

Segmentarea și clasificarea imaginilor medicale

Huminiuc Simona

Universitatea "Gheorghe Asachi" Iași

Facultatea de Automatică și Calculatoare

1. Abstract

În acest studiu de caz este prezentată o metodă de segmentare și clasificare a imaginilor medicale, în special imagini X-Ray ale plămânilor sănătoși și bolnavi. În procesul de segmentarea s-au folosit diferite metode cum ar fi următoarele: thresholding, detectarea conturului, operații morfologice precum deschidere, eroziune pentru a izola secțiunea plămânilor de restul imaginii și pentru o vizualizare mai bună a acestora. De asemenea sunt utilizate măștile pentru a identifica "casete de încadrare" (eng. bounding box). Codul a fost implementat în Python utilizând librăriile OpenCV și Numpy pentru o procesare mai eficientă.

Pentru clasificare se utilizează deviația standard pentru a calcula contrastul și dimensiunea zonelor cu pixeli albi segmentați. Imaginile sunt clasificate ca sănătos, bolnav sau nedefinit în funcție de rezultatele obținute.

2. Introducere

Pneumonia este o inflamație a țesutului pulmonar cauzată de bacterii, virusuri sau alte microorganisme. Cauza infecției, vârsta și starea generală de sănătate a pacientului poate determina cât de sever îl poate afecta și dacă va fi afectat un singur plămân sau ambii. Câteva simptome ale acestei boli sunt: tușea productivă, febră și frisoane, durere toracică și dificultate la respirație, oboseala.

SARS-CoV-2 cauzează COVID-19 și produce infecții respiratorii cu grade de severitate de la mediu la sever, putând duce inclusiv la deces. Printre cele mai frecvente simptome se numără: febră, tuse uscată, oboseala, dureri de gât, pierderea simțului gustului și a mirosului.

Aceste două boli sunt din ce mai întâlnite, în special în sezonul rece și se manifestă foarte asemănător. Identificarea și tratarea lor în timp util poate face diferența, mai ales în cazuri în care pacienții suferă și de afecțiuni cronice.

Depistarea precoce a cazurilor în primele etape ale ar putea ajuta salvarea multor vieți iar analizarea imaginilor poate dura timp îndelungat. Procesarea imaginilor vine în ajutorul medicilor radiologi.

3. Metode

a. Setul de date

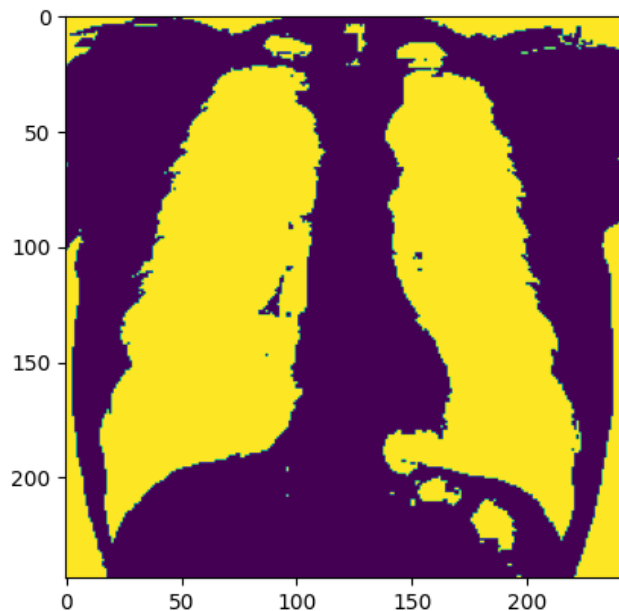
În realizarea acestui proiect am avut nevoie de mai mult seturi de date: plămâni sănătoși [1], covid [2], pneumonie [3] Acest set de date este împărțit în trei categorii pentru a identifica și diferenția plămânii sănătoși și cei afectați de o boală mai bine. Setul de date are un număr foarte mare de imagini.

b. Preprocesare

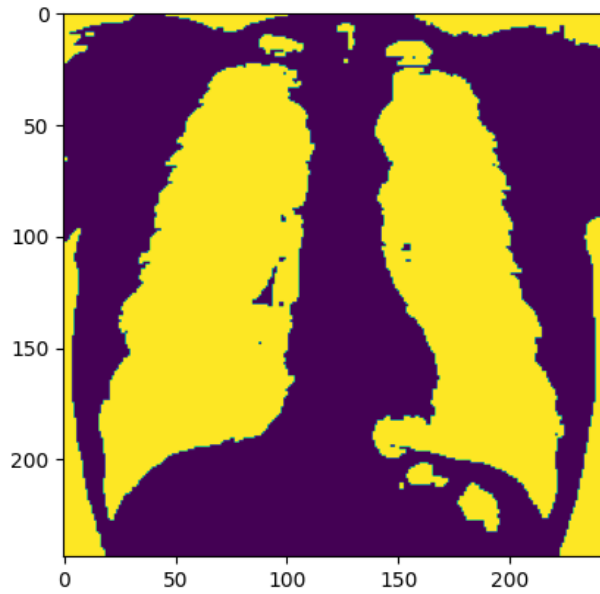
Pentru că algoritmul să aibă o eficiență mai mare și să funcționeze mai rapid, toate imaginile din setul de date au fost redimensionate pentru că toate să respecte aceeași dimensiune 244x244 pixeli. Am utilizat o funcție separată care ia imaginile dintr-un fișier, le procesează și după executare, le avem dispuse într-un fișier nou pentru a nu fi amestecate cu datele originile. În continuare am realizat o funcție pentru vizualizarea histogramei imaginii inițiale pentru a constată diferențe ale nivelurilor de gri ale plămânilor sănătoși și bolnavi. Acest aspect ne va ajuta mai târziu și la partea de clasificare.

c. Segmentare

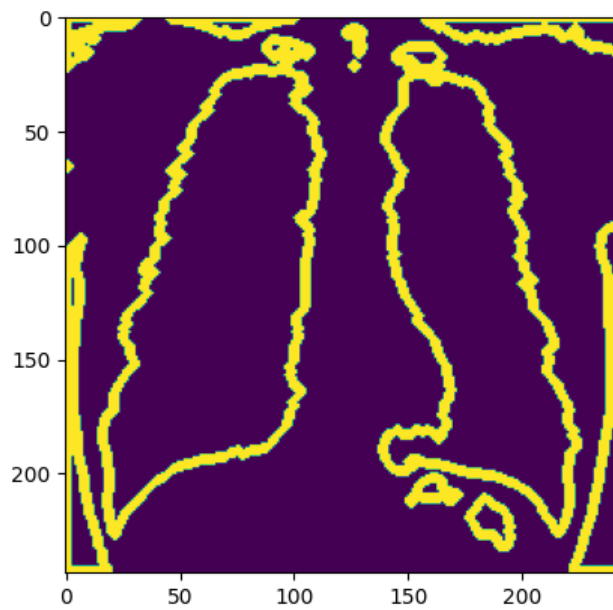
Procesul de segmentare al plămânilor include folosirea diferitor funcții din OpenCV. În primul rând, am utilizat thresholding adaptiv pentru binarizare, astfel obține imagini doar în alb și negru. Metoda presupune înlocuirea fiecărui pixel din imaginea originală cu un pixel negru dacă intensitatea lui este mai mică decât valoarea de thresholding ori cu un pixel alb dacă intensitatea este mai mare. Pentru că metoda este adaptivă, se calculează valoare de thresholding pe porțiuni mai mici din imagine, astfel obținem diferite valori pentru regiuni diferite. Astfel rezultatele sunt mai bune pentru ca imaginile sunt iluminate diferit. S-a încercat și abordarea cu valoare de thresholding fixă pentru tot setul de imagini, dar s-a constatat că o metoda adaptivă are rezultate mai bune

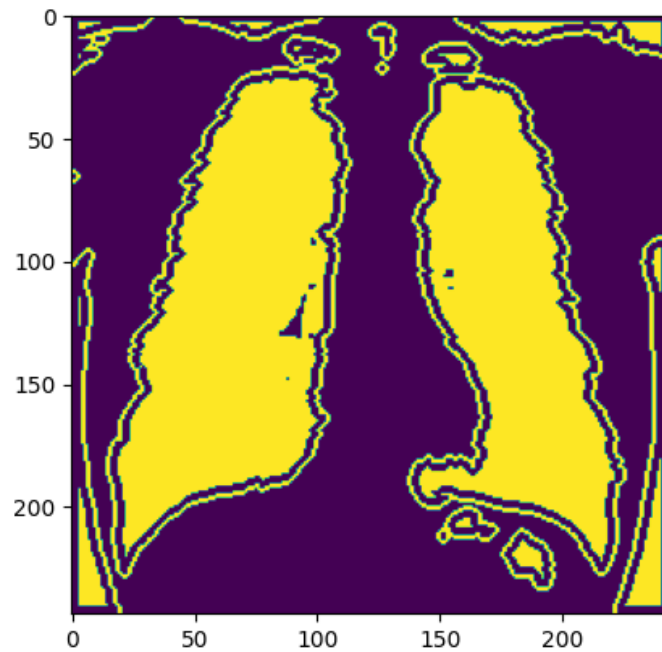


În continuare am utilizat câteva funcții de dilatare, eroziune și deschidere pentru a reduce zgomotul din imagine.

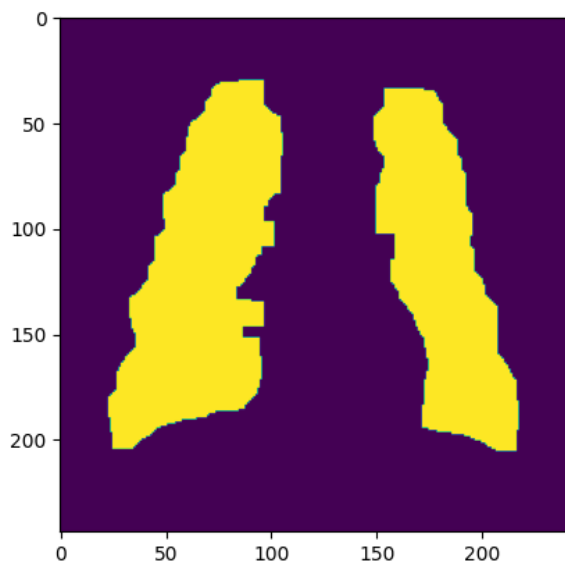


Apoi se creează o mască pentru contur, aplicându-l ulterior peste imaginea binarizată. Acest pas este necesar pentru o mai bună delimitare a zonei de interes de restul obiectelor din imagine



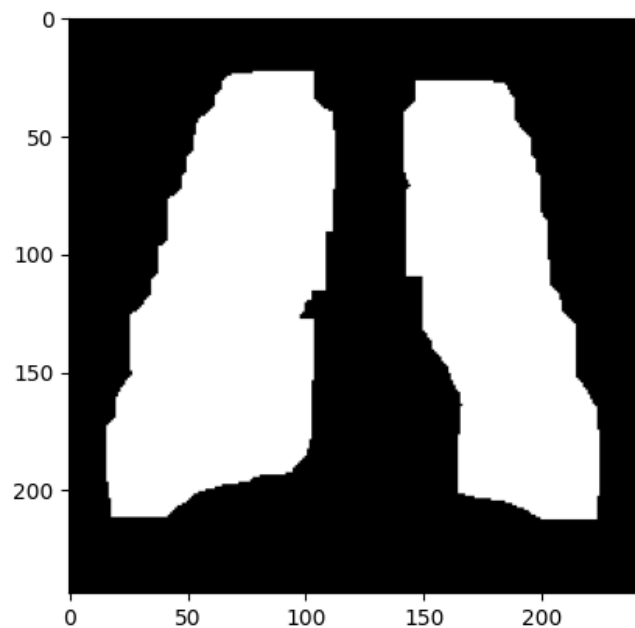
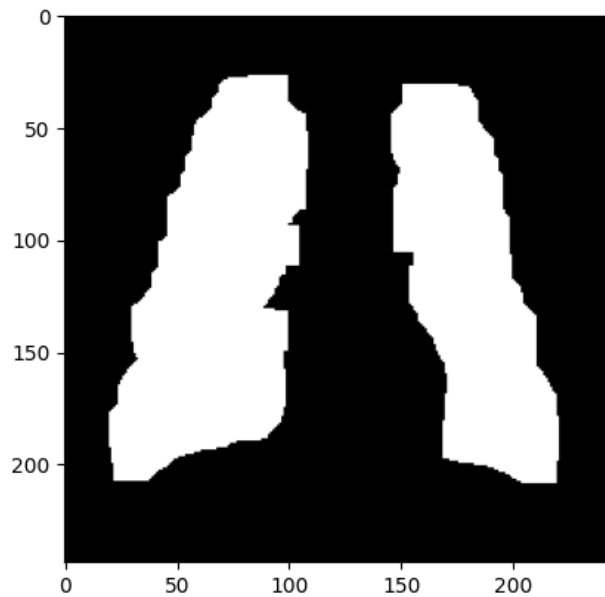


În următorul pas am eliminarea zonele din imagini care sunt prea apropiate de colțurile acesteia astfel: am împărțit imaginea în două zone egale prin axa OY, stânga și dreapta. Fiecare regiune este etichetată cu ajutorul "bounding box". Dacă se găsește o etichetă la jumătatea distanței pe fiecare dintre aceste două regiuni atunci va fi luată în considerare. În general plămânii ocupa un spațiu mare în imagine și sunt poziționați în paralel pe verticală.



Pentru că s-au segmentat plămânii de restul imaginii, în interiorul lor au rămas câteva "petice" (rămășițe de la oase sau bucăți care s-au pierdut în la thresholding) care trebuie

completate pentru că plămânii să fie cât mai întregi cu ajutorul operațiilor morfologice menționate și mai sus.



d. Clasificare

În primul rând am analizat histogramele a mai multor imagini din ambele categorii și am constatat că setul de date cu plămâni sănătoși are ușor un contrast mai ridicat decât de celelalte

deoarece plămânii sunt mai "întregi" și am un nivel de gri mai intens față de ceilalți. În continuare am analizat mărimea plămânilor și am constatat că după procesul de segmentare, plămânii sănătoși sunt mai mari și au mult mai mulți pixeli albi. Deci după aceste trei criterii am realizat o clasificare între plămâni sănătoși/normali și plămâni afectați de pneumonie sau covid-19.

4. Rezultate și concluzii

În urma tuturor procesărilor și analizei de clasificare rezultatele sunt satisfăcătoare pentru procesele prin care au fost puse imaginile și seturile de date de care am dispus. În viitor se va încerca și introducerea unui model de învățare pentru predicție mai bună și eficientă.

5. Referințe

1. <https://data.mendeley.com/datasets/dvntn9yhd2/1>
2. <https://github.com/ieee8023/covid-chestxray-dataset/tree/master/images>
3. <https://www.sciencedirect.com/science/article/pii/S001048252200258X#bib51>
4. <https://www.sciencedirect.com/science/article/pii/S001048252200258X#bib52>
5. <https://www.nature.com/articles/s41598-022-15013-z>
6. <https://www.nature.com/articles/s41597-023-02229-5>