

**SVEUČILIŠTE U SPLITU
FAKULTET ELEKTROTEHNIKE, STROJARSTVA I
BRODOGRADNJE**

**POSLIJEDIPLOMSKI DOKTORSKI STUDIJ
ELEKTROTEHNIKA I INFORMACIJSKA TEHNOLOGIJA**

KVALIFIKACIJSKI ISPIT

**SEMANTIČKA SEGMENTACIJA
MEDICINSKIH SLIKA METODAMA
DUBOKOG UČENJA**

Ana Pinjuh

Split, listopada 2020.

SADRŽAJ

1	Uvod	1
2	Umjetne neuronske mreže	2
3	Konvolucijske neuronske mreže	4
3.1	Arhitektura konvolucijskih neuronskih mreža	4
3.2	Standardne duboke arhitekture	6
4	Primjena dubokog učenja u segmentaciji medicinskih slika.....	10
4.1	Modalitet medicinskih slika.....	10
4.2	Segmentacija medicinskih slika.....	10
4.3	2.5D konvolucijske neuronske mreže.....	12
4.4	3D konvolucijske neuronske mreže.....	12
4.5	Potpuno konvolucijske mreže.....	14
4.5.1	Potpuno konvolucijske mreže za istovremenu segmentaciju više organa.....	15
4.5.2	Kaskadne potpuno konvolucijske mreže.....	15
4.5.3	Multi-Stream potpuno konvolucijske mreže	16
4.5.4	U-Net.....	16
4.5.5	V-Net.....	17
4.6	ResNet	17
4.7	Povratne neuronske mreže.....	19
4.7.1	LSTM	20
4.7.2	CW-RNN.....	21
5	Segmentacija slika magnetske rezonancije koljena	22
6	Proces učenja mreže	25
6.1	Prijenosno učenje.....	25
6.2	Metrike za procjenu rezultata segmentacije	26
7	Izazovi u primjeni metoda dubokog učenja na medicinskim slikama	28
7.1	Izazovi u pripremi podataka i treniranju.....	28
7.2	Izazovi temeljeni na različitim organima i modalitetima	29
8	Zaključak.....	31
	Literatura	32

1 UVOD

Inteligencija ima ključnu ulogu u ljudskom ponašanju, ljudskom razumijevanju okoline i učenju iz novonastalih situacija. Ljudi reagiraju kroz proces učenja i prilagodbe od informacija primljenih kroz raširenu mrežu senzora i kontrolnih mehanizama u našim tijelima. Te sposobnosti, karakteristične za ljude, nastoje se godinama preslikati odnosno postići kod strojeva. Teži se kreiranju sustava koji su sposobni za samostalnu transformaciju i prilagođavanje promjenjivoj okolini kako bi riješili različite vrste problema. Međutim, ono što ljudi smatraju inteligentnim može biti jednostavno realizirati na računalima. S druge strane, godine istraživanja u programiranju računala su pokazala da mnoga ponašanja koja se ne smatraju inteligentnima zahtijevaju velik napor u realiziranju kao što su govor, vid, prepoznavanje uzoraka i ostale svakodnevne radnje prosječnog čovjeka.

Analiza medicinskih slika se radi već desetljećima otkad je moguće skenirati i učitati medicinske slike u računala. Glavni cilj istraživača jeste iskoristiti nove tehnologije kako bi poboljšali brigu o pacijentima i olakšali proces dijagnosticiranja liječnicima. Sredinom 90ih godina prošlog stoljeća konvolucijske neuronske mreže se pojavljuju u analizi medicinskih slika što označava početak upotrebe dubokog učenja na medicinskim slikama. Od tada su konvolucijske neuronske mreže postale standard za zadatke računalnog vida prvenstveno za klasifikaciju, detekciju i segmentaciju medicinskih slika.

Semantička segmentacija slike je postala predmetom interesa istraživača na području računalnog vida te se koristi u razumijevanju 2D slika i videa pa čak i višedimenzionalnih podataka. Međutim, semantička segmentacija medicinskih slika je izazov zbog specifičnosti medicinskih podataka u njihovom formatu, modalitetu i veličini. Također, potrebna je suradnja sa stručnjakom s područja medicine za prikupljanje i označavanje slika kao i procjenu rezultata.

U ovom radu je dat pregled dubokih metoda učenja za analizu medicinskih slika s posebnim osvrtom na semantičku segmentaciju slika. U drugom poglavlju je objašnjen pojam umjetne neuronske mreže. U trećem poglavlju je detaljno objašnjena arhitektura i karakteristike konvolucijske neuronske mreže te neke od standardnih najvažnijih arhitektura. Četvrto poglavlje opisuje doprinos metoda dubokog učenja u zadatku semantičke segmentacije. U petom poglavlju je dat poseban osvrt na segmentaciju slika magnetske rezonancije koljena.

Nadalje, šesto poglavlje je posvećeno procesima učenja mreže. Najčešći izazovi i problemi s kojim se istraživači susreću prilikom obrade medicinskih slika te odabira prikladne arhitekture su predstavljeni u sedmom poglavlju. U posljednjem, osmom poglavlju dat je zaključak.

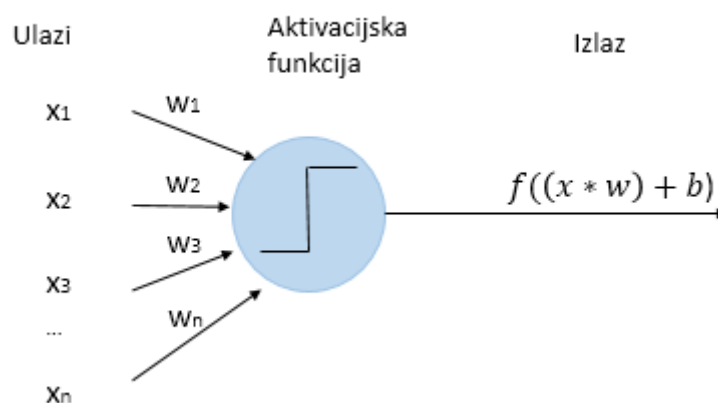
2 UMJETNE NEURONSKE MREŽE

Umjetne neuronske mreže su stvorene po uzoru na prirodne neuronske mreže u ljudskom mozgu, ali nisu ni približno složene kao mozak. Ljudski mozak se sastoji od velikog broja izrazito međusobno povezanih živčanih stanica ili neurona. Funkcioniranje prirodne neuronske mreže je određeno položajem neurona i njihovim međusobnim vezama. Naravno, kako čovjek stari tako se djelovanje njegovog mozga s vremenom mijenja te pojedine veze među neuronima slabe ili jačaju na čemu se zasnivaju svi procesi učenja kod ljudi. Dakle, umjetna neuronska mreža je općeniti matematički model prirodnog neuronskog sustava. Određena je s arhitekturom neurona, snagom veze između neurona (težinama) i svojstvima neurona. Ključne sličnosti između umjetnih i prirodnih neuronskih mreža su njihovi osnovni elementi koji obavljaju neku funkciju te važnost međusobnih veza elemenata u procesu učenja i cjelokupnog ponašanja mreže.

Ulaz u neuron je predstavljen oznakom x (x_1, x_2, \dots, x_n) te se množi sa težinskim faktorom w ($w_1, w_2, \dots, w_n \in \mathbb{R}$) koji definira važnost pojedine ulazne varijable [1]. Produkt se potom zbraja s internim ulazom pomaka b koji je poznat kao sklonost neurona (engl. *bias*) odnosno mjera koja definira lakoću da neuronska mreža predvidi pozitivnu klasu. Na temelju ovih koeficijenata se proizvodi aktivacijsko djelovanje mreže definirano izrazom:

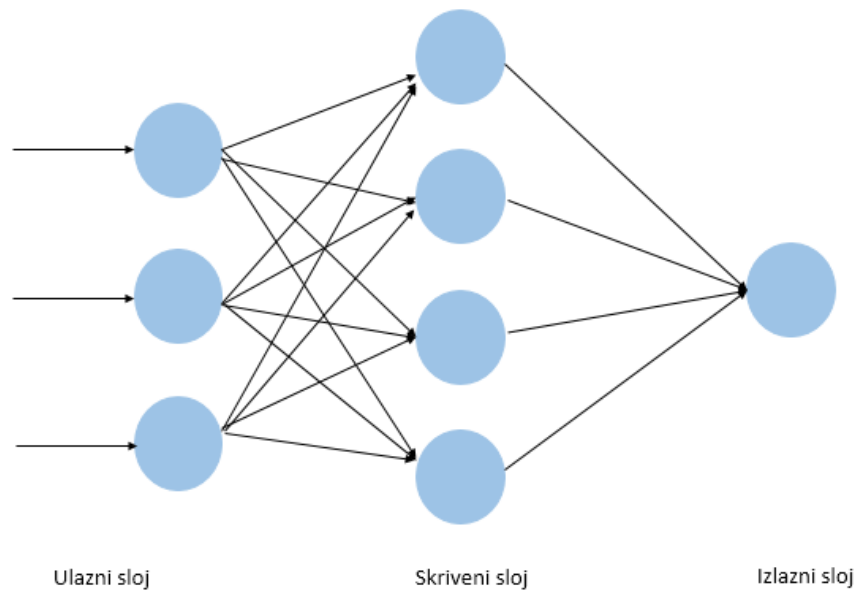
$$h = f(a) = f(x * w + b) \quad (2.1)$$

Slika 2.1 prikazuje umjetni neuron sa osnovnim komponentama.



Slika 2.1 Prikaz umjetnog neurona [1]

Umjetna neuronska mreža (engl. *Artificial Neural Networks*) može sadržavati ulazni sloj (engl. *input layer*), skrivene slojeve (engl. *hidden layer*) i izlazni sloj (engl. *output layer*) kako je prikazano na slici 2.2. Ulazni i izlazni sloj su u direktnoj interakciji s okolinom, a sloj koji nema interakcije naziva se skriveni sloj. Jedna neuronska mreža obično ima više skrivenih slojeva te samo jedan ulazni i izlazni sloj. Broj neurona u ulaznom sloju određen je dimenzijom ulaznih podataka dok je broj neurona u izlaznom sloju jednak broju klasa koje neuronska mreža mora predvidjeti. Svi umjetni neuroni u jednom sloju su potpuno ili djelomično povezani s umjetnim neuronima idućeg sloja. Jakost veze između dva neurona se zove težina. Upravo su te vrijednosti težina parametri mreže i predmet svih procedura učenja kod neuronskih mreža.



Slika 2.2 Grafički prikaz slojeva umjetne neuronske mreže sa jednim skrivenim slojem

Jedna od glavnih značajki neuronskih mreža jeste njihova sposobnost učenja. Umjetne neuronske mreže su i stvorene sa ciljem da pokušaju simulirati postupak učenja biološkog mozga. Proces učenja se u pravilu sastoji od iterativnog prilagođavanja sinaptičkih težina i pragova neurona i ti uobičajeni koraci se nazivaju algoritmima učenja. Mreža nastoji naučiti vezu između prezentiranih uzoraka i njenih izlaza te nastoji postati sposobna generalizirati rješenje. Nakon svake iteracije procesa učenja mreža postaje poučenija o svojoj okolini.

Pri tome se razlikuje nadzirano i nenadzirano učenje. Strategija nadziranog učenja će mreži osim ulaznih podataka dati i željene izlazne vrijednosti. Mreža potom u procesu učenja stalno provjerava svoj rezultat u odnosu na željeni te prilagođuje svoje parametre. Dakle, varijable mreže se podešavaju poznavanjem unaprijed određenog izlaza za taj sustav. Nenadzirano učenje nema znanja o željenim izlazima nego mreža sama mora pronaći sličnosti između ulaznih podataka te podešavati parametre mreže kako bi slične skupove ulaznih podataka grupirala u klastere.

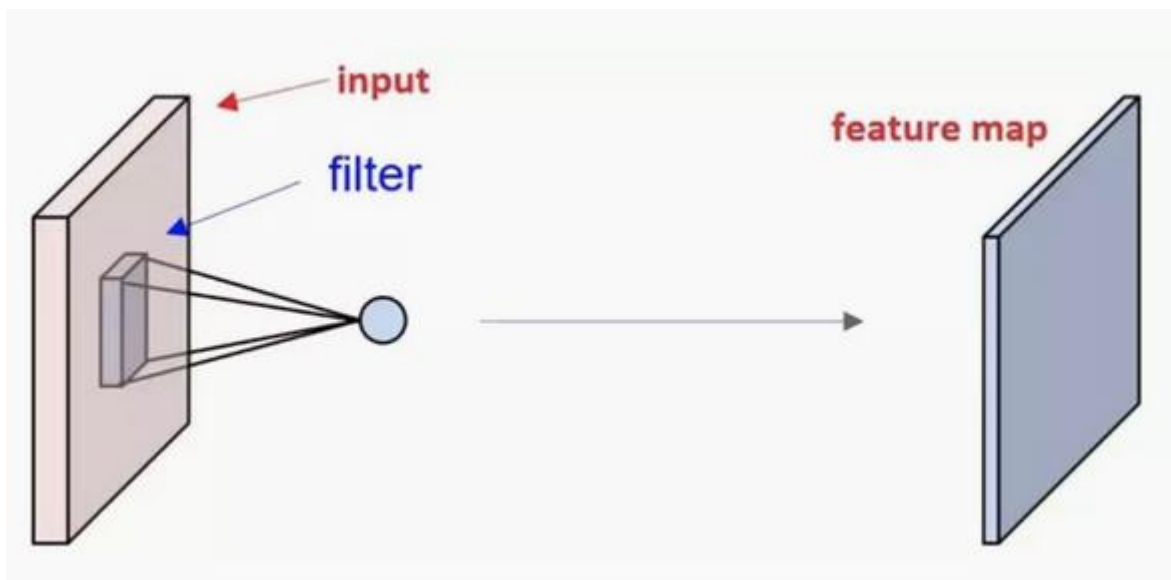
Neuronske mreže kod kojih neuroni svakog sljedećeg sloja primaju izlaz iz prethodnog sloja zovu se unaprijedne neuronske mreže (engl. *feedforward*). Neuronske mreže mogu imati i povratne veze ili petlje među svojim neuronima i takve mreže se nazivaju povratne neuronske mreže (engl. *recurrent*).

3 KONVOLUCIJSKE NEURONSKE MREŽE

Kada neuronska mreža sadrži više skrivenih slojeva smatra se dubokom neuronskom mrežom stoga se često u literaturi nađe pojam „duboke mreže“ ili „duboko učenje“. Osnovu većine algoritama dubokog učenja čini model (neuronska mreža) sastavljen od mnogo slojeva koji transformiraju ulazne podatke u izlazne vrijednosti tako da u tom procesu uči značajke. Konvolucijske neuronske mreže (engl. *Convolutional Neural Networks, CNN*) su najpoznatija, i najkorištenija arhitektura dubokog učenja inicijalno stvorena s ciljem obrade i analize slike [2]. Konvolucijske neuronske mreže su zbog svojih performansi stekle popularnost i koriste se za sve zadatke računalnog vida uključujući klasifikaciju, detekciju, segmentaciju, registraciju i slično [3], [4], [5].

3.1 Arhitektura konvolucijskih neuronskih mreža

Konvolucijske neuronske mreže koriste operaciju konvolucije kako bi izvukli značajke slike korištenjem filtera čije težine se automatski uče u procesu treniranja. Slika 3.1 prikazuje operaciju konvolucije. Za operaciju konvolucije se može koristiti širok spektar različitih filtera gdje svaki filter izvlači određenu značajku sa slike (npr. horizontalni, vertikalni rub i sl.).



Slika 3.1 Primjer izvođenja operacije konvolucije [6]

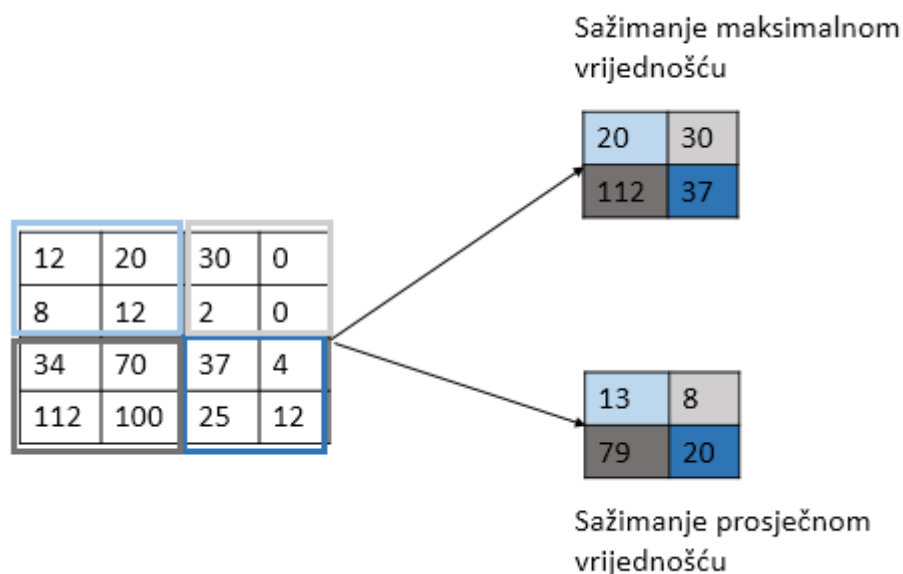
Osnovni dijelovi konvolucijske neuronske mreže su ulazni sloj, konvolucijski sloj, sloj sažimanja i potpuno povezani sloj. Ulazni sloj koji je direktno povezan sa ulaznom slikom koja se sastoji od niza piksela određene visine i širine te broja kanal te je broj neurona jednak broju piksela u slici.

Konvolucijski sloj se koristi za traženje značajki na slici. Zadanim filtrima se obavlja operacija konvolucije preko ulaznog podatka te se filtri pomiču dok se u potpunosti ne obradi cijeli ulazni podatak. Izlaz svakog konvolucijskog sloja je aktivacijska mapa (mapa značajki ili engl. *feature map*) koja naglašava primjenu određenog filtra na ulaz. Svaki neuron sloja odgovara točno određenom području prethodnog sloja što se naziva receptivno polje (engl. *Receptive field*). Svaki konvolucijski sloj mora imati definirane hiperparametre: broj i veličina filtera, dopunjavanje (engl. *Padding*) i korak (engl. *Stride*). Dopunjavanje predstavlja nadodavanje nula na sliku kako bi mogli kontrolirati veličinu izlaznog volumena. Korak određuje pomak filtera po ulaznoj slici pa ukoliko je korak 1 to znači da se filter pomiče za jedan piksel po ulaznoj slici dok se ne obradi cijela slika.

Konvolucijski sloj obično slijedi aktivacijski sloj (engl. *Activation layer*) koji primjenjuje nelinearnost na aktivacijske mape korištenjem aktivacijske funkcije. Najkorištenija aktivacijska funkcija je ReLU (engl. *Rectified Linear Unit*) koja postavlja vrijednost na 0 samo ako je negativna, a u suprotnom vrijednost ostaje ista. Ovaj sloj se također bavi problemima poput nestajućeg gradijenta (engl. *Vanishing gradient*) u propagaciji greške unazad.

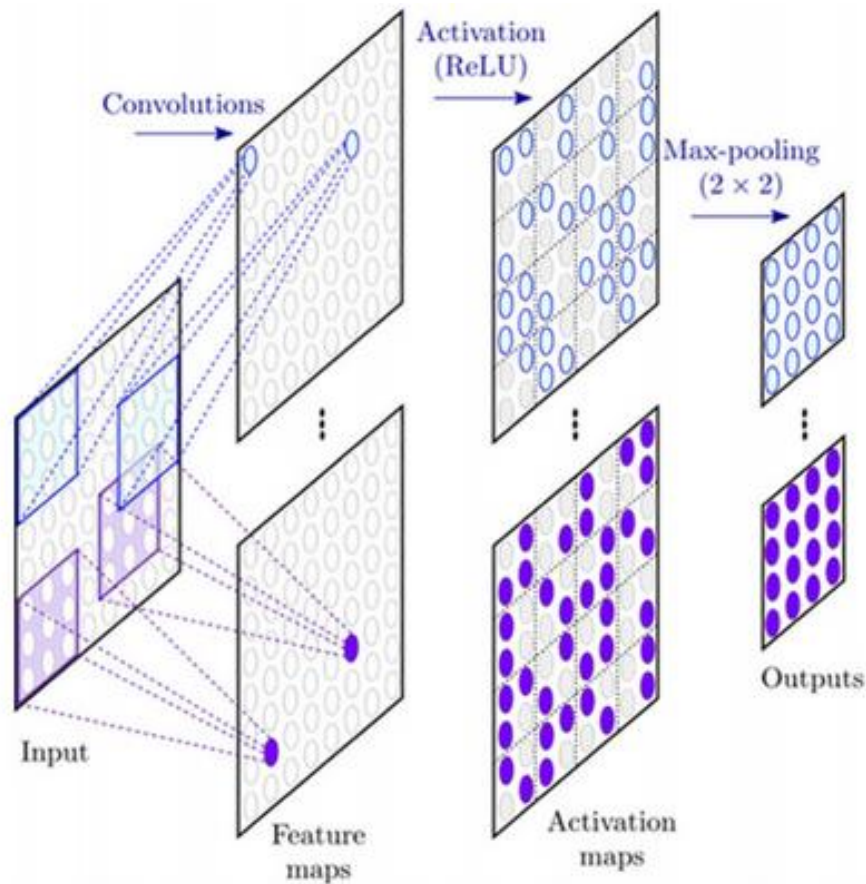
Sloj sažimanja (engl. *Pooling*) djeluje na prostorne dimenzije ulazne slike. Uloga ovog sloja jeste izvršiti redukciju podataka te ujedno smanjiti potrebne računalne resurse za obradu podataka u mreži. Razlikuje se sažimanje maksimalnom vrijednošću (engl. *Max pooling*) i sažimanje prosječnom vrijednošću (engl. *Average pooling*). Sažimanje maksimalnom vrijednošću uzima samo najveću vrijednost iz mape značajki nad kojom se nalazi filter dok sažimanje prosječnom vrijednošću uzima vrijednost aritmetičke sredine svih vrijednosti nad kojima se nalazi filter. Slika 3.2 prikazuje rezultat navedenih dviju vrsta sažimanja nad ulaznim podacima. Konvolucijski sloj i sloj sažimanja zajedno čine *i*-ti sloj konvolucijske mreže te mreža može imati više takvih slojeva.

Potpuno povezani slojevi (engl. *Fully connected layers*) imaju između dva susjedna sloja sve neurone međusobno povezane te kao izlaz iz mreže daju vektor rezultata.



Slika 3.2 Vizualizacija operacije sažimanja (engl. *pooling*)

Postoji neograničen broj mogućnosti kako ove jednostavne slojeve povezati u cjelinu te time omogućiti prilagodbu neuronske mreže odgovarajućem problemu. Uobičajeno je da su prvi slojevi mreže zaduženi za obradu značajki niskog nivoa (npr. rubova, boja, gradijenata itd.) dok se u narednim slojevima obrađuju značajke višeg nivoa. Veći broj slojeva označava višu razinu naučenih značajki jer svaki sloj transformira ulazne podatke u nove značajke. Sve dobivene značajke se na kraju kombiniraju u donošenju konačne odluke odnosno izlazne vrijednosti mreže. Struktura konvolucijske neuronske mreže je prikazana na slici 3.3. Za razliku od običnih neuronskih mreža, slojevi CNN arhitekture imaju tri dimenzije: visinu, širinu i dubinu te su neuroni spojeni s malom regijom slojeva prije njih, umjesto da je svaki neuron spojen sa svim ostalim slojevima.

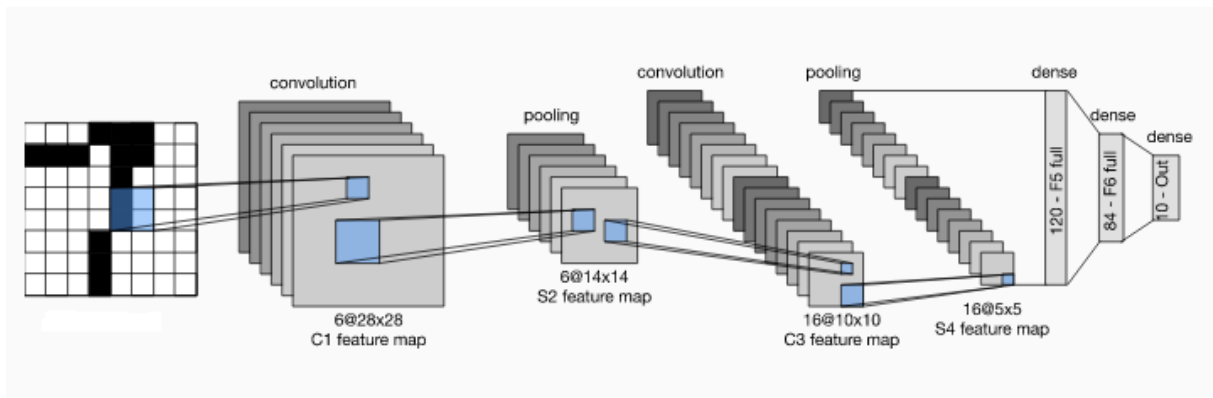


Slika 3.3 Struktura CNN mreže. Preuzeto iz [7]

3.2 Standardne duboke arhitekture

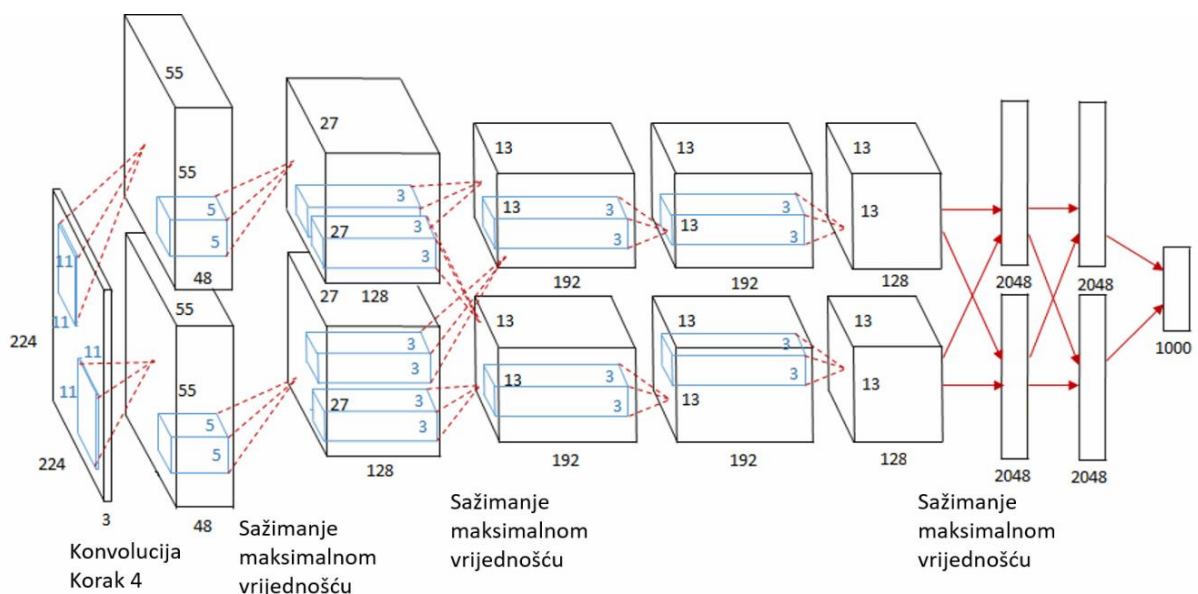
Određene duboke arhitekture su postale svojevrsni standard na području dubokog učenja te su od važnosti za problem segmentacije.

LeNet5 je prva uspješna primjena konvolucijske neuronske mreže za prepoznavanje rukom pisanih znamenki [2]. Sastoji se od dva konvolucijska sloja i tri potpuno povezana sloja što čini skupa 5 slojeva te otud u imenu mreže broj 5. Ova mreža je postala svojevrsni predložak za buduće arhitekture što podrazumjeva slaganje konvolucijskih slojeva sa slojevima sažimanja i završavanje sa potpuno povezanim slojevima. Zadnji dio mreže čini višeslojni perceptron koji se sastoji od skrivenog sloja i sloja logističke regresije. Logistička regresija određuje konačnu klasifikaciju što u ovom slučaju može biti jedan od 10 izlaza (jedna klasa za svaku znamenku). Slika 3.4 prikazuje LeNet5 arhitekturu.



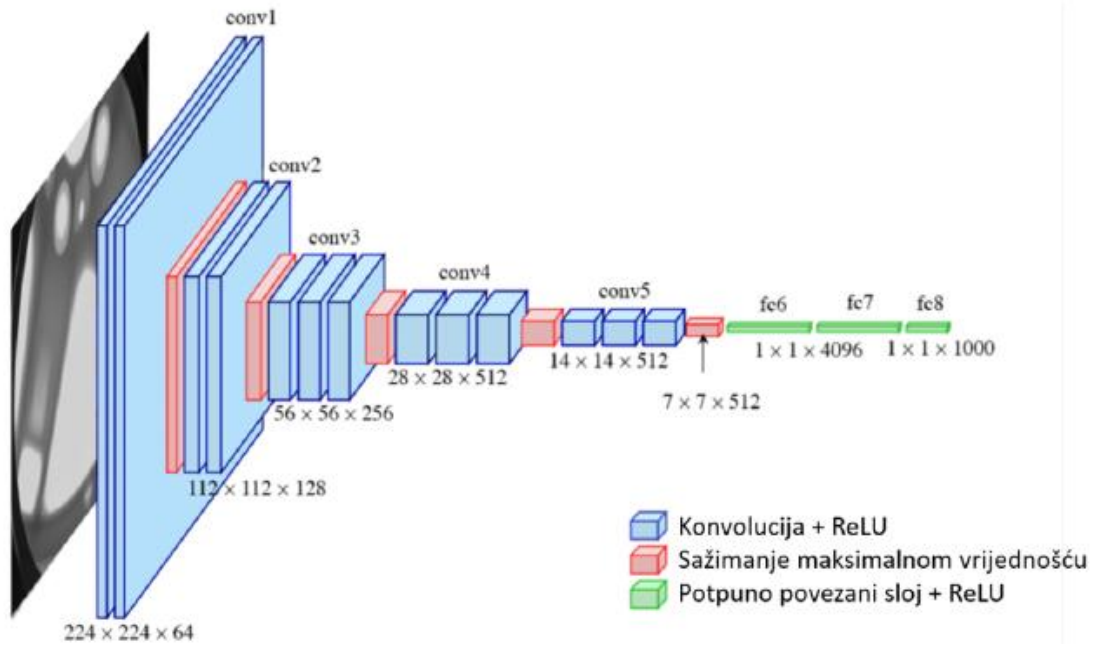
Slika 3.4 Prikaz LeNet5 arhitekture. Preuzeto iz [8]

2012. godine je predstavljena neuronska mreža pod imenom AlexNet [9]. Sličnih je karakteristika kao prethodno objašnjena arhitektura mreže LeNet5 i obje su poprilično plitke arhitekture. AlexNet je imala 5 konvolucijskih slojeva i 3 potpuno povezana sloja. Novost u njihovom radu jeste korištenje ReLU aktivacijske funkcije te spominju da je njihova arhitektura jedna od najvećih do tada korištenih na ImageNet bazi. Arhitektura ove mreže je prikazana slikom 3.5.



Slika 3.5 Arhitektura AlexNet mreže [9]

Nakon 2012 godine je počelo intenzivno istraživanje neuronskih mreža pogotovo dubljih modela neuronskih mreža. 2013. je istražena dublja arhitektura uz korištenje fiksnih manjih filtera na svakom sloju [10]. Taj model neuronskih mreža se naziva VGGNet (*engl. Visual Geometry Group*). Vizualni prikaz arhitekture se nalazi na slici 3.6. Arhitektura je slična kao AlexNet i preuzeli su korištenje ReLU aktivacijske funkcije samo imaju više konvolucijskih slojeva koji su filteri manje veličine. Reduciranjem veličine filtera smanjen je broj parametara koji se moraju naučiti što utječe na bržu konvergenciju i povećanja nelinearnosti te vodi lakšem treniranju mreže. Najpoznatije verzije su VGG16 i VGG19. VGG16 za razliku od VGG19 ima tri konvolucijska sloja manje. Slika 3.7 prikazuje konfiguraciju svih VGG mreža. Model VGG16 je postao popularan zbog postizanja TOP-5 točnosti od 92.7% na ILSVRC-2014 [10].



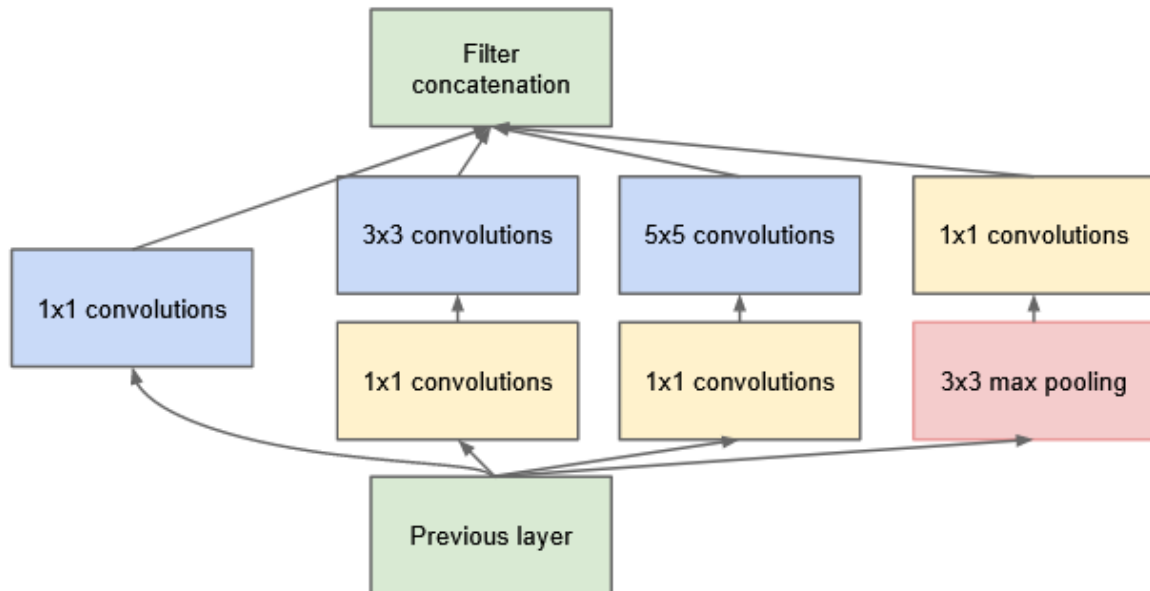
Slika 3.6 Vizualni prikaz arhitekture VGG-16 [10]

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224×224 RGB image)					
conv3-64	conv3-64 LRN	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 conv1-256	conv3-256 conv3-256 conv3-256	conv3-256 conv3-256 conv3-256 conv3-256
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

Slika 3.7 Konfiguracija VGGNet mreža. Stupac D predstavlja mrežu VGG16, dok stupac E predstavlja mrežu VGG19. Ostali stupci su ranije verzije iste mreže. Preuzeto iz [10]

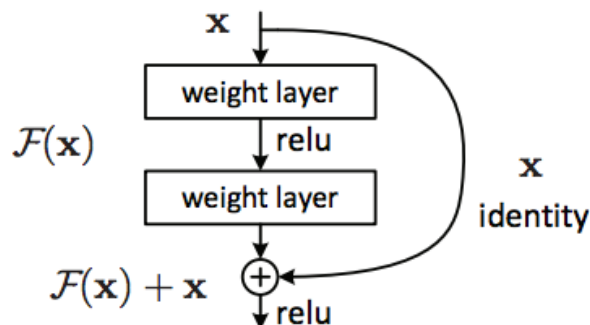
U radu [11] autori su predstavili 22-slojnu mrežu nazvanu GoogLeNet koja koristi takozvane *inception* blokove koji poboljšavaju efikasnost procesa treniranja i smanjuju količinu parametara. Ova arhitektura je osvojila izazov ILSVRC-2015 s TOP-5 točnošću od 93.3%. Novost u odnosu na prethodne arhitekture je u tome što se ne dodaju samo konvolucijski slojevi

nego se dodaju moduli koji u sebi sadrže konvolucijske slojeve. Na taj način autori su pokazali da slojevi dubokih neuronskih mreža ne moraju biti poredani nužno u sekvencijalnom redu kako je to bila prijašnja praksa. *Inception* blok se sastoji od *Network in Network* sloja, sloja sažimanja te velikog i manjeg konvolucijskog sloja. Svaki od slojeva se računa paralelno te je praćen operacijom konvolucije s filterom 1x1 radi smanjivanja dimenzija. Slika 3.8 prikazuje *inception* blok.



Slika 3.8 Prikaz inception bloka. Preuzeto iz [11]

ResNet arhitektura [12] je pobijedila na ILSVRC-2016 izazovu s točnošću od 96.4%. Predstavili su je znanstvenici iz Microsofta. Prvi su koristili vezu za preskakanje blokova takozvani *residual* ili preostali blokovi te su kreirali izrazito duboku arhitekturu od 152 sloja. Preostali blokovi su prikazani na slici 3.9. Uvođenjem *residual* blokova rješavaju problem treniranja dubokih arhitektura i problem nestajućeg gradijenta. Uloga veza za preskakanje jeste da svaki sloj može naučiti nešto novo budući da ima ulaz i izlaz od prethodnog sloja te njegov nepromijenjeni ulaz.



Slika 3.9 Preostali blokovi. Preuzeto iz [12]

4 PRIMJENA DUBOKOG UČENJA U SEGMENTACIJI MEDICINSKIH SLIKA

Konvolucijske neuronske mreže su korištene u analizi medicinskih slika od 1995. godine [13]. Međutim, iako su imale uspješne rezultate, nisu odmah postale popularne. Popularne su postale tek kad su se razvile nove tehnike za efikasnije procese treniranja mreže. Kako bi se bolje objasnila primjena dubokog učenja na medicinskim slikama, navest će se različiti načini prikaza i kreiranja medicinskih slika što se naziva modalitet slike.

4.1 Modalitet medicinskih slika

Medicinske slike omogućuju vizualizaciju anatomskih struktura i prikaz informacija o funkciji ljudskog tijela. Stvaranjem medicinskih slika postiže se razumijevanje procesa unutar tijela, funkcije organa te promjena koje nastanu zbog određenih uvjeta ili bolesti. Sustavi za dobivanje medicinskih slika se razlikuju po vrsti energije koju koriste za prikaz. Vrsta prikaza se naziva i modalitet. Razlikuju se prikazi temeljeni na [14]:

- rendgenskom zračenju. Tu spadaju klasična rendgenska snimanja, mamografija, tomografija, kompjutorizirana tomografija (engl. *computed tomography*, CT)
- magnetskoj rezonanciji (engl. *Magnetic resonance imaging*, MRI)
- raspodjeli radionuklida (pozitronska emisijska tomografija-PET)
- pronalasku, emisiji ili apsorpciji svjetlosti (mikroskopija)
- rasprostiranje ultrazvuka
- i ostali prikazi.

U primjeni metoda dubokog učenja za obradu medicinskih slika najčešće su korišteni modaliteti kompjutorizirane tomografije i magnetske rezonancije [15].

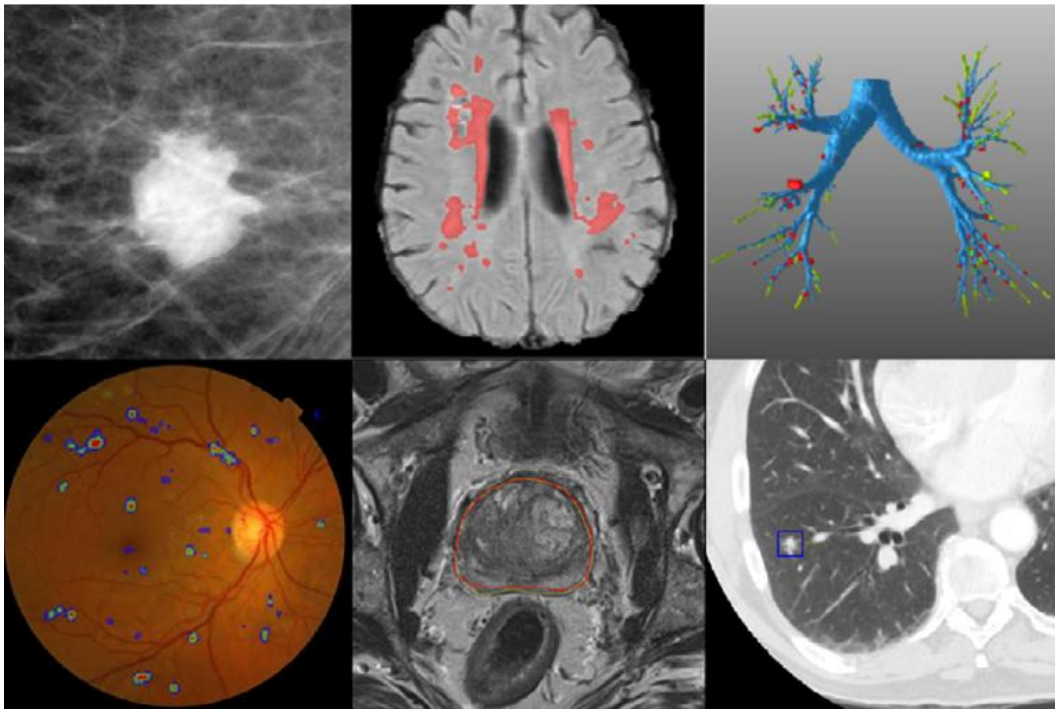
Kompjutorizirana tomografija je najčešći izbor za prikazivanje koštanog sustava međutim koristi se samo ako je nužno jer izlaže pacijente zračenju [16]. Magnetska rezonancija se najčešće koristi za prikaz organa, mekog tkiva, ali i koštanog sustava. Tehnika koristi snažno magnetsko polje i svojstva vodikova atoma prisutna u tijelu kako bi se stvorila slika [17]. Razlikuju se PD (engl. *proton density*), T1 i T2 slike magnetske rezonancije u ovisnosti o korištenim parametrima prilikom stvaranja slike. Konačno, magnetska rezonancija se sastoji od sekvence slika (engl. *slices*) koja prikazuje željenu strukturu.

4.2 Segmentacija medicinskih slika

Neuronske mreže se u analizi i obradi medicinskih slika najčešće koriste za zadatke klasifikacije, detekcije i segmentacije. Klasifikacija slika podrazumjeva prepoznavanje klase objekta unutar slike te dodjeljivanje oznake slici. Klasifikacija medicinskih slika je jedno od prvih zadataka na kojem se duboko učenje pokazalo učinkovito posebno konvolucijske neuronske mreže [18], [19]. U klasifikaciji medicinskih slika je ulazni podatak slika, a konačan rezultat predstavlja jedna varijabla odnosno oznaka prisutnosti/odsutnosti bolesti.

Detekcija objekta podrazumjeva klasifikaciju i detekciju objekta unutar slike. Dakle, nije dovoljna oznaka je li bolest prisutna ili ne, nego mora biti poznato i na kojem području je abnormalnost. To znači da se za svaki objekt mora znati klasa, veličina i pozicija unutar slike. Detekcija objekata od interesa unutar medicinske slike jeste jedan od ključnih dijelova dijagnoze i jedan od zahtjevnijih zadataka za medicinsko osoblje. Obično se sastoji od pronalaska i identificiranja određenih masa ili lezija unutar medicinske slike. Prvi sustav za detekciju objekata koristeći CNN napravljen je još 1995. godine s četiri sloja za detekciju

čvorića na rendgenskim slikama [20]. Slika 4.1 prikazuje neke primjere primjene dubokog učenja na analizu i obradu medicinskih slika.



Slika 4.1 Primjeri primjene dubokog učenja na analizu i obradu medicinskih slika. Od lijevog gornjeg kuta prema dolje desno: klasifikacija masa na mamogramima [21], segmentacija lezija na mozgu (top rangirano na BRATS, ISLES i MRBrain izazovima) [22], detekcija propusnosti dišnih puteva u segmentaciji dijela pluća [23], klasifikacija dijabetičke retinopatije na oku (Kaggle Diabetic Retinopathy izazov 2015) [24], segmentacija prostate (top rangirano na izazovu PROMISE12), klasifikacija čvorića u plućima (CAMELYON16 izazov)

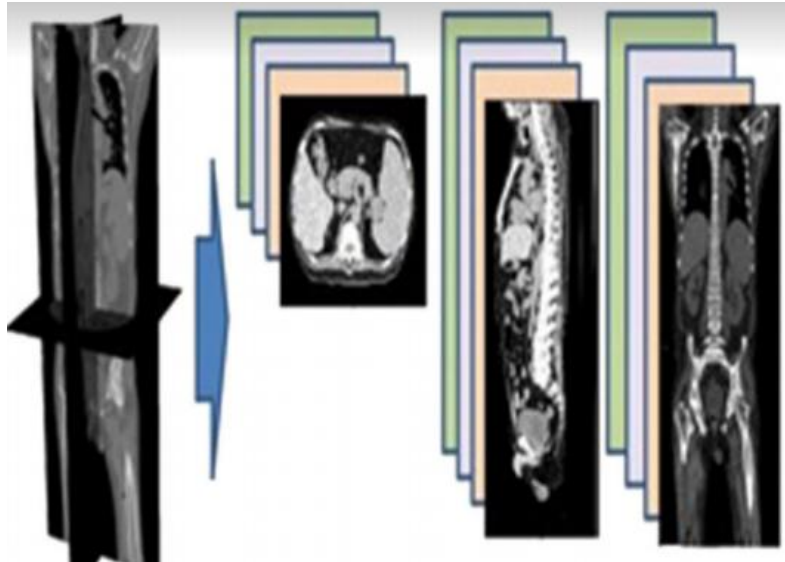
Segmentacija je čest zadatak u obradi medicinskih slika. Identificiranje piksela koji čine organ ili drugu podstrukturu u odnosu na pozadinu slike jeste zahtjevan posao jer rezultira ključnim informacijama o obliku i volumenu tih tkiva. Rezultat segmentacije jeste slika visoke rezolucije i iste veličine kao originalna slika gdje je svaki piksel klasificiran u određenu klasu. Dakle, to je klasifikacija na razini piksela. Prije pojave neuronskih mreža, segmentacija medicinskih slika se temeljila na tradicionalnim tehnikama koje su najčešće uključivale detektore rubova i razne matematičke metode. Potom su se razvile metode strojnog učenja koje su kao najveći problem imale ručno izvlačenje značajki.

Automatska segmentacija tkiva u medicinskim slikama posebno na slikama magnetske rezonancije je od velike važnosti za pomoć u liječenju pacijenata. Stoga, kako bi se potaknuo razvoj automatskih segmentacijskih tehnika specifičnih za medicinske slike održavaju se razni segmentacijski izazovi za segmentaciju određenog organa. Neki od najpoznatijih su: ISLES (engl. *Ischemic Stroke Lesion Segmentation*) [25], BRATS (engl. *Multimodal Brain Tumor Image Segmentation*) [26], MRBrainS (engl. *MR Brain Image Segmentation*) [27] i LV izazov (engl. *MR Left Ventricle*) [28].

Prema vrsti ulaznih podataka u svrhu segmentacije medicinskih slika konvencionalne konvolucijske neuronske mreže se mogu podijeliti na: 2D, 2.5D i 3D konvolucijske neuronske mreže. Kad se spominje izraz konvolucijske neuronske mreže u pravilu se misli na 2D konvolucijske mreže. Takve mreže izvode segmentaciju nad 2D ulaznim podacima korištenjem 2D filtera kako je opisano u poglavlju 3.

4.3 2.5D konvolucijske neuronske mreže

Jedan od izazova u segmentaciji medicinskih slika jeste prilagoditi postojeće arhitekture za različite ulazne formate kao što su podaci u volumenu odnosno trodimenzionalni podaci npr. snimci magnetne rezonance. U ranim primjenama CNN mreža za trodimenzionalne podatke koristila se ideja o dijeljenju volumena (engl. *Volume of interest, VOI*) na pojedinačne 2D slike (engl. *slice*) te se 2D slike potom predaju kao ulaz u neuronsku mrežu. Prva primjena tog pristupa je izvršena za segmentaciju koljena još 2013. godine [29]. Autori primjenjuju tri posebne 2D CNN mreže kako bi svaka obradila prikaz izvučen iz volumena u odgovarajućoj ortogonalnoj ravnini xy, yz i zx. Slika 4.2 prikazuje način podjele volumena u tri ravnine.



Slika 4.2 Prikaz dijeljenja 3D volumena prema ortogonalnoj ravnini na 2D prikaz. Preuzeto iz [30]

Nadalje, autori u radu [31] koriste istu ideju tako što različite ortogonalne ravnine kombiniraju i tretiraju kao tri kanala ulazne slike.

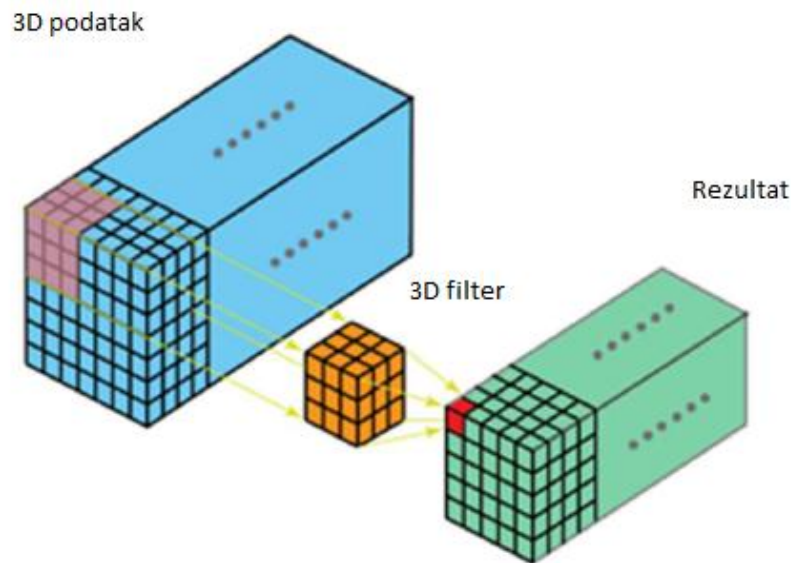
Ovi pristupi se nazivaju 2.5D konvolucijske neuronske mreže zbog više prostornih informacija koje sadrže slike iz istog volumena o susjednim pikselima u odnosu na 2D pristup dok je istovremeno potrebno manje resursa za obradu informacija u odnosu na obradu cijelog 3D volumena. Zbog ublažavanja problema dimenzionalnosti medicinskih podataka, a istodobnog zadržavanja obrade 2D slika, ovakav pristup koristi više autora [30], [32] i [4].

Iako se ovakav pristup čini optimalan s obzirom na performanse, pojedini autori misle da nije dovoljno ni optimalno koristiti samo 3 ortogonalna prikaza iz 3D podataka [33]. Primjena 2.5D strukture je bio dobar pokušaj uključivanja prostornih informacija u obradu međutim i dalje se u arhitekturi koriste 2D filteri.

4.4 3D konvolucijske neuronske mreže

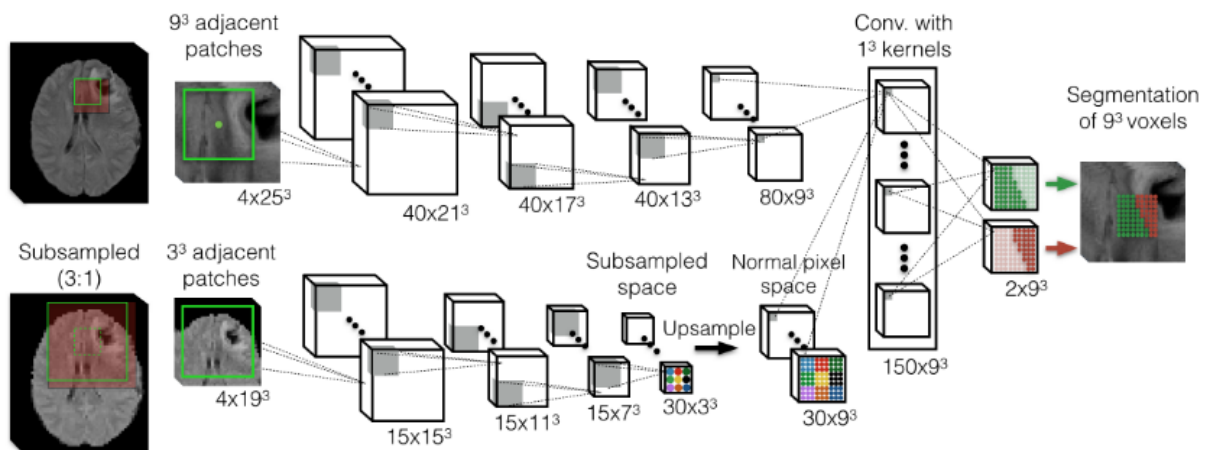
Iako su 2D i 2.5D konvolucijske neuronske mreže postigle značajan uspjeh u segmentaciji medicinskih slika, segmentacija 3D organa i struktura još predstavlja težak zadatak zbog kompliciranih anatomskih prikaza u volumenu, teškoća u optimiziranju 3D mreža te neadekvatnih uzoraka za treniranje. Struktura mreže je općenito slična 2D strukturi konvolucijske neuronske mreže sa razlikom u korištenju 3D konvolucijskih slojeva. Slika 4.3 prikazuje trodimenzionalni filter koji se kreće u sva tri pravca pri obradi ulaznog trodimenzionalnog podatka. Poboljšanje računalnih resursa i općenito medicinskih uređaja za snimanje 3D podataka dovelo je do toga da se pokušaju prostorne informacije u potpunosti

iskoristiti te segmentirati i obrađivati u punom volumenu. Volumetrijske slike mogu pružiti informacije u svim pravcima za razliku od 2D prikaza i tri ortogonalna prikaza u korištenju 2.5D pristupa.



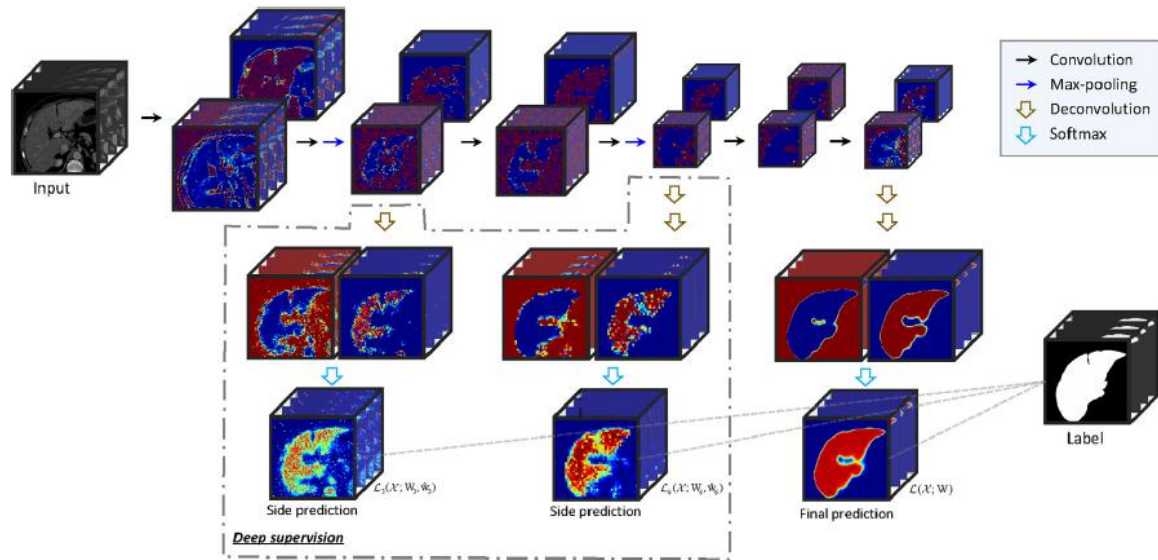
Slika 4.3 Prikaz obrade 3D podatka korištenjem 3D filtera. Preuzeto i prilagođeno iz [34]

Jedan od prvih modela koji u potpunosti koriste 3D pristup je uveden za segmentaciju tumora na mozgu [35]. Sličnu ideju koriste u radu [36] koji je predstavio multiskalarnu dualnu CNN mrežu u kojoj postoje dva paralelna tijeka obrade s istom veličinom receptivnog polja. Druga grana omogućuje obradu dijelova sa slike tako da se obrađuju veća područja oko *voxela* čime se postiže bolje shvaćanje i obrada konteksta. *Voxel* predstavlja jedinicu informacije odnosno točku u trodimenzionalnom prostoru. Ove modifikacije skupa sa korištenjem manjih filtera veličine 3x3 postižu bolju točnost sustava i manje vrijeme obrade u usporedbi sa prijašnjim pristupima. Slika 4.4 predstavlja objašnjenu arhitekturu.



Slika 4.4 Multiskalarna dualna CNN iz rada [36]

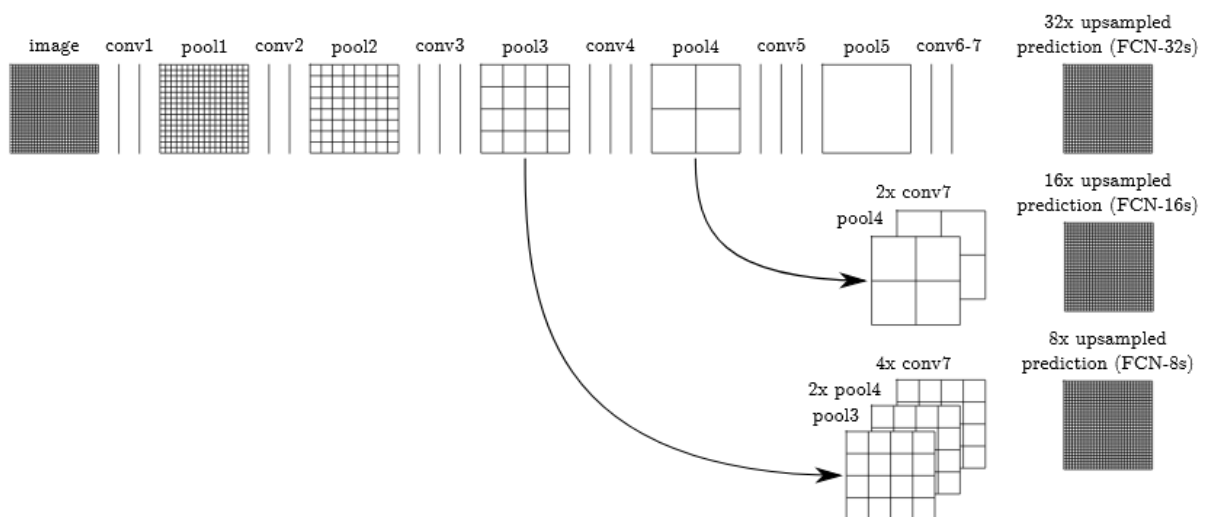
Dou u radu [37] predlaže korištenja 3D filtera koji dijele težine prostorno kako bi se smanjio broj parametara te time poboljšalo vrijeme obrade. Arhitektura je prikazana na slici 4.5



Slika 4.5 Prikaz 3D arhitekture korištene za segmentaciju jetre [37]

4.5 Potpuno konvolucijske mreže

Potpuna konvolucijska mreža (engl. *Fully Convolutional Network, FCN*) [38] se razlikuje od obične konvolucijske mreže po zadnjem sloju unutar mreže gdje FCN ima potpuno konvolucijski sloj umjesto potpuno povezanog sloja. Ova izmjena dopušta mreži da radi predviđanja na razini piksela za cijelu sliku umjesto za dijelove slike (engl. *Patch-wise prediction*) te se cijela slika obradi u samo jednom prolazu kroz mrežu. Za postizanje boljih performansi aktivacijske mape visoke rezolucije s početnih slojeva slojeva se kombiniraju s rezultatima dubljih slojeva i prosljeđuju konvolucijskom sloju kako bi imali točniji rezultat. Arhitektura mreže je predstavljena na slici 4.6.



Slika 4.6 Potpuna konvolucijska mreža. Slika preuzeta iz [38]

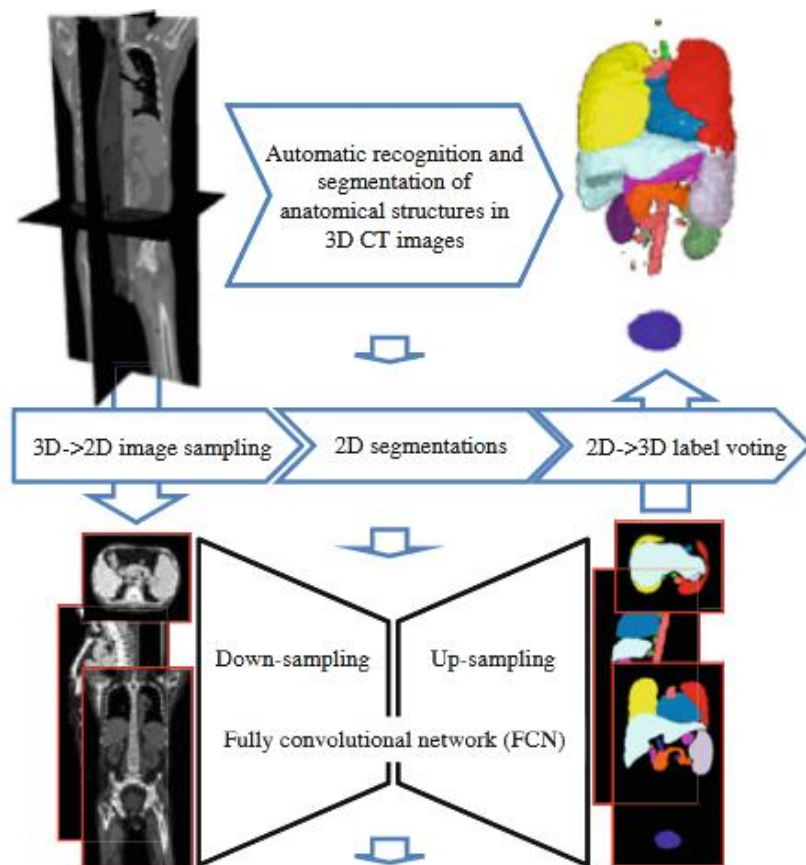
Superiornost FCN mreže nad CNN mrežom je jasno prikazana u radovima [39] i [40] koji izvode potpuno isti eksperiment koristeći istu bazu slika za izvođenje istog zadatka. Autori rade segmentaciju tkiva mozga kod djece starosti 6-8 mjeseci na bijelu tvar, sivu tvar i moždanu

tekućinu u svrhu procjene razvoja zdravog mozga te ranih dijagnoza bolesti. U radu [39] se segmentacija vrši korištenjem konvolucijske neuronske mreže na više različitih modaliteta magnetske rezonancije mozga. Iste ciljeve segmentacije na istoj bazi slika imaju autori u radu [40] s tim da je njihov odabir arhitekture potpuno konvolucijska mreža i postignuti rezultati segmentacije su bolji.

4.5.1 Potpuno konvolucijske mreže za istovremenu segmentaciju više organa

Mreža FCN se pokazala učinkovitom za segmentaciju više organa istovremeno (engl. *Multi-organ segmentation*). Segmentacija više organa podrazumijeva istovremenu segmentaciju više različitih tkiva te se koristi najčešće u obradi slika abdomena [41]. Zhou i ostali [30] koriste FCN za 2.5D pristup segmentaciji 19 organa na 3D slikama kompjutorizirane tomografije. U radu koriste ukupno 3 FCN mreže, po jednu za svaki 2D prikaz iz volumena. Za svaki piksel se uzmu u obzir svi segmentacijski rezultati da bi se generirao konačan rezultat. Ovakav pristup je pokazao veću točnost kod velikih organa (Dice vrijednost 0.937), dok je manja točnost segmentacije manjih organa (Dice vrijednost 0.553). FCN mreže se koriste i za segmentaciju više organa s 3D slika [42].

U radu [43] autor segmentira 17 različitih područja organa korištenjem potpuno konvolucijske mreže na vlastitoj bazi podataka koja se sastoji od 240 CT slika. Postupak segmentacije je izvršen korištenjem 2D i 3D FCN mreže i prikazan na slici 4.7.



Slika 4.7 Prikaz postupka segmentacije 17 različitih regija organa iz rada [43].

4.5.2 Kaskadne potpuno konvolucijske mreže

Christ i ostali [44] su pretpostavili da postavljanjem FCN mreža kaskadno mogu poboljšati točnost segmentacije lezija na jetri. Osnovna ideja jeste da se serija FCN mreža složi na način da svaki model koristi kontekstne značajke iz mape značajki prethodnog modela. Autori koriste

dvije mreže kaskadno tako da prva FCN obavlja segmentaciju područja od interesa na jetri za drugu mrežu koja potom segmentira lezije. Prednost korištenja ovog dizajna jeste u tome da se mogu primijeniti različiti filteri za svaki stadij obrade te time se može značajno povećati točnost segmentacije. Isti postupak su koristili u radu [45] gdje je postignut Dice rezultat 0.823 za segmentaciju lezija na CT slikama te 0.85 na slikama magnetske rezonancije. Slično, u radu [46] Wu i ostali istražuju primjenu kaskadnih FCN na poboljšanje detekcije granica fetusa na ultrazvučnim slikama.

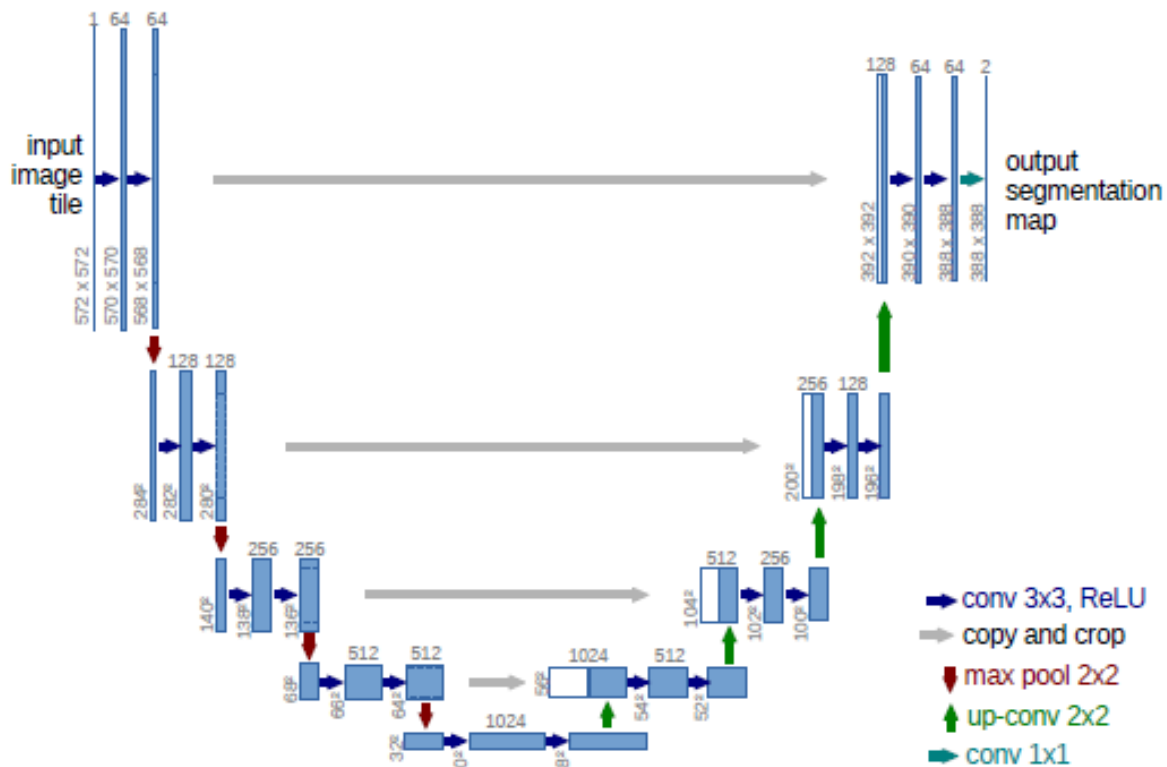
4.5.3 Multi-Stream potpuno konvolucijske mreže

Za obradu medicinskih slika mogu se upotrijebiti različite tehnike s obzirom da slike mogu biti različitih modaliteta (engl. *Multi-modality techniques*) i rezolucija (engl. *Multi-scale techniques*). Multi-stream dizajn neuronskih mreža može dopustiti sustavu da koristi različite formate slike za isti organ kako bi se poboljšali ukupni rezultati obrade. U radu [47] su primijenjene multi-stream tehnike na 3D FCN kako bi maksimalno iskoristili kontekstualne informacije za slike različitih rezolucija i različitih modaliteta kako bi se poboljšala robusnost sustava na različite varijacije struktura i oblika organa. Problem FCN mreža jeste detekcija objekata ukoliko se njihova veličina mijenja jer je receptivna veličina fiksna. Stoga se koriste multi-scale mreže gdje se slikama prije ulaza u mrežu mijenja veličina [5], [33], [48]. Međutim, dijeljenje parametara iste mreže na slikama različitih veličina nije potpuno učinkovito jer objekti različitih veličina zahtijevaju različite parametre za obradu. Dodatno rješenje za receptivna polja fiksne veličine jeste da se za slike koje su veće od predviđenog pogleda FCN primjenjuje tehnika *sliding-window* preko cijele slike [49].

4.5.4 U-Net

Najpoznatija FCN arhitektura korištena u obradi medicinskih slika je U-Net [50]. Arhitektura mreže je simetrična i sastoji se od dva glavna dijela: lijevi dio se naziva enkoder (u praksi često nazvan i *contracting* put ili *downsampling*), a desni se naziva dekoder (u praksi nazvan i *expansive* put ili *upsampling*). Enkoder znači da se slika visoke rezolucije smanjuje na nisku rezoluciju. Prolaskom kroz mrežu slika postaje sve manja zbog slojeva sažimanja. Međutim, broj filtera se povećava što omogućuje izvlačenje složenijih značajki iz slike. Može se zaključiti da nakon prolaska kroz enkoder dio mreže, mreža postaje poučenija o tome što se nalazi na slici, ali se gubi informacija gdje se to točno nalazi. Stoga, se javlja potreba za vraćanjem originalne veličine slike odnosno za prolazak kroz dekoder dio mreže. Dakle, korištenjem regularne konvolucijske mreže sa slojevima sažimanja nikada se neće dobiti izlaz visoke rezolucije kao što je bio ulaz u mrežu. Kako je već navedeno, rezultat semantičke segmentacije jeste cijela slika visoke rezolucije gdje su svi pikseli klasificirani, a ne samo jedna klasa ili oznaka na slici.

Postoji više interpolacijskih tehnika za realizaciju dekoder dijela mreže, međutim sa tim tehnikama mreža ništa ne uči nego se iste matematičke operacije ponavljaju svaki put. Stoga se u praksi najčešće koriste takozvane *transposed* konvolucije za realizaciju dekoder dijela mreže. Iako ovo nije prvi rad gdje se u konvolucijske mreže uvodi dekoder dio [38], autori uvode veze s preskakanjem (engl. *skip connections*) te time povezuju izravno enkoder sa odgovarajućim dekoder dijelom mreže. U-net radi na principu da su ulazna i izlazna slika jednake veličine. Iz perspektive treniranja mreže ova arhitektura omogućava obradu slike u samo jednom prolazu kroz mrežu i rezultat je segmentacijska mapa što ima prednosti u odnosu na mreže kojima je ulaz samo dio slike (engl. *patch*). Arhitekture U-Net mreže je predstavljena na slici 4.8.



Slika 4.8 U-net arhitektura [50]

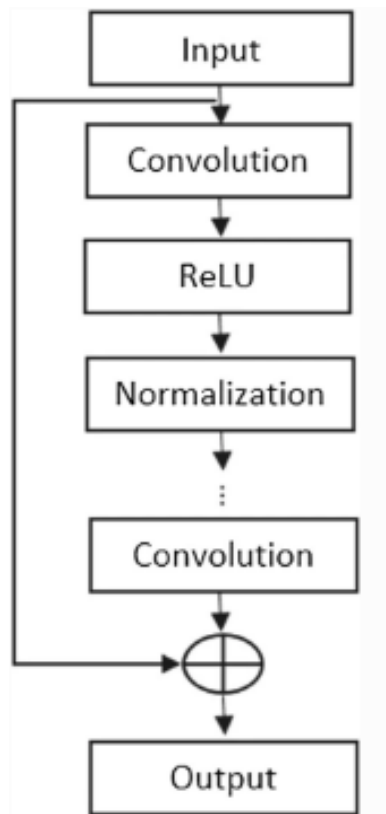
4.5.5 V-Net

Jedna od najpoznatijih inačica U-net mreže je V-net mreža predstavljena u radu [51] za 3D segmentaciju volumena. Primijenili su konvolucije na enkoder dijelu mreže za izvlačenje značajki i smanjenje rezolucije odabirom prikladnog filtera veličine $2 \times 2 \times 2$ i vrijednosti koraka 2. Konvolucije služe kao sažimanja sa prednošću korištenja manje memorije nego klasični slojevi sažimanja. Faza dekodera će izvući dodatne značajke i proširiti mape značajki niske rezolucije te konačno proizvesti volumetrijsku segmentaciju na zadnjem konvolucijskom sloju. Zatim, rezultat se pretvori u probabilističku segmentacijsku mapu i proslijedi *softmax* funkciji za segmentaciju *voxela* na objekt i pozadinu.

Gibson [52] koristi V-net mrežu s velikim receptivnim poljem koje pokriva od 50-100% ulazne slike te četiri različite rezolucije (*multi-scale*) za segmentaciju organa abdomena. Dobiveni rezultati imaju za 12% veću vrijednost Dice koeficijenta u odnosu na originalne V-net mreže.

4.6 ResNet

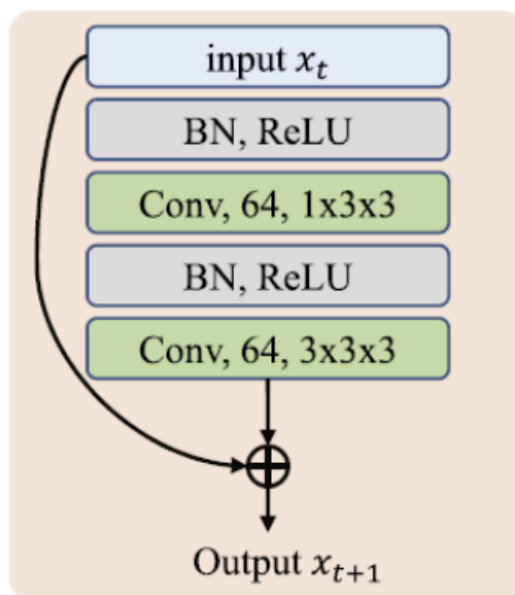
Kao što je već spomenuto, dublje arhitekture neuronskih mreža imaju veće sposobnosti učenja, ali i problem nestajućeg gradijenta. Dakle, sa povećavanjem dubine arhitekture dolazi do zasićenja točnosti te potom ubrzanog pada točnosti. Stoga, ResNet arhitekture koristi preostale blokove što omogućuje veću točnost i sve prednosti koje nosi dublja arhitektura. Slika 4.9 prikazuje preostale blokove korištene u arhitekturi ResNet.



Slika 4.9 Preostali blokovi ResNet mreže. Blok može imati različite kombinacije i broj slojeva zavisno od dizajna mreže. Preuzeto iz [7]

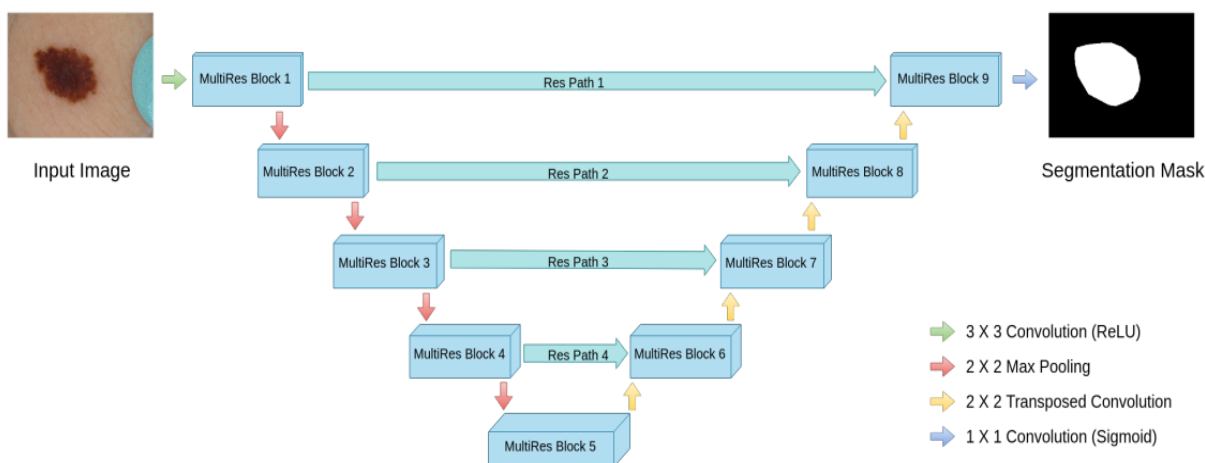
ResNet se uspješno koristi u segmentaciji medicinskih slika [53] međutim razvijene su i različite varijante ove arhitekture. Ovu arhitekturu su uzeli kao osnovu rada [5] te je dodatno modificirali tako da dobiju potpunu konvolucijsku residual mrežu (engl. *Fully convolutional residual networks-FCRN*). Prednost njihove arhitekture je sposobnost izvođenja segmentacije te korištenje dublje arhitekture za izvlačenje lokalnih i globalnih kontekstualnih značajki [54]. FCRN se sastojala od 50 slojeva te postigla rezultate sa Dice koeficijentom od 0.897 u segmentaciji lezija na koži u odnosu na 0.794 za VGG-16 model.

VoxResNet [55] je predložena verzija ResNet mreže za segmentaciju 3D podataka magnetne rezonance mozga. Sastoji se od 25 slojeva te je mreža sposobna izvući značajke iz više modaliteta kao i različitih skala (viša razina kontekstualnih informacija je integrirana u mrežu). Slika 4.10 prikazuje korišten VoxRes modul u arhitekturi mreže. VoxResNet je postigla segmentaciju s Dice koeficijentom 0.8696, 0.8061 i 0.8113 za T1, T1-IR i T2 modalitete.



Slika 4.10 Prikaz VoxRes modula korištenog u radu [55]

MultiResUNet je arhitektura predstavljena kao poboljšanje popularne U-Net mreže [56]. Novost u ovoj arhitekturi jeste u tome da se MultiResUNet sastoji od MultiRes blokova umjesto pojedinih konvolucijskih slojeva U-Net mreže. Nadalje, obične veze za preskakanje zamjenjuju s takozvanim *res* putevima (engl. *Res Path*) koji su sastavljeni od sekvence konvolucijskih slojeva čime pokušavaju smanjiti semantičku prazninu (engl. *semantic gap*) između odgovarajućih razina enkoder-dekoder arhitekture. Slika 4.11 prikazuje MultiResUNet arhitekturu. Postigli su bolje rezultate u odnosu na klasičnu U-Net arhitekturu na 5 poznatih baza slika različitih modaliteta (mikroskopske slike, dermatološke i magnetne rezonance).

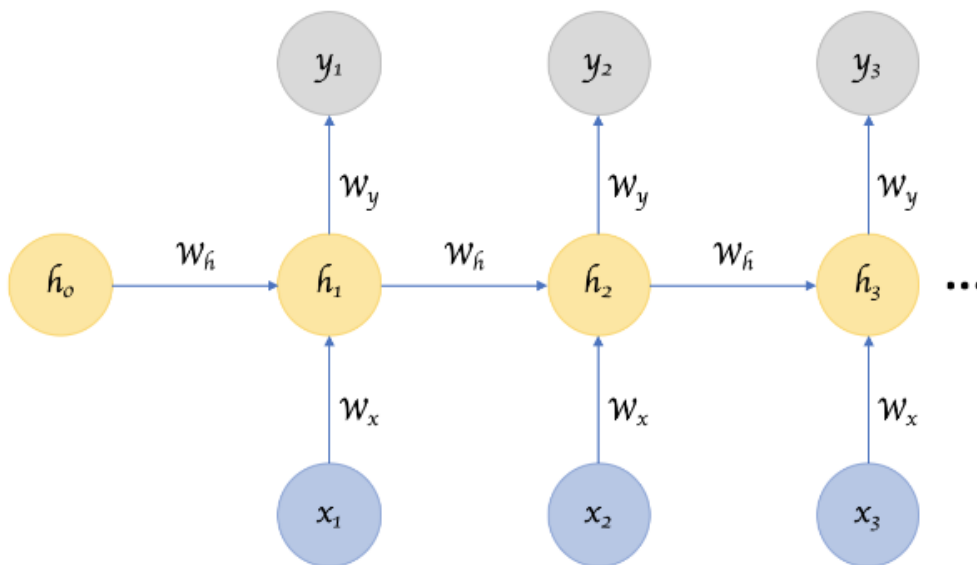


Slika 4.11 Prikaz MultiResUNet arhitekture korištene u radu [56]

4.7 Povratne neuronske mreže

Povratne neuronske mreže (engl. *Recurrent neural networks-RNN*) imaju povratne veze koje omogućavaju mreži pamćenje uzoraka iz prethodnih ulaza. Činjenica je da se područje od interesa (engl. *Region of interest, ROI*) na medicinskim slikama obično pojavljuje na više susjednih 2D slika (engl. *slices*) na CT ili magnetskoj rezonanciji. RNN su sposobne izvući *inter-slice* kontekst između ulaznih slika jer su specijalizirane da obrađuju sekvencu podataka

$x(t) = x(1), x(2), \dots, x(n)$ gdje je t vrijeme u rasponu od 1 do n . Za svaki element sekvence se izvršava ista obrada i izlaz ovisi od prethodnih uzoraka. RNN mogu imati jedan ili više ulaznih vektora te proizvesti jedan ili više izlaznih vektora koji ne ovise samo o težinama primijenjenim na ulazne podatke u procesu treniranja nego i statusu „skrivenog“ vektora koji predstavlja kontekst prethodnih ulaza. Stoga, mreža može za isti ulaz proizvesti različit izlaz na temelju svojih prethodnih ulaznih podataka u sekvenci. Slika 4.12 prikazuje arhitekturu RNN mreže.



Slika 4.12 Prikaz RNN mreže sa "skrivenim" stanjem h koje prenosi informaciju iz jednog ulaza prema ostalima u sekvenci [57]

4.7.1 LSTM

RNN mreže pamte prethodnu obradu i koriste tu informaciju za generiranje izlazne vrijednosti trenutnog ulaznog podatka. Međutim, za neke zadatke kao što su prepoznavanje govora ili predviđanje riječi u rečenici nekada treba znati cjelokupni kontekst prethodnih ulaza, a ne samo zadnji obrađeni podatak. Stoga se za učenje tih dugoročnih ovisnosti koristi najpoznatiji tip RNN mreža koje se zovu LSTM (engl. *Long Short Term Memory*). LSTM mreže su predstavljene 1997. godine [58]. Struktura se sastoji od tri jedinice nazvane vrata (ulazna vrata, izlazna vrata i *forget* vrata) te jedne memorijske jedinice. Stanje memorijske jedinice se mijenja kroz tri navedena vrata i na taj način LSTM odlučuje o odbacivanju nevažnih informacija i pamćenju korisnih informacija. Medicinske slike se mogu prikazati kao sekvenca *sliceova* (MRI) koji prikazuju isti organ te povezanost informacija na susjednim *sliceovima* i cjelokupan kontekst mogu pomoći u boljoj segmentaciji organa ili nepravilnosti. Upravo zbog sposobnosti pamćenja prethodnih obrada i učenja dugoročnih ovisnosti, LSTM mreža se koristi u segmentaciji medicinskih slika [59].

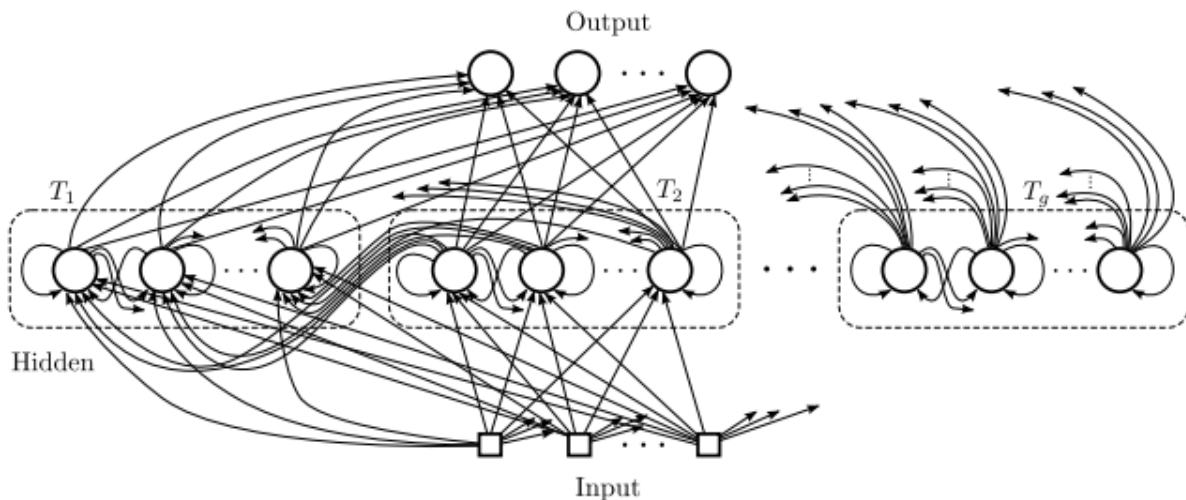
U standardnoj LSTM mreži ulazni podatak mora biti vektoriziran što je nedostatak zbog toga što će se za segmentaciju slika izgubiti prostorne informacije. Stoga je u radu [60] primijenjena konvolucijska LSTM mreža gdje je množenje vektora zamjenjeno s operacijom konvolucije. Tseng [61] je predložio korištenje konvolucijskog LSTM za modeliranje sekvence 2D *sliceova* za 3D biomedicinsku segmentaciju slika različitih modaliteta. U svom radu prikazuje koristi kombiniranja CNN i RNN arhitektura u segmentaciji tumora na mozgu na slikama magnetske rezonancije javno dostupne baze podataka BRATS-2015 [62].

U radu [59] je primijenjena kontekstualna LSTM na zadnji sloj duboke konvolucijske neuronske mreže kako bi se postigla bolja segmentacija korištenjem kontekstualnih informacija na susjednim *sliceovima*. Cilj istraživanja je segmentacija organa gušterače korištenjem dvije različite baze podataka koje se sastoje od magnetske rezonancije i kompjutorizirane tomografije. Njihova metoda je postigla Dice vrijednost 0.8247 što je znatno poboljšanje u odnosu na 0.7976 što je rezultat korištenja poznate U-net arhitekture [50].

Poznata verzija LSTM mreže je GRU (engl. *Gated Recurrent Unit*) koja uklanja memorijske jedinice te time pojednostavi strukturu bez degradacija u performansama [63]. U GRU strukturi izlazna i *forget* vrata su spojena u jedna vrata imena *update*. *Update* vrata se koriste za dobivanje trenutnog stanja izlaza korištenjem linearne interpolacije. Poudel [64] primjenjuje GRU u kombinaciji sa potpuno povezanom mrežom (FCN) za segmentaciju lijeve klijetke srca.

4.7.2 CW-RNN

CW-RNN (engl. Clockwork RNN) je verzija povratnih neuronskih mreža u kojoj je skriveni sloj podjeljen u module koji vrše obradu ulaznog podatka samo prateći vlastitu stopu rada (engl. *clock rate*) [65]. Između ulaznog, skrivenog i izlaznog sloja postoje unaprijedne veze. Neuroni iz skrivenog sloja su podjeljeni u g modula veličine k . Svaki modul ima dodijeljen *clock period* $T_n \in \{T_1, \dots, T_g\}$ i svaki modul je interno potpuno povezan, ali povratne veze od modula j do modula i postoje samo ako je period T_i manji od perioda T_j . CW-RNN smanjuje količinu potrebnih RNN parametara i poboljšava performanse. Skriveno stanje se prenosi od sporijih modula prema bržim modulima. Ova struktura je prikazana na slici 4.13.

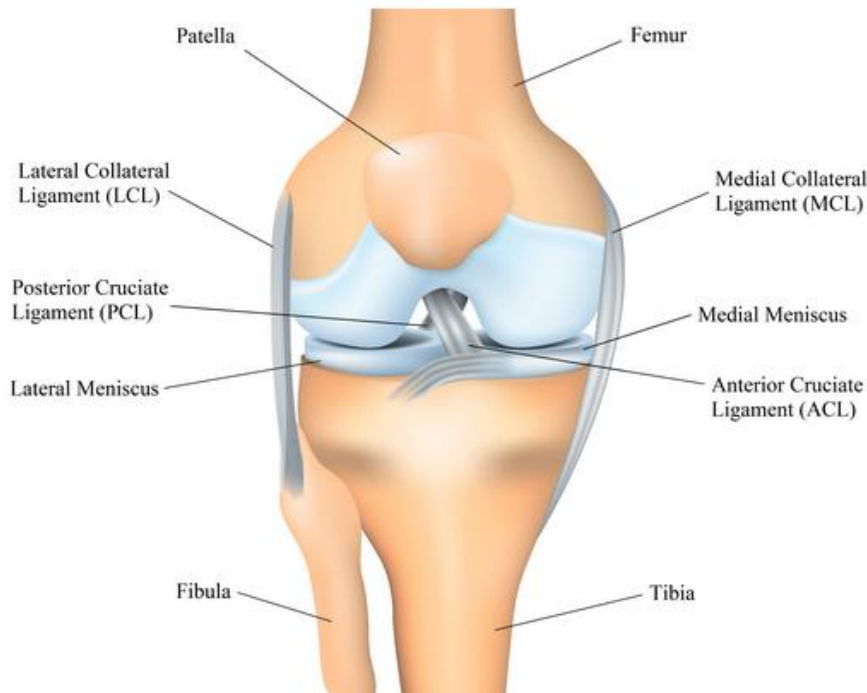


Slika 4.13 Prikaz CW-RNN strukture. Preuzeto iz [65].

CW-RNN je efikasniji u usporedbi s drugim pristupima budući da je samo dio aktivan u određenom trenutku [66]. CW-RNN mreža je korištena u radu [67] za segmentaciju mikroskopskih slika mišićnog tkiva. Općenito, RNN arhitekture imaju bolje performanse kada je cilj segmentacija većih organa jer takvi organi imaju više *inter-slice* informacija za razliku od manjih organa koji se mogu pojaviti samo na jednom *sliceu*.

5 SEGMENTACIJA SLIKA MAGNETSKE REZONANCIJE KOLJENA

Koljeno je najveći i jedan od najsloženijih zglobova ljudskog tijela. Koljeno podupire cjelokupnu tjelesnu težinu te je zbog opterećenja sklono čestim ozljedama. Glavni dijelovi koljena su *femur* (bedrena kost), *tibia* (goljenična kost), *fibula* (lisna kost), zglobne hrskavice, meniskus i ligamenti. Anatomija koljena je prikazana na slici 5.1.



Slika 5.1 Anatomija koljena. Preuzeto iz [68].

Većina radova istraživača o segmentaciji koljena odnosi se na segmentaciju hrskavice ili meniskusa u svrhu otkrivanja određenih bolesti [69], [16]. Osteoartritis je najčešća bolest mišićno-koštanog sustava posebno kod ljudi srednje i starije životne dobi. Za ranu dijagnozu osteoartritisa je vrlo važna segmentacija hrskavice na slikama magnetske rezonancije jer takve slike imaju veliki kontrast između tkiva što olakšava procjenu oštećenja hrskavice. Meniskus je poseban tip hrskavice koji nije dio kosti, nego se nalazi između kostiju i služi ublažavanju pritiska među krajevima kostiju. Automatska segmentacija koljena s naglaskom na segmentaciju kostiju i hrskavice je istražena u više radova [70], [71].

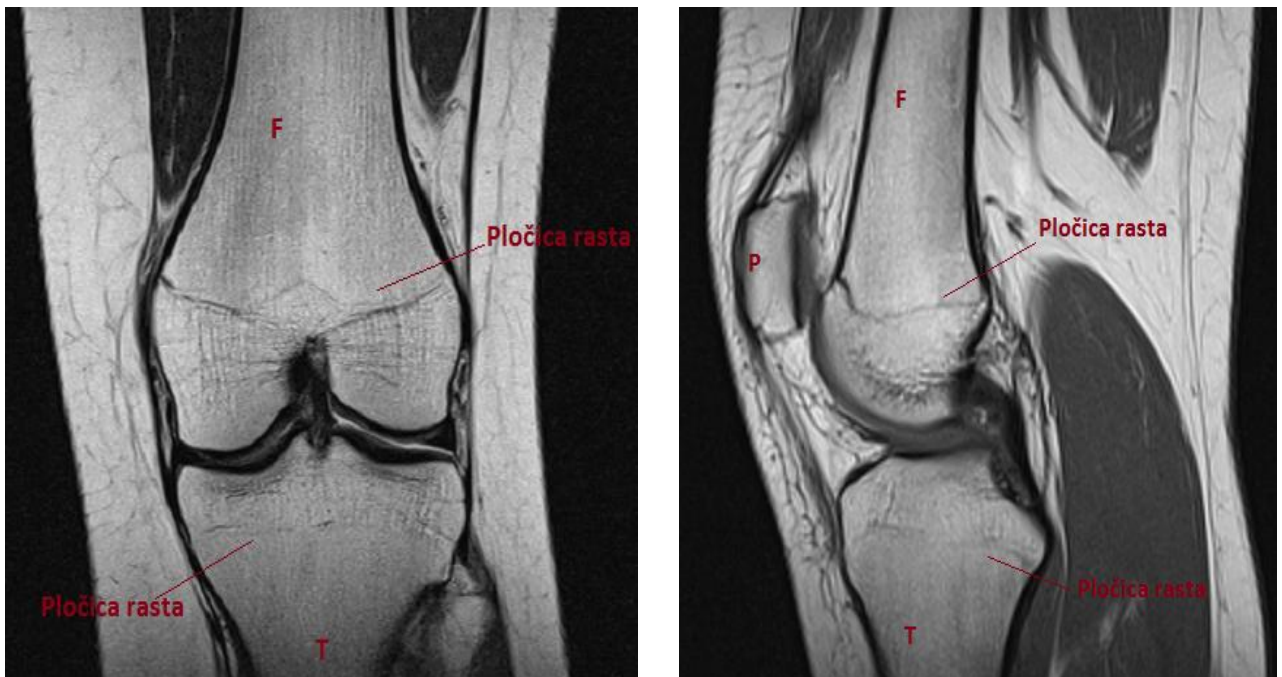
Najčešće korištene javne baze za segmentaciju koljena su SKI10 (engl. *Segmentation of Knee Images 2010 MICCAI*) [72] i OAI (engl. *Osteoarthritis Initiative*) [73].

Norman [71] istražuje kako se automatska segmentacije 638 T1 slika magnetske rezonancije koljena razlikuje po točnosti i preciznosti od ručne segmentacije koju je izveo radiolog. Njihov skup podataka uključuje pacijente s osteoartritisom i bez bolesti te rezultati pokazuju kako točnost i preciznost segmentacije hrskavice i meniskusa ne ovisi o tome ima li pacijent osteoartritis. Automatsku segmentaciju su izveli korištenjem 2D U-Net neuronske mreže, a rezultat su predstavili pomoću Dice koeficijenta sličnosti od 0.770 do 0.878 za segmentaciju hrskavice *femura*, *patelle* i *tibie*. Byra [74] u svom istraživanju predstavlja automatsku segmentaciju koja se zasniva na U-Net arhitekturi sa dodanim *attention* slojevima. Dodani slojevi služe za posebnu obradu mapi značajki koje se prenose putem veza s preskakanjem za poboljšanje segmentacije. Istraživanje se temelji na podacima magnetske rezonancije od 61 pacijenta gdje dva radiologa ručno segmentiraju slike i određuju ROI koji se smatra zlatnim

standardom. Točnost segmentacije za medialni meniskus varira od 83.1% do 87.2% dok za lateralni meniskus iznosi od 80.5% do 84.7% zbog razlike u ručnoj segmentaciji doktora. Gaj [75] radi segmentaciju na skupu podataka od 88 slika magnetske rezonancije iz OAI baze korištenjem generativne protivničke mreže (engl. *Generative adversarial networks*, GAN) mreže sa U-Net mrežom kao generatorom. GAN mreže su specifičan model koji je sposoban implicitno naučiti distribuciju podataka iz skupa podataka kako bi generirao nove podatke korištenjem naučene distribucije. Sastoji se od dvije konvolucijske mreže koje se natječu: generatora i diskriminatora. Točnost predstavljene segmentacije je 89.5% za lateralni i 87% za medialni meniskus.

Liu [76] primjenjuje 2D SegNet mrežu za segmentaciju kostiju i hrskavice koljena na bazi SKI10. Također za poboljšanje rezultata koriste algoritam 3D *simplex* modeling za ublažavanje granica deformacije i kostiju te očuvanje cjelokupnog oblika anatomije koljena. Slično, u radu [70] koriste konvolucijsku neuronsku mrežu te rezultate segmentacije poboljšaju korištenjem 3D uvjetnih slučajnih polja (engl. *condition random fields*, CRF) i 3D *simplex* modelinga. CRF koriste za poboljšanje veza među *voxelima* koji pripadaju istom tkivu, dok 3D *simplex* modeling služi za očuvanje cjelokupnog oblika i površina struktura.

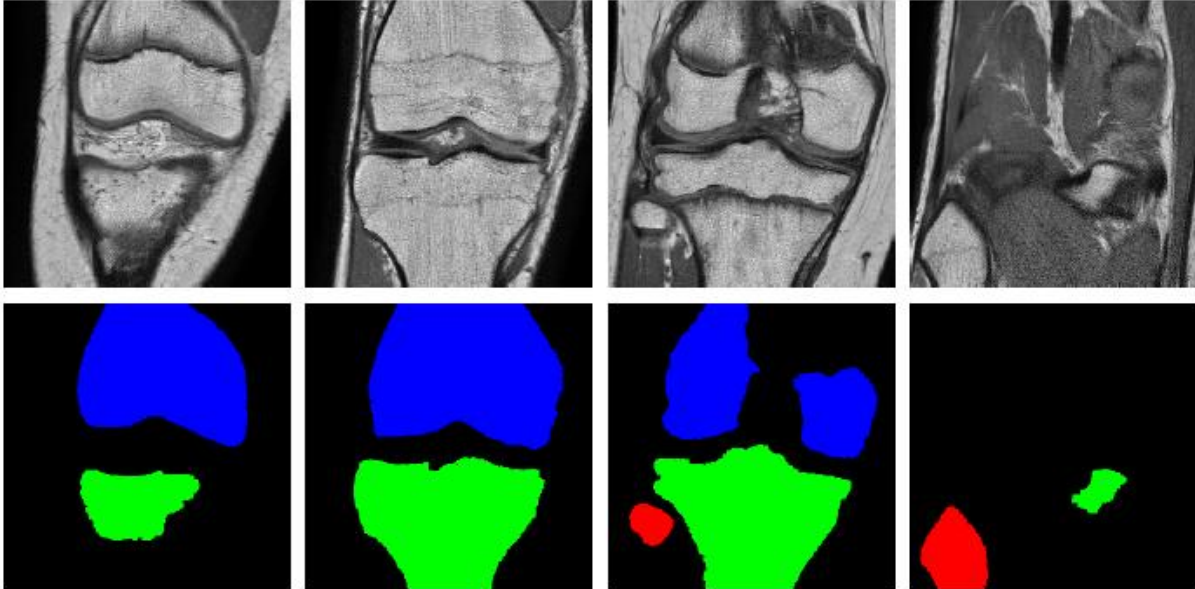
Navedena istraživanja su uglavnom usmjerena na otkrivanje neke bolesti i deformacija unutar koljena te su dostupne baze podataka sastavljene od magnetske rezonancije pacijenata starije životne dobi. U novije vrijeme tek par istraživanja se radi u smjeru procjene starosti pacijenata na temelju fizičkih karakteristika zgloba koljena [77], [78]. Naime, do određene starosti ljudi imaju na *femuru*, *tibiji* i *fibuli* takozvanu pločicu rasta (engl. *growth plate*) koja svojom površinom i volumenom ukazuje na određenu kronološku starost pacijenta. Slika 5.2 prikazuje pločice rasta na koronarnom i sagitalnom prikazu koljena.



Slika 5.2 Prikaz pločice rasta na koljenu. Lijevo: T2 koronarni prikaz magnetske rezonancije. Desno: Sagitalni prikaz istog pacijenta. Oznake: F-femur, T-Tibia i P-Patella

U radu [78] procjena starosti pacijenta se vrši na temelju klasifikacije konvolucijskim neuronskim mrežama predtreniranim na prirodnim slikama. Testirali su 7 različitih arhitektura konvolucijskih neuronskih mreža i najbolje rezultate postignu s predtreniranom GoogLeNet na

ImageNet bazi slika. Za razliku od ovog pristupa temeljenog na klasifikaciji, u radu [77] autori segmentiraju slike magnetske rezonancije koljena na tri glavne kosti (*femur*, *tibia* i *fibula*) korištenjem samo koronarnih *sliceova* iz sekvence. Koriste U-Net neuronsku mrežu za segmentaciju. Nakon segmentacije procjenjuju starost pacijenata. Slika 5.3 prikazuje rezultate segmentacije za različite *sliceove*.



Slika 5.3 Prikaz rezultata segmentacije koljena iz rada [77]. Plava boja predstavlja femur, zelena tibu i crvena predstavlja fibulu.

Glavni nedostatak ovih istraživanja jeste korištenje vlastitih baza podataka koje su izrazito male (samo 14 testnih subjekata u istraživanju [77]) zatim podjela pacijenata po spolu [78] i odabir točno određenih *sliceova* iz sekvence magnetske rezonancije.

6 PROCES UČENJA MREŽE

Učenje mreže se može odvijati na tri načina:

- nadzirano učenje,
- nenadzirano učenje i
- prijenosno učenje

Osnovna ideja nadziranog učenja je omogućiti direktni nadzor nad skrivenim slojevima umjesto samo na izlaznom sloju [79]. Kod GoogLeNet arhitekture nadzor je uključen u dva skrivena sloja od ukupnih 22 [11]. Dou [37] je koristio nadzirano učenje u svom pristupu segmentiranja CT snimaka jetre. U radu je prikazana bolja konvergencija, ali i manje vrijeme treniranja te manja greška validacije. Sličan pristup je objašnjen u radu [80]. Tri klasifikatora su određena za klasifikaciju rezultata srednjih slojeva na arhitekturi koja je slična U-net mreži. Klasificirani izlazi služe kao regulatori za fazu treniranja. Kontekstualne informacije s više razina u mreži su pomogle poboljšati sposobnosti lokalizacije i diskriminacije mreže.

Nenadzirano učenje ne zahtjeva znanje o željenom izlazu mreže. Mreža mora sama zaključiti sličnosti u skupu podataka. Postojeći nadzirani pristupi za automatsku segmentaciju slika zahtijevaju označavanje slika na razini piksela ili u slučaju 3D pristupa na razini *voxela* što nije uvijek dostupno u pojedinim slučajevima. Također, označavanje takvih slika podrazumijeva veliki utrošak vremena i energije stručnog kadra. Ideja uvođenja djelomičnih oznaka je predstavljena u radu [3] gdje autori označe samo jedan pixel kao prisutnost čvorića umjesto cjelokupni volumen. Ovaj pristup pokušava smanjiti ovisnost sustava o potpuno označenim slikama. Na temelju označenog piksela se izvuče okolni volumen i koristi kao pozitivan uzorak za treniranje mreže. Obično su čvorići u rasponu od 3 do 28 piksela široki te se nalaze na 3-7 susjednih *sliceova*. Metoda je pokazala razumnu osjetljivost od 80% sa slabo označenim podacima.

6.1 Prijenosno učenje

Treniranje dubokih konvolucijskih neuronskih mreža može biti težak zadatak jer zahtjeva velike količine označenih podataka za treniranje i stručnost za postizanje odgovarajuće konvergencije. Alternativa treniranju od početka jeste iskoristiti prethodno treniranu mrežu uz određene preinake. Prijenosno učenje je definirano kao sposobnost sustava da prepozna i upotrijebi znanje naučeno na prethodnim domenama za nove zadatke [81]. Treniranje mreže sa setom predtreniranih težina se naziva *fino* poboljšavanje (engl. *fine-tunning*). *Fino* poboljšanje počinje tako što se prenesu težine sa predtrenirane mreže na novu mrežu koja se želi trenirati. Iznimka je zadnji potpuno povezani sloj. Učestala praksa jeste zamjeniti zadnji potpuno povezani sloj sa novim potpuno povezanim slojem nakon toga se krene poboljšavati zadnji sloj te potom svi ostali slojevi mreže. Općenito početni slojevi CNN mreže su značajke niže razine koje su primjenjive na većinu zadataka računalnog vida, ali dublji slojevi uče značajke više razine koji su specifični za određene zadatke. Stoga, *fino* poboljšanje zadnjih par slojeva je u pravilu dovoljno za prijenosno učenje i u literaturi se naziva *plitko fino* poboljšanje (engl. *shallow tuning*). Međutim, razlika između primjene nove mreže i prethodno trenirane mreže određuje koliko će prilagođavanje biti potrebno. Stoga, nekada će se poboljšanje od zadnjeg sloja kretati prema ostalim slojevima sve dok se ne postignu odgovarajuće performanse. *Fino* poboljšanje svih konvolucijskih slojeva se naziva i *duboko poboljšanje* (engl. *deep tuning*).

Jedan od izazova primjene metoda dubokog učenja na medicinske podatke jeste u tome da se većina metodologija razvija za obradu prirodnih slika te se postojeće arhitekture trebaju prilagoditi medicinskoj domeni.

Prijenosno učenje se može realizirati kao *fino* poboljšanje mreže trenirane na općenitim slikama ili kao *fino* poboljšanje mreže prethodno trenirane na medicinskim slikama za različit zadatak ili na različitim organima. Prijenosno učenje ostvaruje bolje rezultate kada su predtrenirane mreže slične ciljanim zadacima i čak u slučajevima potpuno različitih zadataka pokazuju bolje performanse od nasumične inicijalizacije težina [82], [83].

Međutim, odabir pristupa za realizaciju prijenosnog učenja ovisi i o izboru strukture mreže te veličine skupa podataka koji se koristi. Tako za „plitke“ mreže bolje performanse daje duboko poboljšanje dok u „dubljim“ arhitekturama djelomična prilagodba poboljšava performanse i smanjuje vrijeme konvergencije [83]. Za manje baze podataka *fino* poboljšanje je bolje od *dubokog* zato što za mali skup podataka vrlo lako može doći do pretreniranosti (engl. *overfittinga*) što znači da se mreža previše prilagodi podacima koje ima te nije sposobna generalizirati. Nadalje, jedan od ključnih elemenata za odabir prijenosnog učenja je i modalitet slike (npr. MRI, CT i slično) te organ koji se želi proučavati. Različiti modaliteti mogu imati dosta šuma i drugih smetnji te mogu biti velike varijacije u prikazu ciljanog organa.

U radu [46] autori koriste predtreniranu standardnu arhitekturu VGG16 i potom rade *fino* poboljšanje težina mreže kako bi je prilagodili za segmentaciju prenatalnih slika ultrazvuka. Slično u radu [83] autori koriste prijenosno učenje za detekciju polipa. Autori su radili duboko poboljšanje svih slojeva te uočili da time poboljšavaju osjetljivost mreže za 25% u odnosu na samo poboljšanje zadnjeg sloja. Također, postoje i istraživanja koja pokazuju kako treniranje cijele mreže od početka daje bolje rezultate od predtreniranih mreža korištenih za prijenosno učenje [84].

6.2 Metrike za procjenu rezultata segmentacije

Usporedba različitih tehnika segmentacije nije jednostavan zadatak. Postoji tek nekoliko javno dostupnih baza podataka koje imaju ručno označene slike i mogu služiti kao zlatni standard (engl. *ground truth data*) za usporedbu različitih metoda. Najčešće korištene javne baze za segmentaciju koljena su SKI10 [72] i OAI [73]. Međutim, većina autora koristi vlastite podatke za evaluaciju što otežava usporedbu istraživanja. Čak i ako koriste istu bazu podataka, rezultati se predstavljaju različitim metrikama jer ne postoji standardna mjera za evaluaciju rezultata. Prema literaturi većina rezultata se predstavlja kao mjera udaljenosti ili preklapanja volumena.

Najčešće korištena mjera procjene rezultata temeljena na preklapanju volumena je Dice koeficijent sličnosti (engl. *Dice similarity coefficient*, DSC) koji je zadan izrazom [85]:

$$DSC = \frac{2|V_A \cap V_B|}{|V_A| + |V_B|} \quad (6.1)$$

V_A predstavlja volumen koji je ručno segmentirao radiolog i služi kao zlatni standard, dok V_B predstavlja volumen koji je segmentacijski rezultat određenog algoritma.

Za segmentaciju kosti i hrskavice koljena na bazi SKI10 neki rezultati su predstavljeni mjerom razlike volumena (engl. *volume difference*, VD) [86]:

$$VD = 100 \cdot \frac{V_B - V_A}{V_A} \quad (6.2)$$

Također, spominje se mjera pod nazivom greška u preklapanju volumena (engl. *volume overlap error*, VOE) [16]:

$$VOE = 100 \cdot \left(1 - \frac{DSC}{200 - DSC}\right) \quad (6.3)$$

Nadalje, pojedini rezultati segmentacije se opisuju kao mjera specifičnosti algoritma (engl. *specificity*, SP) i mjera osjetljivosti (engl. *sensitivity*, SN) korištenjem izraza:

$$SP = \frac{TN}{TN+FP}, \quad SN = \frac{TP}{TP+FN} \quad (6.4)$$

gdje su FP lažno pozitivni, TP istinitno pozitivni, FN lažno negativni i TN istinitno negativni rezultati [16].

Točnost algoritma (engl. *accuracy*, A) je zadana izrazom:

$$A = \frac{TP+TN}{TP+TN+FP+FN} \quad (6.5)$$

7 IZAZOVI U PRIMJENI METODA DUBOKOG UČENJA NA MEDICINSKIM SLIKAMA

Primjena metoda dubokog učenja je izazovna sama po sebi zbog složenosti metoda i potrebnih računalnih resursa. Međutim, primjena takvih arhitektura za obradu medicinskih slika, koje su same po sebi specifične, može uzrokovati dodatne probleme i izazove.

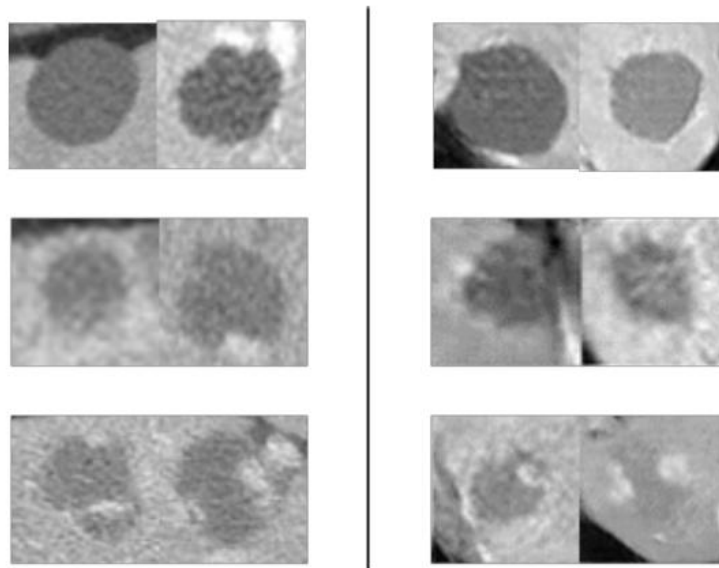
7.1 Izazovi u pripremi podataka i treniranju

Jedan od glavnih izazova je priprema podataka odnosno označavanje baze za treniranje neuronske mreže. Tehnike dubokog učenja su uvelike poboljšale točnost segmentacije zbog svojih sposobnosti koje se prvenstveno temelje na procesu učenja mreže. Mreže tipično zahtijevaju velike količine označenih uzoraka za treniranje što je izazovan zadatak za medicinske podatke. Čak i ako podatke označi stručnjak iz određene domene, može doći do problema u razvoju algoritama za učenje. Na primjer, u radu [87] na često korišenoj bazi podataka LIDC-IDRI za evaluaciju algoritama za detekciju čvorića u plućima na slikama CT-a su četiri radiologa neovisno radila označavanje odnosno anotaciju čvorića. Nakon označavanja mogli su pregledati međusoban rad, ali nisu se morali međusobno usuglasiti. Ispostavilo se da broj čvorića oko kojih se nisu složili jeste 3 puta veći od broja anotacija za koje se slažu. Treniranje modela dubokog učenja na takvim podacima zahtjeva pažljivo promatranje i odluku što učiniti za određene situacije zbog šuma i nesigurnosti u označavanju. U radu [88] predstavljena je strategija treniranja neuronske mreže koja uzima u obzir nepouzdanost označavanja te sa dodatnim slojem unutar arhitekture modeliraju pouzdanost određenih oznaka u procesu treniranja.

Smanjivanje vremena treniranja i ubrzanje konvergencije je ključna tema mnogih istraživanja. Jedno od najranijih rješenja je primjena slojeva sažimanja koji smanjuju dimenzionalnost parametara [37]. Također, korištenje normalizacije (engl. *Batch normalization*) gdje se vrijednosti piksela postavljaju oko nule dijeljenjem njihove početne vrijednosti sa ukupnom *mean* vrijednošću slike je poznat način ubrzanja konvergencije. Normalizacija se češće koristi budući da nema negativnih efekata na performanse za razliku od slojeva sažimanja gdje se gube korisne informacije [54].

Jedan od izazova u radu sa medicinskim podacima je upravljanje neuravnoteženosti klasa. Teško je pronaći dovoljan broj odgovarajućih klasa (ovisno koji je zadatak obrade u pitanju) jer većina slika budu normalne dok tek pojedine sadrže sumnjive tvorevine (područja od interesa). Na primjer u radovima koji istražuju obradu i detekciju tumora grudi većina mamograma je normalna. Tek pojedini imaju neprirodne tvorevine, a i ako sadrže tvorevinu ne znači da su sve maligne. Tako da je razvoj vještih algoritama dubokog učenja koji mogu voditi računa o prilagodbi odnosno ravnoteži klasa važno područje istraživanja. Tipične strategije uključuju transformaciju podataka popularno nazvanu augmentacija podataka, kako bi se dobili novi pozitivni primjeri. Transformacija podataka je standardna tehnika koja daje dobre učinke na treniranje dubokih modela [89]. Služi za izbjegavanje pretreniranosti, za ubrzanje konvergencije i generalizaciju modela. Za povećanje podataka uobičajeno se koriste operacije translacije, rotacije, zamatanja, skaliranja, rezanje, mijenjanja prostora boja i slično. Pereira [90] procjenjuje utjecaj različitih augmentacijskih strategija u borbi sa neuravnoteženosti klasa na segmentaciju moždanih lezija. Mnogi autori istražuju mogućnosti generiranja sintetičkih podataka korištenjem generativne protivničke mreže umjesto standardnih metoda augmentacije podataka [91]. Tijekom treniranja generator poboljšava svoju sposobnost stvaranja realističnih sintetičkih slika dok diskriminator poboljšava svoju sposobnost razlikovanja stvarnih od sintetičkih slika. Slika 7.1 prikazuje generirane podatke koji prikazuju lezije na jetri korištene

u radu [92]. Generiranje sintetičkih podataka omogućuje i neku vrstu anonimnosti zato što takve sintetičke slike onda nisu povezane za određenog pacijenta te se mogu lakše dijeliti sa znanstvenom zajednicom za daljnja istraživanja [93].



Slika 7.1 Prikaz GAN generiranih podataka. Lijevi stupac prikazuje stvarne slike. Desni stupac prikazuje sintetičke slike. Preuzeto iz [92].

Konačno, medicinske slike se spremaju u DICOM format koji osim slike sprema mnoštvo korisnih tekstualnih podataka o pacijentu u takozvanom DICOM zaglavlju. Pojam DICOM označava univerzalni i temeljni standard za prijenos, pohranu i ispis digitalnih slika u medicini [94]. Liječnici često koriste informacije o pacijentu (dob, spol, povijest liječenja i sl.) kako bi donijeli bolju konačnu odluku u procjeni. Korisnost tih podataka u samom procesu dizajniranja i treniranja algoritama dubokog učenja nije dovoljno istražena. Tek pojedini autori spominju kombiniranje tih informacija unutar neuronske mreže [95].

7.2 Izazovi temeljeni na različitim organima i modalitetima

Postoji tek nekoliko javno dostupnih baza medicinskih podataka koje istraživači mogu koristiti za procjenu svojih modela. Međutim, broj slučajeva i u tim bazama nije ni približan veličini baza prirodnih slika koje se popularno koriste te nisu jednako zastupljeni svi organi ni modaliteti slika. U tablici 7.1 su prikazane neke od najpopularnijih baza medicinskih slika sa veličinom baze, ciljnim organom i korištenim modalitetom.

Tablica 7.1 Prikaz popularnih baza podataka za segmentaciju određenih organa. Preuzeto i prilagođeno iz [3]

Organ/Dio tijela	Naziv baze podataka	Veličina	Dimenzija	Modalitet
Abdomen	NIH-CT-82	82 uzorka	3D	CT
	UFL-MRI-79	79 uzoraka	-	-
Mozak	Brain MRI C34	-	-	MRI
	MR Brains	-	-	MRI
	ADNI	339 uzoraka	3D	PET
Grudi	Breast MRI-34	-	-	T1-MRI
	INbreast	116 uzoraka	2D	Mamogramami
	DDSM-BCRP	158 uzoraka	-	-
Srce	ACDC	150 pacijenata	2D	MRI

	Cardic CTA	-	-	CT
Pluća	Lung Nodule Analysis 2016 (LUNA16)	880 pacijenata	2D	CT
	Kaggles Data Science Bowl (DSB)	1397 pacijenata	2D	CT
	LIDC-IDRI	1018 slučajeva	-	CT
Koža	ISBI 2016	1250 slika	2D	-

Odabir metode dubokog učenja kao i način treniranja mreže uvelike ovisi o modalitetu podataka i ciljanom organu. Heterogeni prikaz ciljanog organa je jedan od najvećih izazova prilikom segmentacije medicinskih slika. Moguće je da jedan organ ima potpuno drugačiji prikaz na magnetskoj rezonanciji, CT-u i ultrazvuku te može previše varirati u obliku, veličini i lokaciji od pacijenta do pacijenta. Također, nejasne granice između ciljanog organa i okolnog tkiva zbog lošeg kontrasta su jedan od većih izazova. To obično javlja zbog načina generiranja slika prilikom snimanja CT-a i MRI-a. Tablica 7.2 prikazuje kako različiti autori koriste različite arhitekture u ovisnosti o modalitetu i ciljanom organu.

Tablica 7.2 Prikaz različitih metoda dubokog učenja za segmentaciju različitih ulaznih podataka odnosno modaliteta. MRI-magnetna rezonanca, UZV-ultrazvuk, CT-Tomografija i DP- Digitalno patološke slike (mikroskop)

Referenca	Modalitet	Metoda
Zhang i ostali [39]	MRI	Multi-modalna 2D CNN za segmentaciju tkiva odnosno mozga novorođenčadi
Birenbaum i Greenspan [96]	MRI	2.5D CNN za segmentaciju lezija multiple skleroze
Chen i ostali [55]	MRI	3D ResNet za segmentaciju mozga
Kamnitsas i ostali [36]	MRI	3D CNN za segmentaciju tkiva
Milletari i ostali [97]	MRI, UZV	2D, 2.5D i 3D CNN za segmentaciju regija mozga
Charbonnier i ostali [23]	CT	CNN za segmentaciju dišnih puteva
BenTaieb i Hamarneh [98]	DP	FCN za segmentaciju crijevnih žlijezda
Chen i ostali [99]	DP	Kombinacija LSTM-RNN i U-net za segmentaciju neuronskih membrana i gljivica
Christ i ostali [44]	CT	Kaskadne FCN i 3D uvjetna slučajna polja (CRF)

Svi prethodno navedeni izazovi mogu biti znatno ozbiljniji ako se radi sa volumetrijskim podacima odnosno 3D podacima zbog male razlike ciljanih i susjednih *voxela*, velikog broja parametara, ograničenih podataka za treniranje i potrebnih računalnih resursa.

Konačno, metode dubokog učenja se često opisuju kao „crne kutije“ što u medicini može imati ozbiljne posljedice ukoliko sustav ne predviđa dovoljno odgovorno i precizno. Više strategija je razvijeno s ciljem da se bolje razumije rad i rezultat srednjih slojeva konvolucijske mreže kao što su: vođena back-propagacija [100], dekonvolucijske mreže [101], duboka Taylor kompozicija [102]. Nadalje, neki istraživači su pokušali povezati Bayesovu statistiku sa dubokim mrežama kako bi istinsku procjenu nesigurnosti mreže te time dopustili liječnicima da kontroliraju nepouzdanu rezultate odnosno predviđanja [103].

8 ZAKLJUČAK

U ovom radu dat je pregled popularnih metoda dubokog učenja u svrhu obrade i segmentacije medicinskih slika te su naglašene njihove prednosti. Primjena dubokog učenja u medicini prvenstveno ima ulogu ubrzanja pregleda velikih količina generiranih slika koje liječnici ručno moraju pregledati i kao pomoć u odlučivanju prisutnosti nekih abnormalnosti. Za snimanje medicinskih slika mogu se koristiti različiti uređaji: CT, MRI, ultrazvuk, mikroskop i slično što rezultira različitim modalitetima slika. Slike se potom pohranjuju u DICOM formatu koji ima svoje specifičnosti i mnoštvo korisnih informacija o samoj slici kao i uređaju koji je generirao sliku. Zbog specifičnosti formata, modaliteta i dimenzionalnosti slika (2D ili 3D) je obrada medicinskih slika složenija od obrade prirodnih slika.

Na početku rada su objašnjeni osnovni pojmovi vezani za duboko učenje s posebnim osvrtom na konvolucijske neuronske mreže i potpuno konvolucijske mreže zbog njihove popularnosti. Iako se mreža sastoji od jednostavnih slojeva, postoji neograničen broj kombinacija tih slojeva i odabira željene dubine mreže što utječe na konačan rezultat. Semantička segmentacija nije do kraja istražena bez obzira na količinu radova u području dubokog učenja jer ne postoji arhitektura koja može nadmašiti ostale i koja se može primijeniti za sve modalitete, organe i dimenzije slika. Istraživači intenzivno rade kako bi poboljšali preciznost semantičke segmentacije i riješili probleme gubitka informacija u slojevima sažimanja.

Kroz ovaj rad su također objašnjene glavne tehnike treniranja te najčešće metrike kojima su predstavljeni rezultati segmentacije. Neuronske mreže se specijaliziraju za određen zadatak i svaka promjena značajki ili parametara zahtjeva ponovno treniranje mreže. Stoga se znanje nastoji prenijeti između različitih modela kako bi se izbjeglo treniranje cijelog modela koje zahtjeva veliku količinu podataka i vremena. Zato je najveći naglasak na prijenosnom učenju budući da je jedan od najvećih izazova imati dovoljan broj označenih medicinskih slika te ubrzanje konvergencije.

Nadalje, objašnjene se korištene metode segmentacije na magnetskoj rezonanciji koljena. Segmentacija slika koljena u posljednje vrijeme postaje posebno popularna u svrhu procjene starosti pacijenta. Procjena kronološke dobi osobe je ključna tema forenzičkih znanosti posebno kod nedostatka valjanih legalnih dokumenata, kaznenih postupaka i zahtjeva za azilom zbog trenutne europske migrantske krize [104]. Zbog nepostojanja javno dostupne baze podataka s osobama mlađe životne dobi i sva navedena istraživanja imaju određene nedostatke, očekuje se u budućnosti više istraživanja u tom smjeru.

U zadnjem poglavlju se naglašeni najčešći izazovi i problemi pri korištenju metoda dubokog učenja na medicinskim podacima. Kod odabira određene arhitekture treba uzeti u obzir količinu dostupnih podataka za treniranje, modalitet slika te ciljani organ kao i tehnike predprocesiranja podataka i uravnoteživanja klasa.

Unatoč velikom interesu za istraživanje ovog područja, svi izazovi objašnjeni u prethodnom poglavlju nisu prevaziđeni. Duboko učenje će sigurno u budućnosti imati veliku ulogu u pomaganju zdravstvenim djelatnicima ubrzati proces rada i analizirati stanje pacijenata kao i biti pomoć u odlučivanju o daljnjem liječenju.

LITERATURA

- [1] F. Rosenblatt, "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain," *Cornell Aeronautical laboratory, Psychological Review*, vol. 65, no. 6, pp. 386-408, 1958..
- [2] Y. LeCun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. of the IEEE*, vol. 86, pp. 2278-2324, 1998.
- [3] R. Anirudh , J. Thiagarajan, T. Bremer and H. Kim, "Lung nodule detection using 3D convolutional neural networks trained on weakly labeled data," *Medical Imaging: Computer-Aided Diagnosis*, vol. 9785, 2016.
- [4] H. Roth, L. Lu, J. Liu, J. Yao, A. Seff, K. Cherry, L. Kim and R. Summers, "Improving computer-aided detection using convolutional neural networks and random view aggregation," *IEEE Transactions on Medical Imaging* , vol. 35, no. 5, pp. 1170-1181, 2016.
- [5] L. Yu, H. Chen, Q. Dou, J. Qin and P. Heng, "Automated melanoma recognition in dermoscopy images via very deep residual networks," *IEEE Transactions on Medical Imaging*, vol. 36, no. 4, pp. 994-1004, 2017.
- [6] "Advanced Topics in Deep Convolutional Neural Networks," [Online]. Available: <https://towardsdatascience.com/advanced-topics-in-deep-convolutional-neural-networks-71ef1190522d>. [Accessed 31 03 2020].
- [7] M. H. Hesamian , W. Jia, X. He and P. Kennedy , "Deep Learning Techniques for Medical Image Segmentation: Achievements and Challenges," *Journal of Digital Imaging* , vol. 32, pp. 582-596, 2019.
- [8] "Dive into deep learning," [Online]. Available: http://www.d2l.ai/chapter_convolutional-neural-networks/lenet.html#lenet. [Accessed 09 06 2020].
- [9] A. Krizhevsky, I. Sutskever and G. Hinton, "Imagenet classification with deep convolutional networks," *Proceedings of the Advances in Neural Information Processing Systems*, pp. 1097-1105, 2012.
- [10] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *ICLR*, 2014.
- [11] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke and A. Rabinovich, "Going deeper with convolutions," in *Conference on Computer Vision and Pattern Recognition*, 2015.
- [12] K. He, X. Zhang, S. Ren and J. Sun, "Deep residual learning for image recognition," in *arXiv:1512.03385*, 2015.

- [13] S.-C. Lo, S.-L. Lou, J.-S. Lin, M. Freedman, M. Chien and S. Mun, "Artificial convolution neural network techniques and applications for lung nodule detection," *IEEE Trans.Med. Imaging*, no. 14, pp. 711-718, 1995.
- [14] "Biomedicinska instrumentacija," [Online]. Available: [https://www.fer.unizg.hr/_download/repository/P03_2016_Medical_Imaging_\(1\).pdf](https://www.fer.unizg.hr/_download/repository/P03_2016_Medical_Imaging_(1).pdf). [Accessed 09 06 2020].
- [15] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghahfoorian, . A. J. van der Laak, B. van Ginneken and C. I. Sanchez, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60-88, 2017.
- [16] S. Ebrahimkhani, M. Jaward, F. Cicutini, A. Dharmarante, Y. Wang and A. Seco de Herrera, "A review on segmentation of knee articular cartilage: from conventional methods towards deep learning," *Artificial Intelligence in Medicine*, vol. 106, 2020.
- [17] A. Mamourian , Practical MR physics: And case file of MR artifacts and pitfalls, Oxford University Press Inc. ISBN: 978-0-19-537281-6, 2010.
- [18] A. Menegola, M. Fornaciali, R. Pires, S. Avila and E. Valle, "Towards automated melanoma screening: exploring transfer learning schemes," in *Computer Vision and Pattern Recognition*, *arXiv:1609.01228*, 2016.
- [19] A. Esteva, B. Kuprel, R. Novoa, J. Ko, S. Swetter, H. Blau and S. Thrun, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, *doi:10.1038/nature21056.*, vol. 542, pp. 115-118, 2017.
- [20] S.-C. Lo, S.-L. Lou, J.-S. Lin, M. Freedman , M. Chien and S. Mun, "Artificial convolution neural network techniques and applications for lung nodule detection," *IEEE Transactions on Medical Imaging* , vol. 14, pp. 711-718, 1995.
- [21] T. Kooi, G. Litjens, B. van Ginneken, A. Gubern-Merida, C. Sanchez, R. Mann, A. den Heeten and N. Karssemeijer, "Large scale deep learning for computer aided detection of mammographic lesions," *Medical Image Analysis*, vol. 35, pp. 303-312, 2016.
- [22] M. Ghahfoorian, N. Karssemeijer, T. Heskes, I. van Uden, F.-E. de Leeuw, E. Marchiori, B. van Ginneken and B. Platel, "Non-uniform patch sampling with deep convolutional neural networks for white matter hyperintensity segmentation," *Proceedings of the IEEE International Symposium on Biomedical Imaging* , pp. 1414-1417, 2016.
- [23] J. Charbonnier, E. van Rikxoort, A. Setio, C. Schaefer-Prokop, B. van Ginneken and F. Ciompi, "Improving airway segmentation in computed tomography using leak detection with convolutional networks," *Medical Image Analysis*, vol. 36, pp. 52-60, 2017.
- [24] M. van Grinsven, B. van Ginneken, C. Hoyng, T. Theelen and C. Sanchez, "Fast convolutional neural networks training using selective data sampling: application to hemorrhage detection in color fundus images," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1273-1284, 2016.
- [25] "ISLES," [Online]. Available: <http://www.isles-challenge.org/>. [Accessed 30 03 2020].

- [26] "BRATS Challenge," [Online]. Available: <http://braintumorsegmentation.org/>. [Accessed 30 03 2020].
- [27] "MRBrainS," [Online]. Available: <https://mrbrains13.isi.uu.nl/>. [Accessed 30 03 2020].
- [28] "LV Segmentation Challenge," [Online]. Available: <http://www.cardiacatlas.org/challenges/lv-segmentation-challenge/>. [Accessed 30 03 2020].
- [29] A. Prason, K. Petersen, C. Igel, F. Lauze, E. Dam and M. Nielsen, "Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network," *Proceedings of the Medical Image Computing and Computer-Assisted Intervention, Springer*, pp. 246-253, 2013.
- [30] X. Zhou, R. Takayama, S. Wang, T. Hara and H. Fujita, "Deep learning of the sectional appearances of 3D CT images for anatomical structure segmentation based on an FCN voting model," *Medical Physics*, vol. 44, no. 10, pp. 5221-5233, 2017.
- [31] H. Roth, L. Lu, A. Seff, K. Cherry, J. Hoffman, S. Wang, J. Liu, E. Turkbey and R. Summers, "A new 2.5D representation for lymph node detection using random sets of deep convolutional neural network observations," in *International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer*, 2014.
- [32] P. Moeskops, J. Wolterink, B. van der Velden, K. Gilhuijs, T. Leiner, M. Viergever and I. Išgum, "Deep learning for multi-task medical image segmentation in multiple modalities," *International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer*, pp. 478-486, 2016.
- [33] K. Kamnitsas, C. Ledig, V. Newcombe, J. Simpson, A. Kane, D. Menon, D. Rueckert and B. Glocker, "Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation," *Medical Image Analysis*, vol. 36, pp. 61-78, 2017.
- [34] [Online]. Available: <https://towardsdatascience.com/understanding-1d-and-3d-convolution-neural-network-keras-9d8f76e29610>. [Accessed 04 04 2020].
- [35] G. Urban, M. Bendszus, F. Hamprecht and J. Kleesiek, "Multi-modal brain tumor segmentation using deep convolutional neural networks," in *Proceedings MICCAI braTS (Brain Tumor Segmentation)*, 2014.
- [36] K. Kamnitsas, L. Chen, C. Ledig, D. Rueckert and B. Glocker, "Multi-scale 3D convolutional neural networks for lesion segmentation in brain MRI," in *Ischemic Stroke Lesion Segmentation 13:46*, 2015.
- [37] Q. Dou, L. Yu, H. Chen, Y. Jin, X. Yang, J. Qin and P. Heng, "3D deeply supervised network for automated segmentation of volumetric medical images," *Medical Image Analysis*, vol. 41, pp. 40-54, 2017.
- [38] J. Long, E. Shelhamer and T. Darrell, "Fully convolutional networks for semantic segmentation," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2015.

- [39] W. Zhang , R. Li, H. Deng, L. Wang, W. Lin, S. Ji and D. Shen, "Deep convolutional neural networks for multi-modality isointense infant brain image segmentation," *NeuroImage*, vol. 108, pp. 214-224, 2015.
- [40] D. Nie, L. Wang, Y. Gao and D. Sken, "Fully convolutional networks for multi-modality isointense infant brain image segmentation," *IEEE 13th International Symposium on Biomedical Imaging* , pp. 1342-1345, 2016.
- [41] E. Gibson, F. Giganti, Y. Hu, E. Bonmati, S. Bandula, K. Gurusamy, B. Davidson, S. Pereira, M. Clarkson and D. Barratt, "Towards image-guided pancreas and biliary endoscopy: automatic multi-organ segmentation on abdominal CT with dense dilated networks," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2017.
- [42] P. Hu, F. Wu, J. Peng, Y. Bao, F. Chen and D. Kong, "Automatic abdominal multi-organ segmentation using deep convolutional neural networks and time-implicit level sets," *International Journal of Computer Assisted Radiology and Surgery*, vol. 12, no. 3, pp. 399-411, 2017.
- [43] X. Zhou, "Automatic Segmentation of Multiple Organs on 3D CT Images by Using Deep Learning Approaches," *Advances in Experimental Medicine and Biology*, vol. 1213, pp. 135-147, 2020.
- [44] P. Christ , M. Elshaer, F. Ettliger, S. Tatavarty, M. Bickel, P. Bilic, M. Rempfler, M. Armbruster, F. Hofmann and M. D'Anastasi, "Automatic liver and lesion segmentation in CT using cascaded fully convolutional neural networks and 3D conditional random fields," *Proceedings of the Medical Image Computing and Computer-Assisted Intervention, In: Lecture Notes in Computer Science*, vol. 9901, pp. 415-423, 2016.
- [45] P. Christ , F. Ettliger, F. Grun, M. Elshaera, J. Lipkova, S. Schlecht, F. Ahmaddy, S. Tatavarty, M. Bickel and P. Bilic, "Automatic liver and tumor segmentation of CT and MRI volumes using cascades fully convolutional neural networks," *arXiv: 1702.05970*, 2017.
- [46] L. Wu, Y. Xin, S. Li, T. Wang, P. Heng and D. Ni, "Cascaded fully convolutional networks for automatic prenatal ultrasound image segmentation," *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI)*, pp. 663-666, 2017.
- [47] G. Zeng and G. Zheng, "Multi-stream 3D FCN with multi-scale deep supervision for multi-modality isointense infant brain MR image segmentation," *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018) IEEE*, pp. 136-140, 2018.
- [48] J. Wang , J. MacKenzie, R. Ramachandran and D. Chen, "A deep learning approach for semantic segmentation in histology tissue images," *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, pp. 176-184, 2016.
- [49] S. Hamidian , B. Sahiner, N. Petrick and A. Pezeshk, "3D convolutional neural network for automatic detection of lung nodules in chest CT," *Medical Imaging 2017: Computer-Aided Diagnosis* , vol. 10134, 2017.

- [50] O. Ronneberger, P. Fischer and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2015.
- [51] F. Milletari , N. Navab and S. Ahmadi, "V-net: fully convolutional neural networks for volumetric medical image segmentation," in *2016 Fourth International Conference on 3D Vision (3DV), IEEE*, 2016.
- [52] E. Gibson, F. Giganti, Y. Hu, E. Bonmati, S. Bandula, K. Gurusamy, B. Davidson, S. Pereira, M. Clarkson and D. Barratt, "Automatic multi-organ segmentation on abdominal CT with dense v-networks," *IEEE Transaction on Medical imaging*, vol. 37, no. 8, pp. 1822-1834, 2018.
- [53] E. Gibson, M. Robu, S. Thompson, P. Edwards, C. Schneider, K. Gurusamy, B. Davidson, D. Hawkes, D. Barratt and M. Clarkson, "Deep residual networks for automatic segmentation of laparoscopic videos of the liver," *Medical Imaging 2017: Image-Guided Procedures, Robotic Interventions, and Modeling* , vol. 10135, p. 101351m, 2017.
- [54] J. Kawahara, A. BenTaieb and G. Hamarneh, "Deep features to classify skin lesions," *IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, pp. 1397-1400, 2016.
- [55] H. Chen, Q. Dou, L. Yu, J. Qin and P. Heng, "Voxresnet: deep voxelwise residual networks for brain segmentation from 3D MR images," *NeuroImage*, vol. 170, pp. 446-455, 2017.
- [56] N. Ibtehaz and M. Sohel Rahman, "MultiResUNet: Rethinking the U-Net architecture for multimodal biomedical image segmentation," *Neural Networks*, vol. 12, pp. 74-87, 2020.
- [57] "Recurrent Neural Networks," [Online]. Available: <https://towardsdatascience.com/recurrent-neural-networks-d4642c9bc7ce>. [Accessed 31 05 2020].
- [58] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput*, vol. 9, no. 8, pp. 1735-1780, 1997.
- [59] J. Cai, L. Lu, Y. Xie, F. Xing and L. Yang, "Improving deep pancreas segmentation in CT and MRI images via recurrent neural contextual learning and direct loss function," *arXiv:1707.04912*, 2017.
- [60] S. Xingjian, Z. Chen, H. Wang, D. Yeung, W. Wong and W. Woo, "Convolutional LSTM network: a machine learning approach for precipitation nowcasting," *Advances in Neural Information Processing Systems*, pp. 802-810, 2015.
- [61] K.-L. Tseng, Y.-L. Lin, W. Hsu and C.-Y. Huang, "Joint Sequence Learning and Cross-Modality Convolution for 3D Biomedical Segmentation," *arXiv:1704.07754*, 2017.

- [62] "BRATS-2015," [Online]. Available: http://people.csail.mit.edu/menze/papers/proceedings_miccai_brats_2015.pdf. [Accessed 18 06 2020].
- [63] D. Cheng and M. Liu, "Combining convolutional and recurrent neural networks for Alzheimer's disease diagnosis using pet images," *IEEE International Conference on Imaging Systems And Techniques (IST)*, pp. 1-5, 2017.
- [64] R. Poudel , P. Lamata and G. Montana , "Recurrent fully convolutional neural networks for multi-slice MRI cardiac segmentation," *Reconstruction, Segmentation, and Analysis of Medical Images Springer*, pp. 83-94, 2016.
- [65] J. Koutnik, K. Greff, F. Gomez and J. Schmidhuber, "A clockwork rnn," *arXiv:1402.3511*, 2014.
- [66] D. Ciresan, A. Giusti, L. Gambardella and J. Schmidhuber, "Deep neural networks segment neuronal membranes in electron microscopy images," *Advances in Neural Information Processing Systems* , pp. 2843-2851, 2012.
- [67] Y. Xie, Z. Zhang, M. Sapkota and L. Yang, "Spatial Clockwork Recurrent Neural Network for Muscle Perimysium Segmentation," *Medical Image Computer Assisted Intervention*, vol. 9901, pp. 185-193, 2016.
- [68] [Online]. Available: <https://www.osgpc.com/anatomy-of-the-knee/>. [Accessed 10 06 2020].
- [69] M. Rahman, L. Durselen and A. Seitz, "Automatic segmentation of knee menisci-A systematic review," *Artificial Intelligence In Medicine*, vol. 105, 2020.
- [70] Z. Zhou, G. Zhao, R. Kijowski and F. Liu, "Deep convolutional neural network for segmentation of knee joint anatomy," *Magnetic Resonance in Imaging*, vol. 80, no. 6, pp. 2759-2770, 2018.
- [71] B. Norman, V. Padoia and S. Majumdar, "Use of 2d u-net convolutional neural networks for automated cartilage and meniscus segmentation of knee mr imaging data to determine relaxometry and morphometry," *Radiology* , vol. 288, no. 1, pp. 177-185, 2018.
- [72] "Segmentation of Knee Images 2010," [Online]. Available: <http://www.ski10.org/>. [Accessed 16 06 2020].
- [73] [Online]. Available: <https://nda.nih.gov/oai/>. [Accessed 16 06 2020].
- [74] M. Byra, M. Wu, X. Zhang, H. Jang, Y.-J. Ma, E. Chang and et.al., "Knee menisci segmentation and relaxometry of 3d ultrashort echo time cones MR imaging using attention U-Net with transfer learning," *Magnetic Resonance in Medicine*, vol. 83, no. 3, pp. 1109-22, 2020.

- [75] S. Gaj, M. Yang, K. Nakamura and X. Li, "Automated cartilage and meniscus segmentation of knee MRI with conditional generative adversarial networks," *Magnetic Resonance in Medicine*, 2019.
- [76] F. Liu, Z. Zhou, H. Jang, A. Samsonov, G. Zhao and R. Kijowski, "Deep convolutional neural network and 3d deformable approach for tissue segmentation in musculoskeletal magnetic resonance imaging," *Magnetic Resonance in Medicine*, vol. 79, no. 4, pp. 2379-91, 2018.
- [77] P.-L. Prove, E. Jop-van Well, B. Stanczus, M. Morlock, J. Herrmann, M. Groth, D. Saring and M. Auf der Mauer, "Automated segmentation of the knee for age assessment in 3D MR images using convolutional neural networks," *International Journal of Legal Medicine*, vol. 133, no. 4, pp. 1191-1205, 2019.
- [78] A. L. Dallora, J. Berglund, M. Brogren, O. Kvist, S. Ruiz, A. Dubbel and P. Anderberg, "Age Assessment of Youth and Young Adults Using Magnetic Resonance Imaging of the Knee: A Deep Learning Approach," *JMIR Medical Informatics*, vol. 7, no. 4, 2019.
- [79] C. Lee, S. Xie, P. Gallagher, Z. Zhang and Z. Tu, "Deeply-supervised nets," *Artificial Intelligence and Statistics*, pp. 562-570, 2015.
- [80] H. Chen, X. Qi, J. Cheng, P. Heng and et al., "Deep contextual networks for neuronal structure segmentation," *AAAI*, pp. 1167-1173, 2016.
- [81] C. Shie, C. Chuang, C. Chou, M. Wu and E. Chang, "Transfer representation learning for medical image analysis," in *37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2015.
- [82] J. Yosinski, J. Clune, Y. Bengio and H. Lipson, "How transferable are features in deep neural networks?," *Advances in Neural Information Processing Systems*, pp. 3320-3328, 2014.
- [83] N. Tajbakhsh, J. Shin, S. Gurudu, R. Hurst, C. B. Kendall, M. B. Gotway and J. Liang, "Convolutional neural networks for medical image analysis: Full training or fine tuning?," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, 2016.
- [84] D. Tran, L. Bourdev, R. Fergus, L. Torresani and M. Paluri, "Deep end2end voxel2voxel prediction," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 17-24, 2016.
- [85] L. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297-302, 1945.
- [86] G. Vincent, C. Wolstenholme, I. Scott and M. Bowes, "Fully automatic segmentation of the knee joint using active appearance models," *Medical Image Analysis for the Clinic A Challenge 2010*, vol. 1:224, 2010.
- [87] S. Armato, G. McLennan, L. Bidaut and et al., "The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans," *Medical Physics*, vol. 38, pp. 915-931, 2011.

- [88] Y. Dgani, H. Greenspan and J. Goldberger, "Training a neural network based on unreliable human annotation of medical images," in *15th International Symposium on Biomedical Imaging (ISBI 2018)*, Washington, USA, 2018.
- [89] S. Wong, A. Gatt, V. Stamatescu and M. McDonnell, "Understanding data augmentation for classification: when to warp?," *CoRR*, vol. abs/1609.08764, 2016.
- [90] S. Pereira, A. Pinto, V. Alves and C. Silva, "Brain tumor segmentation using convolutional neural networks in MRI images," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1240-1251, 2016.
- [91] V. Sandford, K. Yan, P. Pickhardt and R. Summers, "Data augmentation using generative adversarial networks (CycleGAN) to improve generalizability in CT segmentation tasks," *Springer Nature* 9:16884, 2019.
- [92] M. Frid-Adar, E. Klang, M. Amitai, J. Goldberger and H. Greenspan, "Synthetic data augmentation using GAN for improved liver lesion classification," *Computer Vision and Pattern Recognition*, *arXiv:1801.02385*, 2018.
- [93] H. Shin, N. Tenenholtz, J. Rogers, C. Schwarz, M. Senjem, J. Gunter, K. Andriole and M. Michalski, "Medical Image Synthesis for Data Augmentation and Anonymization using Generative Adversarial Networks," *Computer Vision and Pattern Recognition*, *arXiv:1807.10225*, 2018.
- [94] P. Mildenerger, M. Eichelberg and E. Martin, "Introduction to the DICOM standard," *European Radiology*, vol. 12, no. 4, pp. 920-927, 2002.
- [95] T. Kooi, B. van Ginneken, N. Karssemeijer and N. den Heeten, "Discriminating solitary cysts from soft tissue lesions in mammography using a pretrained deep convolutional neural networks," *Medical Physics*, vol. 44, no. 3, pp. 1017-1027, 2017.
- [96] A. Birenbaum and H. Greenspan, "Longitudinal multiple sclerosis lesion segmentation using multi-view convolutional neural networks," *Proceedings of the Deep Learning in Medical Image Analysis (DLMIA)*, In: *Lecture Notes in Computer Science*, vol. 10008, pp. 58-67, 2016.
- [97] F. Milletari, S.-A. Ahmadi, C. Kroll, A. Plate, V. Rozanski, J. Maiostre, J. Levin, O. Dietrich, B. Ertl-Wagner, K. Botzel and N. Navab, "Hough-CNN: deep learning for segmentation of deep brain regions in MRI and ultrasound," *arxiv: 1601.07014*.
- [98] A. BenTaieb and G. Hamarneh, "Topology aware fully convolutional networks for histology gland segmentation," *Proceedings of the Medical Image Computing and Computer-Assisted Intervention*. In: *Lecture Notes in Computer Science*, vol. 9901, pp. 460-468, 2016.
- [99] J. Chen, L. Yang, Y. Zhang, M. Alber and D. Chen, "Combining fully convolutional and recurrent neural networks for 3D biomedical image segmentation," *Proceedings of the Advances in Neural Information Processing Systems*, pp. 3036-3044, 2016.

- [100] J. Springenberg, A. Dosovitskiy, T. Brox and M. Riedmiller, "Striving for simplicity: the all convolutional net," *arxiv:1412.6806*, 2014.
- [101] M. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *Proceedings of the European Conference on Computer Vision*, 2014.
- [102] G. Montavon, S. Lapuschkin, A. Binder, W. Samek and K.-R. Muller, "Explaining nonlinear classification decisions with deep taylor decomposition," *Pattern Recognit.*, pp. 211-222, 2017.
- [103] A. Kendall and Y. Gal, "What uncertainties do we need in Bayesian deep learning for computer vision?," *arXiv:1703.04977*, 2017.
- [104] "European Asylum Support Office Age Assessment practice in Europe," 2013. [Online].