

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

DIPLOMSKI RAD br. 2016

**KLASIFIKACIJA HISTOPATOLOŠKIH SNIMAKA
DIJELOVA LIMFNIH ČVOROVA POMOĆU
STROJNOG UČENJA**

Domagoj Plušćec

Zagreb, lipanj 2019.

Zahvaljujem mentoru doc. dr. sc. Marku Čupiću i suradniku dr. sc. Tomislavu Lipiću s Instituta Ruđer Bošković koji su mi pomogli prilikom izrade ovog rada. Ujedno zahvaljujem Institutu Ruđer Bošković i Istraživačkom centru mladih na ustupljenoj računalnoj infrastrukturi za potrebe izrade ovog rada.

Sadržaj

| | |
|--------------------------------------------------------------------------|----|
| 1. Uvod..... | 1 |
| 2. Pregled dosadašnjih istraživanja..... | 2 |
| 3. Klasifikacija histopatoloških slika limfnih čvorova..... | 5 |
| 3.1 Klasifikacija slika | 5 |
| 3.2 Skup podataka | 6 |
| 3.3 Odabir pristupa rješavanja problema klasifikacije | 9 |
| 4. Duboko učenje u analizi histopatoloških slika | 10 |
| 4.1 Korišteni modeli | 11 |
| 4.1.1 Model <i>AlexNet</i> | 12 |
| 4.1.2 Model <i>ResNet</i> | 14 |
| 4.1.3 Model <i>DenseNet</i> | 16 |
| 4.1.4 Model <i>Inception-v3</i> | 18 |
| 4.2 Učenje prijenosom značajki | 19 |
| 4.3 Proširivanje skupa podataka | 21 |
| 5. Rješavanje problema klasifikacije histopatoloških slika | 23 |
| 5.1 Usporedba modela dubokog učenja | 23 |
| 5.2 Ispitivanje utjecaja tehnika proširivanja skupa podataka..... | 24 |
| 5.3 Vrednovanje modela dubokog učenja..... | 24 |
| 6. Rezultati | 27 |
| 6.1 Rezultati usporedbe modela dubokog učenja | 27 |
| 6.2 Rezultati ispitivanja utjecaja tehnika proširivanja skupa podataka . | 32 |
| 7. Zaključak..... | 34 |
| 8. Literatura..... | 36 |
| Sažetak | 41 |
| Summary..... | 42 |

1. Uvod

Računarstvo je promijenilo zdravstvene sustave iz temelja. Računala su se počela koristiti u administraciji u bolnicama, promijenila su način na koji liječnici bilježe pacijentove podatke sustavima kao što je e-karton u hrvatskom zdravstvenom sustavu, počeli su se koristiti za naručivanje lijekova, izdavanje recepata, a između ostalog i u analizi medicinskih slika i u potpomognutoj dijagnostici korištenjem algoritama strojnog učenja.

Računalni sustavi za analizu medicinskih slika omogućili su kvantitativni opis medicinskih slika što je bitno kako bi smanjili pogreške liječnika te kako bi ubrzali proces označavanja slika za potrebe dijagnoze. Također, računalne sustave u obradi medicinskih slika moguće je koristiti u istraživanju, primjerice za proučavanje bioloških mehanizama razvoja bolesti [11].

Cilj ovog rada je dati pregled područja analize histopatoloških slika i izrada programske implementacije za klasifikaciju histopatoloških slika limfnih čvorova pomoću strojnog učenja.

Ovaj diplomski rad strukturiran je kroz šest poglavlja. U ovom (uvodnom) poglavlju predstavljena je motivacija za istraživanje, proučavanje i primjenu računalnih metoda za analizu podataka iz medicinskih snimaka. Nakon motivacije predstavljen je cilj ovog diplomskog rada. U drugom poglavlju dan je pregled područja analize medicinskih slika korištenjem strojnog učenja s naglaskom na analizu histopatoloških slika. U trećem poglavlju opisan je konkretan problem koji će se rješavati, predstavljen je korišteni skup podataka te odabir pristupa rješavanju problema. U četvrtom poglavlju opisane su korištene metode dubokog učenja, podskupa strojnog učenja. Zatim su u petom poglavlju opisane korištene eksperimentalne metode. Rezultati provedenih eksperimenata dani su u šestom poglavlju.

2. Pregled dosadašnjih istraživanja

Analizu medicinskih slika prvenstveno su radili doktori na otisnutim radiološkim nalazima, odnosno na mikroskopima u analizi histopatoloških uzoraka [1]. Napretkom računala i omogućavanjem spremanja i učitavanja medicinskih slika u računalima omogućena je i njihova analiza korištenjem algoritama. Tako su se i algoritmi strojnog učenja počeli koristiti u analizi medicinskih slika već od sredine 60-tih godina prošlog stoljeća [2][3].

Sustavi potpomognute dijagnostike (engl. *computer aided diagnosis*, CAD) predstavljaju cjelovite sustave koji se mogu između ostalog koristiti i za analizu medicinskih slika kako bi pomogli doktoru u donošenju zaključaka temeljem nalaza pretraga i u konačnici u donošenju dijagnoze pacijentove bolesti. Primjenom analize medicinskih slika sustavi potpomognute dijagnostike su 80-tih godina prošlog stoljeća počeli pomagati doktorima u detekciji raka i dijagnostici na radiološkim slikama prsa [3]. Od tada analiza medicinskih slika se proširila i na druge vrste radioloških slika kao što su računalna tomografija (engl. *computed tomography*, CT), magnetska rezonancija (engl. *magnetic resonance imaging*, MRI) i ultrazvuk [3].

Razvojem digitalne mikroskopije omogućeno je skeniranje većih uzoraka, s različitim uvećanjem i s velikom rezolucijom slika. Jedno od takvih dostignuća je tehnika skeniranja čitavog preparata (engl. *whole slide imaging*, WSI) koja je postala tržišno dostupna od 2000. godine [1][22]. Prvo natjecanje koje je ponudilo sudionicima podatke u obliku slika dobivenih metodom WSI je bio *Camelyon16* [4]. Tipična veličina primjera iz skupa podataka *Camelyon16* je $200,000 \times 100,000$ slikovnih elemenata na najvećoj rezoluciji s tri bajta memorije za svaki slikovni element što predstavlja ukupno 55.88 GB podataka s jedne razine za jednog pacijenta [4].

U tablici 1 istaknute su razlike između radioloških i histopatoloških snimaka na konkretnim primjerima snimki te na temelju podataka preuzetih iz [11].

Kako je slike takve rezolucije teško ručno analizirati došlo je do potrebe za razvojem automatiziranih sustava za analizu takvih slika. Zadaci takvih sustava su potpomognuta dijagnoza, dohvaćanje sličnih slika te pronalaženje anomalija [1].

Tablica 1. Primjer razlika između radioloških i histopatoloških snimaka temeljena na podacima iz [11]

| | Radiološka snimka | Histopatološka snimka |
|-------------------|-------------------------------------------------------------|-------------------------------------------------------------------------------------------|
| Opis snimke | CT snimka visoke rezolucije prsnog koša | WSI biopsija prostate |
| Dimenzije snimke | $512 \times 512 \times 512$ | $15,000 \times 15,000$ |
| Vrsta informacije | vrijednosti predstavljaju nijanse sive boje | višespektralni podaci (<i>RGB</i> ili za svaku mjerenu valnu duljinu) |
| Svrha korištenja | određivanje prisutnosti raka i prostorne rasprostranjenosti | određivanje prisutnosti raka i stupanj diferenciranosti raka (engl. <i>cancer grade</i>) |

U okviru potpomognute dijagnoze ističu se dva smjera analize histopatoloških slika. Prvi smjer je usredotočen na klasifikaciju jednog ili više uzorka snimanog metodom *WSI* u kategoriju bolesti. Drugi smjer je detekcija ili semantička segmentacija regije od interesa (engl. *region of interest*, ROI). Primjeri istraživanja su segmentacija regije tumora, detekcija mitoze i određivanje rasprostranjenosti tumora [1].

Dohvaćanje sličnih slika se primjenjuje u edukacijske svrhe i za dijagnozu rijetkih slučajeva bolesti.

Pronalaženje anomalija je bitno za detekciju do sada neotkrivenih bolesti ili oblika postojećih bolesti. Kako se količina podataka sve više povećava, postalo je nužno da se koriste automatizirani sustavi temeljeni na strojnom učenju prilikom istraživanja novih bolesti.

Relevantan pregled metoda strojnog učenja koje se koriste prilikom analize histopatoloških slika može se dobiti analizom rezultata na natjecanju *Camelyon16* [6]. Prvi pristup istraživača je bio korištenje različitih statističkih i teksturalnih značajki uz klasifikator učen nadziranim učenjem. Primjeri korištenih značajki u okviru natjecanja su histogram lokalnih binarnih uzoraka, histogram pojavljivanja iste vrijednosti sive boje na određenoj udaljenosti i značajke *SIFT*.

Uz navedene značajke istraživači su za klasifikaciju često koristili sljedeće algoritme strojnog učenja: stroj s potpornim vektorima (engl. *support vector machine*, SVM), algoritam slučajnih šuma (engl. *random forest*) i algoritam *AdaBoost*.

Drugi pristup je bio korištenje metoda dubokog učenja. Često korištene arhitekture modela su bile *GoogLeNet* [40], *ResNet* [9], *VGG-16* [39] i *AlexNet* [12][13].

U okviru natjecanja metode zasnovane na dubokom učenju pokazale su se kao najuspješnije pri čemu su zauzele prvih 19 mjesta.

3. Klasifikacija histopatoloških slika limfnih čvorova

Histologija je grana biologije čiji je cilj proučavanje tkiva biljnog i životinjskog podrijetla od stanične i međustanične razine do organa. Patološka histologija (histopatologija) je grana histologije koja proučava građu bolesnih tkiva [7]. Predmet proučavanja ovog rada su histopatološke slike limfnih čvorova. Limfni čvorovi su dio limfnog sustava, a služe za filtriranje limfe od mikroorganizama i stranih antigena [8].

Cilj ovog poglavlja je predstaviti zadatak klasifikacije slika, korišteni skup podataka te odabir pristupa rješavanju problema.

3.1 Klasifikacija slika

Pristupe strojnom učenju možemo podijeliti na nadzirano učenje, slabo nadzirano učenje, nenadzirano učenje i podržano učenje [15].

Zadatak ovog rada, klasifikacija, se svrstava u nadzirano učenje. Nadzirano učenje je oblik strojnog učenja u kojem su dostupni podaci i njihove oznake. Kada su oznake diskretne i konačne govorimo o klasifikaciji, a kada su oznake kontinuirane govorimo o regresiji [16].

Formalno zadatak nadziranog učenja možemo definirati kao učenje preslikavanja $f: X \rightarrow Y$ iz skupa primjera za učenje D . Skup primjera za učenje D sastoji se od parova primjera i pripadnih oznaka, $D = \{(x^{(i)}, y^{(i)})\}_{i=1}^N$ gdje je N ukupni broj primjera za učenje. X predstavlja skup svih mogućih primjera, a pojedini primjer x iz X možemo definirati kao vektor značajki $x = (x_1, x_2, \dots, x_n)^T$ pri čemu je n broj značajki, odnosno dimenzija vektora. Y predstavlja prostor mogućih oznaka [15][16].

Binarna klasifikacija, zadatak ovog rada, je vrsta klasifikacije u kojoj se ulazni primjeri svrstavaju u dva razreda pri čemu možemo definirati skup oznaka $Y = \{0, 1\}$.

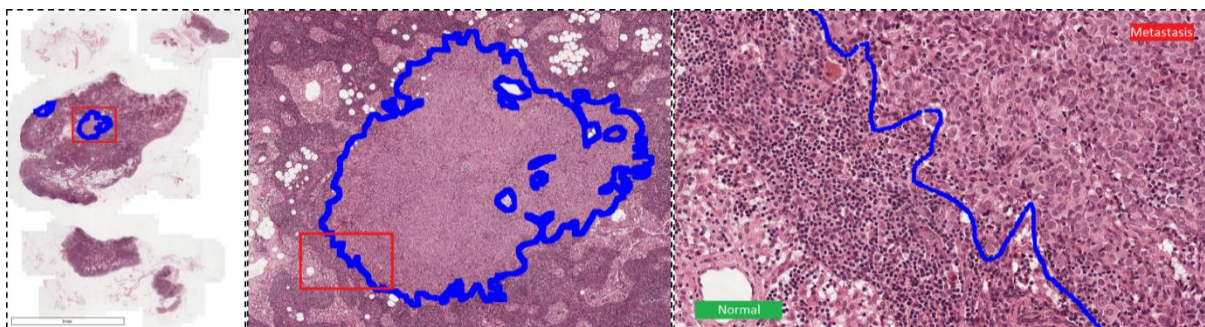
Značajke primjera za učenje prilikom klasifikacije slika mogu biti vrijednosti boja slikovnih elemenata, no moguće je i koristiti drugačije značajke. Problem s izravnim korištenjem vrijednosti slikovnih elemenata je velika dimenzionalnost prostora značajki, dok s korištenjem izvedenih značajki možemo postići manju dimenzionalnost s očuvanom informacijom potrebnom za klasifikaciju. Primjeri korištenih značajki su histogram orijentiranih gradijenata (engl. *histogram of oriented gradients*, HOG),

histogram pojavljivanja iste vrijednosti sive boje na određenoj udaljenosti (engl. *grey level co-occurring matrix*, GLCM), histogram lokalnih binarnih uzoraka (engl. *local binary patterns*, LBP) [1], značajke *SIFT* (engl. *scale-invariant feature transform*) [6] te razne informacije o geometrijskim strukturama na slikama [17]. Navedene značajke su korištene u analizi medicinskih slika, međutim kako su modeli dubokog učenja postigli bolje performanse od prijašnjih modela strojnog učenja danas se često umjesto ručnog oblikovanja značajki prepušta modelu dubokog učenja da nauči izdvojiti značajke izravno iz vrijednosti slikovnih elemenata [1].

3.2 Skup podataka

U okviru ovog rada korišten je skup podataka *PatchCamelyon (PCam)* opisan u radu Bastiaan S. Veeling et al., „*Rotation Equivariant CNNs for Digital Pathology*“ [5]. Skup podataka je nastao na temelju skupa podataka *Camelyon16* [6].

Originalni skup podataka *Camelyon16* [6] se sastoji od 399 slika limfnih čvorova snimanih metodom *WSI* s oznakama na razini slikovnih elemenata. Prilikom snimanja tkiva koristila se metoda bojanja *hemalaun-eozin* (engl. *hematoxylin and eosin*, H&E) kako bi se istaknuli dijelovi limfnih čvorova. Hemalaun ističe jezgre stanica plavom bojom, a eozin boji citoplazmu i spojno tkivo ružičastom bojom [11]. Jedan slikovni element predstavlja uzorak površine $0.243\mu m \cdot 0.243\mu m$. Svaki slikovni element sadrži tri kanala zapisanih u 8 bita po kanalu. Primjer slike preuzete iz skupa podataka *Camelyon16* [6] s oznakom tumorskog tkiva na tri različite razine povećanja dan je na slici 1.

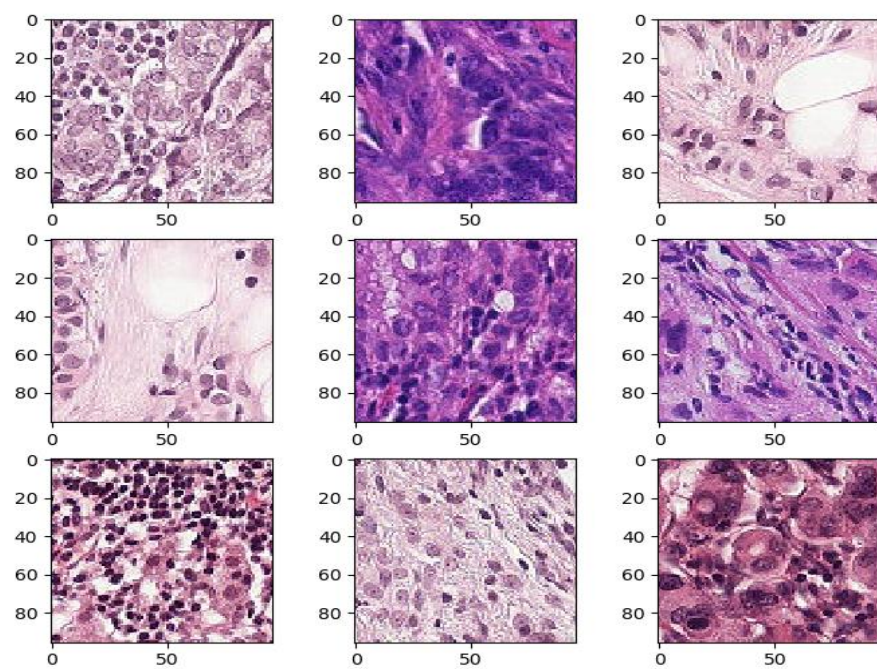


Slika 1. Primjer slike histopatološkog uzorka slikanog metodom WSI preuzet iz Camelyon16 [6] skupa podataka s oznakom tumorskog tkiva na tri razine uvećanja. Želja autora skupa podataka PCam [5] je bila napraviti skup podataka za klasifikaciju koji bi služio kao standardizirani skup podataka za ocjenjivanje modela (engl. *benchmark dataset*) strojnog učenja na medicinskim slikama. Slični skupovi podataka su napravljeni za klasifikaciju prirodnih slika kao što su CIFAR-10 [34], PASCAL VOC [35], ImageNet [33] te primjerice skup podataka za klasifikaciju rukom pisanih znamenki MNIST [38].

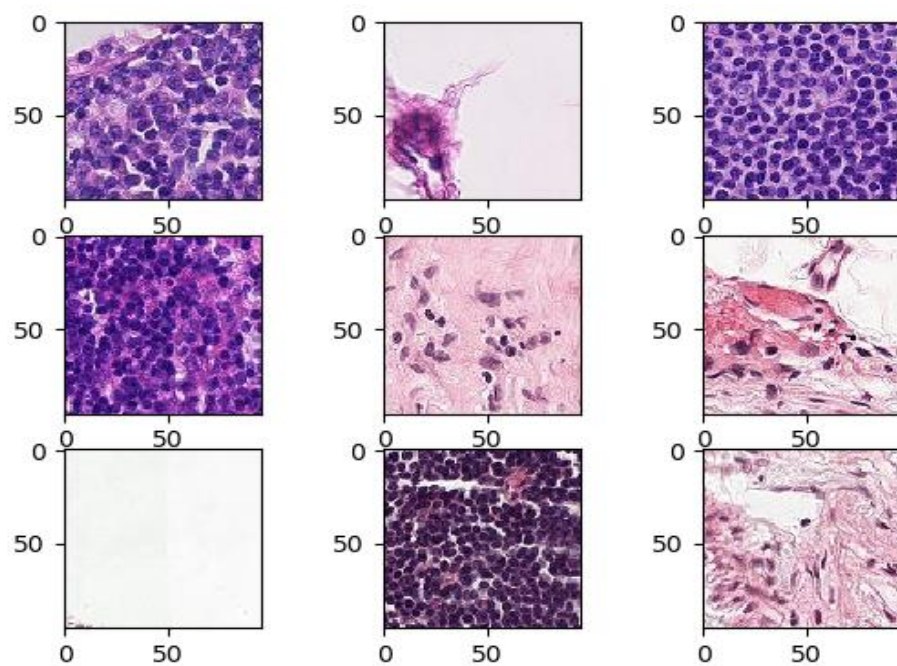
Skup podataka je podijeljen u skup za treniranje, validaciju i testiranje u sljedećim postotnim omjerima: 75%, 12.5% i 12.5%. Ukupno skup podataka sadrži 327,680 slika. Skup podataka i podjele po podskupovima za treniranje, validaciju i testiranje su balansirani što znači da svaki od podskupova ima jednak broj pozitivnih i negativnih primjera. Slike u skupu podataka PCam se sastoje od 96×96 slikovnih elemenata pri čemu su informacije za svaki slikovni element iste kao i kod originalnog skupa podataka.

Oznake primjera u skupu podataka su binarne. Pozitivna oznaka je dodijeljena slikama koje u regiji od 32×32 slikovna elementa u sredini sadrže barem jedan slikovni element koji pripada tumorskom tkivu. Takva definicija pozitivnog primjera je napravljena kako bi se omogućilo korištenje dubokih konvolucijskih modela koji ne koriste ispunjavanje rubova slike nulama (engl. *zero padding*) kako bi osigurali konzistentno ponašanje prilikom primjene na cjelokupne slike dobivene metodom WSI. Primjer pozitivnih primjera dan je na slici 2, a negativnih na slici 3.

Podaci u PCam skupu podataka su spremeni u formatu HDF 5 (engl. *Hierarchical Data Format*) koji omogućava pohranu i organizaciju velikih količina podataka [10].



Slika 2. Primjer histopatoloških slika iz skupa podataka PCam [5] koji sadrže tumor



Slika 3. Primjeri histopatoloških slika iz skupa podataka PCam [5] koji ne sadrže tumor

3.3 Odabir pristupa rješavanja problema klasifikacije

U okviru poglavlja 2 predstavljeni su primjeri pristupa analize medicinskih slika na natjecanju *Camelyon16*. Pristupi rješavanju problema klasifikacije u okviru ovog diplomskog rada ograničeni su na modele dubokog učenja koji su detaljno objašnjeni u poglavlju 4.

Motivacija za takav odabir je odličan rezultat modela dubokog učenja na natjecanjima *Camelyon16* [6] i *Camelyon17* [4][31] te povećani interes istraživača za primjenu dubokih modela u području analize medicinskih slika [3]. Povećani interes istraživača se može vidjeti na slici 4. koja prikazuje broj recenziranih objavljenih radova koji spominju metode dubokog učenja i analizu medicinskih slika kroz godine. Slika je napravljena po uzoru na [3], no uključuje i radove iz područja histopatologije.



Slika 4. Broj recenziranih radova u analizi medicinskih slika koji spominju duboko učenje. Radovi su pretraživani pomoću baze podataka PubMed [26] koristeći sljedeći kriterij: ("deep learning" OR "deep neural network" OR "deep convolution" OR "convolutional neural network" OR "shift-invariant artificial neural network" OR "MTTANN") AND (radiography OR x-ray OR mammography OR CT OR MRI OR PET OR ultrasound OR therapy OR radiology OR MR or mammogram OR SPECT OR histopathology). Za 2019. godinu pretraga uključuje radove objavljene do 1.6.2019 te su zatim rezultati linearno prošireni. Slika je napravljena po uzoru na [3].

4. Duboko učenje u analizi histopatoloških slika

Duboko učenje je grana strojnog učenja temeljena na razvoju višeslojnih neuronskih mreža koje omogućavaju izdvajanje značajki na visokom stupnju apstrakcije iz ulaznih podataka [3]. Algoritmi dubokog učenja, posebice konvolucijske neuronske mreže, su u zadnjih nekoliko godina postale najčešća metoda u analizi medicinskih slika [4].

Prednosti konvolucijskih modela za razliku od potpuno povezanih neuronskih mreža i nekih drugih modela strojnog učenja poput SVM-a (engl. *support vector machine*) je taj što oni uzimaju u obzir raspored slikovnih elemenata i njihovo susjedstvo. Duboki konvolucijski modeli su zbog toga izrazito pogodni za podatke s hijerarhijskom strukturom. Na primjeru prepoznavanja čovjeka na slici prvo možemo prepoznati rubove, pa vidjeti koji rubovi tvore nos, oči, rub lica zatim cjelokupno lice te u konačnici možemo klasificirati lice u određenu klasu. Pretpostavka hijerarhijske strukture podataka pokazala se kao dobra pretpostavka na problemima računalnog vida [18].

Modeli dubokog učenja pokazali su se odličnim izborom za analizu slika na natjecanjima kao što su *ImageNet* [33] i *Pascal VOC* [35] te u analizi medicinskih slika na natjecanjima poput *Camelyon16* [6], *Camelyon17* [4][31], *BraTS* [36][37] i *BreakHis* [32]. Osim uspješnosti koje su duboki modeli pokazali, imaju i prednost što su omogućili automatsko oblikovanje značajki. Zbog toga više nije potrebno veliko domensko znanje specifičnog područja kako bi se mogao razviti prediktivni model.

Međutim bitno je i istaknuti neka od ograničenja koje duboki modeli imaju s naglaskom na analizu medicinskih slika. Prvi problem je što je za učenje dubokih modela potrebna velika količina podataka [1]. Specifičnost prikupljanja medicinskih skupova podataka je da je potrebno da se dobije pristanak pacijenata za sudjelovanje u istraživanju te je potrebno anonimizirati podatke kako se ne bi na temelju nalaza mogao odrediti identitet pacijenta. Takvi problemi ujedno ograničavaju i javnu dostupnost medicinskih podataka. Kako je većina primjene dubokih modela nadzirano učenje uz podatke je potrebno imati i oznake podataka. Zbog navedenog ističe se potreba za označavanjem velikog broja kompleksnih podataka za čije je označavanje potrebno stručno znanje eksperata. Drugi problem je interpretabilnost dubokih modela koja je značajno manja u odnosu na ručno modelirane značajke. Zbog toga se danas sve više istražuju različiti

načini kako omogućiti interpretabilnost dubokih modela. Zainteresirani čitatelj može više pronaći u preglednim radovima [23] i [24].

Treći problem je pronalaženje optimalnih hiperparametara. Primjeri hiperparametara u dubokim modelima su stopa učenja, odabir aktivacijske funkcije, stopa isključivanja čvorova. Promjene u hiperparametrima mogu značajno utjecati na rezultat modela, a za validaciju skupa hiperparametara je potrebno velika količina vremena. Trenutno se provode istraživanja na području automatskog određivanja hiperparametara za modele dubokog učenja [25].

U okviru ovog poglavlja opisane su prednosti i ograničenja dubokih modela prilikom primjene u analizi histopatološkim slikama. Opisane su korištene arhitekture dubokih modela (poglavlje 4.1.), učenje modela korištenjem metode prijenosa značajki (poglavlje 4.2) te metode proširivanja skupa podataka kako bi se poboljšala generalizacijska sposobnost modela (poglavlje 4.3).

4.1 Korišteni modeli

U okviru ovog rada korišteno je nekoliko arhitektura dubokih modela koji su postali popularni u području dubokog učenja. Sažeti pregled korištenih modela i broja njihovih parametara dan je u tablici 2.

Tablica 2. Tablični prikaz korištenih modela, njihovog broja parametara i broja slojeva s aktivacijskom funkcijom

| Naziv modela | Broj parametara | Broj slojeva s aktivacijskom funkcijom |
|---------------------|-----------------|----------------------------------------|
| <i>AlexNet</i> | 57,877,824 | 8 |
| <i>ResNet18</i> | 11,186,645 | 18 |
| <i>DenseNet121</i> | 6,954,881 | 121 |
| <i>Inception-v3</i> | 27,161,264 | 98 |

4.1.1 Model AlexNet

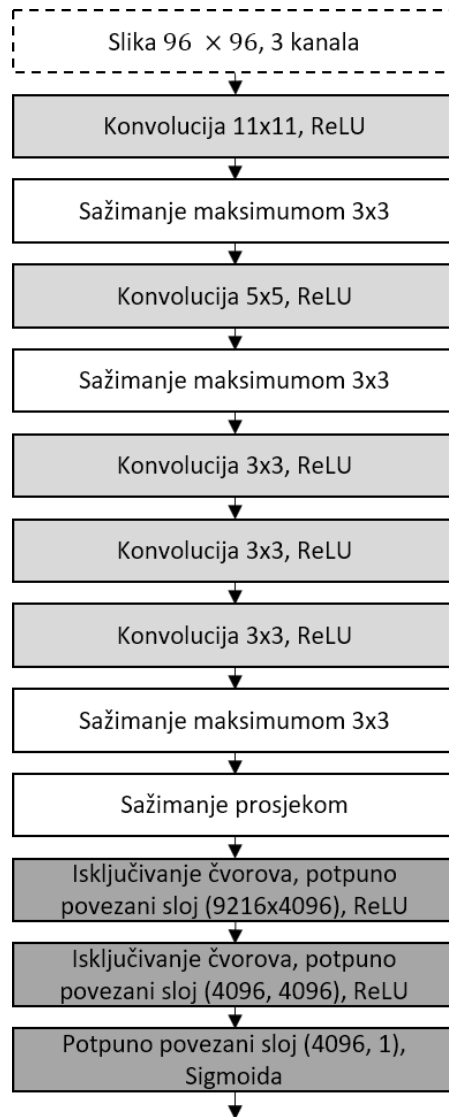
U okviru ovog rada korišten je model *AlexNet* koji je definiran u radu A. Krizhevsky, „*One weird trick for parallelizing convolutional neural networks*“ [13] koji je modifikacija originalnog modela opisanog u A. Krizhevsky et al., „*ImageNet Classification with Deep Convolutional Neural Networks*“ [12]. Korišteni model i parametri modela treniranog na skupu podataka *ImageNet* dostupni su u radnom okviru *PyTorch* [14].

Model se sastoji od osam slojeva s parametrima koji se mogu učiti pri čemu su pet slojeva konvolucijskog tipa, a tri potpuno povezana. Kao aktivacijska funkcija koristi se $ReLU(x) = \max(0, x)$ u svim slojevima osim u zadnjem u kojem se koristi sigmoidalna prijenosna funkcija $\sigma(x) = (1 + e^{-x})^{-1}$.

To je različito od dotadašnjih modela konvolucijskih neuronskih mreža koje su uglavnom koristile $f(x) = \tanh(x)$ i $\sigma(x) = (1 + e^{-x})^{-1}$ kao aktivacijske funkcije.

Model je različit od dotadašnjih modela i zbog veće dubine u usporedbi s primjerice modelom *LeNet5* koji je korišten 90-tih godina 20. stoljeća na skupu podataka *MNIST* koji je imao 5 slojeva, također *AlexNet* je imao čak deset puta više konvolucijskih kanala [18]. Kako bi omogućili treniranje modela toliko većeg kapaciteta autori su odlučili osim kažnjavanja težina normom $L2$ koristiti i tehniku isključivanja čvorova (engl. *dropout*) za regularizaciju. Kako bi omogućili dovoljnu količinu podataka za treniranje autori su u okviru svog rada koristili tehnike povećanja skupa podataka poput transliranja slike, horizontalnog zrcaljenja i mijenjanje intenziteta boja na slikama.

Arhitektura modela grafički je prikazana na slici 5.



Slika 5. Grafički prikaz arhitekture korištenog modela AlexNet iz radnog okvira PyTorch [14]

4.1.2 Model *ResNet*

Istraživači iz istraživačkog odjela korporacije Microsoft su 2015. godine osmislili arhitekturu neuronske mreže *ResNet* [9] koja je omogućila treniranje modela s više slojeva. Prethodna istraživanja su pokazala kako je dubina neuronske mreže važan parametar koji omogućava bolje rezultate. Problem koji se pojavljivao prilikom treniranja dubljih mreža je bio problem nestajućeg gradijenta. Gradijent se između svakog sloja značajno smanjivao jer se gradijent pogreške po težini u trenutnom sloju množi s umnoškom gradijenata pogreške svih idućih slojeva.

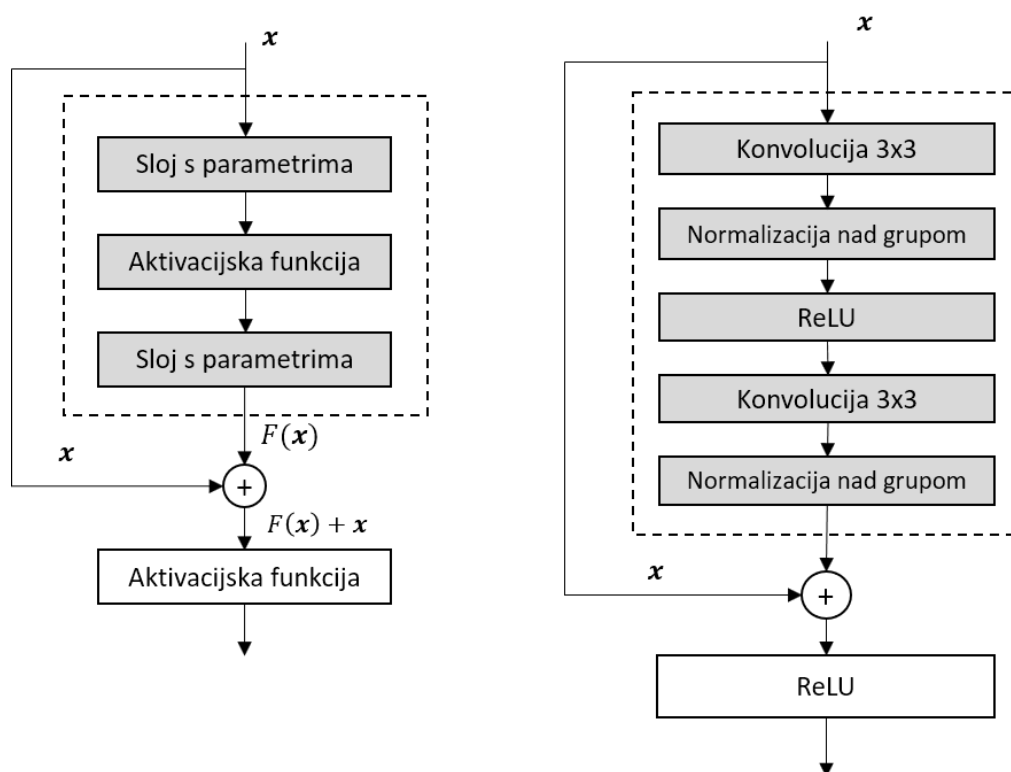
Model je prvotno testiran na skupovima podataka za klasifikaciju prirodnih slika *MS Coco* i *ImageNet* te je osvojio prvo mjesto na natjecanju *ILSVRC 2015* [9].

Novost koju model donosi je učenje rezidualnih preslikavanja. Cilj je naučiti funkciju razlike $F(x) = H(x) - x$ pri čemu $H(x)$ označava očekivani izlaz, a x ulaznu vrijednost. Učenje rezidualnog preslikavanja se zatim postiže tako da uvedemo veze koje preskaču slojeve (engl. *shortcut connections*) i zbrajaju ulaz u slojeve s izlazom iz slojeva. Takav niz slojeva s vezom koja preskače slojeve zovemo rezidualni blok. Matematički izlaz rezidualnog bloka y možemo predstaviti sljedećim izrazom.

$$y = F(x, \{W_i\}) + x$$

U izrazu W_i označava parametre funkcije razlike, a x ulazne podatke u rezidualni blok.

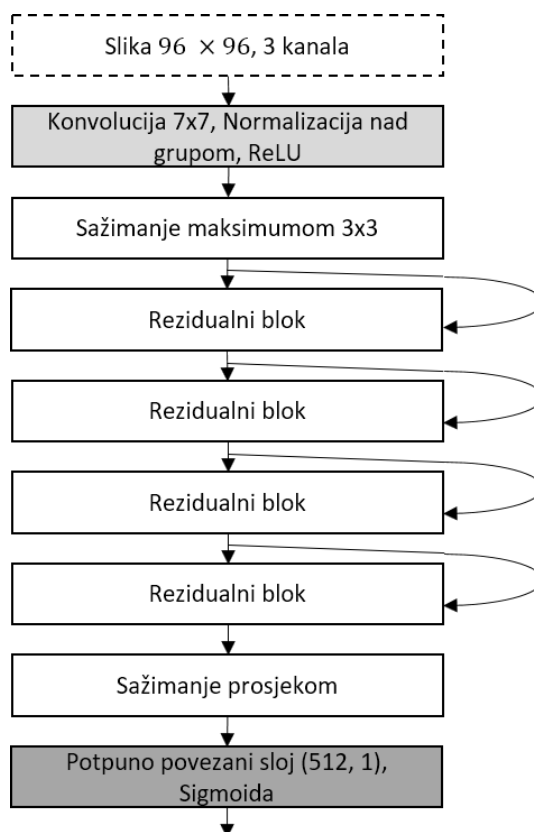
Rezidualni blok kakav su autori predložili te implementacija unutar radnog okvira *PyTorch* [14] su grafički prikazani na slici 6.



Slika 6. Grafički prikaz osnovnog rezidualnog bloka (desno) i implementacije rezidualnog bloka unutar radnog okvira PyTorch napravljenog temeljem [9], [14] i [18] Ako unutar rezidualnog bloka dolazi do promjene dimenzije potrebno je projicirati ulazne podatke u prostor koji odgovara dimenzijama izlaza kako bi se mogli zbrojiti.

U usporedbi s modelom *AlexNet* ovaj model ne koristi tehniku isključivanja čvorova, međutim koristi normalizaciju nad grupama (engl. *batch normalization*). Normalizacija nad grupama se radi na način da se izlaz sloja normalizira na temelju izlaza svih uzoraka za učenje unutar jedne mini-grupe te se zatim na te uzorke primjeni afina transformacija s parametrima za skaliranje i pomak [41]. Parametri skaliranja i pomaka se zatim također uče kao i ostali parametri modela. Nakon što učenje završi vrijednosti koje se koriste za normalizaciju primjera za testiranje računaju se na temelju čitavog skupa za učenje. Normalizacija nad grupama kao i tehnika isključivanja čvorova pokazuje regularizirajući efekt. Korištenjem tehnike normalizacije nad grupama mreža se lakše uči.

Arhitektura korištenog modela je grafički prikazana na slici 7.



Slika 7. Grafički prikaz arhitekture korištenog modela ResNet iz radnog okvira PyTorch [14]

4.1.3 Model DenseNet

U okviru ovog rada korišten je model *DenseNet* opisan u radu G. Huang et al. „*Densely Connected Convolutional Networks*“ [28] implementiran u radnom okviru *PyTorch* [14]. Želja autora rada je bila da se iskoriste sve mape značajki (engl. *feature maps*) izračunate do nekog mjesta u mreži prilikom izračunavanja novih mapa značajki. Motivacija iza takvog modela je bila što mreže koje uvode veze između slojeva prilikom početka mreže i kraja mreže postižu odlične rezultate. Primjeri takvih modela su *ResNet* [9] i *Highway* [30].

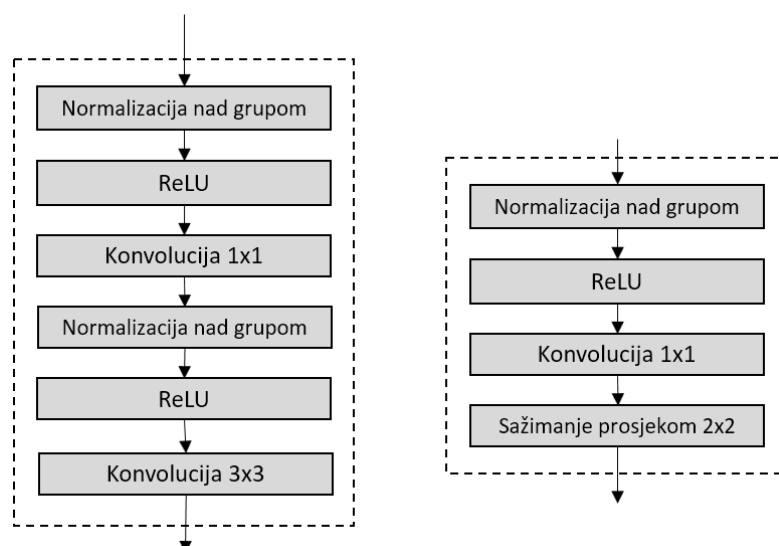
Poput rezidualnih blokova korištenih u modelu *ResNet*, u modelu *DenseNet* koriste se gusto povezani blokovi (engl. *dense block*) unutar kojih se mape značajki propagiraju na način da se izlaz gusto povezanog sloja (engl. *dense layer*) unutar bloka predaje na ulaz svim idućim gusto povezanim slojevima. Ulazi iz različitih slojeva se spoje u jednu mapu značajki, pri čemu je bitno da se unutar bloka ne mijenjaju dimenzije mapa

značajki. Grafički prikaz gusto povezanog sloja dan je na slici 8, a gusto povezanog bloka na slici 9.

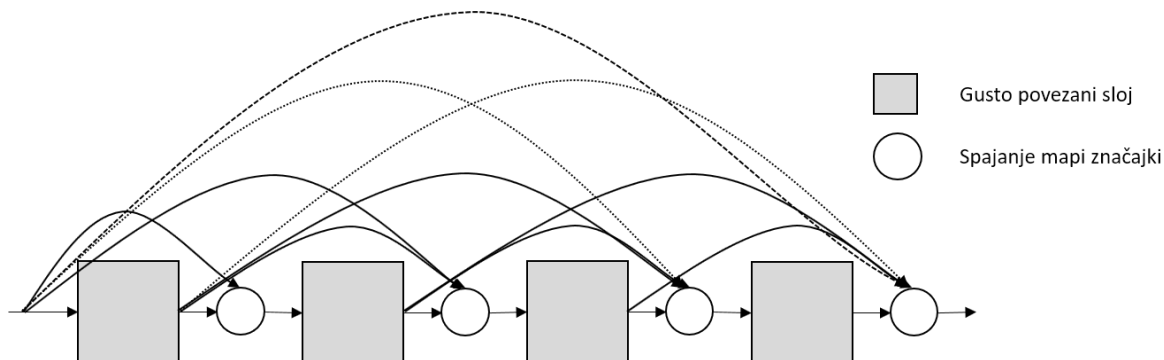
Model se zatim gradi povezivanjem gusto povezanih blokova pomoću prijelaznih blokova (engl. *transition block*) koji se sastoje od konvolucijskog i sažimajućeg sloja. Prijelazni blok služi za smanjenje dimenzionalnost mapi značajki unutar mreže. Grafički prikaz prijelaznog bloka dan je na slici 8.

Analizom tablice 2 može se primijetiti kako uvođenjem dodatnih veza nije povećan broj parametara već je on u konačnom modelu i manji od primjerice modela *ResNet*, a sadrži veći broj aktivacijskih slojeva.

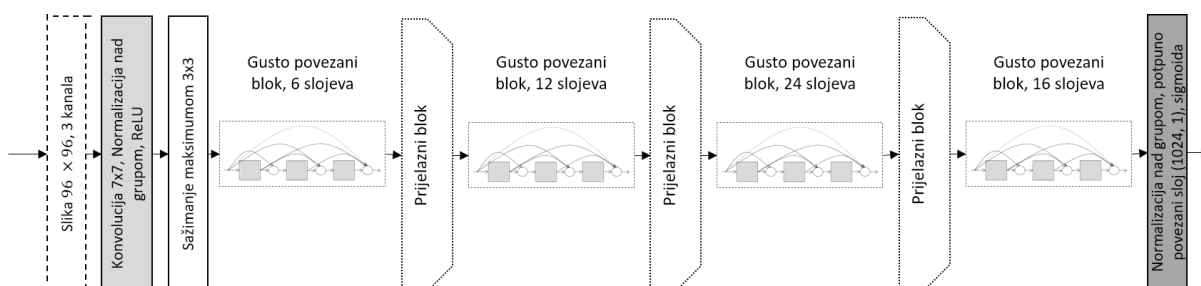
Arhitektura korištenog modela grafički je prikazana na slici 10.



Slika 8. Grafički prikaz gusto povezanog sloja (lijevo) i prijelaznog bloka (desno) temeljem implementacije u radnom okviru PyTorch [14]



Slika 9. Grafički prikaz gusto povezanog bloka s četiri gusto povezana sloja



Slika 10. Grafički prikaz korištene arhitekture DenseNet napravljen po uzoru na prikaz u izvornom radu [28] i temeljem implementacije u radnom okviru PyTorch [14] Arhitektura DenseNet je korištena i u radu koji je predstavio skup podataka PCam [5] kao osnovni model pomoću koje se gradila kompleksnija arhitektura koja je bila invarijantna na rotaciju i zrcaljenje.

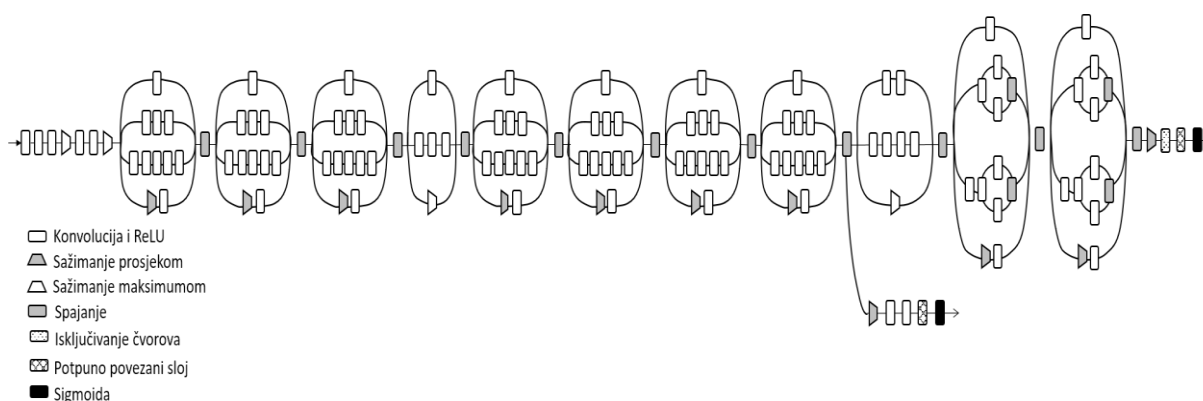
4.1.4 Model Inception-v3

U okviru ovog rada korišten je model *Inception-v3* opisan u radu C. Szegedy et al, „*Rethinking the Inception Architecture for Computer Vision*“ [45] implementiran u radnom okviru PyTorch [14]. Model je nadogradnja *GoogLeNeta* koji je bio često korišten na natjecanju *Camelyon16* [6].

Jedna od novosti koje donosi ova arhitektura dubokog modela je zamjena slojeva koji sadrže veće konvolucijske filtere s više konvolucijskih slojeva koji sadrže manje filtere. Primjerice umjesto jednog sloja s filterima dimenzije 5×5 koriste dva sloja s filterima dimenzije 3×3 te na taj način smanjuju korišteni broj parametara. Autori u model uvode i korištenje asimetričnih konvolucijskih filtera kojima značajno smanjuju broj parametara i vrijeme potrebno za predikciju, međutim pokazalo se kako stavljanje takvih slojeva u početne slojeve mreže pogoršava rezultate [45].

Kako bi omogućili bolju propagaciju pogreške u ranije slojeve uvode dodatni klasifikator. Dodatni klasifikator uzima značajke iz ranijih slojeva te pokušava predvidjeti oznaku ulaznog primjera. Na taj način se u ranije slojeve propagira pogreška od oba klasifikatora.

Arhitektura modela grafički je prikazana na slici 11 koja je prilagođena iz [46].



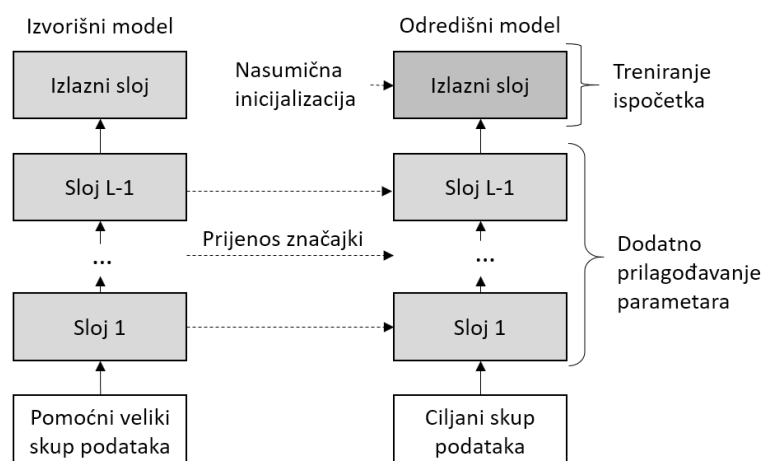
Slika 11. Grafički prikaz modela Inception-v3 prilagođen iz [46]

4.2 Učenje prijenosom značajki

Modeli dubokog učenja su velikog kapaciteta zbog velikog broja parametara. Zbog velikog kapaciteta modela potreban je i veliki skup primjera za učenje kako bi model mogao naučiti dobro generalizirati, odnosno kako bi se izbjegnula situacija u kojoj bi model zapamtio primjere iz skupa podataka za učenje. Jedan od mogućih pristupa je prikupljanje dodatnih podataka međutim to je često nemoguće zbog vremenskih i novčanih ograničenja. Drugi pristup je umjetno povećanje skupa podataka opisano u poglavlju 4.3. Metoda koja je opisana u ovom poglavlju zasniva se na prenošenju značajki (engl. *transfer learning*).

Prenošenje značajki je metoda u kojoj se model uči na velikom skupu podataka koji je sličan skupu podataka na kojem se želi naučiti model. Zatim se napravi novi model kojem se parametri svih slojeva, osim zadnjih nekoliko, postave na vrijednosti prethodno naučenog modela. Zadnji slojevi novog modela se inicijaliziraju s nasumičnim težinama zbog toga što su njihovi parametri specifični za konkretni zadatak na odabranom skupu podataka te zato što se broj izlaza modela može razlikovati od modela kojeg smo učili na drugom skupu podataka. U okviru ovog

diplomskog rada samo je zadnji sloj inicijaliziran nasumičnim težinama prilikom učenja prijenosom značajki. Nakon prenošenja značajki konačni model možemo učiti na više načina. Primjerice model možemo učiti tako da učimo parametre samo zadnjeg sloja modela što se savjetuje u slučajevima kada je skup podataka u zadatku malen i sličan velikom skupu podataka na kojem se učio pomoćni model. Drugi način predlaže da se uči čitav model pri čemu je prijenos značajki inicijalizacija parametara koja se pokazuje boljom od nasumične inicijalizacije. Navedene mogućnosti su dvije krajnosti, no istraživači su pokušali i razne druge tehnike koje zainteresirani čitatelj može pronaći u [17]. Grafički prikaz učenja prijenosom značajki je pokazan na slici 12 koja je nastala na temelju [18].



Slika 12. Grafički prikaz učenja prijenosom značajki napravljen temeljem [18]

Dodatna prednost učenja prijenosom značajki je kraće vrijeme učenja modela zbog brže konvergencije.

U području računalnog vida često se koriste duboki modeli učeni prijenosom značajki s modela učenog na skupu podataka *ImageNet*. Kako je u području analize medicinskih slika malen broj javno dostupnih velikih skupova podataka teško je pronaći prethodno trenirane modele. Jedan od izvora je *NiftyNet* [27], međutim on nudi modele koji su specifičniji za semantičku segmentaciju slika te su trenirani na radiološkim slikama. Razlike radioloških i histopatoloških slika opisane su u 2. poglavlju.

U radu N. Tajbakhsh, et al. „*Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning*“ [17] su pokazali kako prijenos značajki s modela učenog na skupu podataka *ImageNet* na medicinsku domenu može dovesti do boljih

rezultata modela u usporedbi s modelom koji je koristio nasumičnu inicijalizaciju parametara. U radu su koristili arhitekturu duboke neuronske mreže *AlexNet* te su je primijenili na četiri različita zadatka u okviru obrade medicinskih slika: detekcija polipa u kolonoskopskim video zapisima, vrednovanje kvalitete slike u kolonoskopskim video zapisima, detekcija plućnih edema u slikama dobivenim računalnom tomografijom i segmentacija granice krvnih žila koristeći slike dobivene ultrazvukom.

Pregledni rad Komura, D i Ishikawa, S. „*Machine Learning Methods for Histopathological Image Analysis*“ [1] navodi kako se prijenos značajki s modela učenog na skupu podataka *ImageNet* pokazao povoljnim i na histopatološkim slikama te da su neki modeli pokazali do tada najbolje performanse (engl. *state-of-the-art performance*) na zadacima klasifikacije i segmentacije histopatoloških slika mozga i debelog crijeva.

Znanstveni rad na temu učenja značajki [29] na histopatološkim slikama namijenjenim za detekciju raka grudiju je također pokazao kako su dobili bolje rezultate prijenosom značajki s modela učenog na skupu podataka *ImageNet*.

U okviru ovog rada bit će istražen i utjecaj prijenosa značajki za zadatak klasifikacije histopatoloških slika limfnih čvorova.

4.3 Proširivanje skupa podataka

Metode za proširivanje skupa podataka umjetnim primjerima (engl. *data augmentation*) koriste se kako bi se povećao skup podataka s ciljem bolje generalizacije modela te posebice za postizanje otpornosti na transformacije poput rotacije, skaliranja, dodavanje šuma i sl.

Proširenje skupa podataka se postiže tako da se na podatke za učenje iz originalnog skupa podataka primjene željene transformacije.

Histopatološke slike limfnih čvorova imaju svojstva invarijantnosti na translaciju, rotaciju i zrcaljenje [5]. Translacijska invarijantnost označava da ako se predmet na slici pomakne, a da se pritom sam ne promjeni, njegova oznaka klase ostaje ista. Analogno invarijantnost na rotaciju označava da ako se predmet na slici rotira njegova oznaka klase ostaje ista. Invarijantnost na zrcaljenje označava da se predmet može

zrcaliti te da to neće promijeniti njegovu pripadnost klasi. Invarijantnost na zrcaljenje je uzrokovana time što se uzorak može staviti u uređaj za skeniranje na dvije strane.

U radu B. S. Veeling et al. „*Rotation Equivariant CNNs for Digital Pathology*“ [5] korištene su transformacije skaliranja, rotacije i zrcaljenja te je uspoređen model koji je učen s takvim transformacijama s modelom koji je otporan na rotaciju i zrcaljenje ulaznih podataka. Model koji su koristili u radu unutar arhitekture ima ugrađenu pristranost da podaci mogu biti rotirani i zrcaljeni čime koriste dodatne parametre. Zainteresirani čitatelj detaljnije razmatranje takvih modela može pronaći u radu [43] i [44]. Rezultati koje su pokazali upućuju na to da model s takvom pristranošću i sa sličnim brojem parametara može biti bolji u klasifikaciji histopatoloških slika od modela za kojeg su prilikom učenja korištene metode proširivanja skupa podataka.

Osim navedenih transformacija moguće je koristiti i transformacije boje tako da se malom nasumičnom vrijednošću promjeni iznos nijanse boje (engl. *hue*), iznos zasićenja boje (engl. *saturation*), svjetline (engl. *brightness*) i kontrasta (engl. *contrast*) [1].

5. Rješavanje problema klasifikacije histopatoloških slika

U okviru ovog poglavlja opisani su napravljeni eksperimenti, koji su parametri ispitani te kako su vrednovani rezultati (poglavlje 5.3.).

5.1 Usporedba modela dubokog učenja

Prilikom provođenja eksperimenata korišteni su modeli opisani u poglavlju 4.1. Za učenje modela korišten je optimizacijski algoritam *ADAM* sa stopama eksponencijalno umanjujućih momenata $\rho_1 = 0.9$, $\rho_2 = 0.999$. Prilikom učenja optimiran je gubitak unakrsne entropije (engl. *cross entropy loss*). Korištena veličina grupe za učenje (engl. *batch size*) je iznosila 32. Veličina grupe za učenje je određena temeljem preporuke u radu D. Masters i C. Luschi „*Revisiting small batch training for deep neural networks*“ [19] i temeljem dostupne grafičke memorije.

Tijekom usporedbe modela dubokog učenja ispitani su sljedeći parametri.

- Mijenjana je stopa učenja modela kako bi se optimizirali rezultati pojedinog modela.
- Ispitan je utjecaj inicijalizacije i načina treniranja modela. Prvi način je inicijalizacija modela koristeći metodu opisanu u radu K. He et al, „*Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification*“ [21] i učenje ispočetka. Drugi način je inicijalizacija modela prijenosom značajki s modela učenog na skupu podataka *ImageNet* i treniranje čitave mreže. Treći način poput drugog načina inicijalizira model prijenosom značajki, no uči samo zadnje potpuno povezane slojeve modela.

Učenje modela ispočetka nije napravljeno za arhitekturu *Inception-v3* zbog ograničenja u računalnim resursima.

Odabir najboljih parametara modela tijekom učenja je napravljen tako da su zapamćeni parametri modela s kojima je model na skupu podataka za validaciju postignuo najbolju mjeru *F1*.

5.2 Ispitivanje utjecaja tehnika proširivanja skupa podataka

Kako bi se proširio skup podataka za učenje korištene su sljedeće transformacije: rotacija slike za nasumični kut, vertikalno zrcaljenje slike, horizontalno zrcaljenje slike, promjena iznosa nijanse boje, promjena iznosa zasićenja boje, promjena iznosa svjetline i promjena iznosa kontrasta.

Odabran je model *DenseNet* s parametrima koji su se pokazali najbolji u poglavlju 5.1. te su na njemu ispitani utjecaji tehnika proširivanja skupa podataka. Nezavisno je uspoređen utjecaj svake od transformacija na uspješnost modela pri čemu su hiperparametri modela bili fiksirani na najbolje vrijednosti utvrđene u prethodnim eksperimentima.

Konačno rezultati su uspoređeni i s modelom na kojem je korišteno više transformacija s jednakom vjerojatnošću. Prilikom učenja takvog modela korištene su sve transformacije koje nisu pokazale negativni utjecaj na rezultate modela.

5.3 Vrednovanje modela dubokog učenja

Kako bi mogli usporediti modele dubokog učenja potrebno je definirati metrike kojima ćemo ih uspoređivati. U okviru ovog rada za svaki model određena je točnost (engl. *accuracy*), mjera *F1* i površina ispod krivulje *ROC* koje model postiže na skupu za testiranje.

Prilikom klasifikacije primjera razlikujemo četiri vrste predviđanja s obzirom na ispravnost predviđanja: ispravno pozitivna predikcija (engl. *true positive*, TP), lažno pozitivna predikcija (engl. *false positive*, FP), lažno negativna predikcija (engl. *false negative*, FN), ispravno negativna predikcija (engl. *true negative*, TN). Takva podjela je i grafički prikazana matricom zabune (engl. *confusion matrix*) na slici 13.

| | | Stvarno | |
|-------|---|---------|----|
| | | 1 | 0 |
| Model | 1 | TP | FP |
| | 0 | FN | TN |

Slika 13. Grafički prikaz matrice zabune za binarnu klasifikaciju

Točnost (Acc) je definirana kao omjer točno klasificiranih primjera i ukupnog broja svih primjera.

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

Učenjem modela želimo postići da točnost bude što bliža vrijednosti 1, no problem s tom mjerom je taj što nam ne govori koliko je model dobar na svakoj od klasa što je problem kod skupova podataka koji nemaju jednak broj pozitivnih i negativnih primjera.

Mjera $F1$ je definirana kao harmonijska sredina između preciznosti (engl. *precision*) i odziva (engl. *recall*). Preciznost (P) predstavlja omjer ispravnih pozitivnih primjera i broja svih pozitivno klasificiranih primjera, a odziv (R) predstavlja omjer ispravnih pozitivnih primjera i ukupnog broja svih pozitivnih primjera u skupu podataka.

$$P = \frac{TP}{TP + FP}$$

$$R = \frac{TP}{TP + FN}$$

$$\frac{2}{F_1} = \frac{1}{P} + \frac{1}{R}$$

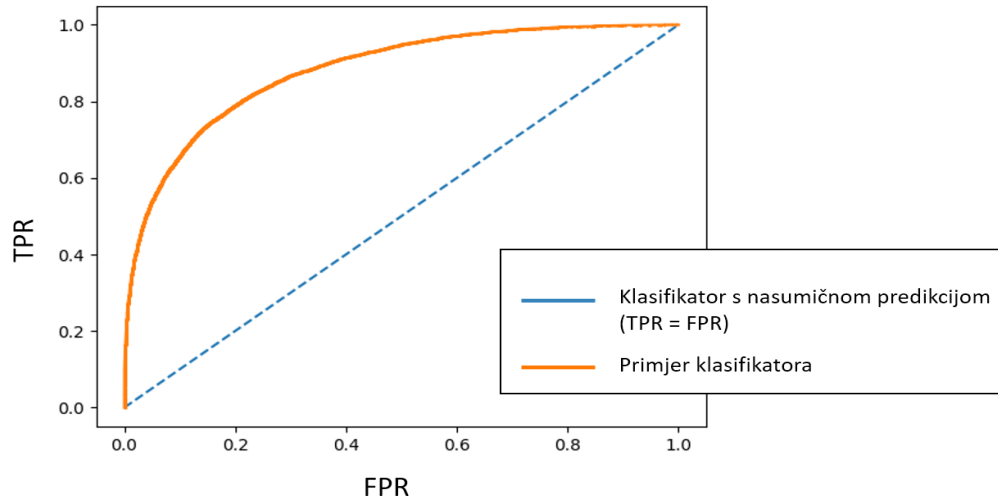
Prilagođavanjem parametara modela želimo postići da mjera $F1$ bude što bliža vrijednosti 1. Za razliku od točnosti koristeći mjeru $F1$ možemo vidjeti ako model klasificira sve primjere u jednu klasu pri čemu će vrijednost mjere $F1$ iznositi 0.

Krivulja ROC (engl. *receiver operating characteristics*) je grafička metoda za vrednovanje klasifikatora s pragom te se često koristi prilikom donošenja odluka u medicini [20]. Klasifikator daje na izlazu vrijednost između 0 i 1, ta se vrijednost zatim uspoređuje s pragom. Ako je vrijednost manja od praga primjer pripada negativnoj oznaci, a inače pripada pozitivnoj oznaci. Metoda se zasniva na praćenju odnosa između odziva i stope neispravno pozitivno klasificiranih primjera (engl. *false positive rate*, *fall-out*) prilikom promjene praga klasifikacije.

Stopa neispravno pozitivno klasificiranih primjera (FPR) predstavlja omjer broja primjera koji su pogrešno klasificirani pozitivnima i ukupnog broja negativnih primjera.

$$FPR = \frac{FP}{FP + TN}$$

Primjer krivulje ROC dan je na slici 14 pri čemu se na apscisi nalazi vrijednost stope neispravno pozitivno klasificiranih primjera, a na ordinati vrijednost odziva klasifikatora.



Slika 14. Primjer krivulje ROC za klasifikator s nasumičnom predikcijom i za klasifikator s rezultatom boljim od nasumične predikcije

Kvantitativna mjera koja se dobiva temeljem analize krivulje ROC je površina ispod krivulje (engl. *area under curve*, AUC) [20]. Prilikom učenja modela cilj je da vrijednost površine ispod krivulje bude što bliža vrijednosti 1.

6. Rezultati

U okviru ovog poglavlja prezentirani su rezultati ispitivanja modela dubokog učenja na zadatku klasifikacije histopatoloških slika limfnih čvorova. Metode provođenja eksperimenata su opisane u poglavlju 5.

6.1 Rezultati usporedbe modela dubokog učenja

U tablici 3 dani su rezultati učenja modela opisanih u poglavlju 4.1 ispočetka s različitim stopama učenja. Može se primijetiti kako učenje modela *AlexNet* nije konvergiralo za stopu učenja 10^{-3} . Modeli *ResNet* i *DenseNet* su se pokazali boljima od modela *AlexNet* te je razlika između njihovih mjera *F1* i mjera točnosti bila mala. Model *ResNet* je imao bolju mjeru površine ispod krivulje *ROC* što ukazuje na to da je model *ResNet* bio sigurniji prilikom predikcije. Međutim može se primijetiti kako model *DenseNet* ima više od pet puta manje parametara od modela *AlexNet* i skoro dva puta manje parametara od model *ResNet* te postiže izvrsne rezultate u usporedbi s ostala dva modela.

U tablici 4 dani su rezultati učenja modela prijenosom značajki s modela učenih na skupu podataka *ImageNet*. Usporedbom rezultata koje su modeli postigli prilikom učenja ispočetka i učenja prijenosom značajki može se primijetiti kako su najbolji rezultati svakog pojedinog modela bili bolji prilikom učenja prijenosom značajki. Razlika kod modela *AlexNet* je najmanja te je takav rezultat u skladu sa zaključkom iznesenim u znanstvenom radu [17] u kojem su pokazali da je moguće dobiti bolje ili iste rezultate učenjem modela *AlexNet* prijenosom značajki sa skupa podataka *ImageNet* u odnosu na učenje modela *AlexNet* ispočetka za zadatke analize radioloških slika.

Postojanje razlike između modela učenih ispočetka i modela učenih prijenosom značajki je djelomično uzrokovana time što prilikom učenja modela ispočetka nisu korištene metode proširivanja skupa podataka koje bi omogućile bolju generalizaciju. Modeli učeni prijenosom značajki su mogli preuzeti značajke koje omogućavaju bolju generalizaciju jer su u radu [42] pokazali kako model *AlexNet* prilikom učenja na skupu podataka *ImageNet* spontano uči reprezentaciju koja je nezavisna prilikom primjene transformacija rotacije, skaliranja i zrcaljenja.

Tablica 3. Rezultati modela učenih ispočetka s različitom stopom učenja

| Naziv modela | Stopa učenja | Točnost | Mjera $F1$ | $AUC\ ROC$ |
|------------------------|-----------------------------|---------------|---------------|---------------|
| <i>AlexNet</i> | 10^{-3} | 0.4998 | 0.6665 | 0.5000 |
| <i>AlexNet</i> | 10^{-4} | 0.7838 | 0.7697 | 0.8711 |
| <i>AlexNet</i> | 10^{-5} | 0.7847 | 0.7763 | 0.8662 |
| <i>AlexNet</i> | 10^{-6} | 0.7688 | 0.7694 | 0.8479 |
| <i>ResNet</i> | 10^{-3} | 0.8123 | 0.7898 | 0.9042 |
| <i>ResNet</i> | 10^{-4} | 0.7853 | 0.7827 | 0.8743 |
| <i>ResNet</i> | 10^{-5} | 0.7892 | 0.7857 | 0.8720 |
| <i>ResNet</i> | 10^{-6} | 0.7613 | 0.7492 | 0.8382 |
| <i>DenseNet</i> | 10^{-3} | 0.8125 | 0.7862 | 0.8788 |
| <i>DenseNet</i> | 10^{-4} | 0.7910 | 0.7583 | 0.8999 |
| <i>DenseNet</i> | 10^{-5} | 0.7872 | 0.7690 | 0.8714 |
| <i>DenseNet</i> | 10^{-6} | 0.7752 | 0.7684 | 0.8581 |

Tablica 4. Rezultati modela inicijaliziranih prijenosom značajki sa skupa podataka ImageNet. Prilikom učenja svi parametri su prilagođavani.

| Naziv modela | Stopa učenja | Točnost | Mjera $F1$ | AUC ROC |
|----------------------------|-----------------------------|---------------|---------------|---------------|
| <i>AlexNet</i> | 10^{-3} | 0.4997 | 0.6664 | 0.4999 |
| <i>AlexNet</i> | 10^{-4} | 0.7991 | 0.7720 | 0.9048 |
| <i>AlexNet</i> | 10^{-5} | 0.7812 | 0.75336 | 0.8815 |
| <i>AlexNet</i> | 10^{-6} | 0.7800 | 0.7477 | 0.8764 |
| <i>ResNet</i> | 10^{-3} | 0.8297 | 0.8144 | 0.9184 |
| <i>ResNet</i> | 10^{-4} | 0.8518 | 0.8334 | 0.9295 |
| <i>ResNet</i> | 10^{-5} | 0.8485 | 0.8386 | 0.9266 |
| <i>ResNet</i> | 10^{-6} | 0.8200 | 0.7978 | 0.9164 |
| <i>DenseNet</i> | 10^{-3} | 0.8150 | 0.7858 | 0.9088 |
| <i>DenseNet</i> | 10^{-4} | 0.8670 | 0.8522 | 0.9409 |
| <i>DenseNet</i> | 10^{-5} | 0.8555 | 0.8408 | 0.9354 |
| <i>DenseNet</i> | 10^{-6} | 0.8201 | 0.7920 | 0.9271 |
| <i>Inception-v3</i> | 10^{-3} | 0.8884 | 0.8872 | 0.9489 |
| <i>Inception-v3</i> | 10^{-4} | 0.8741 | 0.8661 | 0.9430 |
| <i>Inception-v3</i> | 10^{-5} | 0.8501 | 0.8327 | 0.9483 |
| <i>Inception-v3</i> | 10^{-6} | 0.8587 | 0.8436 | 0.9459 |

U tablici 5 dani su rezultati modela koji su učeni prijenosom značajki s *ImageNet* skupa podataka te su zatim prilagođavani samo zadnji potpuno povezani slojevi. Usporedbom rezultata s rezultatima modela učenih ispočetka može se vidjeti kako su modeli *AlexNet* i *ResNet* pokazali značajno bolje rezultate prilikom učenja ispočetka dok je model *DenseNet* postigao nešto bolje rezultate prilikom prijenosa značajki ako uspoređujemo mjeru $F1$ i površinu ispod krivulje ROC.

Usporedbom rezultata s rezultatima učenja modela prijenosom značajki nakon čega su prilagođavani svi parametri modela možemo vidjeti kako su svi modeli postignuli bolje rezultate kada su prilagođavani svi parametri modela.

Dobiveni rezultati ukazuju kako se značajke koje su potrebne modelu u ranijim slojevima za ispravnu klasifikaciju razlikuju od značajki koje model nauči na skupu podataka *ImageNet*.

Najboljim se pokazao model *Inception-v3* koji je imao značajno bolju mjeru točnosti i mjeru $F1$ u odnosu na ostale modele. Drugi najbolji model je *DenseNet* te je on postignuo sličan iznos površine ispod krivulje *ROC*. Jedan od razloga zašto je model *Inception-v3* postignuo bolje rezultate je taj što je imao značajno više parametara od modela *DenseNet* uz nešto manji broj nelinearnih aktivacija.

Najbolji model u radu [5] na skupu podataka *PCam* je postigao točnost u iznosu 0.898 i iznos površine ispod krivulje *ROC* u iznosu 0.963. Takvi rezultati su usporedivi s rezultatima koje postiže model *Inception-v3* (tablica 4.) te model *DenseNet* prilikom učenja s tehnikama proširivanja skupa podataka (poglavlje 6.2.).

Tablica 5. Rezultati modela inicijaliziranih prijenosom značajki sa skupa podataka ImageNet. Prilikom učenja samo zadnji potpuno povezani slojevi su prilagođavani.

| Naziv modela | Stopa učenja | Točnost | Mjera $F1$ | AUC ROC |
|----------------------------|--------------|---------------|---------------|---------------|
| <i>AlexNet</i> | 10^{-3} | 0.7569 | 0.7398 | 0.8308 |
| <i>AlexNet</i> | 10^{-4} | 0.7453 | 0.7061 | 0.8287 |
| <i>AlexNet</i> | 10^{-5} | 0.7589 | 0.7359 | 0.8395 |
| <i>AlexNet</i> | 10^{-6} | 0.7592 | 0.7301 | 0.8496 |
| <i>ResNet</i> | 10^{-3} | 0.7760 | 0.7767 | 0.8547 |
| <i>ResNet</i> | 10^{-4} | 0.7822 | 0.7676 | 0.8637 |
| <i>ResNet</i> | 10^{-5} | 0.7795 | 0.7685 | 0.8599 |
| <i>ResNet</i> | 10^{-6} | 0.7492 | 0.7447 | 0.8156 |
| <i>DenseNet</i> | 10^{-3} | 0.8065 | 0.8009 | 0.8874 |
| <i>DenseNet</i> | 10^{-4} | 0.7988 | 0.7906 | 0.8818 |
| <i>DenseNet</i> | 10^{-5} | 0.7904 | 0.7783 | 0.8766 |
| <i>DenseNet</i> | 10^{-6} | 0.7630 | 0.7546 | 0.8406 |
| <i>Inception-v3</i> | 10^{-3} | 0.8270 | 0.8204 | 0.9018 |
| <i>Inception-v3</i> | 10^{-4} | 0.8210 | 0.8149 | 0.9022 |
| <i>Inception-v3</i> | 10^{-5} | 0.8121 | 0.8059 | 0.8948 |
| <i>Inception-v3</i> | 10^{-6} | 0.7728 | 0.7704 | 0.8536 |

6.2 Rezultati ispitivanja utjecaja tehnika proširivanja skupa podataka

U tablici 6 dani su rezultati usporedbe primjene različitih transformacija prilikom proširenja skupa podataka. Usporedbom pojedinačnih transformacija pokazalo se kako vertikalno zrcaljenje najviše poboljšava rezultate modela te kako rotacija slike do 360° i promjena zasićenja boje utječu negativno na rezultat modela. Rotacija do 360° utječe negativno jer prilikom veće rotacije dio slike izlazi iz slike što može uzrokovati gubitak informacije te se vrijednosti slikovnih elemenata računaju interpolacijom pri čemu se mijenja početna informacija spremljena na slici. Moguće je primijetiti kako mala rotacija slike utječe pozitivno na rezultate što ukazuje na to da prilikom manje rotacije ne dolazi do velike promjene ili gubitka informacije te da zato model može naučiti bolje generalizirati.

Usporedbom rezultata pojedinačnih transformacija i kombinacije transformacija može se primijetiti kako vertikalno zrcaljenje i kombinacija transformacija postižu podjednake rezultate. Izuzev vertikalnog zrcaljenja, kombinacija transformacija postiže bolje rezultate od ostalih pojedinačnih transformacija.

Prilikom kombinacije transformacija korištene su sve transformacije koje nisu negativno utjecale na rezultat modela: rotacija slike do 5° , horizontalno i vertikalno zrcaljenje, promjena nijanse boje, kontrasta i svjetline.

Tablica 6. Rezultati korištenog modela prilikom primjene različitih metoda transformacije za proširenje skupa podataka.

| Transformacija | Točnost | Mjera $F1$ | $AUC\ ROC$ |
|-----------------------------------|---------------|---------------|---------------|
| Bez transformacija | 0.8670 | 0.8522 | 0.9409 |
| Rotacija slike 360° uniformno | 0.8447 | 0.8266 | 0.9110 |
| Rotacija slike 5° uniformno | 0.8719 | 0.8619 | 0.9407 |
| Vertikalno zrcaljenje | 0.8881 | 0.8816 | 0.9563 |
| Horizontalno zrcaljenje | 0.8766 | 0.8701 | 0.9419 |
| Promjena nijanse boje | 0.8704 | 0.8577 | 0.9473 |
| Promjena zasićenja boje | 0.8566 | 0.8425 | 0.9271 |
| Promjena svjetline do 5% | 0.8669 | 0.8549 | 0.9346 |
| Promjena kontrasta do 5% | 0.8714 | 0.8637 | 0.9425 |
| Kombinacija transformacija | 0.8867 | 0.8798 | 0.9589 |

7. Zaključak

U okviru ovog diplomskog rada predstavljen je problem klasifikacije histopatoloških slika limfnih čvorova te je napravljena programska implementacija u programskom jeziku *Python*. Cilj klasifikacije histopatoloških slika limfnih čvorova je odrediti sadrži li slika tumorsko ili zdravo tkivo. Tako klasificirane slike se mogu koristiti u postupcima analize snimaka preparata snimanih metodom *WSI*. Rezultate analize doktori koriste u dijagnostici i terapiji.

Analizu histopatoloških slika je potrebno automatizirati kako bi se postignulo ubrzanje i povećanje točnosti procesa. Jedan od pristupa automatske klasifikacije histopatoloških slika je klasifikacija uporabom dubokih konvolucijskih neuronskih mreža koje su se pokazale uspješne u području računalnog vida i šire.

Odabrane su arhitekture dubokih modela *AlexNet*, *ResNet*, *DenseNet* i *Inception-v3* koje su zatim uspoređene prilikom promjene stope učenja i načina inicijalizacije modela. Pokazalo se kako arhitektura *Inception-v3* postiže najbolje rezultate prilikom inicijalizacije prijenosom značajki s modela učenog na skupu podataka *ImageNet*. Osim navedenog modela istaknuo se model *DenseNet* koji je imao značajno manje parametara te postignuo sličan rezultat.

Usporedbom učenja ispočetka, učenja prijenosom značajki s modela učenog na skupu podataka *ImageNet* i prilagođavanjem svih parametara te učenja prijenosom značajki i prilagođavanjem samo zadnjih potpuno povezanih slojeva dobiveni su sljedeći zaključci. Učenje prijenosom značajki i prilagođavanje potpuno povezanih slojeva je pokazalo najlošije rezultate što ukazuje na to da je potrebno mijenjati i način izdvajanja značajki koji se uče u konvolucijskim slojevima. Učenje prijenosom značajki i prilagođavanje svih parametara je pokazalo najbolje rezultate. Razlika između takvog učenja i učenja ispočetka je bila manja te bi se mogla i smanjiti da su se koristile tehnike proširivanja skupa podataka kako bi mreža postigla bolju generalizaciju. To je bitno istaknuti jer je model s kojeg su se prenosile značajke bio učen takvim metodama te je zbog toga mogao postići i bolju generalizaciju na ciljanom skupu podataka. Takav zaključak se slaže i sa zaključkom dobivenim u [17] koji se odnosio na analizu radioloških slika.

Ispitan je utjecaj metoda proširivanja skupa podataka na rezultat učenog modela. Prilikom ispitivanja korištene su sljedeće transformacije rotacija slike za nasumični kut, vertikalno zrcaljenje slike, horizontalno zrcaljenje slike, promjena iznosa nijanse boje, promjena iznosa zasićenja boje, promjena iznosa svjetline i promjena iznosa kontrasta. Rezultat utjecaja pojedinih transformacija je uspoređen i s kombinacijom svih transformacija. Pokazalo se je da je vertikalno zrcaljenje transformacija s najznačajnijim pojedinačnim doprinosom, te da je podjednak rezultat dobiven s kombinacijom transformacija.

U budućem istraživanju potrebno je vrednovati naučene modele na cjelokupnim snimkama snimljenim metodom *WSI*. Takvim istraživanjem bi pokazali koliko su primjenjivi korišteni modeli prilikom potpomognute analize i dijagnostike u medicinskim sustavima. Osim toga, potrebno je proširiti istraživanje u smjeru mjerenja pouzdanosti predikcije modela te analizirati mogućnosti primjene različitih metoda interpretabilnosti. Cilj tih istraživanja bi bio omogućiti doktorima da mogu bolje razumjeti zašto je sustav predložio određenu dijagnozu te olakšati odluku o prihvatanju ili odbijanju predložene dijagnoze.

8. Literatura

- [1] D. Komura i S. Ishikawa, "Machine Learning Methods for Histopathological Image Analysis", Computational and Structural Biotechnology Journal, vol. 16, pp. 34-42, 2018., DOI: 10.1016/j.csbj.2018.01.001.
- [2] G. Litjens et al., "A survey on deep learning in medical image analysis", Medical Image Analysis, vol. 42, pp. 60-88, 2017., DOI: 10.1016/j.media.2017.07.005.
- [3] B. Sahiner et al., "Deep learning in medical imaging and radiation therapy", Medical Physics, vol. 46, no. 1, pp. e1-e36, 2018., DOI: 10.1002/mp.13264.
- [4] G. Litjens et al., "1399 H&E-stained sentinel lymph node sections of breast cancer patients: the CAMELYON dataset", GigaScience, vol. 7, no. 6, 2018. DOI: 10.1093/gigascience/giy065.
- [5] B. Veeling, J. Linmans, J. Winkens, T. Cohen i M. Welling, "Rotation Equivariant CNNs for Digital Pathology", Medical Image Computing and Computer Assisted Intervention – MICCAI 2018, pp. 210-218, 2018. DOI: 10.1007/978-3-030-00934-2_24.
- [6] B. Ehteshami Bejnordi et al., "Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women with Breast Cancer", JAMA - Journal of the American Medical Association, vol. 318, no. 22, pp. 2199-2210, 2017. DOI: 10.1001/jama.2017.14585.
- [7] „histologija“. U: Hrvatska enciklopedija (on-line). Leksikografski zavod „Miroslav Krleža“. Pristup ostvaren 13.5.2019.
<http://www.enciklopedija.hr/Natuknica.aspx?ID=25754>
- [8] „limfni čvor“. U: Hrvatska enciklopedija (on-line). Leksikografski zavod „Miroslav Krleža“. Pristup ostvaren 13.5.2019.
www.enciklopedija.hr/Natuknica.aspx?ID=36573
- [9] K. He, X. Zhang, S. Ren i J. Sun, "Deep Residual Learning for Image Recognition", 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770-778, 2016., DOI: 10.1109/cvpr.2016.90.
- [10] The HDF Group. „Hierarchical Data Format“, version 5, 1997-2019.
<http://www.hdfgroup.org/HDF5/>.

- [11] M. Gurcan, L. Boucheron, A. Can, A. Madabhushi, N. Rajpoot and B. Yener, "Histopathological Image Analysis: A Review", IEEE Reviews in Biomedical Engineering, vol. 2, pp. 147-171, 2009., DOI: 10.1109/rbme.2009.2034865.
- [12] A. Krizhevsky, I. Sutskever, G. E. Hinton, "Imagenet classification with deep convolutional neural networks", Advances in neural information processing systems, pp. 1097-1105, 2012, <http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf>
- [13] A. Krizhevsky, "One weird trick for parallelizing convolutional neural networks", 2014., arxiv: <http://arxiv.org/abs/1404.5997>.
- [14] A. Paszke et al., Automatic differentiation in PyTorch, U NIPS-W, 2017, <https://openreview.net/pdf?id=BJJsrnCZ>
- [15] S. Shalev-Shwartz, S. Ben-David, Understanding machine learning. Cambridge: Cambridge University Press, 2014., <http://www.cs.huji.ac.il/~shais/UnderstandingMachineLearning/understanding-machine-learning-theory-algorithms.pdf>
- [16] C. Bishop, Pattern Recognition and Machine Learning. New York: Springer Science+Business Media, 2006.
- [17] N. Tajbakhsh et al., "Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?", IEEE Transactions on Medical Imaging, vol. 35, no. 5, pp. 1299-1312, 2016., DOI: 10.1109/tmi.2016.2535302.
- [18] A. Zhang, Z. Lipton, M. Li, A. Smola, "Dive into Deep Learning". University of California, Berkley, 2019., <https://d2l.ai/>
- [19] D. Masters i C. Luschi, "Revisiting Small Batch Training for Deep Neural Networks", 2018. [Online]. Arxiv: <http://arxiv.org/abs/1804.07612>.
- [20] T. Fawcett, "An introduction to ROC analysis", Pattern Recognition Letters, vol. 27, no. 8, pp. 861-874, 2006. DOI: 10.1016/j.patrec.2005.10.010
- [21] K. He, X. Zhang, S. Ren i J. Sun, "Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification", 2015 IEEE International Conference on Computer Vision (ICCV), 2015. DOI: 10.1109/iccv.2015.123.

- [22] L. Pantanowitz, "Digital images and the future of digital pathology", *Journal of Pathology Informatics*, vol. 1, no. 1, p. 15, 2010. DOI: 10.4103/2153-3539.68332.
- [23] G. Montavon, W. Samek i K. Müller, "Methods for interpreting and understanding deep neural networks", *Digital Signal Processing*, vol. 73, pp. 1-15, 2018. DOI: 10.1016/j.dsp.2017.10.011.
- [24] S. Chakraborty et al., "Interpretability of deep learning models: A survey of results", 2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCoM/IOP/SCI), 2017. DOI: 10.1109/uic-atc.2017.8397411.
- [25] J. Kim, J. Hong i H. Park, "Prospects of deep learning for medical imaging", *Precision and Future Medicine*, vol. 2, no. 2, pp. 37-52, 2018. DOI: 10.23838/pfm.2018.00030.
- [26] Z. Lu, "PubMed and beyond: a survey of web tools for searching biomedical literature", *Database*, vol. 2011, no. 0, pp. 1-13, 2011. DOI: 10.1093/database/baq036.
- [27] E. Gibson et al., "NiftyNet: a deep-learning platform for medical imaging", *Computer Methods and Programs in Biomedicine*, vol. 158, pp. 113-122, 2018. DOI: 10.1016/j.cmpb.2018.01.025.
- [28] G. Huang, Z. Liu, L. Maaten i K. Weinberger, "Densely Connected Convolutional Networks", 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2261-2269, 2017., DOI: 10.1109/cvpr.2017.243.
- [29] E. Deniz, A. Şengür, Z. Kadiroğlu, Y. Guo, V. Bajaj and Ü. Budak, "Transfer learning based histopathologic image classification for breast cancer detection", *Health Information Science and Systems*, vol. 6, no. 1, pp. 1-7, 2018. DOI: 10.1007/s13755-018-0057-x.
- [30] R. Srivastava, K. Greff i J. Schmidhuber, "Highway Networks", 2015. arxiv: <http://arxiv.org/abs/1505.00387>.

- [31] P. Bandi et al., "From Detection of Individual Metastases to Classification of Lymph Node Status at the Patient Level: The CAMELYON17 Challenge", *IEEE Transactions on Medical Imaging*, vol. 38, no. 2, pp. 550-560, 2019. DOI: 10.1109/tmi.2018.2867350.
- [32] F. Spanhol, L. Oliveira, C. Petitjean i L. Heutte, "A Dataset for Breast Cancer Histopathological Image Classification", *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 7, pp. 1455-1462, 2016. DOI: 10.1109/tbme.2015.2496264.
- [33] O. Russakovsky et al., "ImageNet Large Scale Visual Recognition Challenge", *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211-252, 2015. DOI: 10.1007/s11263-015-0816-y.
- [34] A. Krizhevsky, "Learning Multiple Layers of Features from Tiny Images", University of Toronto, 2012. <https://www.cs.toronto.edu/~kriz/learning-features-2009-TR.pdf>
- [35] M. Everingham, L. Van Gool, C. Williams, J. Winn i A. Zisserman, "The Pascal Visual Object Classes (VOC) Challenge", *International Journal of Computer Vision*, vol. 88, no. 2, pp. 303-338, 2009. DOI: 10.1007/s11263-009-0275-4.
- [36] B. Menze et al., "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)", *IEEE Transactions on Medical Imaging*, vol. 34, no. 10, pp. 1993-2024, 2015. DOI: 10.1109/tmi.2014.2377694.
- [37] S. Bakas et al., "Advancing the Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features", *Scientific Data*, vol. 4, p. 170117, 2017. DOI: 10.1038/sdata.2017.117.
- [38] Y. Lecun, L. Bottou, Y. Bengio i P. Haffner, "Gradient-based learning applied to document recognition", *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278-2324, 1998. DOI: 10.1109/5.726791.
- [39] X. Zhang, J. Zou, K. He i J. Sun, "Accelerating Very Deep Convolutional Networks for Classification and Detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 10, pp. 1943-1955, 2016. DOI: 10.1109/tpami.2015.2502579.

- [40] C. Szegedy et al., "Going deeper with convolutions", 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015. DOI: 10.1109/cvpr.2015.7298594.
- [41] S. Ioffe, C. Szegedy, „Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift“, 2015, arxiv: <https://arxiv.org/abs/1502.03167>
- [42] K. Lenc, A. Vedaldi, "Understanding image representations by measuring their equivariance and equivalence," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, 2015, pp. 991-999., 2015. DOI: 10.1109/CVPR.2015.7298701
- [43] T.S. Cohen, M. Welling, Group Equivariant Convolutional Networks. Proceedings of the International Conference on Machine Learning (ICML), 2016., arxiv: <https://arxiv.org/abs/1602.07576>
- [44] T.S. Cohen, M. Geiger, M. Weiler, A General Theory of Equivariant CNNs on Homogeneous Spaces, 2018, arxiv: <https://arxiv.org/abs/1811.02017>
- [45] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, 2016, pp. 2818-2826., DOI: 10.1109/CVPR.2016.308
- [46] "Running Inception on Cloud TPU", Google Cloud, 2019., <https://cloud.google.com/tpu/docs/tutorials/inception>,
Pristup ostvaren: 10.6.2019.

Sažetak

Prilikom analize histopatoloških slika ukazuje se potreba za automatiziranim sustavom koji bi mogao pomoći doktorima u analizi slika i dijagnostici. Takav sustav bi mogao povećati točnost i brzinu analize i dijagnostike.

U okviru ovog rada dan je pregled područja analize histopatoloških slika te je izrađena programska implementacija za klasifikaciju limfnih čvorova temeljena na strojnom učenju.

Za učenje i testiranje modela strojnog učenja korišten je skup podataka *PatchCamelyon*.

Ispitani su rezultati koje postižu sljedeći modeli dubokog učenja: *AlexNet*, *ResNet*, *DenseNet* i *Inception-v3* te kako na njihove rezultate utječu tehnike proširenja skupa podataka. Konačno najboljim se pokazao model *Inception-v3* koji je na skupu za testiranje postigao točnost od 89%.

Ključne riječi: strojno učenje, analiza medicinskih slika, histopatologija

Summary

When analyzing histopathological images, there is a need for an automated system that could help doctors in image analysis and diagnostics. Such a system could increase the accuracy and speed of analysis and diagnostics.

Within this paper, an overview of the histopathological image analysis area is provided and a program implementation for the classification of lymph nodes based on machine learning has been developed.

PatchCamelyon dataset has been used for training and testing of chosen machine learning models.

The results of the following deep learning models have been studied: *AlexNet*, *ResNet*, *DenseNet*, *Inception-v3*. Also, the influence of different data augmentation methods on the model performance was investigated. Finally, the *Inception-v3* model proved to be the best, which reached the 89% accuracy on the test set.

Keywords: machine learning, medical image analysis, histopathology