



Univerzitet u Beogradu - Elektrotehnički fakultet

Katedra za signale i sisteme



Izdvajanje obeležja na segmentiranim kalkulozama u bubrezima

Student

Vukašin Krsmanović 2020/0008

Mentor

dr. Milica Janković, vanredni profesor

Beograd, *jun* 2024. godine

ZAHVALNICA

Ovim putem želim da izrazim iskrenu zahvalnost svima koji su učestvovali u izradi ovog projektnog zadatka.

Pre svega, želim da se zahvalim svom mentoru, prof. dr. Milici Janković, na deljenju svog znanja i korisnim savetima. Vaša nesebična podrška tokom cele izrade projekta je u velikoj meri pomogla u izradi ovog rada ali i u napredovanju mog obrazovanja.

Posebnu zahvalnost dugujem dr. Dimitriju Šarcu i dr. Otašu Durutoviću, za korisne sugestije i povratne informacije koje su značajno doprinele u boljem razumevanju tematike ovog rada.

Zahvaljujem se Klinici za urologiju, Univerzitetskog kliničkog centra Srbije i Laboratoriji za biomedicinsko inženjerstvo i tehnologiju, Elektrotehničkog fakulteta Univerziteta u Beogradu.

Vukašin Krsmanović

U Beogradu, jun 2024.

SADRŽAJ

ZAHVALNICA	2
SADRŽAJ.....	3
1 UVOD	4
1.1 Pregled literature.....	4
2 METODOLOGIJA RADA	5
2.1 Pacijenti i proces izvlačenja obeležja.....	5
2.2 Početna analiza i izbor obeležja.....	7
2.3 Dodatna provera i statistička analiza	10
3 REZULTATI I DISKUSIJA	12
3.1 Rezultati statističke analize	12
3.2 Rezultati drugih istraživanja	14
4 ZAKLJUČAK.....	16
5 LITERATURA.....	17

1 UVOD

1.1 Pregled literature i cilj

Kamen, ili kalkuloza je urološka bolest kod koje dolazi do formiranja kristalnih struktura u urinarnom traktu. Kamen može da se formira u bubrezima, ureterima i mokraćnoj bešici. Ovaj rad će se baviti isključivo kamenom u bubregu.

Većina istraživača smatra da kamen u bubregu nastaje kao rezultat fizičko-hemijskog disbalansa u mokraći. Pomenuti disbalans predstavlja poremećaj u ravnoteži između promotera i inhibitora kristalizacije u uslovima prezasićenosti i promene pH vrednosti mokraće. Pored navedenih razloga, na pojavu kamena u bubregu utiču i prisustvo infekcije izazvane mikroorganizmima koji razlažu ureju, poremećaji u morfologiji mokraćnog sistema i poremećaji u protoku mokraće. Istraživanja sugerišu da gojaznost, dijabetes, navike u ishrani, gensko nasleđe i smanjen unos tečnosti su dodatno povezani sa prisustvom kamena u bubregu. Najčešći simptomi kalkuloze su krv u urinu i bol u stomaku, boku ili preponama. Ovo oboljenje se javlja kod 1 od 11 ljudi u nekom periodu njihovog života, pri čemu muškarci obolevaju dvostruko više nego žene [1].

Statistika prevalencije sa zapadne i istočne hemisfere pokazuje na povećanje prisustva kamena kod stanovništva u toku poslednjih nekoliko decenija. Takođe je zabeležen porast u intervencijama i posetama lekarima koji se bave oboljenjima ovog tipa. Zanimljiv je podatak da je u SAD-a 2000. godine za tretmane koji se tiču lečenja kamena u bubregu potrošeno čak 2 milijarde dolara [2,3].

Jasno je da je data tema vrlo aktuelna i interesatna za istraživanje, sa različitih aspekata. Jedan od tih aspekata jeste radiomika. Radiomika je kvantitativna metoda koja izdvaja veliku količinu podataka iz medicinskih slika. Ovi podaci (obeležja) se dalje koriste za pravljenje statističkih modela kao i modela koji rade po principima veštačke inteligencije. Dati modeli služe za rešavanje različitih problema. Na osnovu mnoogobrojnih eksperimenata, zaključeno je da se radiomika u kontekstu ispitivanja kamena u bubregu može primeniti za predviđanje da li pacijent može spontano izbaciti kamen, razlikovanje kamenca i flebolita, identifikacija vrste i sastava kamena, prisustvo oboljenja hidronefroze [6].

Cilj ovog rada je da se utvrdi koja su značajna obeležja koja se mogu izvući sa medicinskih slika bubrega koji sadrže kalkuloze. Data obeležja se u kombinaciji sa statistikom i/ili algoritmima veštačke inteligencije mogu koristiti u različite namene kao što su predviđanje prisustva kamena i različitih oboljenja, utvrđivanje da li ima potrebe za hirurškom intervencijom, određivanje veličine kamena u bubregu kao i druge primene.

2 METODOLOGIJA RADA

2.1 Pacijenti i proces izvlačenja obeležja

Celokupan projekat je odrađen na medicinskim slikama koje su dobijene od strane Klinike za urologiju Univerzitetskog kliničkog centra Srbije. U pitanju su CT slike koje su nastale pomoću uređaja proizvedenog od strane firme Canon Medical Systems.

Baza podataka obuhvata 10 pacijenata (5 muških, 5 ženskih) i svaki pacijent je imao kalkulozu prisutnu makar u jednom bubregu.

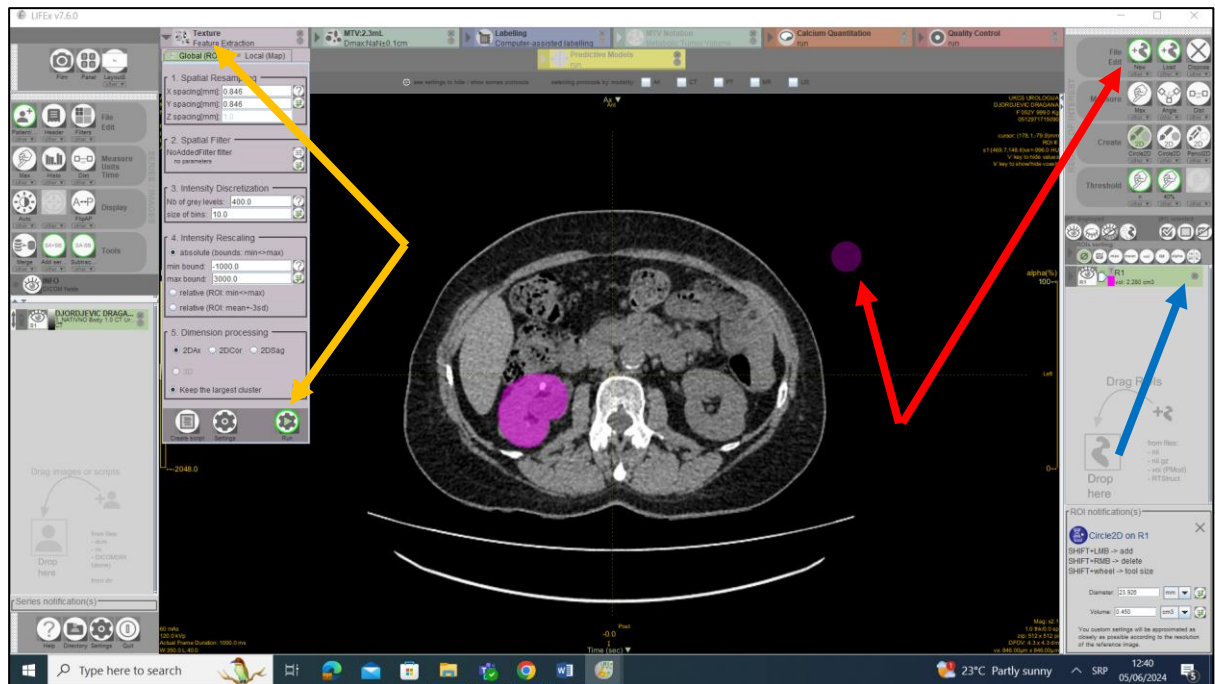
Osnovna ideja je da se izvuku obeležja za zdrav i nezdrav bubreg i da se uporedi da li postoji razlika u vrednostima za data obeležja. Ukoliko postoji statistički značajna razlika između ove dve grupe u odnosu na date vrednosti, onda se može zaključiti da su data obeležja značajna u razlikovanju ove dve grupe i da mogu biti korišćena za dalju analizu i primene.

Obeležja su izvučena uz pomoć komercijalnog softvera LIFEx.org [5]. Proces izvlačenja obeležja se sastoji iz dva koraka:

- 1) manuelno se označi regija od interesa (u ovom slučaju su to bubrezi)
- 2) klikom na opciju Run u jezičku Texture se eksportuju vrednosti obeležja

Pomenute vrednosti obeležja će se nalaziti u Excel fajlu koji softver automatski generiše. Dati Excel fajl će se nalaziti u folderu koji softver LIFEx takođe automatski generiše. Naziv foldera i Excel fajla će biti dodeljen na osnovu podataka iz zaglavlja posmatrane slike.

Na Slici 1 se može videti kako izgleda postupak izvlačenja obeležja za nezdrav bubreg u softveru LIFEx, na konkretnom primeru.



Slika 1. Izgled glavnog prozora u softveru LIFEx

U okviru prozora File Edit bira se opcija sa natpisom New. Ova opcija omogućava da se označi nova regija od interesa (ROI). Nakon toga kursor miša dobija podrazumevani oblik, a to je ljubičasti krug. Veličina kruga se može promeniti pomeranjem tastera Scroll na mišu, istovremeno držeći pritisnut taster Shift na tastaturi. Boja i oblik kursora se takođe mogu promeniti. Pomenuti koraci su ilustrovani strelicama crvene boje na prethodnoj slici.

ROI se na datoj slici boji tako što se prelazi kursom preko regije (klikom na levi taster miša) istovremeno držeći taster Ctrl na tastaturi.

Narandžaste strelice ukazuju na prozor Texture, u okviru koga se automatski otvara jezičak Global (ROI), na čijem dnu se nalazi opcija Run. Dodirom na Run se vrednosti za obeležja eksportuju u Excel fajl koji je softver automatski generisao.

Čitav proces se završava klikom na taster na koji ukazuje plava strelica i sada se postupak može ponoviti, za neku novu ROI.

U prethodnom primeru je objašnjen postupak izvlačenja obeležja za nezdrav bubreg. Analogan proces se može primeniti i za zdrav bubreg. Softver radi na takav način da obeležja izvučena za istu sliku čuva u istom Excel fajlu, iako se obeležja odnose na dve različite ROI. Ukoliko se prvo izvuku obeležja za nezdrav i potom za zdrav bubreg, Excel tabela će imati 3 reda. U poljima tabele će se nalaziti različiti parametri, podaci iz zaglavlja slike kao i obeležja za datu ROI. Prvi red tabele je uvek isti, i on se odnosi na nazive parametara koji su izvučeni za datu sliku – zaglavlje tabele. Ukoliko se prvo pronađu obeležja za nezdrav, a potom za zdrav bubreg, drugi red će sadržati vrednosti parametara za nezdrav bubreg dok će se treći red odnositi na iste vrednosti samo za zdrav bubreg.

Za svakog pacijenta je primenjen pomenuti postupak, pronađene su vrednosti obeležja za nezdrav i zdrav bubreg, redom. Na ovaj način se dobija 10 tabela u kojima postoje podaci za nezdrav i zdrav bubreg.

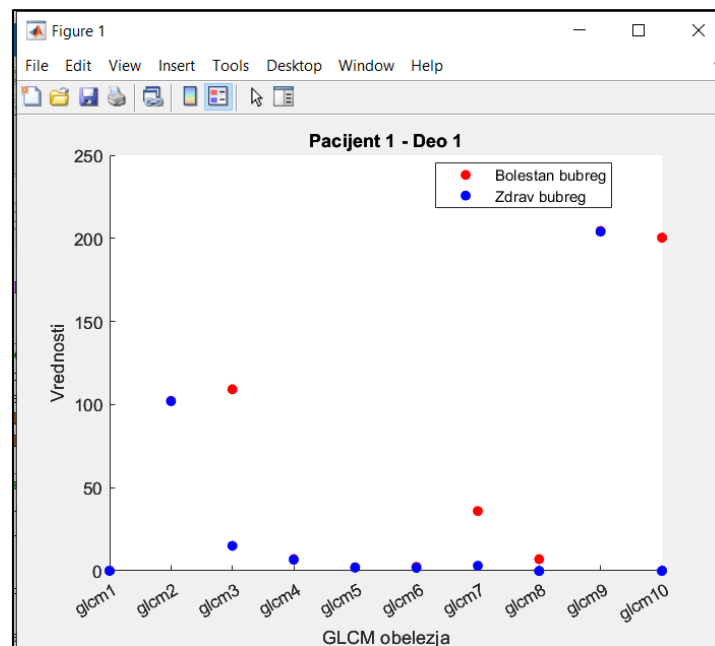
2.2 Početna analiza i izbor obeležja

Ukupan broj izvučenih obeležja pomoću LIFEx za svakog pacijenta je 179. Direktno posmatrajući Excel tabele je zaključeno da postoji značajna razlika u obeležjima koja su izvedena iz GLCM matrice datih ROI. GLCM (Gray Level Co-occurrence Matrix) matrica je alat za analizu tekstone slike. Ona beleži koliko puta se određene kombinacije piksela iste nijanse sive pojavljuju zajedno na slici u različitim smerovima i udaljenostima. Iz GLCM matrice se mogu izvući važna obeležja kao što su kontrast, energija, korelacija, itd. Ova obeležja pomažu da se bolje razume strukture slike i mogu otkriti fine detalje i razlike koje nisu vidljive golim okom, što je korisno u medicinskoj dijagnostici.

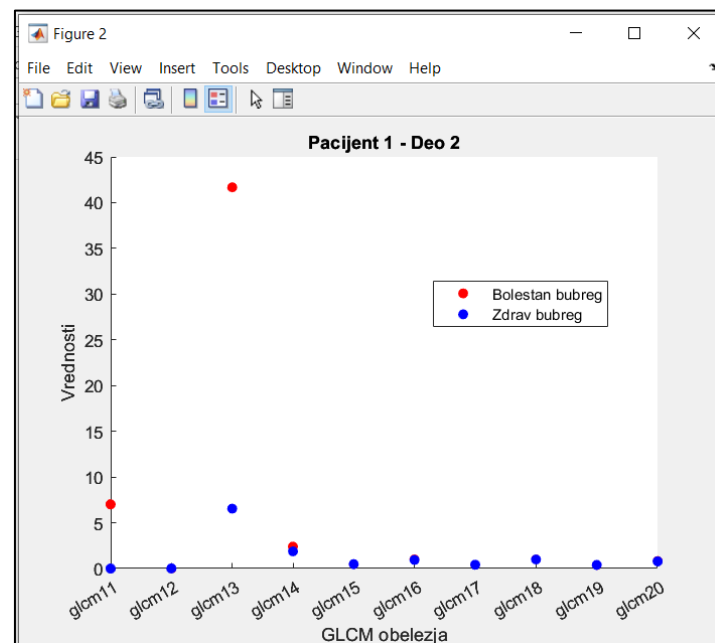
To je bila motivacija da se izbor značajnih obeležja suzi na GLCM obeležja. Potrebno je da se ovaj zaključak koji je intuitivan dodatno potkrepi analizama. Kompletna analiza zajedno sa prpratnim graficima je urađena u programskom paketu MatlabR2022b (The Mathworks, Natick, USA).

Softver izdvaja ukupno 24 GLCM obeležja. Posmatrani su grafici ova 24 obeležja, ponaosob za svakog pacijenta. Svakom pacijentu je odgovaralo po 6 grafika, 1 grafik za prvih 10 obeležja, 1 grafik za drugih 10 obeležja i 4 grafika za po svako od preostala 4 obeležja. Ovo je urađeno zbog preglednosti samih grafika i data podela po graphicima je odgovarajuća, usled različitih redova veličine za posmatrana obeležja. Paralelno su pravljene box plot prikazi ponaosob za svako od 24 obeležja, gde su se na box plot prikazu mogli videti box-ovi i za zdrav i za nezdrav bubreg. Dakle to je ukupno 24 box plot grafika. Ovo je dodatno pomoglo u tome da se primeti po kojim se obeležjima ove dve grupe statistički razlikuju ali i da se vidi raspršenost (varijansa) posmatranog obeležja kroz obe grupe.

Radi ilustracije, na Slici 2 i Slici 3 se redom mogu videti prvih 10 i drugih 10 obeležja za pacijenta pod rednim brojem 1.

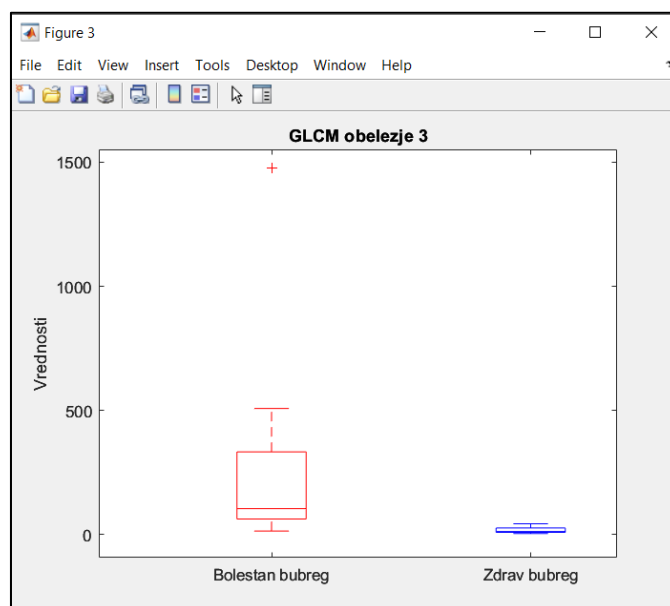


Slika 2. Prikaz vrednosti 10 obeležja iz prve grupe obeležja za pacijenta pod rednim brojem 1



Slika 3. Prikaz vrednosti 10 obeležja iz druge grupe obeležja za pacijenta pod rednim brojem 1

Na Slici 4 se može videti box plot grafik za GLCM obeležje pod rednim brojem 3.



Slika 4. Boxplot grafik za GLCM obeležje pod rednim brojem 3

Sa grafika koji prikazuju vrednosti 24 obeležja pojedinačnog pacijenta, za oba bubrega, se može primetiti koja se obeležja izdvajaju za svakog ponaosob pacijenta, dok se sa box plot grafika za svako pojedinačno obeležje pokazuje kako se dato obeležje karakteriše u globalnom smislu. Posmatrajući date grafike za svakog pacijenta i za svako obeležje, izabrano je ukupno 9 GLCM obeležja koja mogu biti značajna za razlikovanje nezdravog i zdravog bubrega. To su GLCM obeležja pod rednim brojem 3,7,8,10,11,13,22,23,24 i ta obeležja redom imaju nazive, Joint Variance, Difference Variance, Difference Entropy, Sum Variance, Sum Entropy, Contrast, Cluster Tendency, Cluster Shade i Cluster Prominence. Dodatnom analizom je primećeno da postoje i druga obeležja koja mogu dati dobre rezultate. U literaturi se pored GLCM obeležja za analizu koriste NGDTM i GLSZM obeležja [4]. Iz pomenutih razloga su u ovom projektu dodata još dva obeležja za analizu – NGDTM Complexity i GLSZM Gray Level Variance. U Tabeli 1 se može videti šta svako od pomenutih obeležja predstavlja [7].

Tabela 1. Opis svih 11 obeležja

Joint Variance	Mera varijanse u zajedničkoj raspodeli nijansi sive parova piksela u slici.
Difference Variance	Varijansa razlika nijansi sive između parova piksela u GLCM matrici.
Difference Entropy	Entropija razlike nijansi sive, koja meri nepredvidljivost tih razlika.

Sum Variance	Suma varijanse sivih tonova između parova piksela u GLCM matrici.
Sum Entropy	Suma entropija nijansi sive, koja meri nepredvidljivost tih suma.
Contrast	Mera razlike između najtamnijih i najsvetlijih nijansi sive u slici.
Cluster Tendency	Mera grupisanja nijansi sive u GLCM matrici, indikativna za homogenost teksture.
Cluster Shade	Mera asimetrije raspodele nijansi sive u GLCM matrici, koja može ukazivati na prisustvo struktura ili obrazaca.
Cluster Prominence	Mera izraženosti grupisanja nijansi sive u GLCM matrici, koja može ukazivati na značajnu teksturalnu karakteristiku.
NGTDM Complexity	Mera složenosti teksture slike, koja se izračunava na osnovu razlika između srednje vrednosti susednih nijansi sive i vrednosti centralnog piksela.
GLSZM Gray Level Variance	Varijansa nijanse sive unutar zona iste veličine na slici.

2.3 Dodatna provera i statistička analiza

Iz prethodne analize je primećeno da su kod svakog pacijenta obeležja za nezdrav bubreg značajno veća nego obeležja za zdrav bubreg, osim kod dva pacijenta. Pomenuta dva pacijenta su imala slične vrednosti za obeležja i za zdrav i za nezdrav bubreg. Ovo je dovelo do sumnje da pomenuti pacijenti imaju oba bubrega nezdrava. U razgovoru sa lekarima sa Klinike za urologiju je potvrđeno da pomenuti pacijenti imaju oba bubrega nezdrava. Pošto je za dalju statističku analizu od značaja da pravilno podelimo bubrega u grupe zdravih i nezdravih bubrega, napravljena je provera koja će da ima ulogu da bubreg za koji je pretpostavljeno da je zdrav proglasi nezdravim ukoliko posmatrani bubreg prođe datu proveru.

Posmatra se jedno obeležje. Pošto se očekuje da nezdrav bubreg ima svako obeležje veće u odnosu na zdrav, utvrdi se kolika je apsolutna razlika za posmatrano obeležje za oba bubrega i ukoliko je ta razlika manja u odnosu na 50% vrednosti datog obeležja za nezdrav bubreg, to znači da su oba bubrega "bliska" po datom obeležju i inkrementira se brojač. Ukoliko ovo važi za najmanje 5 obeležja (za više od 50% obeležja, pošto ih ima 9), proglašava se da su oba bubrega nezdrava. Data provera je primenjena u daljoj analizi.

Do sada je čitava analiza bila zasnovana na posmatranju grafika i proceni koja bi to obeležja bila od značaja. Potrebno je datu analizu potkrepiti matematički, kroz statistiku. Cilj je da se pokaže da zaista postoji statistički značajna razlika između nezdravih i zdravih bubrega. Potrebno je uraditi analizu za svako obeležje, ukoliko se pokaže da ne postoji statistički značajna razlika između zdravih i nezdravih bubrega, u odnosu na dato obeležje, onda se to obeležje odbacuje.

Kao statističku analizu datih podataka je iskorišćen Mann-Whitney test. Ukratko, Mann-Whitney test, je neparametarski test koji se koristi za poređenje dve nezavisne grupe. U ovom slučaju, koristi se za poređenje obeležja između zdravih i nezdravih bubrega. Ovaj test se koristi iz razloga što nije zadovoljena normalna raspodela ulaznih podataka. Kao rezultat ovog testa dobija se p vrednost i ona je mera verovatnoće da su dobijeni rezultati slučajni. U kontekstu Mann-Whitney testa, p-vrednost pokazuje verovatnoću da ne postoji prava razlika između grupa, već da je uočena razlika rezultat slučajnosti. Kao standard u statističkoj analizi, a i u analizi koja je sprovedena u ovom projektu, p-vrednost manja od 0.05 ($p < 0.05$) smatra se statistički značajnom. To znači da postoji manje od 5% šanse da su rezultati slučajni, što daje visok nivo pouzdanosti da postoji stvarna razlika između posmatranih grupa. Dodatno je tabelarno za svako obeležje prikazana minimalna vrednost, medijana i maksimalna vrednost, ponaosob za zdrave i nezdrave bubrege.

Tokom prvobitne analize obeležja je primećeno da postoji značajna razlika i za druga obeležja. Odabrana su dodatna dva obeležja za analizu – NGTDM Complexity i GLSZM Gray Level Variance. Analiza ovih obeležja je značajna jer povećava broj mogućih obeležja koja se mogu pokazati kao dobra u razdvajanju grupa bubrega. Za oba obeležja je urađen Mann-Whitney test.

3 REZULTATI I DISKUSIJA

3.1 Rezultati statističke analize

Nakon posete Klinici za urologiju, utvrđeno je da pacijenti pod rednim brojem 5 i 10 imaju oba bubrega nezdrava. Implementirana provera je uspešno našla da pacijenti pod rednim brojem 5 i 10 imaju oba bubrega nezdrava.

Tabelarni prikaz minimalne vrednosti, medijane i maksimalne vrednosti GLCM obeležja odvojeno za zdrave i nezdrave bubrege se može videti u Tabeli 2 i u Tabeli 3, redom.

Tabela 2. Vrednosti minimuma, medijane i maksimuma za svako obeležje kod grupe zdravih bubrega

Redni broj obeležja	Grupa zdravih bubrega		
	Min	Medijana	Max
3	3.911	11.702	43.280
7	1.530	2.711	6.010
8	0	0	0.122
10	0	0	58.766
11	0	0	0.137
13	3.541	6.052	10.860
22	0	0	58.761
23	-6260.971	-510.439	-7.170
24	549.387	12785.658	674420.575

Tabela 3. Vrednosti minimuma, medijane i maksimuma za svako obeležje kod grupe nezdravih bubrega

Redni broj obeležja	Grupa nezdravih bubrega		
	Min	Medijana	Max
3	13.670	103.960	1476.751
7	5.827	24.141	128.855
8	0	6.393	7.558
10	0	86.510	1151.799
11	0	6.557	7.604
13	10.554	30.065	159.417
22	0	86.510	1151.799
23	1453.600	42089.266	825040.918
24	115981.843	5210136.903	166090477.467

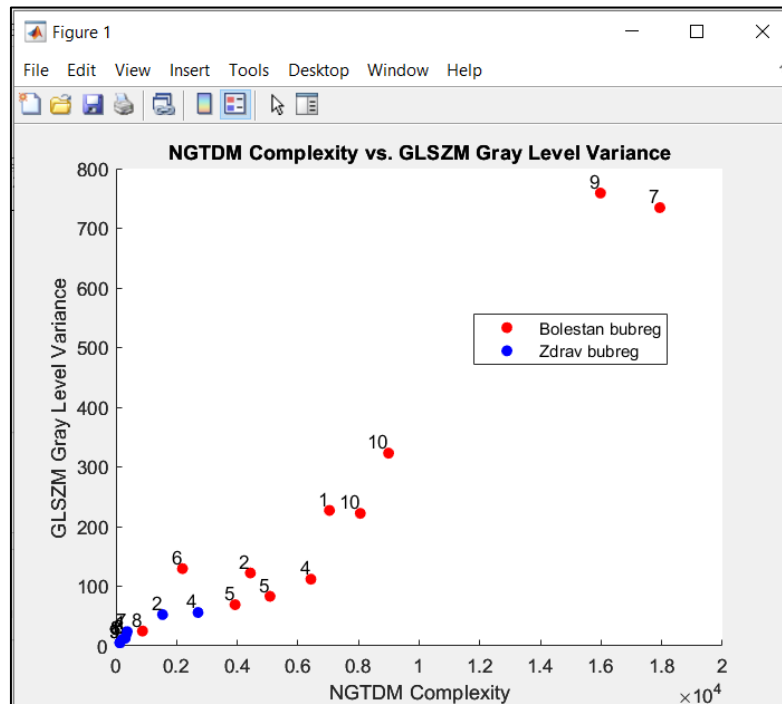
U Tabeli 4 se mogu videti p-vrednosti za svako od 9 GLCM obeležja.

Tabela 4. P-vrednosti za svako GLCM obeležje

Redni broj GLCM obeležja	p - vrednost
3	0.00104
7	0.00033
8	0.00507
10	0.00507
11	0.00507
13	0.00033
22	0.00507
23	0.00025
24	0.00045

Kako za svaku p vrednost iz prethodne tabele važi da je manja od 0.05 zaključak je da postoji statistički značajna razlika između zdravih i nezdravih bubrega u odnosu na svih 9 GLCM obeležja.

Potrebno je prikazati kako izgledaju vrednosti dva dodatna obeležja, za zdrav i nezdrav bubreg. Na Slici 5 možemo videti taj prikaz.



Slika 5. Grafički prikaz za dodatna dva obeležja

Očigledno je da data obeležja imaju značajno različite vrednosti obeležja NGTDM Complexity i GLSZM Gray Level Variance, za zdrav i nezdrav bubreg. P vrednosti koje su dobijene za ova dva obeležja su redom 0.00032 i 0.00018. Zaključujemo da i u odnosu na ova dva obeležja postoji statistički značajna razlika između zdravih i nezdravih bubrega.

3.2 Rezultati drugih istraživanja

Tokom izrade ovog projekta nađen je rad koji je takođe proučavao teksturalna obeležja koja su dobijena za bubrege, a ne za kalkulozu. Pomenuti rad je imao više zadataka, među kojima je bilo da se na osnovu izvučenih obeležja i statističkog modela donese zaključak da li pacijent ima kalkulozu ili ne, što je upravo rezultat ovog projekta. Važno je napomenuti da su istraživači izabrali obeležja zasnovana na GLCM matrici - Inverse difference moment normalized i NGDTM matrici - Coarseness. Pomenuto GLCM obeležje se nije pokazalo kao značajno u analizi ovog projekta, dok pomenuto NGDTM obeležje nije razmatrano. Kao rezultat su dobili p vrednost manju od 0.0001 [4].

U većini radova se ispostavlja da se obeležja izvlače za kalkulozu, a ne za sam bubreg. Iz tog razloga je teško uporediti rezultate ovog projekta sa onim iz literature, ali svakako treba napomenuti da je korišćenje GLCM obeležja u svrhu izučavanja kamena u bubregu standard. Dva obeležja koja su naknadno dodata (NGTDM Complexity i GLSZM Gray Level Variance) u ovom slučaju služe radi povećanja redundantnosti same klasifikacije, ukoliko iz nekog razloga za to bude imalo potrebe.

4 ZAKLJUČAK

Na osnovu rezultata statističke analize, uspešno su identifikovana obeležja koja efikasno razlikuju zdrave od nezdravih bubrega, u smislu prisustva kalkuloze. Konkretno, identifikovano je devet značajnih GLCM obeležja. Dodatno je analizirano po jedno obeležje iz grupe NGDTM obeležja i iz grupe GLSZM obeležja, od kojih su oba takođe dali dobre rezultate.

Ovaj projekat pokazuje potencijal za dalja istraživanja i primene. Identifikovana obeležja se mogu koristiti za različite vrste klasifikacije, kao što su određivanje da li ima potrebe da se pacijent privrgne medicinskom tretmanu, predviđanje prisustva bolesti u bubrezima i slične medicinske primene. Važno je napomenuti da pored statističkih modela, data obeležja bi mogla da budu uspešno iskorišćena za pravljenje modela koji radi na principu nekog od algoritama veštačke inteligencije.

Ukupno gledano, ovaj rad predstavlja dobru osnovu za dalja istraživanja kamena u bubregu i srodnih oboljenja, što svakako može naći primenu u praksi.

5 LITERATURA

- [1] Leslie SW, Sajjad H, Murphy PB. Renal Calculi, Nephrolithiasis. 2024 Apr 20. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2024 Jan–. PMID: 28723043.
- [2] Litwin MS, Saigal CS, Yano EM, et al. Urologic diseases in America Project: analytical methods and principal findings. *J Urol*. 2005;173(3):933-937.
- [3] Fwu CW, Eggers PW, Kimmel PL, et al. Emergency department visits, use of imaging, and drugs for urolithiasis have increased in the United States. *Kidney Int*. 2013;83:479-86.
- [4] Homayounieh, F., Doda Khera, R., Bizzo, B. C., Ebrahimian, S., Primak, A., Schmidt, B. & Kalra, M. K. (2021). Prediction of burden and management of renal calculi from whole kidney radiomics: a multicenter study. *Abdominal Radiology*, 46, 2097-2106.
- [5] C Nioche, F Orlhac, S Boughdad, S Reuzé, J Goya-Outi, C Robert, C Pellot-Barakat, M Soussan, F Frouin, and I Buvat. LIFEx: a freeware for radiomic feature calculation in multimodality imaging to accelerate advances in the characterization of tumor heterogeneity.
- [6] Lim, E.J.; Castellani, D.; So, W.Z.; Fong, K.Y.; Li, J.Q.; Tiong, H.Y.; Gadzhiev, N.; Heng, C.T.; Teoh, J.Y.-C.; Naik, N.; et al. Radiomics in Urolithiasis: Systematic Review of Current Applications, Limitations, and Future Directions. *J. Clin. Med*. 2022, 11, 5151.
- [7] Hall-Beyer, M. (2000). GLCM texture: a tutorial. *National Council on Geographic Information and Analysis Remote Sensing Core Curriculum*, 3(1), 75.