



Univerzitet u Beogradu  
Elektrotehnički fakultet  
Katedra za signale i sisteme



Projekat iz predmeta Analiza biomedicinske slike

**Klasifikacija metastaza kancera dojke na histopatološkim snimcima  
limfnih čvorova primenom konvolucionih neuralnih mreža**

**Studenti**

Marija Milutinović, 2015/0012

Janko Šušteršič, 2015/0114

**Mentor**

Doc. dr. Milica Janković

Maj 2019.

## Sadržaj

1.	Rezime.....	3
2.	Uvod.....	4
2.1.	Podaci .....	4
2.2.	Konvolucione neuralne mreže.....	5
2.2.1.	Sloj konvolucije .....	6
2.2.2.	Kerneli.....	8
2.2.3.	Slaganje više mapa obeležja .....	9
3.	Metod .....	11
3.1.	Obrada podataka .....	11
3.2.	Arhitektura neuronske mreže .....	12
3.2.1.	Osnovna arhitektura .....	13
3.2.2.	Arhitektura sa ReLu aktivacionom funkcijom.....	15
3.2.3.	Arhitektura sa 96 neurona u jako povezanom sloju .....	16
3.2.4.	Arhitektura sa otežanom funkcijom greške .....	18
4.	Rezultati.....	20
5.	Zaključak .....	22
6.	Literatura .....	23

## 1. Rezime

U ovom radu realizovan je sistem za automatsku klasifikaciju histopatoloških snimaka tkiva limfnih čvorova u zavisnosti od prisustva odnosno odsustva metastaza na snimku. Baza podataka predstavljala je skup histopatoloških snimaka u .tif formatu dimenzija 96x96x3 obezbeđena od strane takmičenja Histopathologic Cancer Detection na sajtu Kaggle. Ulazne slike skalirane su na dimenzije 48x48 i konvertovane grayscale sliku radi jednostavnijeg treniranja.

Zatim su u cilju ekstrakcije obeležja sa histopatoloških snimaka primenjene konvolucione neuralne mreže sa 2D konvolucionim slojevima i ReLu (RectifiedLinearUnit) aktivacionim funkcijama i max pool slojevima sažimanja. Za optimizaciju je primenjen Adam optimizer, a za određivanje funkcije greške cross entropija softmax sloja. Na izlazu ove mreže nalazili su se dva jako povezana sloja tako se dimenzija izlaza poslednjeg sloja pretvara u niz čija je dužina broj klasa. Evaluacija performansi mreža izvršena je na delu ulaznih snimaka i pokazano je da najbolje performanse pokazuje mreža sa sigmoidnom funkcijom aktivacije i 48 filtera. Sve mreže trenirane su u 24 epohe, a najbolji rezultati su dobijeni između 16 i 22 epohe. Najveća postignuta tačnost na testirajućem skupu iznosi 84.7 %.

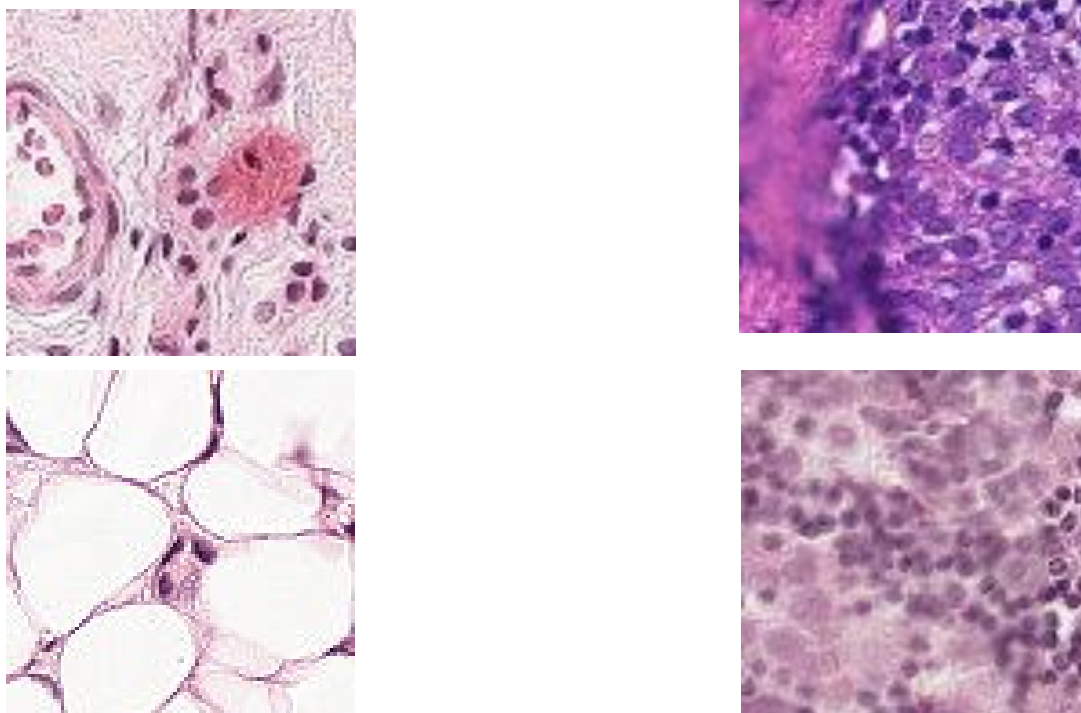
## 2. Uvod

### 2.1. Podaci

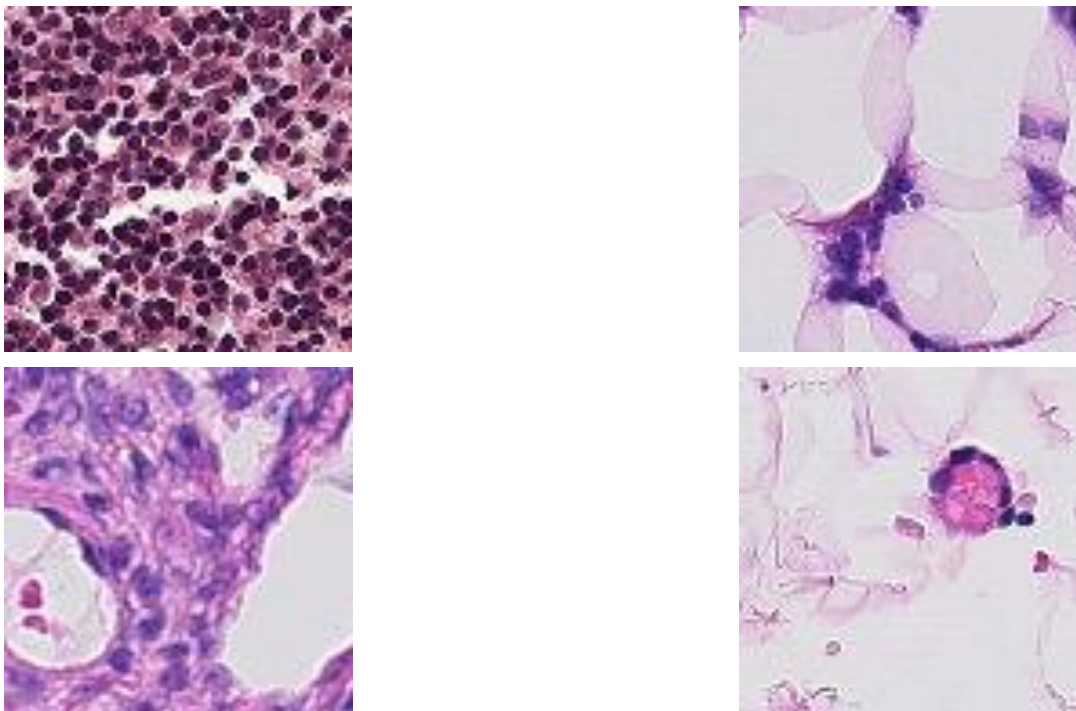
Bazu slika predstavljaju histopatološki snimci dimenzija 96x96 za takmičenje pod nazivom Histopathologic Cancer Detection sa sajta Kaggle koje predstavljaju delove većih snimaka limfnih čvorova dojki. Baza podataka je ujedno modifikovana verzija PatchCamelyon (PCam) baze podataka i zbog postojanja samo dve labelle (prisustvo (1) odnosno odsustvo (0) metastaze), njihova klasifikacija se može izvršiti obučavanjem modela na jednoj grafičkoj procesorskoj jedinici (GPU) za nekoliko sati [1].

Prisustvo metastaze na histopatološkom snimku ogleda se u narušenom odnosu veličine jedra, jedarceta i citoplazme (nukleocitoplazmatskom odnosu) u ćeliji koja je deo tkiva zahvaćenog malignim tumorom. Ćelije ovih tkiva deluju kao da su “rasute” po preparatu odnosno jasna granica između zdravog tkiva i tkiva na kom su prisutne morfološke promene se ne može sa sigurnošću odrediti.

Histopatološki snimci koji su obeleženi kao snimci na kojima se je zabeleženo prisustvo metastaze prikazane su na *slici 2.1.* dok su snimci na kojima metastaze odsustvuju prikazani na *slici 2.2.*



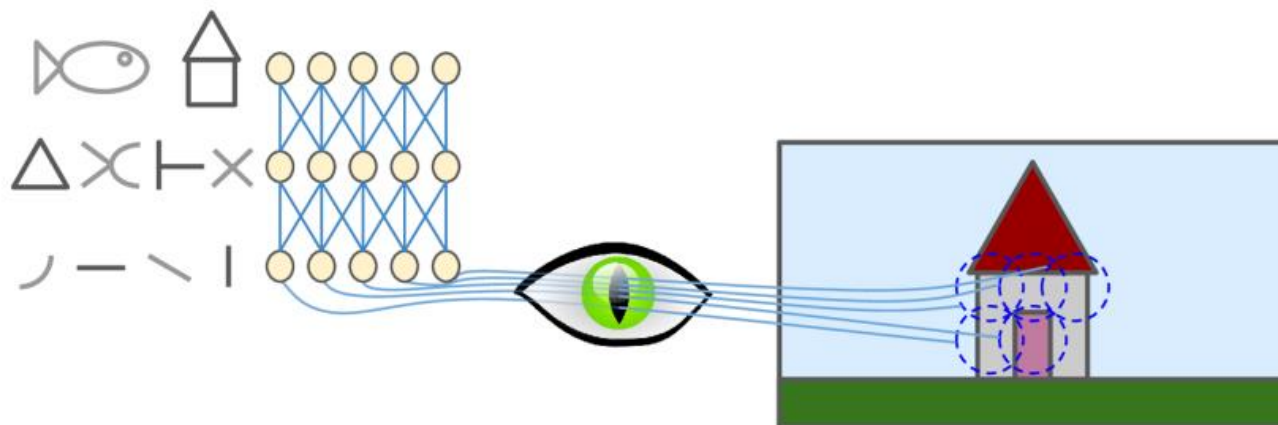
*Slika 2.1. Histopatološki snimci na kojima je označeno postojanje metastaze*



*Slika 2.2. Histopatološki snimci na kojima metastaze nisu prisutne*

## 2.2. Konvolucione neuralne mreže

Naučnici David H. Hubel i Torsten Wiesel su 1981. godine dobili Nobelovu nagradu iz oblasti medicine za svoj doprinos u analizi strukture vizuelnog korteksa. Oni su pokazali da većina neurona u vizuelnom korteksu poseduje svoje malo lokalno vizuelno polje, nazvano receptivno polje. Naime, svaki od neurona reaguje na vizuelne stimulse koji potiču samo iz ograničenog regiona vizuelnog područja, odnosno iz njegovog receptivnog polja, što je prikazano na *slici 2.3*. Receptivna polja različitih neurona se mogu međusobno preklapati, ali ona zajedno pokrivaju čitav vizuelni prostor. Naučnici su takođe pokazali da dva neurona mogu imati isto receptivno polje, ali da mogu reagovati na linije različitih orijentacija u tom polju (horizontalne, vertikalne, pod tačno određenim uglom). Sa druge strane, određeni neuroni reaguju na stimulse iz šire oblasti vizuelnog polja i kompleksnije strukture koje predstavljaju kombinaciju jednostavnijih struktura. Ovi zaključci doveli su do formiranja ideje da određeni neuroni koji raspoznaju složenije objekte, predstavljaju izlaze susednih neurona koji reaguju na jednostavnije strukture što se uočava na *slici 2.3*. gde je svaki neuron povezan samo sa nekoliko neurona iz prethodnog sloja.

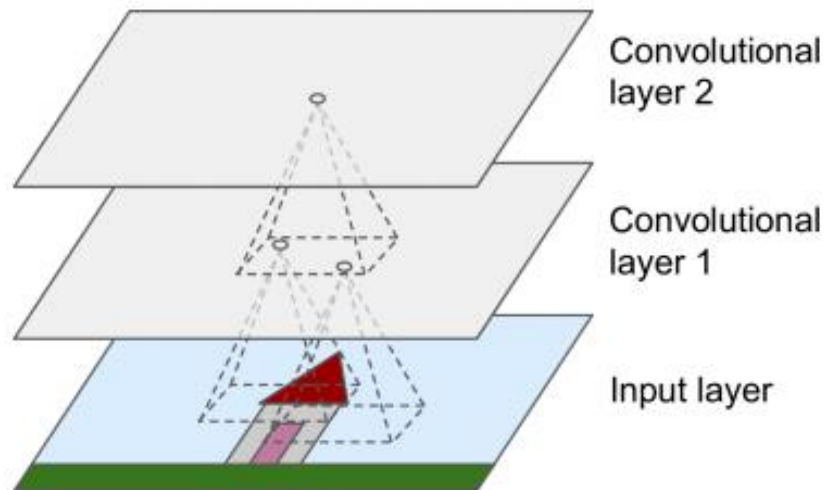


*Slika 2.3. Prikaz neurona (žuto) i njihovih lokalnih receptivnih polja (plavo) – preuzeto iz [4]*

Arhitektura poput ove je u mogućnosti da prepozna sve tipove kompleksnih struktura iz vizuelnog polja. Yann LeCun je sa svojim timom naučnika 1988. godine predložio LeNet-5 arhitekturu koja se primenjuju prilikom prepoznavanja rukom pisanih cifara. Ova mreža sadrži potpuno povezane neurone i sigmoidnu funkciju aktivacije, kao i dve nove gradivne strukture, a to su sloj konvolucije (convolutional layer) i sloj sažimanja (pooling layer). Prednost konvolucione neuralne mreže u odnosu na klasičnu duboku neuralnu mrežu je što sa porastom dimenzija slike broj parametara mreže postaje ogroman (Npr. za sliku dimenzija 100x100, odnosno 10000 piksela i za 1000 neurona u prvom sloju mreže, broj veza između neurona iznosi 10 miliona). CNN prevazilazi ovaj problem uvođenjem delimične povezanosti neurona između dva sloja.

### 2.2.1. Sloj konvolucije

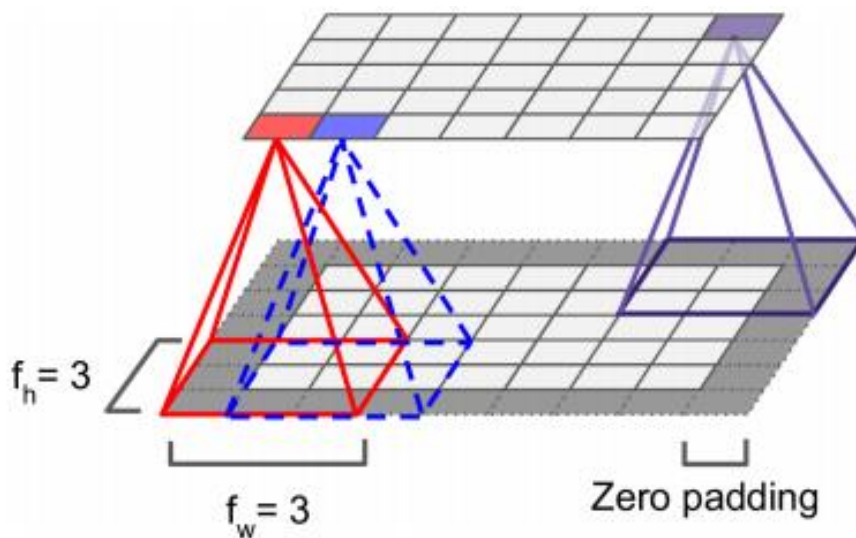
Najvažniji gradivni element konvolucione neuralne mreže - CNN predstavlja konvolucioni sloj. Neuroni u prvom konvolucionom sloju nisu povezani sa svakim pojedinačnim pikselom ulazne slike, već samo sa pikselima koji pripadaju njihovom receptivnom polju. U skladu sa pomenutim, svaki neuron iz drugog konvolucionog sloja je povezan samo sa neuronima koji pripadaju malom regionu u prvom konvolucionom sloju (odnosno receptivnim poljima neurona iz drugog sloja) što je prikazano na slici 2.4. Ova arhitektura omogućava mreži da u prvom skrivenom sloju obradi jednostavna obeležja na slici, da bi ih kasnije sastavljala u složenija obeležja u narednim skrivenim slojevima i tako dalje. Hijerahijska struktura odgovara slikama iz realnog života, što je razlog zašto su se CNN-ovi pokazali kao veoma uspešni prilikom prepoznavanja slika.



Slika 2.4. Konvolucioni slojevi i ulazni slojevi sa pravougaonim receptivnim poljem (preuzeto iz [4])

Klasične neuralne mreže poseduju jednodimenzionalne slojeve neurona, dok ulaz neurona u prvom sloju predstavlja izravnatu matricu pikseli slike. Kod konvolucionih neuralnih mreža, svaki sloj predstavlja matricu neurona koji su “stimulisani” pikselima iz odgovarajućeg receptivnog regiona prethodnog sloja, odnosno pikselima ulazne slike za prvi konvolucioni sloj.

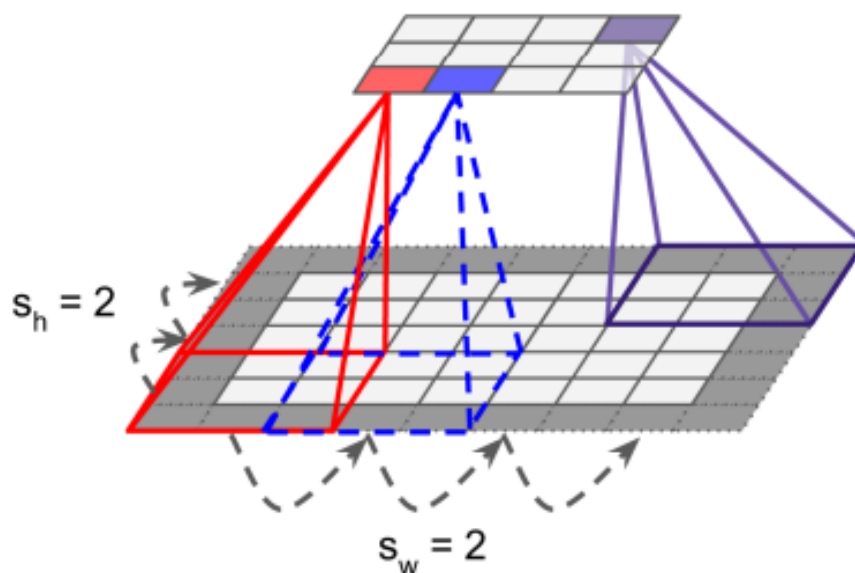
Neuron u  $i$ -toj vrsti i  $j$ -toj koloni određenog sloja je povezan sa izlazima neurona iz prethodnog sloja koji se nalazi u pravougaoniku ograničenom  $i$ -tom i  $i + f_h - 1$  vrstom i  $j$ -tom i  $j + f_w - 1$  kolonom, gde su  $f_h$  i  $f_w$  visina i širina receptivnog polja. Sa ciljem postizanja da naredni sloj neurona ima istu visinu i širinu kao prethodni sloj neurona, uobičajeno je dodavanje nulnih vrednosti na granice matrice, kao što je prikazano na slici 2.5.



Slika 2.5. Prikaz načina povezanosti između slojeva i dopunjavanje matrice nulama (preuzeto iz [4])



Pored toga, moguće je izlaze više neurona povezati sa ulazima manjeg broja neurona u narednom polju, redukcijom veličine razmaka između dva receptivna polja. Rastojanje između dva receptivna polja se još naziva i korakom. Na slici 2.6. prikazan je ulazni sloj dimenzija  $5 \times 7$  proširen nulama koji je povezan sa narednim slojem dimenzija  $3 \times 4$  sa dimenzijama receptivnog polja  $3 \times 3$  i veličinom koraka 2. Neuron na poziciji  $(i, j)$  u narednom sloju je povezan sa neuronima između  $i \times s_h$  do  $i \times s_h + f_h - 1$  vrste i  $j \times s_w$  do  $j \times s_w + f_w - 1$  kolone, gde su  $s_h$  i  $s_w$  vertikalni i horizontalni razmaci između receptivnih polja.

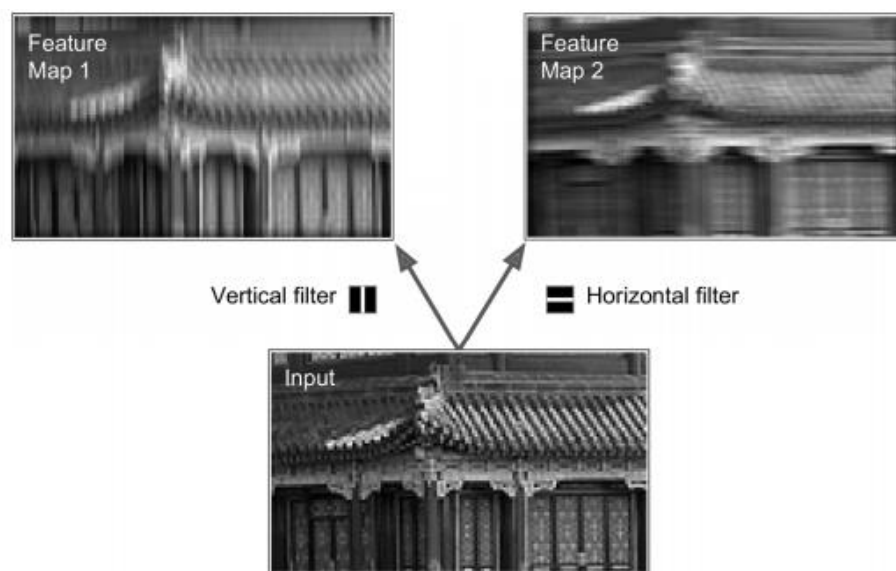


Slika 2.6. Redukcija dimenzije narednog sloja povećanjem rastojanja između receptivnih polja (preuzeto iz [4])

### 2.2.2. Kerneli

Težine neurona mogu se predstaviti kao slike čije su dimenzije jednake dimenzijama odgovarajućeg receptivnog polja. Na slici 2.7. prikazane su dva moguća skupa težina koje se nazivaju filtri(kerneli). Prvi predstavlja crni kvadrat sa vertikalnom linijom u sredini što znači da će neuron koji poseduju ove težine biti osetljiv samo na postojanje centralnih vertikalnih linija u svom receptivnom polju. Drugi filtar predstavljen je crnim kvadratom sa belom horizontalnom linijom u sredini što znači da će svi neuroni sa ovim vrednostima težima biti osetljivi samo na postojanje horizontalnih linija u njihovom receptivnom polju. Izlazi neurona koji koriste iste filtre rezultovaće stvaranjem mapa obeležja u kojima su naglašeni delovi slike najslabiji filtru, da bi tokom treniranja CNN kombinovala jednostavne filtre u nešto složenije.





*Slika 2.7. Primena dva različita kernela za dobijanje mapa obeležja (preuzeto iz [4])*

### 2.2.3. Slaganje više mapa obeležja

Do sada je svaki konvolucionni sloj bio predstavljen kao 2D mreža, ali u realnosti se on sastoji od više mapa obeležja istih dimenzija, te se može posmatrati kao trodimenzionalna struktura što je prikazano na slici 2.8. Svi neuroni jedne mape obeležja imaju iste vrednosti parametara težina i bias-a. Receptivno polje neurona se sada proširuje duž svih mapa obeležja, odnosno u konvolucionom sloju dolazi do primene više filtara na ulaznu sliku čime se obezbeđuje detekcija određenih obeležja bilo gde na ulaznoj slici.

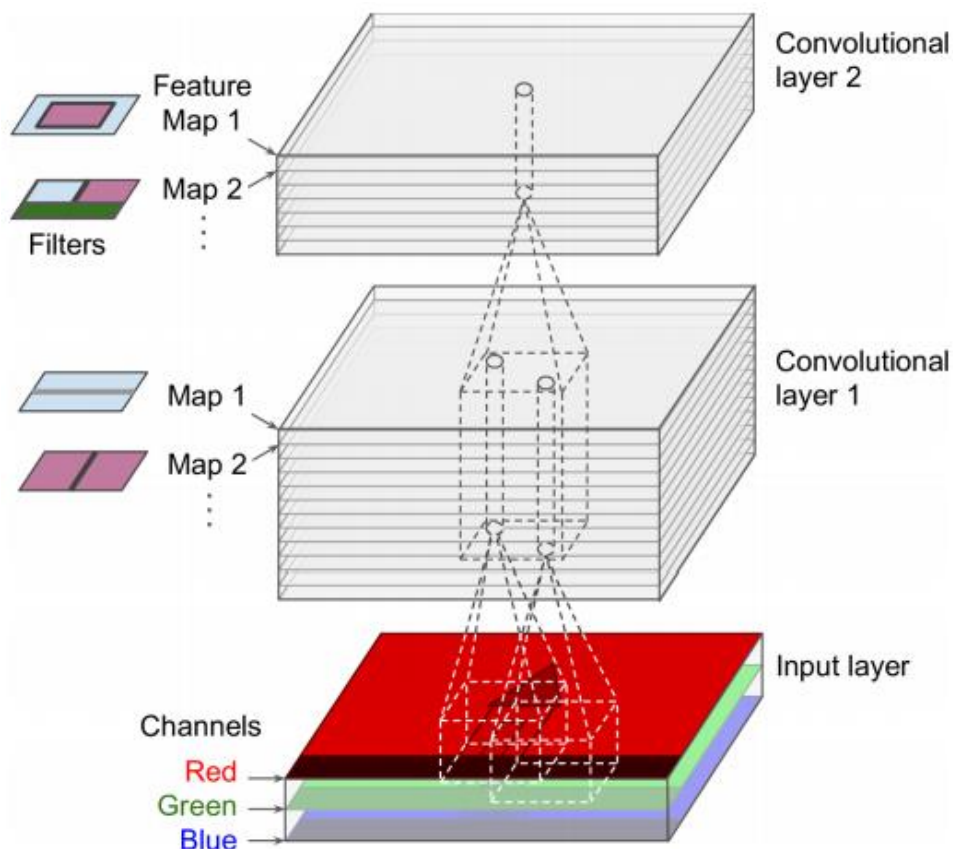
Činjenica da svi neuroni koji su vezani za jednu mapu obeležja poseduju iste parametre značajno redukuje broj parametara modela. Za razliku od klasične neuralne mreže koja kada nauči da prepozna određenu strukturu na jednom delu slike, nadalje može tu strukturu prepoznati samo na toj poziciji, CNN imaju mogućnost prepoznavanja naučenih oblika bilo gde na slici.

Slike u boji mogu se reprezentovati sa više ulaznih slojeva (za svaki od RGB kanala), dok su sive slike predstavljene jednim slojem. Neuron na poziciji  $(i, j)$  mape obeležja  $k$  u konvolucionom sloju  $l$ , povezan je sa izlazima neurona u  $l - 1$  sloju na pozicijama od  $i \times s_w$  do  $i \times s_w + f_w - 1$  reda i od  $j \times s_h$  do  $i \times s_h + f_h - 1$  kolone duž svih mapa obeležja u  $l - 1$  sloju. Svi neuroni koji se nalaze u istom redu i istoj koloni, ali su vezani za različite mape obeležja su povezani sa izlazima istih neurona u prethodnim slojevima.

Sledeća formula opisuje način izračunavanja vrednosti izlaza neurona u konvolucionom sloju odnosno izračunavanje otežinih suma svih ulaza uz dodavanje bias člana.

$$z_{i,j,k} = b_k + \sum_{u=1}^{f_h} \sum_{v=1}^{f_w} \sum_{k'=1}^{f_{n'}} x_{i',j',k'} \cdot w_{i',j',k'}, \text{ gde je } \begin{cases} i' = u \cdot s_h + f_h - 1 \\ j' = v \cdot s_w + f_w - 1 \end{cases}$$

- $z_{i,j,k}$  predstavlja izlaz neurona na poziciji  $(i, j)$  mape obeležja  $k$  u konvolucionom sloju  $l$
- $s_h$  i  $s_w$  predstavljaju visinu i širinu koraka, a  $f_h$  i  $f_w$  visinu i širinu receptivnog polja, dok je  $f_{n'}$  ukupan broj mapa obeležja u prethodnom odnosno  $l - 1$  konvolucionom sloju
- $x_{i',j',k'}$  predstavlja izlaz neurona u  $l - 1$  konvolucionom sloju na poziciji  $(i', j')$  mape obeležja  $k$
- $b_k$  je bias član mape obeležja  $k$  u konvolucionom sloju  $l$
- $w_{i',j',k'}$  je vrednost težina između bilo kog neurona mape obeležja  $k$  u nivou  $l$  i ulaza na poziciji  $(u, v)$  u  $k'$  mapi obeležja



Slika 2.8. Dva konvoluciona sloja i sa više mapa obeležja i ulazna slika sa 3 kanala (preuzeto iz [4])

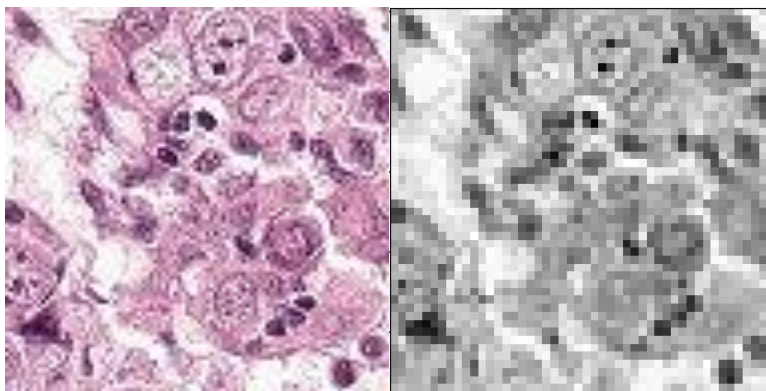
### 3. Metod

#### 3.1. Obrada podataka

Bazu slika predstavlja 220 025 obeleženih i 57458 neobeleženih histopatoloških snimaka dimenzija 96x96x3u .tif formatu, za takmičenje pod nazivom Histopathologic Cancer Detection sa sajta Kaggle koje predstavljaju delove većih snimaka limfnih čvorova dojki. Labele obeleženih slika su date u .csv fajlu u dve kolone odvojene zarezom. U prvoj koloni se nalazi ime slike, a u drugoj binarna labela 0 ili 1, u zavisnosti od prisustva tumora na odgovarajućoj slici. Zastupljenost klase je približno jednaka, slike koje sadrže tumor čine oko 40% ukupnog broja podataka, dok na ostalim slikama tumor nije prisutan.

Obeleženi snimci su nasumično podeljeni na trenirajući i validacioni skup tako da 160 000 slika pripada trenirajućem, a ostatak validacionom skupu, pri čemu se odnos zastupljenosti klase od 60-40% približno održava. Neobeleženi skup je tretiran kao testirajući, a evaluacija tačnosti na ovom skupu je izračunata u formi submitovanja .csv fajla poput datog fajla koji sadrži labele na sajtu Kaggle. Testirajući skup je balansiran, što znači da je zastupljenost obe klase podjednaka, za razliku od obeleženog skupa.

Fajl koji sadrži labele učitani su pomoću *csv reader*-a u *python*-u, a svakoj labeli pridružena je slika sa odgovarajućim nazivom. Slike su učitane pomoću biblioteke *Image* iz paketa *PIL*. Kako boja slike ne nosi dodatne informacije slike su pretvorene u *grayscale* ugrađenim konvertorom iz iste biblioteke. Zbog velikog broja slika i ograničenih kompjuterskih resursa slike su skalirane koeficijentom 2 koristeći ugrađeni *antialiasing* metod. Dobijene slike dimenzija 48x48 kao i labele sačuvane su u formi kompaktnih *numpy* nizova radi jednostavnog i brzog učitavanja pri treniranju i validaciji. Isti postupak primenjen je i na testirajuće slike, samo što su umesto labela u niz sačuvani nazivi slika u svrhu budućeg ispisivanja izlaznog .csv fajla.



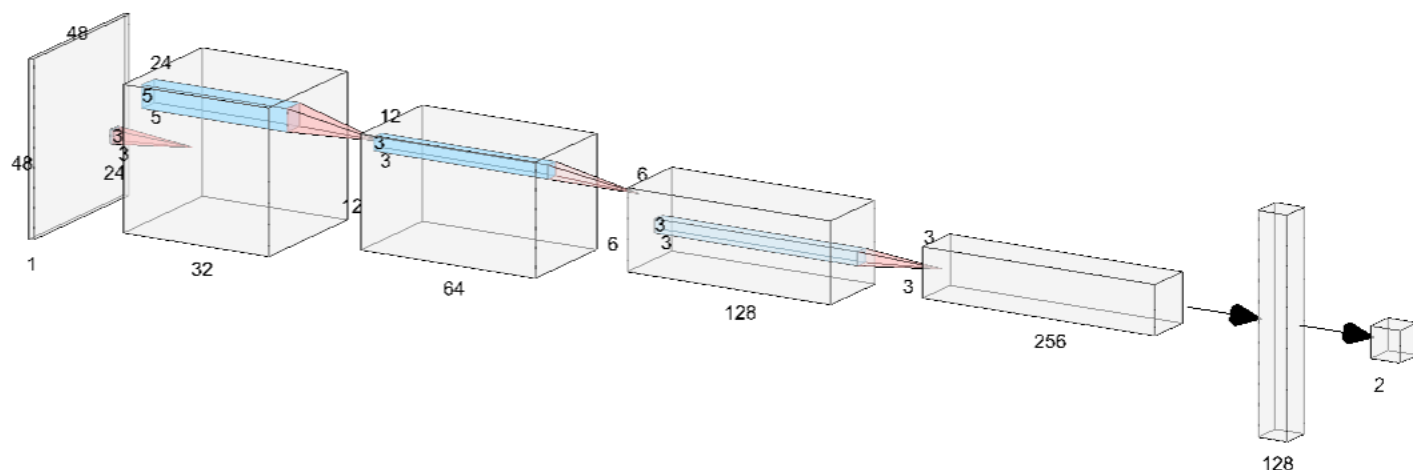
*Slika 3. 1. Prikaz originalne slike i slike nakon ulazne obrade*

### 3.2. Arhitektura neuronske mreže

Osnovna neuronska mreža koja je ispitivana u ovom radu bila je klasična konvoluciona neuronska mreža čiji je ulaz slika dimenzija 48x48. Kako su osvetljenja slika predstavljena celobrojnima vrednostima od 0 do 255 nepogodna za konvolucionu neuronsku mrežu, vrednosti piksela skalirani su na opseg realnih brojeva od -0.5 do 0.5. Arhitektura neuronske mreže sastojala se od četiri dvodimenzionalna konvoluciona sloja sa *ReLU* aktivacionom funkcijom između kojih su se nalazili 2x2 *maxpool* slojevi *pooling*-a. Filtri konvolucionih slojeva bili su dimenzija 3x3, osim drugog koje je bio 5x5. Broj filtara po sloju linearno raste od 32 u prvom do 256 u četvrtom sloju, pri čemu se slike popunjavaju nulama tako da slika zadržava iste dimenzije nakon filtriranja.

Izlaz konvolucione mreže dimenzija 256x3x3 se poravna u jednodimenzioni niz koji prestavlja ulaz dva jako povezana sloja tako da je dimenzija izlaza drugog sloja jednaka broju klasa – u ovom slučaju dva. Vrednosti na izlazu jako povezanog sloja su ulaz u *softmax* sloj koji daje verovatnoće pripadnosti slika svakoj klasi. Za izračunavanje funkcije greške korišćena je *cross entropija softmax sloja*, a za optimizaciju korišćen je *Adam optimiser*.

Izbor osnovne arhitekture neuronske mreže dobijen je ispitivanjem nekoliko arhitektura, variranjem broja slojeva i veličina filtra i testiranjem na validacionom skupu. Dobijeno je da se slaganjem više konvolucionih slojeva između dva *pool* sloja ne poboljšava performans mreže, a da sa povećanjem broja parametara mreže tačnost klasifikatora raste. Konačna arhitektura je ograničena kompjuterskim resursima, kako memorijskim tako i vremenom potrebnim za treniranjem mreže do konvergencije, što se pokazalo da nastupa oko 16-22 epohe. Grafička kartica korišćena u ovom radu je *GeForce GTX 950M 2GB*, a treniranje izabrane mreže od 24 epoha traje oko 1h. Ostali parametri neuronske mreže podešavani su metodom krosvalidacije.

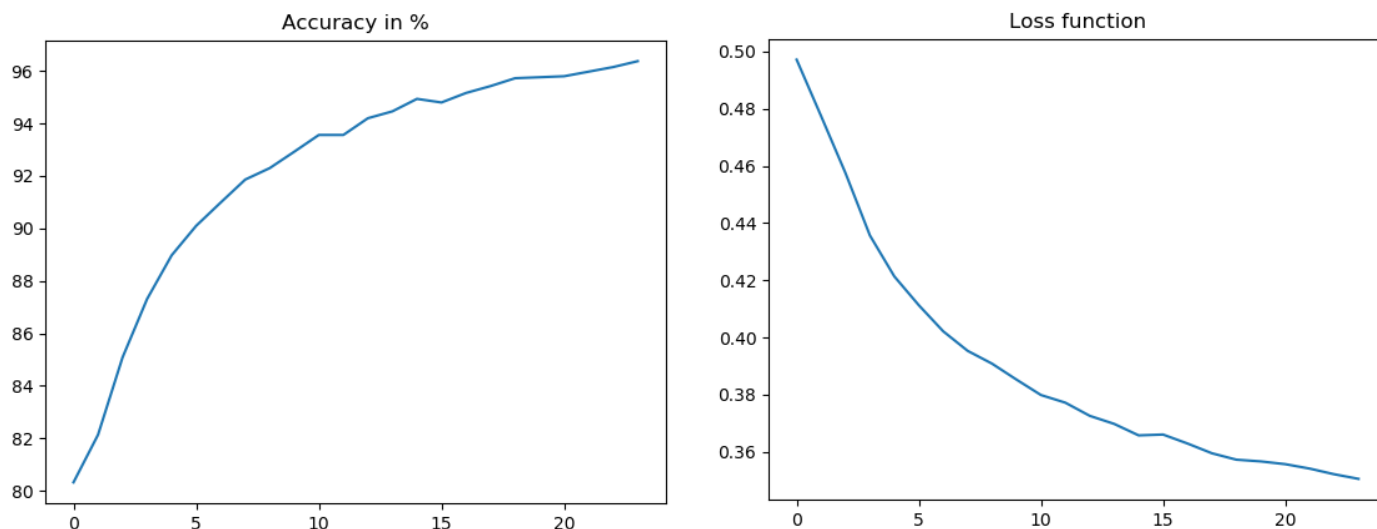


Slika 3. 2. Prikaz osnovne arhitekture konvolucione neuralne mreže

### 3.2.1. Osnovna arhitektura

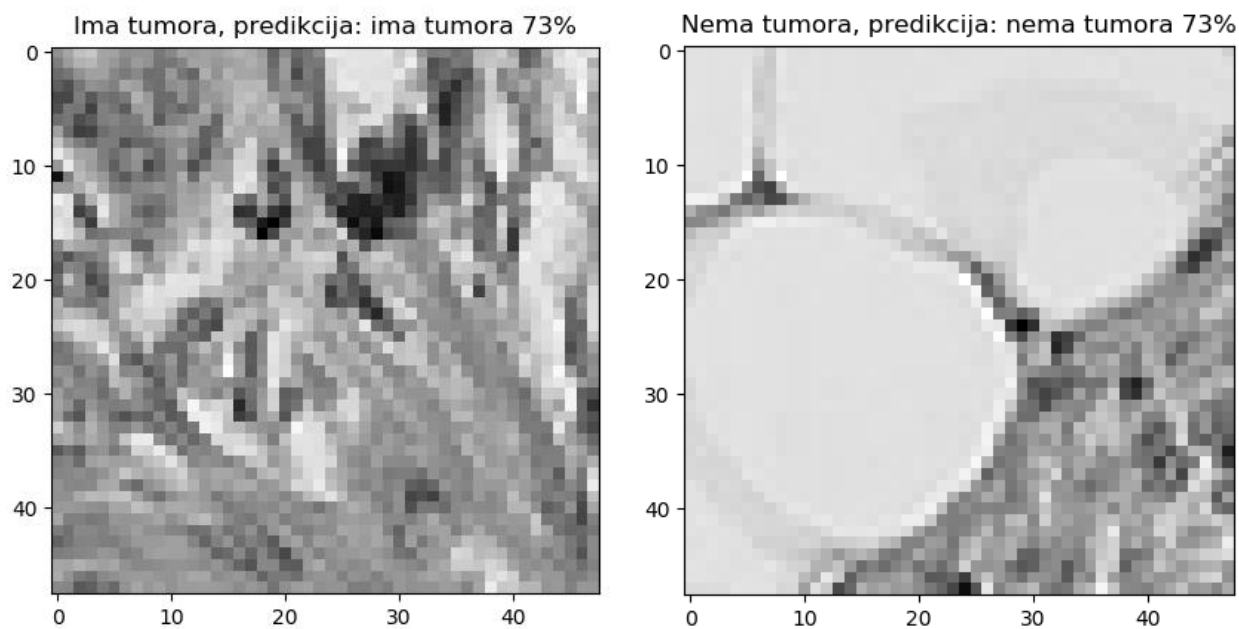
Prva, ujedno i referentna arhitektura konvolucione neuronske mreže ispitivana u ovom radu jeste arhitektura opisana u 3.2. Podesivi parametri obuhvataju broj neurona u prvom i funkciju aktivacije na drugom jako povezanom sloju. Broj neurona u prvom jako povezanom sloju izabran je kao koren broja parametara na ulazu – 48. Za funkciju aktivacije na izlazu izabrana je funkcija *sigmoid*-  $f(x) = \frac{1}{1+e^{-x}}$ . Neuronska mreža je trenirana kroz 24 epoha, dostižući najbolji rezultat na validacionom skupu nakon svega 16 epoha. Tačnost predikcija, izražena u procentima, kao i funkcija greške tokom treninga, u zavisnosti od broja epoha prikazani su na slici 3.3.

Vrednosti funkcije greške i tačnosti za prikaz računate su na oko 5% ukupnog broja podataka na kojim se vrši treniranje, što je zbog velikog broja podataka dovoljno da se postigne njihov pravilan, monoton izgled.

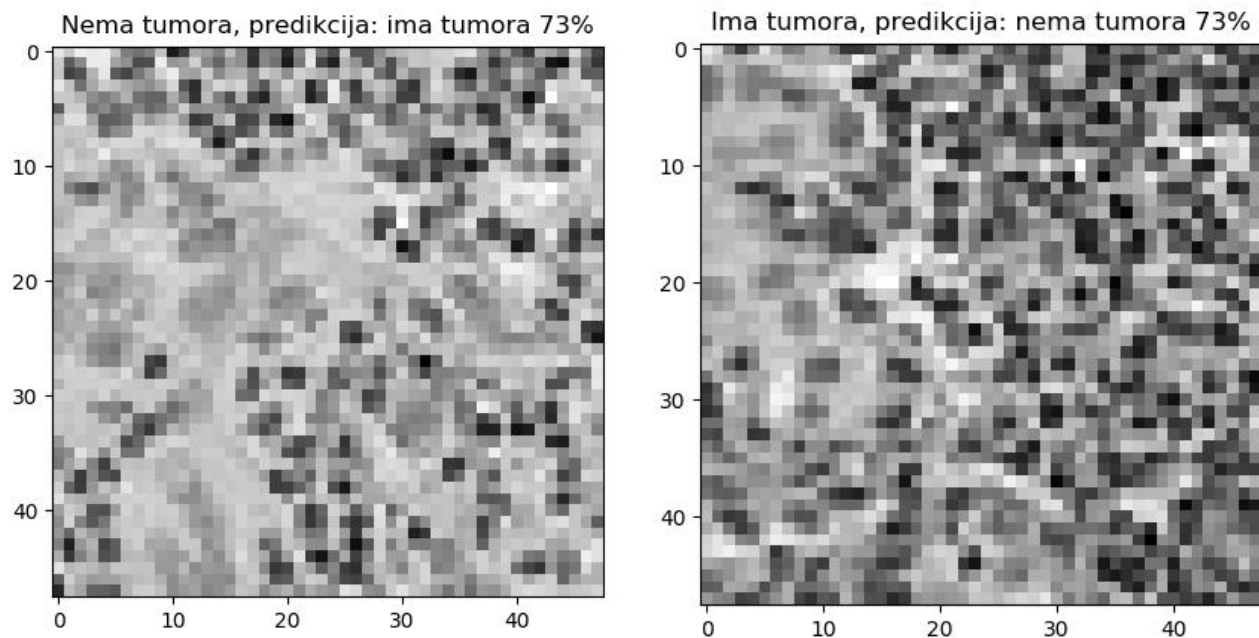


Slika 3. 3. Prikaz tačnosti predikcija u % (levo) i funkcije greške (desno) u zavisnosti od broja epoha tokom treninga

Grafik zavisnosti tačnosti predikcija tokom treniranja mreže od broja epoha nam pokazuje da nakon samo prve epohe prosečna tačnost prelazi 80%. Osim što je ovo dokaz da je arhitektura mreže pogodno izabrana, posledica je i toga što je problem jednostavan, pa je potrebno obratiti pažnju da ne dođe do preobučavanja mreže. Iz tog razloga postoji validacioni skup na kome se evaluiraju mreže trenirane na različitom broju epoha, a zatim se mreža sa najvećom tačnošću evaluira na test skupu. Najbolja tačnost ove mreže na validacionom skupu postiže se nakon 16 epoha i iznosi 90.0%, dok je njena tačnost na testirajućem skupu 84.7%. Primeri slika tačno i netačno klasifikovanih slika iz validacionog skupa dati su na slikama 3.4. a) i 3.4. b)



*Slika 3. 4. a) Primeri tačno klasifikovanih odbiraka iz validacionog skupa*



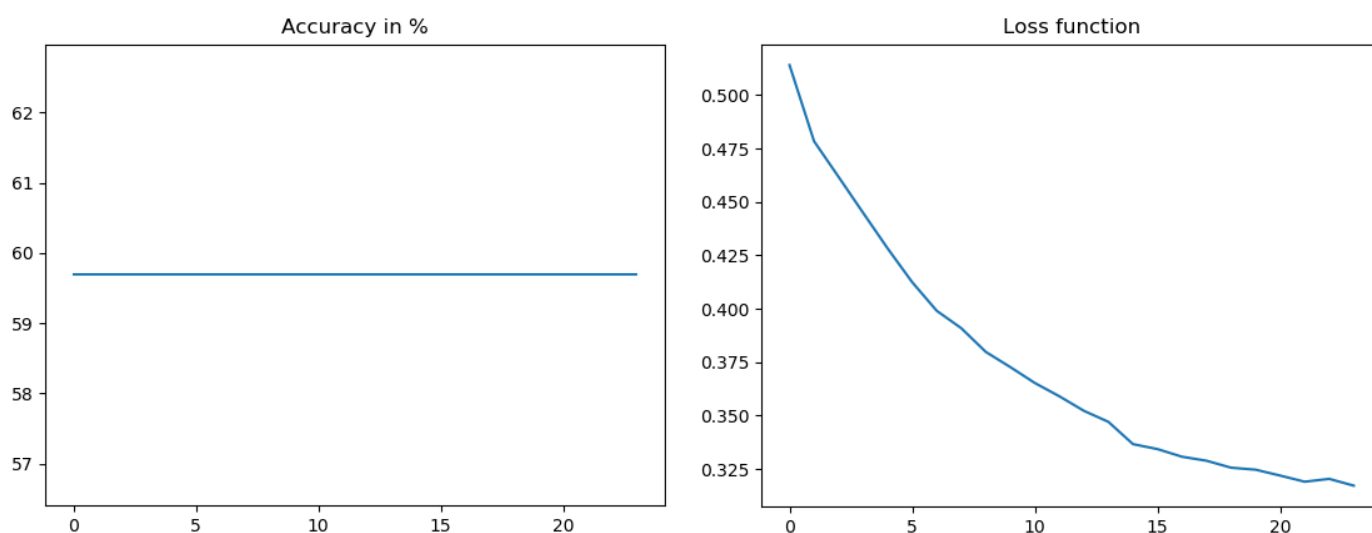
*Slika 3. 4. b) Primeri netačno klasifikovanih odbiraka iz validacionog skupa*



### 3.2.2. Arhitektura sa ReLu aktivacionom funkcijom

Nedostatak osnovnog modela sa *sigmoid* aktivacionom funkcijom na izlazu se može uočiti na primerima klasifikacija na validacionom skupu. Ulaz u poslednji, *softmax*, sloj je ograničen sa obe strane – minimalna vrednost je 0, a maksimalna 1, što je posledica korišćenja *sigmoid* - ograničene aktivacione funkcije. Kao posledica primene *softmax* sloja javlja se ograničen opseg sigurnosti u odluku koja ne može da bude veća od 73% - odnos maksimalnog i minimalnog izlaza iznosi  $e = 2.718$ . Funkcija *ReLU* nije ograničena sa gornje strane, tako da ovo ograničenje ne postoji. Još jedan razlog primene *ReLU* aktivacione funkcije je njihova rasprostranjenost u konvolucionim neuralnim mrežama zbog dobrih rezultata koje postižu.

Uprkos potencijalnim prednostima *ReLU* funkcije, rezultati pokazuju da se njihovom upotrebom na izlazu jako povezanog sloja dobijaju besmisleni rezultati. Naime, neuronska mreža brzo iskonvergira tako da uvek vraća istu klasu, što je u slučaju nebalansiranih podataka u trenirajućem skupu klasa 0, odnosno nepostojanje tumora na slici. Izvod funkcije rampe je funkcija znaka koja nije pogodna za *gradient decent* algoritam, tako da je učenje kroz epohe malo, tako da nema uticaja na tačnost, što se može uočiti sa grafikama tačnosti i funkcije greške na slici 3.5. Rezultati na validacionom i testirajućem skupu su očigledni, 59.5% odnosno tačno 50%, što je posledica predviđanja uvek iste klase. Ova arhitektura nam je pokazala da je bolje primenjivati *sigmoid* funkciju, što važi za svaku narednu grupu parametara koja je ispitivana.



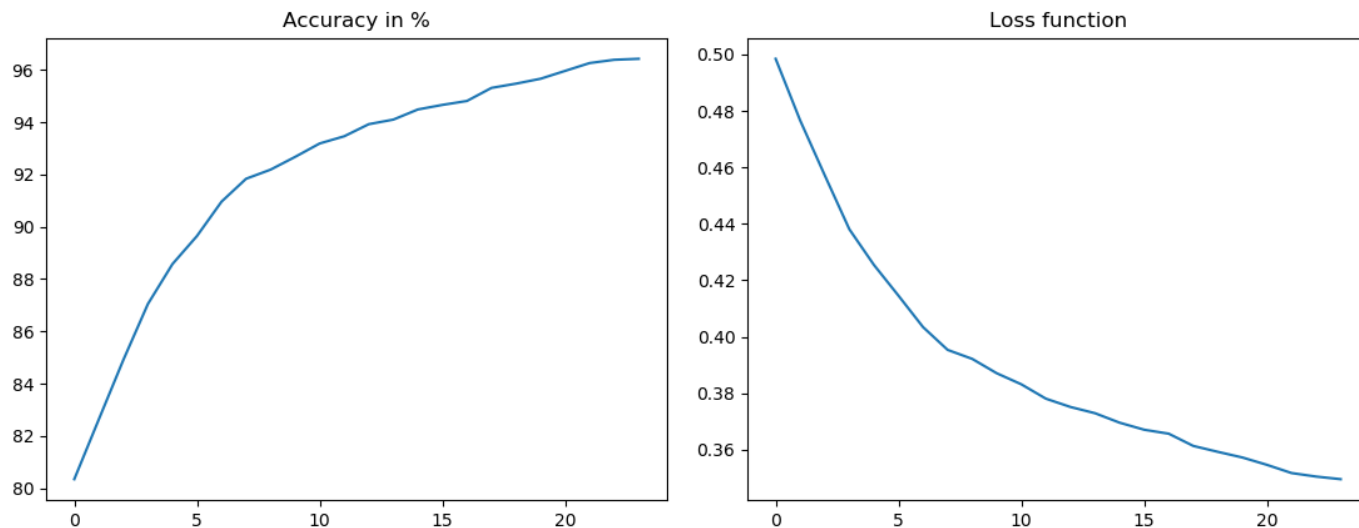
Slika 3. 5. Prikaz tačnosti predikcija u % (levo) i funkcije greške (desno) u zavisnosti od broja epohe tokom treninga



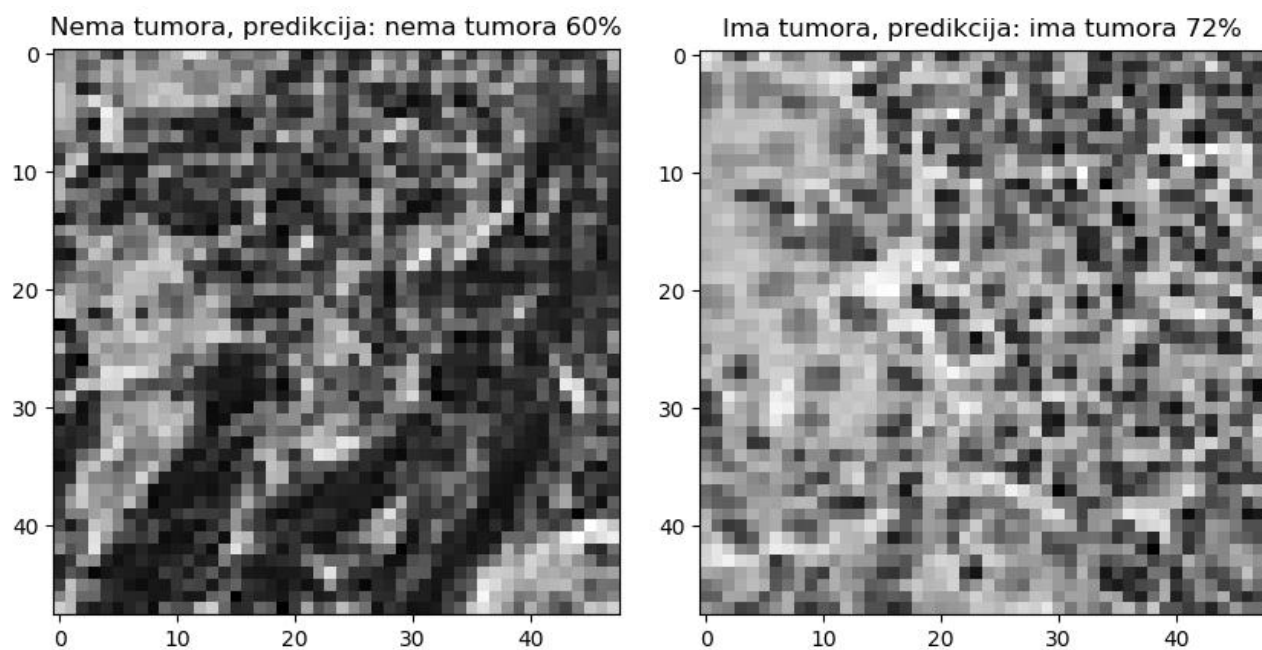
### 3.2.3. Arhitektura sa 96 neurona u jako povezanom sloju

Sledeća arhitektura koja je ispitivana jeste osnovna arhitektura sa *sigmoid* aktivacionom funkcijom na izlazu, sa promenjenim brojem neurona u prvom jako povezanom sloju – umesto 48 neurona postavljeno je 96 neurona. Ideja ove promene je ispitati kako povećanje broja parametara u jako povezanom sloju utiče ne celokupne performanse – tačnost predikcija i vreme treninga, kao i potreban broj epoha do konvergencije. Tačnost predikcija, izražena u procentima, kao i funkcija greške tokom treninga, u zavisnosti od broja epoha prikazani su na *slici3.6*.

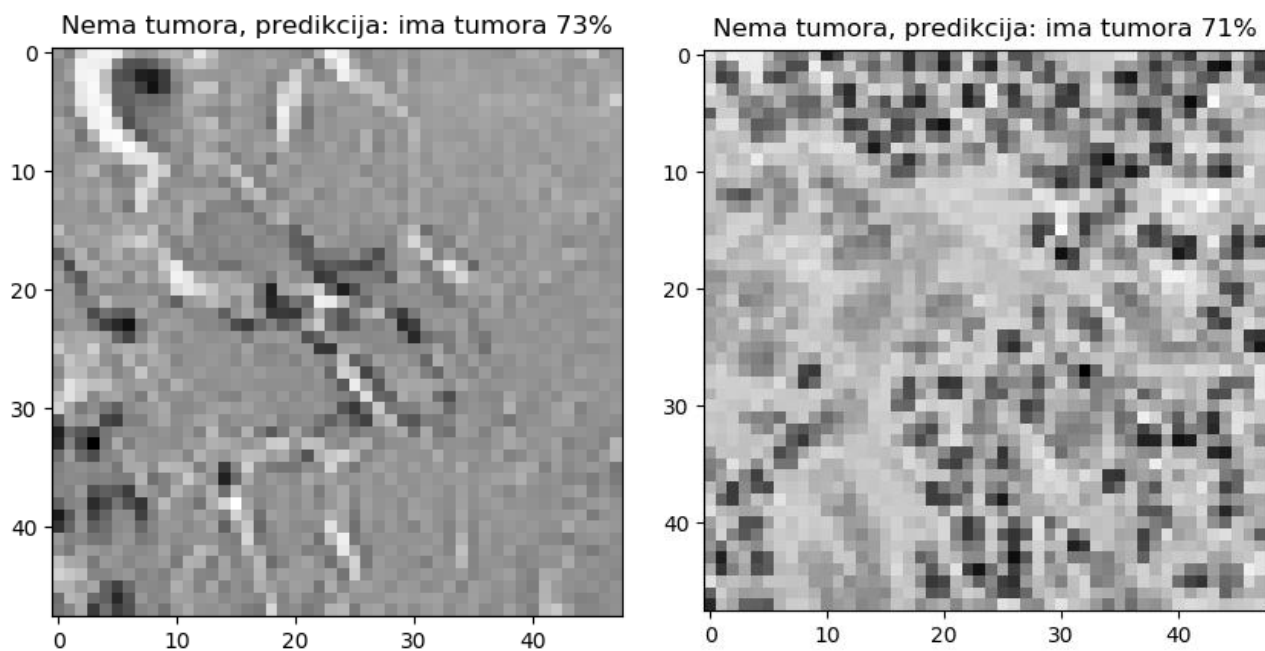
Treniranje ovakve arhitekture konvolucione neuronske mreže ne zahteva značajno dodatno vreme i potreban broj epoha, ali primećujemo da se ni performanse nisu značajno promenile. Najbolja tačnost ove mreže na validacionom skupu postiže se nakon 22 epoha i iznosi 89.7%, dok je njena tačnost na testirajućem skupu 84.1%. Ispitivana je i mreža sa 64 neurona u prvom jako povezanom sloju, ali je zaključak da je uticaj dodatnih neurona manji od uticaja nasumične inicijalizacije matričnih koeficijenata, po cenu nešto većeg vremena treniranja, te se u narednim primerima opredeljujemo za korišćenje 48 neurona. Primeri slika tačno i netačno klasifikovanih slika iz validacionog skupa dati su na slikama 3.7. a) i 3.7. b). Na *slici3.7. a)* prikazani su tačno klasifikovani snimci koji su inicijalnom arhitekturom bili netačno klasifikovani, a na *slici3.7. b)* prikazani su primeri netačno klasifikovanih slika.



*Slika 3. 6. Prikaz tačnosti predikcija u % (levo) i funkcije greške (desno) u zavisnosti od broja epoha tokom treninga*



*Slika 3. 7. a) Primeri tačno klasifikovanih odbiraka iz validacionog skupa*



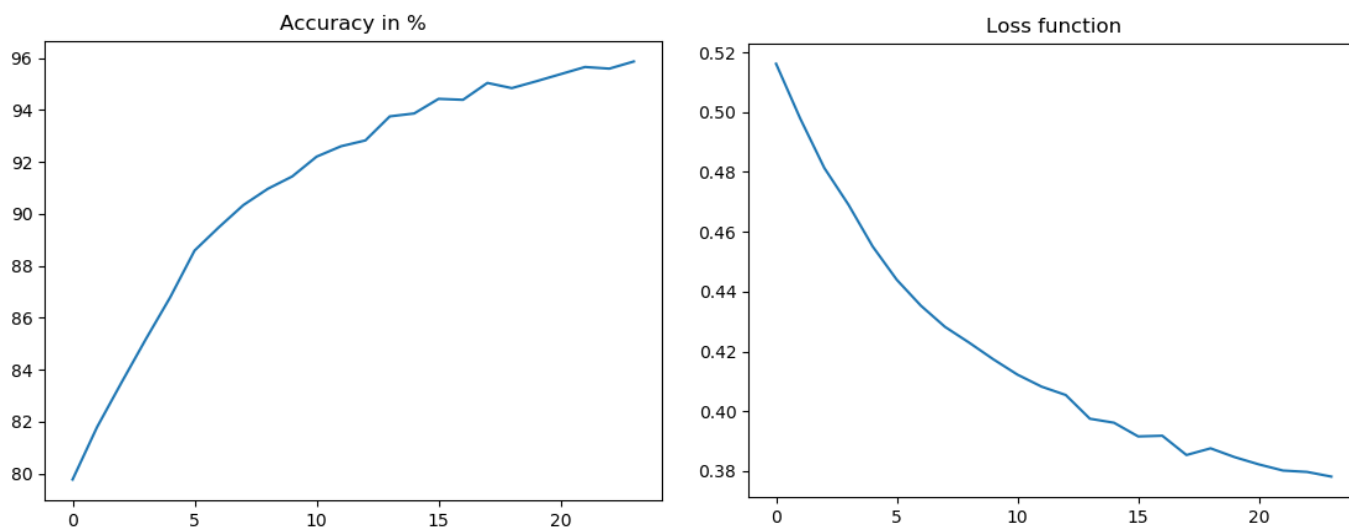
*Slika 3. 7. b) Primeri netačno klasifikovanih odbiraka iz validacionog skupa*

### 3.2.4. Arhitektura sa otežanom funkcijom greške

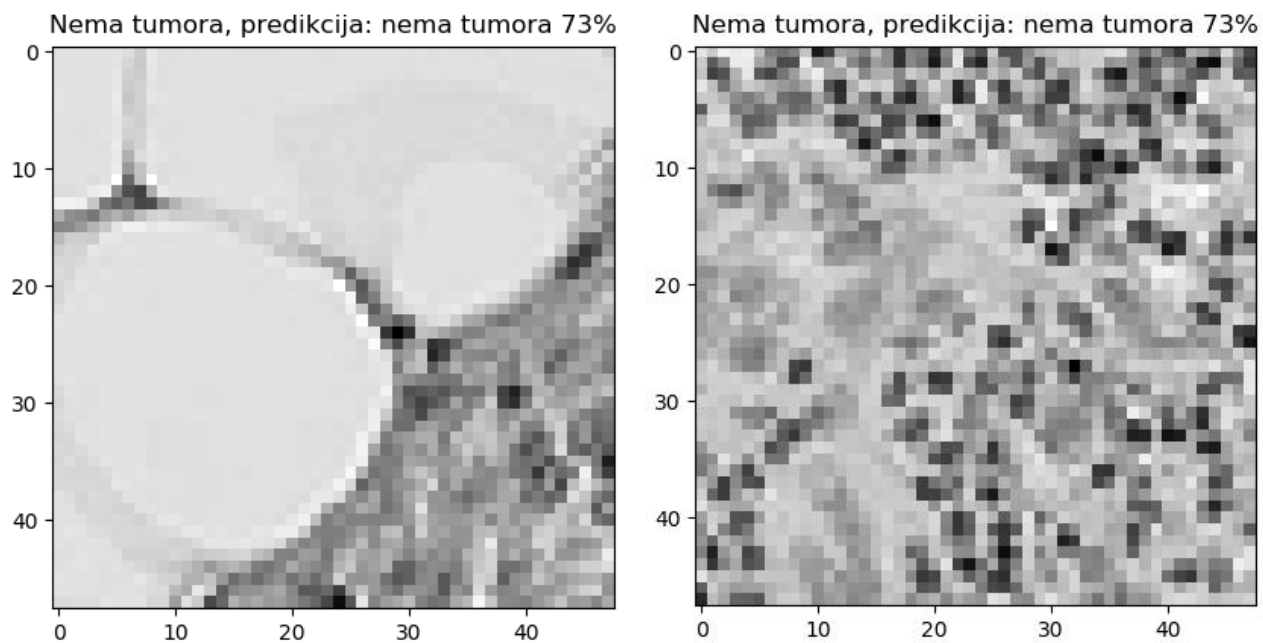
Nedostatak svih prethodno navedenih arhitektura se nalazi u tome što one ne uzimaju u obzir nebalansiranost trenirajućeg skupa. Odnos broja slika na kojima nema tumora i slika na kojima se tumor nalazi je približno 3:2. Posledica neujednačenog broja podataka je nejednak broj grešaka tipa *false positive* i *false negative* – broj *false positive* predikcija je veći nego broj *false negative* predikcija. Broj slučajeva kada postoji tumor na slici, a nije detektovan prelazi 13%, dok je broj netačnih predikcija kada tumora nema skoro duplo manji. To je bila motivacija za uvođenje sledeće izmene u arhitekturu neuronske mreže: zadržana je osnovna arhitektura koja do sada daje nabolje rezultate s tim što *cross entropija softmax funkcije* otežana različitim koeficijentima za različite klase.

Koeficijenti kojima su izlazi *sigmoid* aktivacione funkcije pomnoženi bili su 1.3 i 0.7. Oprobane su obe varijante otežavanja – kada je izlaz 0 pomnožen većim koeficijentom, a izlaz 1 manjim, kao i kada je izlaz 1 pomnožen većim koeficijentom, a izlaz 0 manjim. Tačnost predikcija, izražena u procentima, kao i funkcija greške tokom treninga, u zavisnosti od broja epoha prikazani su na slici 3.8. Rezultati se međutim nisu ponašali očekivano. U oba slučaja došlo je do još veće razilke u tipovima grešaka, pri čemu je celokupna tačnost ostala približno konstantna.

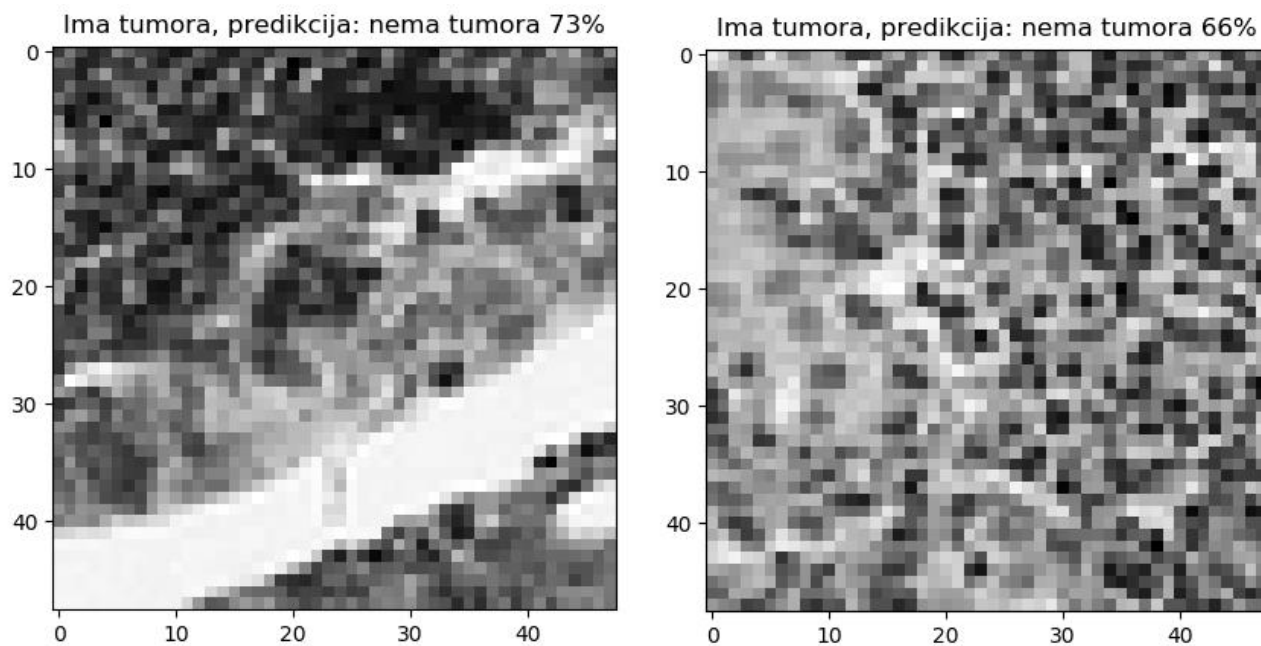
Najbolja tačnost ove mreže na validacionom skupu postiže se nakon 16 epoha i iznosi 89.8%, dok je njena tačnost na testirajućem skupu 83.8%. Primeri slika tačno i netačno klasifikovanih slika iz validacionog skupa dati su na slikama 3.9. Na slici 3.9. a) prikazani su tačno klasifikovani snimci koji su inicijalnom arhitekturom bili netačno klasifikovani, a na slici 3.9. b) prikazani su primeri netačno klasifikovanih slika koji su prethodnim arhitekturama tačno klasifikovani.



Slika 3. 8. Prikaz tačnosti predikcija u % (levo) i funkcije greške (desno) u zavisnosti od broja epoha tokom treninga



*Slika 3. 9. a) Primeri tačno klasifikovanih odbiraka iz validacionog skupa*



*Slika 3. 9. b) Primeri netačno klasifikovanih odbiraka iz validacionog skupa*

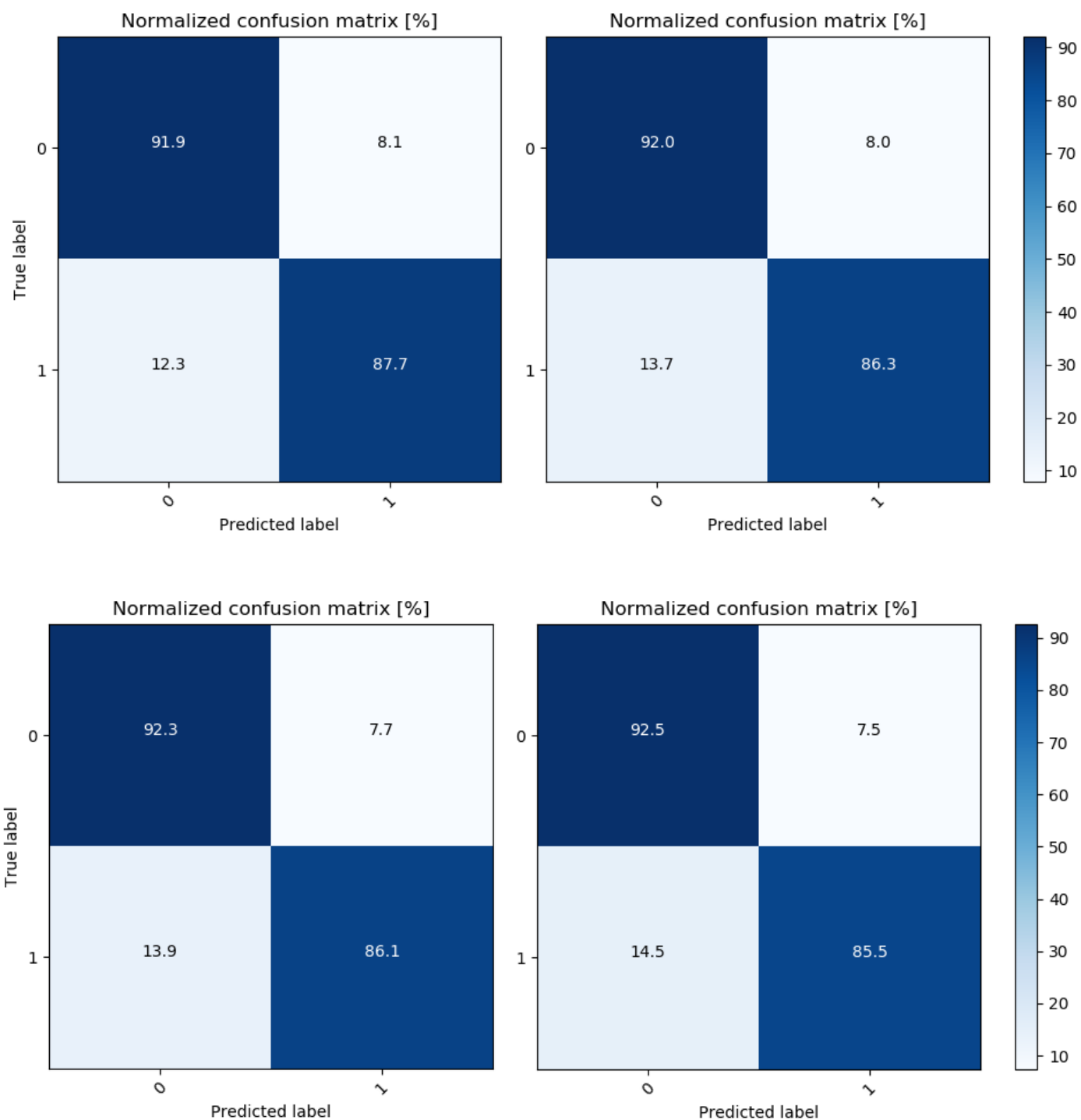
## 4. Rezultati

Komparativna analiza performansi različitih arhitektura konvolucionih neuralnih mreža data je u *tabeli 4.1*. Uočavamo da je najveća tačnost na validacionom i testirajućem skupu ostvarena za prvu arhitekturu mreže, dok približno dobre rezultate pokazuje i treća arhitektura mreže.

Arhitektura	Tačnost na validacionom skupu	Tačnost na test skupu
CNN1	90.0 %	84.7%
CNN2	59.5 %	50.0 %
CNN3	89.7 %	84.1 %
CNN4.1	89.8 %	83.8 %
CNN4.2	89.7 %	83,7 %

*Tabela 4.1. Komparativna analiza performansi različitih CNN arhitektura*

Osim tačnosti dobar pokazatelj performansi arhitektura neuralnih mreža jesu konfuzione matrice. Konfuzina matrica za problem klasifikacije u dve klase je matrica 2x2 na čijim osama se nalaze tačne labele i predikcije. Na glavnoj dijagonali nalaze se tačne predikcije, dok ostala polja predstavljaju greške u predviđanju neuronske mreže. Razlikujemo dve vrste grešaka – neuspešnu detekciju tumora kao i „lažnu uzbunu“, detekciju tumora kada na slici on zapravo ne postoji. U zavisnosti od problema, jedna od njih može da predstavlja veći problem od druge, kada težimo da je minimalizujemo, po cenu povećanja geške drugog tipa. U ovom radu za prag klasifikacije je uzeta vrednost 0.5, međutim ova vrednost se može dodatno podešavati pomoću ROC krive. Normalizovane konfuzione matrice izražene u procentima, na podacima iz validacionog skupa, za različite arhitekture konvolucionih neuronskih mreža date su na *slici 4.2*. Možemo da primetimo da osim najbolje celokupne tačnosti, prva mreža daje najmanju *false negative* grešku, odnosno greške prvog i drugog tipa se najmanje razlikuju.



Slika 4. 2. a)Konfuziona matrica za CNN1 b)Konfuziona matrica za CNN3

c)Konfuziona matrica za CNN4.1 d)Konfuziona matrica za CNN4.2

Početna arhitektura koja ujedno daje i najbolje rezultate odabrana je za treniranje na celokupnom trenirajućem i validacionom skupu. Validacioni skup je služio za podešavanje hiperparametara mreže i odabir epohe treniranja koja će se koristiti za testiranje određene arhitekture kako ne bi došlo do preobučavanja, tako da kada jednom odaberemo arhitekturu koja se najbolje pokazala možemo je trenirati na celokupnom labeliranom skupu. Proširivanjem skupa podataka za oko 30% očekivali smo da dobijemo veću tačnost na testirajućem skupu, međutim to se nije desilo. Najbolja tačnost koju je ova mreža postigla tokom treniranja iznosi 90.0%, dok je njena tačnost na testirajućem skupu 84.3%, što je čak i lošije od iste arhitekture trenirane na manjem skupu podataka.

## 5. Zaključak

Ispitivanjem različitih arhitektura neuralnih mreža zaključili smo da sa postojećim kompjuterskim resursima upotrebom klasičnih konvolucionih neuronalnih mreža ne možemo postići tačnost veću od 85% na testirajućem skupu. Mogući su razni pristupi za postizanje veće tačnosti koje bismo podelili na one koji se odnose na podatke i one koje se odnose na neuronsku mrežu. Od pristupa prvog tipa izdvojili bismo upotrebu slika pune rezolucije 96x96 kao i kreiranje dodatnog skupa podataka iz postojećeg rotacijom i osnosimetričnim preslikavanjem (*data augmentation*). Što se tiče naprednijih arhitektura konvolucionih neuronskih mreža pokazano je da se *DenseCNN* kao i konvolucione neuronske mreže ekvivarijantne na rotaciju pokazuju bolje od klasičnih konvolucionih mreža [2] – čiji se rezultati kreću u opsegu 87.6% - 89.8%, što je uporedivo sa najboljim rezultatom dobijenim u ovom projektu koji iznosi 84.7%. Napomenimo da se u radu [2] koriste slike pune rezolucije 96x96x3 kao i *data augmentation*. Rezultati nisu upoređivani sa rezultatima ostalih takmičara na *kaggle* takmičenju kako ono nije bilo validno zbog postojanja javno dostupnih labela testirajućeg skupa na eksternom sajtu, što je rezultovalo velikog broja rešenja sa 100% tačnosti.



## 6. Literatura

- [1] Podaci - <https://www.kaggle.com/c/histopathologic-cancer-detection>
- [2] B. S. Veeling, J. Linmans, J. Winkens, T. Cohen, M. Welling. Rotation Equivariant CNNs for Digital Pathology. [arXiv:1806.03962](https://arxiv.org/abs/1806.03962)
- [3] Ehteshami Bejnordi et al. Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer. JAMA: The Journal of the American Medical Association, 318(22), 2199–2210. [doi:jama.2017.14585](https://doi.org/10.1001/jama.2017.14585)
- [4] Géron A. Hands on Machine Learning with Scikit-Learn and TensorFlow, O'Reilly, 2017