

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

ZAVRŠNI RAD br. 1444

**KLASIFIKACIJA BOLESTI TEMELJEM RENDGENSKIH
SNIMKI PRSNOG KOŠA KORIŠTENJEM METODA
STROJNOG UČENJA**

Andrea Milanović

Zagreb, lipanj 2024.

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

ZAVRŠNI RAD br. 1444

**KLASIFIKACIJA BOLESTI TEMELJEM RENDGENSKIH
SNIMKI PRSNOG KOŠA KORIŠTENJEM METODA
STROJNOG UČENJA**

Andrea Milanović

Zagreb, lipanj 2024.

Zagreb, 4. ožujka 2024.

ZAVRŠNI ZADATAK br. 1444

Pristupnica: **Andrea Milanović (0036543226)**
Studij: Elektrotehnika i informacijska tehnologija i Računarstvo
Modul: Računarstvo
Mentor: izv. prof. dr. sc. Alan Jović

Zadatak: **Klasifikacija bolesti temeljem rendgenskih snimki prsnog koša korištenjem metoda strojnog učenja**

Opis zadatka:

Rendgenske snimke prsnog koša važan su izvor informacija za klasifikaciju većeg broja bolesti organizma. Liječnicima bi od značajne dodatne pomoći bio automatizirani sustav koji bi ispravno klasificirao razne bolesti temeljem rendgenskih snimki prsnog koša. Cilj ovog završnog rada je izrada klasifikacijskog modela za nekoliko vrsta oboljenja temeljem rendgenskih snimki prsnog koša, kao i segmentacijskog modela koji će označiti lokalno područje za veći broj specifičnih oboljenja na temelju istih snimki. Kao skup za učenje modela potrebno je upotrijebiti neki od skupova podataka rendgenskih snimki prsnog koša koje su kvalitetno označili liječnici, kao što je VinDr-CXR, dostupan na web sjedištu PhysioNet (<https://physionet.org/content/vindr-cxr/1.0.0/>). Klasifikacijski i segmentacijski modeli trebaju biti zasnovani na nekom od algoritama strojnog učenja, uključujući algoritme dubokog učenja. Primjeri takvih algoritama su konvolucijske neuronske mreže i transformerske arhitekture. U radu je također potrebno razmotriti metode predobrade snimki kako bi se poboljšala učinkovitost izgrađenih modela. Implementaciju je potrebno napraviti u programskom jeziku po vlastitom izboru, a za izgradnju modela može se u slučaju nedostatka vlastitih sklopovskih resursa koristiti dostupna web rješenja (npr. Google Colab).

Rok za predaju rada: 14. lipnja 2024.

Sadržaj

1. Uvod	2
2. Opis problema i skupa podataka	3
2.1. Prsni koš i pluća	3
2.2. Plućne bolesti	4
2.3. Rendgenske snimke prsnog koša	6
3. Strojno učenje	8
3.1. Neuronske mreže	8
3.2. Prednaučeni modeli	10
3.2.1. Klasifikacija	10
3.2.2. Detekcija	12
4. Izrada rješenja	15
4.1. Predobrada snimki	15
4.2. Klasifikacijski model	17
4.3. Detekcijski model	22
5. Zaključak	27
Literatura	28
Sažetak	31
Abstract	32

1. Uvod

Rendgenske snimke prsnog koša predstavljaju ključan dijagnostički alat u medicini, omogućujući liječnicima da identificiraju i prate različite bolesti i stanja koja zahvaćaju pluća i srce. Međutim, analiza rendgenskih snimaka može biti izazovna i zahtijeva visoku razinu stručnosti i iskustva. Ubrzan razvoj tehnologije umjetne inteligencije, posebno strojnog učenja, nudi mogućnost razvoja automatiziranih sustava koji bi mogli značajno pomoći liječnicima u preciznijoj i bržoj dijagnostici.

Cilj ovog završnog rada je razviti klasifikacijski model sposoban za prepoznavanje nekoliko vrsta oboljenja na temelju rendgenskih snimki prsnog koša, kao i detekcijski model koji će moći precizno označiti lokalna područja zahvaćena specifičnim oboljenjima. Kao temelj za izgradnju ovih modela koristit će se kvalitetno označen skup podataka rendgenskih snimki prsnog koša, VinDr-CXR dostupan na web sjedištu PhysioNet [1].

U izradi modela koristit će se napredni algoritmi strojnog učenja, specifično duboko učenje kroz konvolucijske neuronske mreže. Također će biti obraćena pažnja na metode predobrade snimaka za poboljšanje uspješnosti modela.

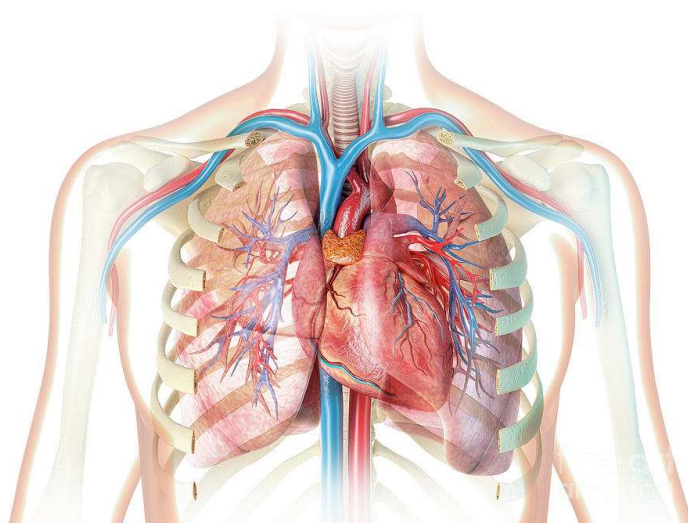
Rad se sastoji od pet poglavlja nakon kojih slijedi popis literature i sažetak. Poglavlja razmatraju građu prsnog koša i njegove bolesti koje nastojimo detektirati, zatim strojno učenje i posebice neuronske mreže, potom slijedi detaljan opis provedenih postupaka te je na kraju iznesen zaključak.

2. Opis problema i skupa podataka

Prije razmatranja razvoja modela potrebno je prvo proučiti značenje, upotrebu i korisnost korištenog skupa podataka. Dakle, u ovom poglavlju razmotrena je građa prsnog koša i pluća, neka oboljenja tog područja te korišteni skup podataka.

2.1. Prsni koš i pluća

Prsni koš ili koštani kavez sastoji se od prsne kosti, rebara i kralježnice, a igra ulogu zaštite pluća i drugih organa u prsištu [2]. Pluća su podijeljena u dva plućna krila do kojih vode dva bronha iz dušnika. Jedno plućno krilo građeno je od manjih dijelova, režnjeva, pri čemu desno ima tri, a lijevo dva. Bronhi se dijele u više dišnih puteva i bronhiola koji konačno vode u alveole. Alveole omogućuju da kisik prijeđe iz zraka u kapilare i krvotok, a ugljikov dioksid obrnuto. Konačno, pluća omeđuje poplućnica, glatka membrana na unutarnoj strani prsnog koša [3].



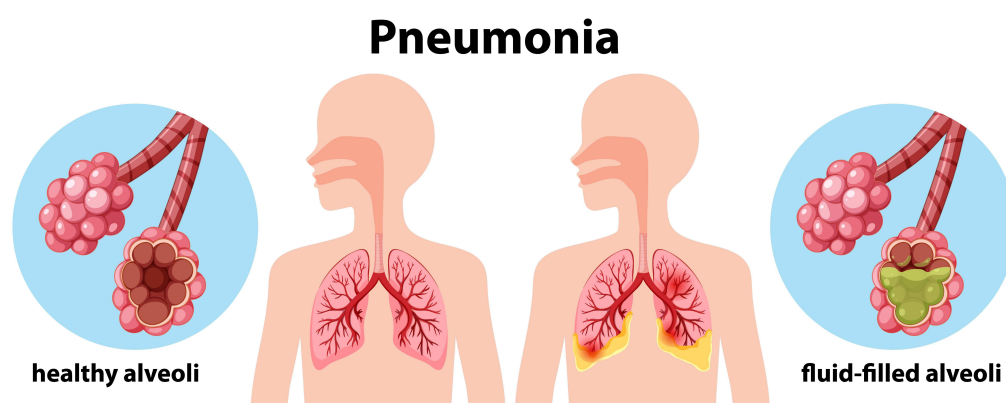
Slika 2.1. Prikaz prsnog koša i njegovih organa [4]

Osim pluća, u sredini grudnog koša, između dva plućna krila nalazi se prostor medijastinuma. U njemu se nalaze srce, timus, limfni čvorovi, dijelovi aorte i šuplje vene, slika 2.1. Upravo zbog važne uloge disanja i dišnog sustava za ljudski život te važnosti ostalih organa unutar prsne šupljine bitno je imati dobre metode nadziranja rada tih organa te detekcije potencijalnih oboljenja.

2.2. Plućne bolesti

Plućna oboljenja koja su razmotrena u ovom radu podijeljena su na globalna i lokalna. Od konkretnih bolesti za klasifikaciju korištene su pneumonija, tuberkuloza, tumor pluća te kronična opstruktivna bolest pluća.

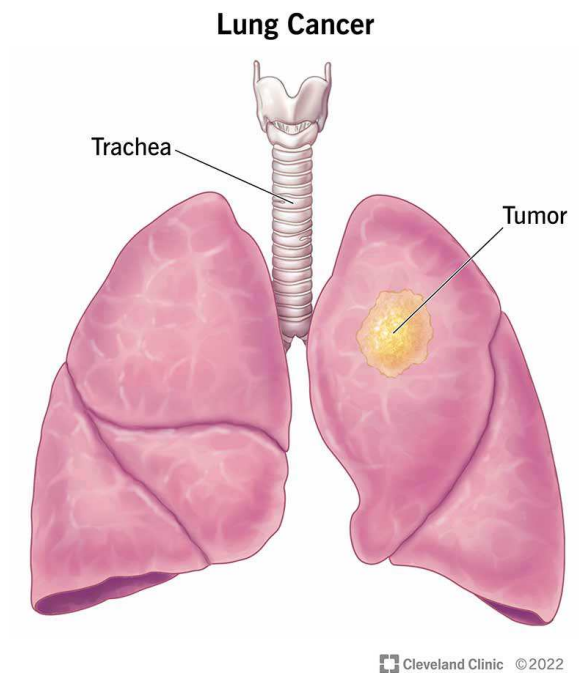
Upalu pluća, odnosno pneumoniju, izazivaju bakterije, virusi ili gljivice, a upala može uzrokovati tekućinu u plućima, slika 2.2. U slučaju sumnje na upalu pluća liječnik pacijenta šalje na rendgen prsnog koša, uzima se kultura ispljuvka, kontrolira se količina kisika u krvi te sama krv. Ako se ne tretira, pneumonija može biti opasna po život [5].



Slika 2.2. Ilustracija pneumonije [6]

Tuberkuloza je zarazna bolest koju izaziva vrsta bakterije te se širi zračnim putem. Dijagnosticira se Mantouxovim testom te pretragama krvi, a rendgenom se provjerava je li aktivna u području pluća [7].

Tumor pluća javlja se kad dođe do mutacija stanica koje se zatim počinju nekontrolirano dijeliti, a tipično započinje u dišnim putevima ili alveolama, slika 2.3. Dijagnoza započinje rendgenom i pretragama krvi, no kasnije se provode biopsije i CT skenovi [8].



Slika 2.3. Tumor pluća [8]

Kronična opstruktivna plućna bolest (KOPB) je zajednički naziv za više progresivnih plućnih bolesti koje izazivaju oštećenja u plućima. Glavni uzročnik je pušenje, a tipično se dijagnosticira spirometrijom, provjerama razina kisika u krvi, elektrokardiogramom te rendgenom.

Od lokalnih oznaka karakterističnih za skup podataka korišten u ovom radu, upotrebljavaju se sljedeće:

- Kardiomegalija (engl. *Cardiomegaly*)
- Zamućenje pluća (engl. *Lung Opacity*)
- Pomak mediastinuma (engl. *Mediastinal shift*)
- Čvorić/Masa (engl. *Nodule/Mass*)
- Atelektaza (engl. *Atelectasis*)
- Infiltracija (engl. *Infiltration*)
- Plućna fibroza (engl. *Pulmonary fibrosis*)
- Konsolidacija (engl. *Consolidation*)
- Intersticijska bolest pluća (engl. *ILD*)
- Kalcifikacija (engl. *Calcification*)
- Zadebljanje pleure (engl. *Pleural thickening*)

- Pleuralni izljev (engl. *Pleural effusion*)

2.3. Rendgenske snimke prsnog koša

Rendgenske zrake vrsta su elektromagnetske radijacije koja prolazi kroz većinu objekata, između ostalog i kroz ljudsko tijelo. U medicini rendgen se koristi za generiranje slika tkiva i struktura unutar tijela [9].

Konkretno, rendgensko snimanje prsnog koša radi se kad postoji sumnja na bolesti pluća i srca. Snimka se najčešće radi s leđa prema naprijed te dobro prikazuje sjenu srca, glavne krvne žile i može otkriti ozbiljniju bolest pluća [10]. Slika 2.4. primjer je izgleda jednog rendgena prsnog koša, a na njoj specifično nema pronađenih abnormalnosti.



Slika 2.4. Rendgenska slika iz skupa podataka [1]

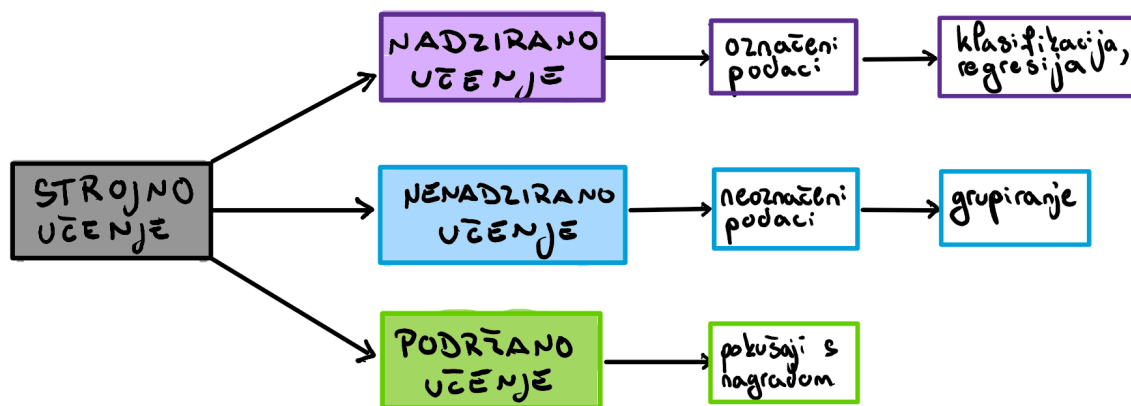
Skup podataka VinDr-CXR korišten za izradu ovog rada sastoji se od rendgenskih snimki prsnog koša iz dvije bolnice u Vijetnamu [1]. Snimke je označilo ukupno 17 radiologa, pri čemu postoje 22 lokalna oboljenja označena pravokutnicima oko abnormalnosti te 6 globalnih oznaka oboljenja. Ukupno je dostupno 18 000 slika koje su autori

podijelili na skup za učenje koji sadrži 15 000 slika i skup za testiranje od preostalih 3 000 slika. Svaku snimku iz skupa za učenje neovisno su označila tri radiologa te su zadržane sve takve oznake, dok se za testni skup oznaka sastoji od jednog konsenzusa pet različitih radiologa.

Same rendgenske snimke pohranjene su u formatu DICOM (engl. *Digital Imaging and Communications in Medicine*). Riječ je o standardu korištenom za prenošenje, spremanje, obradu i prikaz različitih medicinskih snimaka [11]. Osim same slike, DICOM datoteka sadrži i zaglavlje sastavljeno od raznih oznaka. Zaglavlje tipično sadrži informacije o pacijentu poput imena i prezimena, dobi, datuma rođenja, spola, itd., ali također i neke informacije o samoj slici poput dimenzija slike.

3. Strojno učenje

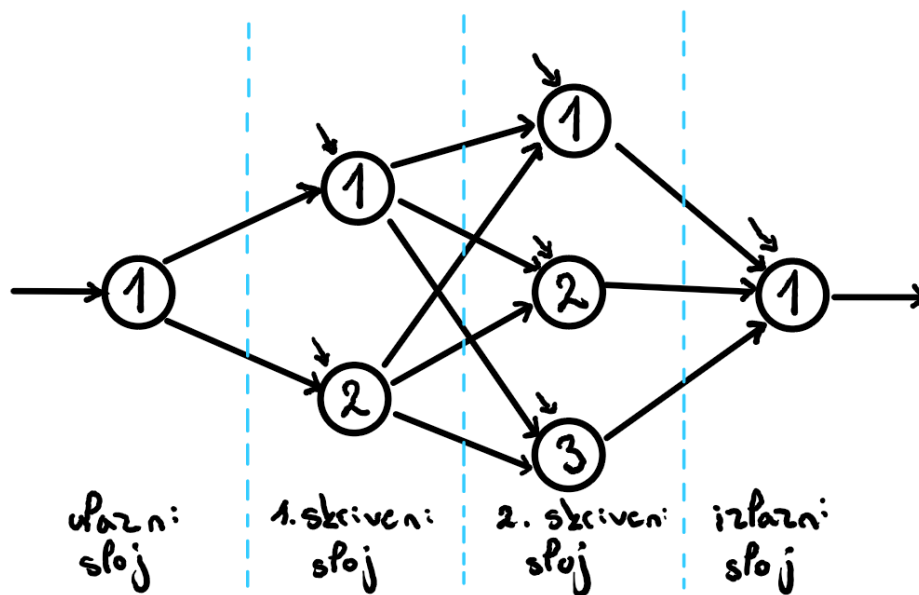
Strojno učenje (engl. *Machine Learning*) je grana umjetne inteligencije koja pokušava imitirati način na koji ljudi uče, postupno poboljšavajući svoju preciznost. To je programiranje računala tako da optimiraju kriterij uspješnosti koristeći pritom primjere podataka ili prethodno iskustvo. Za to se koristi model definiran do nekih parametara, a učenje se svodi na izvođenje algoritma koji optimira parametre modela [12]. Glavni cilj strojnog učenja je dobro generalizirati, odnosno točno predviđati na neviđenim podacima. Postoje tri vrste strojnog učenja: nadzirano, nenadzirano i podržano. Slika 3.1. ilustracija je te podjele i upotrebe svake vrste.



Slika 3.1. Vrste strojnog učenja

3.1. Neuronske mreže

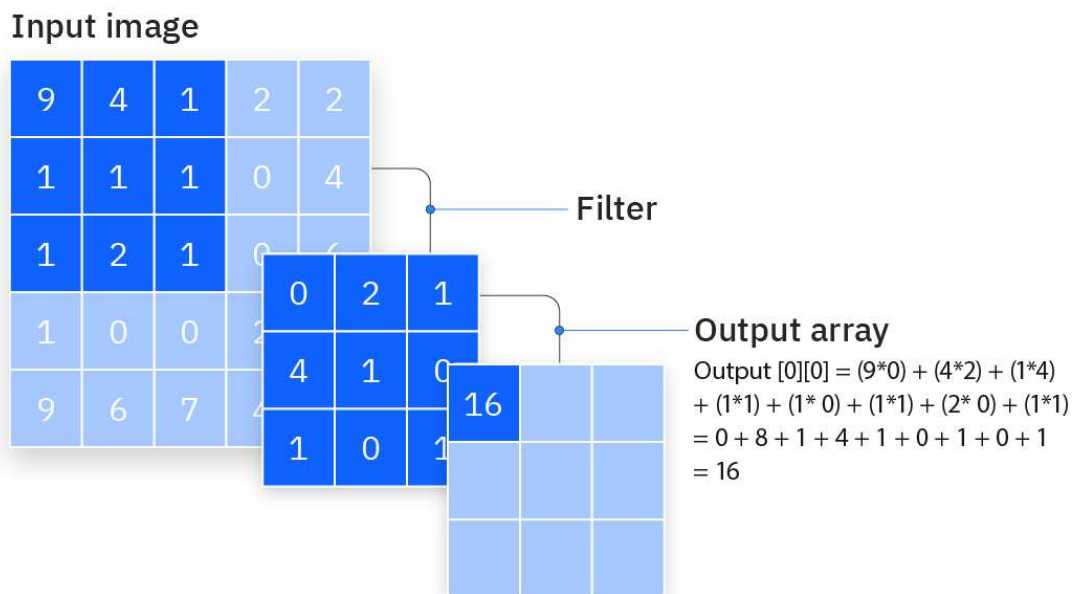
Neuronske mreže su podgrana strojnog učenja. Sama neuronska mreža sastavljena je od više slojeva čvorova: ulazni sloj, zatim neki broj skrivenih slojeva i konačno izlazni sloj. Svaki čvor, odnosno neuron, ima svoju težinu te se spaja s neuronima iz susjednih slojeva pri čemu veze također imaju pridružene težine [13]. Prikaz jedne takve jednostavne mreže s dva skrivena sloja nalazi se na slici 3.2.



Slika 3.2. Prikaz jednostavne neuronske mreže

Same neuronske mreže uobičajeno su unaprijedne, znači teku u jednom smjeru od ulaznog sloja prema izlaznom. Također, mrežu možemo učiti korištenjem postupka propagacije unazad koji nam omogućuje dobro prilagođavanje težina ovisno o pogrešci izlaza [14]. Skup primjera za učenje često dijelimo na skup za učenje, skup za validaciju te skup za testiranje. Skup za učenje iterativno podešava težine veza između neurona, validacija namješta hiperparametre mreže, a testni skup provjerava kvalitetu mreže na prethodno neviđenim podacima. Ne postoji strogi dogovor oko raspodjele između tih skupova no postoci se kreću oko 70 % za učenje, 15 % za validaciju i 15 % za testiranje.

Neuronske mreže tipično se koriste za obradu slike, zvuka i teksta, a pogotovo često se za klasifikaciju slika koriste konvolucijske neuronske mreže (CNN). Riječ je o neuronskoj mreži sastavljenoj od konvolucijskog sloja, sloja sažimanja (engl. *pooling layer*) i potpuno povezanog sloja. Konvolucijski sloj filtrom prelazi preko slike te traži postojanje neke značajke i to rezultira mapom značajki (engl. *feature map*) kao što je ilustrirano na slici 3.3. Sloj spajanja bavi se smanjenjem dimenzionalnosti i broja ulaznih parametara. Konačno, potpuno povezani sloj obavlja zadatak klasifikacije te, kao što i ime kaže, zadnji i predzadnji sloj potpuno su povezani.



Slika 3.3. Rad konvolucijskog sloja [15]

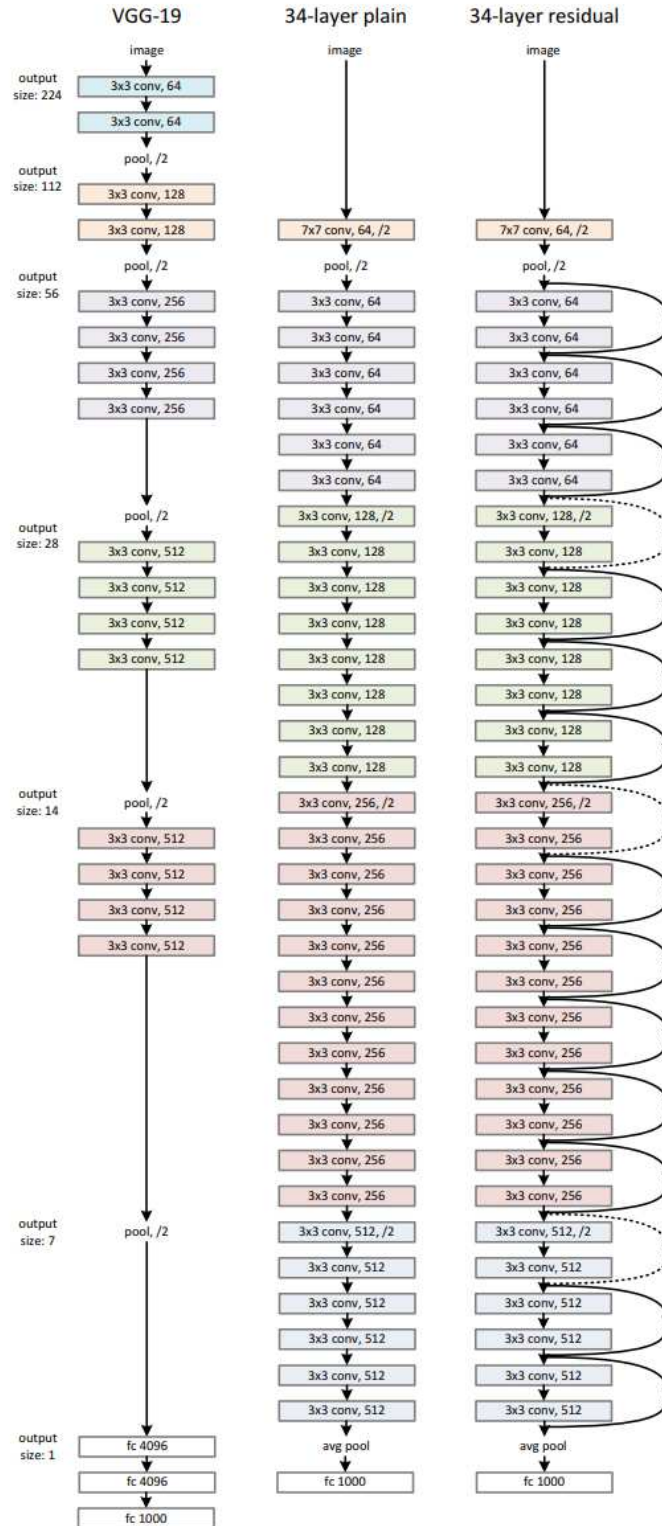
3.2. Prednaučeni modeli

Prednaučeni modeli postoje zbog toga što je za dobar model koji obavlja klasifikaciju ili detekciju slika potreban velik skup podataka te značajna količina resursa. Riječ je modelima koji su prethodno naučeni na nekom većem skupu podataka i nakon toga je omogućeno da ih kasniji korisnik modificira za neki njemu potreban zadatak.

3.2.1. Klasifikacija

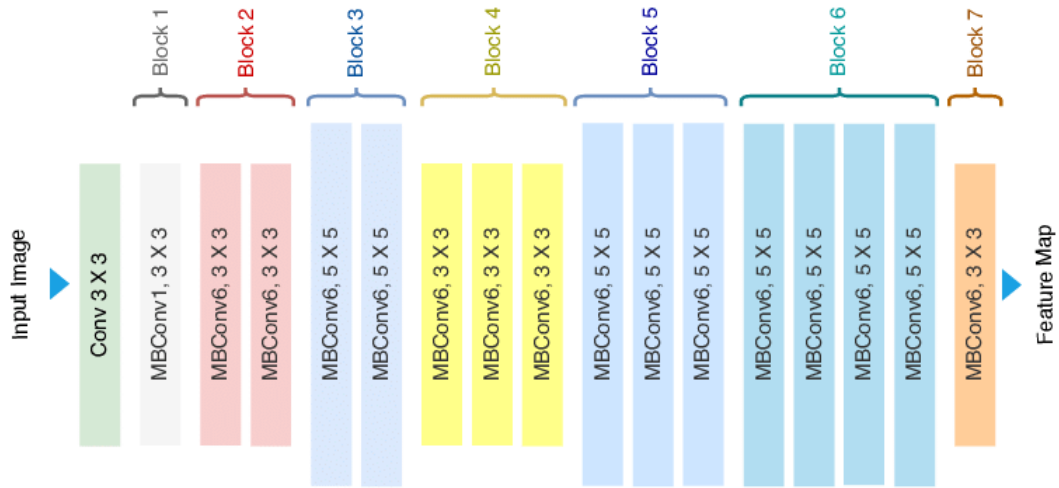
Klasifikacija podrazumijeva da se u model ubacuju slike s nekim globalnim oznakama za tu sliku te model na temelju skupa za učenje nauči određivati oznake za prije neviđene slike. U kontekstu prednaučenih modela u tu svrhu obično se mijenja samo zadnji, izlazni, sloj neuronske mreže koji određuje klasifikaciju specifične slike. U sklopu izrade rada korišten je klasifikacijski model ResNet kao i EfficientNet.

ResNet (engl. *Residual Network*) je arhitektura za duboko učenje koja se koristi za duboke neuronske mreže korištenjem tehnike rezidualnog učenja. Umjesto direktnog preslikavanja od ulaza do izlaza postoje rezidualni blokovi koji dijele taj proces na više dijelova. Time se uklanja problem nestajanja gradijenta koji se pojavljuje kod drugih vrlo dubokih neuronskih mreža. Primjer arhitekture rezidualne mreže nalazi se na slici 3.4. desno.



Slika 3.4. Prikaz arhitekture modela VGG, obične mreže i rezidualne mreže [16]

EfficientNet je konvolucijska neuronska mreža temeljena na konceptu skaliranja dijelova (engl. *compound scaling*) i pritom se uravnoteženo skaliraju širina, dubina i rezolucija. Izgled arhitekture modela EfficientNet vidljiv je na slici 3.5.



Slika 3.5. Prikaz arhitekture modela EfficientNet [17]

Kako bi bilo moguće ocijeniti uspješnost rada klasifikacijskog modela, koriste se neke mjere procjene. Tijekom učenja i testiranja modela gledamo primarno točnost, formula 3.1, koja daje postotak točno klasificiranih slika, te kategorijsku unakrsnu entropiju (engl. *Categorical Crossentropy*), formula 3.2, koja predstavlja mjeru pogreške, odnosno koliko se dobro predviđena vrijednost poklapa sa stvarnom.

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}} \quad (3.1)$$

$$\text{Categorical Crossentropy} = -\frac{1}{n} \sum_{i=1}^n \sum_{c=1}^C y_{ic} \log(\hat{y}_{ic}) \quad (3.2)$$

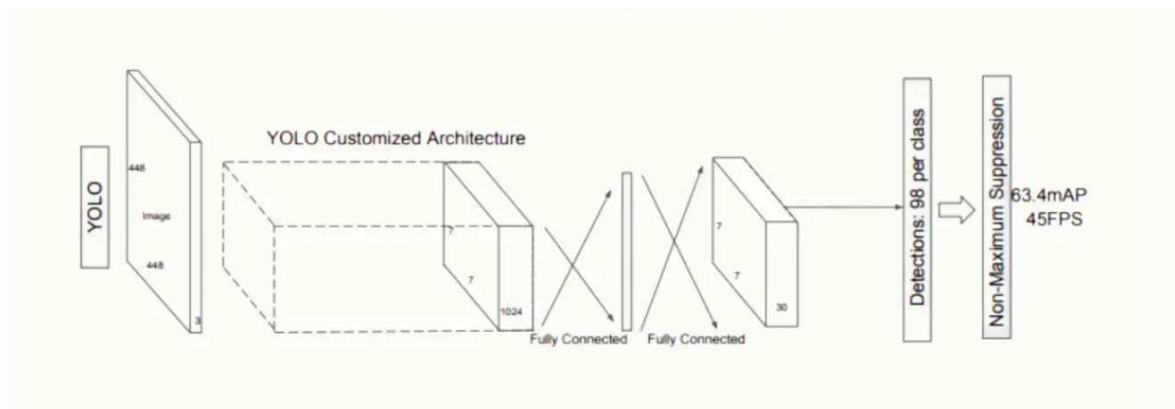
Također u obzir treba uzeti matricu konfuzije, matricu čiji redci predstavljaju stvarne klasifikacije, a stupci predviđene vrijednosti. Dakle vrijednost ćelije n u retku x i stupcu y daje informaciju da je n puta za stvarni razred x predviđena pripadnost razredu y .

3.2.2. Detekcija

Za razliku od klasifikacije, detekcija kao ulaz u model, osim same slike očekuje i neke lokalne oznake. Dakle, uz sliku, modelu se predaje lokaciju označenog pravokutnika u pikselima te kojem razredu ta oznaka pripada. To znači da je na jednoj slici moguća pojava više lokalnih oznaka. Od modela se očekuje da će na neviđenim slikama moći

pronaći, označiti i razvrstati takve lokalne elemente slike. U sklopu izrade rada za taj zadatak korišten je model YOLOv8.

YOLO (engl. *You Only Look Once*) je model dubokog učenja koji se koristi upravo za detekciju elemenata na slici. Koristi jednu konvolucijsku neuronsku mrežu s kojom predviđa okvire oznaka te vjerojatnosti pripadnosti nekom razredu. Algoritam YOLO dijeli sliku na ulazu u mrežu ćelija te za svaku ćeliju zatim predviđa vjerojatnost postojanja objekta i koordinate okvira oko objekta [18]. Također predviđa razred kojem objekt pripada. Arhitektura takvog modela prikazana je na slici 3.6.



Slika 3.6. Prikaz arhitekture modela YOLO [18]

Za ocjenjivanje uspješnosti detekcijskih modela postoji ipak nešto više mjera od klasifikacijskih budući da treba odrediti jesu li pravokutnici dobro označeni te koliko se površine poklapaju sa stvarnim oznakama. Glavne takve evaluacijske mjere su:

- Presjek u odnosu na uniju (engl. *Intersection Over Union*, IoU) koji nam pokazuje udio preklapanja za dvije oznake, a računa se po formuli 3.3

$$\text{IoU} = \frac{\text{Area of Intersection}}{\text{Area of Union}} \quad (3.3)$$

- Preciznost (engl. *Precision*) koji je omjer broja točnih detekcija u odnosu na sve detekcije, formula 3.4

$$P = \frac{TP}{TP + FP} \quad (3.4)$$

- Odziv (engl. *Recall*) koji je omjer prepoznatih objekata u odnosu na sve objekte,

formula 3.5

$$P = \frac{TP}{TP + FN} \quad (3.5)$$

- Prosječna preciznost (engl. *Average Precision, AP*) koja je površina ispod krivulje preciznosti i odziva, formula 3.6

$$AP = \int_0^1 P(R) d(R) \quad (3.6)$$

- Srednja prosječna preciznost (engl. *Mean Average Precision, mAP*) je srednja vrijednost AP vrijednosti za svaki od razreda oznaka, računa se s formulom 3.7

$$mAP = \frac{1}{C} \sum_{c=1}^C AP_c \quad (3.7)$$

- Matrica konfuzije (engl. *Confusion Matrix*) koristi se na isti način kao i kod klasiifikacijskog modela za vizualni prikaz raspodjele između pronađenih i očekivanih detekcija

4. Izrada rješenja

Svrha ovog rada bila je izrada klasifikacijskog modela te detekcijskog modela za rendgenske slike prsnog koša. U sklopu izrade tih modela korišten je programski jezik Python. Također su korištene biblioteke poput pydicom, PIL, cv2 te shutil za baratanje samim slikama i datotekama, biblioteke pandas i matplotlib za baratanje oznakama, analizu sastava podataka i vizualni prikaz, te konačno biblioteke tensorflow, keras, torch i ultralytics za izradu samih modela.

4.1. Predobrada snimki

Prije izgradnje modela potrebno je prvo napraviti predobradu snimki. Budući da su slike u formatu DICOM potrebno ih je prvo prevesti u neki od standardnih formata za slike poput JPEG-a. Na slici 4.1. prikazana je funkcija za pretvorbu slike u odgovarajući format s normalizacijom njezinih piksela na vrijednosti 0-255 što zahtjeva format JPEG.

```
def dicom2jpeg(dicom_file, jpeg_file):
    try:
        dicom = pydicom.dcmread(dicom_file)
        pixel_array = dicom.pixel_array
        # normalizing pixel values
        normalized_array = cv2.normalize(pixel_array, dst=None, alpha=0, beta=255, cv2.NORM_MINMAX)
        image_8bit = np.uint8(normalized_array)
        image = Image.fromarray(image_8bit)
        image.save(jpeg_file, format='JPEG')
        print(f"Successfully converted {dicom_file} to {jpeg_file}")
    except Exception as e:
        print(f"Failed to convert {dicom_file}: {e}")
```

Slika 4.1. Funkcija za pretvaranje iz formata DICOM u JPG

Budući da rendgenske snimke nisu uvijek najbolje moguće kvalitete, jer se nastoji smanjiti količina radijacije kojom se pacijent izlaže, potrebno je poboljšati vidljivost anomalija algoritamski. Stoga je nakon prevođenja slike, nad njom proveden CLAHE (engl.

Contrast Limited Adaptive Histogram Equalization), algoritam za poboljšanje lokalnog kontrasta u slikama. Taj algoritam sliku dijeli u mrežu, za svako polje u mreži prilagođava lokalni kontrast kroz jednačenje histograma, pri čemu je kontrast ograničen kako ne bi došlo do pojačanja šuma, te na kraju spaja polja nazad u jednu sliku pritom koristeći bilinearnu interpolaciju kako ne bi bilo naglih prijelaza između polja [19]. Nakon primjene algoritma kao finalni korak slikama je promijenjena veličina na 224x224 piksela, 380x380 i 640x640 što zahtijevaju kasnije korišteni modeli. Slika 4.2. prikazuje funkciju koja vrši takvu predobradu.

```
def preprocessAndResizeImage(image_path, target_size):
    image = cv2.imread(image_path, cv2.IMREAD_GRAYSCALE) # load
    clahe = cv2.createCLAHE(clipLimit=2.0, tileGridSize=(8, 8))
    clahe_image = clahe.apply(image) # apply CLAHE
    resized_image = cv2.resize(clahe_image, target_size) # resize to target size
    normalized_image = resized_image / 255.0 # normalize image to range [0, 1]
    return normalized_image
```

Slika 4.2. Funkcija za predobradu slika

Osim predobrade samih slika, skup podataka također je bilo potrebno pripremiti za modele. Za početak, budući da je skup velikom većinom bio sastavljen od slika na kojima nisu pronađene nikakve abnormalnosti (njih oko 12 000 od sveukupno 17 000 [1]), uklonjena je velika većina takvih podataka. To je provedeno na način da je slučajno odabrana desetina tih slika, oko 1 200, te su samo one korištene dalje tijekom učenja modela. Osim toga, postoji i globalna oznaka za "ostale bolesti". Nepoznato je što sve ta kategorija može podrazumijevati, stoga su slike gdje su radiolozi jednoznačno sliku označili kao takvu također maknute iz skupa podataka. Potom, slike je potrebno podijeliti na tri podskupa, skup za učenje, validacijski skup te skup za testiranje. Budući da su podaci već bili razdijeljeni na skup za učenje i testni skup, samo je iz oba odvojen dio slika te iz toga stvoren validacijski skup tako da je podjela otprilike 70 %, 15 % i 15 %. Slike su odvojene uz pomoć biblioteke *random* u Pythonu te funkcije *sample* koja od zadanog skupa podataka slučajno odabire uzorak zadane veličine iz te populacije.

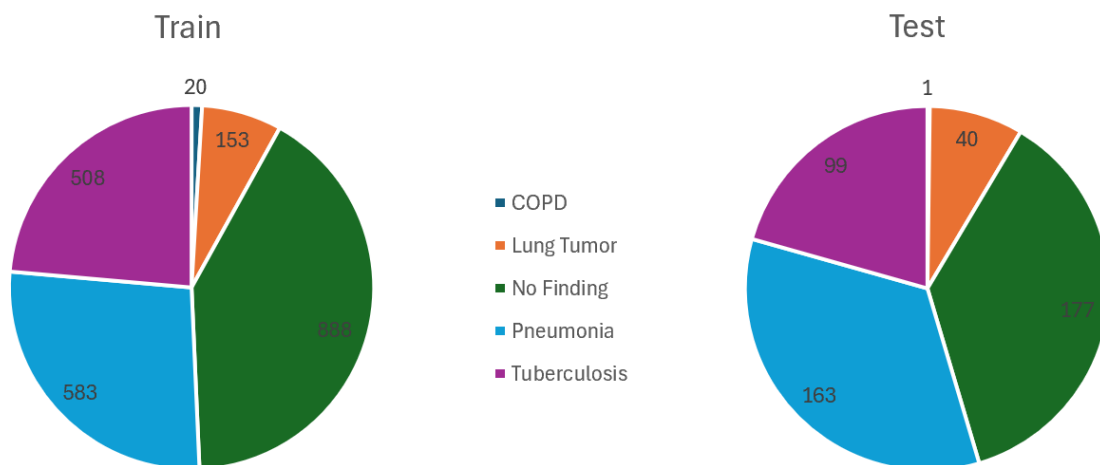
```
def normalizeAnnotations(imagePath, labelPath, annotations, classes):
    with Image.open(imagePath) as image:
        width, height = image.size
    if annotations.empty:
        return
    with open(labelPath, 'w') as file:
        for index, row in annotations.iterrows():
            xMin = row['x_min']
            xMax = row['x_max']
            yMin = row['y_min']
            yMax = row['y_max']
            className = row['class_name']
            classID = classes[className] if className in classes else '12'
            xCenter = (xMin + xMax) / 2
            yCenter = (yMin + yMax) / 2
            anWidth = xMax - xMin
            anHeight = yMax - yMin
            file.write(f'{classID} {xCenter / width} {yCenter / height} {anWidth / width} {anHeight / height}\n')
```

Slika 4.3. Funkcija za normalizaciju lokalnih oznaka

Konačno za pripremu za model trebalo je normirati labele. Modeli za detekciju obično kao ulaz prihvaćaju vrijednosti iz intervala $[0,1]$. Dakle, kao što je prikazano na slici 4.3., za svaku oznaku bilo je potrebno podijeliti vrijednost x-koordinata sa širinom specifične slike, a vrijednost y-koordinata s visinom slike. Također, umjesto krajnjih točaka pravokutnika, model prima središte pravokutnika te visinu i širinu pa je i to uzeto u obzir u korištenoj funkciji.

4.2. Klasifikacijski model

Budući da je svaku sliku koja je došla iz originalnog podskupa za učenje označilo po troje radiologa bez konkretnog konsenzusa, trebalo je smisliti nekakav pristup toj paradoksalnosti. Čak i u podskupu za testiranje, gdje postoji konačna odluka o dijagnozi, ima slika gdje je istovremeno označeno primjerice pneumonija i tuberkuloza. Upravo zbog te neodlučnosti u skupu su ostavljene samo one slike za koje se najviše radiologa složilo za jednu bolest. Konačna raspodijela razreda između skupa za učenje i skupa za testiranje prikazana je na slici 4.4.



Slika 4.4. Raspodjela podataka skupa za učenje i testiranje

Potom su te snimke organizirane u tri direktorija, train, validate i test, pri čemu svaki ima internu strukturu kao što je prikazano na slici 4.5. koju očekuje korišteni ImageDataGenerator iz biblioteke tensorflow.

```
train/
...class1/
.....img1.jpg
.....img2.jpg
...class2/
.....img3.jpg
.....img4.jpg
...class3/
.....img5.jpg
.....img6.jpg
```

Slika 4.5. Očekivana struktura direktorija za ImageDataGenerator

Zatim su za svaki od skupova stvoreni objekti koji u model dovode slike odgovarajuće veličine i u odgovarajućim podskupovima kao što je prikazano na slici 4.6. za skup za učenje. Objekt normalizira piksele na vrijednosti $[0, 1]$, osigurava da su slike očekivane rezolucije specifične za kasnije korišteni model (parametar `targetSize`), podešava veličinu skupine slika koja se šalje odjednom (parametar `batchSize`) i sve te slike uzima iz putanje `train_dir` organizirane kao što je prethodno pokazano na slici 4.5.

```

train_datagen = ImageDataGenerator(rescale=1./255)

train_generator = train_datagen.flow_from_directory(
    train_dir,
    target_size=targetSize,
    batch_size=32,
    class_mode='categorical'
)

```

Slika 4.6. Stvaranje toka podataka iz direktorija za podskup train

Sljedeći korak je inicijalizacija i prevođenje samog modela, slika 4.7. i slika 4.8. Za početak stvara se novi objekt tipa modela koji je korišten (ResNet101 ili EfficientNetB4) i s njim se učitavaju težine bez zadnjeg sloja s prednaučenog skupa podataka imagenet. Zamrzavaju se sve težine početnog modela te se kao zadnji sloj podešava da predviđa vrijednosti za 5 različitih razreda.

```

base_model = ResNet101(weights='imagenet', include_top=False)
#EfficientNetB4(weights='imagenet', include_top=False, input_shape=(380, 380, 3))

for layer in base_model.layers:
    layer.trainable = False

x = base_model.output
x = GlobalAveragePooling2D()(x)
x = Dense(1024, activation='relu')(x)
predictions = Dense(5, activation='softmax')(x)

model = Model(inputs=base_model.input, outputs=predictions)

```

Slika 4.7. Inicijalizacija modela ResNet101, odnosno EfficientNetB4

Za prevođenje modela kao optimizator postavlja se Adam s nekom brzinom učenja te se također definiraju mjere uspješnosti koje će se pritom izračunavati.

```

model.compile(optimizer=Adam(learning_rate=0.001), loss='categorical_crossentropy', metrics=['accuracy'])

```

Slika 4.8. Prevođenje modela s optimizatorom

Kao dodatne funkcionalnosti mogu se definirati točke za spremanje trenutnog modela, no samo ako je bolji od prethodno spremljenog, te opcija ranijeg zaustavljanja koja

zaustavlja učenje ako kroz veći broj epoha nije došlo do poboljšanja rezultata na validacijskom skupu, slika 4.9.

```
checkpoint = ModelCheckpoint(path + 'model.keras', monitor='val_loss', save_best_only=True, mode='min')
early_stopping = EarlyStopping(monitor='val_loss', patience=10, mode='min')
```

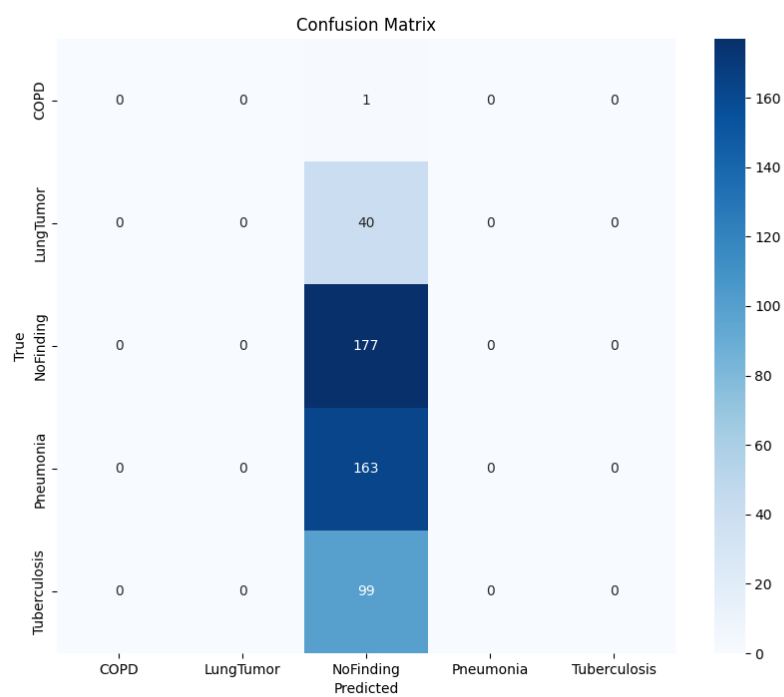
Slika 4.9. Podešavanje parametara za ranije prekidanje

Konačno slijedi samo učenje, slika 4.10., u metodi `fit` definiran je ulazni tok podataka za učenje i validaciju, broj epoha te broj koraka odnosno skupina podataka u jednoj epohi.

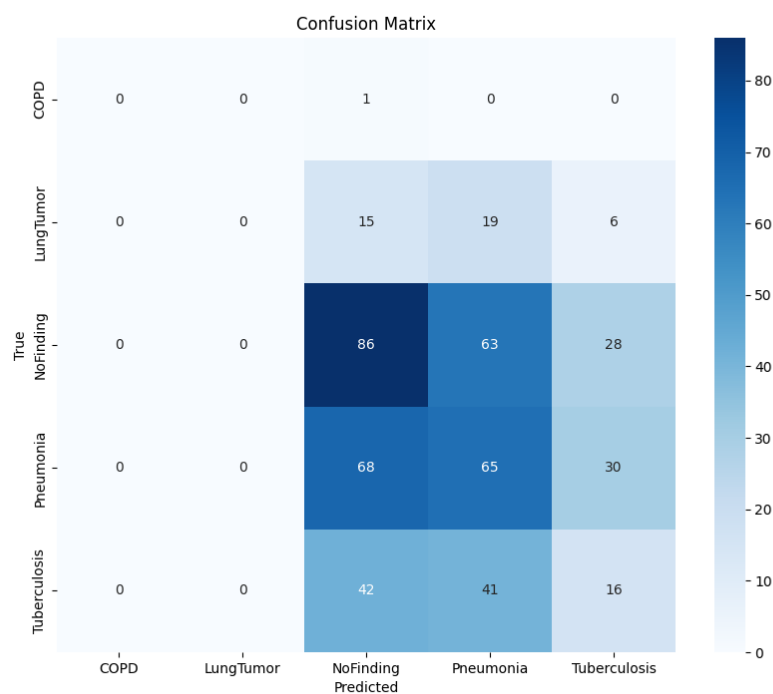
```
model.fit(
    train_generator,
    steps_per_epoch=train_generator.samples // train_generator.batch_size,
    validation_data=val_generator,
    validation_steps=val_generator.samples // val_generator.batch_size,
    callbacks=[checkpoint, early_stopping],
    epochs=50
)
```

Slika 4.10. Učenje modela

Nakon što je model naučen potrebno je evaluirati njegovu uspješnost na testnom skupu podataka naredbom `model.evaluate(test_generator)`, ali i pogledati matricu konfuzije. Cijeli postupak je proveden na dva tipa modela: EfficientNetB4 te ResNet101. Na testnim podacima EfficientNet ostvario je točnost od 36.88 % te gubitak od 1.3049, dok je ResNet zato postigao točnost od 50.42 % i gubitak od 1.1299.



(a) Model EfficientNet



(b) Model ResNet

Slika 4.11. Matrice konfuzije za oba modela

Sama preciznost to možda ne da naslutiti, no iz matrice konfuzije, slika 4.11., može se primijetiti da model EfficientNetB4 samo uvijek klasificira slike u kategoriju "No Findings" i nikad ne svrstava u ništa drugo, dok je model ResNet101 ipak uspio nešto postići iako i dalje nije najbolji. Loše rezultate moguće je objasniti velikom neuravnoteženošću u zastupljenosti između razreda pa modeli više naginju prema razredima koji se češće pojavljuju, ali i činjenicom da nije uvijek postojao konsenzus između doktora pa je teško odlučiti kako podijeliti te slike. Možda bi se moglo postići poboljšanje korištenjem više klasifikacija za jednu sliku. Time bi se svako neslaganje uzelo u obzir te bi se manje morao srezati početni skup podataka.

4.3. Detekcijski model

Za detekcijski model ponovno se trebalo odlučiti za vrstu pristupa podacima koji su više puta označeni. U tu svrhu proveden je algoritam koji prolazi kroz sve oznake za jednu sliku te smanjuje skup oznaka tako da traži one koje pripadaju istom razredu i postoji sjecište između njih i potom uzima presjek kao novu oznaku i miče prethodne dvije. Tako je riješen dio problema, no i dalje su postojale situacije gdje su radiolozi istu nepravilnost, odnosno isto područje na snimci, označili različitim razredima i takvi slučajevi pritom nisu pronađeni.

data.yaml	YAML Document	1 KB
images	Folder	--
train	Folder	--
000000000009.jpg	JPEG image	45 KB
000000000025.jpg	JPEG image	65 KB
000000000030.jpg	JPEG image	22 KB
000000000034.jpg	JPEG image	73 KB
val	Folder	--
000000000036.jpg	JPEG image	49 KB
000000000042.jpg	JPEG image	42 KB
000000000049.jpg	JPEG image	30 KB
000000000061.jpg	JPEG image	100 KB
labels	Folder	--
train	Folder	--
000000000009.txt	Plain Text	312 bytes
000000000025.txt	Plain Text	78 bytes
000000000030.txt	Plain Text	72 bytes
000000000034.txt	Plain Text	39 bytes
val	Folder	--
000000000036.txt	Plain Text	77 bytes
000000000042.txt	Plain Text	35 bytes
000000000049.txt	Plain Text	328 bytes
000000000061.txt	Plain Text	193 bytes

Slika 4.12. Očekivana struktura direktorija [20]

Konačne oznake i slike raspoređene su po direktorijima kao što je prikazano na slici 4.12., budući da je to struktura podataka koju očekuje model YOLO. Uz to modelu je potrebna konfiguracijska datoteka tipa YAML u kojoj se nalaze putanje do direktorija sa slikama te kodovi i imena svih razreda koji će se detektirati.

```
model = YOLO('yolov8l.pt')  
model.info()
```

Slika 4.13. Inicijalizacija detekcijskog modela

Model se inicijalizira kao objekt YOLO pri čemu mu se zadaje konkretna implementacija i razina kompleksnosti unutar konstruktora, slika 4.13. Nakon inicijalizacije odmah kreće učenje prikazano na slici 4.14. U funkciji za učenje modelu se proslijeđuje putanju do konfiguracijske YAML datoteke, veličina slika te broj epoha koje treba provesti.

```
results = model.train(data=path+'data.yaml', imgsz=640, epochs=30)
```

Slika 4.14. Učenje detekcijskog modela

Nakon što je model naučen evaluira se njegovu uspješnost na skupu podataka za validaciju i testiranje, slika 4.15. i slika 4.16. Iz rezultata je vidljivo da model radi gotovo prihvatljivo dobro za neke anomalije poput kardiomegalije, no većinu ipak ne uspijeva detektirati uopće.

```
results = model.val()
```

Class	Images	Instances	Box(P	R	mAP50	mAP50-95):
100%		78/78	[00:32<00:00, 2.42it/s]			
all	1242	1443	0.233	0.243	0.177	0.0694
Cardiomegaly	76	77	0.32	0.74	0.611	0.292
Lung Opacity	144	198	0.172	0.212	0.12	0.0508
Mediastinal shift	15	16	0.229	0.0625	0.123	0.0321
Nodule/Mass	84	169	0.312	0.296	0.227	0.0882
Atelectasis	28	32	0.245	0.219	0.13	0.0565
Infiltration	96	133	0.275	0.288	0.212	0.0783
Pulmonary fibrosis	160	296	0.236	0.145	0.0885	0.0278
Consolidation	57	66	0.135	0.182	0.12	0.0573
ILD	52	78	0.206	0.256	0.165	0.0563
Calcification	27	34	0.252	0.0588	0.0462	0.0141
Pleural thickening	139	233	0.168	0.192	0.0954	0.0234
Pleural effusion	87	111	0.25	0.261	0.19	0.0556

Slika 4.15. Rezultati modela na validacijskom skupu

```
results = model.val(split='test')
```

Class	Images	Instances	Box(P	R	mAP50	mAP50-95):
100%		78/78	[00:24<00:00, 3.24it/s]			
all	1244	1244	0.205	0.185	0.12	0.0411
Cardiomegaly	61	61	0.266	0.754	0.508	0.156
Lung Opacity	45	51	0.0393	0.176	0.0241	0.00897
Mediastinal shift	14	14	0.492	0.0714	0.12	0.046
Nodule/Mass	113	191	0.293	0.251	0.2	0.0644
Atelectasis	44	50	0	0	0.00273	0.00122
Infiltration	47	67	0.13	0.269	0.0765	0.0339
Pulmonary fibrosis	136	238	0.17	0.0924	0.0517	0.0152
Consolidation	83	108	0.137	0.113	0.0637	0.0235
ILD	131	190	0.378	0.216	0.221	0.0865
Calcification	57	69	0.311	0.029	0.0348	0.0143
Pleural thickening	84	121	0.0673	0.0826	0.0315	0.0101
Pleural effusion	69	84	0.174	0.167	0.11	0.0333

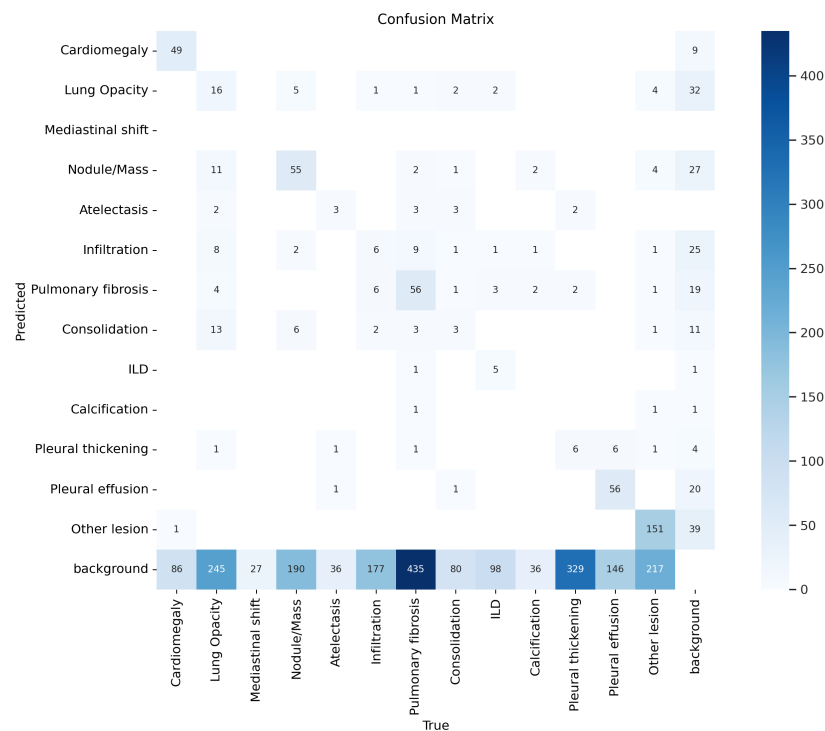
Slika 4.16. Rezultati modela na testnom skupu

Za usporedbu, naučen je još jedan model u kojem je, umjesto traženja presjeka da bi se pronašao zajednički zaključak između radiologa, za svakog liječnika koji je označavao sliku ponovno ta slika dodana u skup podataka s novom datotekom s oznakama. Također, u tom modelu oznake koje su se rjeđe pojavljivale umjesto da su u potpunosti zanemarene pribrojene su razredu za ostale anomalije (*other lesions*). Na slici 4.17. prikazane su matrice konfuzije za oba postojeća modela. Iz njih je vidljivo da je model u kojemu nisu računati presjeci oznaka liječnika radio bolje i detektirao više elemenata od

modela u kojem su presjeci izračunati.



(a) Model s presjecima oznaka



(b) Model bez presjeka oznaka

Slika 4.17. Matrice konfuzije za oba modela

Takve rezultate moguće je objasniti činjenicom da je određivanje presjeka dodatno smanjilo već neuravnoteženi skup podataka. Također, spomenuta činjenica da su različiti liječnici u različite razrede svrstali isto područje na slici što je onda na skupu s presjecima značilo da je isto područje dva puta označeno na jednoj slici, dok se u drugom skupu to nalazilo na dvije slike pa je sam model bio manje zbunjen. Bolje rezultate vjerojatno bi bilo moguće postići izbacivanjem još nekoliko manje zastupljenih anomalija čime bi skup podataka bio još bolje uravnotežen.

5. Zaključak

Rad prikazuje izradu modela za klasifikaciju nekih plućnih bolesti te modela za detekciju lokalnih oboljenja vidljivih na rendgenskim snimkama prsnog koša. Pojašnjena je struktura prsnog koša, neke plućne bolesti te sam rendgen prsnog koša. Također su objašnjeni osnovni pojmovi strojnog učenja i neuronskih mreža. Kao skup podataka korišten je VinDr-CXR, kolekcija rendgenskih snimki prsnog koša pripremljena u Vijetnamu [1].

U sklopu rada izgrađena su ukupno četiri modela, dva za klasifikaciju te dva za detekciju. Za klasifikaciju korišteni su prednaučeni modeli ResNet i EfficientNet, pri čemu je prvi uspio postići veću preciznost i nekakvu mogućnost raspoznavanja, dok je drugi u potpunoj nemogućnosti razlikovanja razreda na slikama. Za detekciju korišten je model YOLOv8, ali sa skupovima podataka prethodno obrađenima na dva različita načina. Slučaj kad su među radiološkim oznakama traženi presjeci radio je lošije nego kad su za svakog liječnika dodane dodatne slike u skup.

Do lošijih rezultata došlo je vjerojatno zbog velike neuravnoteženosti skupa podataka gdje su neka oboljenja jedva zastupljena te zbog činjenice da je skup podataka za učenje modela istovremeno označilo po troje liječnika bez konačnog jedinstvenog zaključka. Uzevši sve navedeno u obzir, može se zaključiti da je problem zaista težak budući da se u nekim slučajevima ni liječnici nisu mogli usuglasiti.

Literatura

- [1] H. Q. Nguyen, K. Lam, L. T. Le, H. H. Pham, D. Q. Tran, D. B. Nguyen, D. D. Le, C. M. Pham, H. T. T. Tong, D. H. Dinh, C. D. Do, L. T. Doan, C. N. Nguyen, B. T. Nguyen, Q. V. Nguyen, A. D. Hoang, H. N. Phan, A. T. Nguyen, P. H. Ho, D. T. Ngo, N. T. Nguyen, N. T. Nguyen, M. Dao, i V. Vu, “Vindr-cxr: An open dataset of chest x-rays with radiologist’s annotations”, *Scientific Data*, sv. 9, br. 1, str. 429, 2022. <https://doi.org/10.1038/s41597-022-01498-w>
- [2] R. Dezube i Željko Ivančević, “Prsni koš”, <https://hemed.hr/Default.aspx?sid=15911>, [posjećeno: lipanj 2024.].
- [3] —, “Prikaz dišnog sustava”, <https://hemed.hr/Default.aspx?sid=15910>, [posjećeno: lipanj 2024.].
- [4] L. Calvetti, <https://sciencephotogallery.com/featured/3-human-chest-anatomy-leonello-calvettiscience-photo-library.html>, [slika].
- [5] C. Clinic, “Pneumonia”, <https://my.clevelandclinic.org/health/diseases/4471-pneumonia>, [posjećeno: lipanj 2024.].
- [6] https://www.freepik.com/free-vector/lung-infection-pneumonia-with-explanation_39265220.htm, [slika].
- [7] C. Clinic, “Tuberculosis”, <https://my.clevelandclinic.org/health/diseases/11301-tuberculosis>, [posjećeno: lipanj 2024.].
- [8] —, “Lung cancer”, <https://my.clevelandclinic.org/health/diseases/4375-lung-cancer>, [posjećeno: lipanj 2024.].

- [9] NIH, “X-rays”, <https://www.nibib.nih.gov/science-education/science-topics/x-rays>, [posjećeno: lipanj 2024.].
- [10] R. Dezube i Željko Ivančević, “Snimanje prsnog koša”, <https://www.hemed.hr/Default.aspx?sid=15897>, [posjećeno: lipanj 2024.].
- [11] intelrad, “What is dicom image format and why is it important in radiology?” <https://www.intelerad.com/en/2023/02/23/handling-dicom-medical-imaging-data>, [posjećeno: lipanj 2024.].
- [12] E. Alpaydin, *Introduction to machine learning*. MIT press, 2020.
- [13] IBM, “What is machine learning?” <https://www.ibm.com/topics/machine-learning>, [posjećeno: lipanj 2024.].
- [14] —, “What are neural networks?” <https://www.ibm.com/topics/neural-networks>, [posjećeno: lipanj 2024.].
- [15] —, “What are convolutional neural networks?” <https://www.ibm.com/topics/convolutional-neural-networks>, [posjećeno: lipanj 2024.].
- [16] K. He, X. Zhang, S. Ren, i J. Sun, “Deep residual learning for image recognition”, *CoRR*, sv. abs/1512.03385, 2015. [Mrežno]. Adresa: <http://arxiv.org/abs/1512.03385>
- [17] T. Ahmed i N. Sabab, “Classification and understanding of cloud structures via satellite images with efficientnet”, 09 2020. <https://doi.org/10.1002/essoar.10507423.1>
- [18] A. Mirkhan, “Yolo algorithm: Real-time object detection from a to z”, <https://kili-technology.com/data-labeling/machine-learning/yolo-algorithm-real-time-object-detection-from-a-to-z#what-is-yolo?>, [posjećeno: lipanj 2024.].
- [19] Zh, G. Omarova, i V. Starovoitov, “Application of the clahe method contrast enhancement of x-ray images”, *International Journal of Advanced Computer Science and Applications*, sv. 13, str. 412–420, 06 2022. <https://doi.org/10.14569/IJACSA.2022.0130549>

- [20] G. Jocher, A. Chaurasia, i J. Qiu, “Ultralytics yolov8”, 2023. [Mrežno]. Adresa:
<https://github.com/ultralytics/ultralytics>

Sažetak

Klasifikacija bolesti temeljem rendgenskih snimki prsnog koša korištenjem metoda strojnog učenja

Andrea Milanović

Ovaj rad prikazuje primjenu konvolucijskih neuronskih mreža za klasifikaciju i detekciju bolesti temeljem rendgenskih snimki prsnog koša. Prikazana je izgradnja dva klasifikacijska te dva detekcijska modela i uspoređena je njihova uspješnost. Analizirane su metrike točnosti, preciznosti i konfuzijske matrice za izgrađene modele. Za izradu rada korišten je programski jezik Python te modeli ResNet, EfficientNet i YOLOv8. Model ResNet postigao je točnost od 50 %. Model YOLO ukupno postiže preciznost od 20.5 % i odziv od 18.5 %, no za neka oboljenja poput kardiomegalije te vrijednosti prelaze 70 %.

Ključne riječi: strojno učenje, konvolucijske neuronske mreže, klasifikacija slika, detekcija na slikama, rendgen prsnog koša

Abstract

Classification of diseases based on chest X-ray images using machine learning methods

Andrea Milanović

This paper presents the application of convolutional neural networks for the classification and detection of diseases based on chest X-ray images. The construction of two classification and two detection models is demonstrated, and their performance is compared. Accuracy metrics, precision metrics and confusion matrices are analyzed for the constructed models. Python programming language was used for the development of the paper, along with ResNet, EfficientNet, and YOLOv8 models. The ResNet model achieved an accuracy of 50 %. The YOLO model achieved an overall precision of 20.5 % and a recall of 18.5 %, but for some diseases such as cardiomegaly, these values exceed 70 %.

Keywords: machine learning, convolutional neural networks, image classification, image detection, chest X-ray