

Primjena umjetne inteligencije u automatskoj dijagnostici rendgenskih snimaka prsnog koša

Kolić, Jan

Master's thesis / Diplomski rad

2023

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, School of Medicine / Sveučilište u Zagrebu, Medicinski fakultet**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:105:383316>

Rights / Prava: [In copyright](#)/[Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2024-07-29**



Repository / Repozitorij:

[Dr Med - University of Zagreb School of Medicine Digital Repository](#)



SVEUČILIŠTE U ZAGREBU
MEDICINSKI FAKULTET

Jan Kolić

**Primjena umjetne inteligencije u automatskoj
dijagnostici rendgenskih snimaka prsnog koša**

Diplomski rad



Zagreb, 2023.

Ovaj diplomski rad izrađen je na Katedri za medicinsku statistiku, epidemiologiju i medicinsku informatiku pod vodstvom doc. dr. sc. Andrije Štajduhara i predan je na ocjenu u akademskoj godini 2022./2023.

Popis kratica

AP (anteriorno-posteriorna) – vrsta projekcije pri snimanju rendgena prsnog koša

AUC („area under the ROC curve“) – površina ispod ROC krivulje, koristi se za procjenu učinka modela strojnog učenja

DAM („Deep AUC Maximization“) – matematička metoda za podešavanje modela kako bi producirao što veću AUC vrijednost

DICOM („Digital Imaging and Communications in Medicine“) – internacionalni standard za medicinske slike i informacije vezane za njih

FN („false negative“) – lažno negativan, rezultat testiranja koji predviđa nešto kao negativno, a da to zapravo nije

FP („false positive“) – lažno pozitivan, rezultat testiranja koji predviđa nešto kao pozitivno, a da to zapravo nije

FPR („false positive rate“) – stopa lažno pozitivnih, omjer lažno pozitivnih rezultata i zbroja lažno pozitivnih i stvarno negativnih

GPU („graphics processing unit“) – vrsta procesora, specijaliziran za paralelnu obradu većeg broja informacija

NIH („National Institutes of Health“) – američki institut za medicinska istraživanja

PA (posteriorno-anteriorna) – vrsta projekcije pri snimanju rendgena prsnog koša

ROC („receiver operating characteristic curve“) – graf koji pokazuje učinak modela za klasifikaciju, uzima u obzir sve moguće vrijednosti klasifikacijskih pragova

TN („true negative“) – stvarno negativan, rezultat testiranja koji predviđa nešto kao negativno, a da to i je tako

TP („true positive“) – stvarno pozitivan, rezultat testiranja koji predviđa nešto kao pozitivno, a da to i je tako

TPR („true positive rate“) – stopa stvarno pozitivnih, omjer stvarno pozitivnih i zbroja stvarno pozitivnih i lažno negativnih

Sadržaj

1	Sažetak	
2	Summary	
3	Uvod.....	1
4	Rendgenske snimke prsnog koša	3
4.1	Tehnike snimanja	3
4.1.1	Posteriorno-anteriorno (PA).....	3
4.1.2	Anteriorno-posteriorno (AP).....	4
4.1.3	Lateralno.....	4
4.2	Analiza snimke	4
4.2.1	A – airway / dišni put (i medijastinum)	4
4.2.2	B – breathing / disanje	5
4.2.3	C – cardiac / srce.....	5
4.2.4	D – disability / invalidnost	5
4.2.5	E – everything else / ostalo.....	5
4.3	Budućnost analize	5
5	Umjetna inteligencija	7
5.1	Općenito.....	7
5.2	Strojno učenje	8
5.3	Duboko učenje	10
5.4	Računalni vid.....	11
5.5	Predtrenirani modeli	14
5.6	Sposobnost generalizacije modela.....	14
5.7	Mjere za procjenu efektivnosti modela	16
6	Baze podataka rendgenskih snimaka prsnog koša	18
7	Predobrada rendgenskih snimaka.....	21
7.1	Poboljšanje kvalitete snimaka	21
7.2	Segmentacija	21
7.3	Supresija kostiju	22
8	Modeli za prepoznavanje stanja.....	23
8.1	Pneumonija	23
8.2	Različita stanja	27

9	Primjena odabranog modela	31
9.1	Općenito.....	31
9.2	Odabrani model.....	31
9.3	Primjeri uporabe modela	31
9.3.1	Pneumonija.....	32
9.3.2	Kardiomegalija.....	33
9.3.3	Tumor pluća.....	34
9.4	Usporedba rezultata prethodnih primjera	34
9.5	Značenje predviđenih vjerojatnosti.....	35
9.6	Procjena efektivnosti modela	36
9.7	Pozadina predviđenih vjerojatnosti.....	37
10	Rasprava	40
11	Zaključak.....	42
12	Zahvale.....	43
13	Literatura	44
14	Životopis	51

1 Sažetak

Primjena umjetne inteligencije u automatskoj dijagnostici rendgenskih snimaka prsnog koša

Autor: Jan Kolić

Umjetna inteligencija se sve više koristi za rješavanje različitih problema. Značajan udio predstavljaju i problemi analize slika za koji se uspješno koriste konvolucijske neuronske mreže po principu dubokog učenja. Takvi problemi postoje i u većini grana medicine, a posebno u radiologiji, gdje analiza snimaka predstavlja velik posao na koji se troši puno vremena i resursa. U ovom radu, fokus je na detekciji patoloških stanja iz rendgenskih snimaka prsnog koša. Opisani su načini i posebnosti snimanja i radiološke analize rendgenskih snimaka prsnog koša. Objasnjeni su temeljni pojmovi i dio teoretske podloge umjetne inteligencije, s posebnim naglaskom na modele dubokog učenja i računalnog vida kao podvrstu strojnog učenja. Treniranje takvih modela zahtijeva velik broj slika, zbog čega danas postoje brojne javno dostupne baze rendgenskih snimaka, a najvažnije su nabrojane u radu. Uz to, navedene su i neke od metoda za predobradu snimaka. Brojni radovi su objavljeni na temu automatske obrade rendgenskih snimaka u svrhu testiranja i poboljšavanja modela. Jednim od najuspješnijih pristupa se pokazala uporaba predtreniranih modela, treniranih na ImageNet bazi slika. Posebno su uspješni bili kombinirani modeli čije konačno predviđanje je konsenzus predviđanja podmodela. Dodavanje neslikovnih podataka nije doprinijelo učinku modela, za razliku od metoda segmentacije snimke te Deep AUC Maximization metode, čime su postignute bolje AUC vrijednosti. Kako bi se pobliže prikazao način njihovog korištenja, kao primjer je odabran jedan od modela te je upotrijebljen na nekoliko snimaka, na kojima je uspješno prepoznao dijelove gdje su pojedina patološka stanja prisutna. Ispravno je predviđena vjerojatnost postojanja određenih stanja bila veća na snimkama gdje su ta stanja stvarno bila prisutna. Prikazani su svi dobiveni rezultati, objašnjeno njihovo značenje te je raspravljeno o podlozi i problemima načina rada takvih modela. Identificiran je potencijalni problem korištenja nemedicinskih podataka sa snimke za prepoznavanje stanja.

Ključne riječi: umjetna inteligencija, klasifikacija slika, automatska dijagnostika, rendgen prsnog koša

2 Summary

The application of artificial intelligence in automated diagnosis from chest X-rays

Author: Jan Kolić

Artificial intelligence is being increasingly used to solve a variety of problems. One of them is image analysis, which is successfully being solved using convolutional neural networks, utilizing the principle of deep learning. Such problems also exist in most medical specialties, especially in radiology, where image analysis takes a lot of time and resources. The focus of this thesis is on the detection of pathological conditions from chest X-rays. Different methods of recording and radiological analysis of chest X-rays are described. Some basic concepts and the theoretical background of artificial intelligence is explained, with special emphasis on deep learning models and computer vision as a type of machine learning. Training such models requires a large number of images, which is why a lot of chest X-ray databases exist today and are publicly available. The most important ones are listed in the paper, along with the most common methods of preprocessing the images. Numerous papers have been published on the topic of developing, testing, and improving models for automated chest X-ray diagnosis. Especially successful were the models whose final prediction is the consensus of the predictions from its submodels. The addition of numerical patient data did not improve model performance, while the usage of image segmentation methods and Deep AUC Maximization methods did. As an example, one of the models was chosen and used on several chest X-rays. The results are presented, their meaning explained, and the background and problems of such models are discussed. A potential problem of extracting non-medical data from the X-rays and using them to make predictions is identified.

Keywords: artificial intelligence, image classification, automated diagnosis, chest X-ray

3 Uvod

Danas svjedočimo eksploziji primjene umjetne inteligencije, koja se sve više uspješno upotrebljava za rješavanje različitih problema iz širokog spektra ljudskih djelatnosti. Prepoznavanje objekata na slikama, prepoznavanje lica i govora, pretraživanje interneta, autonomna vozila, traženje najkraćeg puta samo su neki od njih. Posebno su impresivni i alati generativne umjetne inteligencije kao što je ChatGPT (1), koji komunicira s korisnikom na razini vrlo sličnoj ljudima.

Iako se velik dio teoretske podloge umjetne inteligencije razvio još u prošlom stoljeću (2), tek su se kroz proteklih nekoliko godina zadovoljili uvjeti za njenu efikasnu i raširenu upotrebu. Četiri su glavna razloga za to. (i) Računala su znatno moćnija i brža, što omogućuje bržu obradu podataka i treniranje modela umjetne inteligencije. Ključan razlog tome je ubrzan razvoj grafičkih procesorskih jedinica (GPU), što se može zahvaliti industriji računalnih igara. (ii) Više je podataka dostupno. Modeli umjetne inteligencije uglavnom trebaju veliku bazu podataka iz koje uče, a u današnje vrijeme su digitalizacija i internet omogućili efikasno stvaranje takvih baza. (iii) Razvijaju se novi algoritmi i nove arhitekture modela umjetne inteligencije. Današnje mogućnosti pristupu i obradi podataka potiču stvaranje sve efikasnijih modela, specifičnih za određeni zadatak. (iv) Tehnološki divovi investiraju ogromne svote novca u tržište umjetne inteligencije (3). Predviđanja su da će tržište, s vrijednosti od 86.9 mlrd. USD iz 2022., narasti na 407 mlrd. USD u 2027. (4).

Medicina je kompleksno područje koje obuhvaća mnoge izazove, od teoretskih kao što su dijagnosticiranje bolesti, odluka o liječenju i očitavanja radioloških snimaka, pa do praktičnih kao što su izvođenje operacija. Velik dio tih problema bi mogli biti znatno olakšani alatima koji koriste umjetnu inteligenciju i koji bi služili kao pomoć liječnicima, omogućavajući im da budu brži i precizniji, smanjujući im opseg posla i dajući im veću preciznost pri donošenju odluka. Korist od umjetne inteligencije ne bi bila samo u rješavanju isključivo medicinskih problema, već i u smanjenju administrativnih poslova i poslova na računalu, od čega bi posebno korist imali liječnici primarne zdravstvene zaštite.

Trenutno je jedna od najčešćih (i najperspektivnijih) uporaba umjetne inteligencije ona na medicinskim slikama (bilo radiološkim, dermatološkim, kolonoskopskim...). Iz slika bi se, najčešće pomoću konvolucijskih neuronskih mreža, automatski prepoznavalo određeno stanje te time pomoglo pri dijagnostici bolesti. Uspješnost te tehnike jako varira, a radiologija je, zbog svoje orijentiranosti slikovnim prikazima, područje gdje se ona vjerojatno najčešće primjenjuje ili pokušava primijeniti. Sve vrste radioloških slikovnih pretraga (rendgen, CT, ultrazvuk, magnetska rezonanca...) su potencijalni kandidati za primjenu takvog automatiziranog prepoznavanja stanja.

Motivirano nedavnim značajnim napretkom umjetne inteligencije i potencijalnom uporabom u području medicine (i konkretno radiologije), u ovom radu fokus će biti na primjeni u automatskoj dijagnostici stanja pomoću rendgenskih snimaka prsnog koša.

Analizirat će se trenutna uspješnost takvog pristupa, kao i koristi koje liječnici i sam proces dijagnostike imaju od njega. Poseban naglasak će biti na detekciji pneumonije. Prikazat će se i opisati najuspješniji modeli koji trenutno postoje te će se usporediti njihova uspješnost očitavanja slika s onima od radiologa. Na kraju, prikazat će se upotreba odabranog modela na nekoliko rendgenskih snimaka, prikazati i objasniti rezultati te će se raspraviti o podlozi i problemima načina rada takvih modela.

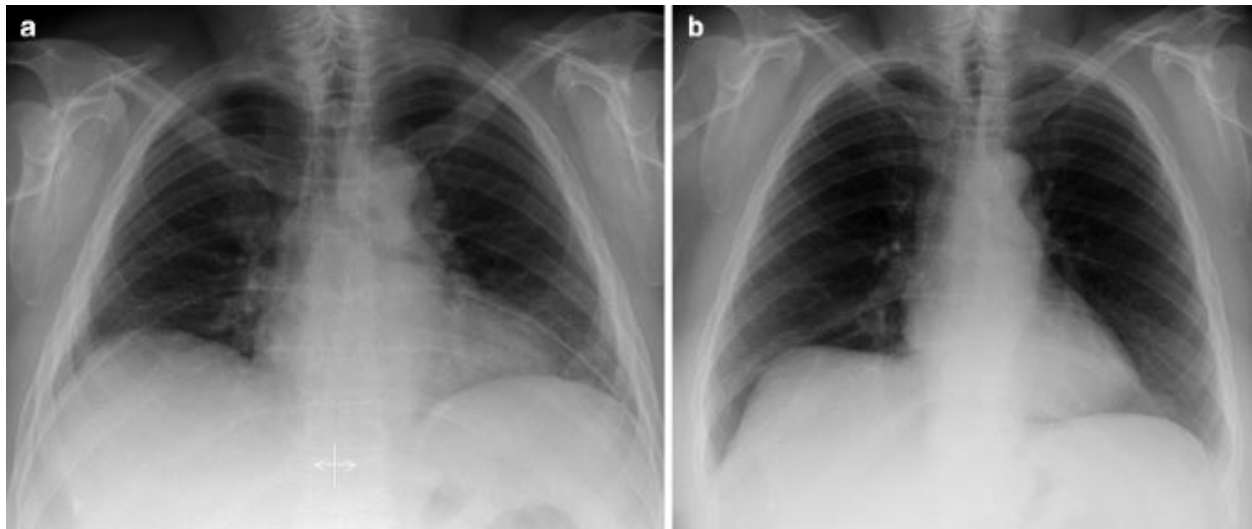
4 Rendgenske snimke prsnog koša

4.1 Tehnike snimanja

Ovisno o položaju tijela pacijenta i odnosu s rendgenskim uređajem, postoje različite vrste snimanja prsnog koša. U nastavku su opisani najčešće.

4.1.1 Posteriorno-anteriorno (PA)

Klasični prikaz kakav se radi na odjelima radiologije. Omogućava jasni prikaz pluća, srca, prsnog koša, medijastinuma, velikih krvnih žila i dijafragme (5). Rendgenska cijev je smještena dorzalno od pacijenta, a film ventralno. Pacijent je u uspravnom položaju, s rukama na bokovima ili oko detektora, obgrljujući ga. U takvom položaju se skapule pomiču lateralno i manje smetaju prolasku rendgenskih zraka kroz pluća. Za dobru snimku, pacijent bi trebao biti paralelno s detektorom, bez rotacije. Za procjenu rotacije, uspoređuju se sternoklavikularni zglobovi. Ukoliko je na jednoj strani njegova sjena šira, ta strana se treba pomaknuti bliže detektoru. Isto tako, rotacija u sagitalnoj ravnini (angulacija) se procjenjuje prema položaju medijalnih krajeva klavikule, koji bi trebali biti iznad stražnjeg dijela 3. do 4. rebra.



Slika 1. Razlika između PA snimke u ekspiraciji (lijevo) i inspiraciji (desno) (6). Prema: Delrue et al. Reproducirano uz dopuštenje Springer Nature.

Iako snimanje u ekspiraciji može imati svoj značaj (7), uobičajeno je da se snimke rade tijekom maksimalne inspiracije, uz zadržavanje daha. Desni dio dijafragme bi se na snimci u srednjoj klavikularnoj liniji trebao nalaziti između prednjeg dijela 5. do 7. rebra (8). Ukoliko to nije zadovoljeno, povećana je sjena srca, izražena je plućna vaskulatura, podignuta dijafragma i nejasan donji dio slike, što može dovesti do pogrešne dijagnoze. Na slici 1 vidimo razliku u snimkama u fazi inspiracije i ekspiracije.

4.1.2 Anteriorno-posteriorno (AP)

Koristi se umjesto PA projekcije kada pacijent iz nekog razloga ne može ustati (9). Najčešće se koristi u jedinicama za intenzivno liječenje, kada su pacijenti npr. intubirani ili u lošem općem stanju. Često se takvo snimanje vrši i prijenosnim rendgenskim uređajem. Rendgenska cijev je u ovom slučaju smještena sprijeda. Pacijenta se nastoji što je više moguće uspraviti, a leđa se naslanjaju na detektor. Zbog povećane udaljenosti srca od filma, sjena srca i medijastinuma je proširena, što je posljedica većeg raspršivanja zraka prije dolaska na detektor. Također, zbog sjedećeg položaja tijela je dijafragma obično smještena više.

4.1.3 Lateralno

Radi se kao nadopuna projekciji u frontalnoj ravnini (10). Bolje od nje prikazuje npr. retrosternalni prostor, retrokardijalni prostor, plućne fisure i posebno plućni hilus (11). Snima se u uspravnom položaju, uz podignute ruke, a ona strana toraksa koja je od interesa se postavlja bliže detektoru. Obično je to lijeva strana jer bolje prikazuje srce i strukture oko njega (12).

4.2 Analiza snimke

Nakon što se utvrdi da je snimka pravilno napravljena, pristupa se njenoj analizi. Pri manualnoj analizi i očitavanju slike, radiolozi se često vode kraticom ABCDE (A – airway / dišni put, B – breathing / disanje, C – cardiac / srce, D – disability / invalidnost, E – everything else) (13). Postoji više verzija značenja te kratice (14–16), a mi ćemo ovdje opisati postupak prema jednoj (13).

4.2.1 A – airway / dišni put (i medijastinum)

Traheja je tipično smještena u samoj sredini prsnog koša ili blago desno. Ako postoji znatna devijacija u jednu stranu, to može značiti postojanje nekog patološkog stanja. Npr. kod pneumotoraksa, tumora ili pleuralnog izljeva može postojati devijacija traheje na suprotnu stranu, dok je kod atelektaze ili nakon pneumonektomije prisutna ipsilateralna devijacija. Gleda se i postoji li suženje traheje na nekom mjestu.

Bifurkacija traheje se obično na snimkama jasno vidi. Osim suženja, treba promotriti i postoji li negdje strano tijelo. Desni bronh se nakon traheje nastavlja okomitije od lijevog, a isto tako je i širi. Zbog toga je strano tijelo češće prisutno u desnom bronhu.

Promatra se položaj i oblik aorte. Lijeva plućna arterija je normalno okomitija i vidi se, dok je desna horizontalnija i uglavnom se ne vidi. Potrebno je obratiti pažnju i na veličinu hilusa. Ukoliko je povećan, može znatići prisutnost sarkoidoze, limfoma ili tumora pluća.

4.2.2 B – breathing / disanje

Promatra se simetrija veličine plućnih režnjeva. Zasjenjenja nad plućima mogu postojati u obliku infiltrata, konsolidacija i masa. Iz njih se procjenjuje ima li pacijent pneumoniju (i je li ona tipična i atipična), tumor, plućni edem, krvarenje... Povećana prozračnost nad plućima može sugerirati na npr. pneumotoraks ili emfizem. Isto tako, promatra se naglašenost venskog sustava, koji je kod zdravih ljudi naglašeniji u nižim dijelovima pluća.

Pleura se kod zdravih ljudi ne vidi, a ukoliko se vidi, može značiti prisutnost fibroze, infekcije, mezotelioma i drugog. Važno je promotriti i kostofrenične kuteve. Njihov nestanak je uglavnom izazvan pleuralnim izljevom.

4.2.3 C – cardiac / srce

Sjena srca bi u klasičnoj PA snimci trebala zauzimati manje od 50% širine prsnog koša. Ukoliko zauzima više, prisutna je kardiomegalija, koja može biti uzrokovana npr. srčanim zatajivanjem, kardiomiopatijom, dugotrajnom hipertenzijom ili bolestima zalistaka. Treba obratiti pažnju i na granice srca. Ako su nejasne, to može značiti konsolidaciju preklapajućeg dijela pluća.

4.2.4 D – disability / invalidnost

Provjeravaju se oblik svih koštanih struktura, uključujući prednje i stražnje dijelove svih rebara, klavikule i kralješke. Treba obratiti pažnju na cjelovitost njihovih rubova, znakove prisutnosti fraktura te prisutnost metastaza i drugih lezija.

4.2.5 E – everything else / ostalo

Na snimci mogu biti prikazani i medicinski uređaji i pomagala, kao što su pacemakeri, umjetni zalisci i npr. žice od EKG uređaja. Posebno, pomoću rendgena se i provjerava odgovarajući položaj nazogastrične sonde, koja bi se trebala nastavljati okomito poslije bifurkacije traheje.

Ispod lijeve dijafragme se može vidjeti povećana prozračnost, čime se otkriva lokacija želuca, koji obično ima određenu količinu zraka u sebi.

Ako je prisutan zrak neposredno ispod dijafragme, a pogotovo ako razdvaja dijafragmu i jetru na desnoj strani, to ukazuje na perforaciju crijeva, jer iz njega izlazi zrak i nakuplja se suprotno od smjera sile teže.

4.3 Budućnost analize

Trenutno je dijagnostika rendgenskih snimki prsnog koša zasnovana na ručnoj analizi, vođenoj ABCDE smjernicama i drugim dijagnostičkim algoritmima. Takav pristup nije savršen. Kao prvo, ručna analiza je vremenski zahtjevna, što je poseban problem kod nejasnih snimki gdje je dijagnoza teže vidljiva te kod velikih skupova podataka. Također, postoje razlike u radu različitih radiologa zbog čega se za istu snimku mogu nekada

postaviti različite dijagnoze. To može biti, primjerice, zbog razlike u iskustvu radiologa ili zbog različitih kriterija prema kojima se određuje prisutnost bolesti. Uz to, čak i iskusniji radiolozi nekada ne mogu zapaziti suptilnije promjene na snimkama, koje bi mogle označavati početke razvoja bolesti.

Zbog svega toga, jasno je da postoji potreba za učinkovitijom i preciznijom analizom rendgenskih snimki koja bi doprinijela i standardizaciji pri donošenju dijagnoze. Kao rješenje se nameće primjena umjetne inteligencije i konkretno dubokog učenja, koje se danas uspješno koristi pri rješavanju sličnih problema. Prednost takvih tehnologija je u brzini analize samih rendgenskih snimki, svrstavanje snimke u velik broj kategorija te mogućnost detekcije suptilnih promjena, a primjena istog modela na različitim snimkama bi značila i da postoji određena standardizacija pri dijagnosticiranju.

U nastavku rada se opisuju današnji doseg i problemi primjene takvih tehnologija na rendgenskim snimkama prsnog koša kao alata koji ima potencijal značajno unaprijediti pristup dijagnostici i liječenju pacijenata.

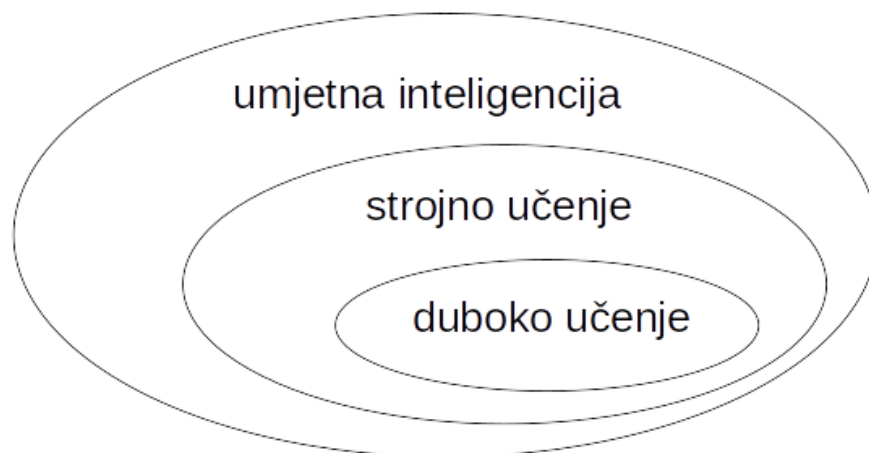
5 Umjetna inteligencija

5.1 Općenito

Umjetna inteligencija se krenula razvijati još 1950-ih godina, a poseban značaj za pokretanje njenog razvoja je imala konferencija u Dartmouthu 1956., koju je inicirao John McCarthy (17). Na njoj je prvi put definiran pojam „umjetna inteligencija” kao znanstvena disciplina koja se bavi izgradnjom sustava čije se ponašanje može tumačiti kao inteligentno. Konkretnije, to je nastojanje da strojevi rješavaju probleme koje trenutno mogu riješiti samo ljudi, i da budu sve bolji u tome.

Važne ideje je iznio i Alan Turing u svome članku iz 1950. objavljenom u časopisu „Mind” (18). Predstavio je ideju o „igri imitacije”, gdje postoje dva igrača A i B – jedan je čovjek a drugi je stroj. Treći sudionik C je čovjek te on postavlja pitanja igračima A i B. Na kraju igre, sudionik C pogađa koji od igrača A i B je računalo. Ukoliko pogriješi u 50% slučajeva igranja igre, stroj (A ili B) bi se mogao smatrati inteligentim. Takav pristup ima svojih mana, ali je zanimljiv kao teorija.

Izuzetno važan koncept korišten u umjetnoj inteligenciji je koncept umjetnog neurona, koji je nastao inspiriran načinom na koji funkcioniraju stvarni, biološki neuroni. Taj koncept su osmislili i iznijeli McCulloch i Pitts u članku iz 1943. godine (19), a ideju o modificiranju veze između neurona je razradio Hebb u knjizi „The Organization of Behavior” iz 1949.



Slika 2. Shema pojmova umjetne inteligencije važnih za ovaj rad.

Razlikujemo dva glavna pristupa umjetnoj inteligenciji. (i) Simbolički pristup manipulira podacima i zaključuje iz njih na temelju točno definiranih pravila prema kojima donosi odluke. Takav pristup je bio puno popularniji između 1950. i 1990. godine, a onda ga je zamijenio konekcionistički, čija popularnost traje i danas. (ii) Konekcionistički pristup ne postavlja tako oštra pravila već zaključuje prema nečemu što više slični ljudskoj intuiciji i načinu na koji funkcionira ljudski mozak. Tipičan primjer su umjetne neuronske mreže, koje uče iz skupine podataka i to „znanje” onda mogu primijeniti na podatke koje još nisu vidjele.

Jasno je iz svega do sada da je umjetna inteligencija širok pojam koji obuhvaća razne pristupe baratanja s informacijama i donošenja odluka i zaključaka. Slika 2 prikazuje shemu važnu za ovaj rad.

5.2 Strojno učenje

Za razliku od tipičnog programiranja koje bi postavilo točna pravila prema kojima se program mora ponašati, strojno učenje funkcionira tako da program iz skupa poznatih podataka sam „nauči” ta pravila. Konkretnije, model strojnog učenja ima određene početne parametre (težine) prema kojima iz zadanih ulaznih podataka računa neko rješenje. Ti parametri (težine) se „učenjem” (treniranjem) podešavaju tako da rješenje bude što točnije. Tijekom učenja se modelu predstavljaju poznate kombinacije ulaznih podataka i onoga što se očekuje da model predvidi za te podatke. Cilj toga je da model za podatke koje do sada nije vidio, može predvidjeti točno rješenje (sposobnost generalizacije modela). Ovakav pristup je prikladan kod problema koji se ne znaju riješiti egzaktnim algoritmima, jer se ne zna ni kako ih ljudski mozak rješava.

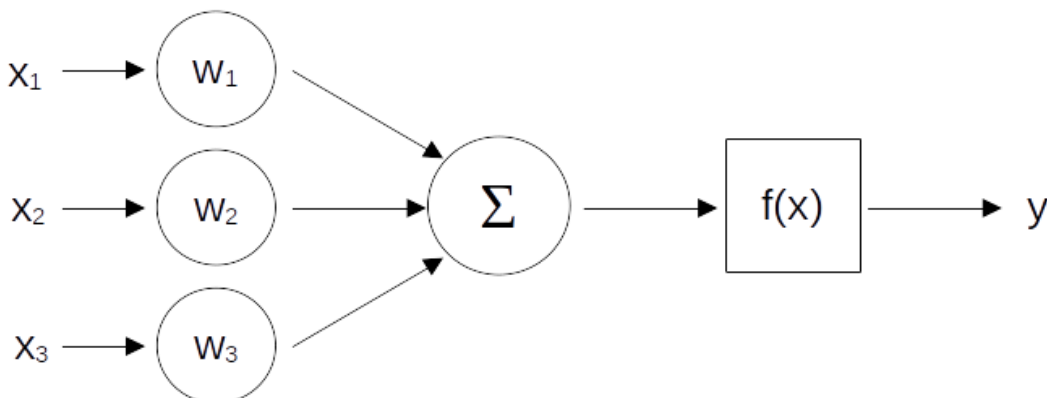
Obično se set dostupnih podataka (koji se sastoji od ulaznih podataka i izlaznih podataka koje model treba predvidjeti) dijeli na tri dijela: set za treniranje, validaciju i testiranje. Set za treniranje se zatim koristi na idući način. Određeni broj ulaznih podataka se provodi kroz model i tada model daje neko rješenje. To rješenje se uspoređuje s onim koje je zapisano kao rješenje za te ulazne podatke („ground truth”). Ovisno o tome koliko je model pogriješio u predviđanju, parametri (težine) modela se podešavaju tako da model u buduću daje rješenje koje je bliže istini. To podešavanje težina se danas kod umjetnih neuronskih mreža uobičajeno radi pomoću matematičkog „backpropagation” algoritma, a upravo je prepoznavanje potencijala tog algoritma 1986. (20) dovelo do popularizacije konekcionističkog pristupa u umjetnoj inteligenciji. Set za validaciju se koristi da bi se provjerila učinkovitost modela i odredili hiperparametri za koje model funkcionira bolje (npr. broj neurona u sloju neuronske mreže), a zatim se na setu za testiranje provodi konačna evaluacija modela.

Kao primjer usko vezan za ovaj rad, izgradimo model za predikciju prikazuje li rendgenska snimka pneumoniju ili ne. Taj model će biti treniran (učiti) na skupu snimki za koju je svaku određeno je li pneumonija ili nije. Zatim, model će biti korišten na način da za novu snimku (koja nije iz skupa na kojoj je treniran) predvidi je li na njoj pneumonija.

Već spomenuti umjetni neuron, odnosno mreže izgrađene od njih, primjer su modela strojnog učenja. Na slici 3 je prikazana shema jednog jednostavnog umjetnog neurona.

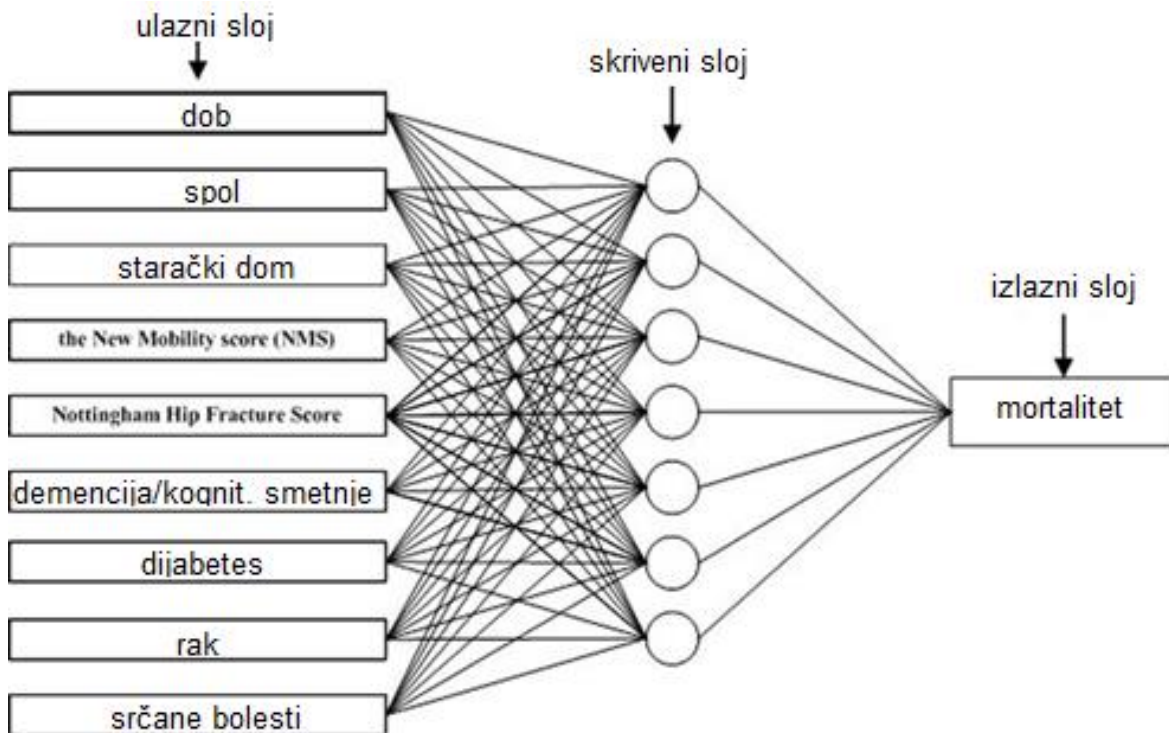
Ulazni podatci su označeni s „ x ”, a predviđena vrijednost s „ y ”. Parametri neurona, odnosno težine, označeni su s „ w ”, a upravo ti parametri se učenjem podešavaju da izlaz iz neurona „ y ” bude odgovarajuć. Nakon što se ulazi pomnože s težinama, sumiraju se i provode kroz aktivacijsku funkciju „ $f(x)$ ”. Često korištene aktivacijske funkcije su sigmoidna, tangens hiberbolni i ReLU. Važno je da ona nije linearna. Uloga te aktivacijske

funkcije je ključna za neuron i upravo je taj dio inspiriran biološkim neuronima. Ako su „x” depolarizacijski signali ulaznih bioloških neurona, a „y” izazvani akcijski potencijal, aktivacijska funkcija na neki način simulira pokretanje akcijskog potencijala tek kad se agregira dovoljan depolarizacijski signal na aksonskom brežuljku. Ipak, u umjernom neuronu „y” neće biti samo 0 ili 1, nego veći raspon vrijednosti, ali to nije važno za njegovu funkciju.



Slika 3. Umjetni neuron. $x_{1,2,3}$ – ulazne vrijednosti; $w_{1,2,3}$ – težine neurona; Σ - suma; $f(x)$ – aktivacijska funkcija; y – izlazna vrijednost

Akcijski potencijal u biološkom neuronu zatim stimulira sljedeće neurone u nizu, pa se tako i u neuronskim mrežama taj „y” dovodi kao ulaz („x”) idućem neuronu.

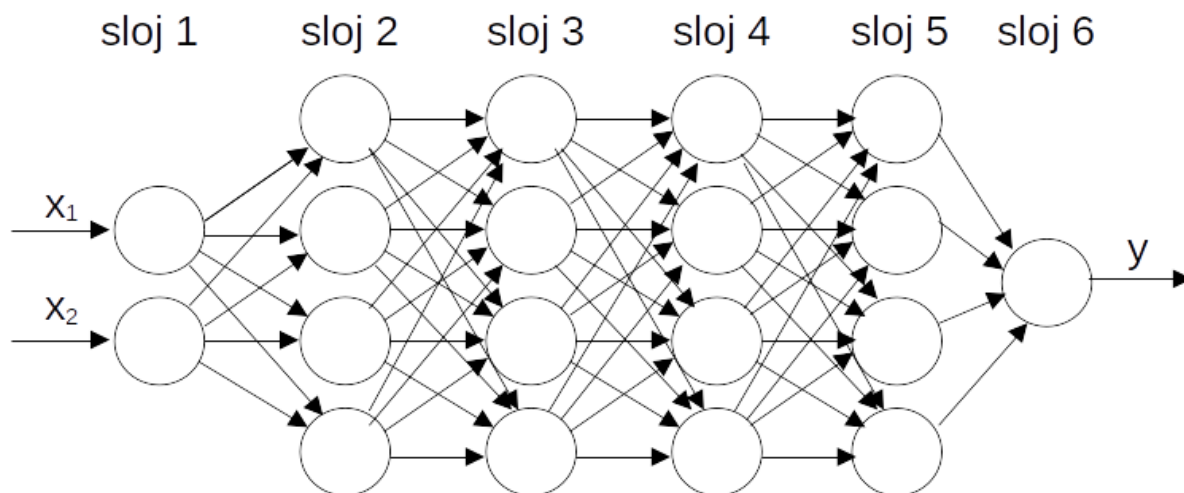


Slika 4. Neuronska mreža za predviđanje 1-godišnjeg mortaliteta pacijenata s intertrohanteričnim prijelomom femura (21). Prema: Shi et al. Dostupno uz licencu: CC BY 4.0

Neuronske mreže su mreže sastavljene od navedenih neurona. Kao primjer neuronske mreže za potrebe medicine, ulazi bi mogli biti simptomi bolesti, a izlaz dijagnosticirana bolest. Na slici 4 je prikazan primjer neuralne mreže korištene u radu (21), koja predviđa 1-godišnje preživljenje u starijih pacijenata s intertrohanteričnom frakturom femura.

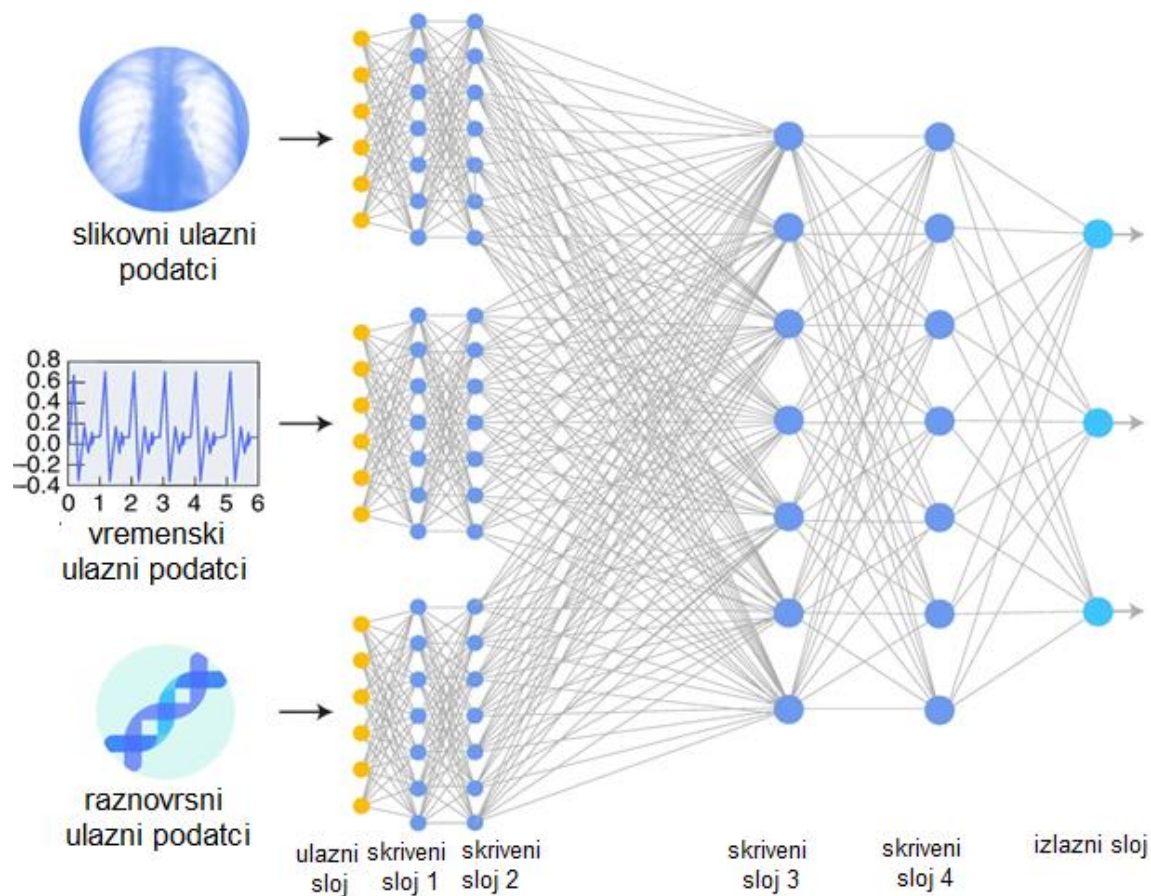
5.3 Duboko učenje

Na slici 4 je prikazan samo jedan sloj umjetnih neurona (hidden layer) koji prima ulazne informacije i u izlazu odmah predviđa mortalitet. Takav jednoslojni pristup ima samo ograničenu uspješnost. Zbog toga se danas modeli grade od puno slojeva nanizanih jedan za drugim, čineći višeslojne, odnosno duboke mreže. Time se znatno povećavaju njihove mogućnosti rješavanja kompleksnih problema. Korištenje takvog pristupa se zove „duboko učenje“. Na slici 5 prikazan je primjer takvog slaganja slojeva.



Slika 5. Primjer modela dubokog učenja (od 6 slojeva).
 $x_{1,2}$ – ulazne vrijednosti; y – izlazna vrijednost

Ovisno o problemu, naravno, moguć je različit broj ulaza, izlaza, slojeva te neurona u sloju. Obično se za neki problem isprobavaju različite arhitekture pa se odabire ona najpovoljnija. Potreban je inženjerski pristup izgradnji takvih modela, a iskustvo je jako važno jer se puno odluka donosi „po osjećaju“ i nema garancije da će neki pristup biti najbolji. Ipak, otprilike se zna kakva arhitektura bi vjerojatno bila najbolja za neki problem. Primjer kompleksnije arhitekture za primjenu u medicini iz rada (22) je prikazan na slici 6. Posebno je zanimljivo obratiti pažnju na različite tipove ulaznih podataka (slikovni, vremenski...) jer se tu može vidjeti širina informacija koje model dubokog učenja može obraditi i nešto iz njih zaključiti.

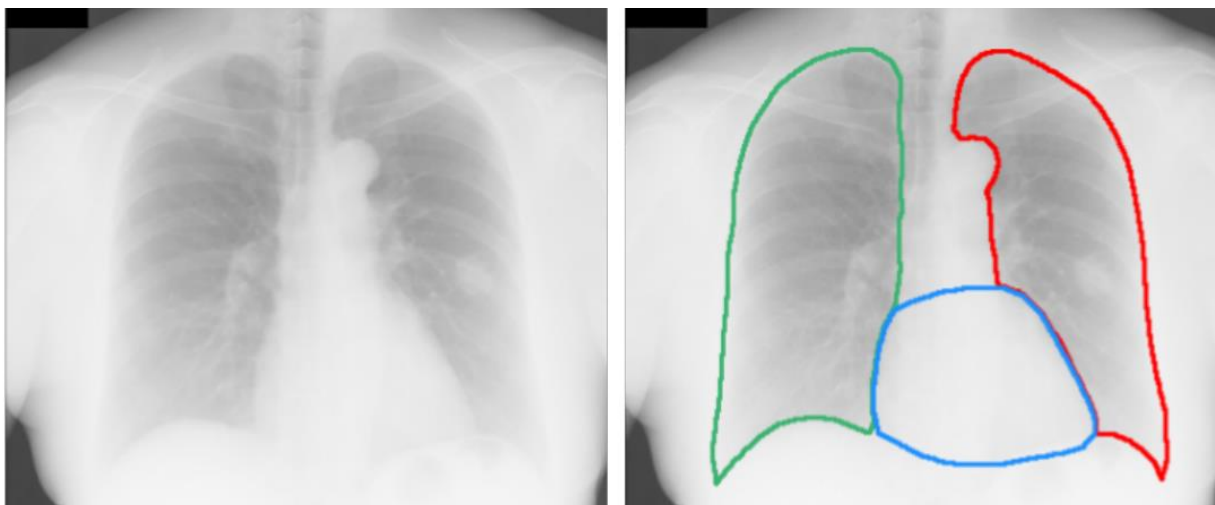


Slika 6. Primjer kompleksne arhitekture kakva se može primijeniti u medicini (22). Prema: Esteva et al. Reproducirano uz dopuštenje Springer Nature.

5.4 Računalni vid

Grana umjetne inteligencije koja se bavi obradom vizualnih informacija, tipično slika i videozapisa, zove se računalni vid. Više je problema kojima se bavi računalni vid. (i) Prvo treba spomenuti klasifikaciju slika. Primjer toga je svrstavanje rendgenske snimke u kategorije s obzirom to je li prisutna pneumonija ili ne. Također je moguće i lokaliziranje dijela zahvaćenog pneumonijom na slici. (ii) Važno je spomenuti i detekciju tj. prepoznavanje objekata na slici, što se odnosi na npr. odvojeno prepoznavanje pneumonije, tumora i nodula na rendgenskoj snimci. (iii) Semantička segmentacija je npr. prepoznavanje dijela snimke gdje se nalazi srce te lijevo i desno plućno krilo od ostatka snimke (prikazano na slici 7), a (iv) segmentacija instanci razdvaja različite pojave na slici koje pripadaju istoj kategoriji, npr. prepoznavanje različitih plućnih nodula.

U medicini postoje brojni problemi u kojima bi se mogao koristiti računalni vid. Samo neki od njih su očitavanja radioloških snimki (rendgenskih, CT, MR, ultrazvuk...), prepoznavanje dermatoloških stanja, analiza kolonoskopskih snimki, analiza snimki retine, detekcija objekata kod robotski asistiranah operacija...



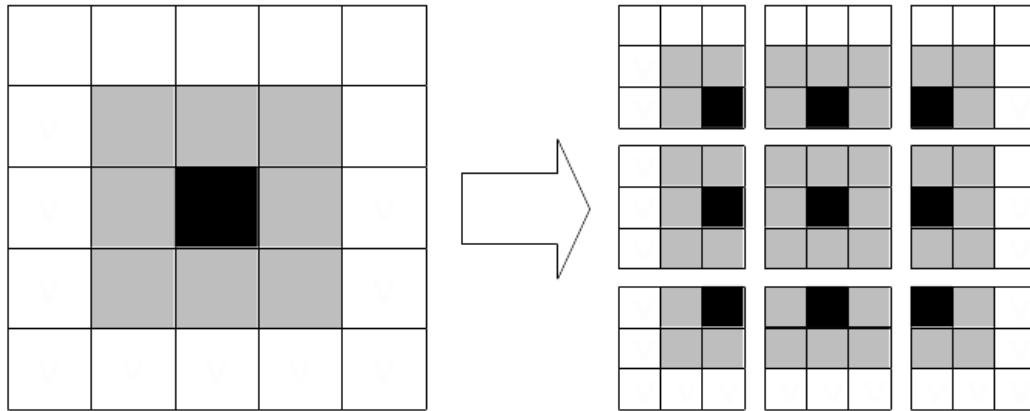
Slika 7. Semantička segmentacija srca te lijevog i desnog plućnog krila (23) Prema: Dai et al. Reproducirano uz dopuštenje Springer Nature.

Već je navedeno kako su arhitekture modela dubokog učenja vrlo raznovrsne. Veliku razliku između njih čini i odabir vrste slojeva neurona. Najjednostavniji tip sloja je „potpuni”, gdje su svi neuroni prethodnog sloja povezani linearnom funkcijom sa svim neuronima idućeg sloja, kao što je prikazano na slici 5. Analizirajmo što bi se dogodilo s obradom slikovnih informacija ukoliko bi se samo ta vrsta slojeva koristila u modelima računalnog vida. Uzmimo da su pikseli slika ulazni podatci za takav model te su označeni s $x_1, x_2, x_3...$ Model bi tada previše uzimao u obzir globalni kontekst i točnu lokalizaciju piksela na slici i ne bi mogao efikasno prepoznavati objekte kad bi oni bili smješteni na različitim mjestima na slici. Taj problem se rješava koristeći konvolucijske neuronske mreže, tip modela kakav je posebno efikasan za rješavanje problema računalnog vida.

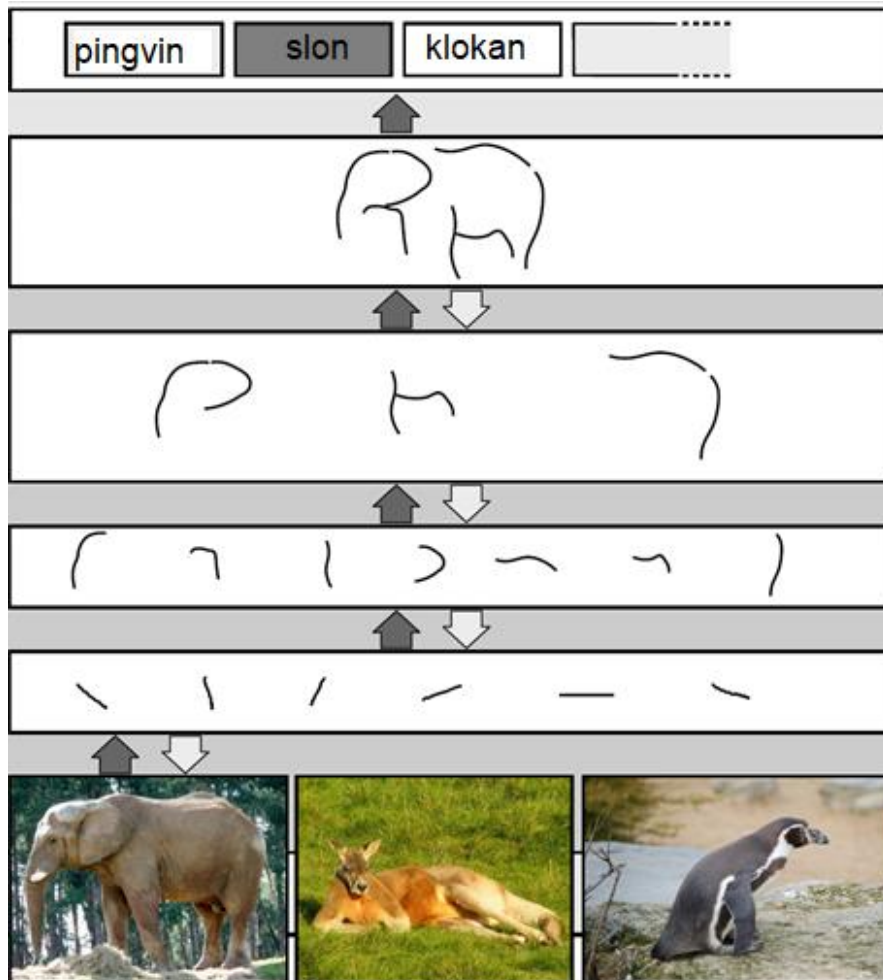
Konvolucijski slojevi funkcioniraju otprilike tako da pomiču manji okvir po početnoj slici i svaki od tih manjih prikaza šalju dalje u neuronsku mrežu. Slika 8 prikazuje primjenu tog principa na slici 5x5 piksela, koristeći okvir 3x3. Na taj način početna lokacija objekta na slici gubi na značaju te se objekt lakše prepoznaje.

Poseban značaj konvolucijskih neuronskih mreža je što slojevi imaju sposobnost hijerarhijskog spajanja jednostavnijih slikovnih obilježja u kompleksija. Time se omogućava da dublji slojevi mreže stvaraju reprezentacije informacija koje su vrlo direktno informativne o konačnom predviđanju. Slika 9 jasnije opisuje navedeni princip.

U modelima se nakon konvolucijskih slojeva koristi i „pooling” sloj. Taj sloj koristi rezultate primjene manjih okvira (desno na slici 8), svakog transformira u manji (obično samo uzme piksel najveće vrijednosti) i zatim sve te spoji u jedan veći. Time omogućava da se informacije zaključuju iz cijele slike tijekom cijele propagacije kroz neuronsku mrežu.



Slika 8. Ulazna slika 5x5 piksela (lijevo) i rezultat pomicanja okvira 3x3 piksela po njoj (desno).



Slika 9. Prikaz različitih reprezentacija slika kroz slojeve (24) Prema: Schultz et al. Uz dopuštenje Springer Nature.

5.5 Predtrenirani modeli

Nekad se u izgradnji novih modela koriste već prethodno izgrađeni i trenirani modeli korišteni za rješavanje sličnih problema. To je posebno izraženo kod računalnog vida, gdje postoji velik broj takvih predtreniranih modela.

Daleko najpopularnija baza slika za treniranje modela računalnog vida je ImageNet (25), a sastoji se od preko 14 mil. slika svrstanih u preko 20 000 kategorija. Većina javno dostupnih predtreniranih modela je trenirano na toj bazi podataka i naučeno da vrlo uspješno raspoznaje objekte na slikama.

Treba imati na umu da takvi modeli imaju već naučene težine/parametre koji im omogućuju da kroz slojeve raspoznaju razne obrasce na slikama (kao na primjer na slici 9). Ima smisla, stoga, uzeti konkretno takav naučeni model i zadržati taj prvi dio modela isti, a promijeniti i naučiti samo dio modela koji je bliže izlazu (i prilagoditi izlaze našim potrebama). Na primjer, ako se želi raspoznati je li na rendgenskoj slici pneumonija ili ne, uzet će se neki predtrenirani model, zamrznut će se njegovi parametri u svim slojevima osim nekoliko zadnjih. Uz to će se promijeniti i izlaz na način da vraća informaciju je li na slici pneumonija ili ne. Zatim će se model trenirati na rendgenskim slikama, ali podešavajući samo parametre zadnjih nekoliko slojeva. Na taj način se može dobiti puno veću uspješnost nego da je cijeli model treniran iz početka.

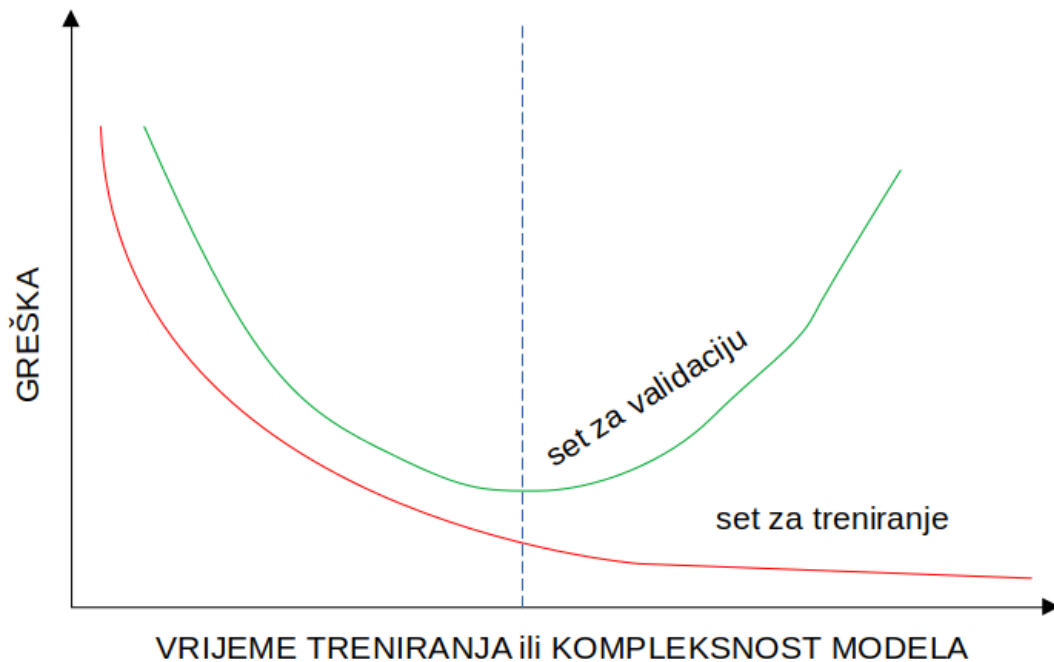
Neki od najpopularnijih predtreniranih modela, koji su korišteni za cijeli spektar problema računalnog vida, pa tako i za analizu rendgenskih snimki, su: VGG-16 (26), ResNet50 (27), Inceptionv3 (28), DenseNet (29)...

5.6 Sposobnost generalizacije modela

Za sve modele umjetne inteligencije jako je važna sposobnost generalizacije. To je pojam koji označava ispravno donošenje zaključaka iz podataka koje model prethodno nije vidio tj. nije na njima treniran.

S druge strane, „overfitting”, odnosno pretreniranje modela je pojava kada je model „prenaučio” podatke koji se koriste za treniranje. Zbog toga radi sve bolja predviđanja na setu za treniranje, ali gubi sposobnost predviđanja na setu za validaciju, odnosno gubi sposobnost generalizacije. Najčešći razlog tome je predugo treniranje modela ili prevelika kompleksnost modela.

Na slici 10 se vidi kako se mijenja greška u predviđanjima modela s obzirom na vrijeme treniranja ili kompleksnost modela. Vidimo kako greška na setu za validaciju (odnosno podatcima koje model vidi po prvi puta) u jednom trenutku krene rasti, dok pogreška na setu za treniranje nastavlja biti sve manja. Optimalno treniranje modela (ili optimalna kompleksnost) je označeno isprekidanom plavom crtom.



Slika 10. Promjena greške na setovima za treniranje i validaciju s obzirom na vrijeme treniranja ili kompleksnost modela.

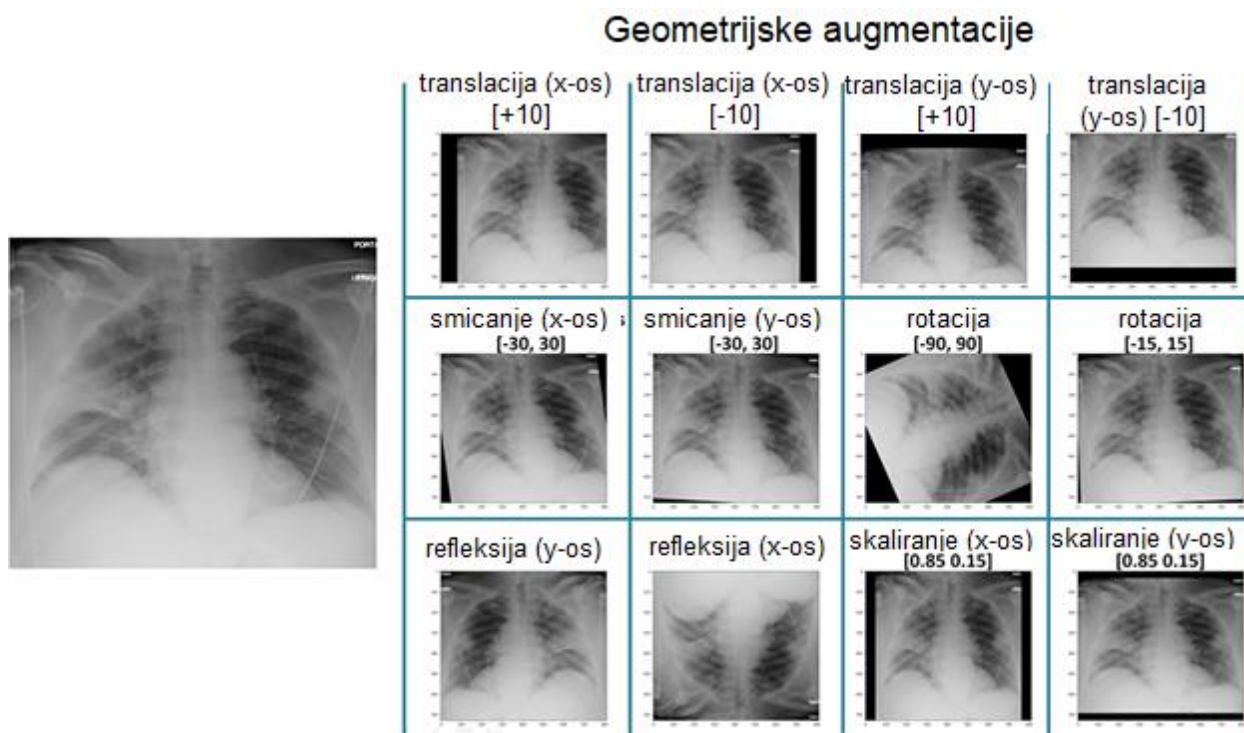
Brojne tehnike se mogu koristiti kako bi se spriječio „overfitting”. Ovdje su navedene neke od njih.

Veći set za treniranje dovodi do smanjivanja overfittinga, ali nekada je jako teško doći do dodatnih podataka za treniranje. Augmentacija podataka je način umjetnog povećavanja seta za treniranje na način da se postojeći podatci na neki način malo promijene, ali i dalje se očekuje od modela da može zaključiti iz njih. Na primjer, ako su ulazni podatci slike, one se mogu zrcaliti, rotirati, odrezati, povećavati, smanjivati, pomicati, sužavati... Model se zatim trenira i na tim promijenjenim slikama. Takav pristup se često koristi i u radovima vezanim za automatsku detekciju stanja na rendgenskim snimkama prsnog koša, a primjer je prikazan na slici 11.

Smanjivanjem kompleksnosti modela također bi se smanjio overfitting. To se može postići na primjer smanjivanjem slojeva neuronske mreže ili smanjivanjem broja neurona u slojevima.

Overfitting će se smanjiti i ako se treniranje modela zaustavi točno u trenutku isprekidane plave crte na slici. To se inače postiže tako da se model naizmjenično trenira i evaluira (kroz više ciklusa - epoha). Model se inače nakon svake epohe sprema u memoriju. U trenutku kad se primijeti da je krivulja greške na setu za validaciju krenula rasti, treniranje se zaustavlja, a kao konačni model se uzima onaj koji je u memoriju spremljen u trenutku najmanje greške na setu za validaciju.

Također se u modelima mogu koristiti „dropout” slojevi, koji privremeno izbacuju doprinose izlaza nekih neurona prema sljedećem sloju.



Slika 11. Primjer augmentacije podataka iz rada o detekciji COVID-19 (30) Prema: Elgendi et al. Uz licencu: CC BY 4.0.

5.7 Mjere za procjenu efektivnosti modela

Kod procjene predikcije binarnih labela (npr. pneumonija nije prisutna - negativno ili pneumonija je prisutna - pozitivno) koriste se sljedeći izrazi:

TP (true positives) – predviđeni su kao pozitivni i zapravo su pozitivni

FP (false positives) – predviđeni kao pozitivni, ali su zapravo negativni

TN (true negatives) – predviđeni kao negativni i zapravo su negativni

FN (false negatives) – predviđeni kao negativni, ali su zapravo pozitivni

TOČNOST: $(TP + TN)/(TP + FP + TN + FN)$

PRECIZNOST: $TP/(TP + FP)$

OSJETLJIVOST: $TP/(TP + FN)$

SPECIFIČNOST: $TN/(TN + FP)$

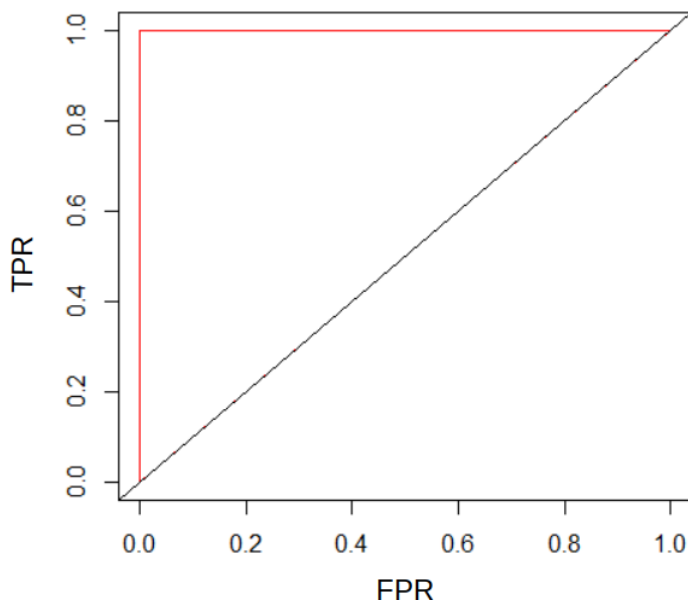
F1 MJERA: $(2 * PRECIZNOST * OSJETLJIVOST)/(PRECIZNOST + OSJETLJIVOST)$

ROC KRIVULJA – govori o odnosu „true positive rate” i „false positive rate” vrijednosti, a te su vrijednosti definirane na sljedeći način:

True Positive Rate (TPR) = OSJETLJIVOST = $TP/(TP + FN)$

False Positive Rate (FPR) = $1 - \text{SPECIFIČNOST} = FP/(TN + FP)$

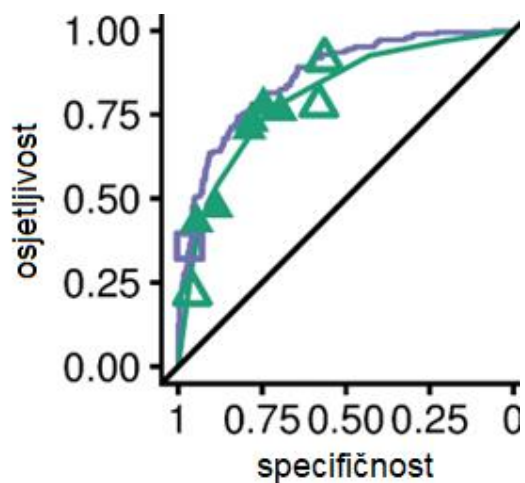
Savršena ROC krivulja ima konstantnu osjetljivost od 100% i FPR od 0%, a prikazana je crvenom bojom na slici 12. Može se reći da što je krivulja više smještena prema gore i lijevo, model je bolji.



Slika 12. Savršena ROC krivulja (prikazana crvenom bojom) (31).
Uz licencu: CC BY-SA 4.0.

Primjeri stvarnih ROC krivulja (32) su prikazani na slici 13. U tom konkretnom slučaju se vidi da je model dubokog učenja nadmašio radiologe u prepoznavanju atelektaze jer je krivulja modela smještena više gore i lijevo.

Mjera AUC (area under the curve) označava veličinu površine ispod ROC krivulje i ta se mjera najčešće koristi pri procjeni modela za detekciju stanja na rendgenskim snimkama prsnog koša.



Slika 13. Usporedba ROC krivulja radiologa (zeleno) i modela dubokog učenja (ljubičasto) pri detekciji atelektaze iz rada (32). Prema: Rajpurkar et al. Uz licencu: CC BY 4.0

6 Baze podataka rendgenskih snimaka prsnog koša

Velika baza podataka je nužna za uspješno treniranje modela, a danas postoje velike i javno dostupne baze podataka rendgenskih slika prsnog koša. Razlikuju se po broju različitih stanja čije slike sadrže, a i načinu na koji su slike anotirane (npr. neke su namijenjene samo za treniranje modela koji prepoznaju pneumoniju, a druge za treniranje modela koji prepoznaju širok spektar različitih stanja). Većina modela koji su danas popularni su trenirani na nekoj od tih baza. Ovdje navodimo najznačajnije i najveće od njih.

ChestX-ray14 (NIH) (33) - sastoji se od 112.120 frontalnih snimki prsnog koša 30.805 različitih pacijenata, prikupljenih između 1992. i 2015. godine. Slike su anotirane s 14 labela različitih stanja (1 - atelektaza, 2 - kardiomegalija, 3 - izljev, 4 - infiltracija, 5 - masa, 6 - nodul, 7 - pneumonija, 8 - pneumotoraks, 9 - konsolidacija, 10 - edem, 11 - emfizem, 12 - fibroza, 13 - pleuralno zadebljanje, 14 - hernija). Jedna slika može biti označena s više oznaka klasa, često u ovom kontekstu nazivanih i labele. Labele su izvučene iz radioloških izvještaja koristeći metode obrade prirodnog jezika, još jedne grane umjetne inteligencije. Pretpostavlja se da točnost takvih labela prelazi 90% što je dovoljno za učinkovito treniranje modela. Originalni radiološki izvještaji nisu javno dostupni. Ova baza je nadogradnja originalne iz rada (34), koja je sadržavala 8 labela.

CheXpert (35) – sadrži 224.316 snimaka od 65.240 različitih pacijenata. Također se sastoji od 14 labela (1 - no finding, 2 - enlarged cardiom., 3 - cardiomegaly, 4 - lung lesion, 5 - lung opacity, 6 - edema, 7 - consolidation, 8 - pneumonia, 9 - atelectasis, 11 - pneumothorax, 12 - pleural effusion, 13 - pleural other, 14 - fracture, 15 - support devices) izvučenih iz izvještaja radiologa koristeći umjetnu inteligenciju. Svaki snimak sadrži informaciju o svakoj labeli je li stanje prisutno (1), nije prisutno (0) ili se ne može reći sa sigurnošću (u - uncertain). Ta nesigurnost uključuje nesigurnost koju je radiolog iznio u izvještaju i neprecizno izražavanje radiologa iz kojeg se ne može konkretno zaključiti.

MIMIC-CXR (36) – sadrži DICOM snimke 227.835 oslikavanja (obično frontalna i lateralna projekcija – ukupno 377.110 snimaka) 65.379 pacijenata koji su zaprimljeni između 2011. i 2016. u „Beth Israel Deaconess Medical Center Emergency Department”. Uz svaku sliku je priložen izvještaj radiologa koji je napisan kratko nakon oslikavanja. Iz izvještaja su izbačeni bilo kakvi osobni podatci o pacijentu. Postoji i verzija sa slikama u JPG formatu [A9], gdje su snimke labelirane s 14 labela (istovjetnima labelama u bazi CheXpert)

CXLSeg (37) – Bazirana na MIMIC-CXR bazi. Sadrži 243.324 frontalnih snimki sa segmentacijskim maskama pluća (nalik maskama za lijevo i desno plućno krilo sa slike 8).

PadChest (38) – Sastoji se od svih snimki interpretiranih od strane 18 radiologa između 2009. i 2017. u „Hospital Universitario de San Juan” u Alicanteu u Španjolskoj. Takvih snimki ima 168.861 (iz 109.932 oslikavanja). U ovom slučaju je 25% izvještaja radiologa bilo manualno labelirano, a za ostatak je korišteno duboko učenje. Za opis snimki su

korištene različite labele, a one se mogu podijeliti u dijagnoze (19 različitih – pneumonija, KOPB, emfizem...), radiološki pronalasci (170 različitih - infiltrati, intersticijski uzorak, mase, noduli, kateteri...) i anatomske lokacije (104 različitih – plućna polja, medijastinum, kosti, hilus, aorta...).

VinDr-CXR (39) – Sadrži 18.000 frontalnih snimaka prsnog koša izdvojenih iz baze od 100.000 snimaka iz dvije bolnice u Vijetnamu. Na slici 14 prikazane su neke od snimki iz baze. Za razliku od svih prethodno navedenih baza, ovdje su slike anotirane ručno, a u tome je sudjelovala grupa od 17 radiologa. Set je podijeljen u dio za treniranje (15.000 snimaka) i za testiranje (3.000). U anotiranju svake snimke seta za treniranje su sudjelovala 3 radiologa, a set za testiranje je anotiran još pažljivije, pa je tako u anotiranju svake snimke sudjelovalo 5 radiologa. Labele su podijeljene u dvije skupine: lokalne (konkretan objekt na snimci, ručno označen pravokutnikom, 22 različite) i globalne (sveukupni dojam radiologa sa snimke, označava bolest, 6 različitih).

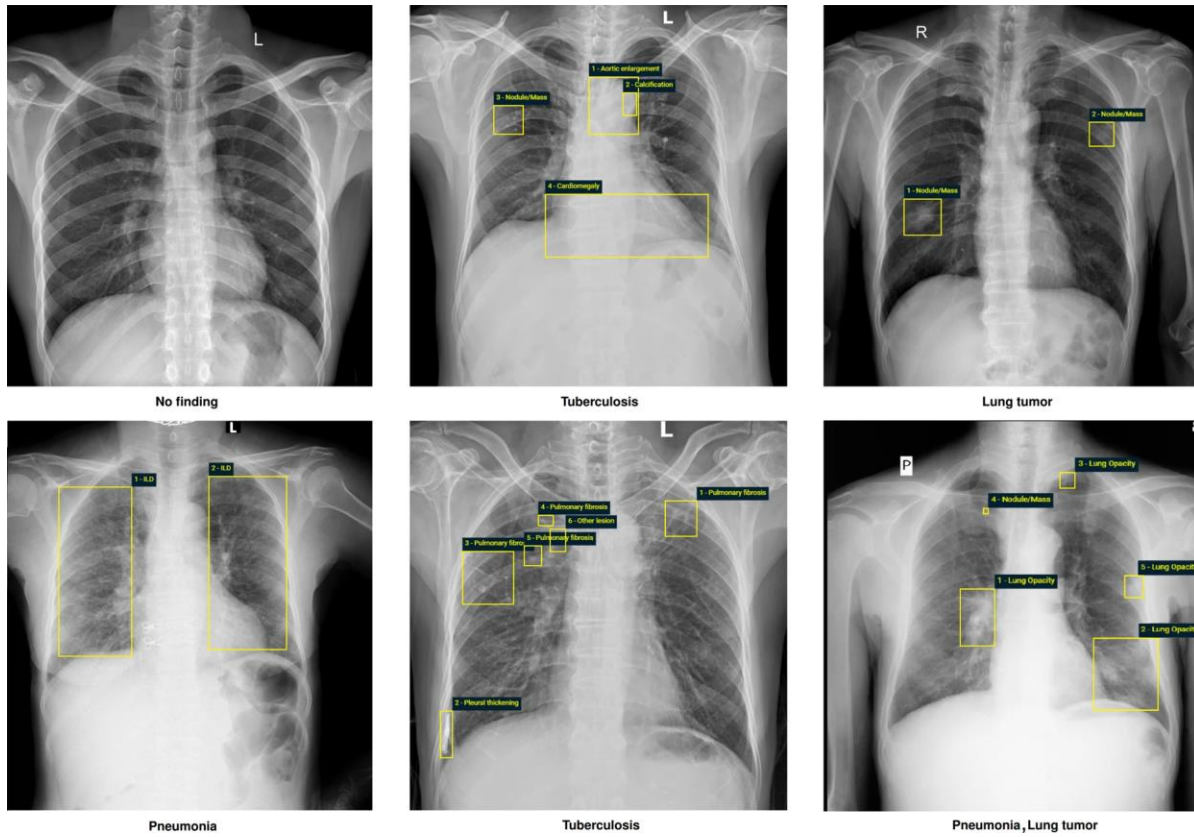
Lokalne: 1 - Aortic enlargement, 2 - Atelectasis, 3 - Cardiomegaly, 4 - Calcification, 5 - Clavicle fracture, 6 - Consolidation, 7 - Edema, 8 - Emphysema, 9 - Enlarged PA, 10 - Interstitial lung disease (ILD), 11 - Infiltration, 12 - Lung cavity, 13 - Lung cyst, 14 - Lung opacity, 15 - Mediastinal shift, 16 - Nodule/Mass, 17 - Pulmonary fibrosis, 18 - Pneumothorax, 19 - Pleural thickening, 20 - Pleural effusion, 21 - Rib fracture, 22 - Other lesion

Globalne: 23 - Lung tumor, 24 - Pneumonia, 25 - Tuberculosis, 26 - Other diseases, 27 - Chronic obstructive pulmonary disease (COPD), 28 - No finding

Indiana (40) – Baza od 8121 snimki i 3996 izvještaja radiologa iz „Indiana Network for Patient Care”. Jedna od prvih baza koja je objavila snimke s izvještajima (nakon deidentifikacije).

Montgomery County (41) – ili „MC” set - sadrži 138 frontalnih snimaka iz programa za prevenciju tuberkuloze Montgomery County-ja. 80 snimki iz baze je normalno, 58 su slučajevi manifestirane tuberkuloze.

Shenzhen (41) – iz „Shenzhen No.3 People’s Hospital, Guangdong Medical College, Shenzhen” Sadrži 662 frontalne snimke, od kojih je 326 normalno, a 336 su manifestacije tuberkuloze.



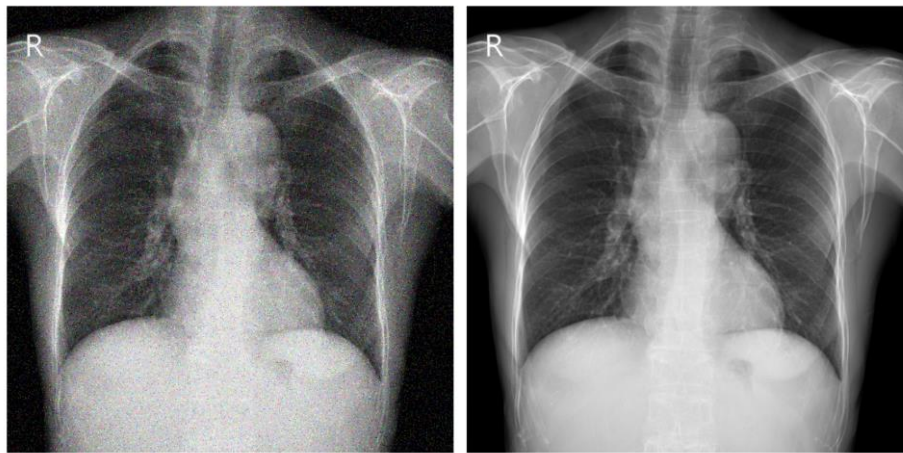
Slika 14. Primjer anotiranih snimaka iz baze VinDr-CXR (39).
 Prema Nguyen et al. Dostupno uz licencu: CC BY 4.0

7 Predobrada rendgenskih snimaka

Kako bi modelu olakšali zaključivanje iz rendgenskih snimki, one se prije provlačenja kroz model raznim metodama nastoje poboljšati, tj. nastoji se na neki način pojačati njihov informativni sadržaj. Ovdje navodimo neke od tih metoda (42,43), koje su često korištene u radovima.

7.1 Poboljšanje kvalitete snimaka

Odnosi se na uporabu raznih efekata na slikovnim podacima kako bi se povećala kvaliteta te bi se snimka na neki način razbistila i bila pregledniji. Tim metodama slike postaju jasnije i ljudskom oku, tj. radiolozima. Na slici 15 vidimo primjer takvog poboljšanja kvalitete snimke, a u nastavku navodimo neke od korištenih tehnika.



*Slika 15. Usporedba originalne snimke (lijevo) i nakon poboljšanja kvalitete (desno) (42).
Prema: Nasser, Akhloufi. Dostupno uz licencu: CC BY 4.0*

Pojačavanje kontrasta označava povećavanje razlike između svjetlijih i tamnijih dijelova snimke i time cijela snimka te neki manji objekti na njoj postaju jasniji i vidljiviji. Provodi se i smanjivanje „buke” na slici tijekom koje se nastoji smanjiti količina podataka koji nisu korisne informacije, a da se te korisne informacije što je moguće više sačuvaju. Raznim tehnikama manipulacije crno-bijelim spektrom se rubovi objekata na slici nastoje prikazati jasnijima. Slike se i provode kroz razne filtre dizajnirane upravo za maksimizaciju kvalitete takvih snimaka.

7.2 Segmentacija

Segmentacijom se prepoznaju dijelovi snimke koji predstavljaju određene objekte, na primjer srce i pluća sa slike 7. Na taj način se omogućava da se na primjer atelektaza ili plućni noduli traže samo na dijelovima snimke gdje se mogu pojaviti, a to su plućna polja. Segmentacijom se zapravo eliminiraju okolna polja koja nepotrebno otežavaju

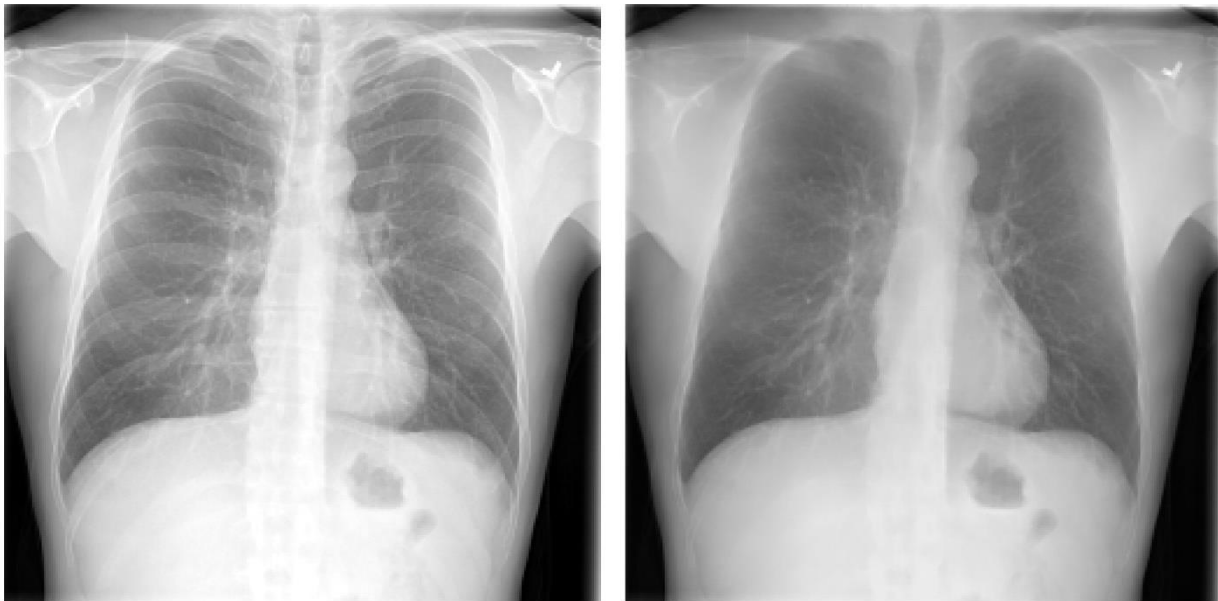
prepoznavanje traženih objekata na slici. Primjer takve segmentacije je prikazan na slici 16.



Slika 16. Prikaz originalne snimke (lijevo), segmentacijske maske pluća (sredina) i snimke nakon segmentacije (desno) (44). Prema: Teixeira et al. Dostupno uz licencu: CC BY 4.0

7.3 Supresija kostiju

Budući da kosti prekrivaju velik dio površine u kojoj se trebaju prepoznati objekti, korisno ih je eliminirati iz snimke različitim metodama. Primjer snimke prije i nakon supresije kosti je prikazan na slici 17.



Slika 17. Usporedba snimke prije (lijevo) i nakon supresije kostiju (desno) (42). Prema: Nasser, Akhloufi. Dostupno uz licencu: CC BY 4.0

8 Modeli za prepoznavanje stanja

Na rendgenskim snimkama se patološka stanja prezentiraju kao zasjenjenja i povećane prozračnosti, a zbog dvodimenzionalnosti snimke se različite pojave često preklapaju. Uz to, postoji značajna varijabilnost između pacijenata. Česta je i istovremena prisutnost različitih patoloških stanja, a sve to čini dijagnostiku i klasifikaciju stanja izazovnom. Danas se razvijaju i testiraju brojni modeli koji bi pomogli liječnicima u dijagnostici i pružili im određenu sigurnost. Konkretnije rečeno, na te modele dubokog učenja ne treba gledati samo kao na metodu koja bi donijela konačnu dijagnozu, već može pružiti značajnu pomoć pri dijagnostici, npr. označavanjem mjesta na snimci koje bi moglo biti pneumonija, a zatim bi se liječnici mogli fokusirati na taj dio i biti sigurniji da im nešto nije promaknulo. Također, analiza rendgenskih snimki prsnog koša predstavlja velik posao na koji se troši puno vremena i resursa. Učinkoviti modeli za očitavanje snimci bi potencijalno značajno smanjili opseg posla i vrijeme koje radiolozi potroše na analizu.

8.1 Pneumonija

Jedno od problema kod analize rendgenskih snimki prsnog koša je prepoznavanje pneumonije. Na rendgenskim snimkama se one tipično prezentiraju kao zasjenjenja, a slična zasjenjenja mogu proizvesti i brojna druga stanja, kao što su karcinom pluća i višak tekućine (45) pa stoga dijagnostika nekada predstavlja velik problem. Pneumonije se s obzirom na radiološki izgled svrstavaju se u četiri tipa: lobarna, lobularna, bronhopneumonija i intersticijska pneumonija. Često su prisutne istovremeno s drugim patološkim stanjima i pokazuju značajnu varijabilnost između pacijenata. Isto tako, obično radiološka slika pneumonije kasni nekoliko dana za kliničkom, što dodatno usporava dijagnosticiranje (46), a kasna dijagnoza usporava početak efikasnog liječenja. S obzirom na poteškoće i dvojbe pri dijagnostici, postoje i velike razlike u dijagnozama različitih radiologa (47–49). Uvođenje automatiziranog modela za dijagnostiku bi elimiralo takvu varijabilnost i omogućilo određenu razinu standardizacije.

Zbog svega navedenog, a i zbog posebne zastupljenosti razvoja modela za prepoznavanje pneumonije, u ovom radu su ti modeli razdvojeni od modela koji prepoznaju i ostale bolesti. U nastavku je navedeno nekoliko modela za prepoznavanje pneumonije te je analizirana njihova uspješnost.

Rajpurkar et al. (49) 2017. su u svrhu detektiranja pneumonije razvili model dubokog učenja nazvan CheXNet. To je zapravo konvolucijska neuralna mreža od 121 sloja, a baziran je na DenseNet-u (29), jednom od ranije spomenutih predtreniranih modela. Rajpurkar et al. su modificirali posljednji sloj originalnog modela tako da kao izlaz vraća samo jednu vrijednost, koju zatim provode kroz još jednu sigmoidnu funkciju (vrstu aktivacijske funkcije). Takav konačni izlaz iz modela predstavlja predviđanje je li na rendgenskoj slici prisutna pneumonija ili nije.

Rajpurkar et al. za treniranje i validaciju CheXNeta koriste bazu rendgenskih slika prsnog koša ChestX-ray14 (33). One snimke koje imaju anotiranu pneumoniju kao prisutnu su označene kao pozitivne, a sve ostale kao negativne (baza originalno sadrži podatke o 14 različitih labela). DenseNet je originalno treniran na bazi slika ImageNet (25), a težine (parametri) mreže nakon tog treniranja su zadržane kao početne težine CheXNeta prije treniranja na ChestX-ray14. Mreža je trenirana u cijelosti (sve težine su bili podešavane), dakle težine početnih slojeva nisu bile zamrznute. U testiranju modela su uspoređivane anotacije četvero radiologa i CheXNet-a na setu od 420 frontalnih rendgenskih slika. Računata je F1 mjera za svakoga od njih, koristeći anotacije ostalih kao „ground truth” (ono što se smatra točnom anotacijom). CheXNet je postigao iznos F1 mjere od 0.435, što je više od prosječne vrijednosti radiologa (0.387).

Iako se model pokazao boljim u klasifikaciji od radiologa, rad ističe nekoliko važnih stvari koje treba imati na umu. Pri testiranju je korištena samo frontalna projekcija, a pretpostavlja se da je za postavljanje 15% dijagnoza potrebna i lateralna. Uz to, radiolozi i CheXNet nisu imali informacije o kliničkoj slici pacijenata, što je inače značajan čimbenik za dobru interpretaciju rendgenskih snimki.

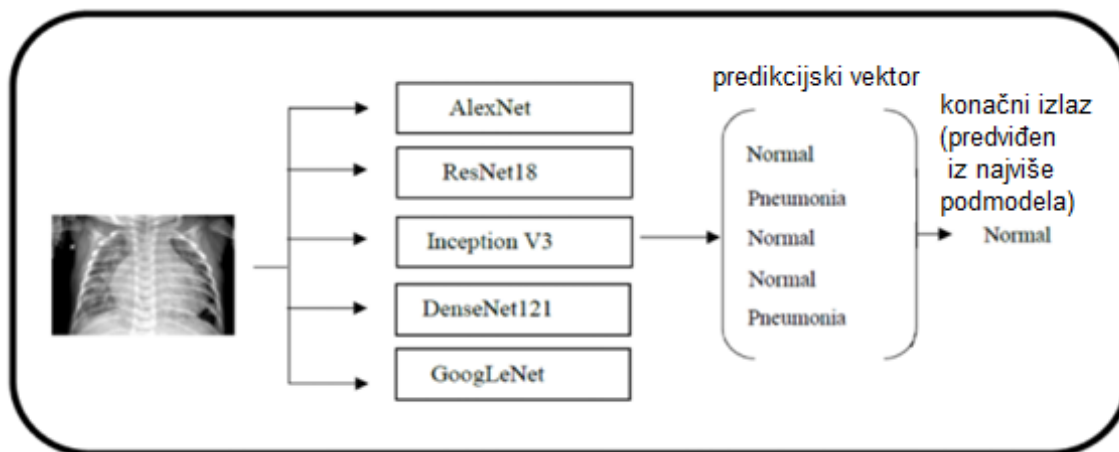
Osim same predikcije je li na snimci pneumonija, CheXNet označava i koji dijelovi snimke su suspektni na pneumoniju, tj. iz kojih dijelova snimke je došlo najviše informacija za predikciju da je pneumonija prisutna. Na slici 18 vidimo kako je srodan model, CheXNeXt (32) ispravno detektirao lokaciju pneumonije u donjem dijelu desnog plućnog režnja.



Slika 18. Ispravno detektirana pneumonija pomoću CheXNeXta (32). Prema: Rajpurkar et al. Uz licencu: CC BY 4.0

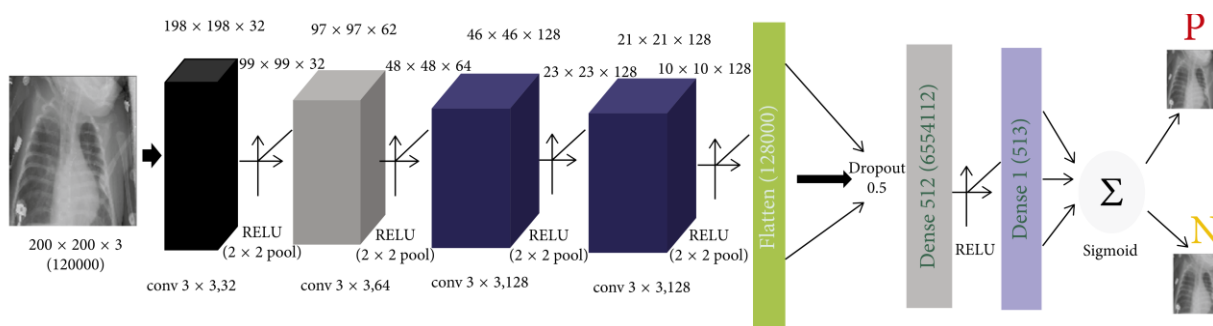
Chouhan et al. (50) 2020. su također dizajnirali model koji binarno predviđa je li na snimci pneumonija ili ne. Treniran je na Chest X-ray bazi snimki, i to koristeći augmentirane podatke. Prvo, dodan je šum u slike te je promijenjena veličina na 224x224x3. Zatim su snimke augmentirane pomoću tri tehnike. Nasumično su horizontalno zrcaljane, nasumično su odrezane („crop“-ane) te je na kraju mijenjan i intenzitet. Sam model je zapravo kombinacija 5 poprilično različitih modela dubokog učenja - predtreniranih modela AlexNet, DenseNet121, ResNet18, Inceptionv3 i GoogLeNet-a. Uz male promjene u arhitekturi svakog od modela, podešeni su da binarno predviđaju prisutnost pneumonije. Navedeni predtrenirani modeli su trenirani na ImageNet (25) bazi, a korišteno je i zamrzavanje težina tih modela u nekim slučajevima. Nakon treniranja, uspoređivana je efikasnost svih 5 modela te modela koji predstavlja njihovu kombinaciju, koji je zapravo zamišljen tako da je konačno predviđanje rezultat većinskog „glasanja” pojedinačnih modela. Slika 19 prikazuje shemu tog modela.

Za evaluaciju je korišten set od 5232 snimki iz „Guangzhou Women and Children's Medical Center”. Najbolje rezultate od pojedinačnih podmodela u smislu točnosti je dobiven s ResNet18 modelom. Rezultat konačnog kombiniranog modela je bio najbolji, uz točnost od 96.39%, osjetljivost 99.62% te AUC (area under the ROC curve) od 99.34%.



Slika 19. Konačni model sastavljen od 5 podmodela iz rada (50).
Prema: Chouhan et al. Uz licencu: CC BY 4.0

Stephen et al. (51) 2019. su krenuli drugačijim pristupom. U uvodu navode da su današnji modeli dubokog učenja vrlo slične arhitekture, a razlog tomu je slučajna uspješnost određenih modela, koji su većinom dizajnirani intuitivno, po principu pokušaja i pogreške. Optimalno podešavanje takvih modela (koristeći neke tehnike kao što su evolucijski algoritmi ili podržano učenje) je izrazito zahtjevno. Kao alternativu predlažu jednostavniju, ali efikasnu arhitekturu dizajniranu iz početka. Konačan model njihove konvolucijske neuronske mreže je prikazan na slici 20. Za ulaz je odabrana veličina slike od $200 \times 200 \times 3$ jer je producirala najbolju točnost na setu za validaciju. Korištena je baza snimki (52). Točnost modela na setu za treniranje je iznosila 0.9531, a na setu za validaciju 0.9373.



Slika 20. Arhitektura konvolucijske neuronske mreže iz rada (51). Vidi se slaganje konvolucijskih slojeva („conv“) te potpunih linearnih („Dense“) sve do konačne predikcije postoji li pneumonija (P) ili ne (N); RELU, Sigmoid – aktivacijske funkcije; conv – konvolucijski sloj; Dense – potpuni linearni sloj; Dropout – dropout sloj; Flatten – prebacuje podatke u vektorski (1d) oblik;
Prema: Okeke Stephen et al. Uz licencu: CC BY 4.0

Irfan et al. (53) 2020. argumentiraju kako postoje različite spoznaje o tome je li korištenje predtreniranih modela znatno učinkovitije od randomizirane inicijalizacije težina modela prije treniranja (54,55). Zbog toga se odlučuju testirati oba pristupa koristeći popularne predtrenirane modele ResNet-50, Inceptionv3 i DenseNet121. Za treniranje i validaciju modela je korišteno 86524 snimki iz ChestX-ray14 baze. Modeli su testirani na dva različita seta. Prvi se sastoji od ostatka snimki iz ChestX-ray14 baze, a drugi od seta za validaciju iz baze CheXpert (na kojima su snimke anotirali radiolozi). Testirano je 6 različitih slučajeva. Svaki od navedenih predtreniranih modela je prvo korišten tako da su težine (parametri) inicijalizirane na one nakon treniranja na ImageNet (25) bazi slika („Configuration B”), a zatim je svaki od njih testiran s nasumično odabranim težinama („Configuration A”), ali je arhitektura ostala ista. Na slici 21 vidimo usporedbu vrijednosti AUC za testirane modele. Modeli s težinama inicijaliziranim na one nakon treniranja na ImageNet-u su postigli značajno bolje rezultate, s poboljšanjem vrijednosti AUC između 4.1% i 52.5%.

baza snimki	eksperimenti		
	Model	Konfiguracija A	Konfiguracija B
1: ChestX-ray14 [4]	ResNet-50 [18]	0.46	0.59
	Inception V3 [19]	0.51	0.55
	DenseNet121 [20]	0.57	0.71
2: CheXpert [5]	ResNet-50 [18]	0.58	0.69
	Inception V3 [19]	0.40	0.61
	DenseNet121 [20]	0.73	0.76

Slika 21. Usporedba modela s nasumično inicijaliziranim ("Konfiguracija A") i predtreniranim ("Konfiguracija B") težinama iz rada (53). Vidi se da su modeli s predtreniranim težinama postigli bolji rezultat. Prema Irfan et al. Uz licencu: CC BY 4.0

Sharma et al. (56) 2020. primjećuju kako većina radova vezanih za temu koristi metode za poboljšavanje generalizacije modela u obliku augmentacije snimki te korištenja „dropout” slojeva. Generalizacija predstavlja sposobnost modela da ispravno predviđa na podatcima na kojima nije treniran. Augmentacijom snimki radimo male promjene na podatcima za treniranje koje ne narušavaju važne karakteristike snimke, a služe svrsi povećavanja baze podataka na kojoj se trenira (što je bitno za generalizaciju). Korištenjem „dropout” slojeva se pokušava spriječiti da se model „prenauci” na setu za treniranje razbijanjem dosad uspostavljenih veza između neurona. Dropout-om se privremeno izbacuju doprinosi nekih neurona prema sljedećem sloju. Sharma et al. u radu koriste konvolucijsku neuronsku mrežu i testiraju njenu efikasnost bez i uz korištenje augmentacije podataka i dropouta. Najboljim se pokazao model s augmentacijom i dropoutom, a najgorim onaj bez obje tehnike.

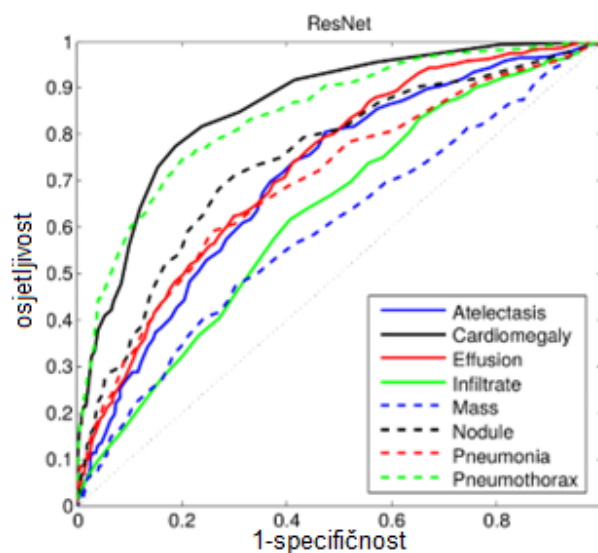
Ayan i Ünver (57) 2019. istražili su efikasnost dva popularna predtrenirana modela konvolucijske neuronske mreže u dijagnostici pneumonije – Xception (58) i VGG16 (26). Za treniranje, validaciju i testiranje je korištena baza (52). VGG16 model se pokazao boljim u smislu točnosti (0.87%), specifičnosti (0.91%) i preciznosti u detekciji pneumonije

(0.91%), dok je Xception model bio učinkovitiji u smislu osjetljivosti (0.85%) i preciznosti u detekciji normalnih snimki (0.86%).

8.2 Različita stanja

Ultimativni cilj uporabe dubokog učenja u ovoj domeni bi, naravno, bila detekcija svih stanja koja se mogu raspoznati na rendgenskim slikama prsnog koša. Ovdje razmatramo modele koji streme k tom cilju te su namijenjeni prepoznavanju većeg broja različitih stanja.

Wang et al. (34) 2017. se u svom radu bave detekcijom 8 različitih stanja, a to je upravo rad u kojem je pri put prezentirana baza rendgenskih slika ChestX-ray8 (8 različitih labela u početku, a kasnije je nadograđena da sadrži 14 labela i poznata je pod nazivom ChestX-ray14 (33)). Objašnjavaju kako su labela snimki prepoznate automatski iz radioloških izvještaja koristeći alate dnorm (59) i MetaMap (60), a zatim su obrađene da se isključi ukoliko radiolog zapravo negira ili nije siguran u njihovo postojanje. Dodatno, baza sadrži podatke o lokalizaciji 200 instanci svakog od 8 patoloških stanja, što obuhvaća 983 snimke. Svrha toga je mogućnost testiranja sposobnosti modela da lokalizira stanja na samoj snimci.



Slika 22. ROC krivulje modela ResNet-50 za različite labela (34). Prema: Wang et al. © 2017 IEEE

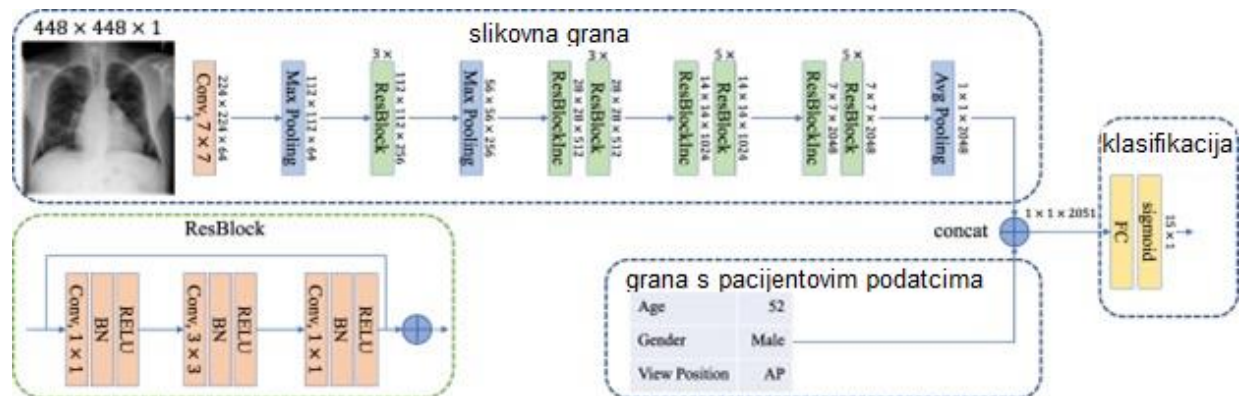
Srž modela testiranih u radu su predtrenirani modeli AlexNet, GoogLeNet, VGG16 i ResNet-50. Na te je dodano još nekoliko slojeva kako bi izlaz sadržavao informaciju o prisutnosti ili odsutnosti svake od 8 stanja. Uspoređivane su AUC vrijednosti različitih modela pri detekciji svih stanja. Kao najbolji model se pokazao ResNet-50. Na slici 22 su prikazane njegove ROC krivulje za svaku od labela. Učinkovitost detekcije jako varira između različitih stanja, pa se tako model pokazao efikasan u detekciji kardiomegalije (AUC=0.8141) i pneumotoraksa (AUC=0.7891), a lošiji u detekciji manjih objekata, kakvi bi, na primjer, pripadali labelama „mass” (AUC=0.5609) i „nodule”. Detekcija pneumonije također nije efikasna, a to je objašnjeno njenom slabijom zastupljenošću u setu za treniranje (<1%).

Yao et al. (61) 2018. argumentiraju kako Wang et al. (34) u svome radu pretpostavljaju da su labela neovisne, dok u stvarnosti one to nisu. Zbog toga pri dizajniranju modela pokušavaju modelirati te ovisnosti. Uz to, svoje modele u ovom radu treniraju iz početka, jer smatraju da su predtrenirani modeli previše prilagođeni prirodnim slikama i kao takvi

su manje povoljni za obradu informacija s radioloških snimaka. Arhitektura njihovog modela je ipak bazirana na DenseNet-u. Baza podataka korištena za treniranje i testiranje je ChestX-ray14, što je nadograđena baza iz rada Wang et al. (34). U usporedbi s tim radom, ovaj model je postigao bolje rezultate predikcije svih labela, osim fibroze.

Rajpurkar et al. (49) 2017. u svom radu predstavljaju CheXNet, već spomenuti model za prepoznavanje pneumonije, ali idu još dalje pa taj model prilagođavaju prepoznavanju stanja za svih 14 labela iz seta ChestX-ray14. Već u istom radu navode kako njihov model pokazuje najbolje rezultate do sada u prepoznavanju svih labela, nadmašujući one od modela Yao et al. (61) i Wang et al. (34). U idućem radu, Rajpurkar et al. (32) 2018. iznose detaljne rezultate o efikasnosti njihovog modela u prepoznavanju svih 14 stanja. Te rezultate zatim uspoređuju s rezultatima radiologa na istom setu, koristeći vrijednost AUC-a kao parametar usporedbe. Za 10 od 14 stanja, nije bilo statističke razlike u učinku. Model je postigao statistički značajno bolji učinak u detekciji atelektaze, postignuvši AUC od 0.862, dok je AUC radiologa iznosio 0.808. Radiolozi su značajno bolji učinak postigli u detekciji kardiomegalije, emfizema i hijatalne hernije, postignuvši AUC vrijednosti 0.888, 0.911 i 0.985, istim redoslijedom (model je postigao vrijednosti: 0.831, 0.704 i 0.851). Pretpostavlja se da je razlog tome niska zastupljenost tih snimki u setu za treniranje, ali s druge strane je model bio efikasan u prepoznavanju pneumonije, fibroze i edema, a te snimke su također bile slabo zastupljene. Radiolozima je za anotiranje svih 420 validacijskih snimki trebalo prosječno 240 minuta, a modelu 1.5 minuta. Dodatnih 40 sekundi je model trebao za označavanje dijelova snimke koji sugeriraju da postoje određena patološka stanja.

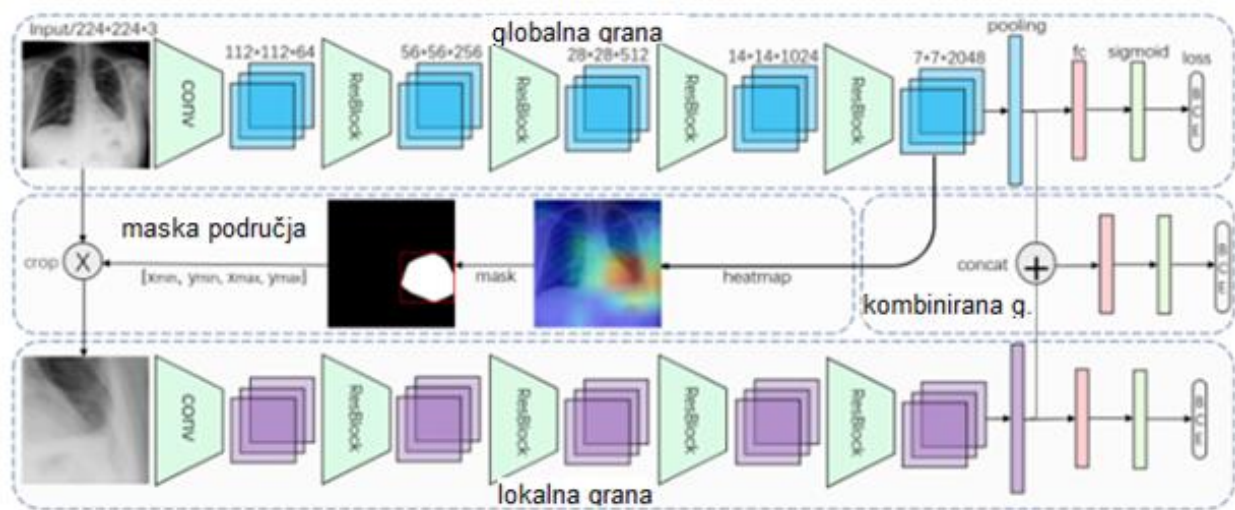
Baltruschat et al. (62) 2019. koriste pristup dodavanja podataka koji nisu u slikovnom formatu kao ulaze u neuronsku mrežu. Konkretno, dodaju dob, spol i vrstu projekcije kao ulaz s namjerom da te informacije pomognu modelu pri donošenju odluke. Arhitektura njihovog modela je prikazana na slici 23.



Slika 23. Arhitektura modela iz rada Baltruschat et al. (62); Conv – konvolucijski sloj; BN – „batch normalization“ / normalizacijski sloj; RELU, sigmoid – aktiv. funkcija.; Max/Avg Pooling – pooling sloj koji zadržava najveću / prosječnu vrijednost; Prema: Baltruschat et al. Uz licencu: CC BY 4.0

Vidi se da pri izgradnji modela koriste dvije podmreže – jednu za analizu snimke („slikovna grana“), a druga za analizu neslikovnih podataka („grana s pacijentovim podacima“). Korištenje tih neslikovnih podataka nije značajno doprinijelo poboljšanju modela, a rad je pokazao i koji je vjerojatan razlog. Pri tome su korištena 3 modela bazirana na modelu ResNet-50, a svi kao ulaz primaju rendgenske snimke prsnog koša. Jedan se koristi za predikciju dobi pacijenta, drugi spola, a treći za predikciju vrste projekcije (AP ili PA). Ti modeli su bili poprilično efikasni. Predviđanje vrste projekcije je postiglo iznos AUC-a od 0.9983 ± 0.0002 , predviđanje spola iznos AUC-a od 0.9345 ± 0.0067 , a prosječna greška u predviđanju dobi je bila 9.13 ± 7.05 . Iz toga se da zaključiti kako same snimke u sebi već sadrže dovoljno informacija od dobi i spolu pacijenta te o vrsti projekcije i nema potrebe da se te informacije još dodatno uključuju kao ulaz.

Guan et al. (63) 2018. naglašavaju problem kod detekcije manjih lezija čija je lokacija na snimkama nepredvidljiva, kao što je na primjer atelektaza. Stoga smatraju da ostatak snimke (dio koji ne uključuje leziju) stvara informacijski šum te otežava prepoznavanje takvih manjih lezija. Zbog toga svoj model dijele na tri grane – jednu za globalni kontekst, drugu za lokalni i treću koja kombinira obje informacije. Arhitektura modela je prikazana na slici 24.



Slika 24. Arhitektura modela iz rada Guan et al. (63) - prikazane su tri grane (globalna, lokalna, kombinirana); fc – „fully connected layer“, potpuno povezani linearni sloj („dense“); BCE – „binary cross entropy“ funkcija gubitka (loss function) koju minimizira model; Prema: Guan et al. Reproducirano uz dopuštenje: Elsevier

Model je ostvario najveći prosječan AUC (0.871) pri klasifikaciji snimki na ChestX-ray14 bazi do sada, nadmašujući onaj CheXNet-ov (49).

Irvin et al. (35) 2019. predstavljaju CheXpert bazu rendgenskih slika te objašnjavaju proces ekstrakcije anotacija iz radioloških izvještaja. Uz to, razvijaju i model za predikciju stanja baziran na mreži DenseNet121, koja se pokazala najboljom od testiranih predtreniranih modela. Pri iznošenju rezultata, fokus je stavljen na 5 labela koje su

procijenjene kao klinički najbitnije i dovoljno zastupljene u setu za validaciju. Uspoređivana je efikasnost modela i tri radiologa. Model se pokazao najboljim u predikciji pleuralnog izljeva (AUC=0.97), a najgori u predikciji atelektaze (0.85). U usporedbi s pojedinačnim radiolozima, model se pokazao bolji u detekciji kardiomegalije, edema i pleuralnog izljeva, ali ne bolji od većinskog glasovanja trojice radiologa. U detekciji konsolidacije, model je imao bolji učinak od 2/3 radiologa, a sva tri radiologa su bila bolja u detekciji atelektaze.

Yuan et al. (64) 2021. koriste metodu maksimizacije AUC-a (DAM – Deep AUC Maximization) pri izgradnji svog modela i s njime pobjeđuju na Stanfordovom natjecanju za najbolji model za predikciju 5 labela (kardiomegalija, edem, konsolidacija, atelektaza, pleuralni izljev) na CheXpert bazi snimaka. Taj model je nazvan DeepAUC-v1, a zapravo je kombinacija modela DenseNet121, DenseNet161, DenseNet169, DenseNet201 i Inception-ResNet-v2. Prosječan AUC postignut za 5 odabranih labela je 0.935.

Blais i Akhloufi (65) 2021. su također izgradili model za predikciju labela na CheXpert bazi snimki, a sklopljen je od arhitektura DenseNet-121,169,201, InceptionResNetV2, InceptionV3, MobileNet, MobileNetV2, ResNet101, ResNet101V2, ResNet152, ResNet152V2, ResNet50, ResNet50V2, VGG16, VGG19 i Xception. Takav model je postigao najveći prosječni AUC (0.9676) na 5 labela (kardiomegalija, konsolidacija, edem, pleuralni izljev i atelektaza). Dodatno, model je testiran za detekciju svih 14 labela i postigao je prosječan AUC od 0.949.

9 Primjena odabranog modela

9.1 Općenito

U ovom dijelu bit će objašnjen konkretan način na koji se odabrani model dubokog učenja koristi kao pomoć pri dijagnostici rendgenskih snimaka prsnog koša. Opisat ćemo kako od same rendgenske snimke uporabom odabranog modela dobiti izlazne vrijednosti (predviđanja) i što te izlazne vrijednosti predstavljaju i znače.

9.2 Odabrani model

Model koji ćemo testirati je ekvivalentan modelu CheXNet iz rada Rajpurkar et al. (49). Centralni dio modela čini konvolucijska neuronska mreža od 121 sloja, bazirana na DenseNet-u. Na internetu je objavljeno više implementacija iste arhitekture (66).

Odabrana implementacija je slobodno dostupna na github-u (67). Na istom mjestu je dostupan i već gotov model, treniran na istoj bazi rendgenskih snimki (ChestX-ray14) kao i izvorni CheXNet model.

Minimalna je razlika u arhitekturi originalnog i reimplementiranog modela. Originalno su težine neuronske mreže podešavane pomoću „Adam“ optimizatora (vrsta algoritma za podešavanje težina modela pri „backpropagation“ algoritmu), a u reimplementiranom modelu je korišten „SGD+momentum“ jer se pokazalo da producira nešto bolje rezultate. Postoje i drugi radovi koji su naveli isto zapažanje (68).

Model je treniran da na snimci prepoznaje 14 različitih stanja. To zapravo znači da snimkama određuje vjerojatnost pripadnosti svakoj od 14 dostupnih labela (atelectasis, cardiomegaly, consolidation, edema, effusion, emphysema, fibrosis, hernia, infiltration, mass, nodule, pleural_thickening, pneumonia, pneumothorax).

Osim predviđanja vjerojatnosti pripadnosti labelama, korištenjem ovog modela je moguće i prikazati koji dijelovi snimke su bili ključni pri donošenju odluke o prisutnosti patološkog stanja. Najčešće se to prikazuje grafički u obliku toplinske karte („heatmap“).

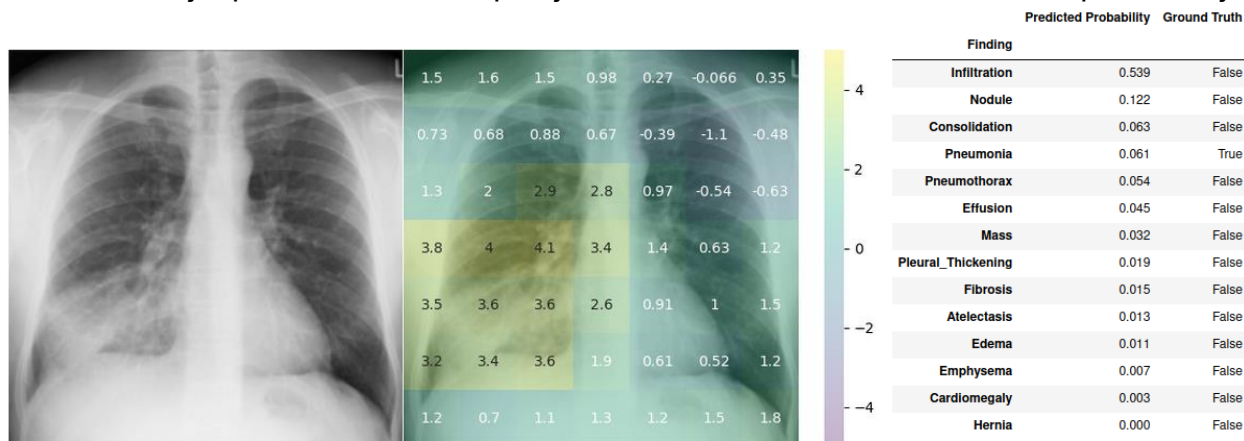
Svako novo treniranje modela dubokog učenja, očekivano, producira modele koji daju malo različite rezultate. Zbog toga (a i zbog korištenja različitog optimizatora) se AUC vrijednosti dobivene ovim gotovim modelom nešto razlikuju od rezultata dobivenih u originalnom radu.

9.3 Primjeri uporabe modela

Na odabranim rendgenskim snimkama primijenit ćemo model te prikazati rezultate. Objasniti ćemo što ti rezultati zapravo znače te kako se interpretiraju i uspoređuju da bi se dobilo konačno predviđanje.

9.3.1 Pneumonija

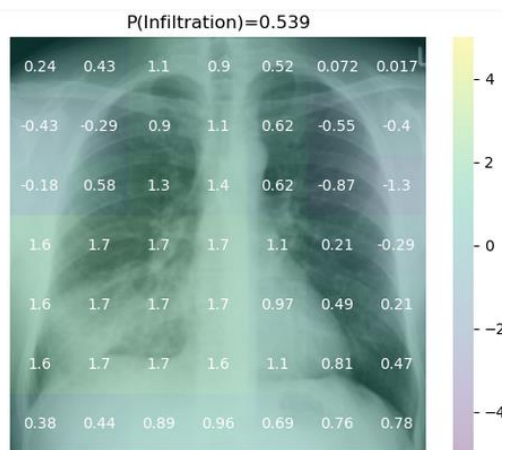
Na slici 25 je prikazan rezultat primjene modela na odabranu snimku pneumonije.



Slika 25. Originalna snimka pneumonije (lijevo) – Ljubaznošću: Sajoscha A Sorrentino, radiopaedia.org, uz licencu CC BY-NC-SA 3.0 (69); značaj pojedinih područja za labelu "Pneumonia" (sredina); tablica svih izlaznih vjerojatnosti predviđenih modelom (desno).

U tablici na slici 25 (desno), „Ground Truth“ predstavlja nešto za što se smatra da je nepobitna istina za tu snimku. Te vrijednosti su postavljene prema očitanjima radiologa već prije provlačenja snimke kroz model, a ovdje su prikazane samo da bi se mogle usporediti s izlazima iz modela („Predicted Probability“). Za ovu snimku je postavljeno samo „Pneumonia“ kao „True“, budući da ne utječe na analizu i rezultat, ali adekvatno bi bilo i postaviti kao pozitivne i ostale preklapajuće pojmove, poput „Infiltration“ i „Consolidation“.

Iz rendgenske snimke na slici 25 (lijevo), model je predvidio određene vjerojatnosti za svako stanje (labelu) za koju je treniran. Te vjerojatnosti su prikazane na tablici na slici 25 (desno) kao „Predicted Probability“. Uz to, na slici 25 (sredina) prikazano je kako su pojedini dijelovi slike utjecali na predikciju vjerojatnosti labele „Pneumonia“ u obliku toplinske mape („heatmap“). Vidimo se da su, u ovom slučaju ispravno, veće vrijednosti u području snimke gdje se stvarno nalazi pneumonija. Isto tako, može se prikazati i kako su pojedini dijelovi slike utjecali na određivanje vjerojatnosti ostalih labele. Na slici 26 je prikazan primjer utjecaja na vjerojatnost labele „Infiltration“ za istu rendgensku snimku. Vidi se da je infiltracija također dobro prepoznata u istom dijelu snimke.

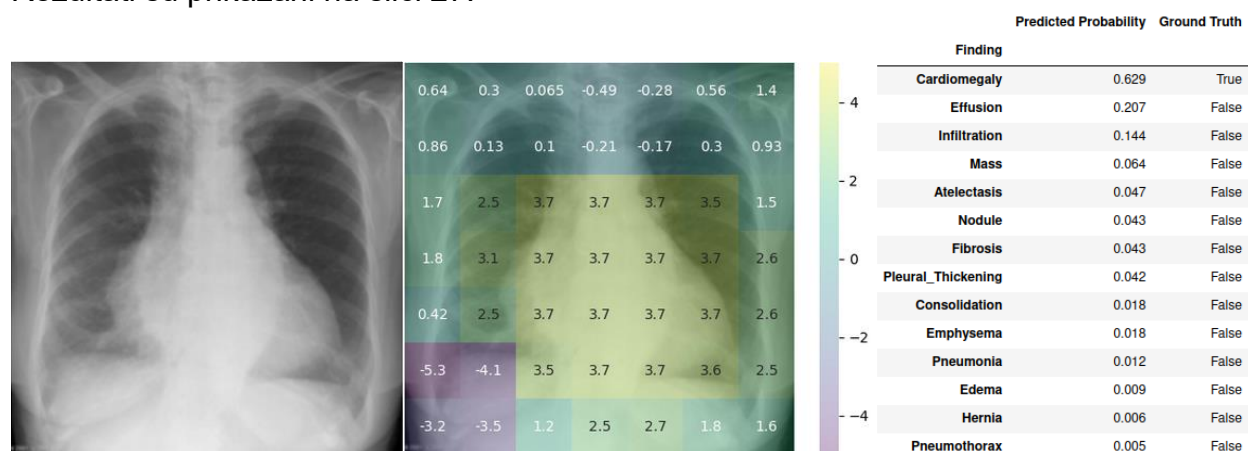


Slika 26. Utjecaj dijelova snimke na predviđenu vjerojatnost za labelu "Infiltration".

Vratimo se sada na analizu izlazne vjerojatnosti („Predicted Probability“ u tablici desno na slici 25) koje model predvidi za svaku labelu. Sama po sebi, ta vjerojatnost nema preveliko značenje, već se mora uspoređivati odnos vjerojatnosti dobivene za različite snimke. To se izvodi na način da se odredi iznos vjerojatnosti koji bi bio granica („threshold“) između toga da se nešto proglašuje pneumonijom (1) ili bez pneumonije (0). Međutim, teško je definirati koji bi bio optimalan iznos te granice, budući da je u nekim slučajevima važnija dobra osjetljivost, a u drugima specifičnost. Upravo zbog toga se efikasnost modela prikazuje ROC krivuljom (objašnjena u 3.7.). Postupak će biti detaljnije objašnjen kasnije.

9.3.2 Kardiomegalija

Kao sljedeći primjer je uzeta snimka kardiomegalije (70) te je i ona provučena kroz model. Rezultati su prikazani na slici 27.

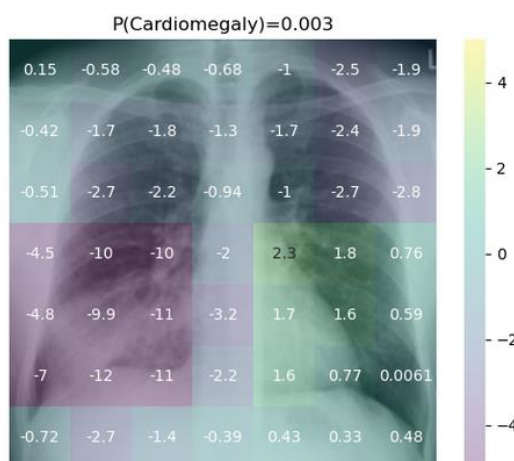


Slika 27. Originalna snimka kardiomegalije (lijevo) – Ljubuznošću: Mohammadtaghi Niknejad, uz licencu CC BY-NC-SA 3.0 (70); značaj pojedinih područja za labelu "Cardiomegaly" (sredina); tablica svih izlaznih vjerojatnosti predviđenih modelom (desno).

U ovom slučaju je također kao „True“ postavljena samo jedna labela - „Cardiomegaly“, ali bilo bi u redu postaviti i npr. „Effusion“.

Ovoga puta, (u sredini slike 27) kao „heatmap“ je prikazan utjecaj pojedinih dijelova snimke na predviđenu vjerojatnost kardiomegalije, koja u ovom slučaju iznosi 0.629. Vidi se da je područje snimke na kojoj se nalazi povećano srce imalo znatan učinak na povećanje predviđene vjerojatnosti labela „Cardiomegaly“.

Za usporedbu, prikazan je i značaj pojedinih dijelova prethodno analizirane snimke pneumonije (slika 25 - lijevo) na za nju predviđenu vjerojatnost labela „Cardiomegaly“, koja je

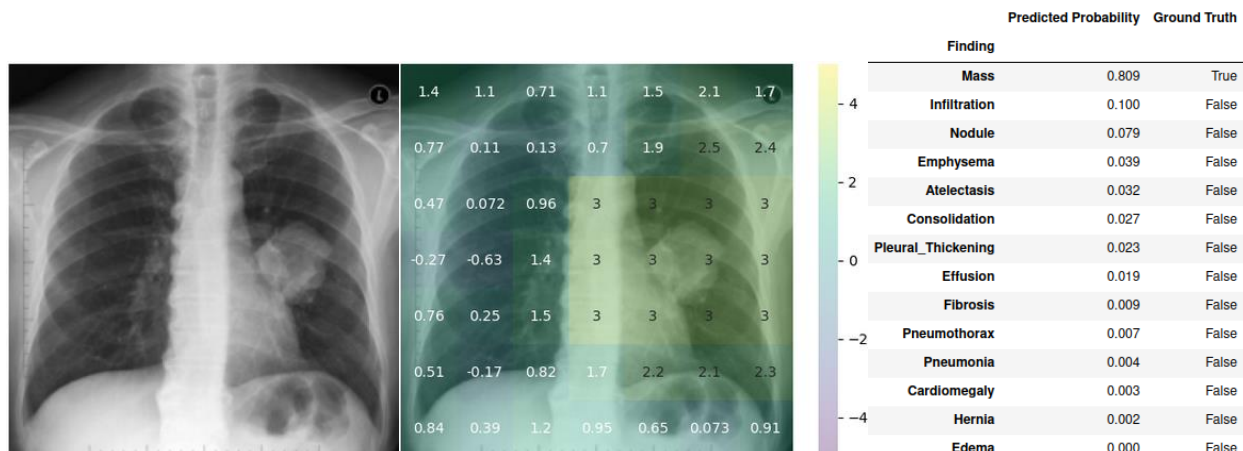


Slika 28. Utjecaj dijelova snimke pneumonije (sa slike 25 - lijevo) na za nju predviđenu vjerojatnost labela "Cardiomegaly".

iznosila 0.003. Vidi se da su prikazane vrijednosti puno manje od onih na slici 27. U desnom plućnom krilu su one i ekstremno negativne, što se dijelom može objasniti time da model očekuje detektirati srce i na desnoj strani ukoliko je kardiomegalija prisutna, a manjak takve detekcije stoga postaje jak signal za predviđanje kardiomegalije kao manje vjerojatnom, tj. konačno predviđen iznos vjerojatnosti kardiomegalije je znatno manji.

9.3.3 Tumor pluća

Iduće će biti procijenjena snimka koja sadrži tumor pluća. Rezultati su prikazani na slici 28.



Slika 29. Originalna snimka tumora pluća (lijevo) – Ljubaznošću: Frank Gaillard, uz licencu: CC BY-NC-SA 3.0 (71); značaj pojedinih područja za labelu "Mass" (sredina); tablica svih izlaznih vjerojatnosti predviđenih modelom (desno).

Tumor pluća prezentirao se kao masa pa je stoga i labela „Mass“ postavljena kao „Ground truth“.

Iz toplinske karte (slika 29 - sredina) je jasno da je signal za povišenje vjerojatnosti prisutnosti mase došao upravo iz dijela snimke na kojoj je masa prisutna.

9.4 Usporedba rezultata prethodnih primjera

Budući da „predviđene vjerojatnosti“ nemaju preveliko značenje same po sebi, nužno je bilo prvo iznijeti rezultate primjene modela na različitim snimkama. Samo tada se može vidjeti značaj tih predviđenih vrijednosti. Usporedit će se rezultati predviđanja za tri snimke dosad prikazane (pneumonija (slika 25), kardiomegalija (slika 27) i tumor pluća (slika 29)). Zbog jednostavnosti, uzimat će se samo one labela relevantne za usporedbu. Za bolju usporedbu, tablice izlaza iz modela (desno na slikama 25, 27 i 29) su spojene u jednu i rezultati su prikazani kao tablica 1.

Tablica 1. Najznačajniji rezultati predviđanja modela na prethodnim primjerima pneumonije (slika 25), kardiomegalije (slika 27) i tumora pluća (slika 29).

	Pneumonija (slika 25)	Kardiomegalija (slika 27)	Tumor pluća (slika 29)
Finding	Predicted Probability	Predicted Probability	Predicted Probability
Pneumonia	0.061	0.012	0.004
Infiltration	0.539	0.144	0.100
Consolidation	0.063	0.018	0.027
Cardiomegaly	0.003	0.629	0.003
Effusion	0.045	0.207	0.019
Mass	0.032	0.064	0.809

Prije svega, očito je da ove predviđene vjerojatnosti nisu povezive s onim kako bi ih čovjek (radiolog) izražavao za pojedino stanje. Vjerojatnost kardiomegalije od 62.9% ili tumora pluća od 80.9% se čini mala s obzirom da bi ih radiolog na istim snimkama dijagnosticirao s vjerojatnošću bližoj 100%. Još je ekstremniji slučaj pneumonije, gdje je ona predviđena s vjerojatnošću od 6.1%.

Ipak, vidljivo je da su iznosi predviđenih vjerojatnosti za određena stanja znatno veći u slučajevima gdje je to stanje stvarno prisutno.

Vrijednosti predviđenih vjerojatnosti labela svih navedenih patoloških stanja koja se preklapaju s pneumonijom su veća kod snimke koja prikazuje pneumoniju od ostalih snimki.

Kod snimke koja prikazuje kardiomegaliju, predviđena vjerojatnost kardiomegalije (0.629) je znatno veća od one na ostalim snimkama (0.003). Vidi se i veća vjerojatnost postojanja izljeva, koji često prati kardiomegaliju, a prisutan je i na ovoj snimci.

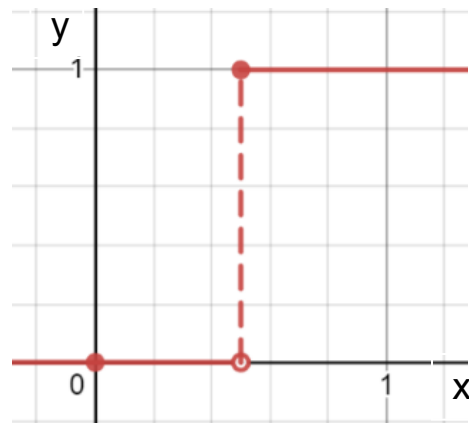
Predviđena vjerojatnost za labelu „Mass“ na snimci tumora pluća je također dosta veća od onih na ostalim snimkama.

9.5 Značenje predviđenih vjerojatnosti

Sukladno navedenom, da bi se moglo zaključivati iz predviđenih vjerojatnosti, moraju se uspoređivati odnosi vjerojatnosti iste labele dobivene za različite snimke. To se izvodi na način da se odredi iznos vjerojatnosti koji bi bio prag („threshold“) između toga da se nešto proglaši prisutnim (1) ili odsutnim (0) patološkim stanjem.

Uzmu li se za primjer vjerojatnosti pneumonije iz tablice 1, može se reći da bi se prag vjerojatno trebao nalaziti između iznosa vjerojatnosti od 0.012 i 0.061.

Ekvivalentna takvom određivanju praga je „step“ funkcija, kakva se može koristiti kao aktivacijska funkcija u modelu dubokog učenja. Ona je mogla biti implementirana i u samom modelu pa bi on vraćao samo vrijednosti vjerojatnosti 0 i 1, ali optimalna vrijednost praga nije jednoznačna i bilo bi teže napraviti stvarnu procjenu efikasnosti modela, a u nastavku će biti objašnjeno zašto. Na slici 30 je prikazan primjer „step“ funkcije kakva bi se mogla koristiti za binarno određivanje prisutnosti pneumonije u ovom slučaju.



Slika 30. Step funkcija koja bi se mogla koristiti za binarno predviđanje je li pneumonija u ovom slučaju prisutna (1) ili ne (0). Na „x“ osi su izlazne vjerojatnosti iz modela, a na „y“ binarno određena prisutnost pneumonije.

Definirati optimalan iznos praga je teško jer je u nekim slučajevima važnija dobra osjetljivost, a u drugima specifičnost. Smanjivanjem iznosa praga, podiže se osjetljivost i smanjuje specifičnost, a podizanjem praga smanjuje se osjetljivost i povećava specifičnost. Moguće je i, naravno, postaviti prag tako da model ima najveću moguću točnost.

9.6 Procjena efektivnosti modela

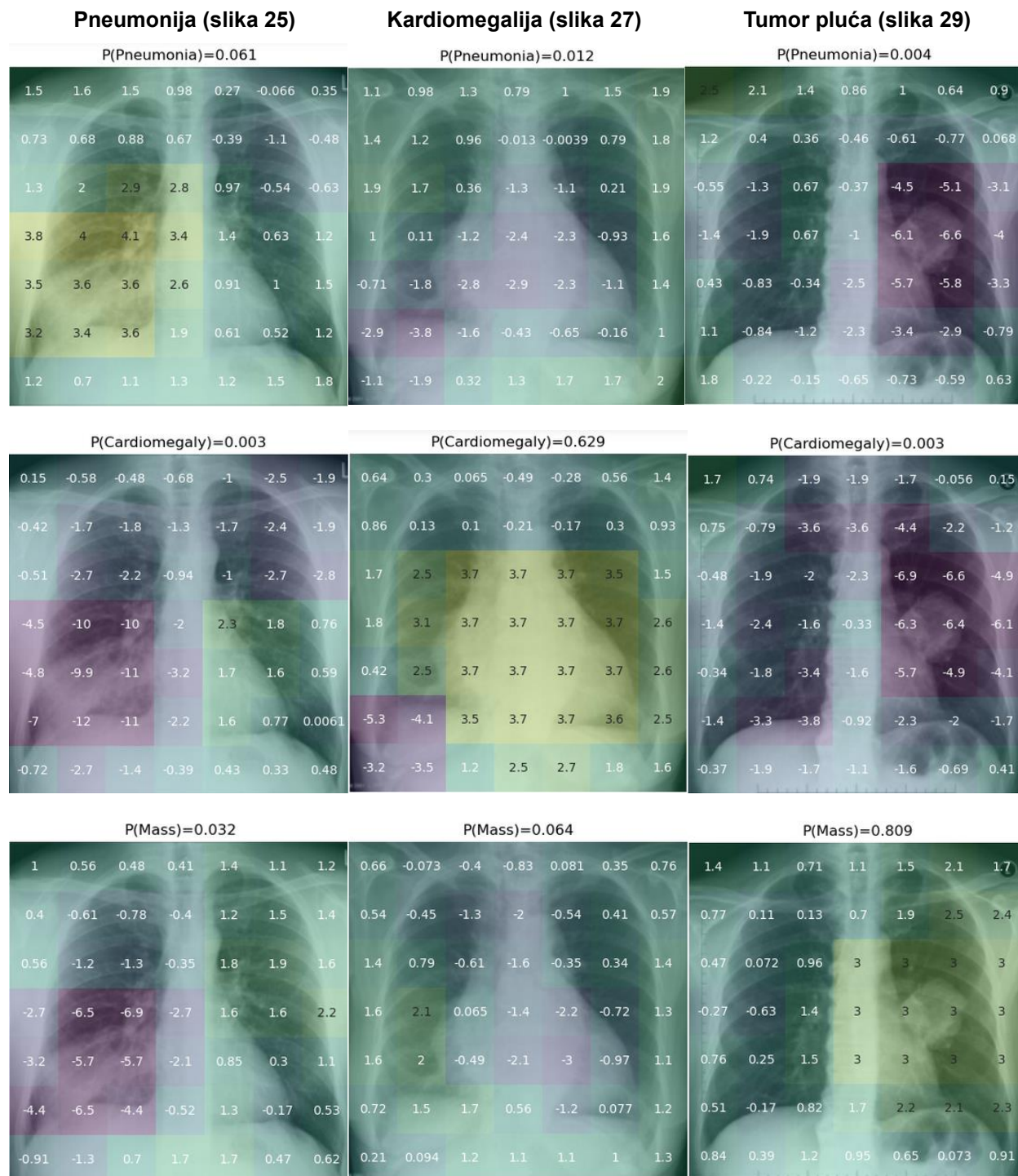
Zbog problema odabira optimalnog praga te potrebe za usporedbom različitih modela, efektivnosti modela se najčešće prikazuju ROC krivuljom (objašnjena u 3.7.) i uz nju vezanom AUC vrijednošću.

ROC krivulja zapravo uzima u obzir sve moguće pragove koje je moguće odabrati, a time i sve moguće stupnjeve osjetljivosti i specifičnosti koji se mogu postaviti.

Vezano s tim, da bi se krivulja prikazala, mora se izračunati TPR (osjetljivost) i FPR (1-specifičnost) za svaki mogući iznos vjerojatnosti koji možemo uzeti kao prag.

Primjerice, testira se model za predviđanje pneumonije na određenom broju snimki koja svaka ima određen „Ground Truth“ postoji li pneumonija na njoj (1) ili ne (0). Svaka od snimki se provuče kroz model te se predvide vjerojatnosti pneumonije. Uzme li se, na primjer, da je najmanja vrijednost predviđene vjerojatnosti 0.005, a najveća 0.150, za moguće pragove treba uzeti sve moguće vrijednosti između 0.005 i 0.150. Zatim se za svaki od tih pragova izračunava koliku osjetljivost i specifičnost ima model te se ta točka upisuje u koordinatni sustav koji na „x“ osi ima 1-specifičnost, a na „y“ osi osjetljivost. Crtanjem krivulje koja povezuje te točke dobivamo konačnu ROC krivulju, a izračunavanjem površine ispod krivulje dobivamo AUC vrijednost kao konačnu mjeru efektivnosti modela koja je adekvantna za uspoređivanje s onom od drugih, sličnih modela.

9.7 Pozadina predviđenih vjerojatnosti



Slika 31. Toplinske karte za primjere snimki pneumonije (slika 25), kardiomegalije (slika 27) i tumora pluća (slika 29). Iznad svake toplinske mape je označena labela za koju je prikazan utjecaj dijelova slike te predviđena vjerojatnost za tu labelu.

Iz prethodno spomenutih istraživanja i primjera je jasno da modeli mogu efikasno prepoznavati stanja na rendgenskim snimkama, u nekim slučajevima i na razini radiologa. Međutim, ne može se odmah pretpostaviti da oni to rade na potpuno medicinski legitiman način. U ovom ulomku će se nešto detaljnije prikazati kako značajke snimki utječu na predviđanje modela te neki potencijalni problemi.

Za početak, na slici 31 su prikazane toplinske mape snimki i labela iz prethodnih primjera.

Na njoj se vidi, a i već je spomenuto u ulomku 7.3, da su ispravni dijelovi svih snimki bili značajni za ispravno predviđanje vjerojatnosti odgovarajućih labela pneumonije („Pneumonia“), kardiomegalije („Cardiomegaly“) i tumora pluća („Mass“).

Međutim, pregledaju li se ostale toplinske karte na slici 31 (one gdje je prikazan utjecaj dijelova snimke na vjerojatnosti labela koje na snimci nisu prisutne), vidi se mnogo negativnih vrijednosti označenih ljubičastom bojom. Te negativne vrijednosti se nalaze upravo u područjima na kojima se nalazi neko drugo patološko stanje, a ne ono za čiju se labelu prikazuje toplinska karta. Može se zaključiti da je prisutnost nekog drugog patološkog stanja smanjila predviđenu vjerojatnost za neku labelu.

Takvo „zaključivanje“, u kojem bi prisutnost jedne bolesti isključilo istovremenu prisutnosti druge sigurno nije medicinski legitimno, a ovaj fenomen se događa jer uglavnom na napravljenim snimkama na kojima je model treniran i postoji samo jedno stanje, odnosno više stanja koja često idu skupa (kardiomegalija i izljev) ili se pojmovi preklapaju (pneumonija, infiltracija).

O sličnim problemima piše i autor odabrane implementacije modela dubokog učenja koju smo koristili, John Zech. U svom članku (72) spominje još neke od elemenata snimki koje mogu navesti model da medicinski nelegitimno „zaključuje“. Kao primjer navodi mogućnost modela da prepozna na kojem rendgenskom aparatu je izvedeno snimanje. Ukoliko na snimci u gornjem desnom kutu piše „portable“, znači da je korišten mobilni rendgenski aparat, koji se uglavnom koristi kad su pacijenti teže bolesni. Stoga i snimke napravljene s njim vjerojatno češće sadržavaju bolesti (a i izrazitije su pa su prepoznatljivije na snimci). Gotovo sigurno će model treniran na takvim snimkama predviđati veću vjerojatnost prisutnosti nekih stanja ukoliko na testiranoj snimci postoji oznaka „portable“. Takav način razmišljanja, iako logičan i poboljšava predviđanja, nije medicinski legitiman. Naravno, može se argumentirati da je to slično tomu da liječnik pri donošenju dijagnoze kombinira saznanja iz kliničke slike i rendgenske slike. U ovom slučaju bi model dubokog učenja zapravo iz rendgenske slike dobio saznanje da je bolesnik vjerojatno nepokretan, što je na neki način slično kliničkoj slici.

Dodatno, u radu Zech et al. (73) primijećeno je da model može iz rendgenskih snimki prepoznati u kojoj je bolnici snimka napravljena. Budući da su neke bolesti češće u različitim područjima, model bi mogao prema bolnici „zaključiti“ da je veća ili manja vjerojatnost prisutnosti bolesti. Takvo „zaključivanje“ također nije medicinski legitimno. Bez obzira na to, ipak poboljšava efektivnost modela, ali samo na bolnicama koje je naučio prepoznavati. Za nove bolnice ili aparate bi tako naučen model lošije funkcionirao.

Važno je imati na umu prethodno navedena saznanja i razumjeti da model ne uči dijagnosticirati kao što uči čovjek. Iako ovdje navedeni načini „zaključivanja“ modela nisu medicinski legitimni, ne može se poreći efektivnost modela dugokog učenja u rješavanju ovog zadatka.

10 Rasprava

Ručna analiza radioloških snimaka je posao na koji se troši mnogo vremena i resursa, a samo očitavanje snimki podliježe problemima neiskustva, subjektivnosti i nedovoljne standardizacije. Primjenom modela umjetne inteligencije može se značajno unaprijediti taj proces.

Poslijednjih nekoliko godina dolazi do ubrzanog razvoja modela umjetne inteligencije. Problemi računalnog vida (npr. klasifikacija slika) se danas uspješno rješavaju konvolucijskim neuronskim mrežama kao podvrstom dubokog učenja. Takvi modeli svoje mjesto pronalaze i u analizi medicinskih slika, a posebno onih radioloških.

S obzirom da je za učinkovito treniranje modela potreban velik broj slika, stvaraju se i mnoge javno dostupne baze podataka rendgenskih snimaka prsnog koša, kao što su ChestX-ray14, CheXpert, MIMIC-CXR, VinDr-CXR... Snimke koje one sadrže su svrstavane u kategorije ručnim unosom ili automatskom ekstrakcijom iz očitavanja radiologa.

Prije same uporabe modela umjetne inteligencije, snimke se predobrađuju raznim metodama kao što su poboljšanje kvalitete snimaka, segmentacija te supresija kostiju.

Brojni su radovi objavljeni na temu automatske dijagnostike rendgenskih snimaka prsnog koša, a u njima su predstavljeni i testirani različiti pristupi dizajniranju modela kojima bi se on optimizirao za navedeni problem. Sam učinak modela se procjenjuje prema iznosu AUC (area under the ROC curve) vrijednosti. Većina radova vrši treniranje i testiranje svojih modela na nekoj od javno dostupnih baza podataka rendgenskih snimaka. Jedan od najčešćih i najuspješnijih pristupa je korištenje predtreniranih modela (uz određene promjene arhitekture), koji su trenirani na javno dostupnoj bazi slika ImageNet, npr. DenseNet121, ResNet18, Inceptionv3... Uspoređen je i učinak takvih modela uz zadržavanje predtreniranih težina neuronske mreže i bez zadržavanja, odnosno uz nasumično inicijalizirane težine. Modeli sa zadržanim težinama su postizali bolje rezultate. Također, posebno uspješnima su se pokazali modeli koji su kombinacija više podmodela, gdje je konačno predviđanje konsenzus predviđanja svih podmodela. Dodavanje neslikovnih podataka (spol i dob pacijenta, vrsta snimke) kao ulaz u model nije doprinijelo učinkovitosti modela, vjerojatno jer model može takve informacije dobiti već iz samih snimki. Kao uspješnom metodom se pokazalo provlačenje snimke s označenim potencijalno prepoznatim stanjem kroz još jednu neuronsku mrežu, čime je omogućena uspješnija detekcija manjih lezija čija je lokacija nepredvidljiva. Osim toga, metodom DAM (Deep AUC Maximization) postignute su još bolje vrijednosti AUC-a pa se time i taj model smatra uspješnijim.

U radu su prikazani i rezultati primjene jedne od implementacija modela za predikciju stanja na nekoliko rendgenskih snimaka prsnog koša. Predviđene vjerojatnosti svake od labela koje nam model vraća se ne mogu gledati same za sebe, već se moraju uspoređivati vrijednosti dobivene za različite snimke. Konačno predviđanje postoji li neko

stanje na snimci ili ne se određuje s obzirom na definirani prag predviđene vjerojatnosti. On se pak definira s obzirom na to je li potrebna veća specifičnost ili osjetljivost modela. Izračuna li se specifičnost i osjetljivost modela za svaki od mogućih pragova, može se dobiti ROC krivulja, a iz nje i AUC vrijednost te se tako procjenjuje učinak modela.

Uz same predviđene vjerojatnosti za svaku od labela, mogu se i prikazati toplinske karte, iz kojih se vidi kako su koji dijelovi snimke utjecali na iznos predviđene vjerojatnosti. U većini slučajeva, ispravni dijelovi snimke su bili ključni za visoki iznos predviđene vjerojatnosti. Ipak, u pojedinim slučajevima se može primijetiti kako model predviđa ispravno, ali iz medicinski pogrešnih razloga. Primjer toga je prepoznavanje da je snimka snimljena s određenim rendgenskim uređajem koji se inače koristi za snimanje samo jedne bolesti, npr. pneumonije. U tom slučaju bi pneumonija bila prepoznata i ukoliko se istim uređajem snimi i zdrav pacijent.

Navedena saznanja i primjeri iz ovog rada služe dubljem razumijevanju tematike, trenutnog dosega i problema primjene umjetne inteligencije u automatskoj dijagnostici rendgenskih snimaka prsnog koša.

11 Zaključak

Primjena modela umjetne inteligencije ima potencijal znatno unaprijediti dijagnostiku rendgenskih snimaka prsnog koša smanjenjem trenutnih problema kao što su neiskustvo, subjektivnost i nedovoljna standardizacija. Korištenje predtreniranih modela (treniranih na ImageNet bazi) se pokazalo uspješnim. Posebno uspješnim su se pokazali modeli koji su kombinacija više podmodela i čije je konačno predviđanje konsenzus predviđanja svih podmodela. Dodavanje neslikovnih podataka o pacijentu nije doprinijelo učinkovitosti modela, ali metode segmentacije slike te Deep AUC Maximization metoda jesu. Konkretni model testiran u radu je ispravno predvidio veću vjerojatnost stanja za one snimke gdje su ta stanja stvarno bila prisutna. Isto tako, ispravan dio slike, tj. onaj gdje se neko stanje stvarno nalazilo, je bio ključan za takva predviđanja. Ipak, u nekim slučajevima se može vidjeti kako su nemedicinski podatci (npr. vrsta uređaja) sa snimke ključni za neko ispravno predviđanje (npr. ako se na istom uređaju snima samo jedno patološko stanje) pa treba biti svjestan eventualnih problema koji proizlaze iz te činjenice.

12 Zahvale

Zahvaljujem mentoru, doc. dr. sc. Andriji Štajduharu na vodstvu, pomoći te izrazitoj pristupačnosti i iskazanom povjerenju pri pisanju ovog diplomskog rada.

Zahvaljujem obitelji, za neizmjernu i nezamijenjivu podršku u svim aspektima svega čime sam se u životu odlučio baviti pa tako i ovog studija.

Zahvaljujem prijateljima, za sve zajednički proživljene trenutke i iskustva koja su ovo razdoblje studiranja učinili predivnim i nezaboravnim.

13 Literatura

1. ChatGPT [Internet]. [citirano 27. kolovoz 2023.]. Dostupno na: <https://openai.com/blog/chatgpt>
2. Haenlein M, Kaplan A. A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence. *California Management Review*. 17. srpanj 2019.;61:000812561986492.
3. Tech Advisor [Internet]. [citirano 27. kolovoz 2023.]. How tech giants are investing in artificial intelligence. Dostupno na: <https://www.techadvisor.com/article/724782/how-tech-giants-are-investing-in-artificial-intelligence.html>
4. Artificial Intelligence Market worth \$1,345.2 billion by 2030 [Internet]. [citirano 27. kolovoz 2023.]. Dostupno na: <https://www.marketsandmarkets.com/PressReleases/artificial-intelligence.asp%20.asp>
5. Murphy A. Radiopaedia. [citirano 27. kolovoz 2023.]. Chest (PA view) | Radiology Reference Article | Radiopaedia.org. Dostupno na: <https://radiopaedia.org/articles/chest-pa-view-1>
6. Delrue L, Gosselin R, Ilsen B, Van Landeghem A, de Mey J, Duyck P. Difficulties in the Interpretation of Chest Radiography. U: Coche EE, Ghaye B, de Mey J, Duyck P, urednici. *Comparative Interpretation of CT and Standard Radiography of the Chest* [Internet]. Berlin, Heidelberg: Springer; 2011 [citirano 27. kolovoz 2023.]. str. 27–49. (Medical Radiology). Dostupno na: <https://link.springer.com/book/10.1007/978-3-540-79942-9>
7. Thomsen L, Natho O, Feigen U, Schulz U, Kivelitz D. Value of digital radiography in expiration in detection of pneumothorax. *Rofo*. ožujak 2014.;186(3):267–73.
8. Chest X-ray Quality [Internet]. [citirano 27. kolovoz 2023.]. Dostupno na: https://www.radiologymasterclass.co.uk/tutorials/chest/chest_quality/chest_xray_quality_inspiration
9. Murphy A. Radiopaedia. [citirano 27. kolovoz 2023.]. Chest (AP erect view) | Radiology Reference Article | Radiopaedia.org. Dostupno na: <https://radiopaedia.org/articles/chest-ap-erect-view-1>
10. Murphy A. Radiopaedia. [citirano 27. kolovoz 2023.]. Chest (lateral view) | Radiology Reference Article | Radiopaedia.org. Dostupno na: <https://radiopaedia.org/articles/chest-lateral-view-2?lang=us>
11. Feigin DS. Lateral Chest Radiograph. *Academic Radiology*. prosinac 2010.;17(12):1560–6.

12. Hall F. Viewing the lateral radiograph. *American Journal of Roentgenology*. 01. prosinac 1982.;139(6):1235–1235.
13. Jones J. Radiopaedia. [citirano 27. kolovoz 2023.]. Chest x-ray review: ABCDE | Radiology Reference Article | Radiopaedia.org. Dostupno na: <https://radiopaedia.org/articles/chest-x-ray-review-abcde>
14. Ward D, Interpretation· DLP. Chest X-ray Interpretation | A Structured Approach | Radiology | OSCE [Internet]. 2016 [citirano 27. kolovoz 2023.]. Dostupno na: <https://geekymedics.com/chest-x-ray-interpretation-a-methodical-approach/>
15. Ryu Y. Radiopaedia. [citirano 27. kolovoz 2023.]. Chest radiograph assessment using ABCDEFGHI | Radiology Reference Article | Radiopaedia.org. Dostupno na: <https://radiopaedia.org/articles/chest-radiograph-assessment-using-abcdefghi>
16. Medicine LibreTexts [Internet]. 2019 [citirano 27. kolovoz 2023.]. 5.2: Approach to the Chest x-ray (CXR). Dostupno na: [https://med.libretexts.org/Bookshelves/Allied_Health/Undergraduate_Diagnostic_Imaging_Fundamentals_\(Burbridge_and_Mah\)/05%3A_Approach_to_Reviewing_X-ray_Imaging/5.02%3A_Approach_to_the_Chest_x-ray_\(CXR\)](https://med.libretexts.org/Bookshelves/Allied_Health/Undergraduate_Diagnostic_Imaging_Fundamentals_(Burbridge_and_Mah)/05%3A_Approach_to_Reviewing_X-ray_Imaging/5.02%3A_Approach_to_the_Chest_x-ray_(CXR))
17. McCarthy J, Minsky ML, Rochester N, Corporation IBM, Shannon CE. A PROPOSAL FOR THE DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE.
18. TURING AM. I.—COMPUTING MACHINERY AND INTELLIGENCE. *Mind*. 01. listopad 1950.;LIX(236):433–60.
19. McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*. 01. prosinac 1943.;5(4):115–33.
20. Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. *Nature*. listopad 1986.;323(6088):533–6.
21. Shi L, Wang XC, Wang YS. Artificial neural network models for predicting 1-year mortality in elderly patients with intertrochanteric fractures in China. *Braz J Med Biol Res*. 18. studeni 2013.;46:993–9. Licenca: <https://creativecommons.org/licenses/by/4.0/> Dostupno na: <https://www.scielo.br/j/bjmb/a/d9kB88m69gbPbDr5mVZ37kp/>
22. Esteva A, Robicquet A, Ramsundar B, Kuleshov V, DePristo M, Chou K, i ostali. A guide to deep learning in healthcare. *Nat Med*. siječanj 2019.;25(1):24–9.; str. 25, Springer Nature
23. Dai W, Dong N, Wang Z, Liang X, Zhang H, Xing EP. SCAN: Structure Correcting Adversarial Network for Organ Segmentation in Chest X-Rays. U: Stoyanov D, Taylor Z, Carneiro G, Syeda-Mahmood T, Martel A, Maier-Hein L, i ostali, urednici. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision*

- Support. Cham: Springer International Publishing; 2018. str. 263–73. (Lecture Notes in Computer Science).; str. 1, Springer Nature
24. Schulz H, Behnke S. Deep Learning. *Künstl Intell.* 01. studeni 2012.;26(4):357–63.; str. 1, Springer Nature
 25. Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L. ImageNet: A large-scale hierarchical image database. U: 2009 IEEE Conference on Computer Vision and Pattern Recognition. 2009. str. 248–55.
 26. Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition [Internet]. arXiv; 2015 [citirano 27. kolovoz 2023.]. Dostupno na: <http://arxiv.org/abs/1409.1556>
 27. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition [Internet]. arXiv; 2015 [citirano 27. kolovoz 2023.]. Dostupno na: <http://arxiv.org/abs/1512.03385>
 28. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the Inception Architecture for Computer Vision [Internet]. arXiv; 2015 [citirano 27. kolovoz 2023.]. Dostupno na: <http://arxiv.org/abs/1512.00567>
 29. Huang G, Liu Z, van der Maaten L, Weinberger KQ. Densely Connected Convolutional Networks [Internet]. arXiv; 2018 [citirano 27. kolovoz 2023.]. Dostupno na: <http://arxiv.org/abs/1608.06993>
 30. Elgendi M, Nasir MU, Tang Q, Smith D, Grenier JP, Batte C, i ostali. The Effectiveness of Image Augmentation in Deep Learning Networks for Detecting COVID-19: A Geometric Transformation Perspective. *Frontiers in Medicine* [Internet]. 2021. [citirano 27. kolovoz 2023.];8. Dostupno na: <https://www.frontiersin.org/articles/10.3389/fmed.2021.629134> Licenca: <https://creativecommons.org/licenses/by/4.0/>, © 2021 Elgendi, Nasir, Tang, Smith, Grenier, Batte, Spieler, Leslie, Menon, Fletcher, Howard, Ward, Parker and Nicolaou
 31. ukrfX.png (500×500) [Internet]. [citirano 01. rujan 2023.]. Dostupno na: <https://i.stack.imgur.com/ukrfX.png> Licenca: <https://creativecommons.org/licenses/by-sa/4.0/> © 2023 Stack Exchange
 32. Rajpurkar P, Irvin J, Ball RL, Zhu K, Yang B, Mehta H, i ostali. Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists. *PLoS Med.* studeni 2018.;15(11):e1002686. Dostupno na: <https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1002686> Licenca: <https://creativecommons.org/licenses/by/4.0/> © 2018 Rajpurkar et al
 33. CXR8 | Powered by Box [Internet]. [citirano 27. kolovoz 2023.]. Dostupno na: <https://nihcc.app.box.com/v/ChestXray-NIHCC>

34. Wang X, Peng Y, Lu L, Lu Z, Bagheri M, Summers RM. ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases. U: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) [Internet]. 2017 [citirano 27. kolovoz 2023.]. str. 3462–71. Dostupno na: <http://arxiv.org/abs/1705.02315>
35. Irvin J, Rajpurkar P, Ko M, Yu Y, Ciurea-Ilicus S, Chute C, i ostali. CheXpert: A Large Chest Radiograph Dataset with Uncertainty Labels and Expert Comparison [Internet]. arXiv; 2019 [citirano 27. kolovoz 2023.]. Dostupno na: <http://arxiv.org/abs/1901.07031>
36. Johnson AEW, Pollard TJ, Berkowitz SJ, Greenbaum NR, Lungren MP, Deng C ying, i ostali. MIMIC-CXR, a de-identified publicly available database of chest radiographs with free-text reports. *Sci Data*. 12. prosinac 2019.;6(1):317.
37. Nimalsiri W, Hennayake M, Rathnayake K, Ambegoda TD, Meedeniya D. CXLSeg Dataset: Chest X-ray with Lung Segmentation. 2023 International Conference On Cyber Management And Engineering (CyMaEn). 26. siječanj 2023.;327–31.
38. Bustos A, Pertusa A, Salinas JM, de la Iglesia-Vayá M. PadChest: A large chest x-ray image dataset with multi-label annotated reports. *Medical Image Analysis*. prosinac 2020.;66:101797.
39. Nguyen HQ, Lam K, Le LT, Pham HH, Tran DQ, Nguyen DB, i ostali. VinDr-CXR: An open dataset of chest X-rays with radiologist's annotations. *Sci Data*. 20. srpanj 2022.;9(1):429. Dostupno na: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9300612/> Licenca: <https://creativecommons.org/licenses/by/4.0/> © Autori 2022
40. Demner-Fushman D, Kohli MD, Rosenman MB, Shooshan SE, Rodriguez L, Antani S, i ostali. Preparing a collection of radiology examinations for distribution and retrieval. *J Am Med Inform Assoc*. ožujak 2016.;23(2):304–10.
41. Jaeger S, Candemir S, Antani S, Wang YXJ, Lu PX, Thoma G. Two public chest X-ray datasets for computer-aided screening of pulmonary diseases. *Quant Imaging Med Surg*. prosinac 2014.;4(6):475–7.
42. Ait Nasser A, Akhloufi MA. A Review of Recent Advances in Deep Learning Models for Chest Disease Detection Using Radiography. *Diagnostics*. siječanj 2023.;13(1):159. Dostupno na: <https://www.mdpi.com/2075-4418/13/1/159> Licenca: <https://creativecommons.org/licenses/by/4.0/> © 2023 Autori
43. Qin C, Yao D, Shi Y, Song Z. Computer-aided detection in chest radiography based on artificial intelligence: a survey. *BioMedical Engineering OnLine*. 22. kolovoz 2018.;17(1):113.

44. Teixeira LO, Pereira RM, Bertolini D, Oliveira LS, Nanni L, Cavalcanti GDC, i ostali. Impact of Lung Segmentation on the Diagnosis and Explanation of COVID-19 in Chest X-ray Images. *Sensors*. siječanj 2021.;21(21):7116. Dostupno na: <https://www.mdpi.com/1424-8220/21/21/7116> Licenca: <https://creativecommons.org/licenses/by/4.0/> © 2021 Autori
45. Li Y, Zhang Z, Dai C, Dong Q, Badrigilan S. Accuracy of deep learning for automated detection of pneumonia using chest X-Ray images: A systematic review and meta-analysis. *Computers in Biology and Medicine*. 01. kolovoz 2020.;123:103898.
46. Alapat DJ, Menon MV, Ashok S. A Review on Detection of Pneumonia in Chest X-ray Images Using Neural Networks. *J Biomed Phys Eng*. 01. prosinac 2022.;12(6):551–8.
47. Neuman MI, Lee EY, Bixby S, Diperna S, Hellinger J, Markowitz R, i ostali. Variability in the interpretation of chest radiographs for the diagnosis of pneumonia in children. *J Hosp Med*. travanj 2012.;7(4):294–8.
48. Hopstaken RM, Witbraad T, van Engelshoven JMA, Dinant GJ. Inter-observer variation in the interpretation of chest radiographs for pneumonia in community-acquired lower respiratory tract infections. *Clin Radiol*. kolovoz 2004.;59(8):743–52.
49. Rajpurkar P, Irvin J, Zhu K, Yang B, Mehta H, Duan T, i ostali. CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning [Internet]. arXiv; 2017 [citirano 27. kolovoz 2023.]. Dostupno na: <http://arxiv.org/abs/1711.05225>
50. Chouhan V, Singh SK, Khamparia A, Gupta D, Tiwari P, Moreira C, i ostali. A Novel Transfer Learning Based Approach for Pneumonia Detection in Chest X-ray Images. *Applied Sciences*. siječanj 2020.;10(2):559. Dostupno na: <https://www.mdpi.com/2076-3417/10/2/559> Licenca: <https://creativecommons.org/licenses/by/4.0/> © 2020 Autori
51. Stephen O, Sain M, Maduh UJ, Jeong DU. An Efficient Deep Learning Approach to Pneumonia Classification in Healthcare. *J Healthc Eng*. 2019.;2019:4180949. Dostupno na: <https://www.hindawi.com/journals/jhe/2019/4180949/> Licenca: <https://creativecommons.org/licenses/by/4.0/> © 2019 Okeke Stephen et al.
52. Kermany D, Zhang K, Goldbaum M. Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification. 06. siječanj 2018. [citirano 27. kolovoz 2023.];2. Dostupno na: <https://data.mendeley.com/datasets/rsbjbr9sj/2>
53. Irfan A, Adivishnu AL, Sze-To A, Dehkharghanian T, Rahnamayan S, Tizhoosh HR. Classifying Pneumonia among Chest X-Rays Using Transfer Learning. *Annu Int Conf IEEE Eng Med Biol Soc*. srpanj 2020.;2020:2186–9. Dostupno na: <https://ieeexplore.ieee.org/document/9175594> Licenca:

<https://creativecommons.org/licenses/by/4.0/> © 2020 IEEE
<https://ieeexplore.ieee.org/document/9175594>

54. Raghu M, Zhang C, Kleinberg J, Bengio S. Transfusion: Understanding Transfer Learning for Medical Imaging [Internet]. arXiv; 2019 [citirano 27. kolovoz 2023.]. Dostupno na: <http://arxiv.org/abs/1902.07208>
55. Kumar MD, Babaie M, Zhu S, Kalra S, Tizhoosh HR. A Comparative Study of CNN, BoVW and LBP for Classification of Histopathological Images [Internet]. arXiv; 2017 [citirano 27. kolovoz 2023.]. Dostupno na: <http://arxiv.org/abs/1710.01249>
56. Sharma H, Jain JS, Bansal P, Gupta S. Feature Extraction and Classification of Chest X-Ray Images Using CNN to Detect Pneumonia. U: 2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence). 2020. str. 227–31.
57. Ayan E, Ünver HM. Diagnosis of Pneumonia from Chest X-Ray Images Using Deep Learning. U: 2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT). 2019. str. 1–5.
58. Chollet F. Xception: Deep Learning with Depthwise Separable Convolutions [Internet]. arXiv; 2017 [citirano 27. kolovoz 2023.]. Dostupno na: <http://arxiv.org/abs/1610.02357>
59. Leaman R, Khare R, Lu Z. Challenges in clinical natural language processing for automated disorder normalization. J Biomed Inform. listopad 2015.;57:28–37.
60. Aronson AR, Lang FM. An overview of MetaMap: historical perspective and recent advances. J Am Med Inform Assoc. 2010.;17(3):229–36.
61. Yao L, Poblenz E, Dagunts D, Covington B, Bernard D, Lyman K. Learning to diagnose from scratch by exploiting dependencies among labels [Internet]. arXiv; 2018 [citirano 27. kolovoz 2023.]. Dostupno na: <http://arxiv.org/abs/1710.10501>
62. Baltruschat IM, Nickisch H, Grass M, Knopp T, Saalbach A. Comparison of Deep Learning Approaches for Multi-Label Chest X-Ray Classification. Sci Rep. 23. travanj 2019.;9:6381. Dostupno na: <https://www.nature.com/articles/s41598-019-42294-8>
Licenca: <https://creativecommons.org/licenses/by/4.0/>
63. Guan Q, Huang Y, Zhong Z, Zheng Z, Zheng L, Yang Y. Diagnose like a Radiologist: Attention Guided Convolutional Neural Network for Thorax Disease Classification [Internet]. arXiv; 2018 [citirano 27. kolovoz 2023.]. Dostupno na: <http://arxiv.org/abs/1801.09927>
64. Yuan Z, Yan Y, Sonka M, Yang T. Large-scale Robust Deep AUC Maximization: A New Surrogate Loss and Empirical Studies on Medical Image Classification [Internet]. arXiv; 2021 [citirano 27. kolovoz 2023.]. Dostupno na: <http://arxiv.org/abs/2012.03173>

65. Blais MA, Akhloufi MA. Deep Learning and Binary Relevance Classification of Multiple Diseases using Chest X-Ray images. Annu Int Conf IEEE Eng Med Biol Soc. studeni 2021.;2021:2794–7.
66. Papers with Code - CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning [Internet]. [citirano 27. kolovoz 2023.]. Dostupno na: <https://paperswithcode.com/paper/chexnet-radiologist-level-pneumonia-detection>
67. Zech J. reproduce-chexnet [Internet]. 2023 [citirano 27. kolovoz 2023.]. Dostupno na: <https://github.com/jrzech/reproduce-chexnet>
68. Wilson AC, Roelofs R, Stern M, Srebro N, Recht B. The Marginal Value of Adaptive Gradient Methods in Machine Learning [Internet]. arXiv; 2018 [citirano 27. kolovoz 2023.]. Dostupno na: <http://arxiv.org/abs/1705.08292>
69. Hacking C. Radiopaedia. [citirano 29. kolovoz 2023.]. Pneumonia | Radiology Reference Article | Radiopaedia.org. Licenca: <https://creativecommons.org/licenses/by-nc-sa/3.0/> Dostupno na: <https://radiopaedia.org/articles/pneumonia>
70. Radswiki T. Radiopaedia. [citirano 30. kolovoz 2023.]. Cardiomegaly | Radiology Reference Article | Radiopaedia.org. Licenca: <https://creativecommons.org/licenses/by-nc-sa/3.0/> Dostupno na: <https://radiopaedia.org/articles/cardiomegaly>
71. Gaillard F. Radiopaedia. [citirano 30. kolovoz 2023.]. Lung cancer | Radiology Reference Article | Radiopaedia.org. Licenca: <https://creativecommons.org/licenses/by-nc-sa/3.0/> Dostupno na: <https://radiopaedia.org/articles/lung-cancer-3>
72. Zech J. What are radiological deep learning models actually learning? [Internet]. Medium. 2018 [citirano 30. kolovoz 2023.]. Dostupno na: <https://jrzech.medium.com/what-are-radiological-deep-learning-models-actually-learning-f97a546c5b98>
73. Zech JR, Badgeley MA, Liu M, Costa AB, Titano JJ, Oermann EK. Confounding variables can degrade generalization performance of radiological deep learning models. PLoS Med. 06. studeni 2018.;15(11):e1002683.

14 Životopis

Rođen sam u Zagrebu, 27.2.1999.

Još tijekom pohađanja OŠ Dragutina Domjanića u Zagrebu sam otkrio svoju ljubav prema svim granama znanosti i tehnologije. Zbog toga sam upisao, a kasnije i završio, (prirodoslovno-matematičku) V. gimnaziju u Zagrebu, u kojoj se moj interes prema znanosti nastavio razvijati. Među ljepšim iskustvima u životu su mi bila sudjelovanja na državnim natjecanjima iz biologije. Osim toga, učenjem Pythona na satima informatike se razvijala i moja ljubav prema programiranju.

Zbog mnogih interesa, jako sam se dvoumio koji fakultet upisati (molekularnu biologiju, FER, FSB, fiziku). Ipak sam se prvo odlučio za medicinu, koju sam upisao 2017. godine nakon matura od kojih sam neke napisao među najboljima u državi (1. iz biologije, 7. iz kemije, jedna greška iz fizike).

Na 2. godini studija (2019.) mi je dodijeljena Dekanova nagrada za uspjeh.

Tijekom studija sam obavljao dužnost demonstratora na predmetu „fizika i biofizika“. Volontirao sam za udruhu CroMSIC te sudjelovao u njenim projektima „Pogled u sebe“ i „The Talk“ u sklopu kojih sam držao edukacijske radionice učenicima srednje škole. Mjesec dana sam proveo na razmjeni na odjelu kardiologije u Cataniji na Siciliji.

Nakon 3 godine studija medicine, 2020., paralelno upisujem FER, u želji da jednog dana mogu u radu kombinirati znanja stečena na oba fakulteta.

Preddiplomski studij računarstva na FER-u sam završio u srpnju 2023. te sam odmah, netom prije predaje ovog diplomskog rada, upisao i diplomski studij „znanost o podacima“, također na FER-u.