UNIVERSITATEA ”ŞTEFAN CEL MARE” DIN SUCEAVA
FACULTATEA DE INGINERIE ELECTRICĂ ŞI ŞTIINȚA CALCULATOARELOR
Centrul integrat de cercetare, dezvoltare şi inovare pentru Materiale Avansate, Nanotehnologii şi Sisteme Distribuite de fabricație şi control (MANSiD)

Contribuții în recunoaşterea formelor utilizând descompuneri tensoriale

Domeniul Calculatoare şi Tehnologia Informației

Conducător ştiințific:
Prof. univ. dr. ing. Ştefan-Gheorghe PENTIUC

Doctorand:
Laura-Bianca BILIUS

Suceava, 2022

Această lucrare de doctorat a fost realizată în cadrul şi cu suportul Laboratorului de Maşini Inteligente şi Vizualizarea Informației (MintViz) din cadrul Centrului integrat de cercetare, dezvoltare şi inovare pentru Materiale Avansate, Nanotehnologii şi Sisteme Distribuite de fabricație şi control (MANSiD, 671/09.04.2015) al Universității "Ştefan cel Mare" din Suceava, cu suportul proiectelor:

"*Excelență academică şi valori antreprenoriale - sistem de burse pentru asigurarea oportunităților de formare şi dezvoltare a competențelor antreprenoriale ale doctoranzilor şi postdoctoranzilor (ANTREPRENORDOC)*", cofinanțat din Fondul Social European prin Programul Operational Capital Uman, 2014-2020, Contract nr. 36355/23.05.2019, POCU/380/6/13 - Cod SMIS: 123847,

"WearSkill: Interacțiuni fluide cu dispozitive wearable smart pentru abilități motorii specifice", cod proiect PN-III-P2-2.1-PED-2019-0352, contract nr. 276PED/ 2020,

"Centru pentru transferul de cunoştințe către întreprinderi din domeniul ICT – CENTRIC", codul proiectului POC-A1-A1.2.3-G-2015, Cod SMIS 2014+ 119722 (ID _40_305), Contract de finanțare nr. 5/AXA 1/1.2.3/G/13.06.2018, Contract subsidiar nr. 15.875/04.09.2020/acronim Study AI/ IMM:MiCS Software SRL,

"*Platformă hibridă de comunicații prin lumină vizibilă şi realitate augmentată pentru dezvoltarea de sisteme inteligente de asistență şi siguranță activă a autovehiculelor*", cod proiect PN-III-P1-1.2-PCCDI-2017-0917, contract nr. 21PCCDI/2018.

Cuprins

1 Introducere **7**

2 Descompuneri tensoriale **9**
2.1 Tensori 9
2.2 Descompunerea canonică poliadică 9
2.3 Descompunerea Tucker 10

3 Metode şi tehnici de recunoaştere a formelor **11**

4 Procesarea imaginilor hiperspectrale **12**
4.1 Imagini hiperspectrale 12
4.2 Metode şi tehnici de recunoaştere a imaginilor hiperspectrale 12
4.3 Modele Spectral Unmixing 13
4.4 Legătura dintre modelul Linear Spectral Unmixing şi descompunerile tensoriale 13
4.5 Concluzii 13

5 Analiza imaginilor hiperspectrale utilizând tehnici de învățare automată **14**
5.1 Tehnici de învățare supravegheată 14
5.2 Tehnici de învățare nesupravegheată 14
5.3 Segmentarea matricilor de abundență 15
5.4 Concluzii 16

6 Metodă eficientă de clasificare a imaginilor hiperspectrale folosind diagrame Voronoi şi forme tari **17**
6.1 Descrierea metodei 17
6.2 Diagrame Voronoi 18
6.3 Descompunerea Parafac şi construirea matricii de abundență 19
6.4 Identificarea nodurilor Voronoi 19
6.4.1 Metoda M_1 de identificare a nodurilor Voronoi 19
6.4.2 Metoda M_2 de identificare a nodurilor Voronoi 20
6.4.3 Metoda M_3 de identificare a nodurilor Voronoi 21
6.5 Forme tari şi partiția consens 21
6.6 Evaluarea metodei propuse 22
6.7 Avantajele metodei propuse 24

7 Metodă de recunoaştere a gesturilor **25**
7.1 Introducere în analiza gesturilor executate cu mâna în aer 25
7.2 Studii explorative cu privire la interacțiuni gestuale 25
7.3 Analiza sistematică a literaturii de specialitate 26
7.4 Metodă de recunoaştere a gesturilor executate cu mâna în aer folosind descompunerea Tucker2 şi metoda ansamblu de învățare 26
7.5 Colectarea şi procesarea datelor corespunzătoare gesturilor executate cu mâna în aer 26
7.6 Construirea tensorului de gesturi şi aplicarea descompunerii tensoriale Tucker2 27

7.7 Metoda ansablu de învățare a gesturilor executate cu mâna în aer 28
7.8 Evaluarea performanței metodei DELMet 28
7.9 Avantajele metodei propuse 30

8 Concluzii şi contribuții **31**
8.1 Contribuțiile lucrării de doctorat 31
8.1.1 Contribuții teoretice 31
8.1.2 Contribuții practice 32
8.1.3 Diseminarea rezultatelor 33
8.1.4 Implicarea în proiecte de cercetare 35
8.2 Dezvoltări viitoare 35

References **37**

1 Introducere

Odată cu dezvoltarea continuă a tehnologiei, un volum semnificativ de informații sunt generate din diverse surse, prin urmare, impactul dimensionalității devine o provocare în interpretarea datelor. Datele de dimensiuni mari (*en.*: big data) sunt date digitale eterogene generate de companii şi persoane ale căror caracteristici (volumul mare, diversitatea formelor, viteza de procesare) necesită instrumente de stocare şi analiză informatică specializată [46]. Tehnicile de extragere a informațiilor şi a caracteristicilor trebuie identificate pentru a obține rezultate favorabile într-un timp util. Conform literaturii ştiințifice de specialitate [59], descompunerile tensoriale, cum ar fi Parafac şi Tucker, sunt utilizate pentru extragerea caracteristicilor şi rezolvarea problemelor de clasificare oferind astfel posibilitatea de a studia datele evitând costul de calcul ridicat. Imaginile hiperspectrale (*en.* hyperspectral images (HSI)) devin un instrument valoros cu un domeniu larg de utilizări, cum ar fi cartografierea, monitorizarea mediului, detectarea dimensiunii şi formelor zonelor urbanizate. O imagine hiperspectrală este o imagine cubică tridimensională compusă din sute de imagini spațiale. Deoarece imaginile hiperspectrale sunt de dimensiuni mari, acestea pot afecta analiza datelor datorită costului computațional mare [62]. Metodele şi tehnicile existente de recunoaştere a formelor, dar şi a descompunerilor tensoriale au numeroase aplicații atât în clasificarea imaginilor hiperspectrale, cât şi în clasificarea gesturilor. În ceea ce priveşte diversitatea metodelor de intrare, gesturile executate cu mâna în aer (*en.*: mid-air hand gestures) sunt tehnici de interacțiune în plină dezvoltare. Datele generate şi transmise de dispozitivele care încorporează senzori pentru detecția acestor gesturi tind să crească în mărime şi dimensionalitate oferind astfel o mulțime de informații valoroase. Această creştere este adesea considerată o provocare în studiul datelor de dimensiuni mari. Drept urmare, extragerea caracteristicilor şi procesarea lor folosind algoritmi şi tehnici care îşi propun să înțeleagă gesturile mâinii sunt folosite în sistemele informatice cu un timp de răspuns interactiv.

În acest context, scopul acestei lucrări este circumscris reducerii dimensiunilor mari ale datelor şi extragerii caracteristicilor folosind descompuneri multiliniare algebrice tensoriale în vederea obținerii unei precizii şi acurateți ridicate de clasificare pentru probleme de învățare supravegheată şi nesupravegheată.

Obiectivele acestei lucrări sunt:

- Realizarea unui studiu explorativ despre clasificarea imaginilor hiperspectrale utilizând algoritmi de învățare supravegheată şi nesupravegheată.
- Propunerea de metode de învățare nesupravegheată a imaginilor hiperspectrale utilizând descompunerile tensoriale şi segmentarea imaginilor.
- Propunerea unei metode de învățare nesupravegheată pentru date de tipul imaginilor hiperspectrale utilizând descompunerile tensoriale, diagramele Voronoi şi formele tari care reduc semnificativ dimensiunea datelor.
- Prezentarea unor studii explorative despre interacțiuni gestuale cu inele smart şi interacțiuni în interiorul şi exteriorul vehiculelor.
- Propunerea unei metode de învățare supravegheată a gesturilor executate cu mâna în aer utilizând descompunerile tensoriale care urmăreşte îmbunătățirea acurateții clasificării şi reducerii timpului de calcul.

Teza este structurată în 8 capitole şi bibliografie.

Capitolul 1 începe cu o introducere generală, fiind descrise obiectivele, motivația cercetării, inclusiv modul în care a fost structurată şi organizată teza de doctorat. Capitolele 2 şi 3

prezintă noțiunile teoretice cu privire la descompunerile tensoriale şi algoritmii de învățare automată utilizați în dezvoltarea metodelor specifice recunoaşterii formelor. Capitolul 4 introduce noțiuni teoretice despre imaginile hiperspectrale, metodele şi tehnicile de analiză şi procesare utilizate în literatura ştiințifică de specialitate, inclusiv legătura dintre descompunerile tensoriale şi modelele spectral unmixing) [43, 60]. Explorarea diverşilor algoritmi de învățare supravegheată [7] şi nesupravgheată [6] pentru imagini hiperspectrale preluate de sateliți specializați, dar şi rezultatele obținute în urma extragerii caracteristicilor folosind descompuneri tensoriale sunt prezentate şi discutate în Capitolul 5. De asemenea, două metode de segmentare a matricilor de abundență (*en.*: abundance matrix [43]) [5, 31] care au la bază descompunerea tensorilor imaginilor hiperspectrale sunt propuse în acelaşi capitol.

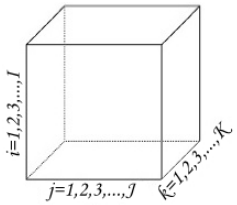
Capitolul 6 introduce o metodă de recunoaştere a formelor din imaginile hiperspectrale care urmăreşte eficientizarea procesului de învățare nesupravegheată [8]. Metoda propusă vizează reducerea semnificativă a cerințelor privind resursele de calcul prin utilizarea descompunerilor tensoriale şi diagramelor Voronoi, dar şi a obținerii unei precizii mai bune prin identificarea formelor tari. Folosind o metodă euristică de clasificare a formelor tari, o partiție optimă este identificată, iar diagrama Voronoi corespunzătoare poate fi construită astfel extinzând clasificarea la întreaga imagine hiperspectrală. În ceea ce priveşte evaluarea performanței metodei propuse, sunt discutate rezultatele obținute pentru seturi de date sintetice şi imagini hiperspectrale preluate de sateliți. Printre altele, este descrisă o metodă de rezolvare a problemelor de restaurare a imaginilor digitale folosind diagrame Voronoi [13].

Capitolul 7 prezintă stadiul actual cu privire la recunoaşterea gesturilor executate cu mâna în aer folosind date colectate de la senzori IMU (*en.*: Inertial Measurement Units). În această direcție, sunt prezentate şi discutate rezultatele unui analize sistematice a literaturii ştiințifice din domeniu care are ca scop identificarea metodelor de procesare a datelor colectate de la senzori inerțiali şi a algoritmilor utilizați în literatură pentru recunoaşterea gesturilor executate cu mâna în aer. Totodată, în Capitolul 7 este introdus intrumentul de tip aplicație web, GestuRING [10, 57] care a fost susținut de o analiza sistematică a literaturii ştiințifice de specialitate cu privire la inele smart realizată cu scopul de a explora diversitatea interacțiunilor gestuale. În strânsă legătură cu gesturile utilizate ca metode de intrare, în acest capitol este prezentat în detaliu un studiu despre interacțiunile dintre vehicule şi utilizatori. Acest studiu constă într-un experiment de solicitare (*en.*: end-user elicitation study) a preferințelor şoferilor şi pasagerilor în ceea ce priveşte gesturile utilizate pentru controlul funcțiilor vehiculelor [30]. Explorarea spațiului din exteriorul vehiculelor a condus la introducerea cadrului conceptual intitulat "smart vehicle proxemics" [11, 12] în cadrul căruia are loc o interacțiune fluentă dintre sistemul de infotainment şi dispozitivele mobile şi purtabile ale şoferilor în timp ce aceştia traversează cinci zone distincte identificabile, din interior la zona personală, zona deschisă distantă, zona deschisă proximală şi zona ascunsă din afara vehiculului. Cadrul conceptual introdus confirmă oportunitățile extinderii interacțiunilor şi în afara vehiculelor. Prin urmare, metoda DELMet propusă în acest capitol se complementează reciproc cu studiile explorative realizate despre inele smart şi interacțiuni în interiorul cât şi în exteriorul vehiculului. Metoda DELMet este dedicată clasificării gesturilor executate cu mâna în aer folosind descompunerile tensoriale pentru extragerea caracteristicilor şi reducerea dimensionalității şi o metodă ansamblu de învățare propusă pentru a creşte acuratețea clasificării.

Capitolul 8 prezintă contribuțiile aduse, concluziile finale, diseminarea rezultatelor şi direcțiile viitoare de cercetare. Teza se încheie cu o listă de referințe bibliografice.

2 Descompuneri tensoriale

2.1 Tensori



Tensorii sunt forme multiliniare care au fost introduse pentru a extinde noțiunile de scalari, vectori și matrici. Tensorii descriu relații dintre vectori, fiind astfel independenți sistemului de coordonate. Ordinul unui tensor $X \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ este N, iar fiecare componentă este notată astfel: $x_{i_1, i_2 \cdots, i_n}$.

2.2 Descompunerea canonică poliadică

Descompunerile tensoriale sunt metode utile de extragere a caracteristicilor care pot fi folosite ulterior în recunoașterea formelor datorită posibilității de a captura structuri multiliniare din seturi de date de dimensiuni mari și de ordin superior [41]. În literatura științifică de specialitate descompunerea canonică poliadică (*en.*: Canonical Polyadic Decomposition (CPD sau CANDECOMP)) este cunoscută și sub numele de analiza factorilor paraleli (*en.*: Paralell Factor Analysis (Parafac)) sau CANDECOMP/PARAFAC [64].

Parafac descompune un tensor ca o sumă de R tensori de rang unu, unde R reprezintă numărul de componente al descompunerii [21]. Descompunerea Parafac pentru un tensor $X \in \mathbb{R}^{I \times J \times K}$ este definită astfel (vezi Ecuația 2.1 și Figura 2.1):

$$X \approx \sum_{r=1}^{R} a_r \circ b_r \circ c_r + E \equiv [\![A, B, C]\!] \tag{2.1}$$

unde $R \in \mathbb{N}$ reprezintă numărul de componente, $a_r \in \mathbb{R}^I$, $b_r \in \mathbb{R}^J$, $c_r \in \mathbb{R}^K$ sunt componentele de rang unu și $\circ$ este produsul exterior. Descompunerea Parafac returnează trei matrici factor care conțin componentele de rang unu, $A = [a_1, a_2, \cdots, a_R] \in \mathbb{R}^{I \times R}$, $B = [b_1, b_2, \cdots, b_R] \in \mathbb{R}^{J \times R}$ și $C = [c_1, c_2, \cdots, c_R] \in \mathbb{R}^{K \times R}$ corespunzătoare celor trei dimensiuni ale tensorului $X \in \mathbb{R}^{I \times J \times K}$ [27].

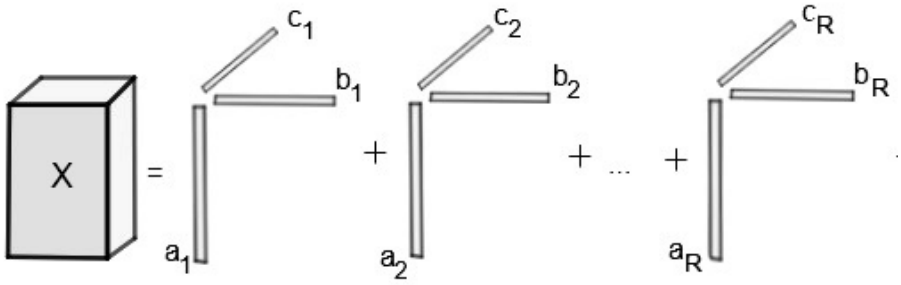


Figura 2.1: Reprezentarea grafică a descompunerii Parafac în R componente corespunzătoare tensorului X de ordin $N = 3$ (conform [27])

.

Pentru a obține o desompunere ideală se alege cel mai mic număr de tensori de rang unu care în sumă generează tensorul X, această problemă NP dificilă implică alegerea unui număr de componente care să reprezinte cât mai bine tensorul X. În acest caz se recomandă experimentarea prin compararea unor numere diferite de componente cu scopul de a alege cea mai mică valoare R care oferă o potrivire cât mai bună [27].

2.3 Descompunerea Tucker

Metoda Tucker [27] descompune un tensor într-un tensor nucleu și mai multe matrici factor care corespund fiecărei dimensiuni a tensorului original. Descompunerea Tucker este o formă de analiză a componentelor principale (*en.*: Principal component analysis (PCA)) de ordin superior. Pentru un tensor de ordin $N = 3$ ($X \in \mathbb{R}^{I \times J \times K}$), descompunerea Tucker este definită astfel:

$$X \approx G \times_1 A \times_2 B \times_3 C = \sum_{p=1}^{P} \sum_{q=1}^{Q} \sum_{r=1}^{R} g_{pqr} a_p \circ b_q \circ c_r = [\![G; A, B, C]\!] \tag{2.2}$$

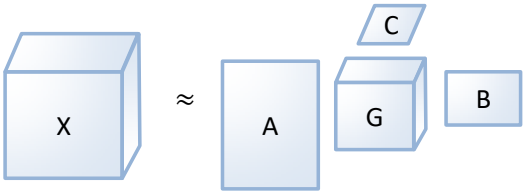


Figura 2.2: Descompunerea Tucker pentru un tensor $X \in \mathbb{R}^{I \times J \times K}$ (conform [27]).

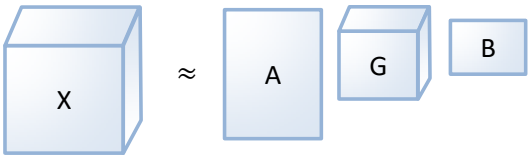


Figura 2.3: Descompunerea Tucker2 pentru un tensor $X \in \mathbb{R}^{I \times J \times K}$ (conform [27]).

$G \in \mathbb{R}^{P \times Q \times R}$ este tensorul central iar valorile sale descriu nivelul de interacțiune dintre componente, $A \in \mathbb{R}^{I \times P}$, $B \in \mathbb{R}^{J \times Q}$ și $C \in \mathbb{R}^{K \times R}$ sunt matricile factor care corespund celor trei dimensiuni ale tensorului X și P, Q și R reprezintă numărul de componente a matricilor factor corespunzătoare celor trei dimensiuni [27] (vezi Figura 2.2). Tensorul nucleu poate fi văzut ca o versiune comprimată a tensorului original dacă $I > P$, $J > Q$ și $K > R$ [27].

Descompunerea Tucker2 [27] a unui tensor de ordin $N = 3$ stabilește una dintre matricile factor ca fiind matricea identitate ($\mathbf{I}$), prin urmare descompunerea este definită astfel (vezi Figura 2.3):

$$X \approx G \times_1 A \times_2 B = [\![G; A, B, \mathbf{I}]\!],\ C = \mathbf{I} \text{ și } K = R. \tag{2.3}$$

Metoda Tucker descompune un tensor în componente fiind de asemenea o metodă de reducere a dimensionalității și compresie [40], un aspect important pentru a reduce costul de calcul. De asemenea, descompunerea Tucker deține avantajul extragerii de caracteristici utile care stau la baza unei acurateți crescute de clasificare [41].

3 Metode şi tehnici de recunoaştere a formelor

Recunoaşterea formelor este o problemă de clasificare unde obiectele sunt descrise cu ajutorul caracteristicilor şi sunt reprezentate ca şi puncte în spaţiul formelor [29]. Rezolvarea problemelor de recunoaştere a formelor presupune utilizarea de diverşi algoritmi, cum ar fi, învăţarea supravegheată, învăţarea nesupravegheată, învăţarea semi-supravegheată, învăţarea profundă (*en.* deep learning) sau învăţarea prin întărire (*en.* reinforcement learning) [2]. Algoritmii de învăţare supravegheată sunt inductivi şi pornesc de la un set de exemple ale unei probleme urmărind identificarea unei funcţii, denumită predictor care să permită clasificarea unor instanţe noi [2]. Astfel, clasificarea datelor folosind algoritmi de învăţare supravegheată presupune gruparea acestora având la bază modele prestabilite cunoscute sub numele de "clase de antrenament". Acestea sunt grupări de referinţă ce corespund realităţii fiind folosite înaintea clasificării de noi instanţe pentru antrenare în aplicarea algoritmilor în practică [51]. Învăţarea nesupravegheată presupune clusterizarea variabilelor în diferite grupuri fără definirea unor predictori [2]. Aceşti algoritmi nu au nevoie de clase de antrenament fiind lăsaţi singuri să identifice gruparea şi ordinea existente intrinsec în cadrul datelor în funcţie de anumite criterii. Comparativ cu învăţarea supravegheată, învăţarea nesupravegheată poate identifica în datele de intrare noi concepte [51]. În ceea ce priveşte învăţarea semi-supravegheată, este un tip de învaţare supravegheat, dar şi nesupravegheat. Algoritmii semi-supravegheaţi se antrenează cu un număr mic de date şi tratatează datele neetichetate ca fiind date de testare [45].

Performanţa algoritmilor este importantă pentru a putea identifica probabilitatea ca aceştia să prezică etichetele corect. Atunci când există un set de date cu clasificare cunoscută se pot calcula diverse mărimi pentru evaluarea performanţei clasificării[1], cum ar fi, acurateţea, precizia, reamintirea, scorul F, matricea de confuzie, sau pierderile validării încrucişate. Atunci când nu există un set de date cu clasificare cunoscută se pot calcula diverse mărimi pentru evaluarea performanţei clusterizării/grupării[1], cum ar fi, scorul Silhouette sau indexul Davies-Bouldin.

În această teză de doctorat sunt aplicaţi algoritmi de învăţare supravegheată, cum ar fi, cei mai apropiaţi k vecini (K-Nearest Neighbor (KNN)), Bayes Gaussian şi Bayes Multinomial Naiv, analiza discriminantă liniară (Linear Discriminant Analysis (LDA)), arbori decizionali (Decision Tree (DT)), maşini cu vector suport (Support Vector Machine (SVM)), propagarea etichetelor (Label Propagation (LP)), algoritmi de învăţare nesupravegheată, cum ar fi, K-means, gruparea ierarhică, deplasarea mediei (Mean-shift), algoritmul punct central (K-medoids), dar şi analiza componentelor principale (*en.*: Principal Component Analysis (PCA)).

[1] `https://scikit-learn.org/stable/modules/clustering.html`

4 Procesarea imaginilor hiperspectrale

4.1 Imagini hiperspectrale

Imaginile hiperspectrale sunt capturate de senzori imagistici hiperspectrali şi sunt date tridimensionale care conțin sute de benzi spectrale continue [55]. Tehnologia de teledetecție preia imaginile în diferite lungimi de undă şi apoi le transmite pe pământ prin telecomunicații [53] (vezi Figura 4.1), ceea ce oferă posibilitatea de a obține informații de pe suprafața pământului fără a avea contact cu elementele din zona de interes.

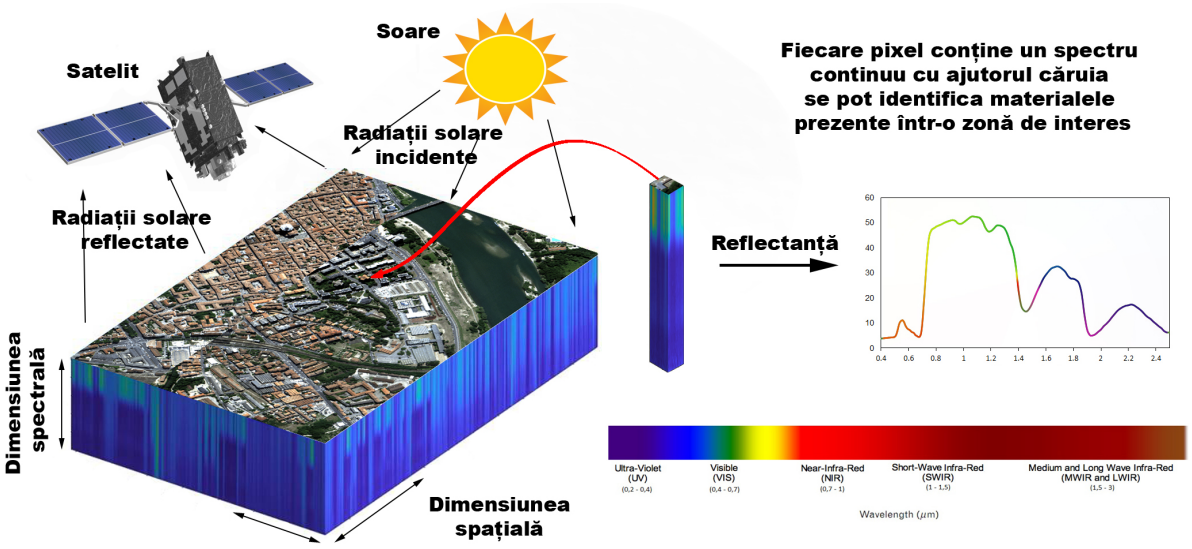


Figura 4.1: Colectarea datelor hiperspectrale utilizând tehnologia de teledetecție.

Semnătura spectrală a unui material este unică fiind un instrument puternic pentru identificarea materialelor care sunt prezente într-o zonă de interes [61]. Un material poate fi identificat din semnătura sa spectrală dacă sistemul de detecție are o rezoluție spectrală şi spațială suficientă pentru a distinge spectrul său de al altor materiale [3] deoarece la diferite lungimi de undă materialele reflectă şi absorb diferit. Imaginile hiperspectrale pot fi văzute ca tensori de ordinul trei, unde primii doi indici reprezintă informația spațială, iar cel de-al treilea indice indică dimensiunea spectrală. După cum imaginile hiperspectrale sunt considerate a fi date de dimensiuni mari, acest lucru poate fi un avatanj în ceea ce priveşte acuratețea rezultatelor, dar şi un dezavantaj datorită cantității mari de date care trebuie procesate. Acest fenomen a fost numit fenomenul Hughes sau blestemul dimensionalității [22], iar evoluția noilor tehnologii şi algoritmi vine în ajutorul analizei şi înțelegerii datelor de dimensiuni mari.

4.2 Metode şi tehnici de recunoaştere a imaginilor hiperspectrale

Recunoaşterea formelor aplicată pentru datele de tipul imaginilor hiperspectrale este un subiect de interes, prin urmare, acuratețea predicțiilor, dar şi complexitatea algoritmilor sunt măsuri importante având în vedere utilizarea sporită în diverse domenii de aplicare. Evaluarea mai multor algoritmi de învățare automată supravegheată [23] şi nesupravegheată [1] a fost studiată în literatura ştiințifică de specialitate cu scopul de a testa şi de a identifica cei mai potriviți algoritmi pentru clasificarea imaginilor hiperspectrale. Analiza datelor de dimensiuni mari

reprezintă o provocare deoarece depăşesc capacitatea software şi hardware şi nu pot fi tratate în mod obişnuit, dar şi datorită prezenței pixelilor cu format mixt. Algoritmii pentru descompunere tensorială utilizați în analiza imaginilor hiperspectrale au scopul de a reduce datele de dimensiuni mari, de a extrage esențialul, de a explica datele, dar şi de a obține modele precise uşor de analizat pentru identificarea materialelor şi estimarea diverşilor parametri dintr-o zonă de interes [59].

4.3 Modele Spectral Unmixing

Procesul de spectral unmixing vizează descompunerea imaginilor hiperspectrale într-o colecție de membri finali şi abundențele fracționate, în acest fel se obțin informații despre materialele pure prezente în zona de interes inclusiv distribuția lor [24]. Dobigeon *et al.* [38] afirmă că analiza spectrală este un instrument valoros deoarece oferă informații spațiale şi spectrale compacte şi esențiale despre imaginile hiperspectrale, pot fi interpretate fizic şi pot fi folosite în scenarii nesupravegheate.

4.4 Legătura dintre modelul Linear Spectral Unmixing şi descompunerile tensoriale

Linear Spectral unmixing este un instrument puternic pentru datele hiperspectrale şi se referă la orice proces care recuperează componentele spectrale pure, numiți membri finali şi un set de abundențe fracționale care indică proporția fiecărui membru final [26]. Membrii finali reprezintă materialele pure prezente în imagine, în timp ce setul de abundențe pentru fiecare pixel reprezintă procentul fiecărui membru prezent în pixel. Conform lui Qian *et al.* [43] informațiile despre poziția spațială sunt păstrate în tensor, desompunerile tensoriale oferind o precizie mai bună pentru problemele de spectral unmixing. Având în vedere că imaginile hiperspectrale sunt forme tridimensionale, descompunerile trilinare, *e.g.*, descompunerea Parafac, pot fi, de asemenea, utilizate pentru analiză şi procesare. Conform modelelor spectral unmixing, tensorul unei imagini hiperspectrale poate fi aproximat printr-o sumă de termeni, unde fiecare termen reprezintă produsul exterior dintre o matrice şi un vector, prin urmare, matricea descrie abundențele fiecărui pixel [43]. Considerând faptul că informațiile despre poziția spațială sunt păstrate în matrice, Qian *et al.* [43] afirmă că produsul dintre matricilor factor A şi B obținute în urma descompunerii Parafac ce corespund dimensiunilor spațiale [27], $A \cdot B^T$ este matricea abundențelor (*en.*: abundances matrix). În consecință, este stabilită o legătură între descompunerile tensoriale şi modelele spectral unmixing [43].

4.5 Concluzii

Imaginile hiperspectrale au o structură multidimensională motiv pentru care analiza tensorială poate fi o soluție în vederea obținerii de caracteristici valoroase care pot fi ulterior utilizate ca date de intrare pentru diverşi algoritmi de învățare automată. Conform Kolda şi Bader [27], analiza şi prelucrarea tensorială constă în două metode de bază, descompunerea unui tensor ca suma a unui număr finit de tensori de rang unu (PARAFAC) şi descompunerea unui tensor într-un tensor nucleu şi matrici factori de-a lungul fiecărei dimensiuni (Tucker). Drept urmare, în capitolele următoare sunt propuse metode de clasificare a imaginilor hiperspectrale şi a gesturilor care au la bază descompunerile tensoriale de bază utilizate pentru extragerea caracteristicilor, reducerea dimensiunilor mari a datelor şi creşterea acurateții clasificării.

5 Analiza imaginilor hiperspectrale utilizând tehnici de învățare automată

5.1 Tehnici de învățare supravegheată

Clasificarea imaginilor hiperspectrale presupune analiza, identificarea şi măsurarea materialelor de pe suprafața pământului fără a avea contact direct cu acestea. În vederea studierii rezultatelor generate de algoritmi pentru învățare automată supravegheată, această secțiune vizează utilizarea algoritmilor celor mai apropiați k vecini, Bayes Multinomial şi Gausian Naiv, analiza discriminantă liniară, arbori decizionali şi maşini cu vector suport folosind date hiperspectrale reale[2]. Pentru evaluarea predicțiilor, diverşi indicatori de performanță, cum ar fi valoarea acurateții, matricea de confuzie, scorurile F1, timpul de execuție şi pierderea validării încrucişate (*en.*: cross-validation loss) au fost calculați [7]. Conform rezultatelor obținute, SVM (*en.*: Polynomial Kernel) deşi este mai costisitor din punct de vedere al timpului de calcul a obținut cele mai bune valori ale acurateții ($\mu = 0.925$). Un algoritm de învățare supravegheată care a obținut valori ale acurateții foarte bune este 3NN ($\mu = 0.905$), cu mențiunea că timpul de execuție prevăzut este mai satisfăcător decât SVM (*en.*: Polynomial Kernel). Dintr-un alt punct de vedere, deşi algoritmii SVM liniar ($\mu = 0.785$) şi Bayes Multinomial ($\mu = 0.725$) au oferit cele mai mici valori ale acurateții dintre algoritmii studiați, timpul de execuție a fost mai scurt.

5.2 Tehnici de învățare nesupravegheată

Având în vedere că algoritmii de învățare supravegheată nu pot fi utilizați în toate cazurile deoarece ne obligă să avem cunoştințe anterioare despre materialele prezente în teren, această secțiune prezintă un studiu al performanței algoritmilor de învățare nesupravegheată, cum ar fi k-means şi gruparea ierarhică pentru clusterizarea imaginilor hiperspectrale. Pe lângă acest studiu, este prezentată şi discutată metoda Parafac-k-means ce presupune utilizarea descompunerii Parafac în vederea reducerii dimensiunii seturilor de date şi a algoritmului k-means pentru etichetarea semnăturilor spectrale [6]. În scopul evaluării algoritmilor de învățare nesupravegheată au fost studiate şi analizate date hiperspectrale care au fost achiziționate de pe site-urile *Grupo de Inteligencia Computational*[2] (S_1, S_2, S_3 şi S_4; rezoluție de 1.3 m/pixel) şi *U.S. Geologic Survey*[3] (S_5; rezoluție de 30 m/pixel [37], nu deține un set de date cu clasificare cunoscută).

În ceea ce priveşte metoda Parafac-k-means, sunt descrise rezultatele obținute pentru setul S_5. Pentru început, descompunerea Parafac a fost aplicată tensorului imaginii hiperspectrale în vederea reducerii dimensiunii mari a datelor şi extragerii caracteristicilor. Deoarece descompunerea Parafac este o metodă de analiză şi descompunere a datelor multi-direcționale, pentru recunoaşterea formelor se pot folosi instrumente de învățare nesupravegheată. În urma descompunerii Parafac datele corespunzătoare dimensiunii spectrale au fost clusterizate folosind algoritmul k-means. Construirea matricii de abundență ilustrată în Figura 5.1 (stânga) a fiecărei clase dintr-o imagine hiperspectrală porneşte de la faptul că descompunerile tensoriale sunt văzute ca modele liniare [59]. În Figura 5.1 (dreapta) se pot vedea hărțile de abundență construite cu scopul de a evidenția clasele identificate; culorile pornesc de la *albastru* pentru o contribuție scăzută, spre *galben* pentru contribuție ridicată [52].

[2] `http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes`
[3] `https://earthexplorer.usgs.gov`

În urma studierii rezultatelor obținute, timpul de execuție al clusterizărilor după utilizarea descompunerii Parafac este vizibil mai mic (în medie, aproximativ 3000 μs, comparativ cu aproximativ 1921s pentru gruparea ierarhică şi k-means a datelor matricizate; în cazul setului S_2 timpul necesar execuției prelucrărilor algoritmilor este foarte mare, 3794s, restul timpilor fiind în intervalul [0.030s, 36.47s]), unde timpul necesar execuției prelucrărilor algoritmului Parafac este în medie 5.4 secunde. Drept urmare, descompunerile tensoriale sunt promițătoare în analiza datelor de dimensiuni mari.

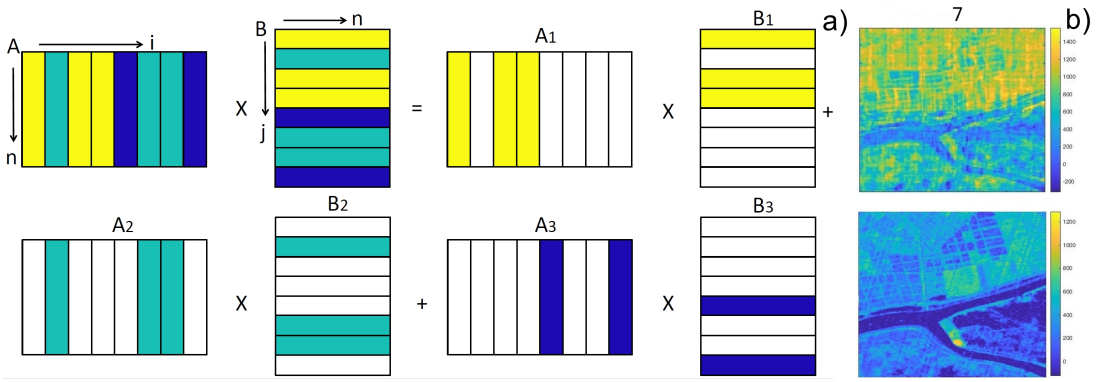


Figura 5.1: Construirea matricilor de abundență pentru fiecare clasă identificată (a) şi reprezentarea grafică a hărții abundențelor corespunzătoare clasei şapte identificată în setul S_5 în urma aplicării metodei Parafac-k-means şi probele 25 din 53 (b).

5.3 Segmentarea matricilor de abundență

Segmentarea imaginilor este o metodă folosită pentru a identifica obiecte şi granițe în imagini fiind un proces nesupravegheat de împărțire a unei imagini în mai multe segmente, astfel suprafața şi regiunile vor fi izolate, devenind mai uşor de analizat [19, 54]. Conform metodei de segmentare în subintervale egale studiată de Bilius şi Pentiuc [31], segmentarea imaginilor hiperspectrale facilitează detectarea materialelor prin analiza vizuală şi compararea spectrelor asociate cu fiecare regiune. Această secțiune propune o metodă nesupravegheată de clusterizare a imaginilor hiperspectrale într-un număr definit de regiuni, care identifică materiale dintr-o zonă de interes folosind descompunerile tensoriale şi segmentarea hărții abundențelor bazată pe algoritmul nucleelor dinamice (*en.*: Dynamic Clustering Algorithm [16, 26]) [5]. În Figura 5.2 sunt ilustrate etapele metodei propuse pentru clusterizarea imaginilor hiperspectrale.

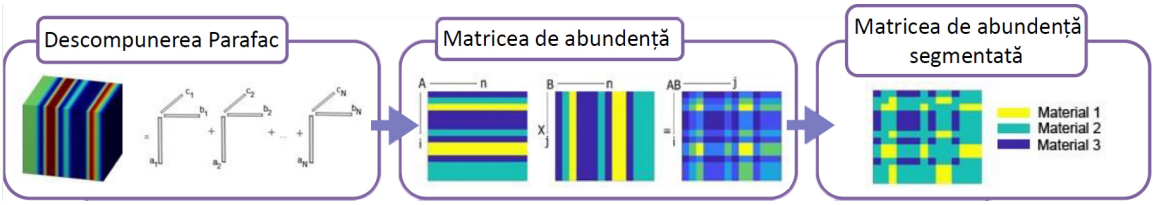
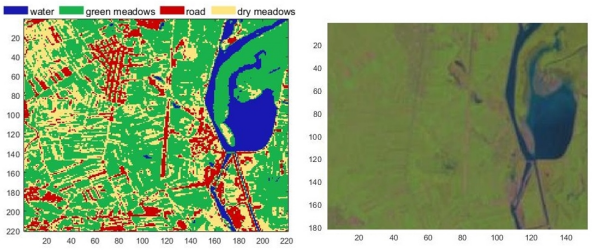


Figura 5.2: Fluxul procesului metodei de segmentare a matricii de abundență.

În dezvoltarea acestei metode [5] au fost utilizate date hiperspectrale preluate de pe site-ul *U.S. Geologic Survey*[3] cu o rezoluție de 30 m/pixel [37]. Tensorul hiperspectral este de dimensiune $221 \times 221 \times 38$. Metoda doreşte obținerea unei reprezentări vizuale a matricii abundențelor

segmentată, numită harta abundențelor segmentată [25]. Matricea abundențelor a fost construită folosind matricile factor corespunzătoare dimensiunilor spațiale obținute în urma descompunerii Parafac. Scopul este de a găsi un algoritm dedicat studiului datelor hiperspectrale utilizând o tehnică rapidă şi eficientă. Construirea matricii de abundență este posibilă datorită legăturii dintre modelele spectral unmixing şi descompunerea tensorială care oferă o interpretare uşoară a datelor, păstrând structura spectrală şi spațială fără pierderi de informații relevante.



Segmentarea imaginii constă în aplicarea unui algoritm ce se bazează pe principiul algoritmului nucleelor dinamice [16]. Pentru a alege numărul potrivit de clase, au fost analizate mai multe cazuri care au fost ulterior comparate cu imaginile satelitare, inclusiv dendrogramele corespunzătoare. Rezultatele experimentale ale metodei au furnizat regiuni omogene datorită descompunerii Parafac şi algoritmului nucleelor dinamice [16, 43]. În urma acestui pas, fiecărui material îi este asociat vizual câte o etichetă. Acest lucru a fost realizat prin analizarea imaginilor satelitare şi a hărții abundențelor. Aceste rezultate nu sunt supravegheate şi pot fi utilizate pentru orice alt set de date hiperspectral, având avantajul că harta abundențelor segmentată este uşor de obținut.

5.4 Concluzii

În urma analizării algoritmilor de învățare supravegheată şi nesupravegheată, clasificarea imaginilor hiperspectrale de dimensiuni mici se poate realiza cu succes. În ceea ce priveşte seturile de date de dimensiuni mari, timpul necesar execuției prelucrărilor algoritmilor este costisitor devenind astfel ineficienți, prin urmare, descompunerea Parafac deține avantajul de a genera modele de dimensiuni mici, uşor de interpretat. Hărțile abundențelor obținute în urma descompunerii Parafac sunt omogene considerând că fracțiunea unui pixel dintr-o anumită clasă este estimată de harta abundenței corespunzătoare. Referitor la segmentarea imaginilor, metoda eficientă care se bazează pe algoritmul nucleelor dinamice prezentat de Diday [16] şi metoda de segmentare care asociază clasa unui material după un anumit prag de segmentare au oferit rezultate foarte bune. Tehnicile de segmentare au fost utilizate pentru a obține o reprezentare care să faciliteze interpretarea informațiilor conținute în imaginea hiperspectrală. Procesul de segmentare are loc în trei etape fiind un mod eficient şi economic din punct de vedere al resurselor imaginilor hiperspectrale. Contribuția acestui studiu constă în utilizarea descompunerii Parafac pentru a extrage abundențele şi utilizarea a două metode segmentare pentru matricile de abundență. Toate aceste prelucrări pot fi utilizate la proiectarea unui instrument specializat pentru analiza imaginilor hiperspectrale. Ținând cont de rezultatele obținute după utilizarea descompunerilor multiliniare algebrice tensoriale, Parafac este utilă pentru analiza datelor de dimensiuni mari ca urmare a avantajelor privind reducerea timpului necesar execuției prelucrărilor algoritmilor şi utilizarea eficientă a memoriei.

Contribuțiile prezentate în cadrul acestui capitol au fost publicate în următoarele lucrări ştiințifice A_2, A_7, A_8 şi A_{13}; vezi Capitolul 8.1.3.

6 Metodă eficientă de clasificare a imaginilor hiperspectrale folosind diagrame Voronoi şi forme tari

Acest capitol introduce o metodă de învățare nesupravegheată ce constă în reducerea dimensiunilor mari ale imaginilor hiperspectrale utilizând descompunerea Parafac şi procesarea unei cantități semnificativ mai reduse de date [8].

6.1 Descrierea metodei

Algoritmii de recunoaştere a formelor consumă multe resurse pentru a analiza datele hiperspectrale datorită prezenței unui număr mare de benzi spectrale. Pentru date de dimensiuni mari, execuția unui algoritm de determinare a partițiilor folosind biblioteca *Scikit-learn Library Machine Learning* din Python [17] poate dura minute, prea mult pentru aplicații care trebuie să răspundă într-un timp interactiv, iar de cele mai multe ori aplicarea unui singur algoritm de învățare nesupravegheată nu este suficient în toate cazurile. Drept urmare, este necesară găsirea unei soluții eficiente pentru clasificarea imaginilor, astfel încât memoria ocupată şi timpul necesar execuției prelucrărilor algoritmilor să fie acceptabile pentru probleme practice.

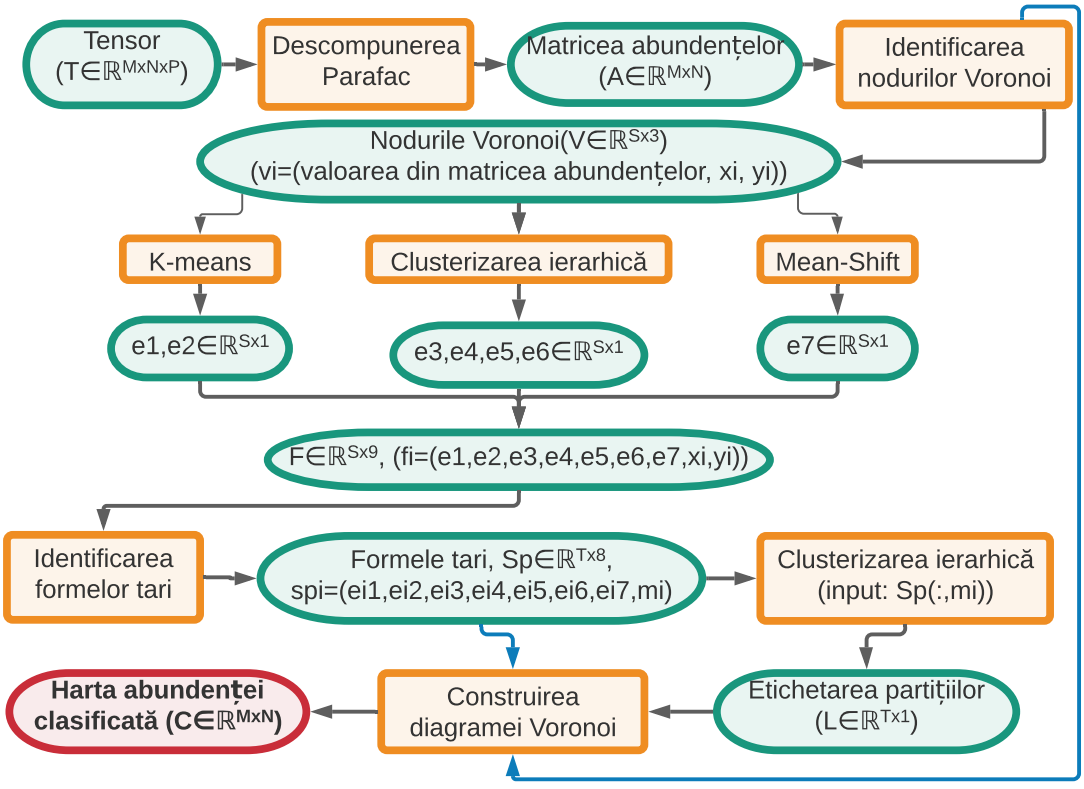


Figura 6.1: Etapele metodei de recunoaştere a formelor din imaginile hiperspectrale utilizând descompunerea Parafac, diagramele Voronoi şi formele tari [8].

În vederea reducerii consumului de resurse de calcul şi îmbunătățirii calității clasificării, metoda propusă se bazează pe algoritmi de învățare automată nesupravegheată folosind descompunerea Parafac, diagramele Voronoi şi formele tari fără a avea cunoştințe anterioare des-

pre conținutul imaginilor. Figura 6.1 descrie etapele metodei de recunoaştere a formelor din imaginile hiperspectrale.

Conform Qian *et al.* [43], există o legătură între descompunerile tensoriale şi modelele spectral unmixing, prin urmare, în dezvoltarea metodei propuse utilizarea descompunerii Parafac a tensorului imaginii hiperspectrale $T \in \mathbb{R}^{M \times N \times P}$ are rolul de a reduce dimensiunea mare a datelor şi de a extrage caracteristicile. Construcția diagramei Voronoi se bazează pe alegerea celor mai reprezentative noduri Voronoi din matricea abundențelor notată $A \in \mathbb{R}^{M \times N}$ în Figura 6.1. Pentru identificarea nodurilor Voronoi, $V \in \mathbb{R}^{S \times 3}$ (unde S reprezintă numărul de noduri) sunt propuse trei metode care sunt evaluate, analizate şi discutate folosind patru seturi de date sintetice şi setul de date, Insula Mică Brăilei, România (vezi Secțiunea 6.4).

Construirea partiției multiple, $F \in \mathbb{R}^{S \times 9}$ se obține în urma aplicării tehnicilor de învățare nesupravegheată, cum ar fi, grupările ierarhic, k-means şi mean-shift. Deoarece nu există un singur algoritm general aplicabil în orice situație, utilizarea mai multor algoritmi de învățare nesupravegheată generează mai multe partiții corespunzătoare pixelilor din imagine. Chiar şi aplicarea aceluiaşi algoritm cu valori diferite ale unor parametri sau moduri de inițializare a soluției de pornire poate duce la obținerea unor partiții diferite. Având aceste partiții apare problema găsirii unei singure partiții care să fie soluție. Găsirea unei partiții consens din multi-partițiile obținute, aşa cum a fost numită de Bejar [4] sau partiție centrală, aşa cum a fost numită de Diday [16] este o problemă NP dificilă. Rezolvarea acestei probleme într-un timp eficient implică aplicarea unei metode euristice. În acest scop, construirea matricii multi-partițiilor a fost utilizată pentru a identifica formele tari, $Sp \in \mathbb{R}^{T \times 8}$ (T reprezintă numărul de forme tari şi ultima coloană reprezintă valoarea corespunzătoare asociată din matricea abundențelor). Formele tari au fost identificate folosind gruparea ierarhică în urma căreia, partiția optimă, $L \in \mathbb{R}^{T \times 1}$ a fost obținută. După identificarea formelor tari, cu ajutorul acestora are loc extinderea clasificării la întreaga imagine hiperspectrală prin construirea diagramei Voronoi obținându-se astfel harta abundențelor clusterizată, $C \in \mathbb{R}^{M \times N}$. Figura 6.1 descrie etapele metodei de recunoaştere a formelor din imaginile hiperspectrale.

Metoda propusă a fost evaluată utilizând seturi de date sintetice obținute prin generarea pentru fiecare clasă de valori uniform sau gausian distribuite. În ansamblu, metoda propusă se bazează pe descompunerea triliniară a tensorului unei imagini hiperspectrale în vederea reducerii dimensiunii datelor, pe o subdiviziune a hărții abundenței în celule uniforme de pixeli numite celule Voronoi şi pe analiza multiplă a partițiilor.

6.2 Diagrame Voronoi

Definim diagrama Voronoi $Vor(P)$ pentru un set de puncte $P := \{p_1, p_2, ..., p_n\}$ care aparțin unui plan α, numite noduri, ca fiind o subdiviziune a planului din care fac parte punctele p_i în n celule $Vor(p_i), i = \overline{1,n}$ [34]. Proprietatea diagramelor Voronoi constă în faptul că orice punct q care se află într-o celulă c_i care corespunde unui punct p_i respectă Ecuația 6.1 [34].

$$q \in \alpha \iff d(q, p_i) < d(q, p_j), \forall\, p_i, p_j \in P, j \neq i \tag{6.1}$$

unde funcția d poate fi, *e.g.*, distanța Euclideană, Manhattan, Chebyshev; vezi Figura 6.3b) pentru o reprezentare grafică a diagramei Voronoi corespunzătoare unui set de puncte.

6.3 Descompunerea Parafac şi construirea matricii de abundență

Evaluarea metodei propuse presupune utilizarea unui set de date hiperspectral care a fost preluat de pe site-ul *USGS (U.S. Geological Survey)*[3] cu o rezoluție de 30 m/pixel [37]; imaginea hiperspectrală a unei zone din Brăila, România, mai exact *Insula Mică a Brăilei* (vezi Figura 6.6d) care nu deține un set de date cu clasificare cunoscută. Modelele spectral unmixing extrag abundențele din imaginile hiperspectrale astfel încât informațiile spectrale şi spațiale sunt păstrate în harta abundențelor alături de corelațiile dintre pixeli [52]; harta abundențelor este ilustrarea grafică a matricii abundențelor. În această direcție, Qian *et al.* [43] au introdus legătura dintre modelele spectral unmixing şi descompunerile tensoriale, ceea ce conduce la posibilitatea de a extrage spectrele constitutive şi abundențele corespunzătoare prin intermediul descompunerilor tensoriale. În consecință, matricea abundențelor a fost construită în urma descompunerii Parafac. Setul de date hiperspectral a fost redus cu aproximativ 97%, matricea de abundență construită în urma descompunerii reprezentând 3% din tensorul imaginii hiperspectrale şi conține cele mai relevante caracteristici.

6.4 Identificarea nodurilor Voronoi

Această secțiune descrie trei metode de identificare a nodurilor Voronoi care reduc semnificativ dimensiunea setului de date. Acest proces are la bază ideea că dacă se găseşte o diagramă Voronoi care aproximează într-un mod acceptabil matricea inițială, atunci această matrice ar putea fi reprezentată de nodurile diagramei Voronoi [8]. Este evident că dacă se aplică o serie de etichete nodurilor Voronoi, va fi posibilă construirea unei diagrame Voronoi care va propaga etichetarea tuturor regiunilor din imaginea hiperspectrală. În continuare sunt descrise şi analizate rezultatele obținute pentru trei seturi de date de dimensiuni $D \in \mathbf{R}^{M \times N}$, *beans gaussian* ($D_1$), *beans uniform* ($D_3$) şi *Insula Mică a Brăilei, România* (D_2). Pentru toate seturile de date $M = 201, N = 301$, iar toate valorile generate aparțin intervalului $[0, 1]$. Eroarea este calculată conform Ecuației 6.2, unde $D \in \mathbf{R}^{M \times N}$ este matricea inițială, iar $VD \in \mathbf{R}^{M \times N}$ este diagrama obținută după găsirea partiției consens.

$$er = \frac{\sum_{i,j=1}^{M,N} |D(i,j) - VD(i,j)|}{\sum_{i,j=1}^{M,N} D(i,j)} \tag{6.2}$$

6.4.1 Metoda $\mathbf{M}_1$ de identificare a nodurilor Voronoi

Fie $E_n \in \mathbf{R}^{M \times N}$ matricea de abundență normalizată (valori între $[0, 1]$) care este discretizată în subintervale de lungime ϵ. În urma discretizării se obține o nouă matrice F, care este folosită pentru a alege nodurile Voronoi. Pentru început, matricea E_d este aranjată în mulțimea $Z \in \mathbf{R}^{MN}$. Folosind mulțimea Z se obține mulțimea $S_u \in \mathbf{R}^{Q}$, care conține valorile unice din Z în ordine crescătoare. În continuare se construieşte matricea $F \in \mathbf{N}^{M \times N}$ asociind valorile pixelilor din E_d cu indexul valorilor pixelilor din S_u (vezi Ecuația 6.3).

$$\begin{aligned} Z = \{e_d(i,j) \in E_d\} = \{z_1, z_2, ..., z_k\},\ \forall \text{i, j},\ i = \overline{1,M},\ j = \overline{1,N},\ k = \overline{1, M \times N} \\ S_u = \{s_{u_1}, s_{u_2}, ..., s_{u_q}\},\ s_{u_{i'}} < s_{u_{j'}},\ i' < j',\ \forall \text{i'}, \text{j'},\ S_u \subset Z \\ f(i,j) = t, \text{dacă } e_d(i,j) = s_{u_t},\ t = \overline{1,q} \end{aligned} \tag{6.3}$$

Pentru a găsi nodurile Voronoi, matricea F este împărțită în blocuri pătratice de dimensiune n (pentru evaluarea metodelor propuse $n = 5$), iar histograma H a fiecărui bloc $F(i, j)$ este construită [39]. Numărul de intervale n_ϵ este calculat astfel: $[(max(E_d) - min(E_d))/\epsilon]$. Dacă valoarea maximă din histograma pentru blocul $F(i, j)$ este mai mare decât $p\% * nr$ (nr este numărul de pixeli din bloc), se poate spune că blocul F este pur și se calculează centrul de greutate al blocului, care va fi un nod Voronoi. Valoarea corespunzătoare nodului Voronoi va fi mediana a pixelilor care aparțin maximului din histogramă. Dacă condiția anterioară nu este îndeplinită, blocul $F(i, j)$ este împărțit în patru și se verifică din nou condiția pentru fiecare celulă nouă. Dacă condiția continuă să nu fie îndeplinită, procesul se poate repeta până când aceasta este îndeplinită sau celula conține un singur pixel care va deveni un nod Voronoi.

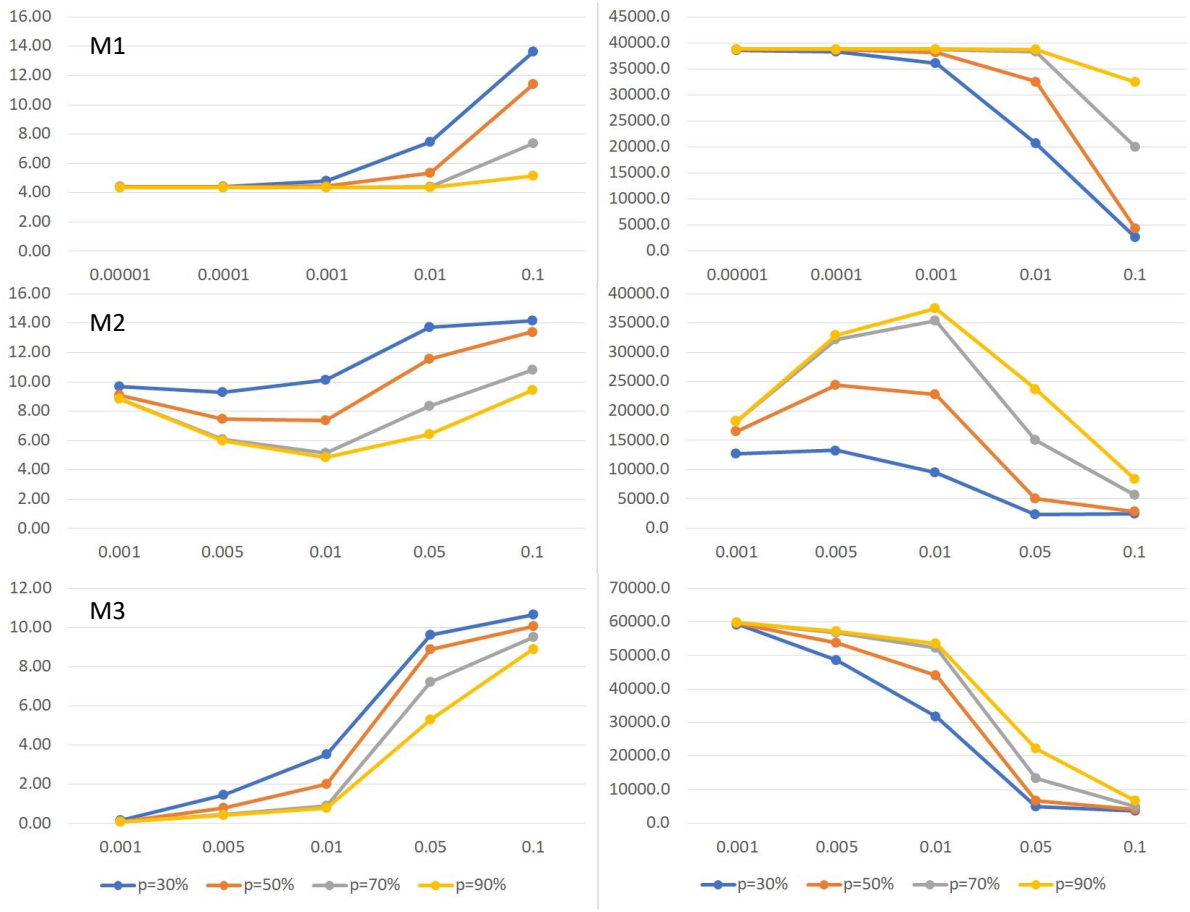


Figura 6.2: Media erorilor (seturile de date D_1, D_2 și D_3) în funcție de gradele de libertate ϵ și puritate (p%) (stânga) și media numărului de noduri Voronoi (seturile de date D_1, D_2 și D_3) obținute pentru harta abundențelor (301 × 201) în funcție de ϵ și puritate (dreapta) folosind metodele M_1, M_2 și M_3.

6.4.2 Metoda M_2 de identificare a nodurilor Voronoi

O altă metodă de a alege nodurile Voronoi reprezentative presupune împărțirea matricii de abundență în blocuri (matrici pătratice) de dimensiune n (valoare aleasă de utilizator). Din fiecare bloc generat se aleg noduri care să reprezinte cât mai bine blocul din care fac parte. Metoda presupune calcularea de medii recursive în fiecare celulă conform Ecuației 6.4; ϵ este o valoare aleasă de utilizator.

$$m_i = \frac{\sum_{k=1}^{q_i}(x_k)}{q_i}, \text{ dacă } \left(\frac{\sum_{k=1}^{q_i-1}(x_k)}{q_i - 1} - x_k\right) < \epsilon, \forall i = \overline{1, nr} \tag{6.4}$$

Conform Ecuației 6.4 se va identifica un număr nr de centre în cadrul unui bloc. Fiecărei medii m_i îi va corespunde un număr q_i de pixeli. Un nod Voronoi este găsit dacă valoarea lui $q_i = max(q_i) \geq p\% * nr$; $\forall i = \overline{1, nr}$, nr este numărul total de pixeli din bloc şi p este o valoare aleasă de utilizator. Nodul Voronoi va avea coordonatele centrului de greutate al pixelilor ce corespund şirului q_i, iar valoarea asociată va fi valoarea medianei valorilor din matricea abundenței a setului de pixeli q_i. Dacă valoarea lui $q_i < p\% * nr$, atunci blocul se va împărți în patru şi se vor recalcula mediile conform Ecuației 6.4 până când condiția va fi îndeplinită. Dacă blocul va fi dimensiune 1×1 atunci acesta va deveni un nod Voronoi.

6.4.3 Metoda M_3 de identificare a nodurilor Voronoi

A treia metodă de a alege nodurile Voronoi reprezentative presupune împărțirea matricii de abundență în blocuri (matrici pătratice) de dimensiune n. Metoda propusă caută în fiecare bloc toate valorile nr_ϵ care aparțin intervalului $[mean(x_i) - \epsilon, mean(x_i) + \epsilon]$, pentru $x_i, i = \overline{1, nr}$, unde nr este numărul total de pixeli din blocul curent. Dacă $nr_\epsilon \geq p\% * nr$, atunci se va calcula centrul de greutate (x_g, y_g) ai pixelilor care aparțin intervalului. Nodul Voronoi (x_v, y_v) va avea coordonatele pixelului din blocul curent care este cel mai apropiat de (x_g, y_g). Valoarea asociată nodului va fi valoarea din matricea de abundență corespunzătoare lui (x_v, y_v). Dacă condițiile continuă să nu fie îndeplinite, procesul se poate repeta până când aceasta este îndeplinită sau celula conține un singur pixel care va deveni un nod Voronoi.

În Figura 6.3b) se poate vedea o diagramă Voronoi corespunzătoare setului de date *Insula Mică a Brăilei* (D_2) generată pentru evidențierea nodurilor şi celulelor Voronoi. De asemenea, diagrama din Figura 6.3c) este ilustrată pentru a realiza o comparație vizuală cu harta abundențelor având o eroare de 9.45%. Rezultatele obținute în urma aplicării metodelor M_1, M_2 şi M_3 se pot vedea în Figura 6.2. Un rezultat acceptat de utilizator după aplicarea algoritmilor descrişi pentru identificarea nodurilor Voronoi folosind diverse grade de libertate (*e.g.*, parametrii ϵ, puritate sau dimensiunea blocului) are loc atunci când există un compromis între numărul de noduri şi eroare. Metodele propuse reduc semnificativ atât numărul de forme, cât şi memoria ocupată. Datele reduse au reprezentat în medie 52.9% (M_1), 27.4% (M_2), respectiv 55.9% (M_3) din matricea abundențelor sau a datelor sintetice în cazul celor trei metode de identificare a nodurilor Voronoi. În cazul imaginii hiperspectrale *Insula Mică a Brăilei*, datele utilizate pentru clusterizare au reprezentat în medie 1.55%, 0.81%, respectiv 1.65% din întreg tensorul pentru cele trei metode de identificare a nodurilor Voronoi, cu mențiunea că datele din matricea abundenței reprezintă 2.94% din întreaga imagine hiperspectrală.

6.5 Forme tari şi partiția consens

Odată ce setul de noduri este definit utilizând una din metodele descrise în Secțiunea 6.4, nodurile sunt clusterizate folosind mai mulți algoritmi de învățare nesupravegheată, cum ar fi, grupările ierarhic, k-means şi mean-shift. Pentru fiecare nod s-au concatenat etichetele atribuite de algoritmii de învățare nesupravegheată, astfel s-a construit matricea multi-partițiilor pentru

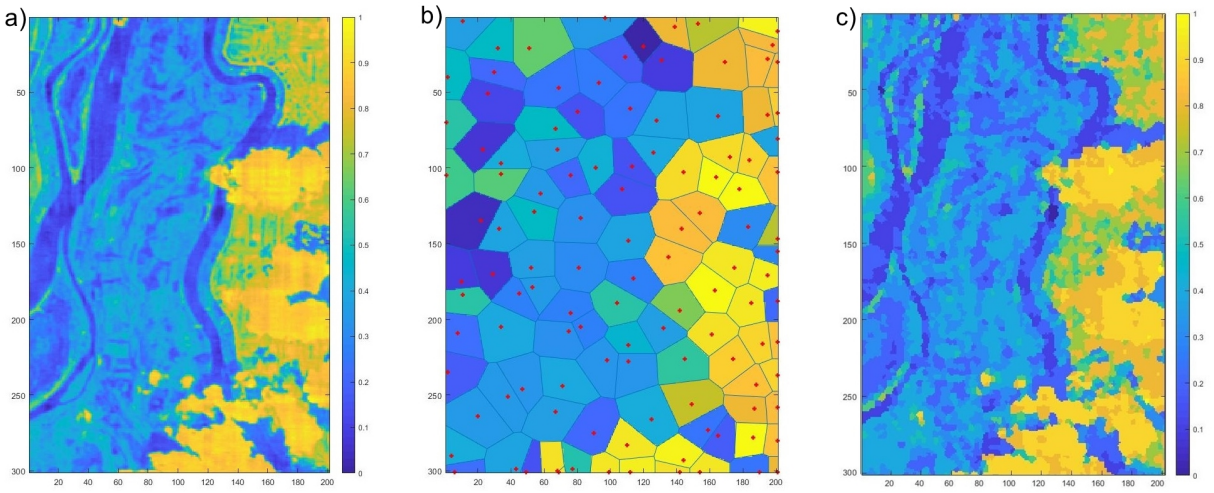


Figura 6.3: Reprezentarea grafică a matricii de abundență (a) şi a diagramei Voronoi corespunzătoare setului de date *Insula Mică a Brăilei* D_2 (dimensiune bloc=25) (b) şi diagrama Voronoi unde valorile nodurilor Voronoi sunt reprezentate de valorile din matricea de abundență (c).

toate nodurile. Formele tari vor fi căutate în matricea multi-partițiilor [42] (vezi Figura 6.1). Formele tari sunt definite ca fiind subseturile de partiții care au fost tot timpul clusterizate împreună [14, 42]; Figura 6.4, pasul I. Scopul este de a găsi o partiție consens pentru multi-partiția obținută anterior, cu mențiunea că această sarcină este o problemă NP dificilă. În vederea rezolvării acestei sarcini, euristica propusă presupune aplicarea grupării ierarhice pentru formele tari extrase din matricea multi-partițiilor (pasul II din Figura 6.4).

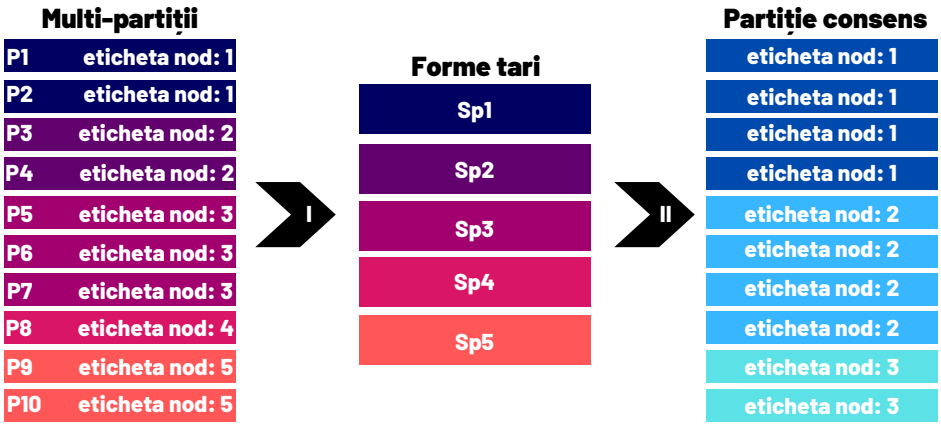


Figura 6.4: Obținerea formelor tari (I) şi a partiției consens (II) din matricea multi-partițiilor.

6.6 Evaluarea metodei propuse

În continuare sunt prezentate rezultatele obținute pentru seturile de date sintetice şi imaginea hiperspectrală *Insula Mică a Brăilei* utilizând nodurile identificate în urma aplicării metodelor M_1, M_2 şi M_3 descrise în Secțiunea 6.4 folosind diverşi parametri.

În vederea obținerii partiției consens pentru setul *beans uniform* (D_3) au fost utilizate nodurile Voronoi (25489 noduri şi eroare de 10.12%) obținute în urma aplicării metodei M_1. Pentru a ajunge la partiția consens s-a construit matricea multi-partițiilor (25489×7) prin aplicarea al-

goritmilor de învățare automată menționați anterior. Din matricea multi-partițiilor au fost identificate 9 forme tari (9×7) care au fost clusterizate folosind gruparea ierarhică, obținând astfel partiția consens. Valorea asociată unei forme tari, Sp_i este media valorilor nodurilor Voronoi care au fost asociate cu Sp_i. În cele din urmă, fiecare partiție din matricea multi-partițiilor este asociată cu eticheta formei tari din care face parte. Odată ce partițiile sunt asociate cu etichete de clasă, nodurile Voronoi sunt şi ele etichetate, prin urmare, va fi posibilă construirea unei diagrame Voronoi care va propaga etichetarea tuturor regiunilor, extinzând astfel etichetarea la întreaga imagine; vezi Figura 6.5. Seturile de date cu clasificare cunoscută au fost folosite doar pentru evaluarea acurateții algoritmului propus. Acuratețea clasificării este de 98.44%. Pentru a calcula acuratețea, corespondența dintre clasele corespunzătoare setului de date cu clasificare cunoscută şi clasele obținute în urma aplicării metodei propuse a fost asociată vizual.

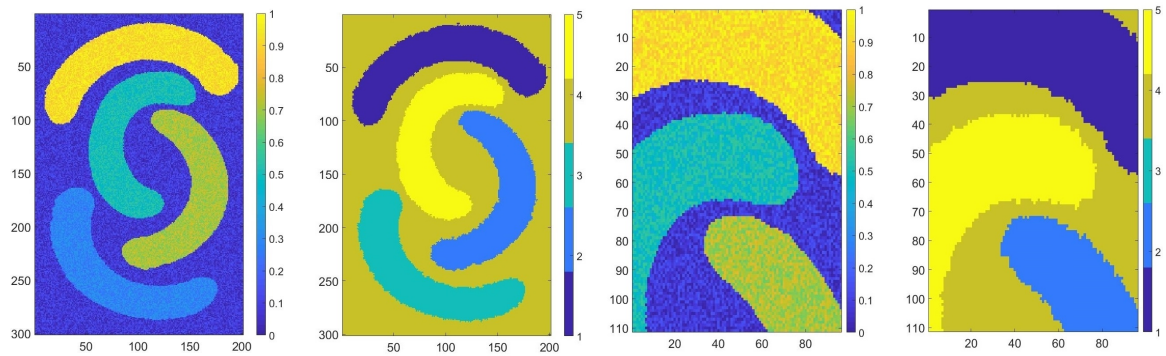


Figura 6.5: Reprezentarea grafică a setului de date *beans uniform*, D_3 (stânga) şi a extinderii clasificării la întreaga imagine folosind şapte algoritmi de învățare nesupravegheată (metoda M_1; $\epsilon = 0.01$, puritate = 30, dimensiunea blocului de cinci pixeli) pentru generarea multi-partițiilor.

Metodele M_1, M_2 şi M_3 au fost testate şi în cazul celorlalte seturi pentru care s-au obținut valori ale acurateții foarte bune, *e.g.*, pentru M_1 şi setul *beans gaussian* (D_1), $\epsilon = 0.05$, puritate = 60, 11886 noduri, eroare = 7.81%, acuratețea clasificării = 98.10%, pentru M_2 şi setul *beans uniform* (D_3), 10027 noduri, $\epsilon = 0.1$, puritate = 90, eroare = 13.16%, acurateț = 97.49%, pentru M_2 şi setul *beans gaussian* (D_1), 10225 noduri, $\epsilon = 0.05$, puritate = 80, eroare = 8.25%, acurateț = 97.60%, pentru M_3 şi setul *beans uniform* (D_3), 20996 noduri, $\epsilon = 0.05$, puritate = 70, eroare = 9.38%, acurateț = 99.39% şi pentru M_3 şi setul *beans gaussian* (D_1), 9920 noduri, $\epsilon = 0.02$, puritate = 50, eroare = 6.77%, acurateț = 99.17%.

În vederea obținerii partiției consens pentru setul de date *Insula Mică a Brăilei* (D_2; $\epsilon = 0.1$, puritate = 60 şi dimensiunea blocului de cinci pixeli), nodurile Voronoi (11032 noduri şi eroare = 6.24%) au fost identificate din matricea abundenței aplicând Metoda M_1. Urmând aceiaşi paşi menționați anterior au fost identificate 25 de forme tari care au fost clusterizate în cinci clase (valoare aleasă în urma analizării dendrogramei şi a imaginii satelitare) folosind gruparea ierarhică, obținând astfel partiția consens; vezi Figura 6.6. Metodele M_2 şi M_3 au fost de asemenea testate, obținându-se rezultate omogene; vezi Figura 6.6 pentru reprezentările grafice ale extinderii clasificării la întreaga imagine în urma aplicării M_2 — $\epsilon = 0.1$, puritate = 90, 10633 noduri, eroare de 6.33%, 25 forme tari şi M_2 — $\epsilon = 0.1$, puritate = 90, 9482 noduri, eroare de 6.32%, 25 forme tari.

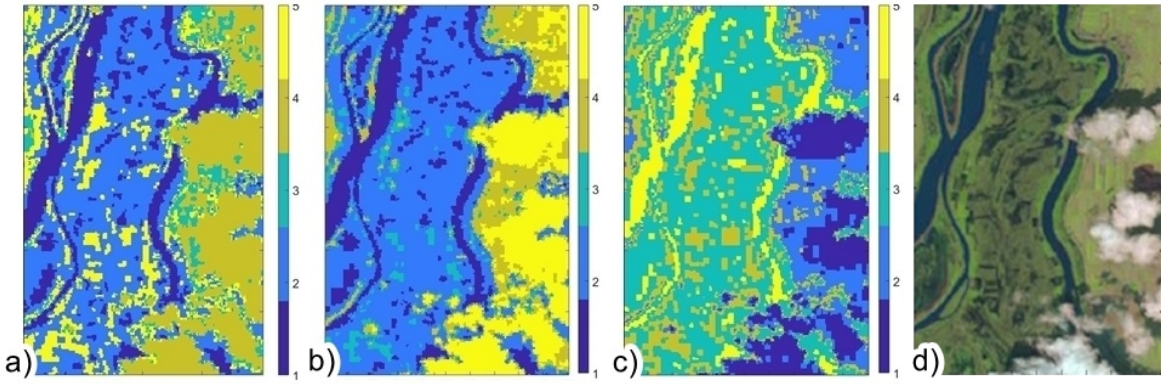


Figura 6.6: Partiția consens folosind şapte algoritmi nesupravegheați aplicând M_1 (a), M_2 (b) şi M_3 (c) şi (d) imaginea satelitară setului de date *Insula Mică a Brăilei*.

6.7 Avantajele metodei propuse

Metoda propusă se bazează pe descompunerea triliniară a tensorului unei imagini hiperspectrale în vederea reducerii dimensiunii datelor, pe o subdiviziune a hărții abundenței în celule uniforme de pixeli numite celule Voronoi şi pe analiza multiplă a partițiilor. Aceasta presupune o reducere a tensorului imaginilor hiperspectrale prin aplicarea descompunerii Parafac şi clasificarea unui număr semnificativ mai mic de forme fără a avea cunoştințe despre materialele prezente în zona de interes şi fără pierderi mari de informații. Pornind de la definiția diagramei Voronoi, ce presupune împărțirea unui plan în subregiuni, acest concept a fost explorat pentru imaginile hiperspectrale. Metoda are la bază ideea că dacă se găseşte o diagramă Voronoi care aproximează într-un mod acceptabil matricea de abundență, atunci această matrice ar putea fi reprezentată de nodurile diagramei Voronoi [8]. Pentru a reduce numărul de forme care vor fi ulterior clasificate cu algoritmi de învățare nesupravegheată au fost propuse trei metode de identificare a celor mai potrivite noduri Voronoi. Conform mediilor erorilor şi mediilor numărului de noduri Voronoi obținute pentru seturile de date D_1, D_2 şi D_3 în funcție de gradele de libertate (ϵ şi puritate) folosind metodele M_1, M_2 şi M_3 trebuie să existe un compromis între numărul de noduri şi eroare. Metodele propuse reduc semnificativ atât numărul de observații, cât şi memoria ocupată. Deoarece nu există un algoritm general valabil care generează o partiție perfectă, au fost construite mai multe partiții şi s-a căutat cea mai bună partiție consens abordând o metodă euristică care se realizează prin gruparea ierarhică a formelor tari. Formele tari conțin submulțimi de elemente care au fost clasificate împreună, acestea fiind identificate din matricea multi-partițiilor. În urma grupării ierarhice a formelor tari, partiția consens este obținută, drept urmare, şi nodurile Voronoi sunt asociate cu etichete de clasă. Clasificarea matricii de abundență implică construirea diagramei Voronoi care extinde clasificarea la întreaga imagine hiperspectrală. Având în vedere că imaginile hiperspectrale sunt date de dimensiuni mari, scopul acestei metode a fost de a reduce timpul de execuție al clasificării prin reducerea modelului şi a numărului de observații la un număr semnificativ mai mic de noduri Voronoi şi forme tari. În concluzie, rezultatele obținute prin aplicarea metodei propuse au fost foarte bune în ceea ce priveşte calitatea clasificării, dar şi din punctul de vedere al utilizării eficiente a resurselor de calcul implicate.

Contribuțiile prezentate în cadrul acestui capitol au fost publicate în următoarele lucrări ştiințifice A_1 şi A_{14}; vezi Capitolul 8.1.3.

7 Metodă de recunoaştere a gesturilor

Achiziția datelor care caracterizează gesturile executate cu mâna în aer (*en.:* mid-air hand gestures) implică un număr de valori numerice colectate de la diverşi senzori cu o frecvență semnificativă de eşantionare pe toată durata articulării, ceea ce conduce la reprezentarea seturilor de gesturi printr-un volum mare de date. În această direcție, descompunerea Tucker este o metodă de extragere a caracteristicilor, fiind de asemenea şi o metodă de compresie şi reducere a dimensionalității [41]. O reducere a volumului datelor de intrare utilizate pentru antrenarea algoritmilor supravegheați poate reduce semnificativ resursele de calcul necesare pentru a rula eficient algoritmii de învățare într-un timp de răspuns interactiv. Intersectând descompunerea Tucker cu metoda ansamblu de învățare (ELMet) propusă, acest capitol introduce metoda DELMet pentru recunoaşterea gesturilor executate cu mâna în aer (*en.*: mid-air gestures) folosind valorile colectate de senzori inerțiali.

7.1 Introducere în analiza gesturilor executate cu mâna în aer

Gesturile executate cu mâna în aer sunt metode de intrare care au loc prin manipularea fără atingere a conținutului digital. Tehnicile de recunoaştere îşi propun să înțeleagă gesturile executate cu mâna în aer, fiind o problemă de învățare automată [63].

7.2 Studii explorative cu privire la interacțiuni gestuale

În literatura ştiințifică inelele smart au fost utilizate ca metode de control [56] deoarece prezintă numeroase avantaje, cum ar fi, versatilitatea, portabilitatea şi dimensiunea mică, care le fac potrivite pentru diverse aplicații, chiar şi ca dispozitive de detectare. Inele smart au aplicații în diverse domenii, *e.g.*, aplicația web, WearSkill implementează metode de intrare personalizabile şi interschimbabile pentru persoane cu diverse dizabilități motorii folosind dispozitive purtabile, *e.g.*, gesturi tactile executate pe un inel smart sau gesturi executate în aer folosind un ceas smart (mai multe detalii în studiile realizate de Schipor *et al.* [49, 50]). Aprofundarea stadiului actual referitor la interacțiunile gestuale cu inele smart a avut loc prin realizarea unei analize sistematice a literaturii (SLR) din care au fost extrase informații relevante pentru proiectarea interfețelor inel smart-utilizator [57]. Informațiile extrase au oferit posibilitatea implementării instrumentului web, GestuRING [10, 57] care oferă următoarele funcționalități: 579 de mapări între gesturi şi funcțiile sistemului pe care îl controlează, descrierile gesturilor în format `JSON`, o bibliotecă YouTube cu înregistrări video ale gesturilor, date numerice şi reprezentările grafice ale gesturilor care au fost colectate cu un accelerometru cu 3 axe. În literatura ştiințifică s-au propus diverse metode de intrare pentru sistemele din vehiculele smart care ar putea avea consecințe pozitive asupra siguranței la volan; mai multe detalii în studiul realizat de Bilius şi Vatavu [9]. În ceea ce priveşte comunicarea dintre utilizatori şi funcțiile vehiculelor, experimentul de solicitare a preferințelor şoferilor şi pasagerilor (mai multe detalii în lucrarea [30]) a arătat o diversitate de gesturi. Deşi interacțiunile din interiorul vehiculelor sunt tot mai studiate, introducerea cadrului conceptual intitulat "smart vehicle proxemics" are rolul de a structura spațiul fizic din jurul unui vehicul în analogie cu zonele lui Hall [20] ale spațiului personal [11, 12]. Deoarece toate comunicațiile cu vehiculul sunt gestionate prin protocoale web şi formate de date standardizate, extinderea aplicațiilor către noi dispozitive, platforme, inclusiv metode de intrare

ce implică obținerea de noi funcționalități sunt ușor accesibile. Rezultatele obținute în această secțiune au fost publicate în cadrul lucrărilor [10–12, 30, 57].

7.3 Analiza sistematică a literaturii de specialitate

Având în vedere diversitatea metodelor și algoritmilor de recunoaștere a gesturilor, în această secțiune este decrisă o analiză a literaturii științifice în domeniu (52 lucrări). Conform informațiilor identificate, cei mai mulți autori au utilizat accelerometrul deoarece permite flexibilitatea de a putea fi atașat direct pe mâna utilizatorului oferind astfel rezultate precise. Dimensiunile seturilor de date raportate în lucrările studiate sunt neomogene, $\mu = 2799.45$ gesturi, SD = 3833.29 gesturi. Cele mai mari seturi de date conțin peste zece mii de gesturi, acestea fiind considerate date de dimensiuni mari. În acest caz, efortul de calcul și timpul necesar execuției prelucrărilor algoritmilor reprezintă un aspect foarte important. Dicționarele de gesturi (μ = 16.46, SD = 18.24) cuprind un număr variabil de gesturi, de la un gest la 130 gesturi. Statisticile extrase în urma acestei analize a literaturii științifice au indicat faptul că descompunerile multiliniare algebrice tensoriale și\sau metodele ansamblu de învățare nu au fost utilizate în lucrările studiate pentru recunoașterea gesturilor executate cu mâna în aer.

7.4 Metodă de recunoaștere a gesturilor executate cu mâna în aer folosind descompunerea Tucker2 și metoda ansamblu de învățare

Odată cu dezvoltarea noilor tehnologii, datele achiziționate de la diverși senzori cresc rapid, atât în volum, cât și în complexitate, în consecință, reducerea dimensiunii datelor și extragerea caracteristicilor fără a pierde informații relevante ar putea conduce la interpretarea lor într-un timp de răspuns interactiv. Având în vedere studiile explorative prezentate în Secțiunea 7.2, gesturile prezintă interes utilizatorilor pentru interacțiuni cu inele și vehicule smart. Întrucât în analiza sistematică a literaturii prezentată în Secțiunea 7.3 nu s-a identificat utilizarea descompunerilor tensoriale în problemele de recunoaștere a gesturilor colectate de senzori interțiali, această secțiune introduce o metodă, DELMet bazată pe descompunerea Tucker2 utilizată pentru reducerea dimensiunii datelor și extragerea caracteristicilor și o metodă ansamblu de învățare (ELMet) utilizată pentru a crește precizia clasificării; vezi Figura 7.1.

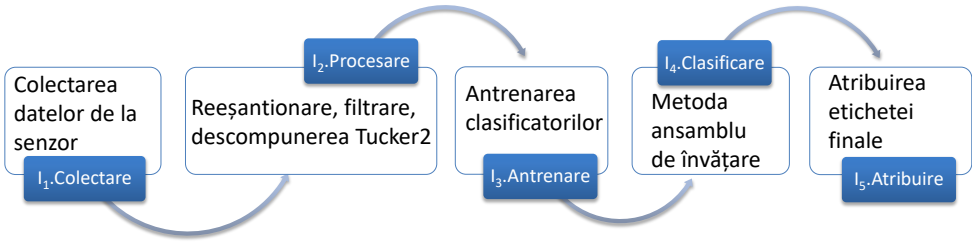


Figura 7.1: Fluxul procesului metodei DELMet.

7.5 Colectarea și procesarea datelor corespunzătoare gesturilor executate cu mâna în aer

În vederea evaluării metodei DELMet au fost utilizate patru seturi de date; două seturi de date proprii și două seturi de date publice [15][4]. Seturile de date S_3 și S_4 au fost evaluate în

[4] https://web.cs.wpi.edu/~claypool/mmsys-dataset/2012/6dmg/

acest studiu deoarece dicționarul de gesturi acoperă o mare parte din gesturile colectate în cadrul studiului prezentat în Secțiunea 7.2. Datele brute colectate de la senzori sunt caracterizate de poziție, orientare, accelerație și viteza unghiulară (un total de treisprezece măsurători diferite). Dicționarele de gesturi conțin douăzeci de gesturi diferite pentru ambele seturi de date și au fost colectate de la 28 de participanți (setul S_3 conține gesturile executate de 7 participanți cu mâna stângă și setul S_4 conține gesturile executate de 21 participanți cu mâna dreaptă); vezi Figura 7.2. În cazul seturilor de date proprii (S_1 și S_2), gesturile au fost descrise de valorile colectate de la accelerometru și giroscop cu o frecvență de 30Hz.

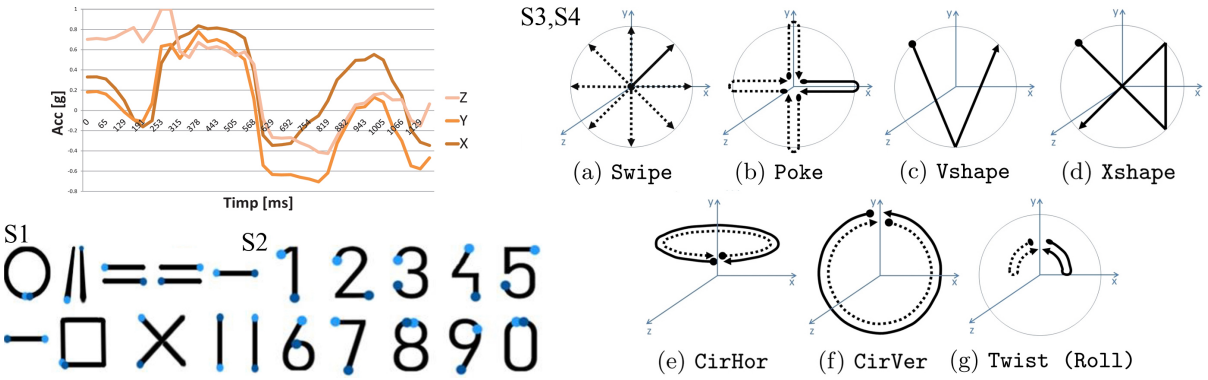
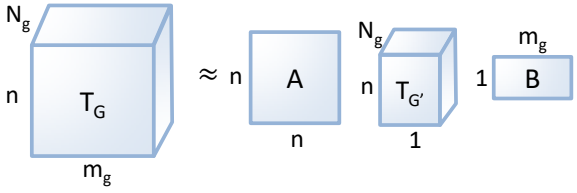


Figura 7.2: Dicționarele de gesturi ale seturilor S_1, S_2, S_3 și S_4 (imaginea corespunzătoare seturilor S_3 și S_4 este preluată din lucrarea [15]) și reprezentarea grafică a datelor brute corespunzătoare cifrei opt primite de la un accelerometru (stânga sus).

7.6 Construirea tensorului de gesturi și aplicarea descompunerii tensoriale Tucker2

Utilizarea descompunerii Tucker2 [27] presupune construirea tensorului de gesturi. Pentru a putea construi tensorul de gesturi, datele brute care descriu gesturile primite de la senzori trebuie prelucrate în prealabil pentru a avea aceeași lungime. În acest caz, metoda propusă presupune calcularea mediei m_g a tuturor lungimilor gesturilor folosite în procesul de antrenare. Reeșantionarea datelor este utilizată pentru ca toate datele primite de la senzori să aibă aceiași lungime m_g [35]. Ulterior a fost aplicat filtrul moving average pentru a netezi semnalele și elimina informațiile nedorite (zgomotele). În urma acestui pas toate gesturile G_i au aceiași dimensiune $n \times m_g$ (n reprezintă numărul de măsurători primite de la senzori la momentul t), în concluzie, tensorul de gesturi $T_G \in \mathbb{R}^{n \times m_g \times N_g}$ este construit cu succes (N_g reprezintă numărul de gesturi din setul de antrenare). Următorul pas constă în aplicarea descompunerii Tucker2 conform Ecuației 2.3 având datele de intrare obținute anterior. După cum se poate vedea în figura de mai jos, tensorul de gesturi T_G este descompus într-un tensor central $T_{G'} \in \mathbb{R}^{n \times 1 \times N_g}$ și matricile factor $A \in \mathbb{R}^{n \times n}$ și $B \in \mathbb{R}^{1 \times m_g}$. Matricea factor corespunzătoare dimensiunii N_g este setată să fie matricea identitate, celelalte două dimensiuni fiind descompuse în n și 1 componente.



Tensorul central $T_{G'} \in \mathbb{R}^{n \times 1 \times N_g}$ este de fapt o matrice, $M_G \in \mathbb{R}^{n \times N_g}$ care va fi folosită ca și intrare pentru ELMet (metoda ansamblu de învățare). Fiecare gest constă într-un vector de lungime n care re-

prezintă numărul de componente pentru Tucker2. Reducerea dimensionalității este necesară pentru a permite o performanță crescută de învățare, mai ales din punctul de vedere al costului de calcul. Pe de altă parte, este necesară extragerea caracteristicilor şi aducerea tututor gesturilor într-un spațiu unic de reprezentare.

7.7 Metoda ansablu de învățare a gesturilor executate cu mâna în aer

Metodele ansamblu de învățare sunt tehnici promițătoare deoarece utilizează algoritmi statistici şi computaționali de învățare pentru a lua decizia finală [44]. Scopul acestor metode este de a colecta mai multe etichete de la mai mulți algoritmi şi de a o extrage pe cea mai reprezentativă. Metoda ELMet este un pas important în cadrul metodei DELMet şi presupune atribuirea unui gest eticheta cu cea mai mare frecvență. Provocarea constă în alegerea etichetei corecte dintre etichetele prezise de predictori, chiar şi în situații în care două sau mai multe clase au acelaşi număr de voturi sau când fiecare clasă are un singur vot. În caz de egalitate, eroarea de clasificare greşită a validării încrucişate pe k-subseturi egale ale datelor (*en.*: k-fold cross-validation (CV))[5] a datelor de antrenare este calculată pentru fiecare algoritm de învățare care face parte din comitetul de vot şi care corespund frecvenței maxime din vectorul de etichete. Dintre valorile calculate se va alege eticheta care corespunde celei mai mici medii a erorilor de clasificare. Membrii metodei ELMet pot fi aleşi de utilizator sau se poate stabili un prag pentru acuratețea validărilor încrucişate calculate.

În ceea ce priveşte setul de testare, gesturile sunt procesate în acelaşi mod ca şi datele de antrenare (reeşantionare şi filtrul moving average). Aceste date sunt înmulțite cu matricile factor obținute din descompunerea Tucker2. Cu alte cuvinte, fiecare gest este transformat liniar prin înmulțirea cu matricile factor, astfel este adus în acelaşi spațiu ca şi datele de antrenare cu un număr mai mic de dimensiuni (vezi Ecuația 2.3).

7.8 Evaluarea performanței metodei DELMet

Acuratețea validării încrucişate (*en.*: leave-one-out cross-validation (LOOCV)) [48] a seturilor de date S_1, S_2, S_3 şi S_4 a fost calculată pentru o multitudine de algoritmi de învățare folosind diverşi parametri cu scopul de a alege membrii comitetului de vot din cadrul metodei ELMet. Algoritmii care au obținut valori ale acurateții mai mari de 85% au fost aleşi membri în comitetul de vot al metodei DELMet, mai exact, 1NN, 1NN, SVM, Label Propagation, 5NN şi LDA. În urma analizării rezultatelor, faptul că unii algoritmi de învățare au etichetat corect unele gesturi, dar alții nu şi vice-versa, aplicarea metodei ansamblu de învățare a extras cele mai relevante etichete oferind rezultate foarte bune; mai multe detalii în Tabelul 7.1.

Pentru a evalua eficiența metodei DELMet, au fost testate alte două metode, mai exact, clasificarea datelor reeşantionate şi filtrate (ELMet-Scls) şi clasificarea caracteristicilor obținute în urma aplicării descompunerii PCA (ELMet-PCA); vezi Tabelul 7.1. Acest studiu presupune realizarea unui studiu comparativ a rezultatelor obținute după utilizarea descompunerii Tucker2, dar şi fără utilizarea acesteia. De asemenea, au fost comparate rezultatele DELMet utilizând o altă metodă de reducere a dimensionalității, PCA. Pentru toate cele trei metode comitetul de vot a fost acelaşi, singura diferență majoră a constat în datele de intrare folosite pentru metoda ELMet. O observație importantă constă în faptul că metoda DELMet este alcătuită atât din

[5] `https://machinelearningmastery.com/k-fold-cross-validation/`

Tabelul 7.1: Detalii despre seturile de date, timpii de execuție [s] şi acuratețea metodelor testate.

Setul de date	$\mathbf{S}_1$	$\mathbf{S}_2$	$\mathbf{S}_3$	$\mathbf{S}_4$
Numărul de gesturi (antrenare/testare)	100/40	100/50	1304/100	4111/100
Lungimea gesturilor (μ)	61	46	76	67
Tucker2 [s]	0.034	0.042	0.376	0.943
Antrenare DELMet [s]	0.757	0.823	1.973	2.787
Acurateţe DELMet	100%	100%	100%	100%
Predicţie [s]	0.065	0.036	0.280	0.431
PCA [s]	0.024	0.025	0.360	0.551
Antrenare ELMet-PCA [s]	0.430	0.634	2.122	3.038
Acurateţe ELMet-PCA	100%	94%	97%	95%
Predicţie [s]	0.069	0.103	0.393	0.415
Antrenare ELMet-SCls [s]	0.756	0.688	7.938	28.030
Acurateţe ELMet-SCls	100%	100%	100%	100%
Predicţie [s]	0.125	0.123	1.110	2.427

descompunerea Tucker2, cât şi din metoda ansamblu de învățare (ELMet), dar metoda ELMet a fost tratată separat în acest studiu pentru ELMet-PCA şi ELMet-SCls cu scopul de a evidenția importanța reducerii dimensionalității folosind descompunerile tensoriale.

O altă evaluare a performanțelor metodei DELMet presupune compararea rezultatelor obținute cu ELMet-PCA şi ELMet-SCls, dar şi cu factorizarea matriceală nenegativă (*en.*: Non-negative matrix factorization) —ELMet-NNMF. Setul de date evaluat conține 5615 gesturi colectata de la de 28 persoane (S_3 şi S_4) [15]; vezi Tabelul 7.1. În cadrul acestei evaluări, metodele PCA, Tucker şi NNMF au descompus datele în aşa fel încât datele de intrare pentru metoda ELMet să aibă aceiaşi dimensiune, *i.e.* 5615×13. Datele metodei ELMet-SCls au avut o dimensiune de 5615×897 după reeşantionare şi filtrare. Optimizarea hiperparametrilor a algoritmilor de învățare supravegheată din librăria Scikit-learn [17] a constat în căutarea lor folosind tehnica "grid search" utilizând 5-fold cross-validation pentru setul de date S_{3+4}. Membrii comitetului de vot au fost reprezentați de top-5 modele care au obținut cea mai bună performanță în urma optimizării hiperparametrilor, mai exact, SVM (ovr, kernel rbf), Extra Trees (50 estimatori, 2 samples split), SVM (ovr, kernel liniar), SVM (date standardizate, ovr, kernel liniar) şi Extra Trees (20 estimatori, 2 samples split). Tabelul 7.2 prezintă mediile acurateților în urma validării încrucişate "leave-one-subject-out" (LOSO). Conform Tabelului 7.2, DELMet a obținut rezultate mai bune comparativ cu ELMet-PCA şi ELMet-NNMF, drept urmare, caracteristicile extrase de descompunerea Tucker au fost mai potrivite[6]. Acuratețea LOSO-CV a fost mai mare în cazul metodei ELMet-SCls, dar timpul necesar execuției prelucrărilor algoritmilor pentru antrenare şi predicție este mai mare nefiind avantajoasă pentru sistemele informatice care necesită un timp de răspuns interactiv. Tabelul 7.2 conține valorile acurateților în urma validărilor încrucişate LOOCV. După cum se poate observa, atunci când modelul are cunoştințe despre un participant (independentă de utilizator), valoarea acurateții este mai mare. Metoda DELMet a obținut o valoare foarte apropiată de ELMet-SCls, deşi dimensiunea setului de date utilizat pentru antrenare este semnficativ mai mic. Figura 7.3 ilustrează rezultatele obținute în urma calculării acurateții a două metode diferite de validare încrucişată. Conform Figurii 7.3, dimensiunea setului de date şi numărul de participanți sunt direct proporționale, în consecință, odată ce setul de antrenare este mai voluminos, predicțiile gesturilor executate de noi participanți devin mai precise. În consecință, deşi acuratețea metodei ELMet-SCls este mai mare decât a metodei DELMet, trebuie să existe un compromis între acurateţe şi timpul necesar antrenării şi predicțiilor.

[6] `https://scikit-learn.org/stable/modules/cross_validation.html`

Tabelul 7.2: Mediile acurateților validărilor încrucișate "leave-one-subject-out" (LOSO-CV)/"leave-one-out" (LOO-CV). *Notă:* SSE a descompunerii Tucker = 38.39/35.21, suma valorilor proprii a descompunerii PCA = 88.5%/89.2%, rădăcina medie pătrată reziduală a factorizării NNMF = 0.89/0.91.

	Model	DELMet (μ)	ELMet-PCA (μ)	ELMet-NNMF (μ)	ELMet-SCls (μ)
LOSO	Descompunere [s]	1.144	0.963	105.464	—
	Învățare [s]	0.701	0.722	0.988	7.887
	Testare [s]	1.234	1.457	1.282	1.933
	ELMet	92.9%	89.9%	76.6%	96.2%
LOO	Descompunere [s]	0.877	0.321	124.34	—
	Învățare [s]	0.753	0.825	0.105	8.127
	Testare [s]	1.351	1.456	1.313	2.003
	ELMet	99.1%	97.7%	91.5%	99.6%

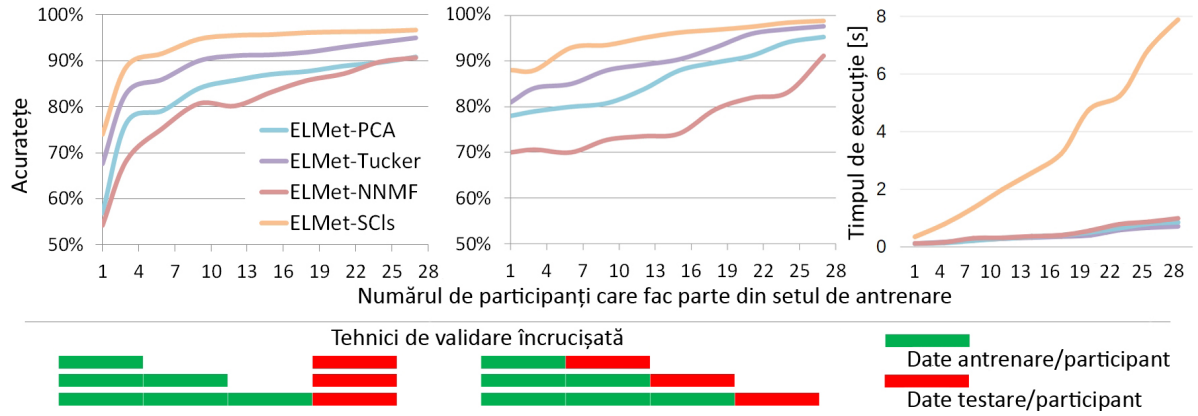


Figura 7.3: Acuratețea obținută în urma validărilor încrucișate și timpul [s] necesar antrenării algoritmilor de învățare în funcție de numărul de participanți a căror date fac parte din setul de antrenare.

7.9 Avantajele metodei propuse

Interacțiunile gestuale presupun colectarea de informații de la diverși senzori, procesarea și clasificarea lor. Având în vedere studiile explorative prezentate în Secțiunea 7.2 care au demonstrat că gesturile prezintă interes utilizatorilor pentru interacțiuni cu inele smart și interacțiuni în interiorul și în exteriorul vehiculelor, acest capitol a introdus o metodă de recunoaștere a gesturilor în mișcare executate cu mâna în aer, DELMet bazată pe descompunerea Tucker2 utilizată pentru reducerea dimensiunii datelor și extragerea caracteristicilor și o metodă ansamblu de învățare (ELMet) utilizată pentru a crește precizia clasificării. Descompunerea a oferit o performanță crescută de învățare, în același timp reducând timpul costisitor necesar antrenării și predicției, ceea ce o face potrivită pentru clasificarea datelor de dimensiuni mari. Pe de altă parte, extragerea caracteristicilor permite reprezentarea într-un spatiu unic a tuturor gesturilor. Aplicarea descompunerii Tucker2 pentru tensorii de gesturi a condus la reducerea semnificativă a dimensiunii datelor utilizate pentru predicție, în medie aceste date reprezintă 1.56% din întreg setul de date. Caracteristicile obținute în urma utilizării descompunerii sunt aplicate metodei ansamblu de învățare propuse, ELMet ce presupune crearea unui comitet de vot care va lua decizia finală de clasificare. Ținând cont de rezultatele obținute utilizând metoda propusă, DELMet oferă avantaje valoroase prin reducerea vizibilă a costului de calcul atunci când datele de antrenare sunt mari, păstrând în același timp acuratețea clasificării foarte bună datorită metodei ansamblu de învățare.

Contribuțiile prezentate în cadrul acestui capitol au fost publicate în următoarele lucrări științifice A_3, A_4, A_5, A_{10} și A_{11}; vezi Capitolul 8.1.3.

8 Concluzii şi contribuții

Contribuțiile acestei lucrări constau în utilizarea descompunerilor multiliniare algebrice tensoriale în recunoaşterea formelor. Descompunerile tensoriale sunt utilizate în recunoaşterea formelor datorită capacității de a extrage caracteristicile relevante şi de a reduce dimensiunile mari ale datelor. Aplicațiile avute în vedere se pot clasifica în două categorii, utilizarea descompunerilor tensoriale pentru imaginile hiperspectrale cu scopul de a extrage caracteristicile relevante şi de a reduce costul de calcul ridicat şi utilizarea descompunerilor tensoriale pentru interfețe gestuale care necesită un timp de răspuns interactiv. Imaginile hiperspectrale sunt date de dimensiuni mari, deci algoritmii de clasificare necesită mult timp pentru a le analiza şi clasifica. În acest scop, dezvoltarea unei metode de reducere a dimensionalității datelor şi eficientizare a rezultatului clasificării utilizând alături de descompuneri tensoriale, diagramele Voronoi şi formele tari au condus la rezultate eficiente [8]. De asemenea, au fost propuse metode ce au la bază segmentarea matricilor de abundență (*en.*: abundance matrix [43]) utilizând descompuneri tensoriale [5, 31]. În al doilea rând, datele produse şi transmise de dispozitivele care încorporează diverşi senzori tind să crească de la o zi la alta şi să ofere un volum semnificativ de informații valoroase. Această creştere este adesea considerată o provocare în studiul datelor de dimensiuni mari. În acest scop, dezvoltarea unei metode denumită DELMet de recunoaştere a gesturilor executate cu mâna în aer colectate de senzori inerțiali aplicând descompunerile multiliniare algebrice tensoriale şi o metodă ansamblu de învățare a condus la obținerea de rezultate eficiente într-un timp de răspuns interactiv. Aplicabilitatea metodei DELMet este promițătoare, fiind evidențiată în cadrul studiilor explorative privind interacțiuni cu inele smart [10, 57] şi interacțiuni în interiorul [30] şi exteriorul vehiculelor [11, 12].

8.1 Contribuțiile lucrării de doctorat

8.1.1 Contribuții teoretice

- T_1 **Specificarea şi evaluarea unor metode de identificare a nodurilor Voronoi în matricea de abundență folosind descompunerea Parafac (Capitolul 6.4).** Pentru a putea utiliza cât mai eficient resursele de calcul implicate, reducerea numărului de forme se obține prin aplicarea unei metode din cele propuse în Secțiunea 6.4 de identificare a celor mai reprezentative noduri Voronoi din matricea de abundență. Datorită legăturii dintre descompunerile multiliniare algebrice tensoriale şi modelele spectral unmixing, matricea de abundență a fost construită în urma descompunerii Parafac.
- T_2 **Propunerea unei metode euristice pentru determinarea partiției consens (Capitolul 6.5).** Identificarea partiției consens corespunzătoare multi-partiției obținute este o problemă NP dificilă, prin urmare, s-a utilizat o euristică care presupune gruparea ierarhică a formelor tari.
- T_3 **Propunerea unei metode eficiente, folosind descompuneri tensoriale, diagrame Voronoi şi forme tari pentru analiza imaginilor spectrale (Capitolul 6).** Metoda propusă constă în reducerea numărului de forme utilizând diagramele Voronoi şi obținerea partiției consens a nodurilor Voronoi care reprezintă întreaga imagine hiperspectrală.
- T_4 **Propunerea metodei DELMet de recunoaştere a gesturilor executate cu mâna în aer aplicând descompunerile multiliniare algebrice tensoriale şi o metodă ansamblu de**

învățare (Capitolul 7.4). A fost dezvoltată o metodă bazată pe descompunerea algebrică multiliniară a tensorilor pentru reducerea dimensionalității şi a extragerii caracteristicilor şi o metodă ansamblu de învățare pentru a îmbunătăți acuratețea clasificării.

T_5 **Analiza oportunității aplicative a interfețelor gestuale folosind inele smart (Capitolul 7.2).** Sunt descrise şi discutate studii explorative cu privire la gesturile în mişcare executate cu mâna în aer, astfel evidențiidu-se aplicabilitatea metodei de recunoaştere a gesturilor, DELMet dezvoltată în Capitolul 7.4.

8.1.2 Contribuții practice

C_1 **Evaluarea metodelor de segmentare a matricilor de abundență corespunzătoare imaginilor hiperspectrale (Capitolul 5.3)**. Metodele dezvoltate presupun segmentarea matricii de abundență având la bază algoritmul nucleelor dinamice şi segmentarea în subintervale egale. În urma testelor s-au obținut segmentări bine definite, atât pentru metoda de segmentare bazată pe algoritmul nucleelor dinamice, cât şi pentru metoda de segmentare în subintervale egale.

C_2 **Cercetări experimentale privind clasificarea supravegheată şi nesupravegheată a imaginilor hiperspectrale (Capitolul 5.1 şi Capitolul 5.2).** S-a studiat şi analizat performanța mai multor algoritmi de învățare automată aplicați imaginilor hiperspectrale satelitare de dimensiuni şi rezoluții diferite.

C_3 **Cercetări experimentale privind clusterizarea matricilor de abundență (Capitolul 5.2).** Clasificarea imaginilor hiperspectrale folosind descompunerile tensoriale şi algoritmi nesupravegheați au oferit o reducere semnificativă a dimensiunii datelor şi o analiză uşoară a materialelor prezente în zona de interes.

C_4 **Cercetări experimentale privind clusterizarea imaginior hiperspectrale folosind descompuneri tensoriale, diagrame Voronoi şi forme tari (Capitolul 6).** S-a realizat o analiză a performanțelor metodei dezvoltate pentru metodele de identificare a nodurilor Voronoi propuse. În urma testelor de performanță s-a observat că numărul de forme a fost redus semnificativ, astfel costul de calcul a fost mult mai mic, iar partiția consens obținută a crescut calitatea clusterizării.

C_5 **Un studiu de solicitare în urma căruia am colectat un set de gesturi de la 40 participanți (Capitolul 7.2).** Studiul a fost realizat cu scopul de a identifica preferințele şoferilor şi pasagerilor cu privire la interacțiuni gestuale în interiorul vehiculelor.

C_6 **Proiectarea şi dezvoltarea unui instrument web, GestuRING (Capitolul 7.2).** GestuRING pune la dispoziție informații şi resurse uşor accesibile care au scopul de a informa deciziile de proiectare cu privire la ce gesturi executate cu inele smart trebuie utilizate pentru a controla o funcție a unui sistem interactiv sau pentru a inspira noi modele de gesturi. Instrumentul web se bazează pe un corp de cunoştințe structurate sistematic extrase în urma unei analize a literaturii ştiințifice de specialitate cu privire la gesturi executate cu inele smart.

C_7 **Cercetări experimentale privind clasificarea supravegheată gesturilor executate cu mâna în aer (Capitolul 7).** S-a realizat o comparație a metodei DELMet cu alte metode de clasificare cu perspective diferite şi s-au analizat performanțele obținute. Conform rezultatelor obținute în urma testelor efectuate, costul de calcul implicat a fost redus semnificativ păstrând în acelaşi timp o acuratețe crescută a clasificării.

8.1.3 Diseminarea rezultatelor

Contribuțiile prezentate în cadrul acestei teze de doctorat au fost diseminate la nivel național şi internațional, în cadrul revistelor şi conferințelor de specialitate într-un număr de **17** articole ştiințifice, dintre care **9** sunt indexate Web of Science, după cum urmează:

a) Publicarea a **4** articole în reviste cotate ISI, dintre care:

- 1 articol Q1 Quartile (zona roşie)
- 2 articole Q2 Quartile (zona galbenă)

b) Publicarea a **10** articole la conferințe internaționale a căror volume sunt indexate ISI, Scopus, IEEE sau ACM, dintre care:

- 2 articole prezentate la conferințe de rang ARC CORE A*
- 2 articole prezentate la conferințe de rang ARC CORE B

c) Publicarea a **3** articole în reviste indexate BDI.

Lista lucrărilor publicate

A_1 **Laura-Bianca Bilius** şi Ştefan Gheorghe Pentiuc, **2020**. "*Efficient Unsupervised Classification of Hyperspectral Images Using Voronoi Diagrams and Strong Patterns*". Sensors 20, no. 19: 5684, **WOS:000587236000001**, **IF(2020): 3.576, 5-YEAR IF(2020): 3.735, Q1 Quartile (zona roşie)**, DOI: `https://doi.org/10.3390/s20195684`.

A_2 **Laura-Bianca Bilius** şi Ştefan Gheorghe Pentiuc, **2020**. "*Unsupervised Clustering for Hyperspectral Images*". Symmetry 12, no. 2: 277, **WOS:000521147600138**, **IF(2020): 2.713, 5-YEAR IF(2020): 2.612, Q2 Quartile (zona galbenă)**, DOI: `https://doi.org/10.3390/sym12020277`.

A_3 **Laura-Bianca Bilius** şi Radu-Daniel Vatavu, **2021**. "*A multistudy investigation of drivers and passengers' gesture and voice input preferences for in-vehicle interactions*", Journal of Intelligent Transportation Systems, 25:2, 197-220, **WOS:000590370400001**, **IF(2020): 4.277, 5-YEAR IF(2020): 4.134, Q2 Quartile (zona galbenă)**, DOI: `https://doi.org/10.1080/15472450.2020.1846127`.

A_4 **Laura-Bianca Bilius**, Radu-Daniel Vatavu şi Nicolai Marquardt, **2021**. "*Smart Vehicle Proxemics: A Conceptual Framework Operationalizing Proxemics in the Context of Outside-the-vehicle Interactions*", În: Ardito C. *et al.* (eds) Human-Computer Interaction – INTERACT 2021. INTERACT 2021. Lecture Notes in Computer Science, vol 12933. Springer, Cham, **WOS:000697565900011**, **Conferință de rang ARC CORE B**, DOI: `http://dx.doi.org/10.1007/978-3-030-85616-8_11`

A_5 **Laura-Bianca Bilius**, Radu-Daniel Vatavu şi Nicolai Marquardt, **2021**. "*Exploring Application Opportunities for Smart Vehicles in the Continuous Interaction Space Inside and Outside the Vehicle*", În: Ardito C. *et al.* (eds) Human-Computer Interaction – INTERACT 2021. Lecture Notes in Computer Science, vol 12933. Springer, Cham, **WOS: 000697565900010**, **Conferință de rang ARC CORE B**, DOI: `http://dx.doi.org/10.1007/978-3-030-85616-8_10`

A_6 **Laura-Bianca Bilius**, **2020**. "*A Smartwatch-based User Interface for In-Vehicle Interactions*", 2020 International Conference on Development and Application Systems

(DAS), Suceava, Romania, 2020, pp. 169-172, **WOS:000589776100031**, DOI: https://doi.org/10.1109/DAS49615.2020.9108911.

A_7 **Laura Bianca Bilius** şi Ştefan Gheorghe Pentiuc, **2020**. "*Improving the Analysis of Hyperspectral Images Using Tensor Decomposition*", 2020 International Conference on Development and Application Systems (DAS), Suceava, Romania, pp. 173-176, **WOS:000589776100032**,
DOI: https://doi.org/10.1109/DAS49615.2020.9108935

A_8 **Laura Bianca Bilius**, Ştefan Gheorghe Pentiuc, David Brie şi Sebastian Miron, **2019**. "*Analysis of Hyperspectral Images Using Supervised Learning Techniques*", 2019 23rd International Conference on System Theory, Control and Computing (ICSTCC), Sinaia, Romania, 2019, pp. 675-680, **WOS:000590181100114**, DOI: https://doi.org/10.1109/ICSTCC.2019.8885627.

A_9 **Laura Bianca Bilius**, **2018**. "*The User Behavior Analysis Based on Text Messages Using Parafac and Block Term Decomposition*". International Journal of Advanced Computer Science and Applications (IJACSA), 9(10), **WOS:000454189600007**, DOI: http://dx.doi.org/10.14569/IJACSA.2018.091007.

A_{10} Radu-Daniel Vatavu şi **Laura-Bianca Bilius**, **2021**. "*GestuRING: A Web-based Tool for Designing Gesture Input with Rings, Ring-Like, and Ring-Ready Devices*", Proceedings of UIST '21, the 34th Annual ACM Symposium on User Interface Software and Technology (Virtual Event), USA: ACM, **Conferință de rang ARC CORE A***, DOI: https://doi.org/10.1145/3472749.3474780.

A_{11} **Laura-Bianca Bilius** şi Radu-Daniel Vatavu, **2021**. "*Demonstration of GestuRING, a Web Tool for Ring Gesture Input*". The Adjunct Publication of UIST '21, the 34th Annual ACM Symposium on User Interface Software and Technology. ACM, New York, NY, USA, 2 pages, **Conferință de rang ARC CORE A***, DOI: https://doi.org/10.1145/3474349.3480199

A_{12} **Laura-Bianca Bilius** şi Radu-Daniel Vatavu, **2020**. "*A Synopsis of Input Modalities for In-Vehicle Infotainment and Consumption of Interactive Media*". In ACM International Conference on Interactive Media Experiences (IMX '20). Association for Computing Machinery, NY, USA, 195–199, DOI: https://doi.org/10.1145/3391614.3399400.

A_{13} **Laura Bianca Bilius** şi Ştefan Gheorghe Pentiuc, **2020**. "*Segmentation Method for Hyperspectral Images using Tensor Decomposition*", Journal of applied computer science & mathematics, Issue 2, (Vol. 14) / 2020, DOI: https://doi.org/10.4316/JACSM.202002001.

A_{14} **Laura Bianca Bilius** şi Ştefan Gheorghe Pentiuc, **2018**. "*Image Restoration using Voronoi Diagrams*", Journal of applied computer science & mathematics, Issue 2, Vol. 12, DOI: https://doi.org/10.4316/JACSM.201802003.

A_{15} Ştefan Gheorghe Pentiuc şi **Laura Bianca Bilius**, **2021**. "*The Analysis of Students Opinion About the Teachers Skills using Supervised Learning*", Journal of Social Sciences, Volume V, no. 1, 2022. DOI: https://doi.org/10.52326/jss.utm.2022.5(1).01

A_{16} Ovidiu-Andrei Schipor, **Laura-Bianca Bilius** şi Radu-Daniel Vatavu, **2022**. "*WearSkill: personalized and interchangeable input with wearables for users with motor impairments*", Proceedings of the 19th International Web for All Conference (W4A '22). Association for Computing Machinery, New York, NY, USA, Article 10, 1–5.

DOI: https://doi.org/10.1145/3493612.3520455

A_{17} Ovidiu-Andrei Schipor, **Laura-Bianca Bilius**, Ovidiu-Ciprian Ungurean, Alexandru-Ionuț Şiean, Alexandru-Tudor Andrei şi Radu-Daniel Vatavu, (2022). "*Personalized wearable interactions with WearSkill*", Proceedings of the 19th International Web for All Conference (W4A '22). Association for Computing Machinery, USA, Article 8, 1–2, **ACCESSIBILITY CHALLENGE JUDGES' AWARD** şi **ACCESSIBILITY CHALLENGE DELEGATES' AWARD**, DOI: https://doi.org/10.1145/3493612.3520474

8.1.4 Implicarea în proiecte de cercetare

$\mathbf{P}_1$ Proiect PNIII, *"Platformă hibridă de comunicații prin lumină vizibilă şi realitate augmentată pentru dezvoltarea de sisteme inteligente de asistență şi siguranță activă a autovehiculelor"*, cod proiect PN-III-P1-1.2-PCCDI-2017-0917, contract nr. 21PCCDI/2018, perioada de derulare 01.2018—11.2022;

$\mathbf{P}_2$ Proiect POCU, "*Excelenta academica şi valori antreprenoriale - sistem de burse pentru asigurarea oportunitatilor de formare si dezvoltare a competentelor antreprenoriale ale doctoranzilor si postdoctoranzilor (ANTREPRENORDOC)*", cofinantat din Fondul Social European prin Programul Operational Capital Uman, 2014-2020, Contract nr. 36355/23.05.2019, POCU/380/6/13 - Cod SMIS: 123847, perioada de derulare 09.2019—08.2020;

$\mathbf{P}_3$ Proiect PNIII, *"WearSkill: Interacțiuni fluide cu dispozitive wearable smart pentru abilități motorii specifice"*, cod proiect PN-III-P2-2.1-PED-2019-0352, contract nr. 276PED/2020, perioada de derulare 08.2020—08.2022.

$\mathbf{P}_4$ Proiect *"Centru pentru transferul de cunoştinţe către întreprinderi din domeniul ICT – CENTRIC"*, codul proiectului POC-A1-A1.2.3-G-2015, Cod SMIS 2014+ 119722 (ID P_40_305), Contract de finanțare nr. 5/AXA 1/1.2.3/G/13.06.2018, Contract subsidiar nr. 15.875/04.09.2020/acronim Study AI/ IMM:MiCS Software SRL, perioada de derulare 11.2020 —11.2022;

8.2 Dezvoltări viitoare

Dezvoltările viitoare ale cercetărilor efectuate în domeniul recunoaşterii formelor utilizând descompuneri tensoriale au în vedere următoarele:

I Îmbunătățirea acurateții clasificării nodurilor Voronoi utilizând algoritmi de învățare profundă în scopul creşterii şanselor de a obține o partiție consens mai generală a nodurilor. În această direcție, pentru rezolvarea problemelor de identificare a materialelor din imaginile hiperspectrale se vor explora performanțele utilizării rețelelor neuronale, cum ar fi rețeaua celulară neliniară (Cellular Nonlinear Network [36]), Restricted Boltzmann Machine [18], Neural turing machines [33].

II O altă direcție de explorări viitoare constă în implementarea unui algoritm specializat de descompunere şi antrenare tensorială a datelor hiperspectrale utilizând algoritmi de învățare prin întărire (*en.* Reinforcement Learning) [47]. Aceste tehnici dețin avatanjul de a putea învăța prin încercare şi eroare folosindu-se de propriile acțiuni şi experiențe, urmărind creşterea acurateții predicțiilor [47].

III Construirea partiției consens formalizată ca o problemă a unui proces de optimizare care să urmărească maximizarea măsurii de similaritate a spectrelor dominante din fiecare clasă cu spectrele existente, aceste tehnici fiind integrate într-o bibliotecă pentru identificarea materialelor susceptibile să apară în zona reprezentată de imaginea hiperspectrală.

IV Aprofundarea descompunerilor tensoriale pentru recunoaşterea gesturilor ce presupun mişcarea mâinii utilizând date colectate de la senzori inerțiali prezentată în Secțiunea 7 prin implementarea unui model dedicat ce presupune extragerea caracteristicilor şi învățare automată explicativă (*en.* Explainable AI) [32] a datelor, modelele explicative având capacitatea de a interpreta rezultatele şi a creşte încrederea utilizatorilor. În ceea ce priveşte îmbunătățirea acurateții metodei ansamblu de învățare, se poate continua prin utilizarea algoritmilor generici de recunoaştere, cum ar fi $3 [28] şi algoritmilor de învățare prin întărire (*en.* reinforcement learning). Aceşti algoritmi dețin avantajul de a fi flexibili în rezolvarea problemelor de clasificare, caracteristicile fiind deduse automat şi ajustate pentru a obține rezultatul dorit.

V Dezvoltarea şi implementarea unei librării pentru recunoaşterea gesturilor colectate de diverşi sensori inerțiali având la bază metoda DELMet. În funcție de caracteristicile datelor, algoritmul foloseşte datele antrenate corespunzătoare, generalizând în acest fel libertatea utilizatorilor de a lucra cu senzorii doriți. Această direcție presupune crearea unei baze de date care să conțină gesturi atât bidimensionale, cât şi tridimensionale colectate de la utilizatori prin intermediul studiilor de solicitare în condiții reale de utilizare [58].

VI Creşterea acurateții clasificărilor ar putea fi realizată prin complementarea caracteristicilor extrase de decompunerile tensoriale cu statistici descriptive cantitative extrase din date. Această direcție poate fi implementată atât pentru recunoaşterea gesturilor, cât şi a materialelor prezente în imaginile hiperspectrale.

References

[1] Sahar A. El Rahman, *"Hyperspectral image classification using unsupervised algorithms"*, in International Journal of Advanced Computer Science and Applications, **7 (2016)**, https://doi.org/10.14569/IJACSA.2016.070425.

[2] Asad Abdi, *"Three types of machine learning algorithms"*, **(2016)**, https://doi.org/10.13140/RG.2.2.26209.10088.

[3] Nicolas Audebert, Bertrand Saux, and Sébastien Lefèvre, *"Deep learning for classification of hyperspectral data: A comparative review"*, in IEEE Geoscience and Remote Sensing Magazine, **7 (2019)**, https://doi.org/10.1109/MGRS.2019.2912563.

[4] Javier Bejar, *"Consensus clustering (course slides)"*, in String CS-MAI, **(2020)**, https://www.cs.upc.edu/~bejar/URL/material/06-Consensus.pdf.

[5] Laura-Bianca Bilius and Stefan Gheorghe Pentiuc, *"Improving the analysis of hyperspectral images using tensor decomposition"*, in 2020 International Conference on Development and Application Systems (DAS), **(2020)**, pp. 173-176, https://doi.org/10.1109/DAS49615.2020.9108935.

[6] Laura-Bianca Bilius and Stefan-Gheorghe Pentiuc, *"Unsupervised clustering for hyperspectral images"*, in Symmetry, **12(277) (2020)**, https://doi.org/10.3390/sym12020277.

[7] Laura-Bianca Bilius, Stefan-Gheorghe Pentiuc, David Brie, and Sebastian Miron, *"Analysis of hyperspectral images using supervised learning techniques"*, in 2019 23rd Int. Conf. on System Theory, Control and Computing (ICSTCC), **(2019)**, pp. 675-680, https://doi.org/10.1109/ICSTCC.2019.8885627.

[8] Laura Bianca Bilius and Ştefan Gheorghe Pentiuc, *"Efficient unsupervised classification of hyperspectral images using voronoi diagrams and strong patterns"*, Sensors, **20(19) (2020)**, ISSN 1424-8220, Article No. 5684, https://doi.org/10.3390/s20195684.

[9] Laura-Bianca Bilius and Radu-Daniel Vatavu, *"A synopsis of input modalities for in-vehicle infotainment and consumption of interactive media"*, in ACM International Conference on Interactive Media Experiences, IMX '20, New York, NY, USA, **(2020)**, Association for Computing Machinery, Cornella, Barcelona, Spain, pp. 195-199, ISBN 9781450379762, https://doi.org/10.1145/3391614.3399400.

[10] Laura-Bianca Bilius and Radu-Daniel Vatavu, *"Demonstration of gesturing, a web tool for ring gesture input"*, in The 34th Annual ACM Symposium on User Interface Software and Technology, New York, NY, USA, **(2021)**, Association for Computing Machinery, p. 124–125, No. pages ; 2 pages, ISBN 9781450386555, https://doi.org/10.1145/3474349.3480199.

[11] Laura-Bianca Bilius, Radu-Daniel Vatavu, and Nicolai Marquardt, *"Exploring application opportunities for smart vehicles in the continuous interaction space inside and outside the vehicle"*, in Human-Computer Interaction - INTERACT 2021 - 18th IFIP TC 13 International Conference, Bari, Italy, 2021, Proceedings, Part II, Lecture Notes in Computer Science, Springer,, **12933 (2021)**, pp. 140-149, https://doi.org/10.1007/978-3-030-85616-8_10.

[12] Laura-Bianca Bilius, Radu-Daniel Vatavu, and Nicolai Marquardt, *"Smart vehicle proxemics: A conceptual framework operationalizing proxemics in the context of outside-the-vehicle interactions"*, in Human-Computer Interaction - INTERACT 2021 - 18th IFIP TC 13 International Conference, Bari, Italy, 2021, Proceedings, Part II, Lecture Notes in Computer Science, Springer, **12933 (2021)**, pp. 150-171, https://doi.org/10.1007/978-3-030-85616-8_11.

[13] Bilius Laura-Bianca, Pentiuc Stefan-Gheorghe, *"Image restoration using voronoi diagrams"*, **12(2) (2018)**, pp. 22-24, https://doi.org/10.4316/JACSM.201802003.

[14] Gilles Celeux, Edwin Diday, Govaert Govaert, Yves Lechevallier, and Henri Ralambondrainy, *"Classification automatique des donnees"*, **(1989)**, ISBN-13 978-2040187989; ISBN-10 2040187987.

[15] Mingyu Chen, Ghassan AlRegib, and Biing-Hwang Juang, *"6dmg: A new 6d motion gesture database"*, in Proceedings of the 3rd Multimedia Systems Conference, MMSys '12, ACM, **(2012)**, pp. 83-88, ISBN 9781450311311, https://doi.org/10.1145/2155555.2155569.

[16] Edwin Diday, *"The dynamic clusters method in nonhierarchical clustering"*, in International Journal of Computer & Information Sciences, **2 (1973)**, pp. 61-88, https://doi.org/10.1007/BF00987153.

[17] Fabian Pedregosa, Gael Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot and Edouard Duchesnay, *"Scikit-learn: Machine learning in python"*, in Journal of Machine Learning Research, **12 (2011)**, pp. 2825-2830.

[18] Asja Fischer and Christian Igel, *"An introduction to restricted boltzmann machines"*, **(2012)**, pp. 14-36, https://doi.org/10.1007/978-3-642-33275-3_2.

[19] Vairaprakash Gurusamy, Subbu Kannan, and G.Nalini, *"Review on image segmentation techniques"*, **(2014)**.

[20] Edward T. Hall, *"The hidden dimension"*, Garden City, N.Y.: Doubleday, **(1966)**, ISBN 0385084765 9780385084765.

[21] Samantha Hansen, Todd Plantenga, and Tamara Kolda, *"Newton-based optimization for kullback-leibler nonnegative tensor factorizations"*, in Optimization Methods and Software, **30 (2013)**, https://doi.org/10.1080/10556788.2015.1009977.

[22] Lida Hosseini and Ramin Kandovan, *"Hyperspectral image classification based on hierarchical svm algorithm for improving overall accuracy"*, in Advances in Remote Sensing, **06 (2017)**, pp. 66-75, https://doi.org/10.4236/ars.2017.61005.

[23] K. Huang, S. Li, X. Kang, and L. Fang, *"Spectral–spatial hyperspectral image classification based on knn"*, in Sensing and Imaging, **17 (2015)**, https://doi.org/10.1007/s11220-015-0126-z.

[24] Tales Imbiriba, Ricardo Augusto Borsoi, and José Carlos Moreira Bermudez, *"Low-rank tensor modeling for hyperspectral unmixing accounting for spectral variability"*, in IEEE Transactions on Geoscience and Remote Sensing, **58 (2020)**, pp. 1833-1842, https://doi.org/10.1109/TGRS.2019.2949543.

[25] Jose Manuel Amigo, *"Hyperspectral imaging"*, 1st Edition, Elsevier, **32 (2019)**, ISBN 9780444639776, eBook ISBN 9780444639783.

[26] Nirmal Keshava and John F. Mustard, *"Spectral unmixing"*, in IEEE Signal Processing Magazine, **19 (2002)**, pp. 44-57, https://doi.org/10.1109/79.974727.

[27] Tamara G. Kolda and Brett W. Bader, *"Tensor decompositions and applications"*, in SIAM Review, **51 (2009)**, pp. 455-500, https://doi.org/10.1137/07070111X.

[28] Sven Kratz and Michael Rohs, *"A $3 gesture recognizer: Simple gesture recognition for devices equipped with 3d acceleration sensors"*, in Proceedings of the 15th International Conference on Intelligent User Interfaces, IUI '10, New York, NY, USA, **(2010)**, Association for Computing Machinery, p. 341–344, https://doi.org/10.1145/1719970.1720026.

[29] Ludmila Kuncheva and Chris Whitaker, *"Pattern recognition and classification"*, **(2015)**, pp. 1-7, ISBN 9781118445112, https://doi.org/10.1002/9781118445112.stat06503.pub2.

[30] Laura-Bianca Bilius and Radu-Daniel Vatavu, *"A multistudy investigation of drivers and passengers' gesture and voice input preferences for in-vehicle interactions"*, in Journal of Intelligent Transportation Systems, **25(2) (2021)**, pp. 197-220, https://doi.org/10.1080/15472450.2020.1846127.

[31] Laura-Bianca Bilius, Stefan Gheorghe Pentiuc, *"Segmentation method for hyperspectral images using tensor decomposition"*, in Journal of Applied Computer Science & Mathematics., **14(2) (2020)**, https://doi.org/10.4316/JACSM.202002001.

[32] Pantelis Linardatos, Vasilis Papastefanopoulos, and Sotiris Kotsiantis, *"Explainable ai: A review of machine learning interpretability methods"*, Entropy, **23 (2021)**, https://doi.org/10.3390/e23010018.

[33] Soroor Malekmohammadi Faradounbeh, Faramarz Safi, and Morteza Karimian-Kelishadrokhi, *"A review on neural turing machine (ntm)"*, SN Computer Science, **1 (2020)**, p. 333, https://doi.org/10.1007/s42979-020-00341-6.

[34] Mark de Berg, Otfried Cheong, Marc van Kreveld and Mark Overmars, *"Computational geometry, algorithms and applications"*, Computationa Geometry, **(2008)**, ISBN 978-3-540-77974-2, https://doi.org/10.1007/978-3-5-540-77974-2.

[35] MathWorks, *"MATLAB & Simulink R2018b"*, Available online, https://www.mathworks.com.

[36] P. Mazumder, S.-R Li, and Idongesit Ebong, *"Tunneling-based cellular nonlinear network architectures for image processing"*, Very Large Scale Integration (VLSI) Systems, IEEE Transactions on, **17 (2009)**, pp. 487 - 495, https://doi.org/10.1109/TVLSI.2009.2014771.

[37] E. Middleton, S. Ungar, D. Mandl, L. Ong, S. Frye, P. Campbell, D. Landis, J. Young, and N. Pollack, *"The earth observing one (eo-1) satellite mission: Over a decade in space"*, in IEEE Journal Of Selected Topics In Applied Earth Observations And Remote Sensing, **6 (2013)**, p. 243, https://doi.org/10.1109/JSTARS.2013.2249496.

[38] Dobigeon N., Y. Altmann, N. Brun, and Moussaoui S., *"Chapter 6 - linear and nonlinear unmixing in hyperspectral imaging"*, in Resolving Spectral Mixtures, Elsevier, **30 (2016)**, Data Handling in Science and Technology, Cyril Ruckebusch (ed), https://doi.org/10.1016/B978-0-444-63638-6.00006-1.

[39] NUMXL SUPPORT HISTBINS, *"Number of histogram bins"*, **(2016)**, https://support.numxl.com/hc/en-us/articles/216034163-HISTBINS-Number-of-Histogram-Bins.

[40] Ivan Oseledets, D.V. Savostianov, and E. Tyrtyshnikov, *"Tucker dimensionality reduction of three-dimensional arrays in linear time"*, in SIAM Journal on Matrix Analysis and Applications, **30 (2008)**, https://doi.org/10.1137/060655894.

[41] Anh-Huy Phan and Andrzej Cichocki, *"Tensor decompositions for feature extraction and classification of high dimensional datasets"*, in Nonlinear Theory and Its Applications, IEICE, **1 (2010)**, https://doi.org/10.1587/nolta.1.37.

[42] Joaquim Pinto da Costa and P. R. Rao, *"Central partition for a partition-distance and strong pattern graph"*, in REVSTAT, **2 (2004)**.

[43] Y. Qian, F. Xiong, S. Zeng, J. Zhou, and Y. Y. Tang, *"Matrix-vector nonnegative tensor factorization for blind unmixing of hyperspectral imagery"*, in IEEE Transactions on Geoscience and Remote Sensing, **55(3) (2017)**, pp. 1776-1792, ISSN 1558-0644, https://doi.org/10.1109/TGRS.2016.2633279.

[44] Matteo Re and Giorgio Valentini, *"Ensemble methods: A review"*, in Advances in Machine Learning and Data Mining for Astronomy, **(2012)**, pp. 563-594, http://homes.dsi.unimi.it/valentini/papers/ens.review.revised.pdf.

[45] Y Reddy, Viswanath Pulabaigari, and Eswara B, *"Semi-supervised learning: a brief review"*, International Journal of Engineering Technology, **7 (2018)**, p. 81, https://doi.org/10.14419/ijet.v7i1.8.9977.

[46] Youssra Riahi, *"Big data and big data analytics: Concepts, types and technologies"*, **5 (2018)**, pp. 524-528, https://doi.org/10.21276/ijre.2018.5.9.5.

[47] Richard S. Sutton and Andrew G. Barto, *"Reinforcement learning, second edition an introduction"*, in A Bradford Book, The MIT Press, Cambridge, Massachusetts, London, England, **(2014)**.

[48] Claude Sammut and Geoffrey I. Webb, *"Leave-one-out cross-validation"*, in Encyclopedia of Machine Learning, Springer US, Boston, MA, **(2010)**, pp. 600-601, ISBN 978-0-387-30164-8, https://doi.org/10.1007/978-0-387-30164-8_469.

[49] Ovidiu-Andrei Schipor, Laura-Bianca Bilius, Ovidiu-Ciprian Ungurean, Alexandru-Ionuț Șiean, Alexandru-Tudor Andrei, and Radu-Daniel Vatavu, *"Personalized wearable interactions with wearskill"*, in Proceedings of the 19th International Web for All Conference, W4A '22, New York, NY, USA, **(2022)**, Association for Computing Machinery, https://doi.org/10.1145/3493612.3520474.

[50] Ovidiu-Andrei Schipor, Laura-Bianca Bilius, and Radu-Daniel Vatavu, *"Wearskill: Personalized and interchangeable input with wearables for users with motor impairments"*, in Proceedings of the 19th International Web for All Conference, W4A '22, New York, NY, USA, **(2022)**, Association for Computing Machinery, https://doi.org/10.1145/3493612.3520455.

[51] Gabriela (Serban) Czibula and Florina Codrean, *"Modele de invatare nesupervizată"*, in Babes-Bolyai University of Cluj-Napoca, Faculty of Mathematics and Computer Science, **(2012)**.

[52] Y. Shao, J. Lan, Y. Zhang, and J. Zou, *"Spectral unmixing of hyperspectral remote sensing imagery via preserving the intrinsic structure invariant"*, in Sensors, **18 (2018)**, https://doi.org/10.3390/s18103528.

[53] Shefali Aggarwal, *"Principles of remote sensing, satellite remote sensing and gis applications in agricultural meteorology"*, in Proceedings of a Training Workshop, Dehra Dun, India, **(2003)**, pp. 22-38, eds. Sivakumar, M.V.K., Roy, P.S. ; Harmsen, K.; Saha S.K.; AGM-8, WMO/TD-No. 1182 Dun, https://library.wmo.int/doc_num.php?explnum_id=5030.

[54] Ying Tan, *"Chapter 11 - applications, gpu-based parallel implementation of swarm intelligence algorithms"*, Elsevier, Morgan Kaufmann, **(2016)**, pp. 167-177, ISBN 9780128093627, https://doi.org/10.1016/B978-0-12-809362-7.50011-X.

[55] Yuan Tang, Yang Lu, and Haoliang Yuan, *"Hyperspectral image classification based on three-dimensional scattering wavelet transform"*, in IEEE Transactions on Geoscience and Remote Sensing, **53 (2015)**, https://doi.org/10.1109/TGRS.2014.2360672.

[56] Radu-Daniel Vatavu, *"User-defined gestures for free-hand tv control"*, in Proceedings of the 10th European Conference on Interactive TV and Video, EuroITV '12, New York, NY, USA, **(2012)**, ACM, Berlin, Germany, pp. 45-48, ISBN 9781450311076, https://doi.org/10.1145/2325616.2325626.

[57] Radu-Daniel Vatavu and Laura-Bianca Bilius, *"Gesturing: A web-based tool for designing gesture input with rings, ring-like, and ring-ready devices"*, in The 34th Annual ACM Symposium on User Interface Software and Technology, New York, NY, USA, **(2021)**, Association for Computing Machinery, pp. 710-723, No. pages ; 14 pages, ISBN 9781450386357, https://doi.org/10.1145/3472749.3474780.

[58] Radu-Daniel Vatavu and Jacob O. Wobbrock, *"Between-subjects elicitation studies: Formalization and tool support"*, in Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, CHI '16, New York, NY, USA, **(2016)**, Association for Computing Machinery, p. 3390–3402, https://doi.org/10.1145/2858036.2858228.

[59] Miguel A. Veganzones, Jeremy E. Cohen, Rodrigo Cabral Farias, Jocelyn Chanussot, and Pierre Comon, *"Nonnegative tensor cp decomposition of hyperspectral data"*, in IEEE Transactions on Geoscience and Remote Sensing, **54 (2016)**, pp. 2577-2588, https://doi.org/10.1109/TGRS.2015.2503737.

[60] Alberto Villa, *"Advanced spectral unmixing and classification methods for hyperspectral remote sensing data"*, in Theses, Universite de Grenoble ; 102 Univ of Iceland, Reykjavik, **(2011)**, https://tel.archives-ouvertes.fr/tel-00767250.

[61] Sowmya Vishvanathan, Soman Kp, and M. Hassaballah, *"Hyperspectral image: Fundamentals and advances"*, in Recent Advances in Computer Vision. Studies in Computational Intelligence, Cham Springer, **804 (2019)**, pp. 401-424, ISBN 978-3-030-02999-9, https://doi.org/10.1007/978-3-030-03000-1_16.

[62] Zhongliang Wang and Hua Xiao, *"Distributed compressed hyperspectral sensing imaging based on spectral unmixing"*, in Sensors, **20 (2020)**, p. 2305, https://doi.org/10.3390/s20082305.

[63] J. O. Wobbrock, A. D. Wilson, and Y. Li, *"Gestures without libraries, toolkits or training: A $1 recognizer for user interface prototypes"*, in Proc. of the 20th Symp. on User Interface Software and Technology, UIST '07, ACM, **(2007)**, pp. 159-168, https://doi.org/10.1145/1294211.1294238.

[64] Guoxu Zhou, Zhaoshui He, Yu Zhang, Qibin Zhao, and Andrzej Cichocki, *"Canonical polyadic decomposition: From 3-way to n-way"*, in 2012 Eighth International Conference on Computational Intelligence and Security, **(2012)**, pp. 391-395, https://doi.org/10.1109/CIS.2012.94.