



Metode bazate pe textură pentru analiza și recunoașterea structurilor anatomice din imagini ultrasonografice

Teză de Doctorat
Octombrie 2011

Autor: Asis. Ing. Delia-Alexandrina Mitrea, Universitatea Tehnică din Cluj-Napoca,
Facultatea de Automatică și Calculatoare

Conducător științific: Prof. Dr. Ing. Sergiu Nedevschi

~ Rezumat ~

1. Obiectivele cercetării

Textura este o proprietate vizuală fundamentală a suprafețelor obiectelor din lumea exterioară, supusă percepției directe a ochiului uman, dar și a țesuturilor organelor interne, vizibile și pe cale noninvazivă, prin intermediul tehnologiilor aferente imagisticii medicale. Aceasta este în măsură a diferenția obiectele în funcție de natura lor, furnizând, de asemenea, informații privind tipul formațiunilor anatomo-patologice, respectiv starea patologică a țesuturilor. Metodele computerizate oferă posibilitatea unei caracterizări statistice și multi-rezoluționale a tuturor tipurilor de suprafețe, conducând la sesizarea unor diferențe subtile între texturi aparent similare [15].

Rolul științei calculatoarelor și al electronicii în domeniul cercetărilor interdisciplinare fiind unul foarte important, iar rolul medicinei fiind în mod evident vital din punct de vedere socio-uman, conjugarea celor două discipline - care reprezintă domenii de vârf ale științei atât prin natura cât și prin misiunea lor, oferă contextul ideal pentru obținerea de rezultate de mare finete menite a marca progrese semnificative în domeniul cunoașterii.

Aplicabilitatea acestor rezultate are o importanță semnificativă mai ales în cazul abordării problematicei afecțiunilor medicale severe, având în vedere faptul că incidența acestora prezintă tendința de a atinge cote alarmante. Din acest motiv, implicarea potențialului tehnologiei informației în perfecționarea actului medical, bazat pe fundamentele robuste oferite de domeniul matematicii, constituie efectiv un demers necesar pentru progrese în ceea ce privește diagnosticul precoce și întreaga conduită medicală aferentă acestor afecțiuni.

Tumorile maligne reprezintă o problemă de mare interes în zilele noastre, întrucât acestea pot afecta în mod iremediabil organismul uman. Detectia precoce sau cel puțin anterior etapei de metastazare a acestor tumori, ar putea conduce la o îmbunătățire a șanselor de vindecare, cât și la prevenirea metastazelor. Biopsia se numără printre putinele metode de încredere în ceea ce privește diagnosticarea tumorilor maligne, reprezentând „standardul de aur” în acest sens, fiind însă o metoda invazivă, periculoasă pentru pacient, putând conduce la infecții și chiar la răspândirea în organism a tumorilor. Din acest motiv, interesează punerea la punct a unor metode non-invazive de diagnosticare, putându-se înlocui, astfel, biopsia. O astfel de metodă s-ar putea axa doar pe imagistica medicală, dar capacitatea de interpretare a informației vizuale de către subiectul uman este deseori limitată, deci dezvoltarea

unor metode computerizate adecvate care să vină în completarea analizei medicului ar putea constitui o soluție viabilă.

În ceea ce privește metodele specifice imagisticii medicale, Tomografia Computerizată (CT), Rezonanță Magnetică-Nucleară (RMN) și endoscopia reprezintă tehnici care comportă un anumit grad de risc, fiind considerate invazive, costul acestora fiind de asemenea ridicat. Ultrasonografia (ecografia) este, în schimb, o metodă non-invazivă, cu cost redus, putând fi repetată mult mai frecvent decât celelalte metode, în scopul urmăririi evoluției stării pacientului.

Ne propunem, astfel, dezvoltarea atât a unor metode de caracterizare cât mai rafinată a texturilor, cât și a unor tehnici non-invazive de diagnosticare a afecțiunilor severe, bazate pe prelucrări computerizate menite să proceseze și să analizeze informația preluată din imagini ultrasonografice. Metodele elaborate vor contribui la realizarea unei **biopsii virtuale**, îndeplinind următoarele sarcini: caracterizarea țesutului analizat prin trăsături imagistice, corelate cu proprietățile fizico-chimice ale acestuia; stabilirea clasei de apartenență a țesutului, în funcție de boala considerată.

Pe lângă tumorile maligne, există și alte clase de boli de importantă majoră, diagnosticate prin intermediul biopsiei sau al endoscopiei, care necesită o urmărire atentă a evoluției. Acestea sunt bolile hepatice difuze [7] (steatoza, fibroza, hepatita, ciroza), care, prin evoluția lor, pot conduce la apariția carcinomului hepatocelular (HCC) – cea mai frecventă tumoră hepatică malignă [7], bolile inflamatorii intestinale [59], care conduc la o degradare treptată a stării de sănătate a întregului organism. Aceste afecțiuni vor fi luate de asemenea în considerare în cadrul cercetării noastre.

În această ordine de idei, am elaborat **modelul imagistic textural** al afecțiunilor considerate, constând în:

- Setul exhaustiv și neredondant al trăsăturilor texturale relevante în ceea ce privește diferențierea țesutului afectat de boală în raport cu țesuturi vizual similare;
- Valori statistice specifice pentru trăsăturile texturale relevante: media, deviația standard, distribuția de probabilitate, intervale de încredere.

Necesitatea elaborării unui model imagistic textural optim a condus la următoarele obiective specifice acestei cercetări, toate aceste obiective fiind atinse prin investigație științifică directă, rezultatele obținute fiind, deci, originale în integralitatea lor:

- Experimentarea și compararea metodelor destinate analizei texturii și elaborarea unor metode noi; obținerea unor trăsături texturale potențial relevante, independente de orientare, iluminare, scală, acestea fiind experimentate atât în cazul texturilor obiectelor obișnuite, din lumea reală, cât și în cazul texturilor organelor interne, vizualizate prin intermediul imaginilor ultrasonografice.

- Experimentarea și compararea metodelor destinate reducerii dimensionalității datelor (metode de selecție a trăsăturilor și metode de extragere a trăsăturilor), combinarea acestora, elaborarea unor metode noi; stabilirea metodelor celor mai adecvate obținerii unui set final de trăsături texturale care să conducă la o caracterizare optimă a clasei de țesut supusă analizei și la o performanță maximă în ceea ce privește recunoașterea automată a acesteia.

- Experimentarea și compararea clasificatorilor, realizarea unor combinații de clasificatori; determinarea metodelor de clasificare care sunt cele mai adecvate obținerii celor mai bune performanțe în procesul de clasificare, în acest mod realizându-se și o evaluare de mare acuratețe a modelului imagistic textural.

- Impărțirea țesuturilor în subclase, în funcție de caracteristicile vizuale ale acestora, dar și în mod nesupervizat, în funcție de valorile trăsăturilor texturale.

Sintetizând, obiectivul principal al tezei îl constituie studiul texturii, dezvoltarea metodelor noi de analiză a acestora vizând în mod direct creșterea acurateții de discriminare între texturi.

Concretizarea rezultatelor prin aplicarea acestora în domeniul medical a condus la definirea și generarea modelului imagistic textural al afecțiunilor tisulare benigne și maligne, cu experimente și aplicații pentru bolile ficatului (maladii hepatice difuze, hepatocarcinom, tumori hepatice benigne), adenocarcinom prostatic, boli inflamatorii intestinale cât și cancer colo-rectal. Importanța abordării aplicabilității în domeniul medical decurge, în mod special, din incidența acestor maladii, mai ales în țările dezvoltate.

2. Stadiul actual al cercetărilor. Realizări semnificative existente în domeniile abordate.

În ceea ce privește metodele bazate pe textură, pentru caracterizarea naturii suprafețelor obiectelor de diferite tipuri, metode de tipul statisticilor de ordinul întâi au fost utilizate în scopul caracterizării texturilor arheologice [1], statisticile de ordinul al doilea fiind implementate pentru recunoașterea tipurilor de lemn, în conjuncție cu un clasificator de tip Learning Vector Quantization (LVQ) [9]. Indicele de autocorelație, statisticile bazate pe trăsături locale ale imaginii, de tipul muchiilor, respectiv pe microstructuri texturale detectate utilizând nucleele de convoluție ale lui Laws, au fost evaluate cu succes în contextul recunoașterii diferitelor tipuri de texturi din lumea reală [11], prin intermediul unui clasificator bazat pe metoda vecinului celui mai apropiat (k-nearest neighbor: k-nn). Metodele bazate pe fractali au fost studiate atât în contextul caracterizării formei obiectelor [51], cât și în cel al discriminării între texturi [4],[67]. Autorii subliniază faptul că metodele bazate pe fractali nu sunt în măsură a distinge între primitivele de bază ale texturii precum muchiile, colțurile și regiunile uniforme. În ceea ce privește utilizarea metodelor bazate pe transformări, în lucrarea [50] este propusă o nouă metodă pentru calculul matricei de coocurență a nivelurilor de gri (GLCM) pe baza transformatei Haar Wavelet. S-a obținut o acuratețe de recunoaștere maximă pentru de 93%, în cazul imaginilor de dimensiune 320x320 pixeli.

În scopul detecției maladiilor hepatice difuze, metodele bazate pe textură au fost utilizate pe scară largă, pentru extragerea unor valori caracteristice distribuției nivelurilor de gri și deci stării țesutului, aceste valori constituind, în cele mai multe cazuri, intrările unor clasificatori. Primele și cele mai simple metode utilizate pentru analiza bazată pe textură a imaginilor ultrasonografice, în scopul determinării maladiilor hepatice difuze, au fost cele bazate pe statistici de ordinul întâi. Histograma nivelurilor de gri și parametrul GLHW (Grey Level Histogram Width) – denotând diferența dintre nivelul maxim, respectiv cel minim, măsurată în număr de bare ale histogramei, au fost utilizate în [25], în scopul cuantificării ecogenității. Statisticile de ordinul întâi au fost considerate insuficiente pentru caracterizarea completă a proprietăților țesutului hepatic, în scopul de a diferenția maladiile hepatice difuze. Astfel, s-au utilizat statisticile de ordinul doi, precum matricea de coocurență a nivelurilor de gri (GLCM) și parametrii Haralick asociați: media GLCM, varianța GLCM, omogenitatea, entropia, energia (momentul unghiular de ordinul doi), contrastul, corelația [19], care au constituit intrările unor clasificatori precum **k-nn (k nearest neighbor)**, Support Vector Machine (SVM), respectiv rețele neuronale artificiale (ANN) [16]. În [16], în scopul de a se face diferențierea între maladiile hepatice difuze, a

fost utilizat un arbore structurat în funcție de perechile de clase analizate, parametrii texturali folosiți fiind media GLCM, energia GLCM, media și varianța GLCM, entropia, entropia sumă, entropia diferență, derivate din GLCM. Pentru clasificările realizate la fiecare nivel, a fost utilizată metoda Perceptronului Multinivel (MLP). [16]. Transformata Wavelet a fost aplicată în [76] și în [114]. În [76] a fost utilizată transformata **quincunx wavelet non-separabilă** pentru a analiza modificările semnalului la diferite scale. Această transformată a avut, de asemenea, rolul de a înlătura zgomotul diagonal din imaginile ultrasonografice.

În [57] Raeth a utilizat pentru prima dată parametrii texturali în scopul diferențierii ficatului normal de maladiile hepatice difuze, respectiv de țesutul hepatic malign. Au fost utilizați următorii parametri texturali: parametrii descriind distribuția nivelurilor de gri, obținuți pe baza histogramelor de intensitate; matricea run-length; parametrii de ordinul întâi ce caracterizează gradientul imaginii, alături de matricea de coocurență a nivelurilor de gri (GLCM) și parametrii asociați (contrast, corelație, energie, entropie); matricea de coocurență a direcțiilor asociate punctelor de muchie (gradient nenul), împreună cu parametrii Haralick corespunzători, au fost utilizate în scopul caracterizării texturii țesutului și a formei conturului inomogenităților localizate în imaginea ultrasonografică, în special în interiorul formațiunilor maligne. Alte trăsături texturale importante în diferențierea stărilor patologice considerate au fost: parametrii aparținând spectrului putere bi-dimensional, obținut prin aplicarea transformatei Fourier, parametrii ce caracterizează distribuția zgomotului de speckle în imagine. Tot în scopul stabilirii naturii maligne sau benigne a formațiunilor tumorale prezente în ficat, o metodă bazată pe transformata Wavelet a fost utilizată cu succes în [67]. Caracterizarea tumorilor prostatice s-a realizat în [28] prin intermediul transformatei Gabor și al clasificatorilor Bayesieni, pentru diferențierea țesutului prostatic malign în imagini de rezonanță magnetică (MRI).

În ceea ce privește caracterizarea și recunoașterea bolilor inflamatorii intestinale și a cancerului colonic prin intermediul texturii, în [30] autorii au utilizat Matricea de Coocurența a Nivelurilor de Gri (GLCM), împreună cu trăsături morfologice referitoare la formă și orientare, pentru a caracteriza țesutul malign și benign al pacienților bolnavi de cancer colonic, pe baza slide-urilor obținute la biopsie; în [6] autorii au analizat imaginile fluorescente ale peretelui colonic pe baza parametrilor texturali derivați din matricea de coocurență a nivelurilor de gri (GLCM), pentru a distinge mucoasa colonică sănătoasă de cea afectată de adenocarcinom. De asemenea, trăsături texturale precum asimetria și excesul histogramei de intensitate a imaginii, împreună cu trăsături de culoare de tip RGB, au fost implementate cu succes în scopul segmentării lumenului intestinal pe baza imaginilor endoscopice. Tot în acest context, s-a realizat detecția pe cale computerizată a maladiei Crohn și segmentarea leziunilor peretelui intestinal [10].

În ceea ce privește matricele de coocurență a nivelurilor de gri de ordin superior, există studii și implementări ale matricei de ordinul trei. Astfel, în [5], autorii au evaluat posibilitatea de discriminare a diferitelor tipuri de culturi de cereale utilizând trăsături texturale și informație de multipolarizare, pe baza imaginilor radar. Pentru determinarea trăsăturilor texturale, s-au utilizat histograme de ordinul întâi, doi și trei, respectiv parametrii rezultați pe baza acestora. Analiza texturală a radiografiilor femurale a fost realizată prin intermediul matricei de coocurență a nivelurilor de gri de ordinul trei (3rd order GLCM) în [20]. Această matrice a fost calculată considerând cei trei pixeli ca fiind coliniari, sau formând un triunghi dreptunghic, fiind situați la distanțele 1, 2, 3 sau 4 unul față de celălalt.

Trăsăturile texturale finale au fost obținute prin calculul mediei aritmetice a trăsăturilor Haralick rezultate din fiecare matrice de ordinul trei determinată. Doar corelații moderate au fost găsite între trăsăturile GLCM de ordinul trei și proprietățile bio-mecanice ale oaselor ($R^2=0.38$, $p<0.05$), în timp ce corelația dintre trăsăturile GLCM de ordinul trei și densitatea minerală a oaselor a fost nesemnificativă ($p>0.05$). Matricea de coocurență generalizată de ordinul n ($n>2$) a fost descrisă din punct de vedere teoretic în [3], autorii experimentând matricele de ordinul trei, patru și cinci în scopul recunoașterii gradului de ocupare al solurilor regiunilor vulcanice, pe baza imaginilor radar.

Matricele de coocurență generalizate, de ordinul doi, ale direcțiilor asociate punctelor de muchie, au fost utilizate în [65] în scopul detecției imperfecțiunilor procesului de osteogeneză pe baza imaginilor microscopice, iar în [57] pentru caracterizarea formei neregulate a conturului în interiorul formațiunilor maligne. De asemenea, în [57] au fost utilizate matrice de coocurență de ordinul doi ale valorilor gradientului de intensitate al imaginii.

Cercetări privind metodele de reducere a dimensionalității datelor s-au realizat în [27], unde arborii decizionali au fost utilizați ca metodă de selecție a trăsăturilor, în scopul caracterizării și recunoașterii cancerului la sân. Algoritmii genetici în combinație cu metode bazate pe rețele neuronale artificiale, au fost utilizați cu succes în [27], pentru diferențierea cancerului cervical pe baza imaginilor citologice. Metodele bazate pe Analiza Componentelor Principale (PCA) au fost utilizate atât în contextul segmentării [26], cât și al recunoașterii imaginilor medicale [53], în scopul eliminării zgomotului prezent în datele de intrare, precum și al punerii în evidență a principalelor moduri de variație ale acestora. Metoda Analizei Discriminanțelor Lineare (LDA) a fost utilizată în [54], sub forma unei variante modificate, care realizează totodată și selecția trăsăturilor relevante.

Schemele de combinare a clasificatorilor cele mai des experimentate au fost acelea de votare [31], bagging[16], boosting [16], stacking [61], dar și suma rezultatelor clasificatorilor individuali [31].

Metodele de clasificare multiclasă au fost elaborate și implementate în scopul recunoașterii unor tipuri multiple de cancer, atât pe baza imaginilor medicale [29], cât și utilizând date referitoare la expresii genetice [2].

➤ Probleme nerezolvate. Concluzii

Din cele prezentate, rezultă importanța metodelor bazate pe textură și relevanța parametrilor texturali în domenii aplicative multiple, în special în ceea ce privește caracterizarea stării patologice a țesuturilor diverselor organe, în particular în domeniul detecției bolilor hepatice difuze, a bolilor inflamatorii intestinale, respectiv a tumorilor maligne. Astfel, metodele imagistice computerizate bazate pe textură, pot contribui substanțial la identificarea precoce, pe cale neinvazivă, a unor afecțiuni ce sunt în măsura a pune în pericol viața umană. Atât metodele bazate pe trăsături statistice de ordinul întâi și doi, cât și metodele bazate pe fractali, respectiv pe transformări, furnizează rezultate satisfăcătoare în ceea ce privește caracterizarea și recunoașterea tumorilor maligne ale ficatului, prostatei, sânelui, glandelor salivare și respectiv a altor organe. Cele mai performante metode s-au dovedit a fi: matricea de coocurență a nivelurilor de gri (*GLCM*), statisticile bazate pe trăsături de tip gradient și pe microstructuri texturale obținute în urma aplicării nucleelor de convoluție ale lui Laws, matricele de coocurență generalizate bazate pe trăsături de tip gradient, trăsăturile multirezoluție obținute (în multe cazuri) după aplicarea transformatei Wavelet, trăsăturile referitoare la distribuția zgomotului în imaginea ultrasonografică. O arie practic neexplorată în ceea ce privește caracterizarea tumorilor maligne prin intermediul

texturii, o reprezintă aceea a matricelor de coocurență de ordin superior, mai mare decât doi ($n > 2$), care, din punct de vedere teoretic, au capacitatea de a depăși limitele percepției umane și de a mări gradul de acuratețe al analizei structurii țesutului, odată cu creșterea lui n . De asemenea, se mai pot aduce completări în ceea ce privește statisticile bazate pe trăsături de tip gradient, respectiv statisticile referitoare la forma 2D și 3D a microstructurilor texturale.

Privind prin prisma problemei biopsiei virtuale, se pot aduce contribuții importante referitor la următoarele aspecte legate de caracterizarea și recunoașterea afecțiunilor pe baza imaginilor ultrasonografice :

- Creșterea în continuare a acurateței procesului de recunoaștere automată, utilizând metode adecvate în scopul analizei texturii, al reducerii dimensionalității datelor, precum și metode de clasificare performante.
- Determinarea sub-claselor tisulare ce corespund unor stadii variate de evoluție a maladiei și corelarea trăsăturilor texturale relevante cu proprietățile fizico-chimice ale țesutului tumoral, corespunzătoare fiecărei astfel de clase.

Referitor la bolile inflamatorii intestinale și la tumorile colonice, există puține studii care să implice utilizarea parametrilor texturali determinați pe baza imaginilor ultrasonografice, deși ultrasonografia este o metodă non-invazivă, de cost scăzut, care poate fi repetată fără probleme în scopul monitorizării evoluției acestor maladii.

Cercetarea de față va realiza studiul și compararea unor metode reprezentative destinate analizei texturii, experimentate în situații diverse, în special în contextul analizei imaginilor ultrasonografice. Îmbunătățirea metodelor existente, generalizarea acestora, precum și dezvoltarea unor metode noi, care să confere metodelor existente un grad mai înalt de rafinament, crescând performanțele acestora în ceea ce privește discriminarea texturilor vizual asemănătoare, constituie unul dintre obiectivele principale ale tezei.

Contribuțiile aduse în domeniul aplicativ vor converge în scopul definirii modelului imagistic textural al unor afecțiuni abdominale reprezentative, pe baza imaginilor ultrasonografice. Acest obiectiv presupune combinarea, în manieră optimă, a celor mai performante metode destinate analizei texturii, a metodelor destinate selecției și extragerii trăsăturilor, a clasificatorilor și combinațiilor de clasificatori, respectiv a metodelor de clasificare multiclasă. Scopul urmărit este acela de a obține o caracterizare cât mai adecvată a afecțiunilor considerate, care să releve proprietăți subtile ale structurii țesuturilor afectate și să conducă la o acuratețe optimă de diagnoză automată și asistată de calculator.

Modelul imagistic textural al unei anumite afecțiuni este constituit pe baza identificării *setului trăsăturilor texturale* relevante atât pentru caracterizarea maladiei considerate cât și pentru diferențierea acesteia de alte maladii similare, acestor trăsături asociindu-li-se și *intervalele de valori măsurabile* identificate ca fiind specifice afecțiunilor considerate.

3. Contribuții proprii, care stau la baza metodologiei elaborate și a realizării aplicative

3.1. Contribuții teoretice

➤ Contribuții în domeniul metodelor destinate analizei texturii

În scopul de a mări gradul de acuratețe al analizei texturale, pentru a adăuga un grad cât mai înalt de rafinament caracterizării texturilor, astfel încât să fie depășită percepția vizuală umană atât în ceea ce privește domeniul imagisticii medicale cât și domeniul analizei celorlalte tipuri de imagini, am definit un set de metode noi, originale, de analiză a texturii. În mod concret, am definit în primul rând statistici de ordinul întâi ale trăsăturilor imaginii,

capabile de a surprinde gradul de dezordine locală prin variabilitatea gradientului direcțional, respectiv variabilitatea în orientarea muchiilor. Cele mai complexe și semnificative metode dezvoltate, în măsură a determina proprietăți imagistice dificil sau imposibil de detectat cu ochiul liber, sunt matricele de coocurență de ordin superior, bazate pe niveluri de gri, respectiv pe orientările muchiilor.

▪ **Definirea unor statistici de ordinul întâi ale orientărilor muchiilor**

Pentru caracterizarea variației statistice a orientării muchiilor în regiunea de interes, precum și a direcției de variație a intensității imaginii au fost elaborate noi trăsături, anume **variabilitatea în orientarea muchiilor** și, respectiv, **variabilitatea gradientului direcțional**. Eficiența acestora, în ceea ce privește capacitatea de diferențiere a texturilor, a fost testată atât utilizând texturi din lumea reală (colecția Brodatz), cât și texturi din domeniul imagisticii medicale. Puterea de discriminare a acestor trăsături a fost comparată atât cu cea a trăsăturilor texturale anterioare, cât și cu cea a unor trăsături cu semnificație similară, anume cu varianța orientării muchiilor, cu varianța gradientului direcțional, precum și cu trăsături derivate din histograma orientării muchiilor: asimetria și excesul [21] acestei histogramă, varianța valorilor (orientări ale muchiilor) corespunzătoare maximelor locale din histogramă. S-a considerat că există un maxim local atunci când valoarea curentă din histogramă, $H(v)$, a fost mai mare decât $\mu + Th$, unde μ a reprezentat media valorilor histogramei pe un interval de dimensiune k ($k=10$) în jurul punctului curent, iar Th ($Th=3$) a reprezentat un prag ales de utilizator.

• **Variabilitatea în orientarea muchiilor**

Variabilitatea în orientarea muchiilor este luată în considerare pentru a pune în evidență structura haotică a texturii. Orientarea muchiilor este determinată în fiecare punct al imaginii utilizând formula clasică (arctangenta unghiului format de gradientul orizontal cu cel vertical). Astfel, este obținută matricea orientărilor locale ale muchiilor, mapată pe matricea intensităților pixelilor, iar variabilitatea în orientarea muchiilor este determinată utilizând următoarea formulă:

$$\text{Variabilitate_orientare_muchii} = \# \{ (x, y) \mid \exists (x', y') \in V(x, y), \\ | \text{Orient_muchie}(x, y) - \text{Orient_muchie}(x', y') | \geq \text{prag} \} / \text{Nr_pixeli_muchie} \quad (1)$$

În formula (38), simbolul „#” reprezintă cardinalul mulțimii considerate, V reprezintă vecinătatea pixelului curent, **prag** este o valoare prag care poate fi setată funcție de natura problemei, iar **Nr_pixeli_muchie** este numărul total de pixeli de muchie din regiunea de interes. Astfel, variabilitatea în orientarea muchiilor reprezintă frecvența de apariție a unor pixeli de muchie pentru care orientarea muchiei în punctul respectiv diferă semnificativ față de orientările pixelilor de muchie vecini [D. Mitrea, 2011].[44]

• **Variabilitatea gradientului direcțional**

Variabilitatea gradientului direcțional este determinată în mod similar cu variabilitatea orientării muchiilor (D. Mitrea, 2008): [45]

$$\text{Variabilitate_grad_directional} = \# \{ (x, y) \mid \exists (x', y') \in V(x, y), \\ | \text{Grad_directional}(x, y) - \text{Grad_directional}(x', y') | \geq \text{prag} \} / \text{Nr_pixeli_muchie} \quad (2)$$

Gradientul direcțional a fost determinat pe baza raportului dintre vectorul gradient și modulul său, după cum este descris în Capitolul III, prin ecuația (54).

▪ **Definirea matricelor de coocurență generalizate de ordin superior și a parametrilor Haralick asociați**

În abordarea sa, Akono [3] nu furnizează o definiție a matricei de coocurență a nivelurilor de gri de ordinul n și definește doar o parte dintre parametrii Haralick asociați. De asemenea, Davis în [14] definește doar matricele de coocurență generalizate de ordinul doi.

În cele ce urmează, prezentăm definiția matricei de coocurență generalizată de ordinul n pe care am elaborat-o pentru a completa fundamentele teoretice ale domeniului, precum și definiția parametrilor Haralick corespunzători.

Extinzând definiția matricei de coocurență generalizată de ordinul doi furnizată în paragraful anterior, rezultă expresia de definiție dată în (3). Astfel, un element al acestei matrice, definește probabilitatea de apariție simultană a n elemente ale imaginii (ex. pixeli, muchii), notate cu $f_1, f_2, \dots, f_n$, aflate în relația spațială definită de predicatul S , valorile atributelor A asociate fiind $v_1, v_2, \dots, v_n$, [D. Mîrea, 2011] [48].

$$p_{S,A}(v_1, v_2, \dots, v_n) = \frac{\#\{(f_1, f_2, \dots, f_n) : A_{f_1} = v_1, A_{f_2} = v_2, \dots, A_{f_n} = v_n, S(f_1, f_2, \dots, f_n) = TRUE\}}{\#\{(f_1, f_2, \dots, f_n) : S(f_1, f_2, \dots, f_n) = TRUE\}} \quad (3)$$

Considerând elementele f_i ale imaginii ca fiind pixeli de coordonate (x_i, y_i) , S fiind un predicat ce definește relațiile spațiale dintre pixeli prin intermediul unor vectori de deplasament, iar $v_1, \dots, v_n$ fiind valorile atributelor determinate pentru fiecare pixel, definiția matricei de coocurență de ordinul n va fi cea descrisă în (15).

$$\begin{aligned} C_D(v_1, v_2, v_3, \dots, v_n) &= \#\{((x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_n, y_n)) : \\ &A(x_1, y_1) = v_1, A(x_2, y_2) = v_2, \dots, A(x_n, y_n) = v_n, \\ &|x_2 - x_1| \models \vec{dx}_1, |x_3 - x_1| \models \vec{dx}_2, \dots, |x_n - x_1| \models \vec{dx}_{n-1}, \\ &|y_2 - y_1| \models \vec{dy}_1, |y_3 - y_1| \models \vec{dy}_2, \dots, |y_n - y_1| \models \vec{dy}_{n-1}, \\ &\text{sgn}((x_2 - x_1)(y_2 - y_1)) = \text{sgn}(\vec{dx}_1 \cdot \vec{dy}_1), \dots \\ &, \text{sgn}((x_n - x_1)(y_n - y_1)) = \text{sgn}(\vec{dx}_{n-1} \cdot \vec{dy}_{n-1})\} \end{aligned} \quad (4)$$

Mulțimea vectorilor de deplasament este definită conform (16):

$$\vec{d} = ((\vec{dx}_1, \vec{dy}_1), (\vec{dx}_2, \vec{dy}_2), \dots, (\vec{dx}_{n-1}, \vec{dy}_{n-1})) \quad (5)$$

În practică, se utilizează matricea probabilităților de coocurență a n pixeli, descrisă în (17):

$$p(v_1, v_2, \dots, v_n) = \frac{C_D(v_1, v_2, \dots, v_n)}{\sum_{v_1=0}^{M-1} \sum_{v_2=0}^{M-1} \dots \sum_{v_n=0}^{M-1} C_D(v_1, v_2, \dots, v_n)} \quad (6)$$

În [3], Akono reprezintă matricea de coocurență a nivelurilor de gri sub forma unui arbore. Acesta definește, de asemenea, câțiva parametri pe baza matricei de coocurență a nivelurilor de gri de ordin superior, precum asimetria, a cărei formulă de definiție este redată în Capitolul III, descriind un algoritm de calcul care conduce la o complexitate computațională redusă, dar nu tratează definițiile parametrilor descriși mai jos, aceștia din urmă furnizând informații despre proprietățile vizuale și fizice globale ale texturii.

- Media calculată pe baza **GCM** de ordinul **n**, corespunzătoare pixelului de referință (μ_1), respectiv pixelilor de rang **k** și **n** din matricea de coocurență (μ_k , respectiv μ_n), se definește conform expresiilor matematice (18) - (20) [48] [D. Mitrea, 2011]:

$$\mu_1 = \sum_{v_1=0}^{M-1} v_1 \sum_{v_2=0}^{M-1} \dots \sum_{v_n=0}^{M-1} p(v_1, v_2, \dots, v_n) \quad (7)$$

.....

$$\mu_k = \sum_{v_k=0}^{M-1} v_k \sum_{v_1=0}^{M-1} \dots \sum_{v_{k-1}=0}^{M-1} \sum_{v_{k+1}=0}^{M-1} \dots \sum_{v_n=0}^{M-1} p(v_1, v_2, \dots, v_n), k \geq 2 \quad (8)$$

.....

$$\mu_n = \sum_{v_n=0}^{M-1} v_n \sum_{v_1=0}^{M-1} \sum_{v_2=0}^{M-1} \dots \sum_{v_{n-1}=0}^{M-1} p(v_1, v_2, \dots, v_n) \quad (9)$$

- Deviația standard definită pe baza **GLCM** de ordinul **n**, corespunzătoare pixelului de referință (σ_1), respectiv pixelilor de rang **k** și **n** din matricea de coocurență a nivelurilor de gri (σ_k , respectiv σ_n), se definește conform expresiilor matematice (21) - (23) [48]:

$$\sigma_1^2 = \sum_{v_1=0}^{M-1} (v_1 - \mu_1)^2 \sum_{v_2=0}^{M-1} \dots \sum_{v_n=0}^{M-1} p(v_1, v_2, \dots, v_n) \quad (10)$$

.....

$$\sigma_k^2 = \sum_{f_k=0}^{M-1} (v_k - \mu_k)^2 \sum_{v_1=0}^{M-1} \dots \sum_{v_{k-1}=0}^{M-1} \sum_{v_{k+1}=0}^{M-1} \dots \sum_{v_n=0}^{M-1} p(v_1, v_2, \dots, v_n) k \geq 2, k < n \quad (11)$$

.....

$$\sigma_n^2 = \sum_{v_n=0}^{M-1} (v_n - \mu_n)^2 \sum_{v_1=0}^{M-1} \sum_{v_2=0}^{M-1} \dots \sum_{v_{n-1}=0}^{M-1} p(v_1, v_2, \dots, v_n) \quad (12)$$

- Mărimile ce caracterizează proprietățile globale ale imaginii, determinate pe baza matricei de coocurență a nivelurilor de gri de ordinul **n**, rezultate prin extensia trăsăturilor Haralick de ordinul doi, se definesc conform expresiilor matematice (24)-(29) [48].

$$Entropia = \sum_{v_1=0}^{M-1} \sum_{v_2=0}^{M-1} \dots \sum_{v_n=0}^{M-1} p(v_1, v_2, \dots, v_n) (-\ln p(v_1, v_2, \dots, v_n)) \quad (13)$$

Entropia pune în evidență gradul de dezordine existent în aranjamentul nivelurilor de gri, cuantificând gradul de incertitudine privind apariția unei combinații de valori, $(g_1, g_2, \dots, g_n)$ ale intensităților pixelilor, aflați într-o relație spațială definită prin intermediul vectorilor de deplasament.

$$Energia = \sum_{v_1=0}^{M-1} \sum_{v_2=0}^{M-1} \dots \sum_{v_n=0}^{M-1} p(v_1, v_2, \dots, v_n)^2 \quad (14)$$

La fel ca în cazul matricei de coocurență a nivelurilor de gri de ordinul doi, și în cazul matricei de coocurență a nivelurilor de gri de ordinul n energia - fiind opusul entropiei, exprimă gradul de ordine, de uniformitate a texturii.

$$Omogenitatea\ locala = \sum_{v_1=0}^{M-1} \sum_{v_2=0}^{M-1} \dots \sum_{v_n=0}^{M-1} \frac{p(v_1, v_2, \dots, v_n)}{1 + \sum_{u=1}^{n-1} \sum_{v=u+1}^n dif(v_u, v_v)^2} \quad (15)$$

Omogenitatea locală caracterizează textura din punctul de vedere al disimilarității nivelurilor de gri corespunzătoare pixelilor aflați unul în apropierea celuilalt, fiind direct proporțională cu probabilitatea de apariție a unei combinații de niveluri de gri $(g_1, g_2, \dots, g_n)$ a n pixeli aflați într-o relație spațială definită prin intermediul vectorilor de deplasament, și invers proporțională cu suma valorilor absolute ale diferențelor nivelurilor de gri corespunzătoare tuturor prechilor de pixeli ce corespund acelei combinații.

$$Varianta = \sqrt{\sigma_1^2 \sigma_2^2 \dots \sigma_n^2} \quad (16)$$

Varianța a fost determinată ca fiind valoarea rădăcinii pătrate extrase din produsul deviațiilor standard corespunzătoare pixelilor de rang k , $1 \leq k \leq n$, din matricea de coocurență a nivelurilor de gri.

$$Corelatia = \sum_{v_1=0}^{M-1} \sum_{v_2=0}^{M-1} \dots \sum_{v_n=0}^{M-1} \left(\sum_{u=1}^{n-1} \sum_{v=u+1}^n \frac{(v_u - \mu_u)(v_v - \mu_v)}{\sqrt{\sigma_u^2 \sigma_v^2}} \right) p(v_1, v_2, \dots, v_n) \quad (17)$$

Corelația matricei de coocurență a nivelurilor de gri de ordinul n a fost determinată ca fiind suma corelațiilor existente între oricare două niveluri de gri, de rang u și v , din matricea de coocurență a nivelurilor de gri de ordin superior.

$$p_max = Max(p(v_1, v_2, \dots, v_n), 0 < v_k < M, 0 \leq k \leq n) \quad (18)$$

Probabilitatea maximă, **p_max**, definită în formula (18), reprezintă valoarea maximă care apare în matricea de coocurență de ordinul n și indică cea mai frecventă combinație de valori ale atributelor $(v_1, v_2, \dots, v_n)$ care apare în regiunea de interes, corespunzând unui număr de n pixeli aflați într-o relație spațială definită de către vectorii de deplasament. Pe baza acestor definiții, matricea de coocurență a nivelurilor de gri de ordin superior se definește considerând ca atribute ale pixelilor valorile nivelurilor de gri, iar matricea de coocurență a orientărilor muchiilor rezultă prin înlocuirea atributelor **A** cu valorile orientărilor muchiilor, determinate în punctul respectiv [44] [D.Mitrea, 2011]. Definițiile matematice aferente acestor matrice de coocurență sunt redată mai jos, în (30), respectiv (31).

$$\begin{aligned}
C_D(g_1, g_2, g_3, \dots, g_n) = & \#\{((x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_n, y_n)) : \\
& f(x_1, y_1) = g_1, f(x_2, y_2) = g_2, \dots, f(x_n, y_n) = g_n, |x_2 - x_1| = |\vec{dx}_1|, \\
& |x_3 - x_1| = |\vec{dx}_2|, \dots, |x_n - x_1| = |\vec{dx}_{n-1}|, |y_2 - y_1| = |\vec{dy}_1|, |y_3 - y_1| = |\vec{dy}_2|, \dots, \\
& |y_n - y_1| = |\vec{dy}_{n-1}|, \text{sgn}((x_2 - x_1)(y_2 - y_1)) = \text{sgn}(\vec{dx}_1 \cdot \vec{dy}_1), \dots \\
& , \text{sgn}((x_n - x_1)(y_n - y_1)) = \text{sgn}(\vec{dx}_{n-1} \cdot \vec{dy}_{n-1})\}
\end{aligned} \tag{19}$$

În (30), f este funcția imagine, iar $g_1, g_2, \dots, g_n$ sunt valorile intensităților pixelilor în punctele de coordonate $(x_1, y_1), (x_2, y_2)$, respectiv (x_n, y_n) .

$$\begin{aligned}
C_D(o_1, o_2, o_3, \dots, o_n) = & \#\{((x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_n, y_n)) : \\
& \text{orient_muchie}(x_1, y_1) = o_1, \text{orient_muchie}(x_2, y_2) = o_2, \dots, \\
& , \text{orient_muchie}(x_n, y_n) = o_n, |x_2 - x_1| = |\vec{dx}_1|, |x_3 - x_1| = |\vec{dx}_2|, \dots, \\
& |x_n - x_1| = |\vec{dx}_{n-1}|, |y_2 - y_1| = |\vec{dy}_1|, |y_3 - y_1| = |\vec{dy}_2|, \dots, \\
& |y_n - y_1| = |\vec{dy}_{n-1}|, \text{sgn}((x_2 - x_1)(y_2 - y_1)) = \text{sgn}(\vec{dx}_1 \cdot \vec{dy}_1), \dots \\
& , \text{sgn}((x_n - x_1)(y_n - y_1)) = \text{sgn}(\vec{dx}_{n-1} \cdot \vec{dy}_{n-1})\}
\end{aligned} \tag{20}$$

În (31), $o_1, o_2, \dots, o_n$ reprezintă orientările muchiilor în punctele de coordonate $(x_1, y_1), (x_2, y_2)$, respectiv (x_n, y_n) .

În studiul de față, s-a realizat implementarea matricelor de coocurență a nivelurilor de gri de ordinul doi, trei și cinci, respectiv a matricelor de coocurență a orientărilor muchiilor de ordinul doi și trei [44], [48].

➤ Îmbunătățirea metodelor destinate selecției trăsăturilor

Pentru a modela densități de probabilitate complexe, în [22] este propusă metoda de descompunere a acestora în mixturi finite de densități de probabilitate elementare. Cel mai adesea, densitățile de probabilitate se reprezintă prin intermediul mixturilor de distribuții gaussiene. Mixturile de densități de probabilitate se consideră a fi rezultatul următorului scenariu [22]: fie K surse ce generează date de diferite categorii. Fiecare categorie de date are asociată o densitate de probabilitate gaussiană de forma $p_m(y|\theta_m)$, unde $m=1,2,\dots,K$. Se presupune că datele provenite de la o anumită sursă sunt luate în considerare cu o probabilitate α_m . Astfel, densitatea de probabilitate a întregii mixturi va fi reprezentată matematic în felul următor:

$$p(y | \Theta_{(K)}) = \sum_{m=1}^K \alpha_m p_m(y | \theta_m) \tag{21}$$

Pentru a avea o determinare completă a mixturii, se pune problema atât a estimării parametrilor, cât și a numărului de componente. Algoritmi specifici învățării nesupervizate, precum cel de **Maximizare a Așteptării (Expectation Maximization)**, respectiv **Markov Chain Monte Carlo** sunt aplicați în acest scop [22].

Dacă un atribut asociat unui set de date poate fi reprezentat printr-o mixtură de k distribuții gaussiene, se cheamă că atributul respectiv este **k-modal**, deci relevant, adecvat separării între clase. În caz contrar, atributul este irelevant.

Această idee a fost exploatată și în scopul determinării parametrilor relevanți în contextul clasificării nesupervizate [69].

Propunem implementarea unei metode similare, atât în domeniul clasificării supervizate, cât și în cel al clasificării nesupervizate, redefinind indicele de relevanță astfel încât acesta să nu depindă doar de varianța atributului considerat, ci să indice mărimea suprafeței de intersecție a densităților de probabilitate corespunzătoare claselor considerate [D. Mitrea, 2008] [34], [35] .

Indicele de relevanta al atributului, în cazul separarii între două clase, se determină utilizând următoarea expresie matematica:

$$Indice_relevanta_bin = \frac{\min(p_1, p_2)}{\max(p_1, p_2)} \cdot \left(1 - \frac{(\mu_1 + 2\sigma_1) - (\mu_2 - 2\sigma_2)}{(\mu_2 + 2\sigma_2) - (\mu_1 - 2\sigma_1)} \right) \quad (22)$$

În formula de mai sus, μ_1 reprezintă valoarea minimă a mediilor per clasă ale atributului considerat, $\mu_1 < \mu_2$, σ_1 reprezintă deviatia standard asociată clasei cu media cea mai mică, μ_2 reprezintă valoarea maximă a mediilor per clasă ale atributului considerat, σ_2 reprezintă deviatia standard asociată clasei cu media cea mai mare. Scăzătorul expresiei din paranteză reprezintă dimensiunea zonei de suprapunere a densităților de probabilitate gaussiene ale celor două clase, fiind necesar ca această zonă să fie cât mai restrânsă, respectiv nulă, în cazul unei separări optime între clase (vezi Fig. 4.4). În ceea ce privește primul factor al expresiei, p_1 și p_2 reprezintă ponderile celor doi clusteri determinați prin algoritmul de clusterizare **Expectation Maximization**, aceste ponderi fiind, în cazul optim, egale între ele, având valoarea de 50%. Astfel, în cazul optim, indicele de relevanta va avea valoarea 1. Formula (37) se utilizează doar în cazul în care $(\mu_1 + 2\sigma_1) \geq (\mu_2 - 2\sigma_2)$, în caz contrar indicele de relevanță fiind egal cu primul factor (raportul între proporțiile clusterilor).

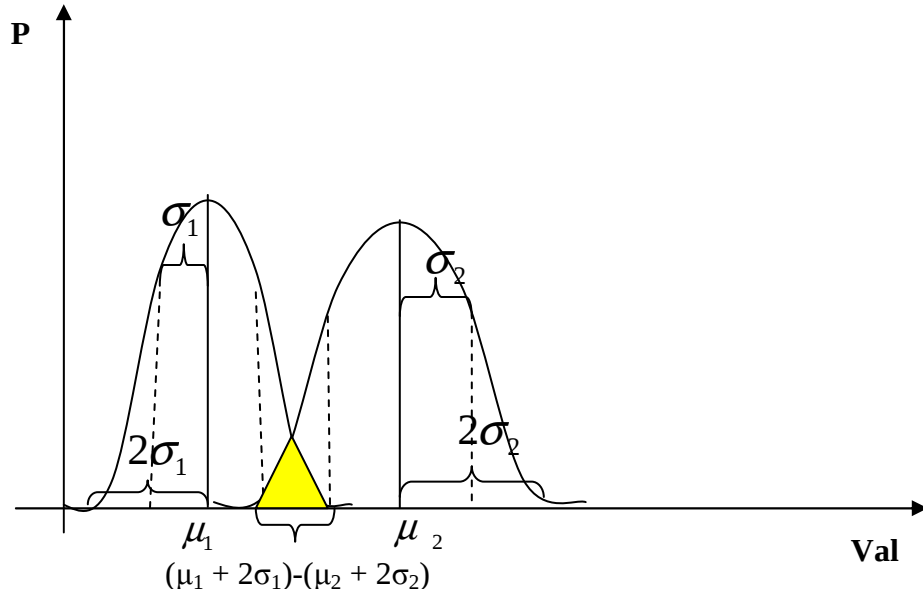


Figura 1. Reprezentarea zonei de intersecție a densitatilor de probabilitate gaussiene

În cazul general, al existenței a k clase, indicele de relevanță pentru un anumit atribut se va determina ca o sumă a suprafețelor de intersecție a tuturor perechilor de densități de probabilitate vecine. Astfel,

$$Indice_relevanta = \sum_{j=1}^k Indice_relevanta_bin(j, j') \quad (23)$$

unde j' este indicele clasei (densității de probabilitate din mixtură) care are media trăsăturii considerate cea mai apropiată de media clasei j .

Pentru a maximiza probabilitatea ca cele k distribuții să corespundă claselor reale, în cazul clasificării supervizate în care aceste clase se cunosc, parametrii mixturii trebuie inițializați cu parametrii asociați claselor (atât cu probabilitățile a-priori de apariție a claselor, cât și cu media și varianța fiecărei clase). În urma analizei fiecărui atribut în parte, acesta va fi reținut în sub-setul trăsăturilor relevante dacă este multi-modal și va fi respins dacă are mai puține componente decât numărul așteptat de clase [34], [35].

➤ **Stabilirea corelațiilor existente între trăsăturile texturale**

Utilizând metoda regresiei, se stabilește existența unor corelații de diferite tipuri între trăsăturile texturale. Sunt considerate redondante acele attribute texturale care sunt dependente de celelalte attribute texturale, printr-o anumită regulă sau lege [36] [D. Mitrea, 2008].

➤ **Definirea Modelului Imagistic Textural**

Astfel, **modelul imagistic textural al carcinomului hepatocelular (HCC)** constă în mulțimea trăsăturilor texturale relevante pentru caracterizarea acestei tumori maligne, considerate în reprezentarea lor optimă, respectiv în valorile statistice, specifice asociate acestor trăsături.

Pentru a defini modelul imagistic textural pornim de la un spațiu inițial al trăsăturilor texturale potențial relevante, determinate prin metode specifice de analiză a texturii,

$$F = \{f_i\}_{i=1,...,n} \quad (24)$$

urmărind să transpunem acest set de trăsături într-o reprezentare cât mai adecvată, astfel încât acestea să conducă la o acuratețe de recunoaștere cât mai bună. Avem în vedere, în această scop, următoarele demersuri:

- Transpunerea spațiului inițial de trăsături într-o reprezentare adecvată, care să evidențieze caracteristici ale acestor trăsături, precum varianta sau capacitatea lor de a separa între clase; vom folosi, pentru realizarea acestui obiectiv, metode de extragere a trăsăturilor, precum Metoda Analizei Componentelor Principale (PCA), metodele Analizei Discriminantilor Linari (LDA), respectiv a Analizei Discriminantilor Multipli (MDA).

- Selectia acelor trăsături ce pot fi considerate relevante pentru procesul de clasificare (separarea trăsăturilor relevante de cele nerelevante), prin metode specifice selecției trăsăturilor.

Aceste demersuri sunt ilustrate formal în (25):

$$F = \{\text{Trasaturi_texturale_potentiale_relevante}\};$$

$$F_R \subset F - \text{subspatiul trasaturilor or texturale relevante, in reprezentarea cea mai adecvata} \quad (25)$$

Ambele demersuri implică utilizarea metodelor destinate reducerii dimensionalității datelor, pentru a obține setul final de trăsături, F_R , atasat modelului imagistic textural, conform expresiei (26):

$$F_R = \{\text{trasaturi_texturale_relevante}\};$$

$$F_R \leq \text{Reducerea_Dimensionalitatii}(F) \quad (26)$$

Acestor trăsături le vom asocia valori specifice, precum valoarea medie, deviatia standard, distributia de probabilitate, precum și o măsură a relevanței fiecărui atribut. Modelul imagistic textural rezultat va avea reprezentarea formală descrisă în (27) [34], [35]:

$$TM = \{V_{f_r} \mid V_{f_r} = [\text{Relevanța}, \text{Media}, \text{Deviația standard}, \text{Distributia de probabilitate}]\} \quad (27)$$

3.2. Contribuții aplicative: implementarea Modelului Imagistic Textural

În vederea construirii modelului imagistic textural, este necesară, mai întâi, o **faza de generare** a modelului, constând din următorii pași: *analiza imaginii* și *învățare*, aceasta fiind urmată de o **faza de validare** a modelului. Într-o **etapa premergătoare**, sunt construite clasele de imagini ce vor fi supuse analizei și învățării, pentru o afecțiune dată, acestea constituind setul de antrenament. Clasele vor conține un număr egal de imagini. În timpul **etapei de analiza a imaginii**, se realizează determinarea trăsăturilor imagistice texturale, prin aplicarea metodelor specifice analizei texturii, descrise în cele ce urmează. În timpul **etapei de învățare a modelului**, este determinată relevanța trăsăturilor în ceea ce privește separarea între clase, utilizând metode specifice, precum modelarea densității de probabilitate prin mixturi de distribuții gaussiene, metode specifice selecției trăsăturilor de tip filtru sau wrapper, metoda rețelelor Bayesiene, arbori de decizie, metoda mașinilor cu vectori suport, dar și meta-clasificatori, precum Ada-Boost [16]. Metoda regresiei [23], precum și alte metode specifice sunt utilizate pentru a determina corelațiile dintre trăsături, în scopul stabilirii unui set final de trăsături relevante independente. Apoi, trăsăturile imagistice sunt analizate, stabilindu-se valori statistice specifice acestora, precum media valorilor, deviatia standard a acestora – acestea fiind calculate pentru un număr suficient de mare de cazuri din setul de antrenament. Parametrul *relevanță* este determinat în funcție de metoda utilizată pentru selecția trăsăturilor. În cadrul aceleiași etape, se are în vedere transpunerea reprezentării trăsăturilor texturale într-un spațiu de dimensiuni reduse, punându-se totodată în evidență proprietăți ale datelor precum capacitatea acestora de a separa între clase, respectiv principalele moduri de variație ale datelor. În acest scop, se vor folosi metode de extragere a trăsăturilor precum Analiza Componentelor Principale, Analiza Discriminanțelor Lineare, respectiv Analiza Discriminanțelor Multiple [70]. În **faza de validare a modelului**, se face determinarea acuratetei clasificatorilor aplicați pe un set de imagini de test, diferit de setul de învățare. La intrările clasificatorilor sunt furnizate trăsăturile relevante determinate în cadrul etapei de analiza ca fiind parte a modelului imagistic textural. Se evaluează, prin parametri specifici (rata de recunoaștere, specificitate, sensibilitate, suprafața de sub curba ROC) acuratetea recunoașterii automate în fiecare caz, pentru clasificatori de categorii diferite, iar rezultatele se compară cu cele obținute în cazul aplicării clasificatorilor, utilizând la intrare întreg setul inițial de trăsături [16]. Dacă acuratetea obținută nu este satisfăcătoare, se fac demersuri pentru îmbunătățirea modelului imagistic textural, atât prin perfecționarea la nivelul formării setului de antrenament, cât și la nivelul metodelor de analiză a imaginii, respectiv al celor destinate selecției trăsăturilor relevante [48] [D. Mitrea, 2011].

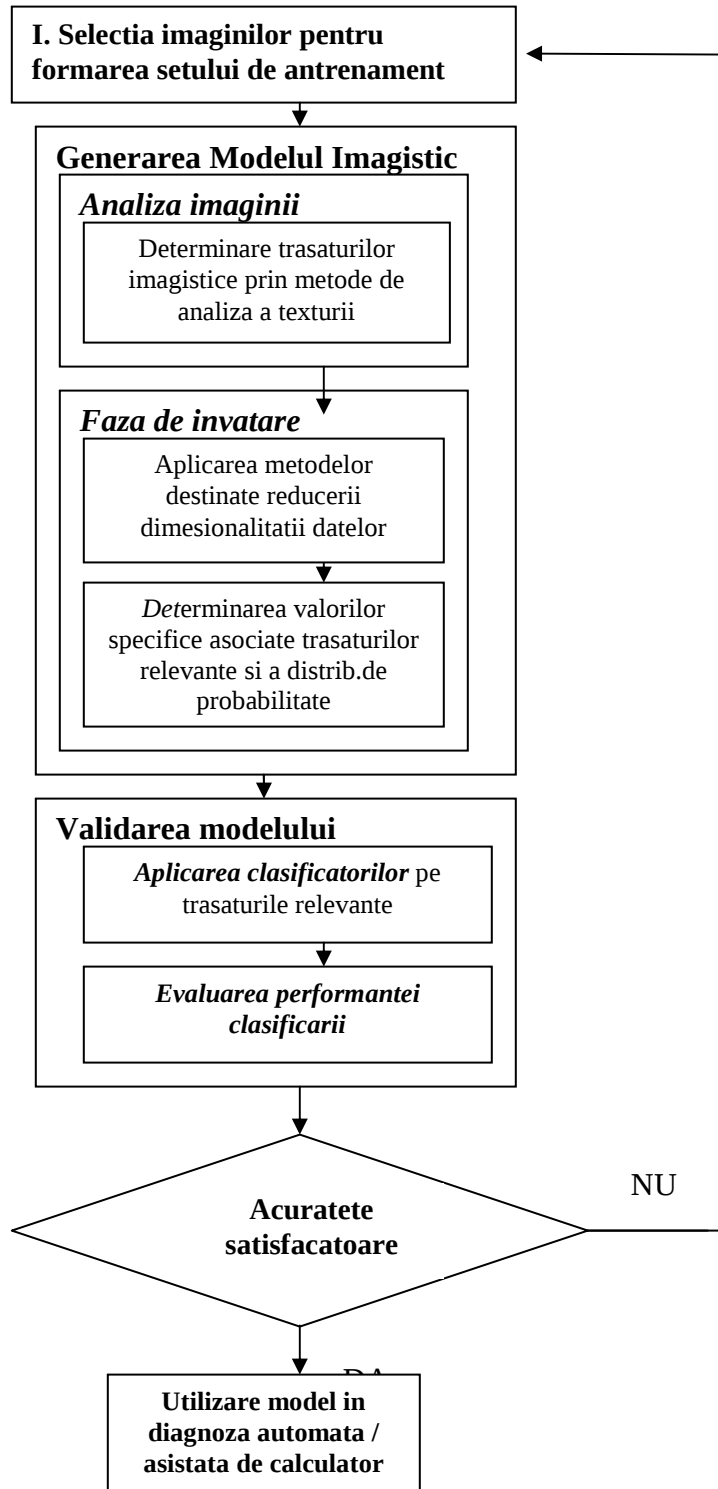


Figura 2. Etapele necesare generării modelului imagistic textural

▪ Etapa de analiză a imaginii

În cadrul etapei de analiză a imaginii, au fost aplicate metode specifice destinate analizei texturii, precum și combinații ale acestora. Într-o primă fază, au fost studiate, analizate și experimentate metodele existente, stabilindu-se rolul fiecărei clase de parametri texturali în caracterizarea afecțiunilor considerate. S-au dezvoltat și implementat următoarele metode de analiză a imaginilor, în scopul determinării parametrilor texturali ce pot caracteriza tumora malignă considerată:

- o Statisticile de ordinul întâi ale nivelurilor de gri: valoarea medie, valoarea maximă, valoarea minimă, deviația standard [21].

- o Matricea de co-ocurență a nivelurilor de gri (GLCM) și parametrii Haralick asociați: media GLCM, varianța GLCM, omogenitatea locală (momentul unghiular de ordinul doi), contrastul GLCM (inerția), corelația GLCM, energia, entropia [25]. S-a implementat și experimentat atât matricea de coocurență a nivelurilor de gri de ordinul doi, cât și matricele de ordin superior – trei, cinci și șapte. S-a studiat posibilitatea de apariție repetată a unei anumite grupări de pixeli, având anumite valori ale intensității (nivelurilor de gri), aflându-se într-o relație spațială bine definită, prin intermediul parametrului probabilitate maximă.

- o Funcția de autocorelație, pentru caracterizarea granularității texturale a țesutului, și pentru a studia dacă există diferențe de această natură între țesutul sănătos și țesutul bolnav, între țesutul hepatic și țesutul tumoral.

- o Statistici ale trăsăturilor de tip gradient: frecvența muchiilor, contrastul mediu al muchiilor, orientarea medie a muchiilor, variabilitatea în orientarea muchiilor, mărimea gradientului, varianța gradientului, mărimea gradientului direcțional, varianța gradientului direcțional, variabilitatea gradientului direcțional [21].

- o Determinarea microstructurilor texturale specifice unei anumite clase de țesut. Statistici ale microstructurilor texturale (nivel, muchie, pată, undă, undișoară), obținute în urma aplicării nucleelor de convoluție ale lui Laws [21].

- o Matrice de coocurență generalizate, precum și parametrii Haralick asociați. S-au elaborat metode specifice pentru construirea matricei de coocurență a orientărilor muchiilor, de ordinul doi și trei.

- o Metode bazate pe fractali: coeficientul Hurst [52].

- o Aplicarea Transformatei Wavelet [63], la două niveluri de rezoluție, și calculul entropiei Shannon pe fiecare componentă a imaginii astfel obținută.

Trăsăturile enumerate mai sus au fost determinate pe regiuni rectangulare, situate în interiorul formațiunii analizate (tumoră malignă sau benignă, parenchim cirotic, parenchim hepatic afectat de maladii hepatice difuze, perete intestinal), în mod independent față de iluminare și orientare [44].

▪ Îmbunătățirea performanțelor etapei de învățare a modelului

Îmbunătățirea etapei de învățare a modelului imagistic implică următoarele:

- o Evaluarea efectului celor mai performante metode de selecție a trăsăturilor (metoda de modelare a densității de probabilitate prin mixturi de distribuții gaussiene, metodele Correlation-based Feature Selection, Consistency-based Feature Subset Evaluation, Information Gain Attribute Ranking) asupra performanțelor procesului de clasificare: vom lua în considerare acele metode de selecție a trăsăturilor care au furnizat

rezultate optime conform cercetărilor anterioare și vom studia efectul pe care îl au aceste metode asupra îmbunătățirii procesului de clasificare, nu numai în ceea ce privește acurately clasificării, dar și în ceea ce privește forma curbelor de învățare (viteza de învățare a clasificatorului), stabilitatea și timpul de răspuns al clasificatorului [40], [41] [D. Mitrea, 2010].

O Evaluarea unor combinații ale metodelor de selecție a trăsăturilor, precum aplicarea succesivă a acestora, reuniunea și intersecția: vom studia și compara acurately clasificatorilor care primesc la intrare subseturile de trăsături relevante rezultate în fiecare dintre aceste cazuri [37] (D. Mitrea, 2009).

O Implementarea metodelor destinate extragerii trăsăturilor (metoda Analizei Componentelor Principale – PCA și extensiile acesteia la cazul non-liniar, kernel PCA, precum și Expectation-Maximization PCA; metoda Analizei Discriminanțelor Linare – LDA, metoda Analizei Generalizate a Discriminațelor – GDA, metoda Analizei Discriminanțelor Multiple – MDA) și realizarea unor combinații ale acestora, precum și combinarea acestora cu metode pentru selecția trăsăturilor [47], respectiv evaluarea efectului acestor metode asupra performanțelor procesului de clasificare: vom studia efectul acestora asupra îmbunătățirii acurately, vitezei și respectiv stabilității procesului de învățare [38], [39] [D. Mitrea, 2009].

▪ **Descrierea fazei de validare a modelului**

Validarea modelului imagistic textural implică furnizarea setului de trăsături relevante rezultate în urma fazei de învățare la intrările unui clasificator adecvat și evaluarea performanțelor acestui clasificator. În acest scop, este necesară determinarea acelor clasificatori care sunt cei mai performanți în contextul recunoașterii maladiilor considerate, acest demers conducând, de asemenea, la rafinarea procesului de recunoaștere automată a acestor maladii. Avem în vedere compararea următorilor clasificatori, cunoscuți pentru performanțele lor, aparținând unor categorii variate de metode de clasificare [16]: Metoda Rețelelor Bayesiene de Încredere, Arbori Decizionali – Metoda C4.5, Rețele Neuronale Artificiale – Metoda Perceptronului Multinivel, Mașini cu Vectori Suport (Support Vector Machines – SVM), cu nucleu polinomial și radial, RBF (Radial Basis Function), Combinații ale acestor clasificatori conform următoarelor scheme de combinare: votare, bagging, boosting, generalizare prin stivuire (stacking). Acești clasificatori vor fi comparați utilizând subseturi de date variate din categoria maladiei analizate și a Țesuturilor cu caracteristici vizuale asemănătoare.

• **Îmbunătățirea fazei de validare a modelului imagistic textural**

Îmbunătățirea etapei de validare a modelului imagistic textural și a procesului de recunoaștere automată, implică următoarele demersuri:

- o Realizarea unor combinații de clasificatori care să îmbunătățească performanțele procesului de clasificare, atât sub aspectul acurately, cât și sub aspectul stabilității, al vitezei de învățare, respectiv al timpului de răspuns al acestora [40] [D. Mitrea, 2010].
- o Evaluarea clasificatorilor individuali pe seturi extinse de date, din punctul de vedere al acurately, vitezei procesului de învățare și respectiv de recunoaștere, dar și al stabilității acestui proces [41]

- o Impartirea carcinomului hepatocelular (HCC) in subclase; realizarea unei clasificari multiclase [28], în scopul creșterii performanțelor clasificării [81] [D. Mitrea, 2010].
- o Utilizarea metodelor de grupare (clustering) in scopul detectiei si caracterizarii prin parametri texturali specifici a subclaselor texturale ale HCC care sunt dificil de diferentiat cu ochiul liber [D. Mitrea, 2010]. [42]

4. Rezultate și discuții

Prin intermediul aplicației software implementate în vederea construirii modelului imagistic textural al unor afecțiuni tisulare reprezentative, s-au generat următoarele modele imagistice texturale: modelul imagistic textural al adenocarcinomului prostatic (ADKP) [D. Mitrea, 2008], modelele imagistice texturale corespunzătoare stadiilor de evoluție care premerg apariția carcinomului hepatocelular (ficat normal, hepatită, ciroză) [33], precum și cel corespunzător carcinomului hepatocelular [D. Mitrea, 20011] [44]. De asemenea, s-au construit modelele imagistice texturale ale bolilor inflamatorii intestinale (boala Crohn, recto-colita ulcero-hemoragică) [46], precum și al tumorilor colonice [48] [D. Mitrea, 2011]. Cele mai semnificative rezultate în acest sens, implicând, de asemenea, noile trăsături texturale, sunt ilustrate în cele ce urmează

4.1. Descrierea fomei finale a modelului imagistic textural în cazul unor afecțiuni semnificative

➤ Forma finală a modelului imagistic textural al carcinomului hepatocelular (HCC)

Setul final al trăsăturilor texturale relevante în ceea ce privește caracterizarea structurii carcinomului hepatocelular (HCC), care asigură diferențierea, cu maximă acuratețe a acestuia în raport cu țesuturile asemănătoare (parenchim cirotic înconjurător, tumori hepatice benigne), obținut după reuniunea seturilor de date corespunzătoare tuturor claselor și aplicarea unora dintre cele mai eficiente metode de selecție a trăsăturilor (Correlation based Feature Selection, respectiv Information Gain Attribute Evaluation), este descris mai jos, în (28).

$$F_R = \{GLCM_Entropie, GLCM3_Omogenitate, EOCM3_Entropie, GLCM7_Energie, GLCM5_Contrast, GLCM_Corelație, Variabilitatea_gradientului_direcțional, Frecvență_pete, Media_nivelurilor_de_gri, Wavelet_entropia_7_hh, Wavelet_Entropia7_hl, Wavelet_entropia8_hl, Wavelet_Entropia6_hh\} \quad (28)$$

Distribuția de probabilitate a celor mai reprezentative trăsături, derivată prin intermediul metodei Rețelelor Bayesiene de Încredere, este redată în tabelele 1, 2 și, respectiv 3.

Tabelul 1. Distribuția de probabilitate în cazul parametrului *Corelația GLCM*

Corelația GLCM	(-∞, 0.698]	(0.698, ∞)
HCC	0.603	0.397
Tumori benigne	0.361	0.639
Parenhim cirotic	0.699	0.301

Se observă, conform Tabelului 106, că în cazul parenchimului cirotic și, respectiv al carcinomului hepatocelular, corelația GLCM este mai mică, deci și granularitatea țesutului este mai mică, probabil datorită prezenței unor noduli de mici dimensiuni, în ambele cazuri.

Tabelul 2. Distribuția de probabilitate în cazul parametrului *Energia GLCM7*

Energia GLCM7	$(-\infty, 0.000267]$	$(0.000267, 0.000455)$	$(0.000455, \infty)$
Tumori benigne	0.051	0.401	0.549
HCC	0.158	0.724	0.118
Parenhim cirotic	0.003	0.865	0.131

Conform Tabelului 2, energia GLCM, indicând uniformitatea distribuției nivelurilor de gri în cadrul texturii, are valoare minimă în cazul HCC, fapt relevant pentru caracterul haotic al țesutului tumoral.

Tabelul 3. Distribuția de probabilitate în cazul parametrului *Variabilitatea gradientului direcțional*

Variabilitatea gradientului direcțional	$(-\infty, 0.000834]$	$(0.000834, 0.002036]$	$(0.002036, \infty)$
Tumori benigne	0.297	0.644	0.058
HCC	0.057	0.819	0.084
Parenhim cirotic	0.01	0.98	0.01

Din Tabelul 3, rezultă că variabilitatea gradientului direcțional are valori maxime în cazul tumorii HCC. În tabelul 4 sunt descrise valorile specifice clasei HCC pentru cele mai reprezentative trăsături texturale, acestea completând modelul imagistic textural al HCC.

Tabelul 4. Valorile specifice trăsăturilor texturale relevante, în cazul clasei HCC

Trăsătura texturală	Indice de relevanță	Valoare medie	Deviație standard	Interval de variație
Corelația GLCM	0.0575	0.61	0.036	$(-\infty, 0.698]$
Energia GLCM7	0.2019	0.00019	0.00008	$(-\infty, 0.000267]$
Variabilitatea gradientului direcțional	0.1817	0.00212	0.0002	$(0.002036, \infty)$
Wavelet_Entropia7_hl	0.1014	-35.84	5.53	$(-43.525, \infty)$

➤ Forma finală a modelului imagistic textural al tumorilor colo-rectale

În (29) este descris setul final al trăsăturilor texturale relevante în ceea ce privește caracterizarea și recunoașterea tumorilor colo-rectale. Acesta a rezultat ca o reuniune a seturilor de trăsături obținute prin aplicarea metodelor individuale.

$$\begin{aligned}
 \mathbf{F_R} = \{ & \text{GLCM_Entropy, Mean_grey_levels, Autocorrelation_index, Edge_Contrast,} \\
 & \text{Wavelet_Entropy1, Wavelet_Entropy2, Wavelet_Entropy4, Laws_level_mean,} \\
 & \text{Laws_edge_frequency, Laws_spot_frequency, Laws_wave_mean,} \\
 & \text{Laws_wavelet_frequency, GLCM5_homogeneity_horiz_vert,} \\
 & \text{GLCM5_energy_horiz_vert, GLCM3_homogeneity_0_270,} \\
 & \text{Variability_directional_gradient, Laws_Edge_mean, GLCM5_Correlation_horiz_vert,} \\
 & \text{GLCM3_Contrast_0_270 } \}
 \end{aligned}
 \tag{29}$$

Conform Tabelului 5, care descrie densitatea de probabilitate a parametrului *Indice de autocorelație*, rezultă că acesta are valoare minimă în cazul tumorilor colo-rectale și valoare maximă în celelalte cazuri.

Tabelul 5. Distribuția de probabilitate în cazul parametrului *Indicele de autocorelație*

Indicele de autocorelație	$(-\infty, 0.00031]$	$(0.00031, 0.00073]$	$(0.00073, \infty)$
IBD	0.043	0.54	0.416
Tumori colo-rectale	0.453	0.54	0.006

Conform Tabelului 6, care descrie densitatea de probabilitate a parametrului *omogenitate*, derivat din matricea GLCM de ordinul trei, rezultă că acesta are valoare minimă în cazul tumorilor colo-rectale și valoare maximă în celelalte cazuri, acest lucru fiind în conformitate cu caracterul eterogen al țesutului tumoral malign.

Tabelul 6. Distribuția de probabilitate în cazul parametrului *GLCM3 Omogenitate*

<i>GLCM3_omogenitate</i>	$(-\infty, 0.0208]$	$(0.0208, \infty)$
IBD	0.144	0.856
Tumori colo-rectale	0.569	0.431

În Tabelul 7 sunt descrise valorile specifice clasei tumorilor colo-rectale, pentru cele mai reprezentative trăsături texturale, care completează modelul imagistic textural al acestor tumori.

Tabelul 7. Valorile specifice trăsăturilor texturale relevante pentru caracterizarea tumorilor maligne colo-rectale

Trăsătura texturală	Indice de relevanță	Valoare medie	Deviație standard	Interval de variație
Indice de autocorelație	0.359	0.00025	0.000146	$(-\infty, 0.00031]$
GLCM3_omogenitate	0.153	0.0186	0.010	$(-\infty, 0.0208]$

➤ Forma finală a modelului imagistic textural al Bolilor Inflamatorii Intestinale (BII): boala Crohn și recto-colita ulcero-hemagoragică

În (30) este descris setul final al trăsăturilor texturale relevante în ceea ce privește diferențierea dintre boala Crohn și recto-colita ulcero-hemoragică. Acesta a rezultat ca o reuniune a seturilor de trăsături obținute prin aplicarea metodelor individuale.

$\mathbf{F_R} = \{ \text{GLCM variance, Edge contrast, Wavelet_Entropy2, Wavelet_entropy3, Wavelet_entropy8_lh, Wavelet_entropy5_hl, GLCM3_contrast_0_90, GLCM3_entropy_0_90, GLCM5_homogeneity_diag, GLCM5_correlation_horiz_vert, GLCM5_variance_diag, Directional gradient magnitude, Directional gradient variance, Edge frequency, Edge contrast, Laws wave frequency Wavelet_entropy2, Wavelet_entropy8_ll, GLCM5_contrast_diag, GLCM5_variance_diag} \}$	(30)
---	------

În Tabelul 8 sunt descrise valorile specifice trăsăturilor texturale reprezentative pentru diferențierea BII, considerând cazul fiecărei clase, acestea completând modelul imagistic textural al acestor afecțiuni.

Tabelul 8. Valorile specifice trăsăturilor texturale relevante în cazul BII

Trăsătura texturală	Indice de relevanță	Valoare medie pt. boala Crohn	Valoare medie pt. recto-colita ulcero-hemoragică
Contrastul muchiilor	0.142	31.67	41.77
Varianța GLCM5	0.395	587.25	622.65

4.2. Rezultate finale și discuții

În urma construirii modelelor imagistice texturale ale afecțiunilor considerate, au rezultat o multitudine de aspecte referitoare la trăsăturile texturale importante în diferențierea maladiilor, care au furnizat informații cu privire la proprietățile structurale ale țesutului, în fiecare caz.

Astfel, în ceea ce privește trăsăturile texturale relevante în caracterizarea maladiilor hepatice difuze, indicele de atenuare a avut un rol special în cazul diferențierii steatozei de ficatul normal și de celelalte afecțiuni hepatice. Media nivelurilor de gri, parametrii matricei de coocurență a nivelurilor de gri (GLCM) precum omogenitatea, corelația, energia și entropia, precum și frecvența muchiilor, s-au dovedit a fi importanți atât în ceea ce privește diferențierea maladiilor hepatice difuze în general, cât și în ceea ce privește caracterizarea gradelor de severitate în cazul steatozei [D. Mitrea, 2008] [33]. Indicele de autocorelație a

devenit important în contextul apariției cirozei și al evoluției acestei maladii spre tumoarea malignă, carcinomul hepatocelular (HCC) [33]. În ceea ce privește trăsăturile texturale relevante în caracterizarea tumorilor maligne, în toate cazurile analizate – carcinom hepatocelular (HCC), respectiv adenocarcinom prostatic (ADKP), tumori colo-rectale, s-a evidențiat indicele de autocorelație, indicând o creștere a granularității texturii în cazul tumorilor maligne, față de stadiile premergătoare acestora, respectiv față de cazul tumorilor benigne; de asemenea, au rezultat ca fiind importante energia GLCM, omogenitatea GLCM, entropia GLCM, indicând scăderea omogenității țesutului în cazul tumorilor și creșterea gradului de entropie (dezordine) al distribuției nivelurilor de gri, în conformitate cu caracterul haotic al acestor tumori. Statisticile referitoare la distribuția trăsăturilor locale (muchii, curburi) și a microstructurilor texturale (în special celor de tip pată, respectiv undisoară) au avut și acestea un rol important în aceste cazuri, tumorile maligne caracterizându-se printr-o frecvență, proeminentă și variabilitate crescută a acestora, conform complexității lor structurale. Media nivelurilor de gri s-a dovedit a fi un parametru relevant în toate cazurile, indicând creșterea ecogenității parenchimului hepatic datorită acumulării de grăsime în cazul steatozei, prezentând, de asemenea, o creștere în cazul apariției fibrozei (la hepatită, respectiv la ciroză), precum și hiperecogenitatea a tumorilor maligne în stadiu evoluat, respectiv scăderea ecogenității în cazul tumorilor maligne incipiente. Totuși, așa cum se poate deduce, media nivelurilor de gri este un parametru nespecific unei afecțiuni date. Trăsăturile multirezolutive (entropia calculată pe componentele Transformatei Wavelet) au indicat o creștere a dezordinii nivelurilor de gri și a caracteristicilor trăsăturilor locale la rezoluții multiple, fiind calificate, de către metoda de modelare a densității de probabilitate prin mixturi de distribuții gaussiene, ca fiind cele mai relevante trăsături (indici de relevanță de valori maxime). Reprezentarea grafică a densității de probabilitate în cazul entropiei calculate pe prima subimagine (componentă) a Transformatei Wavelet (low-low) de la primul nivel, determinată în cazul HCC, se regăsește în Figura 3. Modelarea acesteia prin mixturi de distribuții gaussiene indică faptul că aceasta este o trăsătură bimodală, deci relevantă [D. Mitrea, 2008] [33], [34], [35], [36].

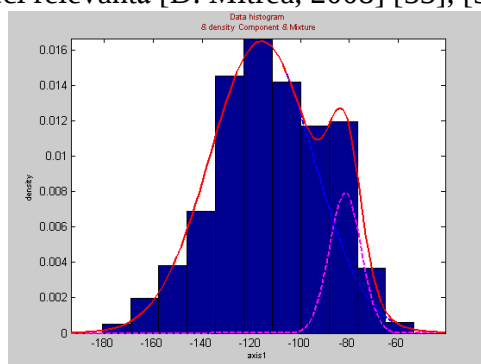


Figura 3. Reprezentarea densității de probabilitate pentru entropia wavelet pe prima componentă (low-low) la primul nivel: trăsătura bimodală (relevantă)

Referitor la metodele destinate reducerii dimensionalității datelor, a rezultat faptul că metodele destinate reducerii dimensionalității trăsăturilor conduc, în general, la diminuarea timpului de execuție al clasificatorilor, la creșterea stabilității, precum și la creșterea acurateții clasificatorilor [D. Mitrea, 2009] [38], [39] [D. Mitrea, 2011] [47].

În urma experimentării și comparației clasificatorilor, au rezultat următoarele concluzii [D. Mitrea, 2010] [40], [41]:

- Scheme de combinare a clasificatorilor precum bagging, boosting și stacking conduc la creșterea acurateții și stabilității clasificatorilor.
- Scheme de combinare a clasificatorilor precum boosting conduc la creșterea acurateții clasificatorilor.

Așa cum a rezultat din experimentele considerate, noile metode dezvoltate pentru analiza texturală, bazate pe matricele de coocurență de ordin superior, au condus la o creștere considerabilă a acurateții de clasificare, precum și la o mai bună evidențiere a caracteristicilor tumorilor maligne, comparativ cu cele ale tumorilor benigne, precum și cu cele ale țesuturilor vizual similare. În ceea ce privește combinațiile de direcții ale vectorilor de deplasament care au condus la o acuratețe de recunoaștere maximă, s-au dovedit a fi eficiente în special combinațiile de direcții orizontale și verticale ale acestora (acestea fiind, de asemenea, paralele sau perpendiculare pe direcțiile de propagare a semnalului ultrasonografic). Combinațiile dintre vechile trăsături texturale, precum și dintre noile trăsături texturale, calculate pe baza matricelor de coocurență de ordin superior, au condus, în majoritatea cazurilor, la o creștere considerabilă a acurateții de clasificare, în comparație cu utilizarea doar a trăsăturilor texturale anterioare. Trăsăturile texturale rezultate din matricele de coocurență a nivelurilor de gri ordinul trei și cinci au condus la o creștere mai accentuată a acurateții decât cele rezultate din matricea de coocurență a nivelurilor de gri de ordinul șapte. Procesul de combinare a trăsăturilor texturale anterioare cu trăsături derivate din matricele de coocurență a nivelurilor de gri de ordinul trei, cinci și șapte, respectiv cu trăsăturile provenite din matricele de coocurență a muchiilor de ordinul doi și trei, urmat de un proces de selecție a trăsăturilor, a condus la o creștere a acurateții de recunoaștere de la 70% la 88% în cazul diferențierii carcinomului hepatocelular de parenchimul cirotic înconjurător și de tumorile hepatice benigne, la o îmbunătățire a acurateții de recunoaștere de la 68% la 75% în cazul diferențierii maladiilor inflamatorii intestinale, respectiv la o creștere a acurateții de recunoaștere de la 80% la 90% în cazul recunoașterii tumorilor colo-rectale. Aceste rezultate sunt completate de către cele referitoare la creșterea acurateții, prin adăugarea trăsăturilor texturale derivate din matricele de coocurență generalizate de ordin superior.

Figura 4 reflectă creșterea ratei de recunoaștere a clasificatorilor celor mai performanți, datorită utilizării noilor trăsături texturale (atât variabilitatea orientării muchiilor, variabilitatea gradientului direcțional, cât și trăsăturile texturale derivate din matricele de coocurență de ordin superior), în toate cazurile considerate (pentru opt seturi de date). Cele opt seturi de date luate în considerare au fost: cele patru perechi de clase de texturi din colecția Brodatz, perechea de clase HCC/Parenchim cirotic din jurul HCC, HCC/Tumori hepatice benigne (TuB), Boli Inflamatorii Intestinale (BII) – boala Crohn și rectocolita ulcero-hemoragică, respectiv tumorile colo-rectale (cancerul colonic –CC) și BII. Probabilitatea de apariție a aceleiași grupări de niveluri de gri în interiorul unei anumite clase de țesut scade odată cu creșterea ordinului de mărime al matricelor de coocurență. Aceasta a avut, de obicei, valori mai mici în interiorul tumorilor maligne și valori mai mari

în cazul claselor de țesut similar, indicând caracterul haotic al tumorilor maligne [D. Mitrea, 2011] [48].

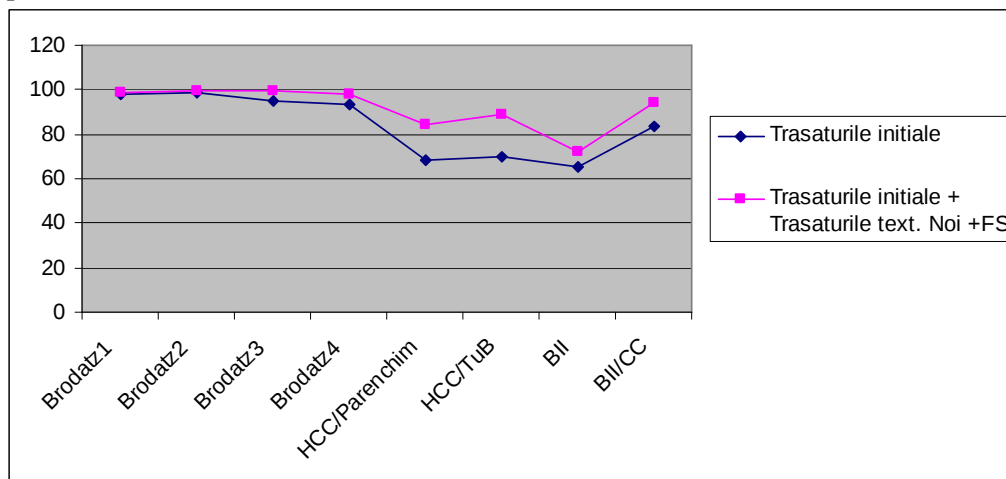


Figura 4. Evaluarea creșterii ratei de recunoaștere a clasificatorilor, în urma adăugării trăsăturilor texturale noi, și a aplicării metodelor pentru selecția trăsăturilor

Grupările de pixeli de valori specifice au corespuns unor zone de țesut tumoral pur sau unor regiuni afectate de necroză în cazul tumorilor maligne, unor lacuri vasculare sau unor zone dense în capilare sanguine în cazul tumorilor benigne, respectiv nodulilor displazici în interiorul parenchimului cirotic din jurul HCC. Astfel, configurația celor n pixeli corespunzători grupărilor cu apariția cea mai frecventă a corespuns, probabil, cu forma unor structuri specifice din componența fiecărui țesut, iar valorile nivelurilor de gri ale acestora au corespuns cu natura fiecărei astfel de structuri. Un număr mai mic de pixeli având configurația spațială corespunzătoare unei forme înscrise în structura respectivă a fost în măsură a caracteriza cu o mai mare exactitate aceste structuri specifice, de formă neregulată și având un aspect textural neomogen. Cele mai eficiente combinații de direcții ale vectorilor de deplasament, precum și distanțele optime stabilite între pixeli, au corespuns dimensiunii și formei celulelor de textură în cazul texturilor din lumea reală reprezentate în colecția Brodatz, respectiv dimensiunii și formei elementelor structurale din componența țesuturilor organelor interne. În cazul imaginilor medicale, cele mai potrivite direcții ale vectorilor de deplasament au rezultat a fi direcțiile orizontală și verticală, datorită orientării în plan orizontal a majorității elementelor structurale ale țesutului organelor analizate, aceste direcții fiind, de asemenea, paralele, respectiv perpendiculare pe direcția de propagare a semnalului ultrasonografic.

5. Concluzii și posibilități de dezvoltare ulterioară

Cercetarea realizată în contextul tezei de doctorat a constatat în următoarele demersuri fundamentale:

- Definirea, experimentarea și validarea unor noi trăsături texturale, obținute în urma dezvoltării unor noi metode de analiză a texturii, respectiv prin extinderea și generalizarea metodelor existente, aceasta constituind contribuția în domeniul teoretic al analizei bazate pe textură a imaginilor. [48]
- Definirea și construirea modelului imagistic textural al unor afecțiuni abdominale reprezentative, în scopul recunoașterii automate a acestora, în manieră optimă, pe cale non-invazivă, pentru a înlocui biopsia reală cu una virtuală. În acest context, s-a realizat

studiul și experimentarea unor metode performante din domeniile analizei bazate pe textură a imaginilor și recunoașterii statistice formelor, precum și îmbunătățirea metodelor existente în aceste domenii și dezvoltarea unor noi metode.

În mod concret, s-au adus următoarele contribuții :

- Studiul stadiului actual al cercetărilor în domeniul metodelor de caracterizare și analiză a texturii, în general, precum și în domeniul aplicativ al recunoașterii imaginilor medicale prin intermediul texturii. Analiza rolului metodelor de reducere a dimensionalității datelor, al clasificatorilor elementari și al combinațiilor de clasificatori în imagistica medicală.
- Studiul metodelor destinate analizei texturii, a metodelor de reducere a dimensionalității datelor, a clasificatorilor elementari, a schemelor de combinare a clasificatorilor, a metodelor de clasificare multiclase.
- Experimentarea și compararea metodelor existente pentru analiza texturii, în special în contextul caracterizării și recunoașterii unor afecțiuni abdominale importante, pe baza imaginilor ultrasonografice.
- Definirea unor metode noi de analiză a texturii: variabilitatea orientării muchiilor, variabilitatea gradientului direcțional, și extinderea matricei de coocurență a nivelurilor de gri, respectiv a matricei de coocurență a orientărilor muchiilor la ordin superior ($n > 2$). Aceste metode s-au dovedit a fi adecvate și în contextul caracterizării și recunoașterii maladiilor care produc modificări tisulare, pe baza imaginilor ultrasonografice.
- Caracterizarea țesutului afectat de maladie prin intermediul texturii; stabilirea corelațiilor existente între trăsăturile texturale, respectiv între valorile trăsăturilor texturale și proprietățile vizuale și structurale ale țesutului; determinarea unor formațiuni specifice structurii țesutului afectat de maladie, prin intermediul microstructurilor texturale, respectiv al coocurenței pixelilor cu proprietăți specifice (grupări de niveluri de gri, grupări de muchii având orientări specifice).
- Experimentarea și compararea metodelor destinate selecției trăsăturilor, precum și a celor destinate extragerii trăsăturilor; combinarea acestor metode; stabilirea metodei, respectiv a combinației celei mai performante.
- Experimentarea și compararea clasificatorilor elementari, a combinațiilor de clasificatori și a metodelor de clasificare multiclase, în scopul maximizării performanțelor recunoașterii unor afecțiuni abdominale reprezentative.
- Împărțirea tumorilor maligne (HCC) în subclase, în funcție de caracteristicile vizuale ale acestora, în scopul îmbunătățirii acurateței procesului de clasificare; studiul și experimentarea metodelor de clasificare multiclase în acest scop; stabilirea trăsăturilor texturale care guvernează procesul de împărțire al HCC în subclase; stabilirea corelației dintre subclasele obținute și fazele de evoluție a tumorii [42].

Asa cum rezultă din experimentele ilustrate în Capitolul V, noile metode dezvoltate în vederea analizei texturale, pe baza cărora au rezultat următoarele trăsăturile texturale *variabilitatea orientărilor muchiilor, variabilitatea gradientului direcțional*, precum și *trăsăturile derivate din matricele de coocurență a nivelurilor de gri și a orientărilor muchiilor de ordin superior*, au contribuit la îmbunătățirea procesului de caracterizare și recunoaștere a texturilor, conducând la creșterea acurateței de clasificare, în comparație cu cazul utilizării doar a metodelor existente. Acest lucru a fost demonstrat utilizând eșantioanele texturale aparținând bazei de date Brodatz [11], disponibilă gratuit prin accesarea pe Internet. Metoda de selecție a trăsăturilor bazată pe modelarea densității de

probabilitate prin mixturi de distribuții Gaussiene s-a dovedit a fi deosebit de eficientă în ceea ce privește determinarea trăsăturilor relevante pentru caracterizarea afecțiunilor abdominale analizate, conducând la îmbunătățirea acurateții recunoașterii acestora.

În ceea ce privește rolul trăsăturilor bazate pe textură în caracterizarea și diagnoza automată a maladiilor hepatice difuze, a tumorilor maligne, a bolilor inflamatorii intestinale și respectiv a cancerului colonic, acestea s-au dovedit a fi deosebit de utile în aceste domenii: ratele de recunoaștere ale clasificatorilor celor mai performanți, în conjuncție cu trăsăturile texturale selectate ca fiind relevante, sunt situate în apropierea valorii de 85%. Noile trăsături texturale definite – variabilitatea în orientarea muchiilor și variabilitatea gradientului direcțional, precum și cele derivate din matricele de coocurență a nivelurilor de gri și a orientărilor muchiilor de ordin superior, au condus la o caracterizare mai rafinată a claselor de țesut analizate, punând mai bine în evidență elementele structurale din componența fiecărui tip de țesut. De asemenea, în combinație cu vechile trăsături texturale, după aplicarea unor metode eficiente destinate selecției trăsăturilor și a unor clasificatori performanți, acestea au condus la o creștere a ratei de recunoaștere de până la 88% în cazul recunoașterii HCC, respectiv peste 90% în cazul recunoașterii tumorilor colo-rectale.

Procesul de selecție a trăsăturilor relevante, ghidat de utilizarea metodelor pentru selecția trăsăturilor, a condus, în majoritatea cazurilor, la îmbunătățirea performanțelor procesului de recunoaștere automată a texturilor pe baza imaginilor ultrasonografice. Cele mai eficiente metode destinate selecției trăsăturilor s-au dovedit a fi metoda bazată pe modelarea trăsăturilor prin mixturi de distribuții gaussiene, metoda de selecție bazată pe corelația trăsăturilor texturale cu parametrul clasă (Correlation-based Feature Selection-CFS), metoda de selecție bazată pe evaluarea consistenței sub-seturilor de attribute (Consistency based Feature Subset Evaluation), respectiv metoda de evaluare a atributelor individuale prin intermediul câștigului informațional (Information Gain Attribute Evaluation). Metodele de extragere a trăsăturilor au contribuit la reducerea timpului de clasificare, precum și la creșterea acurateții clasificatorilor, în unele cazuri, datorită punerii în evidență a unor proprietăți ale datelor de tipul abilității acestora de a separa între clase, realizând totodată și eliminarea zgomotului.

În ceea ce privește procesul de clasificare, după o comparație riguroasă a clasificatorilor, cele mai performante astfel de metode s-au dovedit a fi cea a Mașinilor cu Vectori Suport (SVM) cu nucleu polinomial de grad trei, cea a Perceptronului Multinivel (MLP), cele bazate pe arbori decizionali (J48, echivalentul Weka al metodei C4.5; Random Forest). Schemele de combinare a clasificatorilor au contribuit la creșterea acurateții și stabilității procesului de clasificare, cele mai eficiente în acest sens fiind bagging, boosting (AdaBoost) și metoda de stivuire prin generalizare (stacking).

Posibilitățile de continuare a cercetării în domeniul identificării unor structuri anatomico-patologice, precum și al caracterizării unor stări patologice reprezentative prin intermediul texturii, sunt grupate în jurul următoarelor idei:

- Experimentarea metodelor de analiză texturală, în special a celor nou dezvoltate, precum și a modelului imagistic textural, pe mai multe tipuri de imagini medicale: Imagini de Rezonanță Magnetică (MRI), tomografii computerizate (CT), imagini endoscopice, precum și luând în considerare mai multe maladii. Compararea rezultatelor obținute pe tipurile de imagini considerate.

- Studiul subclaselor texturale care există în interiorul tumorilor maligne. Utilizarea unor metode de explorare care să conducă la o caracterizare cât mai adecvată a structurii tumorale și la identificarea stadiilor de evoluție ale acesteia.
- Localizarea tesutului afectat de maladie, în cadrul organului corespunzător; stabilirea marginilor (limitelor de extindere) pentru afecțiunea considerată, în special în cazul tumorilor maligne, ca parte a procesului de biopsie virtuală; elaborarea, în acest scop, a unor metode de segmentare bazate pe reguli, utilizând trasaturile texturale relevante și valorile specifice asociate acestora, deci, modelele imagistice texturale generate pentru fiecare afecțiune.

Cercetarea descrisă în această teză a fost realizată în contextul a șapte contracte de cercetare – dintre care un grant de tip CNCSIS TD la care am fost director de proiect (2005-2007), cinci granturi de tip CEEEX (2006-2008), respectiv un grant de CNCSIS de tip 'IDEI' (2009-2011). Contribuțiile aduse au fost diseminate sub forma a treizeci și patru de articole științifice, dintre care patru publicații în reviste ISI, nouă publicații în reviste de tip CNCSIS B+, șase articole publicate în volume de conferințe de tip ISI Proceedings românești sau străine, șase articole publicate în volumele unor alte conferințe indexate în baze de date internaționale (IEEE, Springer), respectiv nouă lucrări științifice publicate în reviste și volume de conferințe cu referenți (neindexate).

5. Referințe bibliografice

- [1] M. Adan, J. Barcelo, J. Pijoan-Lopez, "Spatial statistics in archaeological texture analysis", in *The Digital Heritage of Archaeology. Computer Applications and Quantitative Methods in Archaeology* (Hellenic Ministry of Culture), 2003, Online: <http://prehistoria.uab.cat/publication/Adanetal.pdf>
- [2] F. Aiolli, A. Sperduti, "Multiclass classification with multi-prototype Support Vector Machines", *Journal of Machine Learning Research*, Vol. 6, 2005, pp. 817–850
- [3] A. Akono, N. Talla Tankam, E. Tonye, "Nouvel algorithme d'évaluation de texture d'ordre n sur la classification de l'occupation des sols de la région volcanique du mont Cameroun". *Teledetection*, Vol. 5, No. 1, 2005, pp. 227-244
- [4] Nicoleta Angelescu, Daniela Alexandra Crisan, R.Dobrescu, D.Popescu, „Comparing methods for image texture analysis: a survey” , *Proc.of 15-th International Conference on Control Systems and Computer Science*, vol.3, 2005, pp. 69-74
- [5] H. Anyas, D. He, "Evaluation of textural and multipolarization radar features for crop classification". *IEEE Transactions on geoscience and remote sensing*, Vol. 33, No. 5, 1995, pp.1170-1181. [7]
- [6] V. Atlamazoglou, D. Yova, "Texture analysis of fluorescence microscopic images of colonic tissue sections", *Medical and Biol., Engineering and Computing*, vol. 39, no.2, 2001, pp.145-151
- [7] R. Badea "Ficatul". In: "Tratat de ultrasonografie clinica". Badea R., Dudea S.M., Mircea P.A., Stamatian F, Ed Medicală, București, 2000, pp. 105-175
- [8] G. Baudat, F. Anouar, "Generalized discriminant analysis using a kernel approach", *Pattern Recognition Letters*, Vol. 28 , No. 2, 2007, pp. 254-259
- [9] P.A. Bautista, M.A. Lambino "Co-occurrence matrices for wood texture classification", 2001. Online: <http://physics.msuiit.edu.ph/spvm/papers/2001/bautista-lambino.pdf>
- [10] S. Bejakovic, R. Kumar. "Analysis of Crohn's Disease Lesions in Capsule Endoscopy Images", *International Conference on Robotics and Automation, ICRA*, pp.2793-2798, 2009
- [11] Brodatz texture collection, 2011, Online: <http://sipi.usc.edu/database/database.php?volume=textures&image=1>

- [12] D. Cavouras, I. Kandarakis, I. Theotokas, E. Kanellopoulos, D. Triantis, "Computer Image Analysis of Ultrasound Images for Discriminating and Grading Liver Parenchyma Disease Employing a Hierarchical Decision Tree Scheme and the Multilayer Perceptron Neural Network Classifier", *Medical Informatics Europe*, IOS Press, pp. 522-528, 1997.
- [13] D.A. Clausi, "An analysis of cooccurrence texture statistics as a function of grey level quantization", *Canadian Journal of Remote Sensing*, Vol.28, No.1,2002, pp. 45-62
- [14] L. S. Davis, M. Clearman, J.K. Aggarwal: "A Comparative Texture Classification Study Based on Generalized Cooccurrence Matrix", *Proceedings of the 18-th IEEE Conference on Decision and Control*, pp. 71-78, 1979
- [15] L. S. Davis, "Image Texture Analysis Techniques – A Survey", *Digital Image Processing*, Simon and R. M. Haralick (eds.), pp. 189-201, 1981.
- [16] R. Duda, P. Hart , D. Stork, *Pattern Classification*, Wiley Interscience, 2003
- [17] M. Hall, "Benchmarking attribute selection techniques for discrete class data mining", *IEEE Transactions On Knowledge and Data Engineering*, Vol 15, No.3, 2003, pp. 1-16
- [18] J.S. Hallinan, "Feature selection and classification in the diagnosis of cervical cancer", 1999. Online: <http://www.staff.ncl.ac.uk/j.s.hallinan/pubs/Handbook.pdf>
- [19] R. Haralick, "Statistical and structural approaches to texture", *Proceedings of the IEEE*, Vol. 67, No. 5, pp. 786-804, 1979
- [20] M. B. Huber, J. Carballido-Gamio . "Development and testing of texture discriminators for the analysis of trabecular bone in proximal femur radiographs", *Medical Physics*, vol. 36, no. 11, 2009, pp. 5089 – 5098
- [21] A.K. Jain, "*Fundamentals of Digital Image Processing*", Englewood Cliffs, NJ: Prentice Hall, 1989
- [22] A.K. Jain, R. Duin, J. Mao, "Statistical Pattern Recognition: A Review", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No.1, pp. 4 -37, 2000
- [23] C. Jalobeanu, I. Rasa, "*Incertitudine si decizie. Statistica si probabilitati aplicate in management*", Editura UT Press, Cluj-Napoca, 2001
- [24] Y. Kadah, A. Farag, J. Zurada, A. Badawi, A. Youssef, "Classification Algorithms for Quantitative Tissue Characterization of Diffuse liver Disease from Ultrasound Images", *IEEE Transactions on Medical Imaging*, Vol. 15, No. 4, 1996, pp. 466-478.
- [25] M. Kazuo, M. Utsu and P. E. Kihale, "Quantification of sonographic echogenicity with grey-level histogram width: a clinical tissue characterization", *Ultrasound in Medicine & Biology*, Vol. 24, No. 2, 1998, pp. 225-234.
- [26] W. Ku, X. Huang, Y. Jia, „Segmentation in noisy medical images using PCA model based particle filtering”, *In Proc.of SPIE (The International Society for Optical Engineering), Medical Imaging: Image Processing*, Vol. 6914, 2008, pp. 278-283
- [27] W. Kuo, R. Chang, "Data mining with decision trees for diagnosis of breast tumor in medical ultrasonic images", *Springer Netherlands*, Vol. 66, No.1, 2001, pp. 51-57
- [28] A. Madabhushi., M.D. Felman, D.N. Metaxas, J.Tomaszewski, D. Chte, „Automated Detection of Prostatic Adenocarcinoma From High-Resolution Ex Vivo MRI", *IEEE Trans. on Medical Imaging*, Vol. 24, No. 12, December 2005, pp. 1611-1626
- [29] Y. Mao, X. Zhou, D. Pi, Y. Sun, S. Wong, "Multiclass Cancer Classification by Using Fuzzy Support Vector Machine and Binary Decision Tree with Gene Selection", *Journal of Biomedicine and Biotechnology* , Vol. 2, 2005, pp.160–171
- [30] K. Masood, "Co-occurrence and morphological analysis for colon tissue biopsy classification" in *Proceedings of the 4th International Workshop on Information Technology* , Pakistan, pp. 211-216, 2006
- [31] C. Meurie, O. Lezoray, A. Elmoataz, "Multiple pixel classifier combination for bronchial tumors image segmentation", *International Conference on Complex Systems Intelligence and Modern technological Applications (CSIMTA)*, Invited Session, Vol. 1, France, pp. 305-310, 2004.

- [32] J. Minuillon, A. Rosemary Tate, C. Arus, J. Griffiths, "Classifier combination for in vivo magnetic resonance spectra of brain tumors", *Lecture Notes in Computer Science (LNCS 2364)*, pp. 282-292, 2002.
- [33] D. Mitrea, S. Nedevschi, M. Lupșor, R. Badea, "Exploring the Textural Parameters obtained from Ultrasound Images for Modeling the Liver Pathological Stages in the evolution towards Hepatocellular Carcinoma", in Proceedings of IEEE International Conference on Automation, Quality and Testing, Robotics AQTR 2008 - THETA 17th edition, 2008, pp. 478 – 484 (ISI Proceedings)
- [34] D. Mitrea, S. Nedevschi, P. Mitrea, M. Lupșor, M. Socaciu, T. Pop, O. Moșteanu, R. Badea, „The Imagistic Textural Model of the Hepatocellular Carcinoma. Comparisons, discussions, conclusions”, *Automation Computers Applied Mathematics (ACAM)*; vol. 17, nr. 1, 2008, pp. 28-35, (CNCSIS B+)
- [35] D. Mitrea, S. Nedevschi, B. Petrut, I. Coman, „The Textural Imagistic Model of the Prostatic Adenocarcinoma (ADKP)”, *Proceedings of 4-th IEEE International Conference on Intelligent Computer Communication and Processing (ICCP 2008)*, Cluj-Napoca, pp. 99-107, 2008 (IEEE, ISI Proceedings)
- [36] D. Mitrea, Nedevschi S., M. Lupșor, H. Ștefănescu, R. Badea, "Comparing Various Feature Selection Methods for Building the Textural Model of Hepatocellular Carcinoma based on Ultrasound Images", in Proceedings of The Biennial International Conference on Analysis of Biomedical Signals and Images (Biosignal) Brno, Cehia, pp. 11-18, 2008.
- [37] D. Mitrea, S. Nedevschi, M. Lupsor, M. Socaciu, R. Badea, "Classification of the Hepatocellular Carcinoma in Ultrasound Images based on the Imagistic Textural Model of This Tumor", *International Conference on Advancements of Medicine and Health Care Through Technology (Meditech) IFMBE Proceedings Springer*, vol.26, pp. 267-272, 2009
- [38] D. Mitrea, S. Nedevschi, M. Lupsor, M. Socaciu, R. Badea, "Improving the textural model of the hepatocellular carcinoma using dimensionality reduction methods", *Proceedings of 2nd International Congress on Image and Signal Processing (CISP 2009)*, Tianjin, China, pp. 1083-1087, 2009
- [39] D. Mitrea, S. Nedevschi, M. Lupsor, M. Socaciu, R. Badea, "Improving the imagistic textural model of the hepatocellular carcinoma through dimensionality reduction techniques and classifier combinations", *ACAM (Automation, Computers, Applied Mathematics) Scientific Journal*, Vol. 18, pp. 218-223, 2009
- [40] D. Mitrea, S. Nedevschi, M. Lupsor, M. Socaciu, R. Badea, "Experimenting various classification techniques for improving the automatic diagnosis of the malignant liver tumors, based on ultrasound images", *The 3rd International Conference on Image and Signal Processing (CISP)*, Yantai, China, Proceedings, Vol.4, pp. 1853-1859, 2010 [73]
- [41] D. Mitrea, S. Nedevschi, M. Lupsor, M. Socaciu, R. Badea, "Advanced classification methods for improving the automatic diagnosis of the hepatocellular carcinoma, based on ultrasound images", *Proceedings of 2010 IEEE International Conference on Automation, Quality and Testing, Robotics (AQTR)*, Cluj-Napoca, pp. 265-271, 2010
- [42] D. Mitrea, S. Nedevschi, M. Lupsor, M. Socaciu, R. Badea, „Improving the textural model of the hepatocellular carcinoma through multiclass division using clustering methods”, *Automation, Computers, Applied Mathematics (ACAM)*, Nr. 3, Vol. 19, Pp. 409-415, 2010 (CNCSIS B+)
- [43] D. Mitrea, P. Mitrea, R. Badea, M. Socaciu, L. Ciobanu, A. Golea, C. Hagi, A. Seicean, „Relevant feature selection and automatic recognition of the inflammatory bowel diseases and colon cancer from B-mode and contrast-enhanced ultrasound images”, *Automation, Computers, Applied Mathematics (ACAM)*, Nr. 3, Vol. 19, Pp. 403-409, 2010 (CNCSIS B+)
- [44] D. Mitrea, S. Nedevschi, R. Badea, "The role of the superior order GLCM and of the generalized cooccurrence matrices in the caharacterization and automatic diagnosis of the hepatocellular carcinoma, based on ultrasound images", *Proceedings of the 2011 IEEE 7th International*

- Conference on Intelligent Computer Communication and Processing, (ICCP), Cluj-Napoca, pp. 197-204, 2011 (IEEE)
- [45] D. Mitrea, S. Nedevschi, M. Socaciu, R. Badea, “ The role of the superior order GLCM in improving the characterization and recognition of the hepatocellular carcinoma, based on ultrasound images”, Proceedings of the 2011 3rd International Conference on Telecommunications and Signal Processing – (TSP), Budapest, pp. 602-607, 2011 (IEEE, ISI Proceedings) – selectat pentru publicare in RadioEngineering Journal (Science Citation Index - SCI, ISI)
 - [46] D. Mitrea, P. Mitrea, R. Badea, M. Socaciu, L. Ciobanu, A. Golea, C. Hagiu, and A. Seicean, “Texture-Based Methods and Dimensionality Reduction Techniques Involved in the Detection of the Inflammatory Bowel Diseases from Ultrasound Images”, Proceedings of Meditech 2011 Conference, Cluj-Napoca, IFMBE Proceedings Series, No. 36, pp. 220 – 225, 2011 (Springer)
 - [47] D. Mitrea, S. Nedevschi, M. Lupsor, M. Socaciu, R. Badea, “The role of the feature extraction methods in improving the textural model of the hepatocellular carcinoma, based on ultrasound images”, acceptat pentru publicare la International Conference on Digital Information Processing and Communications (ICDIPC 2011), Ostrava, Cehia, publicat in seria Communications in Computers and Information Sciences (CCIS), Lecture Notes in Computer Science, Vol. 188 , pp.496-509, 2011 (Springer)
 - [48] Delia Mitrea, Paulina Mitrea, Sergiu Nedevschi, Radu Badea, Monica Lupşor, Mihai Socaciu, Adela Golea, Claudia Hagiu, Lidia Ciobanu, “Abdominal tumor characterization and recognition using superior order cooccurrence matrices, based on ultrasound images”, acceptat pentru publicare în Computational and Mathematical Methods in Medicine (Hindawi Journal, SCI Index Expanded), 2011 (ISI)
 - [49] A. Mojsilovic, M. Popovic, S. Markovic and M. Krstic, ”Characterization of Visually Similar Diffuse Diseases from B-Scan Liver Images Using Nonseparable Wavelet Transform”, *IEEE Transactions on Medical Imaging*, Vol. 17, No. 4, August 1998, pp.541-549
 - [50] M. Mokji, S. A. Bakar, ”Gray Level Co-occurrence matrix computation based on Haar Wavelet”, IEEE Conference on Computer Graphics, Imaging and Visualization, pp. 124-180, 2007
 - [51] M. Olteanu, M. Tanase, „An algorithm for the analysis of fractal-like structures and miscellaneous applications”, Proc. of IEEE-TTTC Int. Conf. On Automation, Quality and Testing, Robotics, tome 2, pp. 200-206, 2004
 - [52] J. R. Parker, “Algorithms for Image Processing and Computer Vision”, *Wiley Computer Publishing*, USA, 1996
 - [53] M. Pechenizkiy, A. Tsymbal, S. Puuronen, “PCA based feature transformation for classification: Issues in medical diagnosis”, *Proceedings of the 17th IEEE Symposium on Computer-Based Medical Systems*, pp. 535 – 540, 2004
 - [54] Z. Qiao, L. Zhou, J. Huang, “Sparse linear discriminant analysis with applications to high dimensional low sample size data”, *IAENG International Journal of Applied Mathematics*, Vol. 39, Nr. 1, 2009, pp. 1-13
 - [55] S. Ramaswamy, P. Tamayo, R. Rifkin, S. Mukherjee, C. Yeang, M. Angelo, C. Ladd, M. Reich, E. Latulippe, “Multiclass cancer diagnosis using tumor gene expression signatures”, Proceedings of the National Academy of Sciences of United States of America (PNAS), pp. 15149-15154, 2001
 - [56] M. Ranjith, R.M. Balaji, M. Surjith Kumar, J. Dhyaneswaran, A. Baskar, “Content based image retrieval for medical image (cerebrum infarct) using PCA”, National Conference on Recent Trends in Communication and Signal Processing (RTCSP'09), Conference Proceedings, pp.124-126. [98]
 - [57] U. Raeth, D. Schlaps , Diagnostic accuracy of Computerized B-Scan Texture Analysis and Conventional Ultrasonography in Diffuse Parenchymal and Malignant Liver Disease, *Journal of Clinical Ultrasound*, No. 13, pp. 87-89

- [58] S. Roweis, "EM algorithms for PCA and SPCA", <http://www.cs.nyu.edu/~roweis/papers/empca.pdf>
- [59] L. Ruess "Inflammatory bowel disease in children and young adults: correlation of sonographic and clinical parameters during treatment", *AJR Am J Roentgenol.*, Vol.175, No.1, pp.79-84, 2000
- [60] Schölkopf Bernhard, [Alex J. Smola](#), [Klaus-Robert Müller](#) "Kernel Principal Component Analysis", *ICANN* (7th International Conference, Lausanne, Switzerland, October 8-10, 1997, Proceedings), 1997, pp. 583-588
- [61] B. Schuller, B. Schmitt, „Feature selection and stacking for robust discrimination of speech, monophonic singing, and polyphonic music”, *Proceedings of the IEEE International Conference on Multimedia and Expo*, pp. 840-843, 2005
- [62] L. I. Smith, "A tutorial on Principal Component Analysis", February 2002, http://www.cs.otago.ac.nz/cosc453/student_tutorials/principalcomponents.pdf
- [63] E. Stollnitz, T. DeRose, "Wavelets for computer graphics", *IEEE Computer Graphics & Applications*, Vol. 15, No.3, 1995, pp. 76–84
- [64] H. Sujana, S. Swarnamani, S. Suresh, „Application of Artificial Neural Networks for the classification of liver lesions by texture parameters” , *Ultrasound in Med. & Biol.*, Vol. 22, No. 9, 1996, pp. 1177- 1181.
- [65] D. Terzopoulos, S. Zucker, „Detection of osteogenesis imperfecta by automated texture analysis”, *Computer Graphics and Image Processing*, No. 20, 1982, pp.229-243
- [66] Weka, Data Mining Software in Java, Weka 3, <http://www.cs.waikato.ac.nz/ml/weka/>, 2004
- [67] M. Varma, R. Garg, "Locally Invariant fractal features for statistical texture classification", *Proceedings of the International Conference on Computer Vision*, pp. 116-123, 2007
- [68] H. Yoshida, D.Casalino, B. Keserci, A. Coskun, O. Ozturk, A. Savranlar, „Wavelet-packet-based texture analysis for diferentiation between benign and malignant liver tumours in ultrasound images”, *Physics in Med. and Biol.*, No.48, 2003, pp.3735-3753.
- [69] H. Zeng, Y. Cheung, "A new feature selection method for Gauusian mixture clustering”, *Pattern Recognition*, No. 42, 2009, pp.243-250
- [70] Wen-Ming Zheng, Li Zhao, Cai-Rong Zou, "Recognition using extended multiple discriminant analysis (EMDA) method”, *Proceedings of the first International Conference on Machine Learning and Cybernetics*, Beijing, pp. 121-125, 2002