.

Vsi modeli skoraj enako dobro prepoznajo resne okužbe, kar je najbolj bistveno za takšne vrste klasifikacijo. Najbolj se razlikujejo v deležu lažno pozitivnih (hudo bolnih), kar je sprejemljivo, saj pri takšni klasifikaciji predvsem ne želimo zgrešiti kakšnega hudo bolnega in ga prikrajšati za zdravljenje. Povzetek primerjave vseh modelov se nahaja na Sliki 6. Tudi ta razkriva podobne uvide, kot smo jih že opisali. Najbolj presenetljivi so slabi rezultati ResNET modela, ki smo ga trenirali sami, kar je verjetno posledica prekratkega treniranja, saj ima mreža bistveno več skritih parametrov kot gosto povezana in konvolucijska.



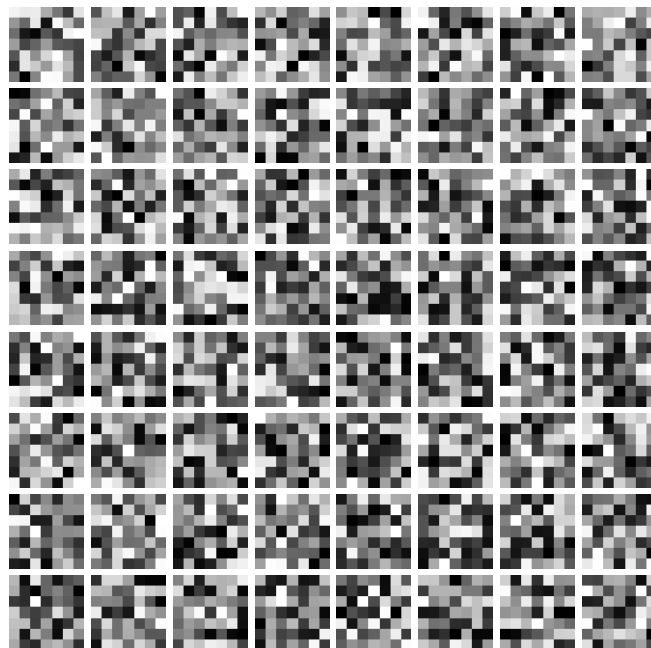
Slika 6: Primerjava vseh 4 modelov v nekaterih pomembnejših značilkah.

2.3 Vizualizacija konvolucijskih filtrov

Za obe konvolucijski mreži si ogledamo, kakšni so (po učenju) njihovi filtri prve konvolucijske plasti. Filtri naše implementacije konvolucijske mreže so prikazana na Sliki 7, mreže **ResNET** pa na Sliki 8. V obeh primerih težko prepoznamo kakšen jasn vzorec. Kakšen izmed filtrov se zdi, da bi lahko predstavljal pomemben rob na vhodnih slikah.



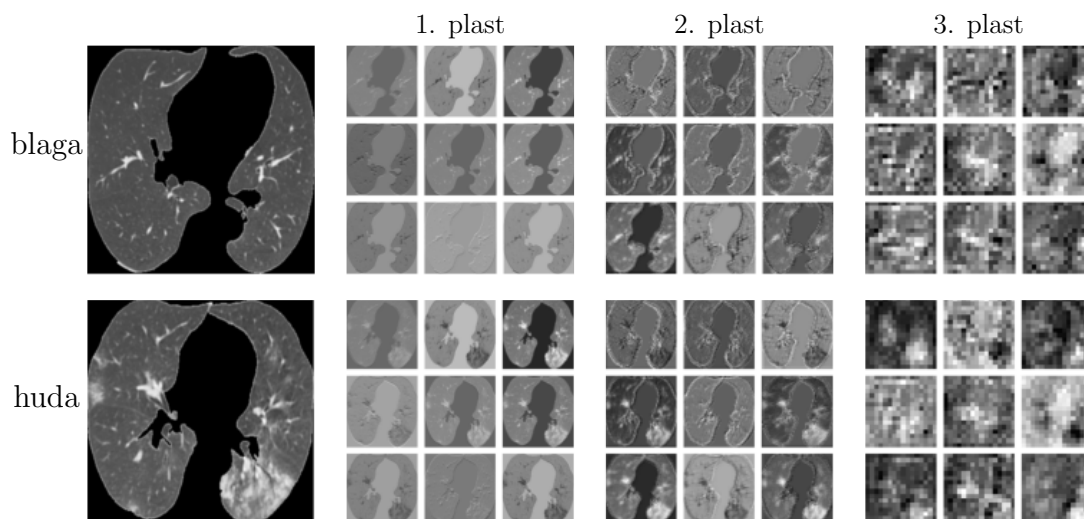
Slika 7: Filtri prve konvolucijske plasti konvolucijske mreže.



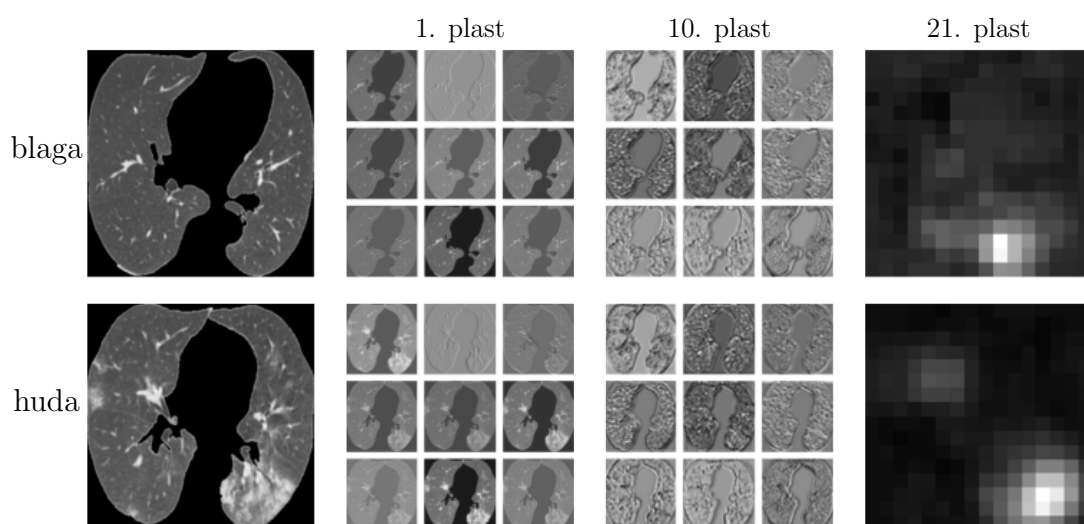
Slika 8: Filtri prve konvolucijske plasti **ResNET** mreže.

2.4 Vizualizacija konvolucij

Poleg filtrov si lahko ogledamo tudi, kaj se dogaja z vhodno sliko na posamezni konvolucijski plasti mreže. V ta namen si pogledamo, kaj se dogaja s sliko z blago in hudo okužbo na plasteh naše konvolucijske mreže (Slika 9) in **ResNET** mreže (Slika 10). Z vizualizacij hude okužbe povezanih z našo konvolucijsko mrežo vidimo se zdi, da se skozi plasti posebej poudarja spodnji desni kot, kjer se nahaj mlečno steklo, srednji še jasneje pokaže na **ResNET** rezultatih, kjer je v 21. plasti prav spodnji desni kot poudarjen, medtem ko je preostal del slike temen. Opazimo celo nekoliko manjši signal v bližini sredine slike, ki pravtako ustreza mlečnemu steklu na sliki s hudo okužbo.



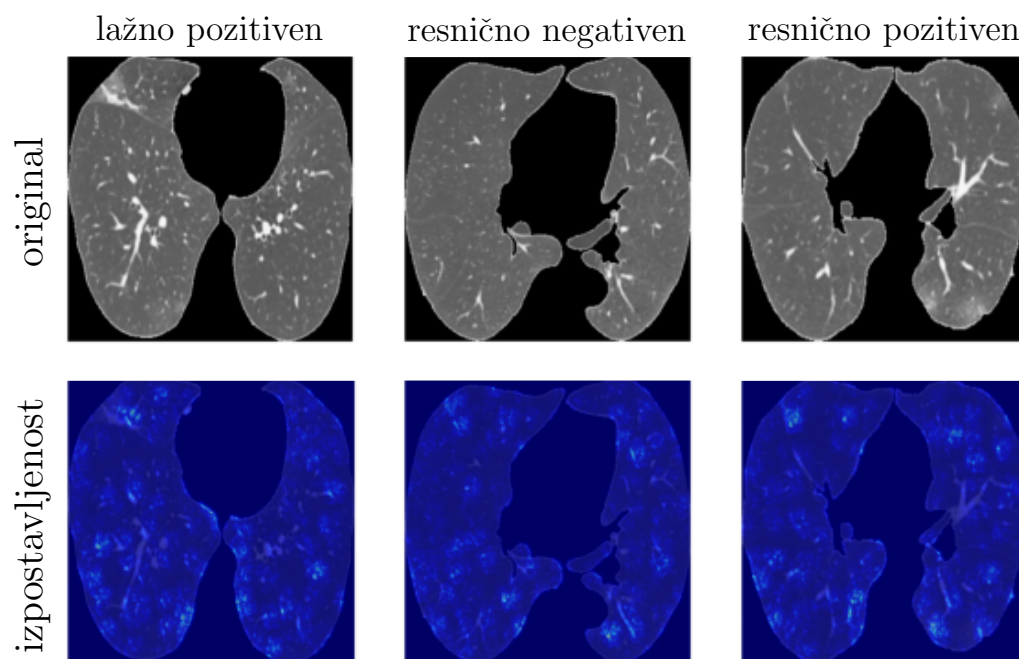
Slika 9: Prikaz prehoda pravilno klasificiranih slik s hudim in blagim obolenjem skozi našo konvolucijsko mrežo. Prikazanih je nekaj filtrov za 1., 2. in 3. plast.



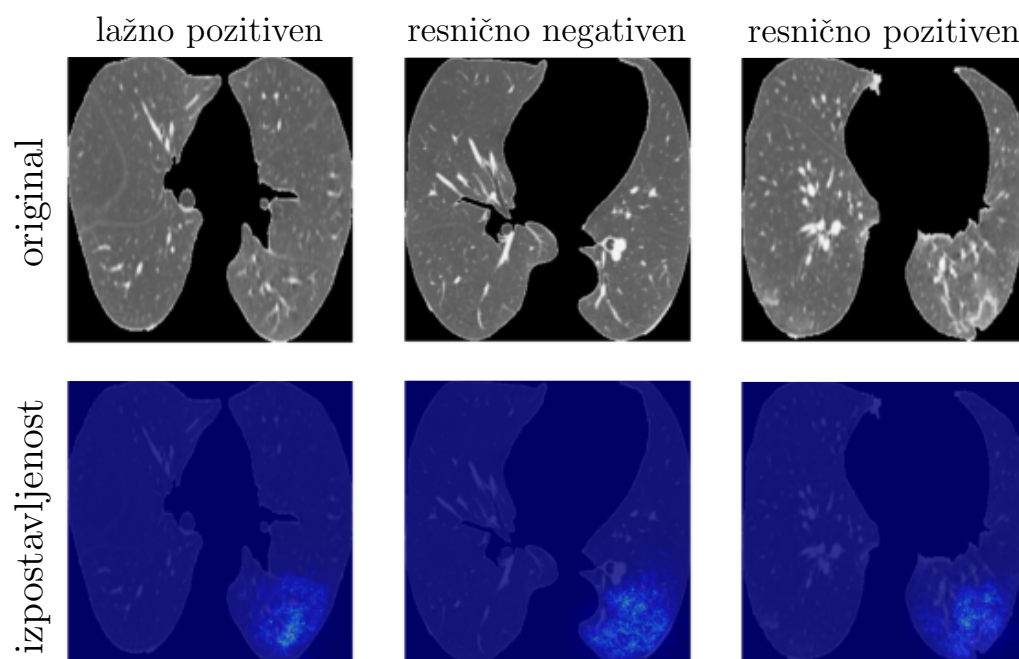
Slika 10: Prikaz prehoda pravilno klasificiranih slik s hudim in blagim obolenjem skozi **ResNET** mrežo. Prikazanih je nekaj filtrov za 1., 10. in 21. plast.

2.5 Vizualizacija map izpostavljenosti

Da bi videli, na podlagi katere informacije posamezna mreža sprejema odločitve glede klasifikacije si lahko pogledamo t. i. mape izpostavljenosti, ki so za našo konvolucijsko mrežo in **ResNET** prikazani na Slikah 11 in 12. Videti je, da se naša konvolucijska mreža odloča glede na signal s celotne slike, saj se piksli, ki bolj prispevajo h klasifikaciji nahajajo v gručah, ki so približno enakomerno razporejene po celotni sliki. Morda je pri mreži prišlo do t. i. prenaučenosti na vhodne podatke. Slike izpostavljenosti **ResNET** mreže prikazujejo drugačno zgodbo, saj se jasno vidi, da so največ prispevali h končni odločitvi piksli, v spodnjem desnem kotu, kjer opazimo mlečno steklo.



Slika 11: Mape izpostavljenosti pri lažno pozitivni, resnično negativni in resnično pozitivni klasifikaciji naše konvolucijske mreže. V zgornji vrstici se nahajajo originalne CT slike, spodaj pa mape izpostavljenosti.



Slika 12: Mape izpostavljenosti pri lažno pozitivni, resnično negativni in resnično pozitivni klasifikaciji ResNET mreže. V zgornji vrstici se nahajajo originalne CT slike, spodaj pa mape izpostavljenosti.