

BABEȘ BOLYAI UNIVERSITY, CLUJ NAPOCA, ROMÂNIA  
FACULTY OF MATHEMATICS AND COMPUTER SCIENCE

# DETECTAREA RISCULUI DE RĂSPÂNDIRE A CANCERULUI DE PROSTATĂ

PE BAZA GRUPEI DE SCOR GLEASON (ISUP GRADE) DIN BIOPSIILE DE  
PROSTATĂ

– MIRPR report –

## Membrii echipei

Păpară Ana - Maria, Informatică, grupa 235

Pop Maria - Iulia, Informatică, grupa 235

2022-2023

# Contents

<b>1</b>	<b>Introducere</b>	<b>1</b>
1.1	De ce? Ce? Cum? . . . . .	1
<b>2</b>	<b>Problema științifică</b>	<b>3</b>
2.1	Definirea problemei . . . . .	3
<b>3</b>	<b>Alte publicații asemenea</b>	<b>5</b>
<b>4</b>	<b>Abordarea aleasă</b>	<b>6</b>
<b>5</b>	<b>Aplicația</b>	<b>7</b>
5.1	Descrierea aplicației și funcționalitățile principale . . . . .	7
5.2	Implementare . . . . .	7
5.3	Testarea codului sursă . . . . .	8
5.4	Numerical validation . . . . .	8
5.4.1	Data . . . . .	8
<b>6</b>	<b>Concluzie și dezvoltare</b>	<b>10</b>
<b>7</b>	<b>Bibliografie</b>	<b>11</b>

# List of Tables

5.1	Loss-ul și acuratețea în funcție de nr de epoci . . . . .	8
-----	---	---

# List of Figures

1.1	Prostata . . . . .	1
1.2	Prelevarea de țesut suspect din prostată . . . . .	2
2.1	Scor și grupe Gleason . . . . .	4
4.1	Procesarea imaginii . . . . .	6
5.1	Flow Diagram . . . . .	7
5.2	Layers . . . . .	8
5.3	Setul de date . . . . .	8

# Chapter 1

## Introducere

### 1.1 De ce? Ce? Cum?

Incidența afecțiunii cancer de prostată este în creștere continuă, fiind a doua cauză de cancer și a patra cauză de deces prin cancer, la bărbat. Reprezintă aproximativ 15% din totalul cancerelor diagnosticate, cu variații geografice și rasiale foarte mari. La nivel global, ocupă locul 4 ca incidență (1276106 cazuri noi în 2018, 7,1% din total - 14,2% la bărbat) și locul 9, la nivel de mortalitate (358,989 decese în 2018, 3,8% din total - 7,6% la bărbat).

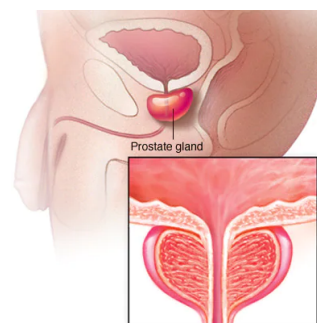
În România, au fost diagnosticate peste 6,000 cazuri noi (6029) și au fost documentate aproximativ 2,500 decese (2479) în anul 2018.

Cancerul de prostată, întâlnit în literatura de specialitate sub denumirea de adenocarcinom de prostată, se dezvoltă lent, însă, există și cazuri în care cancerul de prostată evoluează agresiv. Celulele tumorale se pot extinde prin procesul de metastazare, la oase și ganglioni limfatici de vecinătate.

Cancerul de prostata indica un neoplasm malign la nivelul prostatei, adică o creștere anormală a țesutului cu riscul ca tumoarea să se extindă și la alte țesuturi învecinate. Pentru detectarea cancerului la prostată, pacientul va trebui să realizeze mai multe analize. Dintre acestea enumerăm: testul PSA (test de sânge cu antigen), examinarea rectală digitală (DRE - Digital Rectal Examination), biopsia, imagistica prin rezonanță magnetică nucleară (RMN).

Biopsia este procedura prin care se prelevează o bucată de țesut pentru a determina prezența sau stadiul unei boli. Biopsia de prostată înseamnă prelevarea unei bucăți suspecte de țesut din prostată. Prostata este o glandă aflată la intersecția aparatului urinar cu aparatul genital masculin. Aceasta este situată în fața rectului și sub vezica urinară. Aceasta contribuie, prin secrețiile sale, cu 25-30% din cantitatea de lichid spermatic.

Figure 1.1: Prostata



© MAYO FOUNDATION FOR MEDICAL EDUCATION AND RESEARCH. ALL RIGHTS RESERVED.

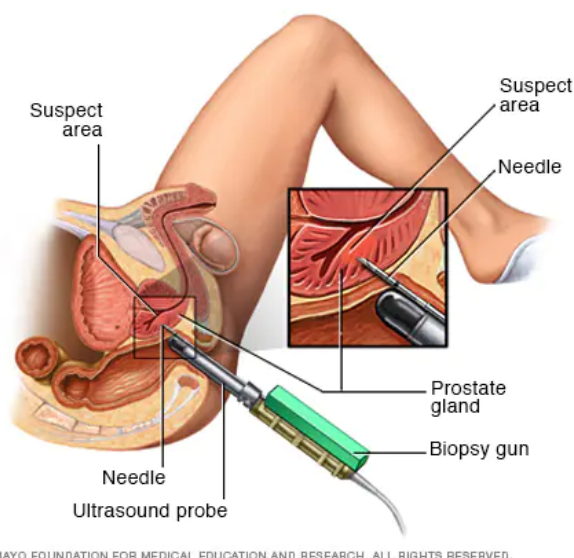


Figure 1.2: Prelevarea de țesut suspect din prostată

În timp ce stadiul cancerului ne prezintă cât de răspândit este cancerul în organism și cum se comportă acesta la nivel macro, gradul lui ne arată cum celulele canceroase se comportă sub microscop.

Este important să cunoaștem gradul cancerului, deoarece există riscul răspândirii acestuia în organism, iar determinarea riscului cu care acesta poate avansa conduce la aplicarea unui tratament corect asupra pacientului. Pentru a putea realiza acest lucru avem nevoie de analiza patologiilor pe baza unei biopsii a prostatei.

În cazul biopsiilor, rezultatul lor este unul supus subiectivității, fiind influențat de experiența pe care o are medicul patolog. Considerăm că intervenția unei mașini automate de învățare este necesară, întrucât poate preveni tratarea superficială a bolii sau exagerarea în tratament.

## Chapter 2

# Problema științifică

### 2.1 Definirea problemei

Biopsia ne ajută să determinăm gradul de răspândire al celulelor canceroase. Astfel există 3 grupuri de risc:

- risc scăzut: Tumoră în stadiu incipient și este puțin probabil să crească sau să se răspândească la alte țesuturi sau organe timp de mai mulți ani.
- risc mediu (favorabil/nefavorabil): Acest grup include cele mai multe tipuri de cancer de prostată. De obicei, cancerul este puțin probabil să se răspândească sau să crească timp de câțiva ani, dar tratamentul poate fi recomandat
- risc ridicat: Tumoră mai avansată și cu risc ridicat de răspândire agresivă

Pentru a putea determina categoria de risc ne vom folosi de scorul Gleason. Acesta cuprinde valori de la 2 la 10. Practic patologul care va analiza exemplele prelevate de țesut din prostată va acorda o notă celui mai predominant tipar din biopsie, și o a doua notă celui de-al doilea cel mai răspândit tipar. El va acorda nota de la 1 la 5 fiecărei probe astfel:

Nota 1 - Celule mici, uniforme, ambalate împreună, care seamănă cu celule normale ale prostatei

Nota 2 - Celule ambalate ușor de diferite forme

Nota 3 - Nereguli suplimentare în dimensiunea și forma celulelor, inclusiv fuziunea celulară

Nota 4 - Celule mari, anormale, topite între ele

Nota 5 - Celule mari, neregulate, s-au contopit și au început să invadeze țesutul conjunctiv din jur

Suma celor două note va fi de fapt scorul Gleason. Un scor mai mic sau egal cu 6 indică prezența unei tumori însă aceasta este într-un stadiu incipient, un scor de 7 prezintă un risc mediu, iar un scor

între 8 și 10 indică prezența unei tumori cu risc mare de răspândire.

În anul 2014, Societatea Internațională de Patologie Urologică (ISUP) au emis o îndrumare în urma căreia au grupat scorurile Gleason în cinci grupe. Acestea poartă numele de grupe de scor ISUP, scor ISUP sau grupe de scor Gleason. În grupa 1 se încadrează scorul Gleason mai mic sau egal cu 6. În grupa 2 se încadrează scorul Gleason 7, 3+4, acesta indică un risc mediu, însă favorabil. În timp ce grupa 3, tot cu scorul Gleason 7, însă 4 + 3, indică un risc mediu, dar nefavorabil, prezicând o creștere moderată a tumorii. În cele din urmă scorul Gleason 8 se încadrează în grupa 4, iar scorul Gleason 9-10 în grupa 5, tumoarea prezentând un risc foarte mare de creștere.

Gleason score	Grade Group	What it means
Gleason score 6 (or 3 + 3 = 6)	Grade Group 1	The cells look similar to normal prostate cells. The cancer is likely to grow very slowly, if at all
Gleason score 7 (or 3 + 4 = 7)	Grade Group 2	Most cells still look similar to normal prostate cells. The cancer is likely to grow slowly
Gleason score 7 (or 4 + 3 = 7)	Grade Group 3	The cells look less like normal prostate cells. The cancer is likely to grow at a moderate rate
Gleason score 8 (or 4 + 4 = 8)	Grade Group 4	Some cells look abnormal. The cancer might grow quickly or at a moderate rate
Gleason score 9 or 10 (or 4 + 5 = 9, 5 + 4 = 9 or 5 + 5 = 10)	Grade Group 5	The cells look very abnormal. The cancer is likely to grow quickly

Figure 2.1: Scor și grupe Gleason

Intervenția subiectivității omului prezintă un factor de influență în aflarea rezultatului unei biopsii. Întrucât analiza se face la microscop de către cadrul medical, experiența și pregătirea acestuia sunt foarte importante. De aceea am considerat că antrenarea unui model de inteligență artificială ar fi binevenit în întâmpinarea acestor probleme. Analizarea imaginilor microscopice de către model ar conduce la diminuarea greșelilor umane.

Probleme intervin des în încadrarea unei biopsii în grupa 2 sau 3, cu scorul Gleason 7. Așadar noi ne propunem să antrenăm un model care poate prezice grupa ISUP în care se încadrează o biopsie.



## Chapter 3

### Alte publicații asemenea

În anul 2020, Radboud University Medical Center și Karolinska Institute au lansat o competiție numită Panda Challenge (Prostate cANcer graDe Assemssment) prin care cereau dezvoltarea unui model de învățare automat pentru diagnosticarea cancerului la prostată. Setul de date a fost pus la dispoziție de aceștia, și constau în imagini ale unor biopsii.

## Chapter 4

### Abordarea aleasă

Pentru problema prezentată am folosit setul de date pus la dispoziție, pe care l-am împărțit în date de antrenament și date pentru testare. Întrucât dimensiunea imaginilor era foarte mare, am optat pentru redimensionarea lor. Am prelucrat label-urile astfel, întrucât pe noi ne interesa doar informația legată de grupa Gleason, adică scorul ISUP. A urmat apoi contruirea și antrenarea modelului, ca în cele din urmă să aibă loc testarea și evaluarea acestuia.

Descriere pseudocod:

Algorithm:

Get\_data

Split\_data

Process\_data

Get\_labels

Constructing model

Train\_model

Test + Evaluate model



Figure 4.1: Procesarea imaginii

# Chapter 5

## Aplicația

### 5.1 Descrierea aplicației și funcționalitățile principale

Predictia grupei de risc în care se află pacientul.

Precizarea caracterului malign sau benign al tumorii.

### 5.2 Implementare

Pentru implementare am folosit limbajul python, întrucât oferă o largă gamă de librării ce sunt utile în antrenarea unor sisteme automate de învățare.

Din librăriile folosite amintim: keras, tensorflow, panda, numpy, PIL. Pentru realizarea modelului am folosit Sequential din keras. Funcția loss de activare folosită a fost crossentropy din libraria keras. Pentru optimizarea modelului am folosit Adadelta din libraria keras.

Modelul este compus din două convoluții 2D, un strat pooling 2D, unul dropout și unul flatten, iar în final două straturi dense separate de unul de tip dropout. În afară de ultimul strat care are ca funcție de activare softmax, restul straturilor din model au funcția de activare relu.

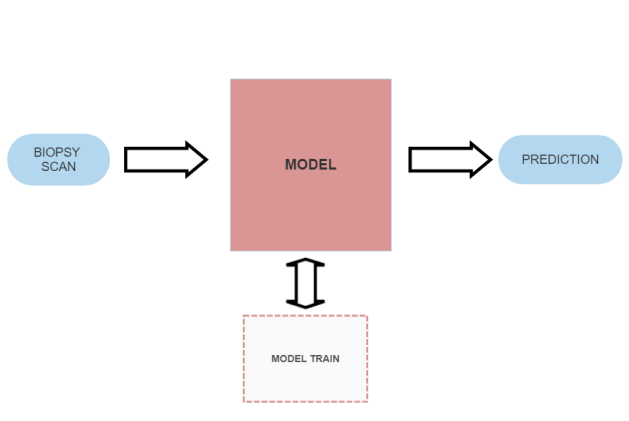


Figure 5.1: Flow Diagram

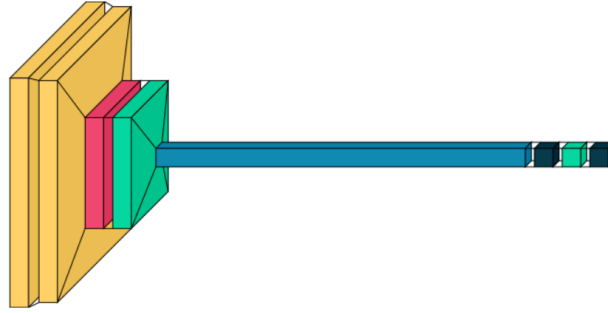


Figure 5.2: Layers

### 5.3 Testarea codului sursă

Pentru testare am folosit un set de imagini separat anterior din setul inițial de imagini. Aceste imagini nu au fost folosite și la antrenare. Selecția a fost făcută astfel încât să aibă o distribuție echilibrată a tipurilor de scor ISUP.

### 5.4 Numerical validation

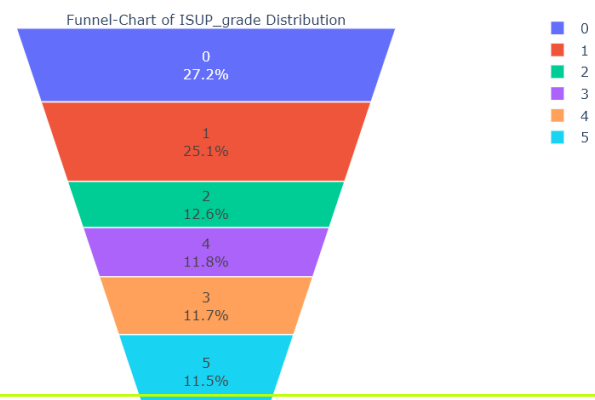
Table 5.1: Loss-ul și acuratețea în funcție de nr de epoci

Epoci	Loss	Acuratete
2	8.100149154663086	0.6000000238418579
5	0.6931336522102356	0.6000000238418579
15	0.6931336522102356	0.6000000238418579

#### 5.4.1 Data

Am folosit setul de date pus la dispoziție de Radboud University Medical Center și Karolinska Institute care conținea aproximativ 11 000 de imagini cu biopsii.

Se poate observa că pacienții cu ISUP Grade 0 și 1 sunt cei mai bine reprezentați, însă aceștia sunt pacienții care nu prezintă cancer. Clasa țintă este destul de slab reprezentată, însă asta se întâmplă des pe datele medicale. Cu toate acestea situația reprezintă și una dintre provocările pe



---

care le aducem modelului de învățare automată.

## Chapter 6

# Concluzie și dezvoltare

Modelul reușește să încadreze în categoria malign sau benign tumoarea pe baza scorului ISUP în care se încadrează biopsia.

Pe viitor ar fi de luat în considerare folosirea scorului Gleason pentru determinarea automată a grupei ISUP și gradul de risc al răspândirii tumorii canceroase.

## Chapter 7

# Bibliografie

1. [<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7837533/>](<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7837533/>) - National Library of Medicine: A Systematic Review of Artificial Intelligence in Prostate Cancer
2. [<https://www.pcf.org/about-prostate-cancer/diagnosis-staging-prostate-cancer/gleason-score-isup-grade/>](<https://www.pcf.org/about-prostate-cancer/diagnosis-staging-prostate-cancer/gleason-score-isup-grade/>) - Prostate Cancer Foundation
3. [<https://www.cancerresearchuk.org/about-cancer/prostate-cancer/stages/grades>](<https://www.cancerresearchuk.org/about-cancer/prostate-cancer/stages/grades>) - Cancer Research UK
4. [<https://www.nature.com/articles/s41591-021-01620-2>](<https://www.nature.com/articles/s41591-021-01620-2>) -Artificial intelligence for diagnosis and Gleason grading of prostate cancer: the PANDA challenge
5. [<https://www.californiaprotons.com/ro/prostate-cancer/stages-grades/>](<https://www.californiaprotons.com/ro/prostate-cancer/stages-grades/>) - California Protons Cancer Terapy Center
6. [<https://www.kaggle.com/competitions/prostate-cancer-grade-assessment/data>](<https://www.kaggle.com/competitions/prostate-cancer-grade-assessment/data>) - kaggle