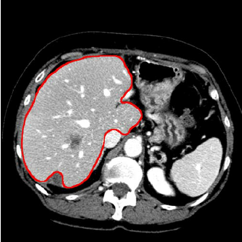
Problema AI Engineer

# Descriere

Echipa noastra lucreaza la dezvoltarea unor produse in zona medicala. Una dintre problemele cu care inginerii nostri AI se confrunta acum este imbunatatirea segmentarii imaginilor DICOM in care, medicii cu care lucram, delimiteaza organele.

Ei lucreaza acum pe imagini DICOM, iar cu ajutorul mouse-ului sau al unui stylus deseneaza conturul acestor organe. Segmentarile si metadatele pe care medicii le furnizeaza sunt folosite pentru a antrena retelele neurale.

Segmentarile pot fi imbunatatite pentru a fi mai precise folosind un algoritm care sa proceseze delimitarea trasata de doctor.

Fig. 1 – Segmentare ficat

Avem nevoie de un **super inginer cu cunostinte de AI,** **ca tine,** care sa ne ajute sa dezvoltam acest algoritm folosind Java sau Python. Daca ai vrea sa abordezi problema folosind alta tehnologie, te rugam sa confirmi cu noi inainte sa te apuci de rezolvare.

# Detalii de implementare

Algoritmul va fi folosit pentru optimizarea segmentarilor pentru imaginile ce compun o investigatie CT abdominala. Fiecare pixel dintr-o imagine este masurat in [unitati Hounsfield](https://en.wikipedia.org/wiki/Hounsfield_scale) ce reprezinta densitatea tesutului masurata de aparatul CT. Mai multe informatii despre cum este folosit in radiologie [aici](https://www.ncbi.nlm.nih.gov/books/NBK547721/).

Organele segmentate de doctori sunt compuse din pixeli care au valori Hounsfield intr-un registru bine definit, dar difera de la organ la organ. Din cauza ca aparatele CT sunt reglate diferit de la o clinica la alta, registru pentru acelasi organ poate sa difere de la un CT la altul (ex. pixelii ce compun ficatul intr-un CT pot avea valori intre [59, 78], iar in alt CT pot avea [67, 87]). Dar, organele segmentate au in jurul lor tesut care masoara valori Hounsfield semnificativ diferite.

A screenshot of a cell phone

Description automatically generated

Fig. – Scala Hounsfield (CT image)

Aplicatia va primi doua seturi de date de intrare:

* O matrice N x N in care fiecare intrare reprezinta valoarea hounsfield pentru pixel-ul corespunzator din imaginea DICOM CT
* O matrice N x N cu valori 0 si 1 in care sunt marcati cu 1 pixelii alesi de medic pentru conturul unui organ

Fisierele ce contin cele doua seturi de date vor fi date ca parametrii la rularea aplicatiei (eg. “optimise 107-HU.in 107-seg.in”), iar formatul acestora este descris in anexa.

Dupa rulare, aplicatia trebuie sa scrie rezultatul (matricea care contine segmentarea optimizata) in fisierul de iesire cu numele “**optim.out**”.

Segmentarea optimizata este cat mai aproape de organul pe care il delimiteaza

* Exclude zone care sunt in afara organului, ce delimiteaza alte tipuri de tesut
* Corecteaza erorile in care cursorul medicului a taiat parte din organ si acea parte nu este inclusa in selectia curenta

# Super bonus

Aplicatia poate sa lucreze pe seturi de date ce contin segmentari de mai multe organe.

# Recomandari

Gandeste structura si codul aplicatiei ca si cand urmeaza sa dezvolte la ea si alti colegi de-ai tai. Iti doresti sa fie un algoritm care va fi imbunatatit timp de 2 ani si va fi folosit mai departe in aplicatii care vor oferi feedback intr-un timp cat mai scurt medicului pentru a putea valida segmentarea optimizata.

Vrem sa scrii cod de care esti foarte mandra/mandru si nu crezi ca mai poate fi imbunatatit :)

Ideal ar fi sa ne scrii si un README in care sa descrii solutia abordata, ce probleme ai intalnit si cum le-ai rezolvat. Vrem sa ne motivezi de ce ai rezolvat-o asa.

Te rugam sa lucrezi sub un repository private de git pe platforma [gitlab.com](https://gitlab.com/), sa faci commit-uri granulare si dese. Cand ai terminat, te rugam sa inviti pe gitlab userul @AB4Systems la repository-ul tau si sa oferi un acces de tip “Developer”. De asemenea, te rugam sa ne dai si un email in care sa ne anunti ca putem face review.

# Ce vom urmari

1. Inventivitatea solutiei – ce algoritm ai dezvoltat si cat de bine adreseaza problema
2. Eficienta si fiabilitate solutiei propuse – **Atentie: nu ne asteptam sa obtii rezultate identice cu noi, in unele aspecte este posibil segmentarea ta sa exceleze, in altele nu**
3. Aplicatia va fi testata automat cu seturi de intrare si iesire; vom compara rezultatul aplicatiei tale cu o segmentarea optima
4. Claritatea codului: cat de usor e de inteles, cat de ordonat este scris (pentru restu exista lint checkere ☺)
5. Timpul de rulare (maxim 1 minut per test)

Anexa

# Format date intrare – Matrice imagine CT

* Matrice 512 x 512 cu valorile hounsfield pentru fiecare pixel din imagine (valori intre -3024 si 3071)

**Exemplu:**

* Fisierul “**107-HU.in**” din directorul “**input**” contine matricea 512x512 pentru un CT abdominal in care medicul urmeaza sa segmenteaze ficatul pacientului
* A picture containing object, black

  Description automatically generatedFisierul “**107-HU.png**” din directorul “**input**” modul in care imaginea CT este afisata de un viewer DICOM (nu trebuie generat acest fisier, a fost adaugat pentru a putea intelege mai bine problema)

Figure 3 – 107-HU.png

# Format date intrare – Matrice segmentare facuta de medic

* Matrice 512 x 512 cu valori de 0 sau 1

**Exemplu:**

* Fisierul “**107-seg.in**” din directorul “**input**” contine matricea 512x512 pentru sementarea ficatului realizata de medic
* Fisierul “**107-seg.png**” din directorul “**input**” modul in care segmentarea ficatului poate fi vizualizata

(nu trebuie generat acest fisier, a fost adaugat pentru a putea intelege mai bine problema)

A close up of a logo

Description automatically generated

Figure 4 - 107-seg.png

# Format date iesire – Matrice segmentare optimizata

* Matrice 512 x 512 cu valori de 0 sau 1

**Exemplu:**

* Fisierul “**107-opt.out**” din directorul “**input**” contine matricea 512x512 cu segmentarea optimizata de algoritm
* Fisierul “**107-opt.png**” din directorul “**input**” modul in care segmentarea ficatului poate fi vizualizata

(nu trebuie generat acest fisier, a fost adaugat pentru a putea intelege mai bine problema)

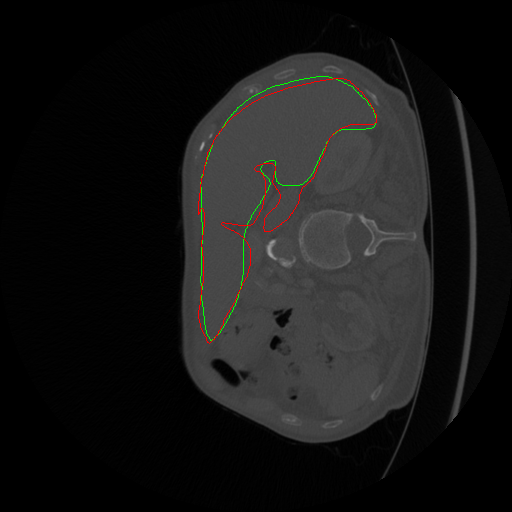
**A silhouette of a person

Description automatically generated**

Figure 5 - 107-opt.png

# Extra

Pentru a intelege diferentele intre cele doua segmentari, am inclus si o imagine care prezinta rezultatele algoritmului in comparative cu segmentarea facuta de medic. Cu rosu este prezentata segmentarea medicului, iar cu verde segmentarea optimizata de catre algoritm. (imaginea “107-compare.png”).

****

# Alte exemple

**In celelalte directoare “input2->input4”, vei gasi alte 3 exemple de date de intrare in acelasi format.**