



UNIVERSITATEA „POLITEHNICA” din BUCUREȘTI

ȘCOALA DOCTORALĂ ETTI-B

Nr. Decizie din

TEZĂ DE DOCTORAT

**CONTRIBUTII LA STUDIUL SI APLICATIILE
MODELULUI PULSE-COUPLED NEURAL NETWORK
PENTRU ANALIZA IMAGINILOR**

**CONTRIBUTIONS TO THE STUDY AND
APPLICATIONS OF THE PULSE-COUPLED NEURAL
NETWORK MODEL FOR IMAGE ANALYSIS**

Doctorand: **Mat. Carată Șerban Vasile**

COMISIA DE DOCTORAT

Președinte	Prof. dr. ing. Gheorghe Brezeanu	de la	Univ. Politehnica București
Conducător de doctorat	Prof. dr. ing. Victor-Emil Neagoe	de la	Univ. Politehnica București
Referent	Col. Prof. dr. ing. Ioan Nicolaescu	de la	Academia Tehnica Militara
Referent	Prof. dr. ing. Alexandru Serbanescu	de la	Academia Tehnica Militara
Referent	Prof. dr. ing. Bogdan-Emanuel Ionescu	de la	Univ. Politehnica București

BUCUREȘTI 2016

Mulțumiri

Aceasta lucrare nu ar fi fost posibilă fără susținerea și ajutorul multor oameni.

În primul rând aș dori să îmi exprim gratitudinea sinceră mentorului meu, Domnul Profesor Victor Neagoe, pentru timpul, inspirația, ideile inovatoare și sfaturile pe care mi le-a acordat pe parcursul întregului stagiu doctoral.

Multe mulțumiri colegilor mei de la departamentul Softtrust, pentru sfaturile lor și sprijinul acordat.

Aș dori să mai mulțumesc familiei mele pentru înțelegerea și suportul acordat pe întreg parcursul studiilor doctorale.

Nu în ultimul rând aș dori să mulțumesc tuturor participanților ca subiecți la constituirea bazei de date necesare pentru experimentările din capitolul 6 ale acestei lucrări V.E. Neagoe, S.V. Carata, „Subject Independent Drunkenness Detection Using Pulse-Coupled Neural Network Segmentation of Thermal Infrared Facial Imagery”, *International Journal of Mathematical and Computational Methods*, pp. 305-312, August 2016, (BDI).

Teza a fost susținută financiar parțial prin proiectul ExcelDOC POSDRU/159/1.5/S/132397.

Cuprins

Mulțumiri	5
Lista tabelelor	11
Lista figurilor	12
Lista abrevierilor	14
CV	17
1. Introducere	20
1.1 Prezentarea domeniului tezei de doctorat	21
1.2 Scopul tezei de doctorat	21
1.3 Conținutul tezei de doctorat.....	21
2. Modelul PCNN.....	24
2.1 Componenta de alimentare	25
2.2 Componenta de conectare	26
2.3 Componenta aditivă.....	27
2.4 Componenta decizională	28
2.5 Neuronul PCNN	29
2.6 Exemplificarea funcționării modelului PCNN.....	30
3. PCNN pentru segmentarea imaginilor în tonuri de gri și a imaginilor color.....	32
3.1 Introducere	32
3.1.1 Segmentarea imaginilor cu algoritmi bazați pe contur activ	33
3.1.2 Segmentarea Imaginilor bazata pe separate și îmbinare	33
3.1.3 Segmentare a imaginilor bazată pe tăieturi normalizate	34
3.2 Algoritmul Mean Shift pentru segmentarea imaginilor	34
3.3 Algoritmul Fuzzy C-Means pentru segmentarea imaginilor	38
3.3.1 Etapele algoritmului Fuzzy C-Means	39
3.4 Rețeaua cu autoorganizare (Self-Organizing Map)	41
3.4.1 SOM pentru segmentarea imaginilor	43
3.5 Optimizarea parametrilor PCNN utilizând algoritmi genetici	44
3.6 PCNN pentru segmentarea imaginilor în tonuri de gri	47
3.7 PCNN pentru segmentarea imaginilor color	47

3.8	Rezultate Experimentale	48
3.8.1	Baza de Date "The Berkeley Segmentation Dataset and Benchmark"	48
3.8.2	Parametrii PCNN pentru segmentarea imaginilor în tonuri de gri.....	49
3.7.3	Parametrii PCNN pentru segmentarea imaginii color.....	49
3.8.4	Parametrii PCNN propus pentru segmentarea imaginilor color	50
3.8.5	Performanțe experimentale.....	50
4.	Recunoașterea plăcuțelor de înmatriculare utilizând modelul PCNN	65
4.1	PCNN pentru segmentarea imaginilor multispectrale	65
4.2	SVM pentru recunoașterea caracterelor	67
4.2.1	Hiperplanul de separare.....	68
4.3	Baza de date	70
4.4	Segmentarea plăcuțelor de înmatriculare	71
4.5	Antrenarea clasificatorului.....	74
4.6	Rezultate experimentale	74
5.	Recunoașterea vehiculelor militare în imagini SAR cu tehnici neuronale	76
5.1	Introducere	76
5.2	Algoritmul propus	77
5.3	Filtrarea Gabor	79
5.4	Analiza Componentelor Principale (PCA).....	81
5.5	SVM antrenat cu date virtuale generate de sistemul neuronal CSOM.....	83
5.6	Rezultate experimentale	84
5.6.1	MSTAR - baza de date a vehiculelor militare terestre	84
5.6.2	Performanțe experimentale.....	85
6.	Detecția stării de ebrietate utilizând modelul PCNN pentru segmentarea imaginilor faciale din spectrul termal.....	87
6.1	Necesitatea identificării stării de ebrietate a conducătorilor auto.....	87
6.2	Descrierea algoritmului propus pentru detecția stării de ebrietate	89
6.2.1	Segmentarea imaginilor cu modelul PCNN	89
6.2.2	Înlocuirea pixelilor activi din masca obținută în urma segmentării cu valorile acestora din imaginea originală.....	89
6.2.3	Extragerea de trăsături.....	90
6.2.4	Clasificarea SVM.....	93

6.3	Rezultate experimentale	93
6.3.1	Baza de date de imagini faciale termale pentru detecția stării de ebrietate	93
6.3.2	Evaluarea performanțelor	95
6.3.3	Performanțe experimentale.....	95
7.	Detecția schimbărilor in imagini multitemporale de observație terestra aplicând modelul PCNN	103
7.1	Metode nesupervizate pentru detecția schimbărilor în serii de imagini de observație terestra.....	103
7.1.1	Metode algebrice	104
7.1.2	Metode bazate pe transformări matematice	104
7.1.3	Metode bazate pe clasificare	104
7.1.4	Metode avansate	105
7.1.5	GIS	105
7.1.6	Analiză vizuală.....	105
7.1.7	Alte abordări	105
7.2	Analiza vectorilor de schimbare (CVA).....	106
7.3	Metoda PCNN - Pacifici pentru detectia schimbarilor in imagini satelitare multitemporale	107
7.4	Modelul PCNN propus pentru detecția schimbărilor în imagini satelitare multitemporale	108
7.5	Rezultate experimentale	109
7.5.1	Baza de date Mexico	109
7.5.2	Performante experimentale.....	111
8.	Concluzii.....	112
8.1	Rezultate obținute	114
8.1.1	Rezultate si concluzii referitoare la segmentarea imaginilor în tonuri de gri și a imaginilor color folosind modelul Pulse Coupled Neural-Network	114
8.1.2	Rezultate si concluzii referitoare la identificarea si recunoașterea plăcuțelor de înmatriculare.....	118
8.1.3	Rezultate si concluzii referitor la recunoașterea vehiculelor militare în imagini SAR	120
8.1.4	Rezultate si concluzii referitoare la recunoașterea starii de ebrietate a subiectilor ..	121
8.1.5	Rezultate si concluzii referitor la detecția schimbărilor în serii de imagini de observație terestra.....	125

8.2 Contribuții originale	126
8.2.1 Contribuții la segmentarea imaginilor in tonuri de gri si a imaginilor color (<i>Capitolul 3</i>):	126
8.2.2 Contribuții la recunoașterea plăcuțelor de înmatriculare (<i>Capitolul 4</i>):.....	127
8.2.3 Contribuții la identificarea și recunoașterea vehiculelor militare în imagini SAR (<i>Capitolul 5</i>):.....	127
8.2.4 Contribuții la diagnosticarea starii de ebrietate (<i>Capitolul 6</i>):	127
8.2.5 Contribuții la identificarea schimbărilor in serii temporale de imagini satelitare (<i>Capitolul 7</i>):.....	128
8.3 Lista lucrărilor originale.....	129
8.4 Perspective de dezvoltare ulterioară	130
Bibliografie	131

Lista tabelelor

Tabelul 2.1 Parametrii Interni PCNN	30
Tabelul 3.1 Matricele M și W	49
Tabelul 3.2 Parametrii PCNN pentru segmentarea imaginilor în tonuri de gri	49
Tabelul 3.3 Matricele M și W	50
Tabelul 3.4 Parametrii PCNN pentru segmentarea imaginilor color	50
Tabelul 3.5 Matricele M și W	50
Tabelul 3.6 Parametrii PCNN pentru segmentarea imaginilor color folosind modelul propus	50
Tabelul 3.7 Precizia Segmentării pentru datele de antrenare	51
Tabelul 3.8 Precizia Segmentarea pentru datele de validare	51
Tabelul 4.1 Acuratețea recunoașterii caracterelor cu modelul SVM in faza de antrenare	74
Tabelul 4.2 Acuratețea modelului LPR	75
Tabelul 5.1 Scorul maxim de recunoaștere pentru setul de date MSTAR	86
Tabelul 6.1 Evaluarea performanțelor independenta de subiect pentru identificarea stării de ebrietate prin utilizarea modelului propus	100
Tabelul 6.2 Evaluarea performanțelor independenta de subiect pentru identificarea stării de ebrietate prin modelul în care nu intervine segmentarea PCNN ci imaginea este data în totalitatea sa spre etapa de extragere de trăsături	101
Tabelul 6.3 Matricea de confuzie a subiectului “i” pentru identificarea stării de ebrietate folosind modelul propus	101
Tabelul 6.4 Matricea de confuzie globala pentru identificarea stării de ebrietate folosind modelul propus	102

Lista figurilor

Fig. 2.1 "Componenta de hrănire" a neuronului PCNN	25
Fig. 2.2 "Componenta de conectare" a neuronului PCNN	26
Fig. 2.3 "Componenta aditivă" a neuronului PCNN	27
Fig. 2.4 "Componenta decizionala" a neuronului PCNN	28
Fig. 2.5 Neuronul PCNN	29
Fig. 2.6 Exemplificarea funcționării modelului PCNN	31
Fig. 3.1 Conceptul de funcționare Mean Shift	37
Fig. 3.2 Algoritmul Mean Shift aplicat pe imaginea clasica Lena	37
Fig. 3.3 Algoritmul FCM	40
Fig. 3.4 Algoritmul FCM aplicat pe imaginea clasică Lena	41
Fig. 3.5 Algoritmul SOM de segmentare aplicat pe imaginea clasică Lena	44
Fig. 3.6 Operațiunea de Încrucișarea în AG	46
Fig. 3.7 Operațiunea de Mutăție în AG	46
Fig. 3.8 Algoritmul PCNN clasic aplicat pe imaginea clasica Lena	47
Fig. 3.9 Algoritmul PCNN propus aplicat pe imaginea clasică Lena	48
Fig. 3.10 Segmentarea imaginii care conține zebra	52
Fig. 3.11 Segmentarea imaginii care conține o floare de cactus	53
Fig. 3.12 Segmentarea imaginii care conține o ciuperca	54
Fig. 3.13 Segmentarea imaginii care conține o pisica sălbatică	55
Fig. 3.14 Segmentarea imaginii care conține cai	56
Fig. 3.15 Segmentarea imaginii de la Acropole	57
Fig. 3.16 Segmentarea imaginii care conține un tigru	58
Fig. 3.17 Segmentarea imaginii care conține o biserică	59
Fig. 3.18 Segmentarea imaginii care conține un jaguar	60
Fig. 3.19 Segmentarea imaginii care conține un înotător	61
Fig. 3.20 Segmentarea imaginii care conține un portret	62
Fig. 3.20 Segmentarea imaginii Apollo	63
Fig. 4.1 Maparea datelor pe un alt spațiu	67
Fig. 4.2 SVM-hiperplanul si vectorii suport	69
Fig. 4.3 Exemple Baza de Date	71
Fig. 4.4 Număr de înmatriculare curat pe timp de zi	72

Fig. 4.5 Număr de înmatriculare curat pe timp de noapte	72
Fig. 4.6 Număr de înmatriculare murdar pe timp de zi	73
Fig. 4.7 Număr de înmatriculare segmentat	73
Fig. 5.1 Trei vehicule militare ale fostei Uniuni Sovietice	77
Fig. 5.2 Modelul ATR propus bazat pe rețele neuronale	78
Fig. 5.3 Segmentarea imaginilor folosind combinația celor două modele PCNN	78
Fig. 5.4. Partea reală a răspunsului filtrului Gabor	80
Fig. 5.5 Partea imaginară a răspunsului filtrului Gabor	80
Fig. 5.6 Organigrama VTDG-CSOM	84
Fig. 5.7 Ilustrarea segmentării cu modelele folosite	85
Fig. 6.1. Diagramă flux a cascadei de detecție a stării de ebrietate	90
Fig. 6.2. Ilustrarea Algoritmului LDA	91
Fig. 6.3. Participanții la baza de date	94
Fig. 6.4. Gabriel	96
Fig. 6.5. Mihaela	97
Fig. 6.6. Adrian.....	98
Fig. 6.7. Cristina	99
Fig. 6.8 Imaginile finale prelucrate pentru a fi clasificata starea de ebrietate	100
Fig. 7.1 Banda 4 din imaginea Landsat-5 TM, din zona Mexic.....	109
Fig. 7.2 Schema logica a algoritmului	110
Fig. 4.3 Vizualizarea rezultatelor experimentale	111

Lista abrevierilor

PCNN – Pulse-Coupled Neural Network
FCM – Fuzzy C-Means
SOM – Self Organizing Map
EA – Algoritmi Evolutivi
AG – Algoritm Genetic
OCR – Optical Character Recognition
SVM – Support Vector Machine
LPR – License Plate Recognition
RGB – Red Green Blue
SAR – Radar cu Apertură Sintetică
GF – Filtrarea Gabor
CSOM – Concurrent Self-Organization Maps
ATR – Recunoaștere Automată a Țintei
VTDG – Virtual Training Data Generation
MSTAR – Moving and Stationary Target Acquisition and Recognition
PCA – Analiza Componentelor Principale
KLT – Kosambi-Karhunen-Loeve
POD – Descompunerea Ortogonală Corespunzătoare
EVD – Descompunerea Valorii Proprii
SVD – Descompunerea la Valoare Singulară
SNL – Sandia National Laboratory
LDA – Analiza Liniară Discriminantă
TP – True Pozitive
TN – True Negative
FP – Fals Pozitive
FN – False Negative
OA – Precizie Generala
FAR – Rata de Alarme False
MAR – Rata de Alarme Ratate

GIS – Sistem Informatic Geografic
DI – Diferenței Dintre Imagini
CVA – Analiza Vectorilor de Schimbare
SCV – Vectorii de Schimbare Spectrali
EM – Expectation Maximization
ANN – Artificial Neural Network
TM – Thematic Mapper


CV

Informatii personale

Carata Serban Vasile

 Bucuresti (Romania)

 +40743066041

 carataserban@gmail.com

Experienta profesionala

1/5/2012–Present

Lead Software Developer

UTI, Bucharest (Romania)

Educatie

2013–Prezent

Phd. Vedere Artificiala

Facultatea de Electronică, Telecomunicații și Tehnologia Informației, Polytechnic University of Bucharest, Bucharest (Romania)

2011–2013

Masterat: Sisteme Intelignte si Vedere Artificiala

Facultatea de Electronică, Telecomunicații și Tehnologia Informației, Polytechnic University of Bucharest, Bucharest (Romania)

2008–2011

Physicist

Facultatea de Fizica Medicala - Universitatea din Bucuresti, Bucuresti (Romania)

2007–2010

Mathematician

Facultatea de Matematica si Informatica - Universitatea din Bucuresti, Bucuresti (Romania)

Limba materna

Romana

Limbi Straine	UNDERSTANDING		SPEAKING		WRITING
	Listening	Reading	Spoken interaction	Spoken production	
English	C2	C2	C2	C2	C1
TOEFL					

Papers

- [1] **S. V. Carata**, V. E. Neagoe, „An Innovative Pulse-Coupled Neural Network Approach to Image Segmentation”, 6th European Conference of Computer Science (ECCS '15) Rome, Italy, pp. 137-141, November 2015, (BDI)
- [2] **S. V. Carata**, V. E. Neagoe , „A Pulse-Coupled Neural Network Approach for Image Segmentation and Its Pattern Recognition Application”, 11th International Conference on Communications (COMM), pp. 61-61, June 2016, (IEEE Xplore, ISI Thomson)
- [3] V.E. Neagoe, **S.V. Carata**, A. D. Ciotec, „Automatic Target Recognition in SAR Imagery Using Pulse-Coupled Neural Network Segmentation Cascaded with Virtual Training Data Generation CSOM-Based Classifier”, International Geoscience and Remote Sensing Conference (IGARSS 2015), Milano, Italy, pp. 3274-3277, July 2015, (IEEE Xplore, ISI Thomson)
- [4] V.E. Neagoe, **S.V. Carata**, A. D. Ciotec, „An Advanced Neural Network-Based Approach for Military Ground Vehicle Recognition in SAR Aerial Imagery ”, Scientific Research and Education in the Air Force-Afases, Brasov, pp.41-47, May 2016, (BDI)
- [5] V.E. Neagoe, **S.V. Carata**, „Subject Independent Drunkenness Detection Using Pulse-Coupled Neural Network Segmentation of Thermal Infrared Facial Imagery”, International Journal of Mathematical and Computational Methods, pp. 305-312, August 2016, (BDI)
- [6] V.E. Neagoe, A. D. Ciotec, **S.V. Carata**, „A new multispectral pixel change detection approach using pulse-coupled neural networks for change vector analysis”, International Geoscience and Remote Sensing Conference (IGARSS 2016), Beijing, China, pp. 3386-3389, July 2016, (IEEE Xplore, ISI Thomson)
- [7] C.A. Mitrea, **S.V. Carata**, M. Ghenescu, SCOUTER: Content-based multiple-instance object retrieval system, IEEE International Conference on Multimedia & Expo Workshops (ICMEW), INSPEC 15330364, DOI 10.1109/ICMEW.2015.7169821, TURIN, 2015.
- [8] C.A. Mitrea, **S.V. Carata**, B. Ionescu, T. Piatrik and M. Ghenescu, Ensemble-based Learning Using Few Training Samples for Video Surveillance Scenarios, fifth International Conference on Image Processing Theory, Tools and Applications IPTA, Orleans, France, 2015

Capitolul 1

1. Introducere

Odată cu evoluția rapidă a tehnologiei din ultimii ani și cu ieftinirea acesteia, preluarea imaginilor digitale a devenit din ce în ce mai frecventă și mai folosită într-o multitudine de domenii. Deși acest lucru duce la o înțelegere mai profundă și la un grad de securitate tot mai mare, volumul imens de informație a devenit imposibil de procesat de către operatorii umani. Din acest motiv studiul procesării automate a imaginilor este din ce în ce mai important și se atribuie fonduri și resurse substanțiale perfecționării metodelor de vedere artificială.

Vederea artificială reprezintă o problemă foarte dificilă care încearcă să ofere calculatoarelor, sau altor dispozitive similare, capacitatea de a extrage automat sau semiautomat informație relevantă din imagini. Vederea este o abilitate care, conform biologilor evoluționiști, a apărut în mod independent de mai multe ori în natură. Cu toate acestea omul încă nu a reușit să simuleze acest proces extraordinar de complex, reușind să simuleze doar anumite componente independente din acest proces.

Printre aplicațiile notabile unde se folosește cu succes vederea artificială, se numără(sunt următoarele):

- Recunoașterea automată a obiectelor.
- Recunoașterea automată caracterelor (OCR).
- Detecția schimbărilor (Change Detection).
- Recunoașterea automată a fețelor.
- Recunoașterea automată a stărilor emoționale.
- Recunoașterea automată a stării de ebrietate.

- Imagistica medicală.
- Securitatea.

1.1 Prezentarea domeniului tezei de doctorat

Scopul tezei a fost dezvoltarea unor tehnici (algoritmi) din domeniul *Vederea Artificială (Computer Vision)*. Studiul nostru s-a bazat pe combinarea metodelor bazate pe rețelele neuronale artificiale cu metode statistice pentru a rezolva probleme de vedere artificială.

În principal această lucrare studiază și aduce îmbunătățiri modelului *Pulse Coupled Neural-Network (PCNN)*. Acest model este inspirat de cortexul vizual al pisicii.

Modelul PCNN este combinat cu alte modele neuronale, evolutive și statistice pentru a obține rezultate cât mai bune în diferite domenii de interes.

Printre cele mai notabile modele neuronale folosite se numără *Concurrent Self-Organization Maps (CSOM)* și *Self-Organization Maps (SOM)*.

Modelele evolutive, mai exact *Algoritmii Genetici*, au fost folosiți în scopul de optimizare și rafinare a parametrilor, iar dintre modelele statistice folosite aici vom aminti modelele *Fuzzy C-Means (FCM)* și *Support Vector Machine (SVM)*.

1.2 Scopul tezei de doctorat

Scopul acestei lucrări este de a aduce contribuții originale modelului *Pulse Coupled Neural-Network (PCNN)* și aplicațiilor sale în domeniul *Computer Vision*.

În cadrul tezei am studiat și conceput tehnici cu elemente inovatoare pentru aplicații ale modelului PCNN în domeniul *Computer Vision*, ca:

- Segmentarea imaginilor în tonuri de gri, color și multispectrale;
- Recunoașterea plăcutelor de înmatriculare;
- Identificarea și recunoașterea vehiculelor militare în imagini SAR (Radar cu Apertură Sintetică);
- Recunoașterea persoanelor în stare de ebrietate folosind imagini termale;
- Identificarea schimbărilor în serii de imagini satelitare temporale.

1.3 Conținutul tezei de doctorat

Această lucrare este alcătuită din opt capitole, fiecare tratând o nouă aplicație a modelului PCNN în combinație cu alte tehnici din domeniul *Computer Vision*:

În Capitolul 1 acest capitol se prezintă obiectivele acestei lucrări, domeniul de aplicabilitate, scopul acesteia, precum și un scurt sumar al conținutului lucrării.

Capitolul 2 prezintă modelul teoretic Pulse Coupled Neural-Network (PCNN), prezentându-se componentele neuronului PCNN și modul de funcționare a acestui model. Sunt descrise detaliat cele trei componente de bază ale neuronului (componenta de hrănire, aditivă, decizională), inclusiv cea care are rolul de a face legătura între acestea (componenta e conectare).

Pe parcursul Capitolului 3 sunt prezentate diferite modificări aduse modelului PCNN astfel încât acest model să obțină un rezultat cât mai bun în domeniul segmentării imaginilor. Tot în acest capitol au fost prezentați alți algoritmi de segmentare a imaginilor cu care am comparat performanțele modelului PCNN în acest capitol și în restul tezei. Printre aceste mențiuni se pot enumera algoritmul de segmentare Fuzzy C-Means și etapele acestuia, modelul Self-Organizing Maps, algoritmul de segmentare a imaginilor Mean Shift, precum și algoritmul genetic pentru optimizarea parametrilor PCNN.

Capitolul 4 propune și testează un model complet de recunoaștere a plăcuțelor de înmatriculare, care are la baza segmentării plăcuțelor de înmatriculare modelul PCNN. Cea mai importantă contribuție adusă în acest capitol este mecanismul de oprire a modelului PCNN. Tot în acest capitol sunt prezentate modelul SVM pentru recunoașterea caracterelor, rolul hiperplanului de separare și maparea liniară și funcțiile nucleu.

În Capitolul 5 se propune un algoritm complet de recunoașterea vehiculelor militare în imagini SAR, care are la bază segmentarea cu modelul PCNN. Acest capitol prezintă o metodă nouă de îmbinare a mai multor modele PCNN, cu scopul rafinării segmentării, printre care filtrarea Gabor, analiza Componentelor Principale (PCA) și clasificatorul SVM folosind o antrenare îmbunătățită bazată pe generarea de date de formare virtuală (VTDG) prin Concurrent Self-Organizing Map (CSOM).

Capitolul 6 demonstrează valabilitatea modelului PCNN pentru identificarea stării de ebrietate a subiecților pe baza analizei imaginilor faciale în spectrul termal. Modelul PCNN are rolul de a extrage zonele semnificante din imaginile termale. Se folosesc analiza liniară discriminantă (LDA) și algoritmul LDA, precum și clasificarea SVM.

Capitolul 7 prezintă o aplicație a modelului PCNN pentru detecția schimbărilor în serii temporale de imagini satelitare de observație terestră.

În Capitolul 8 se prezintă concluziile lucrării, rezultatele relevante ale aplicării metodei PCNN îmbunătățite în fiecare caz, dar în special contribuția personală adusă în fiecare din exemplele prezentate în capitolele 3-7.

Capitolul 2

2. Modelul PCNN

Rețeaua PCNN (Pulse-Coupled Neural Network) este un model neuronal inspirat de modul prin care funcționează cortexul vizual al pisicii și a fost dezvoltat pentru procesarea de înaltă performanță a imaginilor [1][2][3][4].

În 1989, Eckhorn a introdus un model neuronal care să se asemene cu mecanismul cortexului vizual al pisicii. Modelul Eckhorn a oferit un instrument simplu și eficient pentru a studia cortexul vizual al mamiferelor mici, și a fost imediat recunoscut ca având potențial semnificativ de aplicare în procesarea imaginilor [1]. În 1994, modelul Eckhorn a fost adaptat de Johnson pentru a fi un algoritm de procesare a imaginilor, care a numit acest algoritm Pulse-Coupled Neural Network [2][4]. În ultimul deceniu, PCNN a fost propus pentru utilizarea unei varietăți de aplicații de procesare a imaginilor, inclusiv pentru: segmentarea imaginilor, generarea de caracteristici, extragerea feței, detectarea mișcării, reducerea zgomotului, și așa mai departe.

PCNN este o rețea neuronală cu două dimensiuni. Fiecare neuron în rețea corespunde unui pixel într-o imagine de intrare, care primește ca informații culoarea pixelilor săi corespunzători (de exemplu, intensitate) ca un stimul extern. Fiecare neuron, de asemenea, are legături cu neuronii vecini, și primește stimuli locali de la ei. Stimulii externi și locali sunt combinați într-un sistem de activare intern, în care se acumulează stimulii până când se depășește un prag dinamic, rezultând o stare activă a neuronului. Prin calculul iterativ, neuronii PCNN produc serii temporale de impulsuri în ieșiri. Seria temporală de impulsuri în ieșiri conține informații de intrare ale imaginii ce pot fi utilizate pentru diverse aplicații de procesare a imaginii, cum ar fi: segmentarea imaginii și

generarea de caracteristici [3][4]. Comparativ cu mijloacele convenționale de procesare a imaginilor, PCNN are mai multe merite semnificative, inclusiv robustețea împotriva zgomotului, independența variațiilor geometrice din modelele de intrare, capacitatea de a trece peste variațiile minore de intensitate în modelele de intrare, etc.

Modelul PCNN este util pentru prelucrarea imaginilor, așa cum s-a discutat într-o carte scrisă de Thomas Lindblad și Jason M. Kinser. Modelul PCNN (Pulse-Coupled Neural Network) este format dintr-o rețea de neuroni, care se inițializează cu dimensiunea imaginii pe care urmează să o analizeze. Cu alte cuvinte, pentru a analiza o imagine de dimensiunea 10 X 10 este nevoie de o rețea bidimensională, formată din 100 de neuroni, aranjați la fel ca pixelii din imaginea respectivă (10 X 10).

Pentru a înțelege rețeaua PCNN, inițial trebuie înțeleasă componenta sa fundamentală, neuronul. Acesta este format din trei componente de bază și o componentă care are rolul de a face legătura între acestea.

2.1 Componenta de alimentare

Prima componentă, cea care acceptă datele de intrare din exteriorul rețelei, se numește "Componenta de hrănire"[2][4]. Aceasta nu preia din exterior doar valoarea aferentă pixelului de aceleași coordonate cu neuronul, ci primește o mască centrată pe aceste coordonate. Pentru o mai bună ilustrare al acestui concept, se va prezenta o diagramă funcțională a acestei componente:

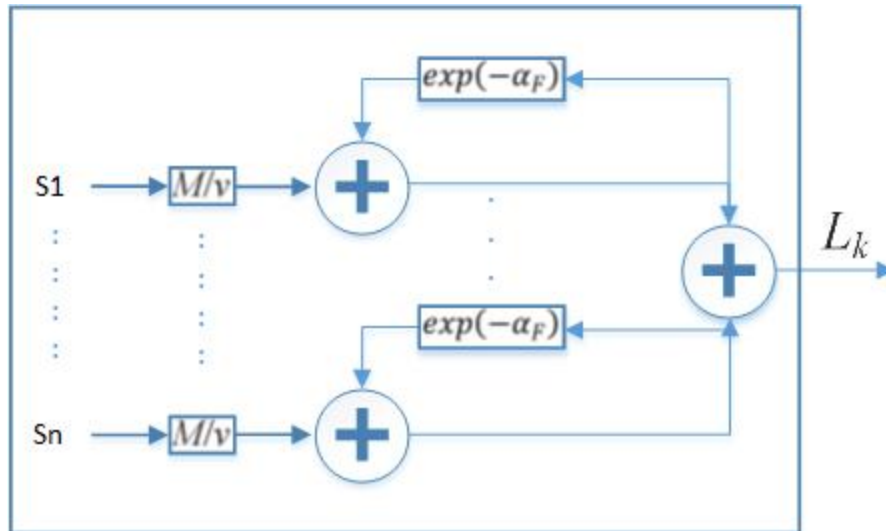


Fig. 2.1 Componenta de „alimentare” a neuronului PCNN.

În aceasta diagramă se observă că rețeaua primește n elemente din exterior și le prelucrează după următoarea ecuație matematică:

$$F_{ij}[n] = \exp(-\alpha_F) * F_{ij}[n - 1] + V_F \sum_k \sum_l M_{ijkl} Y_{kl}[n - 1] + S_{ij} \quad (2.1)$$

unde n este factorul de timp (a câta iterație a rețelei este cea curentă), j este indexul neuronului curent, k reprezintă dimensiunea măștii centrate pe pixelul curent, α_F este o constantă de timp, V_F , este mască de indici care indică puterea de influență a pixelilor din jurul pixelului corespunzător neuronului curent, iar în final M_{ijkl} reprezintă valorile pixelilor din jurul pixelului corespunzător neuronului curent.

Se poate observa din această formulă că etapa curentă este influențată de etapa anterioară. Aceasta influență poate fi controlată cu ajutorul constantei de timp, care se numește și factor de erodare.

2.2 Componenta de conectare

Următoarea componentă a rețelei PCNN se numește "Componenta de conectare (link-are)" [2]. Aceasta are o funcționalitate foarte similară cu cea prezentată anterior, însă are un rol diferit.

Cea mai semnificativă diferență a acestei componente, față de componenta de hrănire, este proveniența datelor de intrare. Deși modul în care acceptă date este identic, primind date de pe o rază stabilită din jurul pixelului curent, aceste date nu sunt provenite dintr-o sursă exterioară, ci datele sunt generate de rețea, în etapa anterioară. Pentru prima iterație, neexistând date anterioare, aceste date se vor genera aleatoriu, cu valori între 0 și 1.

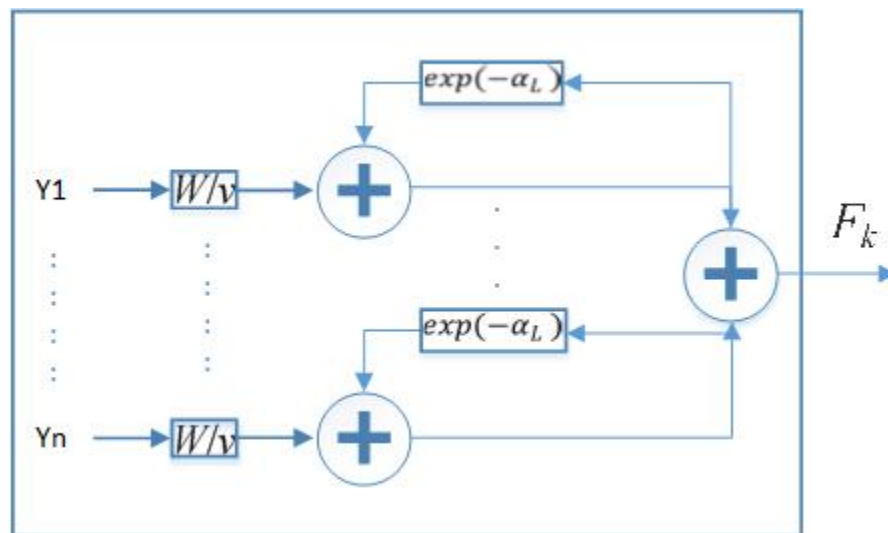


Fig. 2.2 Componenta de „conectare” a neuronului PCNN.

Acestei diagrame îi corespunde formula:

$$L_{ij}[n] = \exp(-\alpha_L) * L_{ij}[n - 1] + V_L \sum_k \sum_l W_{ijkl} Y_{kl}[n - 1] \quad (2.2)$$

unde Y este ieșirea neuronului din iterația anterioară, i este indexul neuronului curent, k reprezintă dimensiunea măștii centrate pe intrarea curentă, α_L este o constantă de timp, V_L este masca de indici care indică puterea de influență a valorilor din jurul valorii corespunzătoare neuronului curent, iar în final M_{ijkl} reprezintă valorile datelor din jurul datei corespunzătoare neuronului curent.

2.3 Componenta aditivă

Componenta următoare face legătura între cele două componente prezentate anterior pentru a genera datele de intrare pentru ultima componentă, cea decizională.

Principiul acestei componente este foarte simplu: se atribuie o pondere valorii generate de componenta de conectare și se înmulțește cu valoarea generată de componenta de hrănire. Schema acestei componente este următoarea, iar acesteia îi corespunde expresia matematică 1.3.

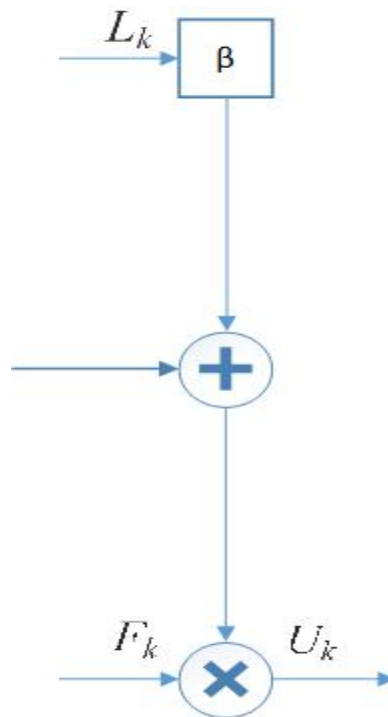


Fig. 2.3 Componenta „aditivă” a neuronului PCNN.

$$U_{ij}[n] = F_{ij}[n](1 + \beta L_{ij}[n]) \quad (2.3)$$

unde β este ponderea de influență a componenteii de conectare. Acest parametru are un rol foarte important, fiind unul din cei mai influenți parametri ai rețelei.

2.4 Componenta decizională

Ultima componentă a rețelei PCNN este cea decizională[2][4]. Aceasta acceptă ca intrare o singură valoare numerică, generată de restul componentelor PCNN, și generează un rezultat final, binar, cu fiecare iterație.

Diagrama acestei componente este următoarea:

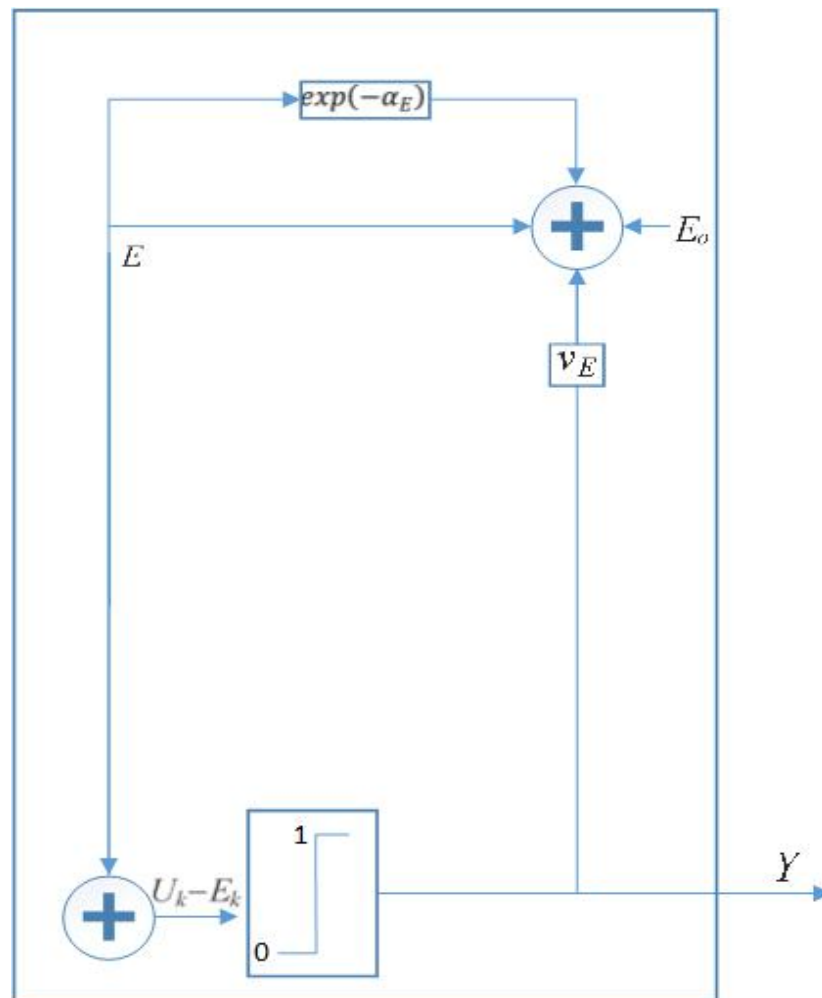


Fig. 2.4 Componenta „decizionala” a neuronului PCNN.

a cărei formule atașate sunt:

$$E_{ij}[n] = E_{ij}[n - 1] \exp(-\alpha_E) + V_E Y_{ij}[n - 1] \quad (2.4)$$

$$Y_{ij}[n] = \begin{cases} 1, & \text{if } U_{ij}[n] > E_{ij}[n] \\ 0, & \text{otherwise} \end{cases} \quad (2.5)$$

unde V_E este un indice care influențează puterea rezultatului din generația trecută. Aici se observă că E_k și Y_k sunt influențați de iterația anterioară, iar la prima iterație se vor genera date aleatoare cu valori între 0 și 1 pentru a le popula. Tot din aceste formule se observă că la prima iterație toți neuronii se vor activa, deoarece E_k invariabil mai mic decât U_k .

2.5 Neuronul PCNN

Odată expuse toate componentele neuronului, se poate prezenta neuronul în întregime sa. Acestuia îi corespunde următoarea schemă funcțională:

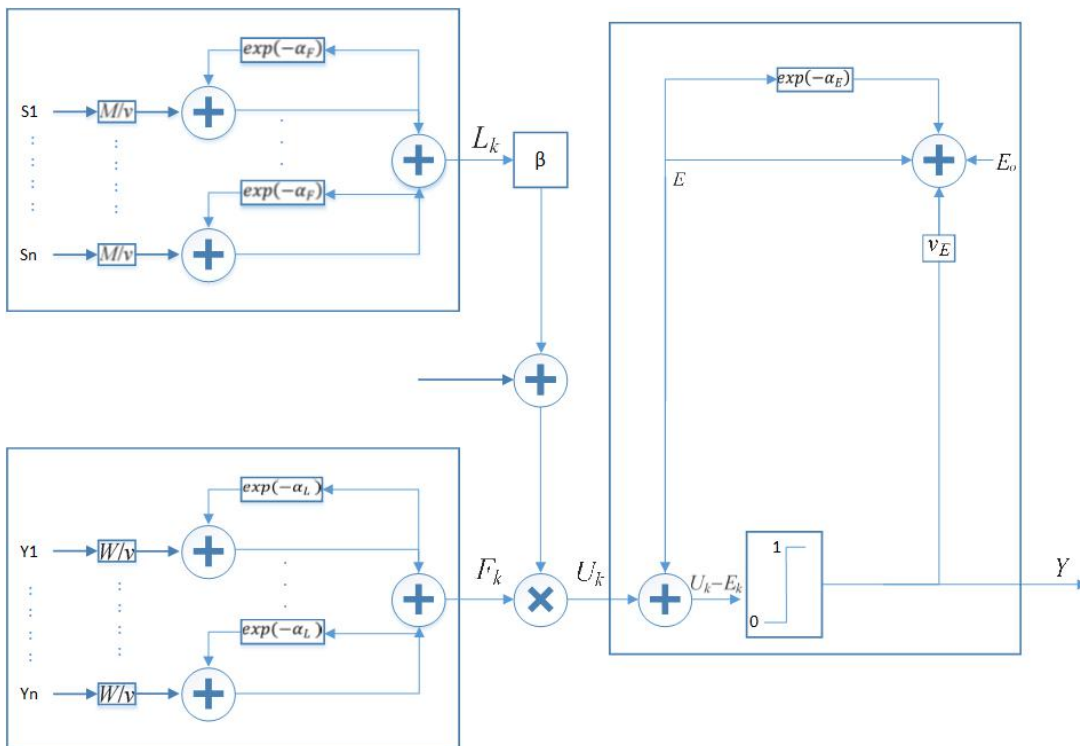


Fig. 2.5 Neuronul PCNN.

Aceștia îi corespund următoarele expresii matematice:

$$F_{ij}[n] = \exp(-\alpha_F) * F_{ij}[n - 1] + V_F \sum_k \sum_l M_{ijkl} Y_{kl}[n - 1] + S_{ij} \quad (2.6)$$

$$L_{ij}[n] = \exp(-\alpha_L) * L_{ij}[n - 1] + V_L \sum_k \sum_l W_{ijkl} Y_{kl}[n - 1] \quad (2.7)$$

$$U_{ij}[n] = F_{ij}[n](1 + \beta L_{ij}[n]) \quad (2.8)$$

$$E_{ij}[n] = E_{ij}[n - 1] \exp(-\alpha_E) + V_E Y_{ij}[n - 1] \quad (2.9)$$

$$Y_{ij}[n] = \begin{cases} 1, & \text{if } U_{ij}[n] > E_{ij}[n] \\ 0, & \text{otherwise} \end{cases} \quad (2.10)$$

Pentru a procesa o imagine cu ajutorul modelului PCNN este suficient să se considere matricea binară de ieșire a rețelei, care grupează pixelii atât după trăsături (valoare), cât și după vecinătatea lor unul față de altul. Acest fapt constituie un avantaj deoarece nu sunt necesare modificări ale modelului standard PCNN, el realizând în forma sa clasică segmentarea imaginilor.

2.6 Exemplificarea funcționării modelului PCNN

Odată expuse bazele matematice și modul în care se inițializează rețeaua, se poate expune modul de funcționare a rețelei. În primul rând trebuie înțeles faptul că rețeaua modifică rezultatul de la fiecare iterație, însă rezultatele nu se repetă, iar influențele anterioare nu sunt cumulative, ele fiind influențate de factorii de descompunere. Pe scurt rețeaua poate merge la nesfârșit generând la fiecare iterație un alt rezultat viabil.

În continuare se vor expune trei iterații consecutive ale rețelei, pentru a exemplifica evoluția rezultatului generat de rețeaua PCNN. Aceste rezultate sunt generate folosind un set de parametri predefiniți, exemplificați în tabelul 2.1.

Tab. 2.1 Parametrii Interni PCNN.

Dimensiunea matricei M si W	$V_L = V_F$	V_E	$\alpha_L = \alpha_F$	α_F	β
3	1.00	1.00	0.07	0.17	0.18

Numărul iterațiilor realizate de modelul PCNN trebuie setat înainte de pornirea rețelei deoarece nu exista nici un mecanism intern pentru oprirea acesteia.



Fig. 2.6 Exemplificarea funcționării modelului PCNN.
(a) Imaginea originală; (b) Ieșirea modelului PCNN - 3 iterații;
(c) Ieșirea modelului PCNN - 5 iterații; (d) Ieșirea modelului PCNN - 7 iterații; (Pixelii albi reprezentand neuronii activați)

În figura 2.6 se poate observa diferența între numărul de iterații realizate de modelul PCNN asupra aceleiași imagini. Este evident că fiecare iterație adițională evidențiază alte trăsături, uneori păstrând și trăsături apărute anterior.

Notabil este faptul că, indiferent de numărul de iterații, rețeaua PCNN nu ajunge la saturație, însemnând activarea întregii imagini, acest lucru datorându-se constantei de timp.

Capitolul 3

3. PCNN pentru segmentarea imaginilor în tonuri de gri și a imaginilor color

3.1 Introducere

Segmentarea imaginilor este un mod specific de prelucrare a semnalelor bidimensionale care identifica grupuri de pixeli care "se potrivesc". În statistică, această problemă este cunoscută sub numele de „analiza cluster-elor” [11]-[15] și este o un domeniul de studiu pe scară largă cu sute de algoritmi diferiți.

În Computer Vision, segmentarea imaginilor este una dintre cele mai vechi și mai larg studiate probleme. Primele tehnici dezvoltate au folosit metode de divizare, sau fuziune, a regiunilor din imagini, aceste tehnici corespunzând algoritmilor de dezbinare și grupare în literatura de specialitate. Algoritmii mai recentți folosesc optimizarea unor criterii globale, cum ar fi texturi similare în cadrul aceleiași regiuni pentru a obține o segmentare cât mai exactă.

Printre algoritmii de segmentare se includ algoritmii bazați pe contururi active, regiuni de separare și concentrare, Mean Shift, tăieturi normalizate (divizarea pe baza valorilor de similaritate a pixelilor), și câmpuri aleatoare binare Markov rezolvate folosind reduceri de grafuri.

Având în vedere că literatura de specialitate privind segmentarea imaginilor este atât de mare, o modalitate bună de a obține o viziune bună asupra unor algoritmi mai performanți este să analizăm comparațiile experimentale asupra bazelor de date indexate manual. Cea mai cunoscută dintre aceste baze de date este Berkeley Segmentation Dataset and Benchmark, care este formată din 1000 de imagini, care au fost marcate de 30 de oameni. Mulți dintre algoritmi de segmentare a imaginilor mai recent raportează rezultate comparative cu privire la această bază de date. De aceea, Unnikrishnan, Pantofaru și Hebert (2007) propun noi metrici pentru compararea acestor algoritmi. Estrada și Jepson (2009) compară patru algoritmi de segmentare bine cunoscuți pe setul de date Berkeley și ajung la concluzia că în timp ce algoritmul propriu SE-MinCut surclasează celelalte metode printr-o marjă mică, există încă un decalaj mare între performanța automată și cea umană.

3.1.1 Segmentarea imaginilor cu algoritmi bazați pe contur activ

În timp ce liniile, punctele de fugă și dreptunghiurile sunt un lucru obișnuit în mediul artificial creat de om, curbele care definesc marginile obiectelor sunt chiar mai frecvente, mai ales în mediul natural. În această secțiune, vom descrie trei abordări legate de localizarea acestor curbe de limită, în imagini.

Prima metoda, numită „snakes” de către inventatorii (Kass, Witkin și Terzopoulos 1988), este o metoda care minimizează energia. Aceasta metoda generează o curbă bi-dimensională, care evoluează (se mișcă) spre caracteristici din imagine, cum ar fi o margine pronunțată. A doua metoda, „foarfece inteligente” (Mortensen and Barrett 1995), permite utilizatorului să schițeze în timp real o curbă, care se „agată” de marginile obiectelor. În cele din urmă, tehnicile „stabilite la nivel” evoluează o curbă ca zero set a unei funcții caracteristice. Acest lucru permite acestor tehnici să își schimbe cu ușurință topologiile și să includă statistici bazate pe regiuni.

3.1.2 Segmentarea imaginilor bazată pe separate și îmbinare

Cea mai simplă tehnică posibilă pentru împărțirea unei imagini în tonuri de gri este de a selecta un prag și de a calcula apoi componentele conectate. Din păcate, un singur prag este rareori suficient pentru întreaga imagine, din cauza dispozitivelor de iluminat și variațiilor statistice din obiecte.

Această tehnică simplistă a fost dezvoltată în o serie de algoritmi care fie împart întreaga imagine în bucăți, pe baza statisticilor de regiune sau invers, care fuzionează pixeli și regiuni împreună într-o manieră ierarhică și apoi aplica un prag.

3.1.3 Segmentare a imaginilor bazată pe tăieturi normalizate

În timp ce metodele prezentate până acum încercau să grupeze pixeli sau regiuni similare, tehnica tăieturilor normalizate introdusă de Shi și Malik (2000) examinează afinități (similarități) între pixeli și încearcă să separe grupurile care sunt legate prin afinități slabe.

O temă comună a algoritmilor de segmentare a imaginilor este dorința de a aduna pixelii în grupul care au aspect similar (statistic) și să obțină limitele dintre regiunile de pixeli diferiți. Putem formula acest lucru ca o funcție clasică de energie bazată pe pixeli, folosind fie o formulare variațională, fie ca un câmp binar aleator Markov.

În continuare se vor prezenta mai amănunțit metodele de segmentare folosite în această lucrare. Pe lângă acestea vom prezenta și modelele propuse, bazate pe modelul neuronal PCNN.

3.2 Algoritmul Mean Shift pentru segmentarea imaginilor

Mean Shift este o procedură simplă iterativă care schimbă fiecare punct de date cu media punctelor de date din vecinătatea sa. S-a arătat că Mean Shift este un proces de căutare, pe o suprafață, a zonelor de înaltă densitate.

Algoritmul Mean Shift este o tehnică de grupare neparametrică care nu necesită cunoașterea prealabilă a numărului de clustere, și nu constrânge forma clusterului.

Dacă se dau n valori x_i , $i = 1, \dots, n$ pe un spațiu d -dimensional R^d , kernelul care estimează densitatea multiplelor variabile $K(x)$ și o matrice H simetrică, cu valori pozitive de dimensiunea $d \times d$ calculează punctul x dat de :

$$f(x) = \frac{1}{n} \sum_{i=1}^n K_H(x - x_i) \quad (3.1)$$

unde

$$K_H(x) = |H|^{-\frac{1}{2}} K\left(H^{-\frac{1}{2}} x\right)$$

Kernelul cu d variabile $K(x)$ este o funcție mărginită care satisface ecuațiile:

$$\int_{R^d} K(x) dx = 1 \quad (3.2)$$

$$\lim_{\|x\| \rightarrow \infty} \|x\|^d K(x) = 0 \quad (3.3)$$

$$\int_{R^d} xK(x)dx = 0 \quad (3.4)$$

$$\int_{R^d} xx^T K(x)dx = c_K I \quad (3.5)$$

unde c_K este o constanta. Kernelul poate fi generat dintr-un kernel simetric cu o singura variabila $K_1(x)$ in doua moduri diferite

$$K^P(x) = \prod_{i=1}^d K_1(x_i) \quad (3.6)$$

$$K^S(x) = a_{k,d} K_1(\|x\|) \quad (3.7)$$

unde $K^P(x)$ este obtinut din produsul nucleelor de o singura variabila, iar $K^S(x)$ prin rotirea kernelului de o singura variabila in R^d . Constanta $a_{k,d}^{-1} = \int_{R^d} K_1(\|x\|)dx$ asigura faptul ca $K^S(x)$ se integreaza la 1.

In acest caz, pentru o mai buna înțelegere, ne concentram pe cazul de kernel simetric care satisface ecuația:

$$K(x) = c_{k,d} k(\|x\|^2) \quad (3.8)$$

in acest caz fiind suficient sa se definească funcția $k(x)$, numita profilul kernelului, numai pentru $x \geq 0$. Constanta de normalizare $c_{k,d}$ este strict pozitiva.

Folosind un H complet parametrizat creste complexitatea estimării si, in practica, matricea H este aleasa ca fiind matrice diagonala sau proporțională cu matricea identitate $H = h^2 I$. Folosind acestea, kernelul devine:

$$f(x) = \frac{1}{nh^d} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right) \quad (3.9)$$

De aici se poate obține estimatorul gradientului de densitate

$$f_{h,K}(x) = \frac{c_{k,d}}{nh^d} \sum_{i=1}^n K \left(\left\| \frac{x - x_i}{h} \right\|^2 \right) \quad (3.10)$$

si prin urmare

$$\nabla f_{h,K}(x) = \frac{2c_{k,d}}{nh^{d+2}} \sum_{i=1}^n (x - x_i) k' \left(\left\| \frac{x - x_i}{h} \right\|^2 \right) \quad (3.11)$$

Definim funcția

$$g(x) = -k'(x) \quad (3.12)$$

presupunând ca derivata lui k exista pentru toți $x \in [0, \infty)$, exceptia unui numar finit de puncte. Folosind aceasta ecuatie, kernelul, $G(x)$, devine

$$G(x) = c_{g,d} g(\|x\|^2) \quad (3.13)$$

Unde $c_{g,d}$ este o contanta de normalizare.

Introducând $g(x)$ in ecuația 3.11 obținem:

$$\begin{aligned} \nabla f(x) &= \frac{qc_{k,d}}{nh^{d+2}} \sum_{i=1}^n (x_i - x) g \left(\left\| \frac{x - x_i}{h} \right\|^2 \right) \\ &= \frac{2c_{k,d}}{nh^{d+2}} \left[\sum_{i=1}^n g \left(\left\| \frac{x - x_i}{h} \right\|^2 \right) \right] \left[\frac{\sum_{i=1}^n x_i g \left(\left\| \frac{x - x_i}{h} \right\|^2 \right)}{\sum_{i=1}^n g \left(\left\| \frac{x - x_i}{h} \right\|^2 \right)} - x \right] \end{aligned} \quad (3.14)$$

unde $\sum_{i=1}^n x_i g \left(\left\| \frac{x - x_i}{h} \right\|^2 \right)$ este presupus a fii un numar pozitiv. Aceasta conditie este usor de satisfacut in practica.

Al doilea termen al acestei ecuații, 3.15, adică:

$$m_h(x) = \frac{\sum_{i=1}^n x_i g \left(\left\| \frac{x - x_i}{h} \right\|^2 \right)}{\sum_{i=1}^n g \left(\left\| \frac{x - x_i}{h} \right\|^2 \right)} - x \quad (3.15)$$

reprezintă deplasarea medie (Mean Shift). Acest vector este orientat mereu spre zona de densitate maxima a clasei. Acest lucru se asigură prin succesiune pașilor următori:

- 1) se calculează $m_h(x^t)$.
- 2) translatarea ferestrei $x^{t+1} = x^t + m_h(x^t)$ astfel încât să tindă la zona de densitate maximă.

Mai jos se va arăta o ilustrare a conceptului - în figura 3.1, și în figura 3.2 este exemplificat un rezultat al segmentării cu ajutorul Mean Shift:

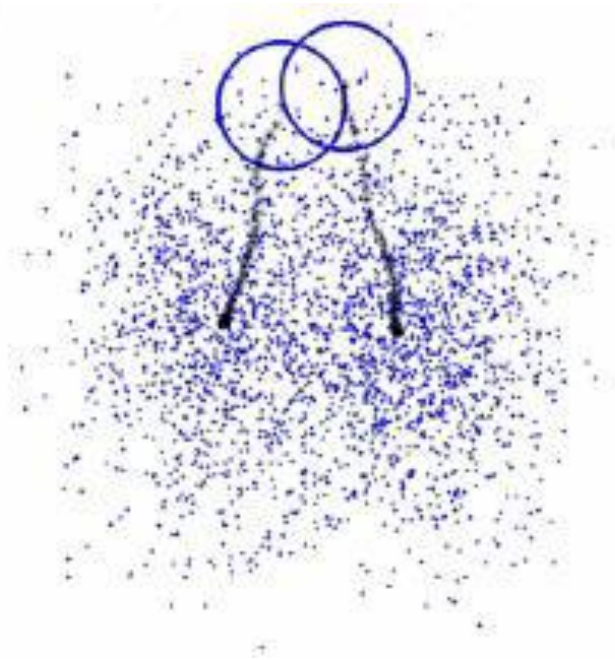


Fig. 3.1 Conceptul de funcționare Mean Shift

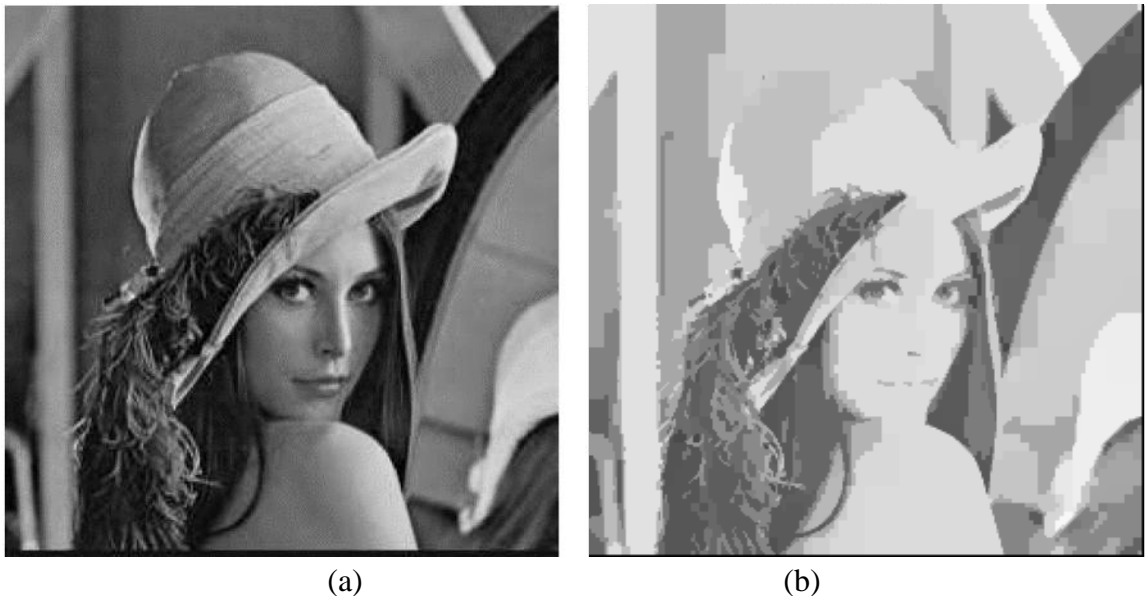


Fig. 3.2 Algoritmul Mean Shift aplicat pe imaginea clasica Lena
(a) Imaginea originală; (b) Imaginea segmentată (8 centroizi)

3.3 Algoritmul Fuzzy C-Means pentru segmentarea imaginilor

Segmentarea fuzzy [5], [6] este un mod eficient de segmentare de obiecte în imaginile care conțin atât zgomot aleatoriu, cât și umbrire. Acest fapt a fost ilustrat atât în fotografiile create matematic, cât și pe unele obținute din imagistica medicală. O teorie de segmentare Fuzzy este următoarea: pentru a efectua segmentarea fuzzy, o "hartă de conectare" trebuie să fie creată. Pentru aceasta se va folosi un algoritm de tip C-Means [5]. Odată ce harta de conectare este creată, pe ea se va aplica logica Fuzzy, astfel obținându-se imaginea segmentată.

Logica fuzzy este o formă de logică probabilistică, ce oferă ca rezultate o plajă de valori în locul unei decizii binare - fixe și exacte. Comparativ cu seturile tradiționale binare (în cazul în care variabilele pot lua valori adevărate sau false), variabilele logicii fuzzy ar putea avea o valoare de adevăr, care variază între 0 și 1. Logica fuzzy a fost extinsă pentru a gestiona conceptul de adevăr parțial, în cazul în care valoarea de adevăr poate varia între complet adevărat și complet fals. În plus, atunci când se utilizează variabile lingvistice, aceste grade pot fi gestionate de funcții specifice. Iraționalitate poate fi descrisă în termeni de ceea ce este cunoscut sub numele de fuzzificare.

Termenul de "logica fuzzy" a fost introdus în anul 1965 la propunerea teoriei mulțimilor fuzzy dezvoltată de Lotfi A. Zadeh. Logica fuzzy a fost aplicată în mai multe domenii, de la teoria controlului, la inteligența artificială.

Dacă se folosește analiza clasică, denumită non-fuzzy sau hard clustering, datele se divid în grupuri bine definite, astfel încât fiecare vector aparține unei singure categorii. În cazul clustering-ului fuzzy, un vector poate aparține mai multor categorii [6].

Asocierea unui vector la o categorie se realizează folosind gradele de apartenență. Altfel spus:

$$\begin{array}{ll} \text{Pentru gruparea non-fuzzy} & \text{Pentru gruparea fuzzy} \\ y_{ij} = \begin{cases} 1 & \text{dacă } x_i \in C_j \\ 0 & \text{dacă nu} \end{cases} & y_{ij} \in [0,1] \end{array}$$

Algoritmul FCM (numit și Fuzzy ISODATA), este bazat pe minimizarea funcției obiectiv

$$J(Y, V) = \sum_{i=1}^n \sum_{j=1}^g (\mu_{ij})^m \|x_i - v_j\|^2 \quad (3.16)$$

unde $X = \{x_1, x_2, \dots, x_n\}$ este o mulțime de vectori, $V = \{v_1, v_2, \dots, v_g\}$ este setul de vectori reprezentând centrele (mediile) grupurilor, unde n este numărul de vectori, g este numărul de categorii, iar y_{ij} reprezintă gradul de apartenență al vectorului x_i la categoria j .

y_{ij} satisface următoarele condiții :

$$y_{ij} \in [0,1], \forall i = 1, \dots, n, \forall j = 1, \dots, g \quad (3.17)$$

$$\sum_{j=1}^g y_{ij} = 1, \forall i = 1 \dots n \quad (3.18)$$

$U = (y_{ij})_{n \times g}$ este matricea gradelor de apartenență.

$\|x_i - v_j\|$ reprezintă distanța Euclidiană dintre x_i și v_j .

m este gradul de fuzificare; este folosit pentru a controla micile diferențe dintre gradele de apartenență. Valoarea lui m este cuprinsă obligatoriu în intervalul $[1, \infty]$. În teorie nu există o regulă pentru alegerea valorii lui m , însă de obicei valoarea $m = 2$ este aleasă, bazându-se pe experiențe anterioare.

3.3.1 Etapele algoritmului Fuzzy C-Means

FCM (Fuzzy C-Means clustering) [5], [6] este o metodă supervizată care a fost aplicată cu succes în multe domenii, printre care și segmentarea imaginilor. Prin aplicarea algoritmilor FCM în probleme de segmentare a imaginilor, imaginile reprezentate în spațiul de valori alese sunt clasificate de FCM prin gruparea punctelor similare, astfel realizându-se și segmentarea. Prin minimizarea funcției de cost care este dependentă de distanțele dintre pixeli și centroizi, funcția își poate modifica centroidul la fiecare iterație.

Primul pas este inițializarea centrelor claselor $V = \{v_1, v_2, \dots, v_g\}$, sau inițializarea gradelor de apartenență u_{ij} cu o valoare arbitrară, care să satisfacă condițiile algoritmului, apoi calcularea centrelor claselor.

Calcularea gradelor de apartenență u_{ij} se face folosind formulele:

$$u_{ij} = \frac{1}{\sum_{k=1}^g \left(\frac{d_{ij}}{d_{ik}}\right)^{\frac{2}{m-1}}} \quad (3.19)$$

unde $d_{ij} = \|x_i - v_j\|, \forall i = 1, \dots, n, \forall j = 1, \dots, g$.

Apoi se calculează centrele v_j folosind:

$$v_j = \frac{\sum_{i=1}^n (u_{ij})^m x_i}{\sum_{i=1}^n (u_{ij})^m} \quad \forall j = 1, \dots, g \quad (3.20)$$

Ultimul pas al algoritmului constituie repetarea celor doi pași anteriori până la minimumul valorii lui J .

În figura 3.3 se ilustrează o imagine care reprezintă rezultatul segmentării cu metoda FCM, iar în figura 3.4 se ilustrează algoritmul FCM printr-o schemă logică.

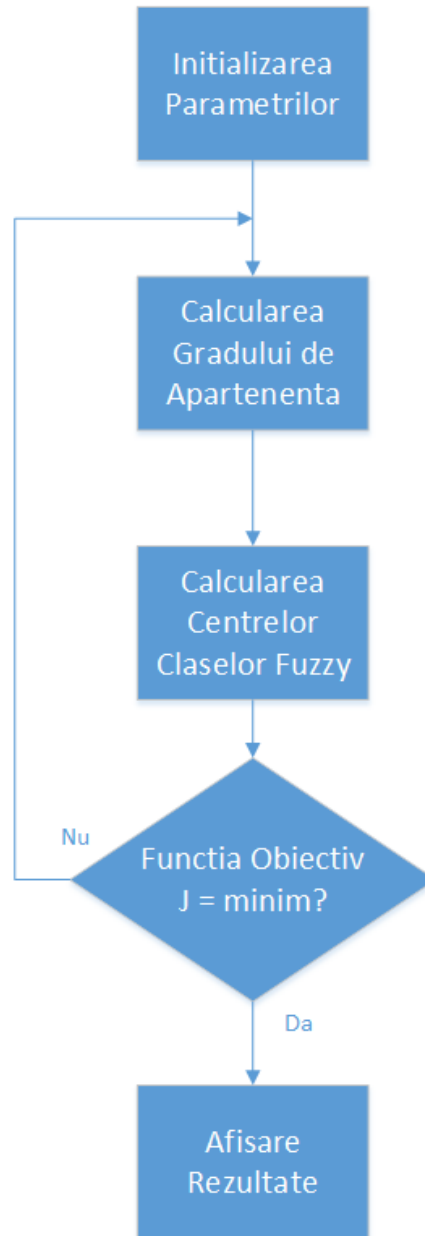


Fig. 3.3 Algoritmul FCM



Fig. 3.4 Algoritmul FCM aplicat pe imaginea clasică Lena
 (a) Imaginea originală; (b) Imaginea segmentată

3.4 Rețeaua cu autoorganizare (Self-Organizing Map)

SOM (Self Organizing Map) [7] este un tip de rețea neuronală artificială, care este instruită folosind metode de învățare nesupravegheate pentru a produce o reprezentare a spațiului de intrare, formând o hartă. SOM este diferită de alte rețele neuronale artificiale, în sensul că utilizează o funcție de învecinatate pentru a păstra proprietățile topologice ale spațiului de intrare.

Modelul a fost descris pentru prima dată ca o rețea neuronală artificială de către profesorul finlandez Teuvo Kohonen, și este uneori numit o hartă sau rețea Kohonen.

Cercetările privind neuro-biologia au arătat că diferite centre de activitate din creier, cum ar fi vorbirea, vederea, auzul, se află în zone specifice ale cortexului. Aceste zone sunt ordonate pentru păstrarea relațiilor topologice dintre informațiile aferente acestora, în timp ce efectuează o reducere a dimensiunii spațiului de reprezentare. O astfel de organizare a condus Kohonen pentru a dezvolta algoritmul SOM [7].

Modelul de SOM explicat în continuare este unidimensional și conține un număr de n noduri. Fiecărui neuron N_i i se asociază un vector de ponderi $w_i = (w_{i1}, w_{i2}, \dots, w_{ip})^t$ unde $w_i \in \mathbb{R}^p$.

În procesul de antrenare a rețelei, se introduce un vector de intrare $x \in \mathbb{R}^p$, ales aleatoriu din lotul de antrenare și este conectat în paralel la toți neuronii rețelei. Acest vector este comparat, folosind distanța Euclidiană, cu toți vectorii de ponderi, w_i , ai

neuronilor din rețea. La etapa k a procesului, vectorul x este atribuit la ponderile neuronului câștigător N_l dacă:

$$\|x - w_l^{[k]}\| = \min \|x - w_i^{[k]}\| \quad (3.21)$$

Toți neuronii dintr-o vecinătate din jurul “liderului” sunt afectați de modificarea ponderii. Considerând faptul ca rețeaua se inițializează cu parametrii aleatorii pentru $w_i^{[0]} (0 \leq i \leq n)$, procesul de antrenare poate fi descris prin următoarea relație iterativă:

$$w_i^{[k+1]} = w_i^{[k]} + \mathcal{H}_{li}^{[k]}(x^{[k]} - w_i^{[k]}) \quad (3.22)$$

Ultima iterație între elementele apropiate topologic este modelată pe o funcție de nivelare, definită asupra neuronului câștigător[7]. Aceasta funcție de nivelare poate fi scrisă similar unei funcții gaussiene.

$$\mathcal{H}_{li}^{[k]} = \alpha^{[k]} \exp\left(-\frac{d^2(l, i)}{2(\sigma^{[k]})^2}\right) \quad (3.23)$$

unde $d(l, i) = \|l - i\|$ reprezintă distanța dintre nodul l și nodul i , $\alpha^{[k]}$ este factorul de învățare iar $\sigma^{[k]}$ definește raza de acțiune a funcției la iterația k . Pentru convergență, este necesar ca $\mathcal{H}_{li}^{[k]} \rightarrow 0$, când $k \rightarrow T$, unde T reprezintă numărul total de pași ai procesului [7]. Prin urmare, pentru prima iterație a procesului, valoarea parametrului $\alpha^{[k]}$ este bine să fie cât mai apropiată de unitate, scăzând pe parcurs ce numărul de iterații crește. Pentru a obține acest lucru se poate folosi ecuația următoare:

$$\alpha^{[k]} = \alpha^{[0]} \left(1 - \frac{k}{T}\right) \quad (3.24)$$

În plus, cu cât avansează mai mult procesul de antrenare, dimensiunea vecinătății trebuie micșorată până când se ajunge la adaptarea unui singur neuron. Luând în considerație acest lucru, valoarea parametrului $\sigma^{[k]}$ poate fi definită folosind ecuația următoare:

$$\sigma^{[k]} = \sigma^{[0]} \left(\frac{\sigma^{[T-1]}}{\sigma^{[0]}}\right)^{\frac{k}{T-1}} \quad (3.25)$$

Se poate observa că ordonarea harții se face în primii pași ai antrenării, în timp ce rafinarea parametrilor se face în pașii ulteriori ai procesului de antrenare.

Etapa de antrenament se face direct pe imaginea ce urmează a fi segmentată. Vectorul de intrare se poate forma prin concatenarea elementelor dintr-o fereastră de 3x3 care este deplasată deasupra imaginii pe niveluri de gri. Prin acest proces fiecare neuron are un număr total de 9 ponderi. La fiecare pas al procesului, poziția ferestrei se modifică și astfel ponderile neuronilor se adaptează conform relație 3.22.

σ are un rol foarte important în calitatea convergenței. Algoritmul trebuie pornit cu o valoare relativ mare pentru acest parametru astfel încât să se poată stabili ordinea globală a hărții. Valoarea inițială a lui σ poate să fie jumătate din dimensiunea hărții. Pe parcursul antrenării σ trebuie să își scadă valoarea până la valoarea dedusă experimentală, $\sigma_{T-1} = 0.1$ pentru a minimiza eroarea de cuantizare definită de ecuația

$$E_{cuant} = \frac{1}{N} \sum_{i=1}^N \|x_i - w_i\| \quad (3.26)$$

unde sumarea se face peste toate cele N ferestre.

3.4.1 SOM pentru segmentarea imaginilor

Sarcina de clasificare constă în rularea unei ferestre glisante deasupra imaginii. Pentru fiecare locație a ferestrei, vectorul de intrare x corespunzător ferestrei este comparat cu toți neuronii folosind ecuația 3.21. Neuronul câștigător care conduce la cea mai mică distanță, oferă clasa pixelului situat în centrul ferestrei. Cu toate acestea, înainte de orice sarcină de clasificare, trebuie să se calibreze harta în scopul de a se asocia etichete pentru fiecare neuron din rețea.

Procesul de segmentare poate fi descris prin următorul algoritm [7] :

1. Se inițializează ponderile interne ale rețelei
2. Se incrementează numărul iterației curente t
3. Se calculează distanțele d_i
4. Se determină nodul a cărui distanță calculată cu ecuația 3.21 este minimă.
5. Se rafinează ponderile aferente neuronului.
6. Dacă nu s-a terminat parcurgerea lotului de vectori, se merge la pasul 2.
7. Dacă s-a terminat de parcurs tot lotul de vectori, se verifică dacă este îndeplinită condiția de stop. Dacă nu este îndeplinită condiția de stop, se merge la pasul 2.
8. Se determină cel mai reprezentativ membru aferent fiecărei clase.
9. Se determină plaje de valori centrate pe cel mai semnificativ membru al claselor.
10. Neuronii care se încadrează în aceste plaje primesc eticheta neuronului dominant.
11. Imaginea finală este segmentată atribuind fiecărui pixel eticheta neuronului aferent.

În general rafinarea rețelei se oprește când gradientul de modificare a ponderilor rețelei devine nesemnificativ sau când se atinge un număr fixat de iterații, dat fiind că rafinarea rețelei este un proces de durată.

În figura 3.5 se poate vedea rezultatul unei segmentări cu ajutorul modelului SOM.



*Fig. 3.5 Algoritmul SOM de segmentare aplicat pe imaginea clasică Lena
(a) Imaginea originală; (b) Imaginea segmentată; (4 clase)*

3.5 Optimizarea parametrilor PCNN utilizând algoritmi genetici

În domeniul inteligenței artificiale, un algoritm genetic [8] (AG) este o metodă euristică de căutare, inspirată din procesul selecției naturale. Acești algoritmi (numiți uneori metaheuristici) sunt folosiți în mod curent pentru a genera soluții utile în vederea optimizării și căutării soluțiilor problemelor. [9] Algoritmii genetici fac parte din clasa mai mare de algoritmi evolutivi (EA), care implică probleme de optimizare combinatorială, asemenea probleme fiind caracterizate prin existența unui set discret de soluții posibile. Calculul evolutiv utilizează progresul iterativ, cum ar fi acela caracterizând creșterea și dezvoltarea unei populații. Această populație este apoi selectată printr-o căutare aleatoare ghidată, utilizând o prelucrare paralelă, pentru a obține o soluție optimă. Asemenea procese sunt inspirate de mecanismele biologice ale evoluției. În biologie, evoluția înseamnă schimbarea materialului genetic al unei populații de organisme de la o generație la alta, astfel încât după un set de generații se poate ajunge la schimbări semnificative ale populației, care pot conduce chiar la noi specii [8].

Într-un algoritm genetic, o populație de soluții-candidat este evoluată spre soluții mai bune printr-o problemă de optimizare. O soluție candidat (phenotype) este reprezentată printr-un cuvânt de cod care constituie cromozomul (genotype). Simbolurile cromozomului sunt reprezentate de *gene*. În mod tradițional, genele sunt reprezentate în cod binar ca șiruri de 0 și 1, dar și alte codificări sunt de asemenea posibile [10].

Evoluția începe de obicei dintr-o populație de indivizi generată în mod aleatoriu, și este un proces iterativ - populația din fiecare iterație numindu-se astfel o generație. În fiecare generație, aptitudinea fiecărui individ din populație este evaluată, și se numește fitness-ul membrului din populație. Indivizii mai potriviți sunt stocastic selectați din populația actuală și genomul fiecărui individ este modificat (recombinat, mutații, etc.) pentru a forma o nouă generație. Noua generație de soluții candidate este apoi utilizată în următoarea iterație a algoritmului. În mod obișnuit, algoritmul se termină atunci când a fost produs fie un număr maxim de generații, sau un nivel satisfăcător de fitness a fost atins pentru populație.

Un algoritm tipic genetic necesită:

- o reprezentare genetică a domeniului soluției,
- o funcție de fitness pentru a evalua domeniul de soluție.

O reprezentare standard a fiecărui candidat este ca un șir de biți. Vectori de alte tipuri și structuri pot fi utilizați, în esență, în același mod. Principala proprietate care face ca aceste reprezentări genetice să fie convenabile este că piesele lor sunt ușor aliniabile, datorită dimensiunii lor fixe, ceea ce facilitează operațiunile simple de crossover. Reprezentările de lungime variabilă pot fi de asemenea folosite, dar implementarea încrucișării lor este mai complexă în acest caz.

Odată ce reprezentarea genetică și funcția de fitness sunt definite, un AG continuă să inițializeze o populație de soluții și apoi să o îmbunătățească prin aplicarea repetată a operatorilor de mutație, încrucișat, inversiune și selecție.

În continuare se vor ilustra câteva concepte ale algoritmilor genetici.

Încrucișarea

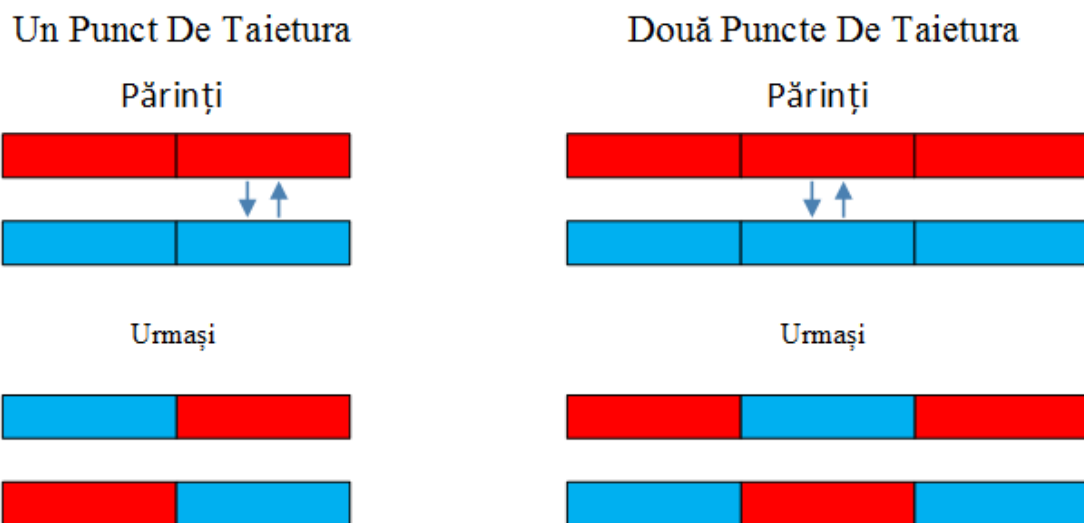


Fig. 3.6 Operațiunea de Încrucișarea în AG.

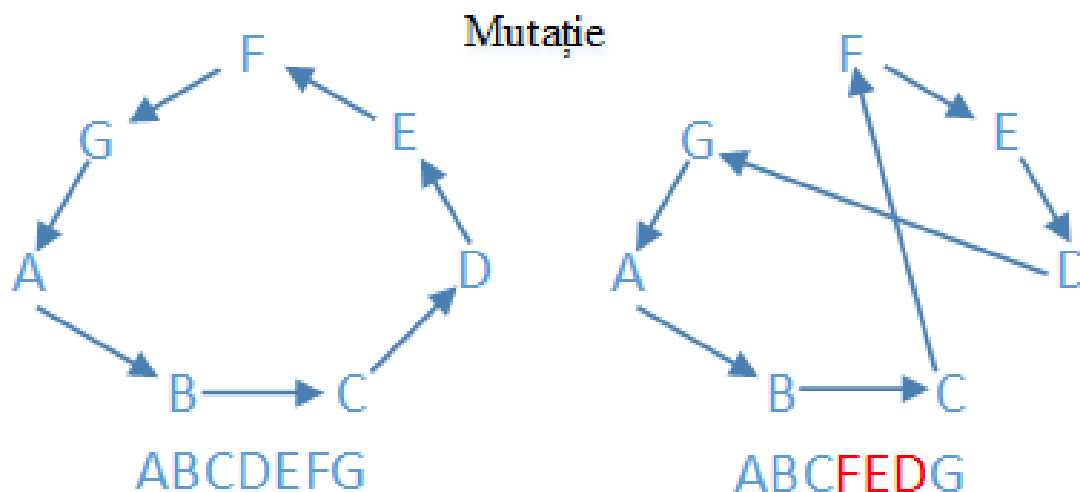
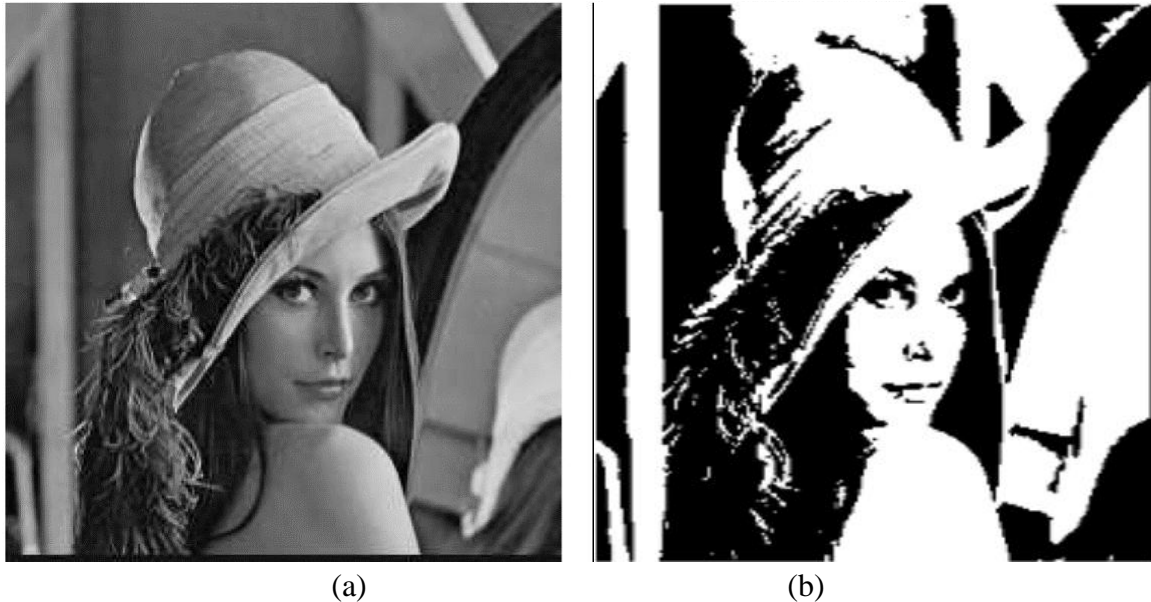


Fig. 3.7 Operațiunea de Mutație în AG

Funcția de fitness a algoritmului genetic maximizează numărul de etichete corect detectate la finalul algoritmului PCNN, în comparație cu referința [4]. Populația este formată din P indivizii. Pentru următoarea generație, descendenți au fost generați prin reproducerea celor mai puternici indivizi și rata de supraviețuire a indivizilor de la o generație la alta este egal cu m . Ca funcție de mutație a fost aleasă o funcție adaptivă. Algoritmul se oprește când ultima generație N este atinsă.

3.6 PCNN pentru segmentarea imaginilor în tonuri de gri

Modelul PCNN sa dovedit a fi foarte potrivit pentru segmentarea imaginilor [11], [12], [13] și [14]. Pentru a obține o performanță de segmentare îmbunătățită pentru imagini la scară gri, ne-am propus să rulăm în paralel două modele PCNN pentru segmentare, una pentru "n" iterații și a doua pentru "m" iterații. Rezultatele acestor două segmentări sunt adunate astfel încât rezultatul obținut accentuează obiectele segmentate și de orice zgomot rămas este stins în fundal. Numărul de iterații, "n" și "m", sunt determinate prin utilizarea unui algoritm genetic (AG). Segmentarea se termină cu trunchierea imaginii rezultate după însumarea celor două module PCNN.



*Fig. 3.8 Algoritmul PCNN clasic aplicat pe imaginea clasica Lena
(a)Imaginea originală; (b) Imaginea segmentată (Pixelii albi reprezinta neuronii activati)*

3.7 PCNN pentru segmentarea imaginilor color

Modelul standard PCNN nu permite segmentarea imaginilor color. Metoda cea mai folosita de segmentare a imaginilor color, folosind modelul PCNN este prin separarea și segmentarea fiecărui spațiu de culoare. Aceasta generează un vector de dimensiune n pentru fiecare pixel al imaginii, unde n este numărul de spații de culoare. Rezultatul final se obține prin selectarea valorii predominante în vector și remiterea acesteia ca rezultat.

Modelul propus de noi rearanjează locațiile pixel, producând o imagine care are toate valorile pixelilor pentru imaginea color conținută într-o singură imagine în tonuri de

gri. Acesta realizează acest lucru prin lărgirea imaginii și punerea valorilor tuturor planurilor componente de culoare, una după alta. De exemplu, într-o imagine RGB, în loc de a avea 3 spații de culoare pixelii sunt reprezentați după cum urmează: RGBRGBR

Această imagine este procesată cu modelul PCNN pentru a obține imaginea binară. Decizia finală se obține dând pixelului final valoarea maximă din grupul de n pixeli.

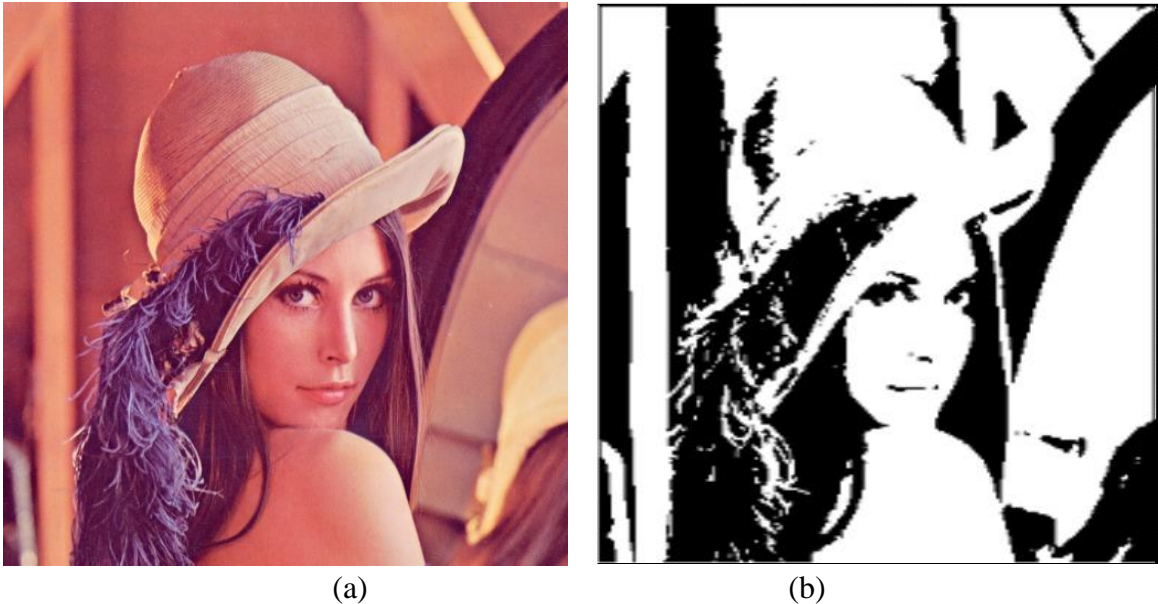


Fig. 3.9 Algoritmul PCNN propus aplicat pe imaginea clasică Lena
(a) Imaginea originală; (b) Imaginea segmentată (Pixelii albi reprezintă neuronii activati)

3.8 Rezultate Experimentale

3.8.1 Baza de Date "The Berkeley Segmentation Dataset and Benchmark"

Am folosit "The Berkeley Segmentation Dataset and Benchmark" pentru a evalua performanțele modelului de segmentare PCNN propus. Această bază de date conține trei seturi de imagini color. Primul set de date este un set de antrenare, al doilea este setul de testare, iar al treilea este setul de validare. Seturile de antrenare și de testare conțin 200 de imagini, iar setul de validare conține 100 de imagini. Din aceste imagini, 8 imagini au fost alese pentru antrenare și 4 imagini au fost selectate pentru validare. Toate imaginile considerate conțin doar două clase de obiecte din imagine.

3.8.2 Parametrii PCNN pentru segmentarea imaginilor în tonuri de gri

Populația este formată din membrii P. Pentru următoarea generație, descendenții au fost generați prin reproducerea celor mai buni indivizi și rata de supraviețuire a indivizilor de la o generație la alta este egal cu m. Funcția de mutație a fost aleasă ca o funcție adaptivă, generând indivizi în funcție de succesul indivizilor din generația anterioară a lui. Condiția de oprire este declanșată atunci când ultima generație, N, este finalizată.

Pentru a identifica setările optime pentru cele două structuri de PCNN, am folosit un algoritm genetic (GA). GA a fost folosit pentru a optimiza un număr de 12 parametri, 6 pentru fiecare dintre cele două rețele PCNN. Acești parametri sunt: n , α_F , α_L , β , V_F și V_L . Parametrii α_E și α_L primesc aceeași valoare numerică, și există o relație similară între parametrii V_E și V_L . Inițializarea rețelei a fost aleasă după cum urmează: matricele F, L, E, U și Y au fost inițializate cu dimensiune imaginii originale și cu valoarea 0 la fiecare poziție, matricele M și W cu dimensiuni egale, de 3×3 , ale căror valori sunt reprezentate în tabelul 3.1. Populația inițială a fost stabilită ca având P = 150 membrii. Pentru reproducere, s-au ales cei mai buni indivizi cu parametrul $n = 0,05$, iar rata de supraviețuire a indivizilor de la o generație la alta este $m = 0,8$. Condiția de oprire a ales ultima generație N = 150.

Această optimizare a fost aplicat unui număr de 8 imagini. Parametrii optimi PCNN sunt prezentate în tabelele 3.1 și 3.2.

Tab. 3.1 Matricele M și W

0.707	1	0.707
1	0	1
0.707	1	0.707

Tab. 3.2 Parametrii PCNN
pentru segmentarea imaginilor în tonuri de gri

	n	α_L	α_F	β	V_L	V_F
PCNN1	8	1.126	9.987	16.592	0.924	63.168
PCNN2	8	6.437	4.468	21.025	0.351	58.670

3.7.3 Parametrii PCNN pentru segmentarea imaginii color

Condițiile inițiale sunt aceleași ca și pentru modelul gri.

Această optimizare a fost aplicată aceluiași 8 imagini utilizate pentru segmentarea imaginilor în tonuri de gri. Parametrii optimi PCNN sunt prezentate în tabelele 3.3 și 3.4.

Tab. 3.3 Matricele M și W

0.707	1	0.707
1	0	1
0.707	1	0.707

Tab. 3.4 Parametrii PCNN
pentru segmentarea imaginilor color

	n	α_L	α_F	β	V_L	V_F
PCNN	34	0.21	43.744	1.316	0.605	10.673

3.8.4 Parametrii PCNN propus pentru segmentarea imaginilor color

Condițiile inițiale sunt aceleași ca și pentru modelul gri.

Optimizarea a fost aplicată pentru aceleași 8 imagini utilizate pentru cele două modele anterioare. Parametrii optimi PCNN sunt prezentate în Tabelele 3.5 și 3.6.

Tab. 3.5 Matricele M și W

0.707	1	0.707
1	0	1
0.707	1	0.707

Tab. 3.6 Parametrii PCNN
pentru segmentarea imaginilor color folosind modelul propus

	n	α_L	α_F	β	V_L	V_F
PCNN	48	0.12	40.294	4.786	0.134	36.938

3.8.5 Performanțe experimentale

Prezentăm în continuare imaginile de antrenare în figurile 3.14 – 3.22, iar rezultatele experimentale în tabelul 3.7.

O dată dovedit că modelul este robust, pasul următor este de a confirma modelul pe imaginile din lotul de validare. Rezultatele obținute pe setul de validare sunt prezentate în tabelul 3.8 și imaginile segmentate corespunzătoare sunt prezentate în figurile 3.10 – 3.13.

Modelul de segmentare PCNN este considerat atât pentru scara gri și, de asemenea,

pentru imagini color. Contribuția noastră pentru segmentare folosind modelului PCNN pe imagini în tonuri de gri constă în rularea în paralel a două structuri de segmentare PCNN, în prima folosind "n" iterații PCNN iar a doua, folosind "m" iterații PCNN. Elementul de noutate al segmentării imaginilor color folosind modelul PCNN consta în a rearanja locațiile pixelilor, producând o imagine care are toate valorile pixelilor din imaginea color într-o singură imagine în tonuri de gri. Se realizează acest desiderat prin lărgirea imaginii și amplasarea pixelilor R, G, B unul după altul. Pentru a optimiza parametrii PCNN (segmentare perfectă), am aplicat algoritmi genetici (AG).

Tab. 3.7 *Precizia Segmentării pentru datele de antrenare*

Imagine	PCNN pentru segmentarea imaginilor în tonuri de gri	PCNN pentru segmentarea imaginilor color	PCNN propus pentru segmentarea imaginilor color	FCM
1(Cal)	90.31%	93.43%	90.34%	40.77%
2(Acropole)	88.15%	72.24%	84.39%	82.45%
3(Tigru)	94.23%	91.04%	94.08%	90.97%
4(Biserica)	90.15%	87.58%	90.88%	85.47%
5(Pantera)	91.67%	87.90%	88.95%	58.16%
6(Înotător)	85.21%	82.84%	86.38%	87.28%
7(Portret)	86.26%	84.45%	85.42%	83.67%
8(Apollo)	89.35%	83.65%	87.58%	85.27%

Tab. 3.8 *Precizia Segmentarea pentru datele de validare*

Imagine	PCNN pentru segmentarea imaginilor în tonuri de gri	PCNN pentru segmentarea imaginilor color	PCNN propus pentru segmentarea imaginilor color	FCM
1(Pisică)	83.07%	77.27%	79.20%	77.13%
2(Cactus)	86.91%	83.77%	85.00%	78.35%
3(Planta)	91.88%	89.66%	94.13%	86.83%
4(Zebra)	83.00%	77.81%	75.31%	66.97%



(a)



(b)



(c)



(d)

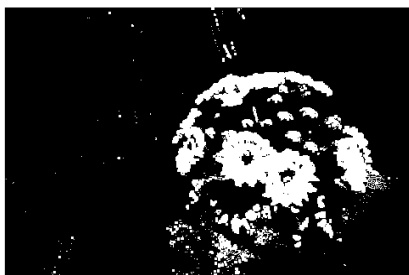


(e)

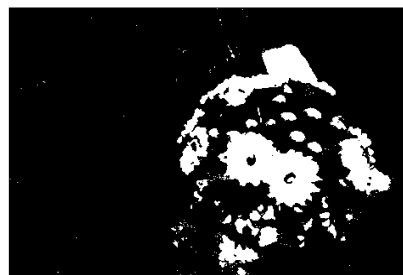
Fig. 3.10 Segmentarea imaginii care conține zebre, din lotul de validare.
(a) Imagine originală; (b) PCNN pentru segmentarea imaginilor în tonuri de gri;
(c) PCNN pentru segmentarea imaginilor color;
(d) PCNN propus pentru segmentarea imaginilor color; (e) referința.



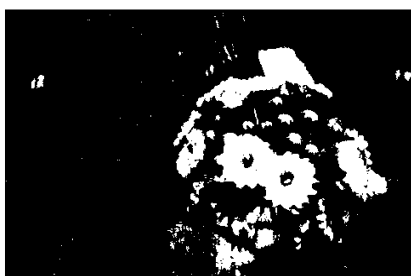
(a)



(b)



(c)



(d)



(e)

Fig. 3.11 Segmentarea imaginii care conține o floare de cactus, din lotul de validare. (a) Imagine originală; (b) PCNN pentru segmentarea imaginilor în tonuri de gri; (c) PCNN pentru segmentarea imaginilor color; (d) PCNN propus pentru segmentarea imaginilor color; (e) referința.



(a)



(b)



(c)



(d)



(e)

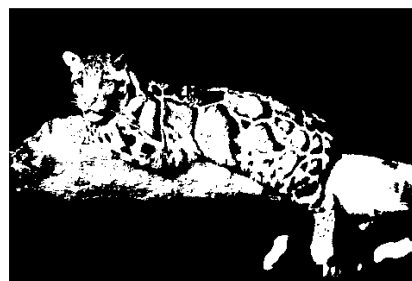
Fig. 3.12 Segmentarea imaginii care conține o ciuperca, din lotul de validare.
(a) Imagine originală; (b) PCNN pentru segmentarea imaginilor în tonuri de gri;
(c) PCNN pentru segmentarea imaginilor color;
(d) PCNN propus pentru segmentarea imaginilor color; (e) referința.



(a)



(b)



(c)



(b)



(c)

Fig. 3.13 Segmentarea imaginii care conține o pisica sălbatică, din lotul de validare. (a) Imagine originală; (b) PCNN pentru segmentarea imaginilor în tonuri de gri; (c) PCNN pentru segmentarea imaginilor color; (d) PCNN propus pentru segmentarea imaginilor color; (e) referința.



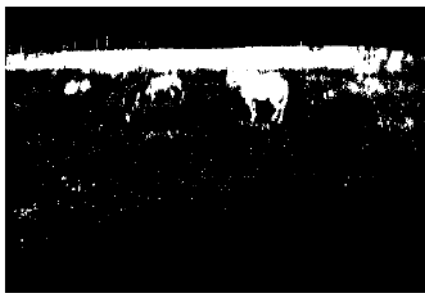
(a)



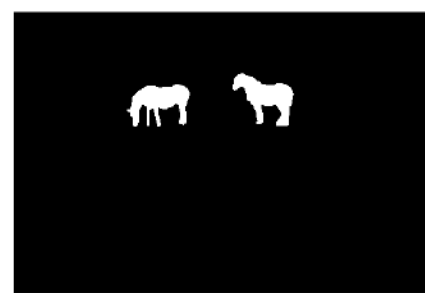
(b)



(c)



(d)



(e)

Fig. 3.14 Segmentarea imaginii care conține cai, din lotul de antrenare.
(a) Imagine originală; (b) PCNN pentru segmentarea imaginilor în tonuri de gri;
(c) PCNN pentru segmentarea imaginilor color;
(d) PCNN propus pentru segmentarea imaginilor color; (e) referința.



(a)



(b)



(c)



(d)



(e)

Fig. 3.15 Segmentarea imaginii de la Acropole, din lotul de antrenare.
(a) Imagine originală; (b) PCNN pentru segmentarea imaginilor în tonuri de gri;
(c) PCNN pentru segmentarea imaginilor color;
(d) PCNN propus pentru segmentarea imaginilor color; (e) referința.



(a)



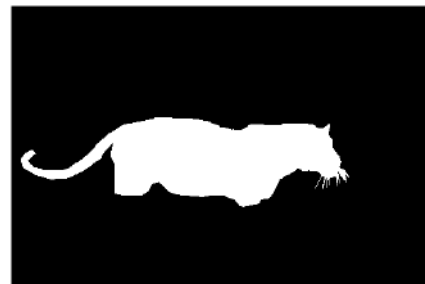
(b)



(c)



(d)



(e)

Fig. 3.16 Segmentarea imaginii care conține un tigru, din lotul de antrenare.
(a) Imagine originală; (b) PCNN pentru segmentarea imaginilor în tonuri de gri;
(c) PCNN pentru segmentarea imaginilor color;
(d) PCNN propus pentru segmentarea imaginilor color; (e) referința.



(a)



(b)



(c)



(d)



(e)

Fig. 3.17 Segmentarea imaginii care conține o biserică, din lotul de antrenare.
(a) Imagine originală; (b) PCNN pentru segmentarea imaginilor în tonuri de gri;
(c) PCNN pentru segmentarea imaginilor color;
(d) PCNN propus pentru segmentarea imaginilor color; (e) referința.



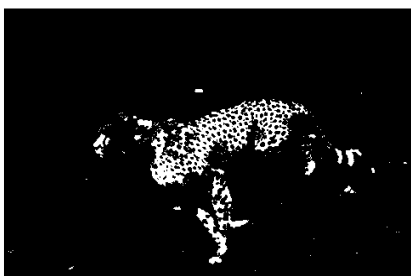
(a)



(b)



(c)



(d)



(e)

Fig. 3.18 Segmentarea imaginii care conține un jaguar, din lotul de antrenare.
(a) Imagine originală; (b) PCNN pentru segmentarea imaginilor în tonuri de gri;
(c) PCNN pentru segmentarea imaginilor color;
(d) PCNN propus pentru segmentarea imaginilor color; (e) referința.



(a)



(b)



(c)



(d)



(e)

Fig. 3.19 Segmentarea imaginii care conține un înotător, din lotul de antrenare.
(a) Imagine originală; (b) PCNN pentru segmentarea imaginilor în tonuri de gri;
(c) PCNN pentru segmentarea imaginilor color;
(d) PCNN propus pentru segmentarea imaginilor color; (e) referința.



(a)



(b)



(c)



(d)

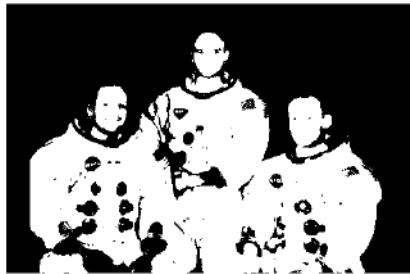


(e)

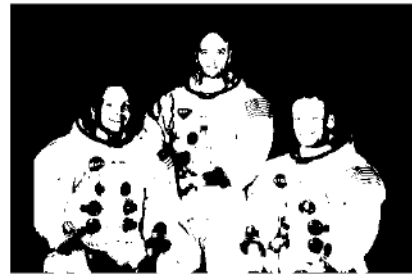
Fig. 3.20 Segmentarea imaginii care conține un portret, din lotul de antrenare.
(a) Imagine originală; (b) PCNN pentru segmentarea imaginilor în tonuri de gri;
(c) PCNN pentru segmentarea imaginilor color;
(d) PCNN propus pentru segmentarea imaginilor color; (e) referința.



(a)



(b)



(c)



(d)



(e)

Fig. 3.20 Segmentarea imaginii Apollo, din lotul de antrenare.
(a) Imagine originală; (b) PCNN pentru segmentarea imaginilor în tonuri de gri;
(c) PCNN pentru segmentarea imaginilor color;
(d) PCNN propus pentru segmentarea imaginilor color; (e) referința.

După cum se observa, atât în tabelele de rezultate prezentate anterior, precum și în seturile de imagini prezentate în detaliu deasupra, modelele PCNN propuse obțin rezultate satisfăcătoare și superioare metodei de referință.

Capitolul 4

4. Recunoașterea plăcutelor de înmatriculare utilizând modelul PCNN

4.1 PCNN pentru segmentarea imaginilor multispectrale

În acest capitol se prezintă o metodă inovativă de recunoaștere a plăcuțelor de înmatriculare (License Plate Recognition-LPR), utilizând un model de segmentare a imaginilor bazat pe rețele neurale cu cuplaj prin impulsuri (Pulse Coupled Neural Network -PCNN) în cascadă cu clasificatorul SVM. Metoda propusă folosește trei etape de prelucrare. Primul pas extrage coordonatele plăcii de înmatriculare din imaginea originală; a doua etapă este metoda de segmentare bazată pe modelul PCNN rezultând astfel o imagine binară care conține doar caracterele plăcuței de înmatriculare; al treilea pas constă dintr-un sistem de recunoaștere optică (Optical Character Recognition - OCR) bazat pe algoritmul SVM aplicat fiecărei litere extrase. Acest model a fost testat pe o bază de date mare deținută de Societatea Softrust, care conține 10 000 de exemple (probe). Rezultatele experimentale

referitoare la scorul de recunoaștere ilustrează avantajele evidente ale modelului PCNN prin comparație cu metodele de segmentare a imaginilor SOM și Mean Shift.

Modelul de segmentare a imaginilor PCNN a fost proiectat inițial pentru a lucra pe imagini în niveluri de gri [4], [11], [12], și [14]. În literatura de specialitate există, de asemenea, diverse opțiuni pentru a modifica algoritmul, astfel încât să acționeze asupra imagisticii multispectrale, așa cum se arată în [4], [12], [13] și [15]. De exemplu, în cazul imaginilor color, segmentarea folosind modelul PCNN se realizează prin separarea și segmentarea fiecărui spațiu de culoare.

Structura de intrare clasică a neuronului PCNN este compusă din valorile pixelilor cu aceleași coordonate ca cele ale neuronului și valorile vecinătății acestei coordonate.

În versiunea propusă de segmentare a imaginii PCNN vom modifica structura de intrare a neuronului, în așa fel încât să conțină valori, nu numai dintr-un singur plan spectral, dar din toate planurile spectrale.

Acest lucru este realizat prin alegerea unui plan spectral de referință, care contribuie valoarea lui S_{ij} din ecuația (4.1), și în loc de valorile vecinătăților vom pune la dispoziția neuronului informații de la celelalte plane spectrale p .

$$S_{ij} = \left\{ \begin{array}{ccc} I_{1_{i-1,j}}, & I_{2_{i-1,j}}, & I_{3_{i-1,j}}, \\ I_{1_{i,j}}, & I_{2_{i,j}}, & I_{3_{i,j}}, \\ I_{1_{i+1,j}}, & I_{2_{i+1,j}}, & I_{3_{i+1,j}} \end{array} \right\} \quad (4.1)$$

unde S_{ij} reprezintă structura de intrare a neuronului PCNN, $I_{q_{i,j}}$ reprezintă valoarea pixelului de la poziția i, j din planul spectral q . în acest caz S_{ij} este $I_{2_{i,j}}$ [21].

Faptul ca modelul PCNN standard nu are un mecanism intern de oprire, poate deveni o problemă importantă, deoarece este imposibil să se decidă la care iterație rezultatul modelului este favorabil.

În această situație, din cauza faptului că unele informații sunt cunoscute din prealabil, referitor la imagine, putem folosi aceste cunoștințe pentru a deduce o condiție de oprire. Luând în considerare că mărimea unei litere este cunoscută, următoarele ecuații pot fi deduse:

$$D = \begin{cases} D + 1 & \text{daca } Y_K[n] = 1 \\ D & \text{daca nu} \end{cases} \quad (4.2)$$

$$F = \begin{cases} 1 & \text{daca } D > r \text{ si } D < R \\ 0 & \text{daca nu} \end{cases} \quad (4.3)$$

unde D este numărul de pixeli activați, r și R sunt dimensiunile minime și maxime a literelor.

Metoda se numește în cascadă, deoarece se parcurg în cascadă diferite seturi de parametri pentru modelul PCNN, până când se îndeplinește condiția de oprire.

Pe scurt, ne propunem două completări principale ale modelului PCNN clasic: un nou vector de alimentare pentru modelul PCNN și o condiție robustă de oprire. Aceste două adaptări permit modelului PCNN să ajungă la performanțe bune în segmentarea imaginilor, fără nici o intervenție umană.

4.2 SVM pentru recunoașterea caracterelor

Algoritmul SVM inițial a fost inventat de către Vladimir N. Vapnik și varianta finală a fost propusă de Vapnik și Corinna Cortes în 1995 [16][17].

Algoritmii de tip SVM sunt bazati pe plane de decizie (hiperplane de separare), acestea definind frontierele claselor. Planele (hiperplanele) de decizie separă seturile de date aferente diferitelor clase[59][71].

În Figura 4.1 se prezintă conceptul de baza al algoritmilor SVM. În partea stanga se vad obiectele originale rearanjate, folosind funcții matematice numite nuclee (kernels). În partea dreaptă a imaginii obiectele sunt linear separabile, astfel construirea unei hipersuprafețe neliniare (în particular, o curba) pentru a le separa nu mai este necesară, este suficientă construirea unui plan "optim" care să separe bilele albastre de cele roșii.

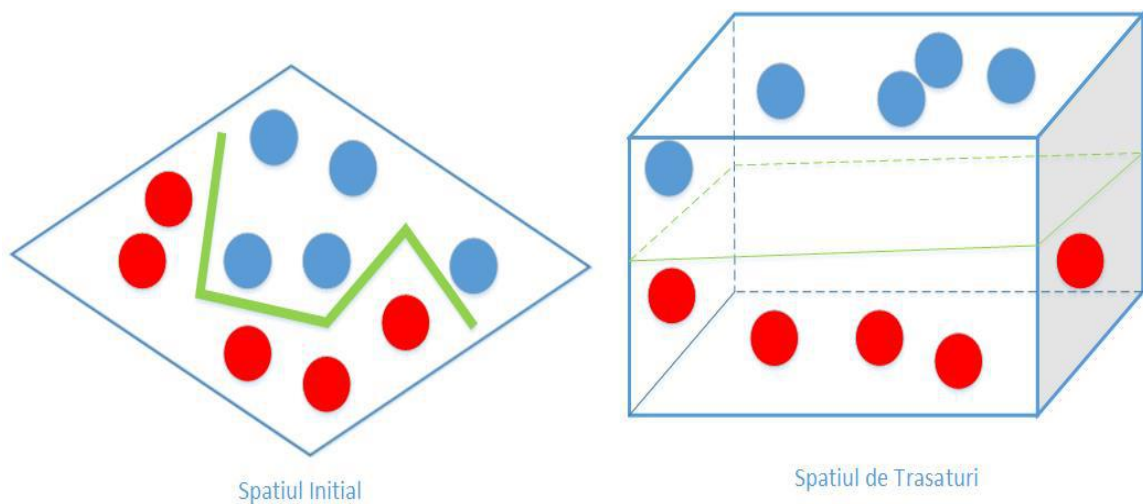


Fig. 4.1 Reprezentarea datelor într-un spațiu superior

Problema clasificării multiple se poate reduce la cazul clasificării a două clase, acest lucru nerestrângând generalitatea. Obiectivul este de a separa două clase, folosind exemplele pe care le avem la dispoziție și generând o funcție de separare. Scopul final este

obținerea unei funcții de separare a claselor cât mai exacte, care funcționează bine și pe exemple noi, asta înseamnă că funcția generalizează bine.

Deși pot exista o multitudine de clasificatori liniari în spațiul transformat, de obicei numai unul dintre ei maximizează marginile (distanța dintre planul de separare și punctele din fiecare clasă). Clasificatorul liniar care îndeplinește această condiție se numește hiperplanul optim de separare. Intuitiv, această separare optimă ajută mult la problema generalizării. Punctele aflate în vecinătatea zonei de frontieră constituie *vectorii suport*.

4.2.1 Hiperplanul de separare

Dacă considerăm problema separării setului de vectori de antrenare, care fac parte din două clase, notăm cu x_i datele de intrare și y_i clasele aferente

$$F = \{(x_1, y_1), \dots, (x_l, y_l)\}, x \in \mathbb{R}, y \in \{-1, 1\}. \quad (4.4)$$

Se dorește obținerea unui hiperplan optim de ecuație

$$\langle w, x \rangle + b = 0 \quad (4.5)$$

Se dorește ca acest hiperplan să separe exact vectorii de antrenare ai celor două clase, maximizând distanța între vectorul cel mai apropiat de hiperplan și hiperplan. Fără a restrânge generalitatea, considerăm un hiperplan canonic [16], astfel încât parametrii w, b să satisfacă relația

$$\min_i |\langle w, x_i \rangle + b| = 1. \quad (4.6)$$

Astfel, norma vectorului trebuie să fie proporțională cu inversa distanței între cel mai apropiat vector din setul de date și hiperplan. Hiperplanul de separare în formă canonică trebuie să satisfacă relațiile

$$y_i[\langle w, x_i \rangle + b] \geq 1, i = 1 \dots l \quad (4.7)$$

Distanța $D(w, b; x)$ între vectorul x și hiperplanul (w, b) se calculează prin expresia

$$D(w, b; x) = \frac{|\langle w, x_i \rangle + b|}{\|w\|} \quad (4.8)$$

Hiperplanul optim se obține prin maximizarea marginii, conform ecuației (4.8).
 Acesta maximizare este echivalentă cu minimizarea relației (4.9)

$$\varphi(w) = \frac{1}{2} \|w\|. \quad (4.9)$$

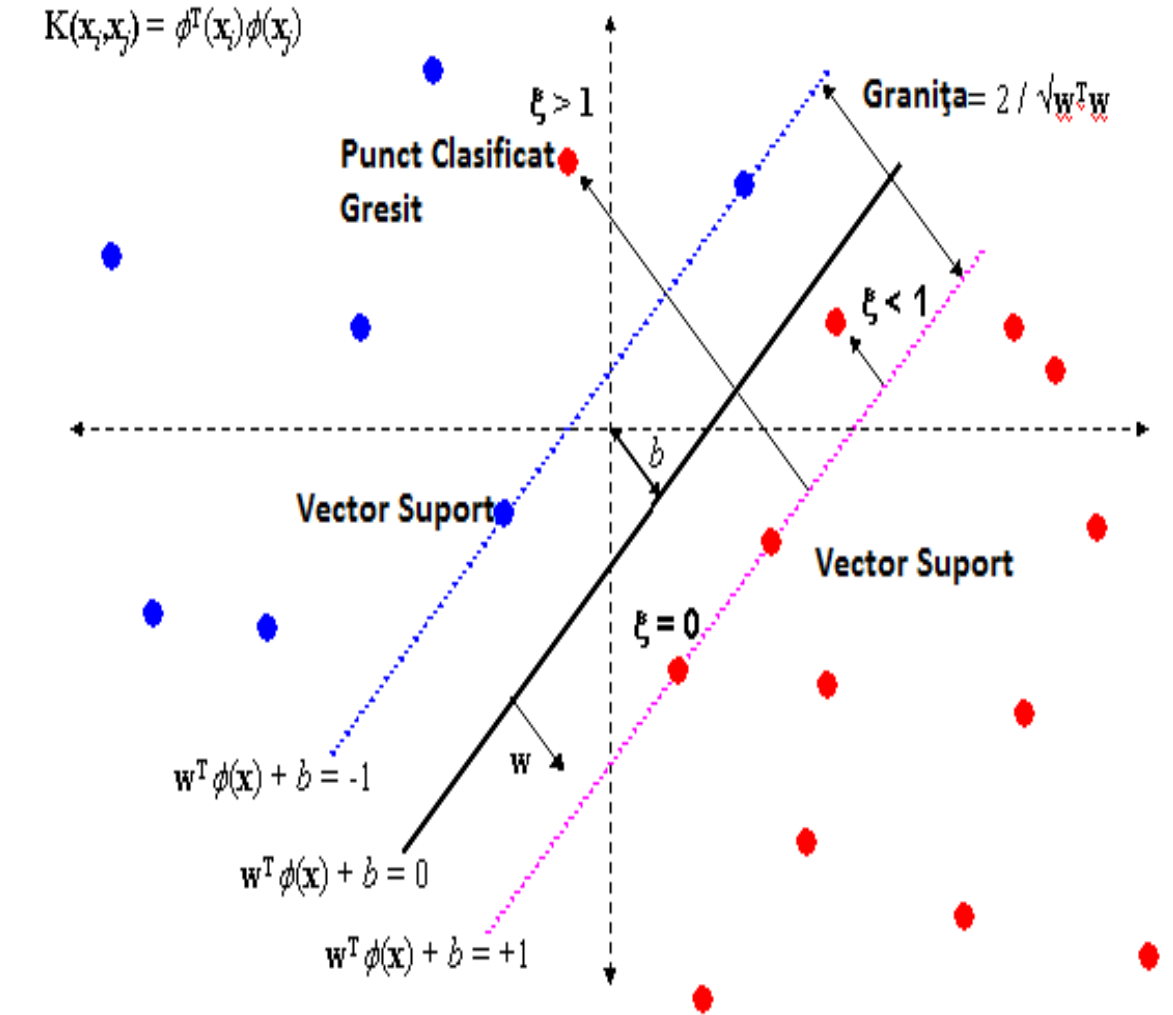


Fig. 4.2 Hiperplanul și vectorii săi suport

Pentru rezolvare se foloseste metoda multiplicatorilor lui Lagrange. Se introduc multiplicatorii $\alpha_i \geq 0, i = 1, 2 \dots l$. Se va minimiza urmatoarea functie

$$\varphi(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^l \alpha_i \{y_i [\langle w, x_i \rangle + b] - 1\}, \quad (4.10)$$

unde $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_l]$.

Derivatele lui $\varphi(w, b, \alpha)$ in raport cu w si b trebuie sa fie nule.

In ipoteza ca cele doua clase sunt liniar separabile, rezultatul final conduce la regula clasificarii vectorilor in functie de semnul expresiei

$$y(x) = \sum_{i=1}^l \alpha_i y_i Ker(x, x_i) + b. \quad (4.11)$$

Interesant este ca vectorii care au $\alpha_i = 0$ nu contribuie niciodata la suma evaluata in relatia (4.11), cu alte cuvinte ei nu au niciun rol in procesul de clasificare. Ceilalti vectori, pentru care $\alpha_i > 0$, contribuie mereu in procesul de clasificare si poarta denumirea de *vectori suport*. Se observa ca multimea vectorilor suport este inclusa in multimea initiala de l vectori.

Se observa ca valoarea lui b din ecuatia nu a fost inca stabilita. O modalitate eleganta de a o afla este propusa de Bishop

$$b = \frac{1}{N_s} \sum_{i \in S} (t_i - \sum_{i=1}^l \alpha_i y_i Ker(x, x_i)) \quad (4.12)$$

unde cu S s-a notat multimea vectorilor suport.

4.3 Baza de date

Baza de date folosită este a Companiei Softrust, la a carei constructii am contribuit. Aceasta a fost utilizată pentru a evalua performanța modelului PCNN propus pentru recunoașterea numărului de înmatriculare. Această bază de date este formată din 10.000 de evenimente de autoturisme, preluate din fața vehiculelor, fiecare eveniment constând din 20 de imagini. Aceste imagini au fost realizate cu o camera specială, conceputa pentru License Plate Recognition (LPR). Aceste camere au fie rezoluție CIF (352 x 288 pixeli), fie 2 CIF (704 x 288 pixeli) și sunt echipate cu un tun IR (infraroșu) și filtrul cu infraroșu. Senzorul

camerei este sensibil la trei lungimi de undă, similar cu senzor RGB, dar lungimile de undă nu sunt identice.

Pentru a înțelege modul în care arată aceste imagini vom prezenta câteva exemple (plăcuțe de înmatriculare curate și "murdare" pe timp de noapte și în timpul zilei).



Fig. 4.3 Exemple Baza de Date; (a) Număr curat pe timp de zi;
(b) Număr murdar pe timp de zi; (c) Număr curat pe timp de noapte;
(d) Număr murdar pe timp de noapte;

4.4 Segmentarea plăcuțelor de înmatriculare

În primul rând, vom prezenta în continuare modul în care segmentarea arată pe o plăcuță de înmatriculare ideală, în timpul zilei, atunci când imaginea nu este afectată de zgomot și placa este curată, figura 4.4. Aceste imagini reprezintă un rezultat brut al metodelor avute în vedere.

În continuare, rezultatele obținute pe o imagine capturată pe timp de noapte, care nu este afectată de zgomot va fi prezentat în figura 4.5.

În cele ce urmează, rezultatele obținute pe o imagine capturată în timpul zilei, care este afectat de zgomot va fi prezentat în figura 4.6.



Fig. 4.4 Număr de înmatriculare curat pe timp de zi;
(a) Imaginea originală; (b) Rezultat PCNN; (c) Rezultat SOM;
(d) Rezultat Mean Shift;

După cum se poate observa toate modelele reușesc sa obțină separare foarte buna între litere și în același timp sa separe fiecare litera de fundal.



Fig. 4.5 Număr de înmatriculare curat pe timp de noapte;
(a) Imaginea originală; (b) Rezultat PCNN; (c) Rezultat SOM;
(d) Rezultat Mean Shift;

În acest caz separarea fata de fundal rămâne perfectă la toate metodele de segmentare însă nici unul dintre algoritmi studiați nu reușește sa separe primele doua

acestea rămânând lipite. Aceasta problema este adresata in postprocesare, folosindu-se metode geometrice.



Fig. 4.6 Număr de înmatriculare murdar pe timp de zi;
(a) Imaginea originală; (b) Rezultat PCNN; (c) Rezultat SOM;
(d) Rezultat Mean Shift;

In acest caz se poate observa ca literele nu mai sunt bine definite si exista situatii in care fundalul se confundă cu zona de interes.

După acest proces, toate imaginile au fost prelucrate în continuare, astfel încât rezultatul final este o imagine binară care conține numai literele și numerele conținute în plăcuța de înmatriculare [18], [19], [20], așa cum se arată în figura 4.7.



Fig. 4.7 Număr de înmatriculare segmentat;
(a) Imaginea originală; (b) Rezultat final segmentare;

Acest proces nu este simplu si necesită o multitudine de cunoștințe apriori despre formatul standard al plăcuțelor de înmatriculare dar. Printre acestea, cele mai notabile sunt următoarele:

- Posibilitatea prezentei unui abțibild rotund între literele de interes. Acesta este folosit ca dovada a trecerii cu succes de inspecției tehnice periodice a autovehiculului. Proprietățile sale fundamentale sunt: forma rotunda si dimensiunea de aproximativ o treime din dimensiunea unei litere.

- Înainte de începerea numărului de înmatriculare exista o regiune care denota tara de origine a autovehiculului. Aceasta se poate elimina. Proprietatea ei fundamentală este forma dreptunghiulara.

4.5 Antrenarea clasificatorului

SVM pe care se bazează clasificarea a fost antrenat pe 35 de clase, 25 litere și 10 simboluri numerice. Sunt doar 25 de clase pentru litere deoarece simbolul literei "O" și pentru numărul "0" sunt instruiți în aceeași clasă.

S-a folosit un număr de 400 de exemple pentru fiecare clasă în faza de antrenare și un număr de 300 de exemple pentru fiecare clasă în etapa de validare. În tabelul 4.1 rezultatele de formare sunt ilustrate.

Tab. 4.1 Acuratețea recunoașterii caracterelor cu modelul SVM in faza de antrenare

	Kernel: liniar; Dimensiunea imaginii: 20x28	Kernel: liniar; Dimensiunea imaginii: 15x21	Kernel: liniar; Dimensiunea imaginii: 10x14
Rezultat	92.30%	95.24%	96.08%
	Kernel: polinomial de gradul 2; Dimensiunea imaginii: 20x28	Kernel: polinomial de gradul 2; Dimensiunea imaginii: 15x21	Kernel: polinomial de gradul 2; Dimensiunea imaginii: 10x14
Rezultat	88.27%	93.87%	95.42%

Dimensiunile imaginilor testate în etapa de antrenare nu au fost alese la întâmplare. Aceste dimensiuni au fost alese astfel încât să păstreze proporțiile inițiale ale literelor. După cum se vede clar în tabelul 4.1, cel mai bun rezultat a fost obținut de nucleu; liniar, care a fost antrenat cu imagini de dimensiune mică, 10 x 14 pixeli.

4.6 Rezultate experimentale

Acest sistem a fost testat pe baza de date mare de 10000 x 20 imagini color.

Criteriile de evaluare au fost următoarele:

Procent de plăcuțe de înmatriculare recunoscute în mod corect - acesta reprezintă situația în care întreg sistemul funcționează. Procesul de segmentare și procesul post de segmentare a obținut literele individuale și OCR returnat numărul corect de înmatriculare.

Numărul de plăcuțe de înmatriculare recunoscute în mod incorect - acesta reprezintă numărul situațiilor în care OCR nu a returnat rezultatul corect. Deoarece rata de recunoaștere OCR este aproape perfectă, acest rezultat este cel mai probabil din cauza segmentare incorecte.

Numărul de plăcuțe de înmatriculare complet nerecunoscute - aceasta reprezintă numărul de cazuri în care segmentarea a eșuat complet, iar OCR nu a returnat nici un rezultat.

Tab. 4.2 Acuratețea modelului LPR

10000 evenimente	PCNN	SOM	Mean Shift
Număr de plăcuțe de înmatriculare complet nerecunoscute	117	322	497
Număr de plăcuțe de înmatriculare nerecunoscute	530	1791	3381
Procentul de plăcuțe de înmatriculare recunoscute corect	93.60%	78.90%	61.30%

Acest capitol a arătat că modelul de segmentare a imaginii PCNN propus, urmat de clasificarea SVM, se comportă ca un sistem de recunoaștere a plăcutelor de înmatriculare.

Metoda de oprire în cascadă asigură faptul că modelul PCNN oprește exact când trebuie. Acest lucru asigură rezultate consistente de înaltă precizie. Metoda care permite modelului să ruleze pe imagini cu mai mult de un spațiu de culoare asigură că neuron PCNN are suficiente informații pentru a activa în mod optim. Tehnicile de segmentare a imaginilor considerate utilizate ca referință pentru evaluarea metodei noastre propuse au arătat în mod clar avantajul modelului propus: SOM - precizie 78,90%, Shift medie - precizie de 61,30%, iar PCNN a propus pe baza modelului - 93,60% acuratețe.

Capitolul 5

5. Recunoașterea vehiculelor militare în imagini SAR cu tehnici neuronale

5.1 Introducere

Pe timp de pace sau de război, soarta a milioane de oameni poate să depindă de analiza imaginilor colectate de radar sau alți senzori de imagistică. Aceste imagini ar putea dezvălui un adevăr semnificativ, la fel ca în 1962, când un avion U-2 a realizat fotografii ale rachetelor sovietice din Cuba [67]. În domeniile de inteligență militară și de recunoaștere militară, cunoașterea reprezintă putere și o modalitate de a obține cunoaștere este să le obținem din aer sau din spațiu. Dar nu este atât de simplu. Contramăsuri inamice, perturbări atmosferice naturale și defecte tehnice complică sarcina. Chiar și atunci când urmărirea inamicului se desfășoară în cele mai bune condiții, nu este ușor să înțelegem ceea ce vedem.

Acest capitol prezintă o abordare nouă, folosind rețele neuronale, de recunoaștere automată a țintei (ATR) în imagini aeriene radar cu apertură sintetică (SAR); această metodă este aplicată pentru identificarea vehiculelor terestre militare [29], [41] Algoritmul propus ATR constă într-o cascadă de prelucrare cu următoarele etape: (a) de detectare a obiectelor cu ajutorul unui modul de Pulse-Coupled Neural Network (PCNN); (b) un prim

modul de selecție a caracteristicilor utilizând filtrarea Gabor (GF); (c) un al doilea modul de selecție a caracteristicilor utilizând analiza componentelor principale (PCA); (d) un clasificator Support Vector Machine (SVM) îmbunătățit prin utilizarea de date generate artificial (Virtual Training Data Generation -VTDG) cu Concurrent Self-Organization Maps (CSOM) propus de Neagoe et al. Modelul propus a fost aplicat pentru recunoașterea a trei clase de vehicule militare ale fostei Uniuni Sovietice, reprezentate în setul de imagini ale bazei de date publice de presă MSTAR, care conține 2987 imagini. Rezultatele experimentale au confirmat eficiența metodei, ceea ce duce la o rată totală de succes de 97.36%.

Pe timp de pace sau de război, soarta a milioane de oameni poate depinde de analiza imaginilor colectate de radar sau alți senzori de imagistică. Aceste imagini ar putea dezvălui un adevăr semnificativ, la fel ca în 1962, când fotografiile obținute cu ajutorul avionului U-2 au dezvăluit rachetele sovietice din Cuba [22]. În inteligența militară și de recunoaștere militară, cunoașterea este putere și o modalitate de a obține informații este să le obții din aer sau din spațiu. Dar nu este atât de simplu. Contramăsurile inamice, perturbările atmosferice naturale și tehnice complică sarcina. Chiar și atunci când inamicul este urmărit în cele mai bune condiții, nu este simplu să înțelegem ceea ce vedem. Progresele recente încearcă să substituie interpretării de imagini prin recunoașterea automată a țintei (ATR) bazată pe inteligență artificială. Obiectele țintă sunt adesea vehicule militare ca cele arătate în figura 5.1 luate în considerare pentru baza de date de eliberare publică MSTAR.



Fig. 5.1 Trei vehicule militare ale fostei Uniuni Sovietice:
(a) BMP 2 (vehicul de luptă - infanterie);
(b) BTR 70 (purtător de personal blindat); (c) T72 (tanc).

5.2 Algoritmul propus

Schema logică a modelului ATR propus este prezentată în figura 5.2.

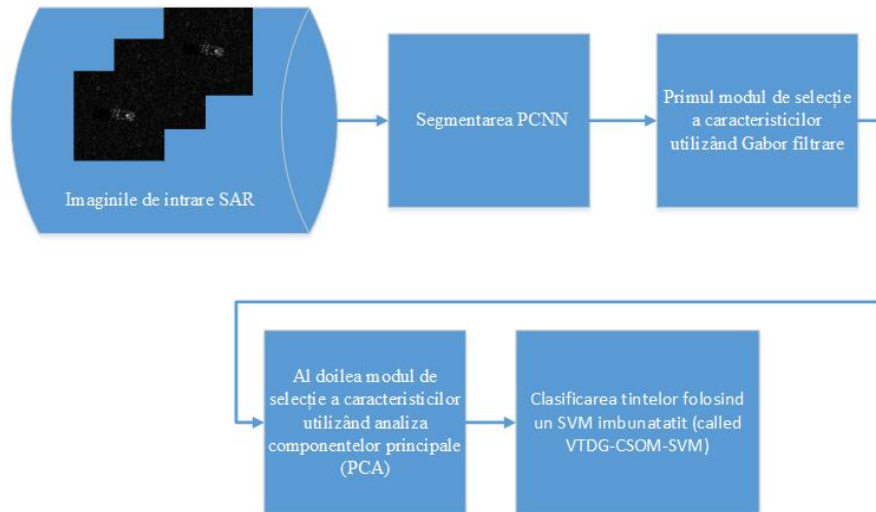


Fig. 5.2 Modelul ATR propus bazat pe rețele neuronale.

Modelul PCNN s-a dovedit a fi foarte potrivit pentru segmentarea imaginilor după cum se arată în [2], [4], [11] și [14]. Pentru a obține o performanță îmbunătățită de segmentare, ne-am propus să rulăm în paralel două modele de segmentare PCNN, una pentru n iterații și în al doilea pentru m iterații (vezi figura 5.2). Rezultatele acestor două segmentări sunt adunate astfel rezultatul final accentuează obiectele și scade influența zgomotului. O relație bună între n și m s-a dovedit a fi $m = n + 3$. Segmentarea este terminată de un filtru „trece sus” pe imaginea de ieșire PCNN.

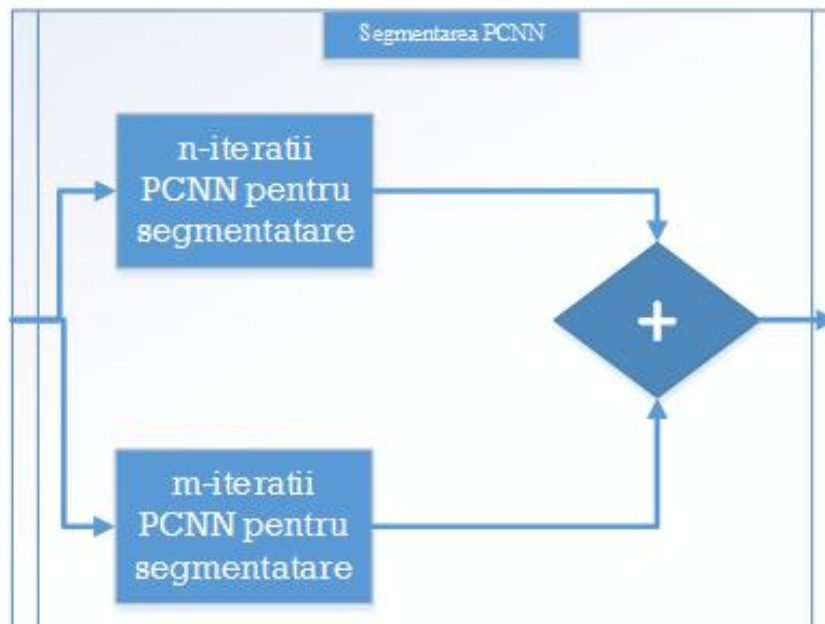


Fig. 5.3 Segmentarea imaginilor folosind combinația celor două modele PCNN.

5.3 Filtrarea Gabor

Filtrele Gabor sunt filtre bandpass care sunt utilizate în procesul de prelucrare a imaginilor pentru extragerea caracteristicilor, analiza texturii [23], și estimarea disparității stereo [23-30]. Răspunsul acestor filtre este creat prin multiplicarea unei funcții Gauss cu o oscilație complexă. Gabor [31] a arătat că aceste funcții elementare minimizează spațiul de incertitudine. Prin extinderea acestor funcții la două dimensiuni este posibilă crearea unor filtre care sunt selective pentru orientare [32][33]. În anumite condiții, faza răspunsului filtrelor Gabor este de aproximativ liniară. Această proprietate este exploatată prin abordări stereo care utilizează diferența de fază dintre răspunsurile de filtrare din stânga și din dreapta pentru a estima diferențele în imaginile stereo.

Fie $\mathbf{x} = [x_1 x_2]^T$ coordonatele imaginii. Răspunsul unui filtru Gabor $g(\mathbf{x})$ este dat de:

$$g_{mn}(\mathbf{x}) = \frac{1}{2\pi a_n b_n} e^{-\frac{1}{2}\mathbf{x}^T A_{mn} \mathbf{x}} e^{jk_{0mn}^T \mathbf{x}} \quad (5.1)$$

Aici, matricea A determină lățimea de bandă și orientarea filtrului.

$$A_{mn} = \begin{bmatrix} \cos \varphi_m & -\sin \varphi_m \\ \sin \varphi_m & \cos \varphi_m \end{bmatrix} \begin{bmatrix} a_n^{-2} & 0 \\ 0 & b_n^{-2} \end{bmatrix} \begin{bmatrix} \cos \varphi_m & \sin \varphi_m \\ -\sin \varphi_m & \cos \varphi_m \end{bmatrix} \quad (5.2)$$

În cazul în care vectorul frecvenței de modulație k_0 este în aceeași direcție ca și una dintre axe, cu

$$k_{0mn} = k_{0n} \begin{bmatrix} \cos \varphi_m \\ \sin \varphi_m \end{bmatrix} \quad (5.3)$$

atunci partea reală și partea imaginară a răspunsului filtrului arată în felul următor:

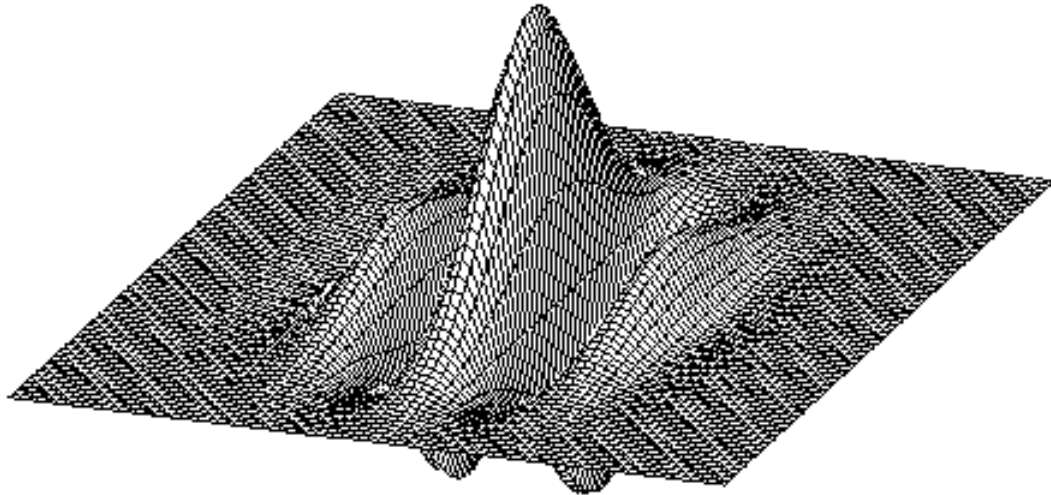


Fig. 5.4. Partea reală a răspunsului filtrului Gabor[33].

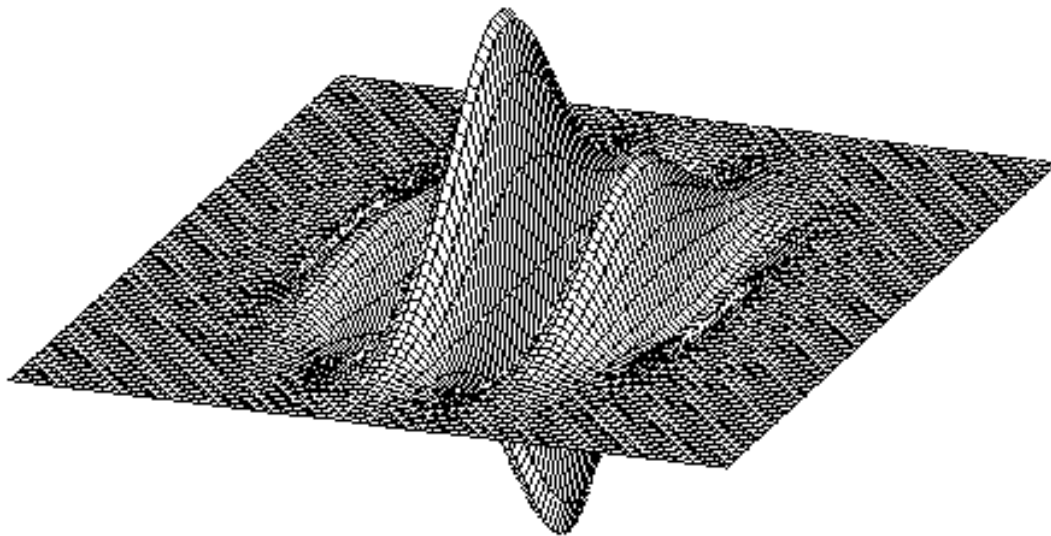


Fig. 5.5 Partea imaginară a răspunsului filtrului Gabor[33].

Funcția de transfer $G(k)$ a unui filtru Gabor (transformata Fourier a răspunsului la impuls) este dată de:

$$G_{mn}(k) = e^{-\frac{1}{2}(k-k_{0mn})^T (A_{mn}^{-1})^T (k-k_{0mn})} \quad (5.4)$$

unde $k = [k_1 k_2]^T$ este frecvența spațială. Imaginea se poate filtra cu un set de N filtre Gabor cu diferite lățimi de bandă și frecvențe de modulare. Dacă frecvențele de modulare sunt date de

$$k_{0_n} = \frac{\pi}{2^{n+1}} ; n \in [0 \dots N - 1] \quad (5.5)$$

și lățimea de bandă relativă este aleasă să fie constantă pentru toate filtrele.

În acest caz filtrul Gabor acceptă imaginile segmentate de modelul PCNN și oferă un vector coloană de ieșire, care este compus din caracteristicile Gabor. Vectorii de caracteristici sunt normalizați variației medii. Vectorul de ieșire GF are o lungime $(m * n * u * N) / (d1 * d2)$, unde m și n reprezintă înălțimea și lățimea imaginii, u este numărul de scala, N este numărul de orientări, $d1$ factorul de prelevare de probe de-a lungul rândurilor și $d2$ este factorul de prelevare a probelor de-a lungul coloanelor.

5.4 Analiza Componentelor Principale (PCA)

Analiza Componentelor Principale (PCA) este o procedură statistică care utilizează o transformare ortogonală pentru a converti un set de variabile, posibil corelate, într-un set de valori liniar necorelate numite componente principale. Numărul de componente principale este mai mic sau egal cu numărul original de variabile. Această transformare este definită în așa fel încât prima componentă principală are cea mai mare varianță posibilă (adică, reprezintă la fel de mult din variabilitatea datelor care este posibil), și fiecare componentă ulterioară, la rândul său, are cea mai mare variația posibilă sub constrângerea că este ortogonală la componentele anterioare. Vectorii rezultați sunt o un set de date necorelate. PCA este sensibil la scalarea relativă a variabilelor originale.

PCA a fost creat în 1901 de Karl Pearson, [34], ca un analog al teoremei axei principale din mecanică; a fost dezvoltat ulterior, în mod independent (și numit) de Harold Hotelling în 1930 [35], [36]. În funcție de domeniul de aplicare, acesta este numit, de asemenea, transformarea discretă Kosambi-Karhunen-Loeve (KLT) în procesarea de semnal, transformarea Hotelling în mai multe variații de control al calității, descompunerea ortogonală corespunzătoare (POD) în inginerie mecanică, descompunerea la valoare singulară (SVD) (Golub și Van Loan, 1983), descompunerea valorii proprii (EVD) în algebra liniară, și altele.

PCA este utilizat în principal ca un instrument de analiză a datelor și pentru a face modele predictive. PCA se poate face prin descompunerea valorilor proprii a unei matrice de covarianță (sau corelare) sau descompunerea la valoare singulară a unei matrice de date, de obicei, după mediere de centrare la zero (și normalizare).

Dacă se considera o matrice de date, X , în care fiecare dintre cele n rânduri reprezintă o repetare diferită a experimentului, iar fiecare dintre coloanele p se dă caracteristică (de exemplu, rezultatele de la un anumit senzor).

Se considera X , un vector stocastic, n -dimensional. Cautăm o transformare liniară ortogonală, cu ajutorul careia să reprezentăm optim vectorul X , în raport cu principiul erorii medii pătratice minime.

Considerăm Q , ca fiind transformarea căutăată:

$$Q = (\Phi_1, \dots, \Phi_n)^T \quad (5.6)$$

unde Φ_n sunt vectori n -dimensionali, încă nedeterminați.

Presupunem de asemenea că vectorii coloană ai lui Q formează un sistem ortonormat, adică:

$$\Phi_i^T * \Phi_j = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \quad (5.7)$$

Astfel, toți vectorii X se devin:

$$Y = QX = (y_1, y_2, \dots, y_n)^T \quad (5.8)$$

cu $y_i = \Phi_i^T X$.

Relațiile (5.7) și (5.8) conduc la $Q^T \cdot Q = Q \cdot Q^T = I_n$ unde I_n este matricea unitate $n \times n$. De aici obținem:

$$X = Q^T Y = \sum_{i=1}^n y_i \Phi_i \quad (5.9)$$

Deci vom reține doar $m < n$ elemente ale lui Y , celelalte $n - m$ elemente fiind înlocuite cu constantele selectate anterior b_i , și se estimează X că:

$$\hat{X}(m) = \sum_{i=1}^m y_i \Phi_i + \sum_{i=m+1}^n b_i \Phi_i \quad (5.10)$$

Eroarea asociată estimării se exprimă în ecuația 5.11:

$$\Delta X(m) = X - \hat{X}(m) = \sum_{i=m+1}^n (y_i - b_i) \Phi_i \quad (5.11)$$

Se considera criteriul erorii pătratice medii:

$$\overline{\varepsilon^2}(m) = E\{\|\Delta X(m)\|^2\} = E\left\{\sum_{i=m+1}^n \sum_{j=m+1}^n (y_i - b_i)(y_j - b_j)\Phi_i^T \Phi_j\right\} \quad (5.12)$$

care, ținând cont de relația (5.7), devine $\overline{\varepsilon^2}(m) = E\{(y_i - b_i)^2\}$. Funcția eroare $\overline{\varepsilon^2}(m)$ se minimizează prin alegerea adecvata a lui Φ_i și b_i . Optimizarea se realizează în două etape:

- A. Minimizarea $\overline{\varepsilon^2}(m)$ în raport cu b_i , punând condiția necesară $\frac{\partial}{\partial b_i} E\{(y_i - b_i)^2\} = -2[E\{y_i\} - b_i] = 0$, de unde, conform relației (5.8), deducem $b_i = E\{y_i\} = \Phi_i^T E(X)$ cu $1 \leq i \leq n$. În acest caz, eroarea pătratică medie este

$$\overline{\varepsilon^2}(m) = \sum_{i=1+m}^n \Phi_i^T \Sigma_x \Phi_i \quad (5.13)$$

unde $E\{\cdot\}$ este media statistica.

- B. Optimizarea vectorilor Φ_i se realizează prin minimizarea $\overline{\varepsilon^2}(m)$ în raport cu Φ_i , punând și condițiile de legatură $\Phi_i^T \Phi_i = 1$ pentru $1 \leq i \leq n$. Se utilizează metoda multiplicatorilor Lagrange, și se obține:

$$\left[\overline{\varepsilon^2}(m)\right]_{min} = \sum_{i=1+m}^n \lambda_i \quad (5.14)$$

Formula (5.9) se mai numește dezvoltarea Karhunen-Loève, iar transformarea definită de relația (5.8) se numește transformarea Karhunen-Loève. Problema minimizării lui $\overline{\varepsilon^2}(m)$ se numește în statistică analiză factorială (sau analiza componentelor principale).

5.5 SVM antrenat cu date virtuale generate de sistemul neuronal CSOM

În scopul de a îmbunătăți setul de antrenare, folosim ideea inovatoare de a construi un set de antrenare compus din eșantioane virtuale, numai că substituind complet probele originale de intrare [37]. Structura VTDG-CSOM este prezentată în figura 5.6. Etapele algoritmului propus sunt următoarele:

- A. Se construiesc M subseturi ale modelelor. Unul împarte datele originale în subseturi de model M corespunzătoare fiecărei clase, unde M este numărul de clase.

- B. Instruirea fiecăruia dintre modelele SOM (k) ($k = 1, \dots, M$). Se folosește algoritmul SOM nesupervizat [37], [38], [39] pentru a instrui fiecare dintre modulele SOM. Mai precis, fiecare modul SOM a fost antrenat doar cu probele care au aceeași etichetă cu eticheta modulului neuronal.
- C. Generarea datelor virtuale de antrenare. După antrenare, vectorii ponderați a modulelor SOM au devenit datele virtuale care substituie probe reale, construind un set de antrenare îmbunătățit pentru a obține un clasificator mai precis.

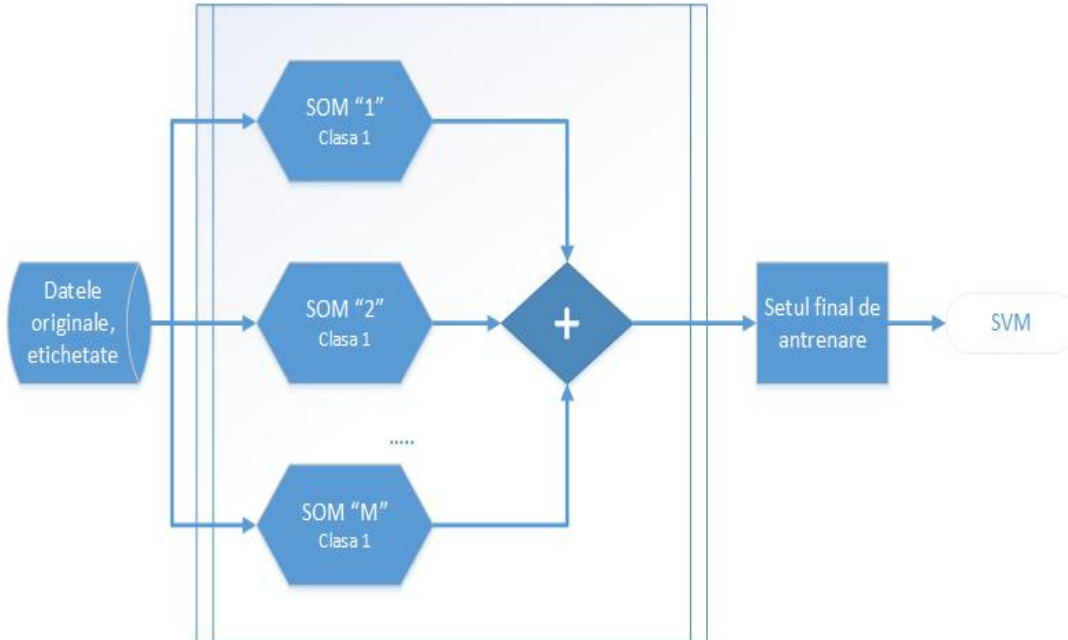


Fig. 5.6 Organigrama VTDG-CSOM.

5.6 Rezultate experimentale

5.6.1 MSTAR - baza de date a vehiculelor militare terestre

Ne-am propus să testăm modelul propus pentru identificarea vehiculelor militare. De aceea am ales baza de date MSTAR, ca un set de date standard pentru recunoașterea automată a țintei (ATR). Este realizată la Redstone Arsenal, Huntsville, AL de către Sandia National Laboratory (SNL), folosind o platformă de senzori radar cu apertură sintetică (SAR). Efortul a fost sponsorizat în comun de către DARPA și Air Force Research Laboratory, ca parte a programului Moving and Stationary Target Acquisition and Recognition (MSTAR). SNL a folosit un senzor SAR cu rezoluție de 0.3 metri. Trei clase de vehicule militare ale fostei Uniuni Sovietice au fost luate în considerare pentru experimentele noastre din baza de date MSTAR: BMP2 (vehicul de luptă de infanterie), BTR70 (blindate de transport de

personal) și T72 (rezervor). Am ales 2987 imagini de 128 x 128 pixeli, folosind două unghiuri de achiziție: 15 grade și 17 grade. Se aleg 1622 imagini corespunzătoare unghiului de 17 grade pentru antrenare, în timp ce pe celelalte 1365 imagini corespunzătoare unghiului de 15 grade sunt luate în considerare pentru testare.

5.6.2 Performanțe experimentale

Pentru experimentele noastre, dimensiunile imaginilor sunt $w = h = 128$. Pentru PCNN, am ales numărul de iterații $n = 21$ și $m = 24$ pentru a obține cele mai bune rezultate. Am ales ca metodă de referință pentru segmentare metoda constând din succesiunea algoritmilor următori: filtrare mediană, histogramă normalizată și binarizare[40]. O comparație a modelului PCNN față de performanțele modelului standard pentru astfel de aplicații poate fi evaluate în mod subiectiv în figura 5.7 [41][70].

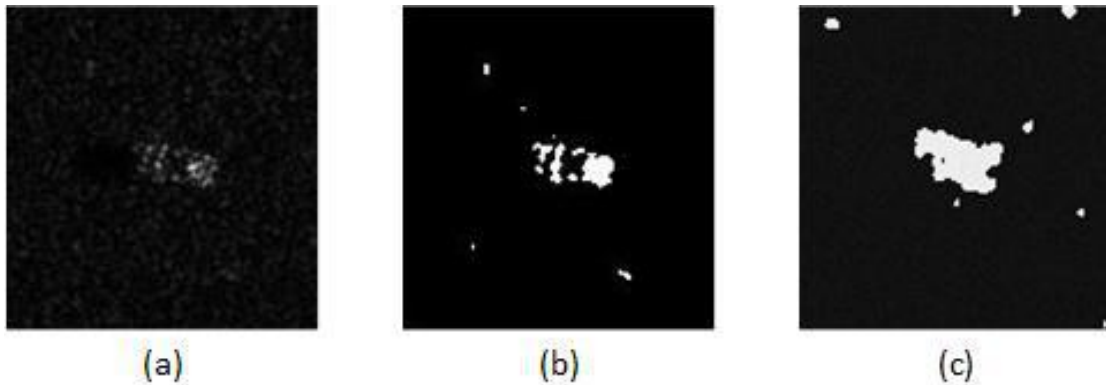


Fig. 5.7 Ilustrarea segmentării cu modelele folosite:
(a) Imaginea originală SAR (BMP 2);
(b) Imaginea obținută prin segmentarea de referință;
(c) Imaginea obținută prin segmentarea PCNN ($n = 21$, $m = 24$).

Parametrii pentru filtrul Gabor aleși sunt $u = 5$, $v = 21$, $d1 = d2 = 4$, și, în consecință, rezultatul returnat de filtrul Gabor este un vector caracteristic 107520 - dimensional. După aceea algoritmul PCA reduce acești vectori la $p = 500$. Performanțele experimentale sunt prezentate în tabelul 5.1.

Tab. 5.1 Scorul maxim de recunoaștere pentru setul de date *MSTAR*.
 $R = (\text{numărul de date de antrenare generate virtual folosind CSOM}) /$
 $(\text{număr de date de antrenare inițială}).$

Algoritmi	Scor de recunoaștere	Parametrii
PCNN-GF-PCA- {VTDG-CSOM}- SVM	97.36%	R=3.25; modele CSOM rectangulare de dimensiuni: (48x47), (28x27), (48x47); SVM cu nucleu RBF; gamma=0.0001
PCNN-GF-PCA- SVM	94.06%	SVM cu nucleu RBF; gamma=0.0001
RS-GF-PCA-SVM	86.15%	SVM cu nucleu RBF; gamma=0.01

În acest capitol s-a arătat o metodă ATR bazată pe rețele neuronale folosind radar cu apertură sintetică (SAR), pentru aplicarea specială a recunoașterii vehiculului militar la sol. ATR-ul propus are următoarele etape de procesare: segmentarea PCNN pentru detectarea obiectelor; prima etapă de selecție a caracteristicilor folosind filtrul Gabor (GF); a doua etapă de selecție caracteristică cu PCA; Clasificarea SVM folosind o procedură originală neuronală, numită VTDG-CSOM. Două dintre aceste etape sunt bazate pe rețele neuronale avansate: segmentarea PCNN și procedura VTDG-CSOM pentru îmbunătățirea performanței clasificării SVM. Experimentele sunt efectuate cu ajutorul setului de date de vehicule militare terestre *MSTAR*.

O comparație a segmentării PCNN față de performanțele metodei de referință sunt subiectiv evaluate în figura 5.7. Pentru clasificatorul clasic SVM, Tabelul 5.1 arată că segmentarea PCNN conduce la un scor de clasificare 94.06% față de 86,15% punctajul obținut prin segmentarea de referință.

Tabelul 5.1 prezintă, de asemenea, avantajul generării de date virtuale (VTDG) folosind CSOM, în scopul de a îmbunătăți clasificarea SVM. Prin substituirea datelor de antrenare originale cu eșantioane virtuale generate de modelul VTDG-CSOM (numărul de probe virtuale fiind de aproximativ $R = 3,25$ ori mai mare decât numărul de exemplare originale), se obține o creștere a scorului de recunoaștere de la 94.06% la 97.36%. Ambele abordări neuronale (segmentare PCNN, precum și cea a clasificării VTDG-CSOM) conduc la creșterea performanței de clasificare de la 86,15% până la 97.36%, ceea ce înseamnă o îmbunătățire a scorului total de 11,21%.

Capitolul 6

6. Detecția stării de ebrietate utilizând modelul PCNN pentru segmentarea imaginilor faciale din spectrul termal

6.1 Necesitatea identificării stării de ebrietate a conducătorilor auto

În toată lumea, guvernele și companiile private pun tehnologiile biometrice în centrul proiectelor de cercetare ambițioase, variind de la controlul accesului în clădiri și alte puncte cheie la pașapoarte high-tech, cărți de identitate, permise de conducere. Una dintre cele mai importante zone ale tehnologiei biometrice este de recunoaștere a feței folosind imagini în domeniul vizibil și / sau termice [42], [43], [44], [45], [46], [47], [48]. În particular, tehnologia de recunoaștere facială ar putea ajuta în urmărirea suspecților de terorism. Un domeniul legat de recunoașterea feței este recunoașterea emoțiilor prin analiza expresiilor faciale; acest lucru face parte din aria de studiu comunicare om-mașină [49].

Recunoașterea feței bazată doar pe imagini în spectrul vizibil are reale dificultăți de funcționare în medii de funcționare necontrolate. Precizia de recunoaștere a feței se degradează rapid atunci când iluminarea este slabă sau când nu se luminează în mod uniform față [49], [45], [18]. Lumina reflectată de chipurile umane, de asemenea, variază în funcție de culoarea pielii. Utilizarea imaginilor termice în infraroșu (IR) poate îmbunătăți performanța de recunoaștere a feței, în condiții de iluminare necontrolată [43], [44], [46]. Spectrul de infraroșu termic care cuprinde benzile de la mijlocul unde IR (3-5 μ m) și cu lungime de undă IR (8-12 μ m) a fost sugerat ca o sursă alternativă de informație pentru detectarea și recunoașterea fețelor [47]. Senzorii IR termici măsoară energia emisă de obiecte, nu cea reflectată. De aceea imaginile termice au mari avantaje în recunoașterea feței [47], în condiții de iluminare scăzută sau chiar și în întuneric total, caz în care tehnicile vizuale de recunoaștere a feței eșuează.

Accidentele rutiere sunt o amenințare la nivel mondial în fiecare țară [50], [51]. Raportul privind starea globală privind siguranța rutieră 2013 întocmit de Organizația Națiunilor Unite (ONU) prezintă informații privind siguranța rutieră din 182 de țări, reprezentând aproape 99% din populația lumii [50]. Raportul indică faptul că la nivel mondial numărul total de decese din traficul rutier rămâne inacceptabil de ridicat la 1,24 milioane pe an. Doar 28 de țări, acoperind 7% din populația lumii, au legi cuprinzătoare în materie de siguranță rutieră pe cinci factori de risc cheie: conducere sub influența alcoolului, excesul de viteză, utilizarea căști de protecție pentru motociclete, centuri de siguranță și scaune speciale pentru copii [50]. Aceste cifre cresc într-un ritm alarmant în fiecare an. În cazul în care nu se iau măsuri, numărul de decese va crește la 2,4 milioane până în 2030 [50]. Un raport semnificativ al acestor accidente sunt cauzate de conducătorii auto intoxicați peste limita legală cu băuturi alcoolice. Cercetarea descrisă aici abordează această problemă prin analiza imaginilor termice asupra feței subiecților. Beția este o stare fiziologică complexă și provocator de identificat fără a fi intrusivă subiectului [52]. Cu toate acestea, majoritatea publicațiilor se referă numai la sistemele auto anti-ebrietate a șoferilor, care folosesc semnale electrice de la nivelul inimii sau creierului [53]. Koukiou și Anastassopoulos au publicat recent mai multe lucrări care au ca scop diagnosticarea bețiilor folosind procesarea imaginilor faciale în domeniul termic [54], [55], [56], [57]. Succesul acestui tip de abordare se bazează pe faptul că arterele și vasele de pe fața unei persoane în stare de ebrietate își cresc capacitatea cu consumul de alcool, intervenind vasodilatația [57].

În cadrul acestui capitol ne propunem o abordare originală de diagnosticare a stării de ebrietate aplicând modelul PCNN. Primul pas este segmentarea imaginii faciale termale, folosind modelul PCNN, apoi selecția caracteristicilor prin analiza componentelor principale (PCA) urmata de Analiza discriminatorie liniară (LDA) și în final un clasificator Suport Vector Mașina (SVM). Am folosit o bază de date creată de noi, formată din 10 subiecți.

6.2 Descrierea algoritmului propus pentru detecția stării de ebrietate

Algoritmul propus pentru identificarea stării de ebrietate are următoarele etape [60], pentru a procesa imaginile termale, care sunt reprezentate grafic în figura 6.1:

- (i) Captarea imaginilor faciale termale
- (ii) Segmentarea imaginilor cu ajutorul modelului PCNN
- (iii) Înlocuirea pixelilor activi din masca obținută în urma segmentării cu valorile acestora din imaginea originală
- (iv) PCA
- (v) LDA
- (vi) Clasificarea cu ajutorul SVM

Etapele de prelucrare menționate mai sus sunt prezentate în continuare în detaliu prin selectarea și gruparea acestora în câțiva pași esențiali.

6.2.1 Segmentarea imaginilor cu modelul PCNN

Pixelii care formează segmentele extrase sunt pixelii care s-au activat în momentul iterației n , și anume, acei pixeli a căror ieșiri din neuronii corespunzătoare devin "1".

Am folosit un algoritm genetic (GA) [58] pentru a optimiza următorii șapte parametri interni ai modelului de segmentare PCNN: numărul de iterații n ; puterea de conectare β ; dimensiunea matricelor M și W ; constantele de timp $\alpha_F = \alpha_L$ și α_E ; ajustările de magnitudine $V_F = V_L$ și V_E . Parametrii AG sunt: populația cromozomilor; rata de crossover; rata de mutație; conta elită; generația de oprire. Funcția de fitness pentru GA este scorul corect de diagnostic/clasificarea a segmentelor extrase de PCNN din clasele normalitate / ebrietate. Ca urmare a segmentării PCNN, obținem un set de pixeli albi aparținând segmentelor plasate într-un fond de pixeli negri.

6.2.2 Înlocuirea pixelilor activi din masca obținută în urma segmentării cu valorile acestora din imaginea originală

După extragerea pixelilor care definesc segmentele, înlocuim pixelii din segmentele extrase de PCNN cu pixelii în niveluri de gri corespunzători.

6.2.3 Extragerea de trăsături

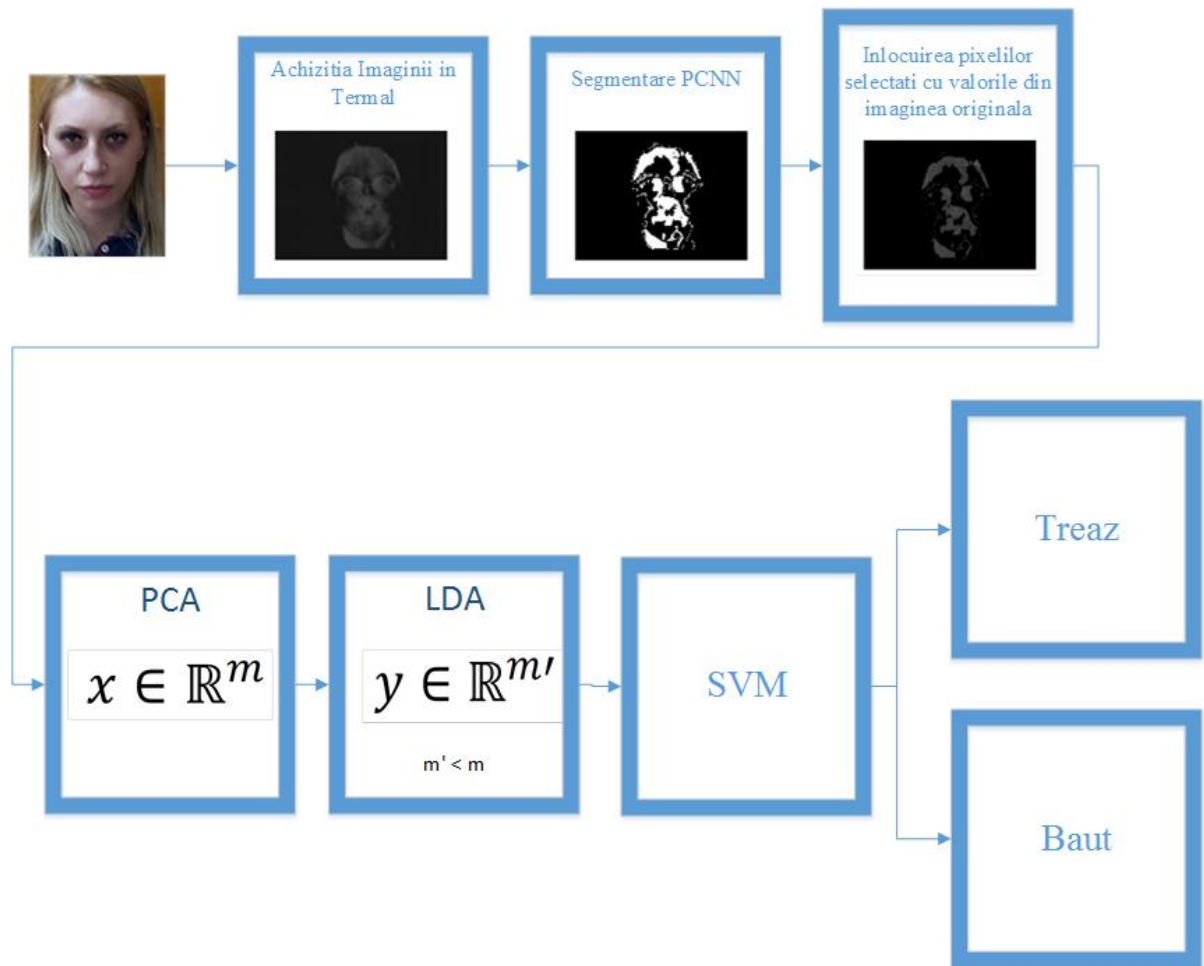


Fig. 6.1. Diagramă flux a cascadei de detecție a stării de ebrietate.

A doua etapă de selecție a caracteristicilor este Analiza Liniar Discriminantă (LDA) [59]. Prin această tehnică, am realizat o nouă reducere a dimensiunii la m la m' , dar în deosebire de PCA, prin LDA am luat în considerare etichetele datelor care aparțin setului de antrenare în vederea creșterii separabilității claselor.

6.2.3.1 PCA

Am folosit Analiza Componentelor Principale (PCA), ca o primă etapă de selecție a caracteristicilor. Am aplicat PCA pentru a transforma imaginea termică cu o dimensiune de S pixeli, unde $S = p \times q$, (aparținând unui spațiu S -dimensional) într-un spațiu redus m -dimensional. (A se vedea paragraful 5.3)

6.2.3.2 Analiza discriminativa liniară (LDA)

Această metodă mai este cunoscută și ca analiza Fisher discriminantă. Această metodă poate fi folosită în două moduri, fie ca o metodă de reducere a dimensiunii, fie o metodă generală de clasificare. Acestea fiind spuse LDA este folosită mai des ca o metodă de reducere a dimensiunilor.

Această metodă este foarte similară cu PCA, însă nu caută doar o proiecție în care să se minimizeze spațiul de reprezentare a datelor, ci, simultan, caută o proiecție care să maximizeze separarea între clase.

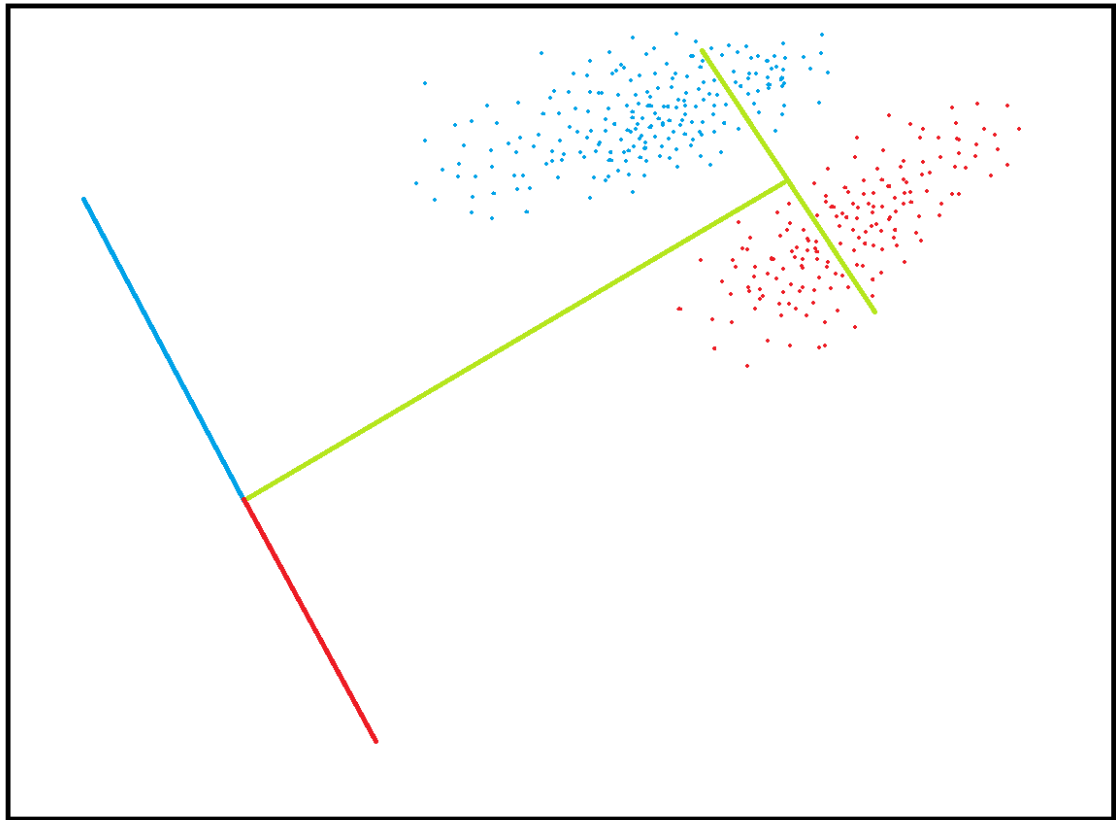


Fig. 6.2. Ilustrarea Algoritmului LDA

Algoritmul LDA

Dacă se pornește de la un număr de N_i vectori n -dimensionali aparținând la C clase se pot defini următorii parametri:

- S_B este matricea de covariație între clase și va fi egală cu:

$$S_B = \sum_{i=1}^C P(\omega_i) (\mu_i - \mu)(\mu_i - \mu)^T \quad (1.1)$$

unde $P(\omega_i)$ = probabilități a priori, μ este media celor C clase, μ_i este media fiecărei clase.

- S_w este matricea de covariație din interiorul claselor și va fi egală cu:

$$S_w = \sum_{i=1}^C P(\omega_i) \frac{1}{N} \sum_{j=1}^{N_i} (x_j^i - \mu_i)(x_j^i - \mu_i)^T \quad (2.2)$$

Se determină valorile proprii ale matricei $S_w^{-1}S_B$ și apoi se ordonează descrescător valorile proprii obținute

$$|S_w^{-1}S_B - \lambda I_n| = 0 \quad (3.3)$$

Se determină vectorii proprii corespunzători valorilor proprii:

$$S_w^{-1}S_B \Phi_i = \lambda_i \Phi_i \quad (4.4)$$

Matricea L de transformare LDA va fi formată din vectorii proprii determinați anterior:

$$L = [\Phi_1, \Phi_2, \dots, \Phi_n]^T \quad (5.5)$$

Fiecare vector de intrare X se va proiecta în spațiul LDA, astfel:

$$Y = L \cdot X \quad (6.6)$$

Se rețin doar $m < n$ componente din Y , restul de $n - m$ componente primind valoarea zero. Se va obține astfel vectorul \tilde{Y} .

Pentru a obține vectorul \hat{Y} în spațiul original, se va înmulți cu transpusa matricei L și se va obține:

$$\hat{X} = L^{-1} \cdot \hat{Y} \quad (7.7)$$

6.2.4 Clasificarea SVM

Am ales Suport Vector Mașina (SVM) (a se vedea paragraful 4.2) pentru a clasifica de imaginile PCNN pe nivel de gri, în două categorii: normalitate și ebrietate. Am considerat mai multe funcții de nucleu (de exemplu liniară, polinomială, sau radială de bază al funcției (RBF)) pentru SVM în timpul procesului de antrenare care selectează vectori de suport. Pentru un nucleu care necesită parametri, cum este RBF, am scanat, de asemenea, o gamă largă de valori ale parametrului sau (γ / spread).

6.3 Rezultate experimentale

6.3.1 Baza de date de imagini faciale termale pentru detecția stării de ebrietate

Pentru evaluarea experimentală, am construit o baza de date cu imagini în domeniul infraroșu, care conțin zona facială pentru identificarea stării de ebrietate. Toate imaginile a fost preluate cu ajutorul unui aparat de fotografiat B2 FLIR ThermoCAM și au o rezoluției de 160 X 120 pixeli, fiind preluate în intervalul de lungime de unda 7.5-13 μm . Baza de date termică conține 400 de imagini, care corespund la $J = 10$ subiecți. Imaginile termale ale fiecărui subiect au fost preluate atât pentru starea de normalitate și, de asemenea, pentru starea sub influența alcoolului, obținută după 30 de minute de la momentul în care persoana a băut o cantitate de 100 ml de whisky (acesta având 40 de grade de alcool). Fiecare subiect este reprezentat de 40 de imagini, 20 de imagini pentru starea de "normalitate" și alte 20 pentru stare de "beție". Pentru fiecare dintre stările subiectului, normalitate / ebrietate, am considerat jumătate din imagini pentru optimizarea parametrilor modelului de segmentare PCNN folosind funcția de fitness a algoritmului genetic și cealaltă jumătate a imaginilor sunt folosite pentru instruire și testare (validare) clasificatorului SVM pentru determinarea stării de beție.

În continuare vom prezenta cei 10 subiecți care au participat la formarea acestei baze de date.

Am ales o procedură de recunoaștere a stării de ebrietate independentă de subiect folosind clasificatorul SVM care implică construirea unor seturi de date de pentru fiecare din $J = 10$ subiecți. Pentru subiectul indicelui "i", am construit setul de date DS_i , unde $i = (1, \dots, J)$; fiecare set DS_i este împărțit în două sub-seturi:

- LDS_i - setul de date de antrenare pentru subiectul "i"; acesta este creat folosind imagini ($J-1$), ale celorlalte persoane din setul de date, și are o dimensiune $k \times (J-1) \times M$ imagini; k este numărul de imagini pentru fiecare dintre $M = 2$ stări ale subiectului (normalitate / ebrietate); pentru baza noastră de date, am

considerat

$k = 10$, ceea ce duce la numărul de 180 de imagini pentru orice LDSi

- TDSi - test (validare) pentru setul de date a subiectul "i", constând din un număr de imagini $M \times k$ ale persoanei "i", ceea ce duce la numărul de 20 de imagini pentru orice TDSi.



Fig. 6.3. Participanții la baza de date; (a) Adrian, (b) Constantin, (c) Gabi, (d) Victor, (e) Serban, (f) Claudia, (g) Cristina, (h) George, (i) Mihaela, (j) Teodora

6.3.2 Evaluarea performanțelor

Am luat în considerare următoarele metrici de performanță pentru a evalua modelul propus:

- TP_i = true pozitive = starea de ebrietate detectată corect pentru subiectul "i" (imagine cu subiect în stare de ebrietate și clasificată corect)
- TN_i = true negative = starea de normalitate detectată corect pentru subiectul "i" (imagine cu subiect în stare de normalitate clasificat ca normal)
- FP_i = fals pozitive = imagini în starea de normalitate ale subiectului "i" diagnosticat în stare de ebrietate
- FN_i = false negative = imagini în starea de ebrietate pentru subiect "i" diagnosticat ca normal (pentru $i = 1, \dots, 10$)
- Precizie generală:

$$OAI = 100 \times (TP_i + TN_i) / (TP_i + TN_i + FP_i + FN_i) \quad (8.1)$$

- Rata de alarme false:

$$FAR_i = 100 \times (FP_i) / (TN_i + FN_i) \quad (6.2)$$

- Rata de alarme ratate

$$MAR_i = 100 \times (FN_i) / (TP_i + FN_i) \quad (6.3)$$

Am mai calculat, de asemenea, indicatorii globali de performanță, OA, FAR, MAR prin calcularea mediei indicilor definiți în relațiile (6.1), (6.2), (6.3) peste toate $J = 10$ subiecți.

6.3.3 Performanțe experimentale

Pentru algoritmul genetic folosit pentru a optimiza parametrii PCNN, am ales următorii parametri:

- Populația $P = 150$
- Rata de crossover = 0.8
- Rata de mutație = 0.01
- Elite count = 0.05
- Generația de oprire $N=50$.

În figurile 6.4, 6.5, 6.6 și 6.7 se prezintă rezultatele modelului propus pentru patru dintre cei zece subiecți.

Decizia finală furnizată de clasificatorul SVM se ia pe imaginile pe nivelului de gri, furnizate de segmentarea cu modelul PCNN și a căror pixeli activați au fost înlocuiți cu valorile pe niveluri de gri din imaginile termale originale, cu cum se arată în figura 6.7.

Pentru etapa de selecție caracteristică PCA, am ales $m = 150$ caracteristici; pentru selecția caracteristicilor cu LDA, am selectat $m' = 1$.

În tabelul 6.1 se regăsește-te evaluarea performanțelor modelului propus iar în tabelul 6.2 se regăsesc rezultatele modelului în care nu intervine segmentarea PCNN și imaginea este data în totalitatea sa spre etapa de extragere de trăsături.

În tabelele 6.3 și 6.4 se arată matricea de confuzie pentru fiecare subiect în parte cât și cea globală.

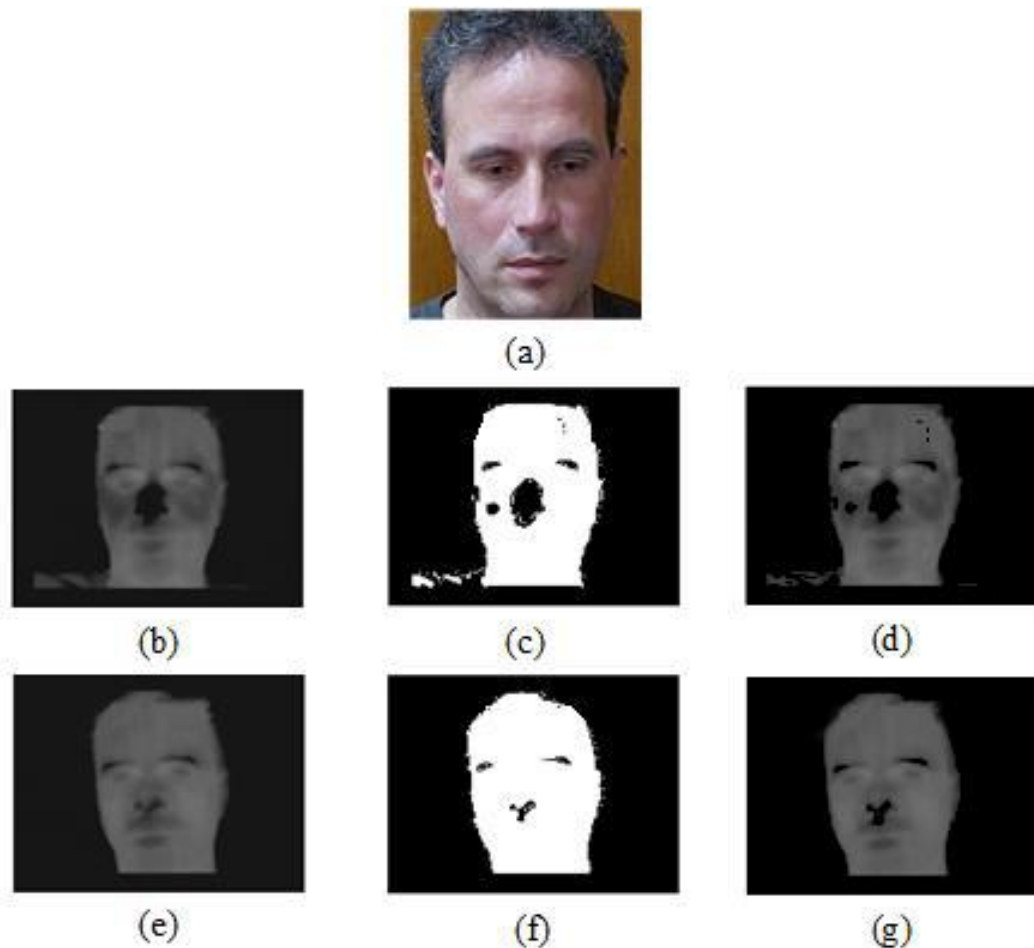


Fig. 6.4. Gabriel: (a) imaginea în spectrul vizibil; (b) imaginea de intrare, în domeniul infraroșu, normalitate; (c) imaginea segmentată PCNN (alb corespunde segmentelor), normalitate; (d) pixelii din segmentele extrase de PCNN înlocuiți cu nivelurile de gri corespunzătoare, normalitate; (e) imaginea de intrare, în domeniul infraroșu, ebrietate; (f) imaginea segmentată PCNN (alb corespunde segmentelor), ebrietate; (g) pixelii din segmentele extrase de PCNN înlocuiți cu nivelurile de gri corespunzătoare, ebrietate;

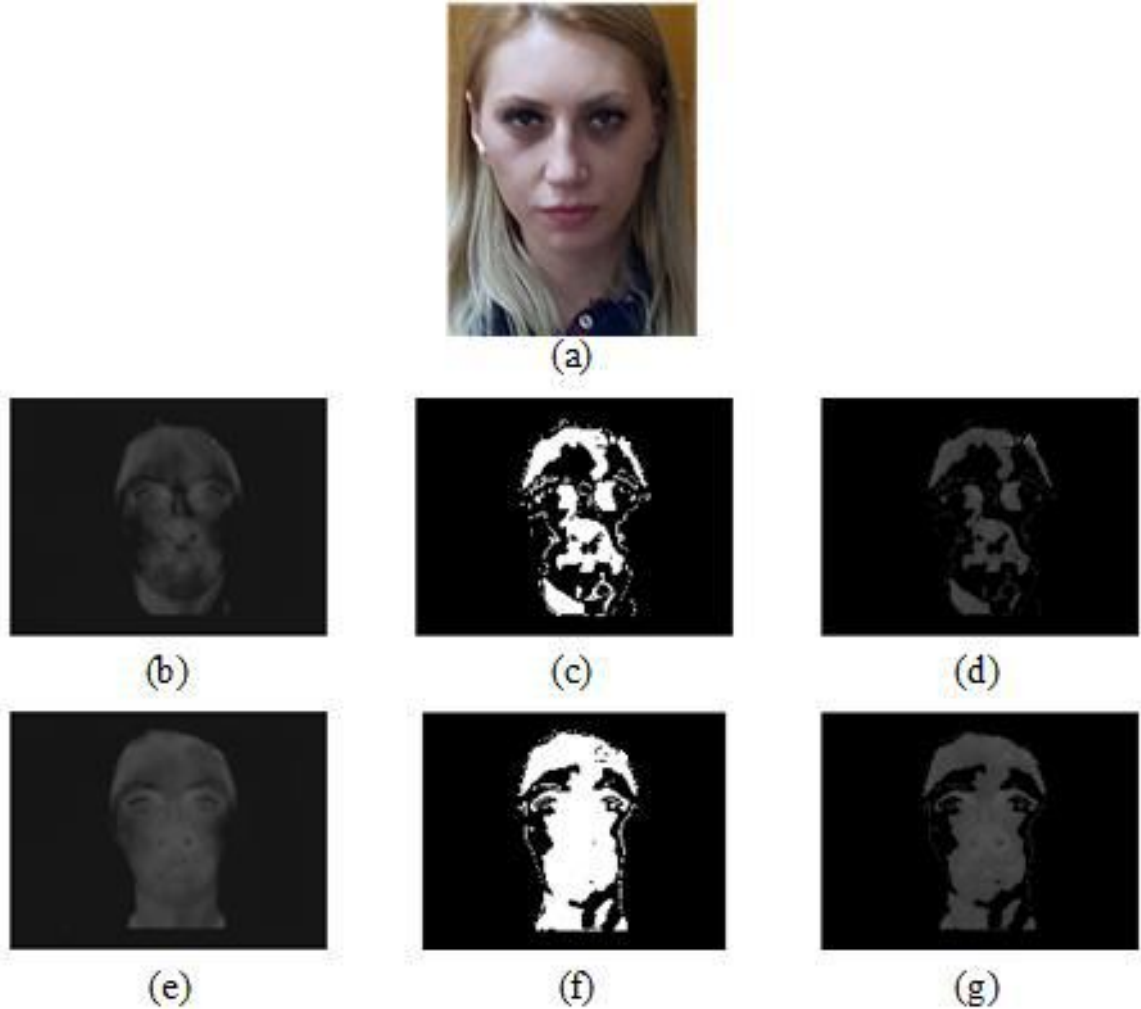


Fig. 6.5. *Mihaela: (a) imaginea în spectrul vizibil; (b) imaginea de intrare, în domeniul infraroșu, normalitate; (c) imaginea segmentată PCNN (alb corespunde segmentelor), normalitate; (d) pixelii din segmentele extrase de PCNN înlocuiți cu nivelurile de gri corespunzătoare, normalitate; (e) imaginea de intrare, în domeniul infraroșu, ebrietate; (f) imaginea segmentată PCNN (alb corespunde segmentelor), ebrietate; (g) pixelii din segmentele extrase de PCNN înlocuiți cu nivelurile de gri corespunzătoare, ebrietate;*

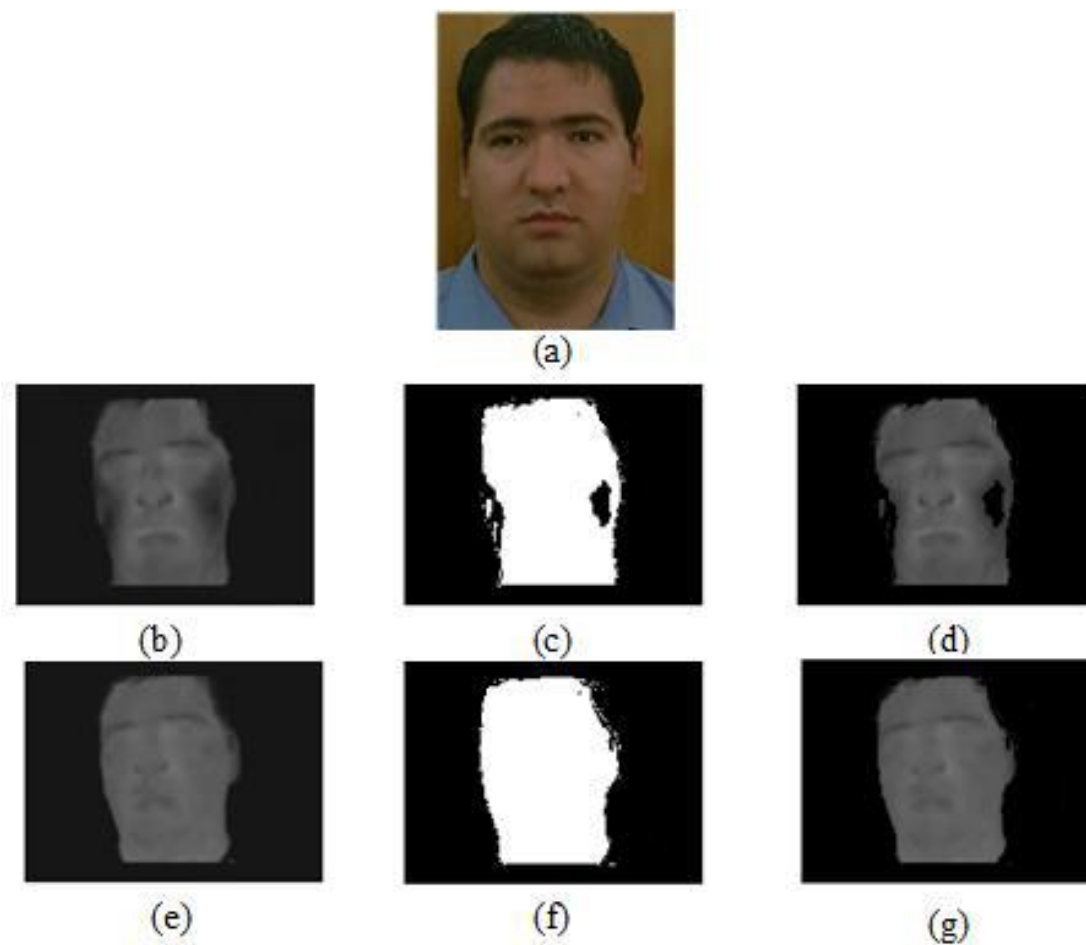


Fig. 6.6. Adrian: (a) imaginea în spectrul vizibil; (b) imaginea de intrare, în domeniul infraroșu, normalitate; (c) imaginea segmentată PCNN (alb corespunde segmentelor), normalitate; (d) pixelii din segmentele extrase de PCNN înlocuiți cu nivelurile de gri corespunzătoare, normalitate; (e) imaginea de intrare, în domeniul infraroșu, ebrietate; (f) imaginea segmentată PCNN (alb corespunde segmentelor), ebrietate; (g) pixelii din segmentele extrase de PCNN înlocuiți cu nivelurile de gri corespunzătoare, ebrietate;

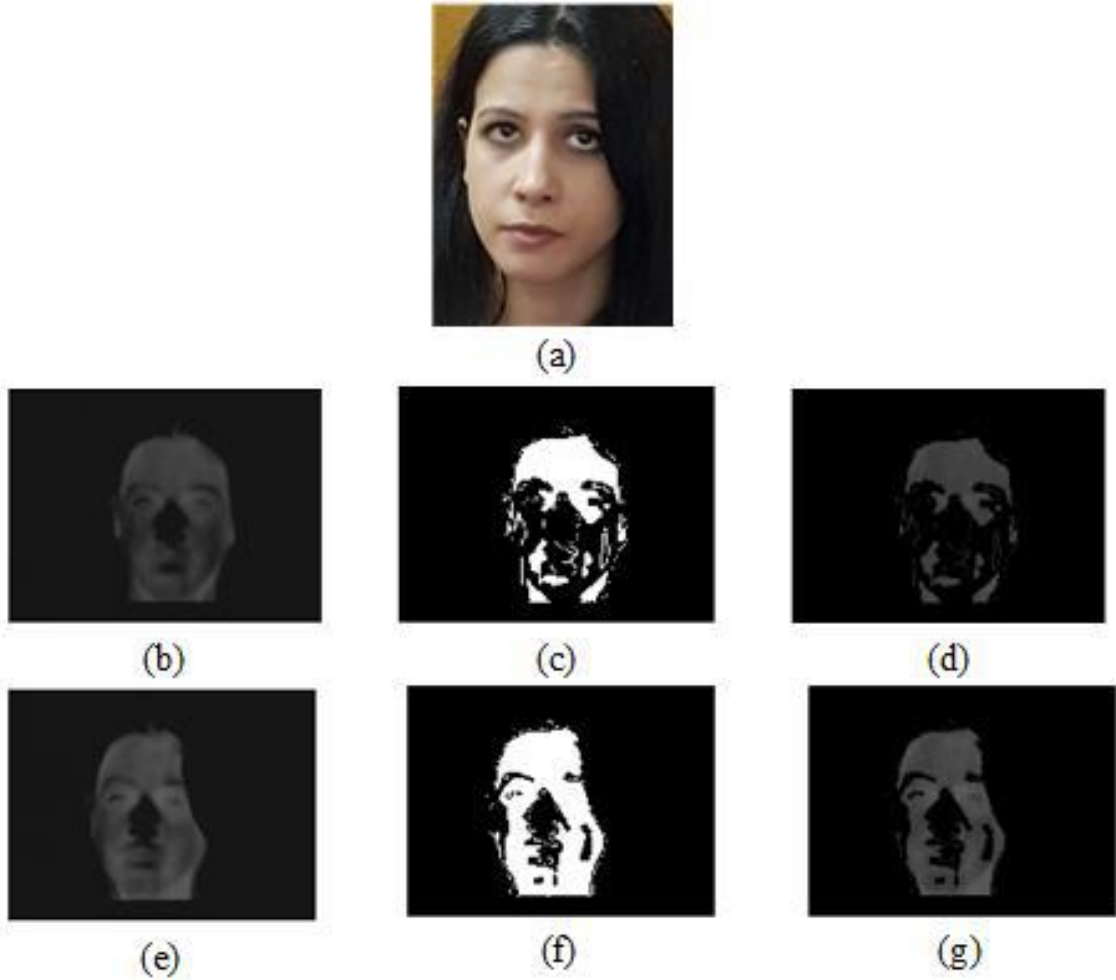


Fig. 6.7. Cristina: (a) imaginea în spectrul vizibil; (b) imaginea de intrare, în domeniul infraroșu, normalitate; (c) imaginea segmentată PCNN (alb corespunde segmentelor), normalitate; (d) pixelii din segmentele extrase de PCNN înlocuiți cu nivelurile de gri corespunzătoare, normalitate; (e) imaginea de intrare, în domeniul infraroșu, ebrietate; (f) imaginea segmentată PCNN (alb corespunde segmentelor), ebrietate; (g) pixelii din segmentele extrase de PCNN înlocuiți cu nivelurile de gri corespunzătoare, ebrietate;



Fig. 6.8 Imaginile finale prelucrate pentru a fi clasificate starea de ebrietate:
(a) normalitate; (b) ebrietate.

Tab. 6.1 Evaluarea performanțelor independenta de subiect pentru identificarea stării de ebrietate prin utilizarea modelului propus,
(m PCA=150; m' LDA=1)

Nucleul si parametrii clasificatorului SVM	Dimensiunea mareicilor M si W	n (Nr. de iterații PCNN)	$V_L = V_F$	V_E	$\alpha_L = \alpha_F$	α_E	β	OA (%)	MAR (%)	FAR (%)
Liniar	4	3	0.1	0.31	2.8	16.6	0.4	97.5	1	4
Polinomial de grad 2	4	3	0.1	0.31	2.8	16.6	0.4	97.5	1	4
Polinomial de grad 3	4	3	0.1	0.31	2.8	16.6	0.4	97.5	1	4
Polinomial de grad 4	4	3	0.1	0.31	2.8	16.6	0.4	97.5	1	4
Polinomial de grad 5	4	3	0.1	0.31	2.8	16.6	0.4	97.5	1	4
RBF sigma = 0.571	6	5	0.28	0.63	36.03	2.05	3.62	97.5	1	4
RBF sigma = 1	6	5	0.28	0.63	36.03	2.05	3.62	97.5	1	4

Tab. 6.2 Evaluarea performanțelor independenta de subiect pentru identificarea stării de ebrietate prin modelul în care nu intervine segmentarea PCNN ci imaginea este data în totalitatea sa spre etapa de extragere de trăsături.

(m PCA=150; m' LDA=1)

Nucleul și parametrii clasificatorului SVM	OA (%)	MAR (%)	FAR (%)
Liniar	77.5	24%	14%
Polinomial de grad 2	76.5	24%	16%
Polinomial de grad 3	78	20%	17%
Polinomial de grad 4	76.5	16%	23%
Polinomial de grad 5	80	8%	23%
RBF sigma=0.411	77	24%	0%
RBF sigma=1	77.5	24%	21%

Tab. 6.3 Matricea de confuzie a subiectului "i" pentru identificarea stării de ebrietate folosind modelul propus (m PCA=150; m' LDA=1)

Subiectul de index "i"	1		2		3		4		5	
	<i>Starea reala</i> → <i>Starea dedusa</i> ↓	normalitate	ebrietate	normalitate	ebrietate	normalitate	ebrietate	normalitate	ebrietate	normalitate
normalitate	100%	0%	100%	0%	100%	0%	90%	10%	100%	30%
ebrietate	0%	100%	0%	100%	0%	100%	10%	90%	0%	70%
Subiectul de index "i"	6		7		8		9		10	
	normalitate	ebrietate	normalitate	ebrietate	normalitate	ebrietate	normalitate	ebrietate	normalitate	ebrietate
<i>Starea reala</i> → <i>Starea dedusa</i> ↓	normalitate	ebrietate	normalitate	ebrietate	normalitate	ebrietate	normalitate	ebrietate	normalitate	ebrietate
normalitate	100%	0%	100%	0%	100%	0%	100%	0%	100%	0%
ebrietate	0%	100%	0%	100%	0%	100%	0%	100%	0%	100%

Tab. 6.4 Matricea de confuzie globala pentru identificarea stării de ebrietate folosind modelul propus ($m_{PCA}=150$; $m'_{LDA}=1$)

<i>Starea reala</i>		
→		
<i>Starea dedusa</i>	normalitate	ebrietate
↓		
normalitate	99%	4%
ebrietate	1%	96%

În acest capitol am prezentat faptul că am construit o bază de date originală, făcută cu scopul de detecție a stării de ebrietate în imagini termale, formată din 400 de imagini aferente la 10 subiecți. Pe lângă aceasta bază de date am efectuat o evaluare detaliată de performanță, ținând seama de precizia generală (OA), rata de ratare a alarmelor (MAR), rata de alarme false (FAR) și matricea de confuzie. Rezultatele experimentale confirmă metoda propusă; ele nu depind de tipul și parametrii clasificatorului SVM. Metoda în cascadă propusă de analiză a imaginilor termale pentru detectarea stării de ebrietate bazată pe PCNN conduce la performanțe foarte bune cu OA de 97,5%, iar MAR de 1%. Scorurile de mai sus sunt mult mai bune decât cele mai bune cele obținute prin metoda de referință folosind analiza imaginilor termale fără segmentarea PCNN, și anume OA de 80%, iar MAR de 8%.

Capitolul 7

7. Detecția schimbărilor in imagini multitemporale de observație terestra aplicând modelul PCNN

7.1 Metode nesupervizate pentru detecția schimbărilor în serii de imagini de observație terestra

Detectarea schimbărilor[39], [42], [61], [62], [63], [69] în imagini folosind metode nesupervizate joacă un rol important în multe domenii legate de exploatarea imaginilor de teledetecție. Disponibilitatea imaginilor achiziționate la aceeași zonă geografică de senzori satelitari la momente diferite face posibilă identificarea și etichetarea posibilelor schimbări ce au avut loc pe teren. În acest context, în scopul de a exploata în mod corespunzător cantitatea de date achiziționate de sateliții de teledetecție, este obligatoriu să se dezvolte tehnici eficiente pentru detecția schimbărilor, iar acestea trebuie sa fie nesupervizate și automate deoarece volumul de date este prea mare pentru a face viabilă intervenția umană la fiecare caz.

Metodele de detectare a schimbărilor sunt grupate în șapte categorii: (1) Metode algebrice, (2) Metode bazate pe transformări matematice, (3) Metode bazate pe clasificare, (4) Modele avansate, (5) Sistem informatic geografic (GIS), (6) Analiză vizuală, și (7) alte abordări. Cea de a șaptea categorie include acele tehnici de detectare a schimbărilor care

nu sunt potrivite pentru a grupa în oricare din cele șase categorii și nu sunt încă utilizate pe scară largă în practică. Prin urmare, această categorie nu este discutată în detaliu. Majoritatea acestor tehnici sunt utilizate pentru detectarea schimbărilor cu o rezoluție relativ fină spațială, cum ar fi Landsat MSS sau radar.

7.1.1 Metode algebrice

Categoria metodelor algebrice include diferența imaginilor, regresia imaginilor, analiza vectorilor de schimbare (CVA) și scăderea fondalului, etc. Acești algoritmi au o caracteristică comună, și anume, selectarea pragurilor, din filtre trece sus sau trece jos, pentru a determina zonele modificate. Aceste metode (cu excepția CVA) sunt relativ simple, ușor de implementat și de interpretat, dar acestea nu pot furniza informații complete de schimbare. CVA este o extensie conceptuală a diferenței de imagini. Această abordare poate detecta toate schimbările mai mari decât pragurile identificate și pot furniza informații detaliate despre schimbare. Un dezavantaj al metodelor algebrice este dificultatea în selectarea pragurilor corespunzătoare pentru a identifica zonele modificate. În această categorie, două aspecte sunt critice pentru rezultatele de detectare a schimbării: selectarea benzilor spectrale adecvate din imagine vegetație și selectarea pragurilor adecvate pentru identificarea zonelor modificate.

7.1.2 Metode bazate pe transformări matematice

Categoria de transformare include PCA, Gramm-Schmidt (GS), și transformările Chi-pătrat. Un avantaj al acestor metode este în reducerea redundanței datelor între benzi și subliniind diferite informații în componente derivate. Cu toate acestea, ele nu pot furniza matrice de schimbare detaliate și necesită selectarea pragurilor pentru identificarea zonelor modificate. Un alt dezavantaj este dificultatea în interpretarea și etichetarea informațiilor.

7.1.3 Metode bazate pe clasificare

Metodele bazate pe clasificare includ compararea post-clasificare, algoritmul de detecție a modificării expectation-maximization(EM), detectarea schimbărilor nesupervizata, detectarea schimbarea hibrid, și ANN. Aceste metode se bazează pe imagini clasificate, în care calitatea și cantitatea de date eșantion de antrenare sunt cruciale pentru a produce rezultate bune de clasificare. Avantajul major al acestor metode este capacitatea de a furniza o matrice de informații și de reducerea impactului schimbărilor externe din diferențele atmosferice și de mediu între imagini multi-temporale. Cu toate acestea, selectarea probelor de înaltă calitate și suficient de numeroase pentru antrenarea clasificării este adesea dificilă, în special, pentru o clasificarea datelor din imagini istorice. Sarcina

este consumatoare de timp și dificilă și de multe ori duce la rezultate nesatisfăcătoare de detectare a schimbărilor, în special atunci când nu există date mostră de antrenare disponibile.

7.1.4 Metode avansate

Modelele avansate de detectare a schimbării includ modelul Li-Strahler, de reflexie, modele bazate pe combinarea planelor spectrale și modele bazate pe estimare a parametrilor folosind metode biofizice. În aceste metode, valorile din imaginile de reflexie sunt adesea convertite în parametri reali sau fracționari prin modele liniare sau neliniare. Parametrii rezultați sunt mai intuitivi și adecvați pentru interpretarea și extragerea informației decât sunt semnăturile spectrale. Dezavantajul acestor metode este timpul necesar rulării și procesul dificil de dezvoltare a modelelor adecvate pentru conversia valorilor imaginii la parametri biofizici.

7.1.5 GIS

Categoria de detecție a modificărilor sistem informatic geografic (GIS) include GIS integrate cu metode de teledetecție și metoda GIS pură. Avantajul folosirii GIS este abilitatea de a încorpora diferite surse de date în aplicații de detectare a schimbărilor. Cu toate acestea, diferite surse de date asociate cu diferite formate de date și precizii afectează adesea rezultatele de detectare a schimbărilor.

7.1.6 Analiză vizuală

Metoda analizei vizuale include interpretarea vizuală a imaginilor compozit multi-temporale și digitizarea zonelor modificate. Această metodă se poate folosi din plin de experiența și cunoștințele unui analist. Textura, forma, mărimea și modelele ale imaginilor sunt elemente cheie utile pentru identificarea schimbării prin interpretare vizuală. Aceste elemente nu sunt adesea folosite în analiza digitală din cauza dificultății de extracție a acestor elemente. Cu toate acestea, în interpretare vizuală, un analist calificat poate încorpora toate aceste elemente în a ajuta să ia decizii cu privire la schimbarea terenului. Dezavantajul acestei metode este timpul consumat pentru o aplicație de detectare a schimbării pe o suprafață mare și este dificilă actualizarea în timp util a rezultatelor de schimbare detectate. De asemenea, este dificil de a oferi traiectorii de schimbare detaliate.

7.1.7 Alte abordări

În plus față de cele șase categorii de tehnici de detectare a schimbărilor discutate mai sus, există, de asemenea, unele metode care nu pot fi atribuite uneia dintre categoriile indicate

mai sus și care nu au fost încă frecvent utilizate în practică. De exemplu, Henebry (1993) a folosit măsuri de dependență spațială cu date de mapare termala pentru a detecta o schimbare în o pășune. Wang (1993) a folosit un sistem bazată pe informații anterioare, pentru a detecta schimbările de teren care la marginea zonelor urbane. Lambin și Strahler (1994b) au folosit trei indicatori, indici de vegetație, temperatura suprafeței terenurilor și structura spațială, pentru a detecta schimbarea terenurilor în Africa de Vest. Lawrence și Ripple (1999) a folosit curbe de schimbare și Hussin et al. (1994) a utilizat un model de producție pe zonă pentru a detecta schimbările de acoperire cu pădure. Morisette et. al. (1999) au folosit modele liniare generalizate pentru a detecta schimbarea terenurilor. O abordare bazată pe teorema curbilor a fost, de asemenea, utilizat pentru detectarea schimbării în delta râului Galben (Yue și colab., 2002).

7.2 Analiza vectorilor de schimbare (CVA)

Să luăm în considerare două imagini multispectrale, X_1 și X_2 , de mărime $I \times J$, achiziționate peste aceeași zonă în momente diferite t_1 și t_2 . Fie x_1 și x_2 două variabile aleatoare multidimensionale care reprezintă distribuțiile statistice ale pixelilor în imaginile X_1 și X_2 . Fie $X_{b,t}$ variabila aleatoare care reprezintă a b -a componenta a imaginii multispectrale X_t ($t = 1, 2$), în spațiul de caracteristici considerat. Fie $\Omega = \{\omega_n, \Omega_C\}$ setul de clase de pixeli schimbați și neschimbați care urmează să fie identificați. Mai în detaliu, ω_n reprezintă clasa pixelilor neschimbați $\Omega_C = \{\omega_{c_1}, \dots, \omega_{c_K}\}$ este setul de K clase posibile de pixeli schimbați dintr-o zonă.

Atunci când se prelucrează imagini multispectrale, cel mai utilizat operator de comparație este diferența. În multe cazuri, diferența este aplicată într-un spațiu de caracteristici n -dimensional, în scopul de a da ca intrare în procesul de detectare a schimbărilor toate informațiile spectrale relevante. Aceasta tehnica poartă numele de analiza vectorilor de schimbare (SVC) [61], [62], [66] și a fost folosită cu succes în multe aplicații [39], [63]. CVA calculează în primul rând, o imagine de diferență multispectrală (X_D), scăzând vectorii caracteristici asociați cu fiecare poziție spațială corespunzătoare în cele două imagini luate în considerare X_1 și X_2 . X_D devine o imagine multidimensională care reprezintă vectorii de schimbare spectrali (SCV) și se obțin astfel:

$$X_D = X_2 - X_1 \quad (7.1)$$

Fiecare SCV este de obicei reprezentată implicit în coordonate polare [62] cu magnitudinea și direcția sa. Cu toate că direcția SCV este bogată de informații (de exemplu, cu privire la tipul de modificări au avut loc pe teren și asupra distribuției zgomotului de înregistrare), în cele mai multe dintre aplicațiile nu este considerată. Printre puținele studii raportate în literatura de specialitate în care magnitudinea și direcția exprimate ca funcții

cosinus sunt luate în considerare împreună pentru detectarea schimbărilor, ne amintim. În 1980, Malila [62] a formulat mai întâi conceptul de vector schimbare și apoi a utilizat magnitudinea și direcția într-un spațiu cu două dimensiuni pentru identificarea schimbărilor datorate plantelor tăiate și apoi recreșterea în pădurea din nordul Idaho (SUA). În [62] variabila direcție a fost împărțită într-un număr fix de sectoare, fiecare dintre ele corespunzând schimbări pozitive sau negative într-una dintre cele B caracteristici luate în considerare (de exemplu, canale spectrale sau combinații liniare ale acestora, cum ar fi transformarea Tasseled - PAC). Acest tip de cuantificare conduce la definirea a maximum 2^B sectoare și, prin urmare, a tipurilor de schimbări. Principalul dezavantaj al acestei abordări este faptul că diferitele tipuri de modificări ar putea să își asume același cod de sector. În [9], abordarea codificarea CVA a fost extinsă ca o soluție cu mai multe variații (adică, aplicații care implică mai mult de două date de achiziție). Allen și Kupfer introduc în tehnica CVA utilizarea cosinusurilor pentru descrierea direcțiilor SCV. Ei au aplicat o analiză discriminantă liniară ierarhică pentru testarea puterii predictive a magnitudinii și unghiurilor vectorilor în rezolvarea problemelor de detecție a schimbărilor.

Amploarea vectorilor spectrali de schimbare (SCV) este derivată în conformitate cu analiza vectorilor de schimbării (CVA):

$$d = \sqrt{\frac{1}{n} \sum_{i=1}^n (b_{1i} - b_{2i})^2} \quad (7.2)$$

unde b_1 și b_2 reprezintă vectorii spectrali ale imaginilor achiziționate la momentele t_1 și t_2 .

7.3 Metoda PCNN - Pacifici pentru detectia schimbărilor în imagini satelitare multitemporale

PCNN poate fi folosit pentru a identifica, într-un mod complet automat, zonele din imagini în care au avut loc schimbări semnificative. De fapt, aplicarea lor în detecție schimbărilor poate fi văzută ca o comparație a undelor corespunzătoare imaginilor obținute la diferite momente [68][69].

Pentru fiecare unitate, adică pentru fiecare pixel al imaginii, PCNN oferă o valoare de ieșire. Considerând $Y_{ij}[n]$ ieșirea modelului PCNN, semnalul de timp $G[n]$, calculat prin:

$$G[n] = \frac{\sum_{ij} Y_{ij}[n]}{N} \quad (7.3)$$

este în general folosit pentru a converti imagini de impulsuri într-un singur vector de informație. În acest fel, este posibil să existe o măsură globală a numărului de pixeli activați la epoca [n] într-o sub-imagine care conține N pixeli. Semnalul asociat la $G[n]$ s-a dovedit a avea proprietăți de invarianță la schimbări de scara și la rotație a unui obiect în scenă. Această ultimă caracteristică face modelul PCNN o abordare adecvată pentru detectarea schimbărilor în imagini cu rezoluție foarte mare, în cazul în care unghiul de vedere al sensorului poate juca un rol important [64][65][70].

Fabio Pacifici, Fabio Del Frate și William J. Emery au realizat și au demonstrat un modelul PCNN pentru detectarea schimbărilor în zonele urbane [68], [69], folosind imagini de rezoluție foarte mare. Aceștia au creat doi vectori, câte unul pentru fiecare imagine, generate de PCNN în timpul fiecărei iterație a algoritmului, a creat semnături specifice ale scenei, care pot fi comparate pentru a decide cu privire la apariția unei schimbări. Metoda este complet automatizată și destul de rapidă, deoarece analizează în mod direct relația dintre cele două semnale de timp asociate imaginilor și evită comparația pe baza pixelilor, puse în aplicare prin abordări mai tradiționale, cum ar fi cele care utilizează analiza vectorilor de schimbare. În plus, nici o pre-procesare, este necesară.

7.4 Modelul PCNN propus pentru detecția schimbărilor în imagini satelitare multitemporale

Modelul propus de Victor Neagoe et al. [43] [70] combină abordarea lui Fabio Pacifici et al. [78], [79] pentru identificarea schimbărilor cu abordarea clasică CVA [66] realizând un algoritm de identificare al schimbărilor în imagini din serii temporale cu rezoluție la nivel de pixel bazat pe modelul PCNN.

Pentru cazul particular al unei imagini multispectrale cu doar două benzi, notate cu B_1 și B_2 , considerăm $x(i, j, B_k, t_h)$ ca luminozitatea corespunzătoare a pixelilor de coordonate (i, j) în banda de B_k în acest t_h ($h = 1$ sau 2). Pixelul din diferență de imagine (DI) este dat de relația:

$$D_{ij} = \sqrt{[x(i, j, B_1, t_2) - x(i, j, B_1, t_1)]^2 + [x(i, j, B_2, t_2) - x(i, j, B_2, t_1)]^2} \quad (7.4)$$

În figura 7.2 este ilustrată schema logică a algoritmului de detectare a schimbărilor propus, acesta având trei etape: (a) calculul diferenței dintre imagini (DI), care corespunde magnitudinilor vectorilor spectrali de schimbare (SCV); (b) evaluarea numărului total de activări din fiecare neuron PCNN pentru un anumit interval de timp, asumând DI ca

imagine de intrare pentru PCNN; (c) realizarea unui clustering cu ajutorul algoritmului K-Means pe doua clase a elementelor Z_{ij} din ecuația (7.5), care reprezinta numarul de activari a unui neuron cu coordonatele (i,j) în intervalul L, pentru identificarea pixelilor modificați.

$$Z_{ij} = \sum_{ij}^L Y_{ij}[n] \quad (7.5)$$

7.5 Rezultate experimentale

7.5.1 Baza de date Mexico

Setul de date luate în considerare este alcătuit dintr-o selecție a două imagini multispectrale achiziționate de senzorul Thematic Mapper (TM) din Landsat-5, si sunt ilustrate în figura 7.1. Cele două imagini, de rezoluție 512x360 pixeli, au fost achiziționate în aprilie 2000 și mai 2002 într-o zonă din Mexic. Între cele două date de achiziție un incendiu a distrus o mare parte a unei păduri vizibile în imagini. Suprafața arsă (29506 pixeli) reprezintă zona modificata în setul de date.

Restul de 154814 pixeli reprezintă aria nemodificată. Am folosit canale spectrale 4 și 5 pentru experimentele noastre.

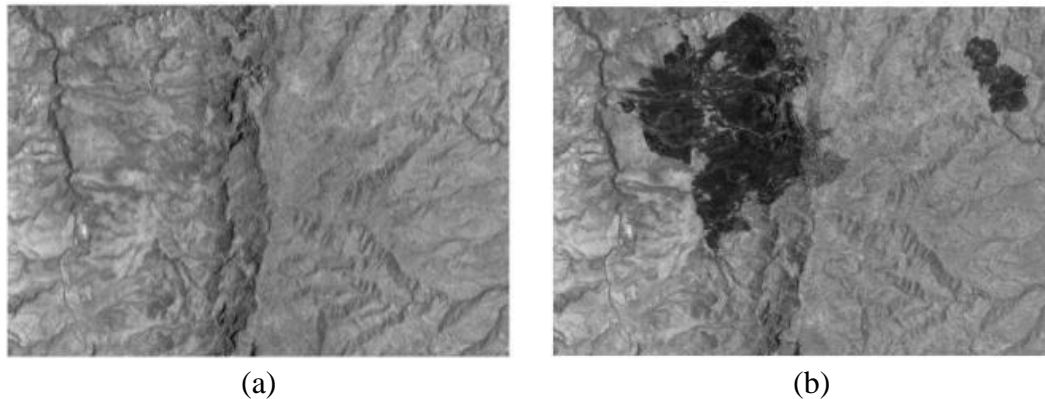


Fig. 7.1 Banda 4 din imaginea Landsat-5 TM, din zona Mexic
(a) aprilie 2000. (b) mai 2002.

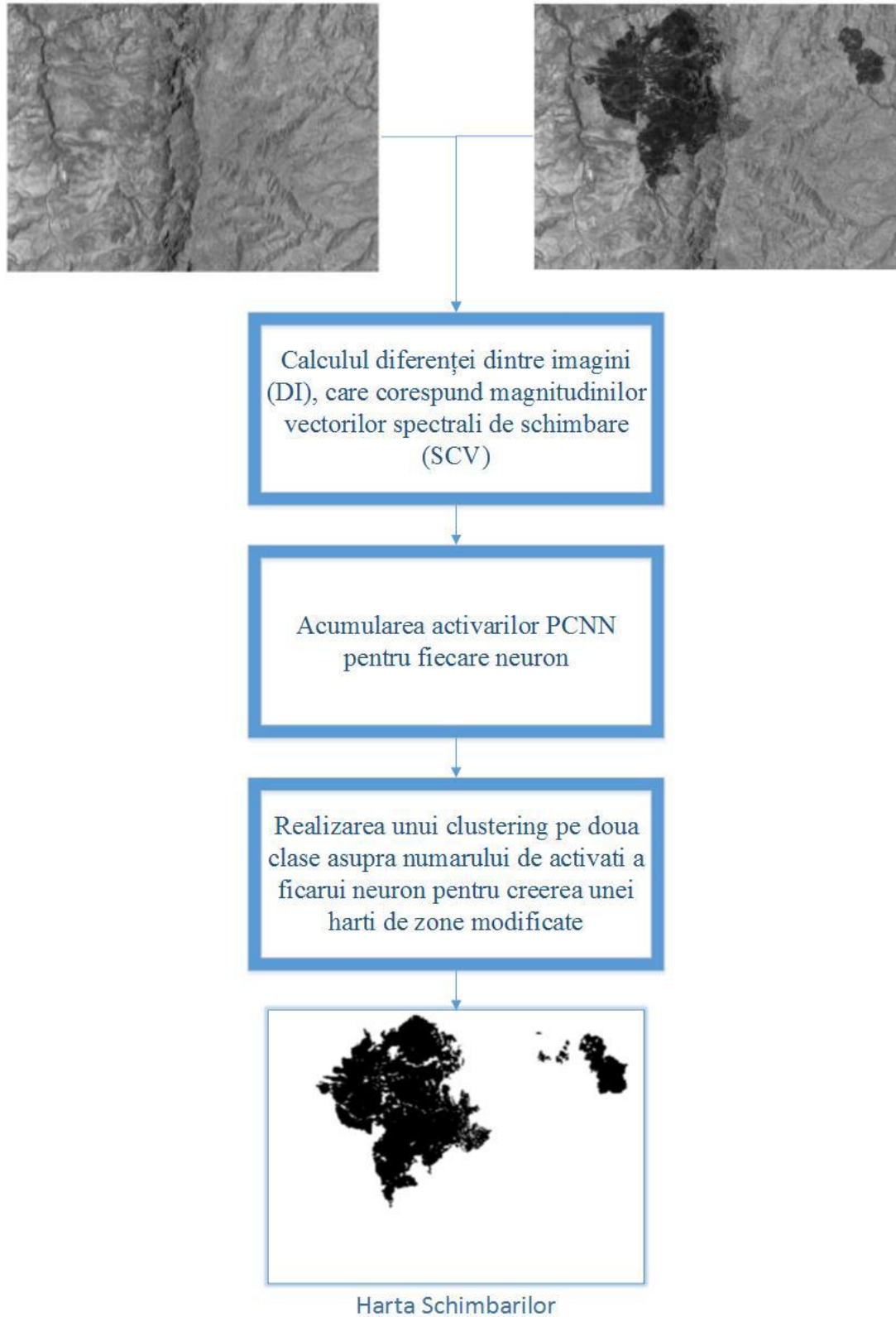


Fig. 7.2 Schema logica a algoritmului

7.5.2 Performante experimentale

Parametrul PCNN numit *link_arrange* controlează raza de corelare între un pixelul central de coordonate (i, j) și vecinii săi de coordonate (k, l).

Pentru a evalua performanța de detectare a schimbărilor, am luat în considerare precizia generală (OA). Cea mai bună performanță a fost obținută prin alegerea $link_arrange = 40$, iar intervalul de evaluare $L = np(iteration_times) = 100$. Harta de schimbare este prezentată în figura 4.3 și corespunde cu un $OA = 97.66\%$. Parametrii de clustering rezultați sunt:

- Valoarea maxima neschimbata = 55
- Valoarea minima schimbata = 56



Fig. 4.3 Vizualizarea rezultatelor experimentale. a) Harta schimbărilor obținută folosind modelul CVA-PCNN cu acumularea activărilor - clustering ($OA = 97.86\%$). (b) harta de referință. (Pixeli negri sunt "schimbate"; pixelii albi sunt "neschimbate").

În acest capitol s-a propus un model inovativ, nesupervizat, de detecție a schimbărilor bazat pe rețele neuronale, constând din CVA (SCV)-PCNN cu acumularea activărilor - clusterizare.

Prin comparație cu metoda Pacifici-Del Frate, bazata pe modelul PCNN, de detectare a schimbărilor, care poate detecta schimbări pe o regiune (zona) de (de exemplu, folosind zone de 100×100 pixeli), modelul nostru poate evalua schimbări la nivel de pixel.

Capitolul 8

8. Concluzii

In capitolul de concluzii vom recapitula succint aplicațiile modelului Pulse-Coupled Neural Network (PCNN) dezvoltate in teza. Aceste aplicații se pot încadra in mai multe domenii de studiu ca Machine Learning, Pattern Recognition, Computer Vision. Fiecare capitol din aceasta lucrare aprofundează o aplicație apartinand unuia din aceste domenii.

In Capitolul 1 se definește contextul acestei tezei.

In Capitolul 2 se expune modelul teoretic PCNN.

Capitolul 3 trateaza modelul PCNN pentru segmentarea imaginilor in tonuri de gri si a imaginilor color. Versiunea propusa pentru segmentarea imaginilor in tonuri de gri se bazează pe combinarea rezultatelor a doua modele PCNN care desfășoară un număr diferit de iterații. Numărul de iterații optim pentru fiecare model se determina folosind un algoritim genetic. Metoda propusa pentru segmentarea imaginilor color bazata pe model PCNN utilizeaza manipularea matricii datelor de intrare, astfel incat toata informatia continuta initial in cele trei plane de culori fundamentale sa poata fi reprezentata intr-un singur plan de culoare. Experimentele au confirmat imbunatatirea preciziei segmentarii aplicand metoda propusa, fata de tehnicile standard.

Capitolul 4 demonstreaza robustetea modelul PCNN de segmentare in cadrul unui sistem de recunoaștere a plăcuțelor de înmatriculare (Licence Plate Recognition - LPR), experimentele fiind realizate pe o baza de date de mari dimensiuni(10 000 de evenimente, a

20 de imagini fiecare). Modelul standard de segmentare PCNN a fost imbunatatit in cadrul acestei aplicatii in 2 moduri: (a) introducerea unei reguli de oprire pentru optimizarea segmentarii; (b) dezvoltarea modelului de segmentare PCNN de la imagini cu tonuri de gri la imagini multispectrale. Faza finala a sistemului de recunoastere consta in Optical Character Recognition (OCR), care se bazeaza pe un clasificator de tip Support Vector Machine (SVM). Sunt comparate performantele sistemului LPR bazat pe segmentare PCNN cu cele ale variantelor LPR utilizand segmentarea SOM si respectiv FCM. Experimentele confirma avantajul versiunii LPR bazata pe metoda de segmentare PCNN propusa.

Capitolul 5 prezinta un sistem de recunoastere a vehiculelor militare terestre (Automatic Target Recognition – ATR) in imagini SAR. Acest algoritm consta dintr-o cascada de prelucrare a imaginilor cu mai multe etape. Prima etapa, detectia obiectelor, se realizeaza cu ajutorul unui modul PCNN. Urmeaza doua module de selectie a caracteristicilor: filtrarea Gabor (GF) si analiza componentelor principale (PCA). Ultima etapa consta din clasificatorul Support Vector Machine (SVM) imbunatatit prin utilizarea de date generate artificial (VTDG) cu Concurrent Self-Organization Map (CSOM). Pentru experimentarea modelului s-a utilizat baza de date Moving and Stationary Target Acquisition and Recognition (MSTAR), care contine 2987 imagini de 128 x 128 pixeli, folosind doua unghiuri de achizitie: 15 grade si 17 grade. Rezultatele experimentale confirma superioritatea cascadei de prelucrare ATR bazata pe segmentare PCNN fata de versiunea ATR care utilizeaza segmentare bazata pe histograma normalizata.

Capitolul 6 propune o abordare originala de diagnosticare a starii de ebrietate utilizand analiza imaginilor faciale din spectrul termal. Secventa de procesare contine urmatoorii pasi: (a) segmentarea imaginii faciale termale cu modelul PCNN; (b) selectia caracteristicilor prin analiza componentelor principale (PCA); (c) analiza discriminatorie liniara (LDA); (d) clasificare tip Support Vector Machine (SVM). In cadrul tezei s-a realizat o baza de date pentru testarea metodelor de detectie a starii de ebrietate, aceasta continand 400 de imagini faciale in spectrul infrarosu indepartat (termal), corespunzand unui lot de 10 subiecti (stare normala/ebrietate). Rezultatele experimentale confirma convingator premisele metodei.

Capitolul 7 aplica modelul PCNN in cadrul unui algoritm pentru detectia schimbarilor in serii temporale de imagini satelitare. Metoda propusa consta din urmatoarele etape: (a) calculul diferentei dintre imagini (DI); (b) evaluarea numarului total de activari pe fiecare neuron PCNN pentru un anumit interval de timp; (c) clustering cu algoritmul K-Means. Setul de date utilizat pentru experimente consta din doua imagini multispectrale Landsat-5 achizitionate in aprilie 2000 si mai 2002, intr-o zona din Mexic. Rezultatele experimentale conduc la performante foarte bune.

8.1 Rezultate obținute

În continuare vom prezenta sinteza rezultatelor obținute pe parcursul acestei teze. Fiecare element (tabel, figură sau imagine) ilustrat în continuare va conține în titlu o trimitere la apariția sa inițială în lucrare.

8.1.1 Rezultate și concluzii referitoare la segmentarea imaginilor în tonuri de gri și a imaginilor color folosind modelul Pulse Coupled Neural-Network

Modelul de segmentare PCNN este considerat atât pentru scara gri și, de asemenea, pentru imagini color. Contribuția noastră pentru segmentare folosind modelului PCNN pe imagini în tonuri de gri constă în rularea în paralel a două structuri de segmentare PCNN, în prima folosind "n" iterații PCNN iar a doua, folosind "m" iterații PCNN. Elementul de noutate al segmentării imaginilor color folosind modelul PCNN constă în a rearanja locațiile pixelilor, producând o imagine care are toate valorile pixelilor din imaginea color într-o singură imagine în tonuri de gri. Se realizează acest desiderat prin lărgirea imaginii și amplasarea pixelilor R, G, B unul după altul. Pentru a optimiza parametrii PCNN (segmentare perfectă), am aplicat algoritmi genetici (AG).

În figura 8.1 se prezintă un exemplu de segmentare a unei imagini din lotul de antrenare a AG, iar rezultatele experimentale în tabelul 8.1.

O dată dovedit că modelul este robust, pasul următor este de a confirma modelul pe imaginile din lotul de validare. Rezultatele obținute pe setul de validare sunt prezentate în tabelul 8.2.

Tab. 8.1 (3.7) Precizia Segmentării pentru datele de antrenare

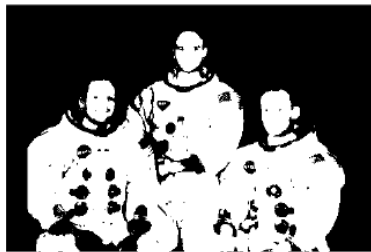
Imagine	PCNN pentru segmentarea imaginilor în tonuri de gri	PCNN pentru segmentarea imaginilor color	PCNN propus pentru segmentarea imaginilor color	FCM
1(Cal)	90.31%	93.43%	90.34%	40.77%
2(Acropole)	88.15%	72.24%	84.39%	82.45%
3(Tigru)	94.23%	91.04%	94.08%	90.97%
4(Biserica)	90.15%	87.58%	90.88%	85.47%
5(Pantera)	91.67%	87.90%	88.95%	58.16%
6(Înotător)	85.21%	82.84%	86.38%	87.28%
7(Portret)	86.26%	84.45%	85.42%	83.67%
8(Apollo)	89.35%	83.65%	87.58%	85.27%

Tab. 8.2 (3.8) Precizia Segmentarea pentru datele de validare

Imagine	PCNN pentru segmentarea imaginilor în tonuri de gri	PCNN pentru segmentarea imaginilor color	PCNN propus pentru segmentarea imaginilor color	FCM
1(Pisică)	83.07%	77.27%	79.20%	77.13%
2(Cactus)	86.91%	83.77%	85.00%	78.35%
3(Planta)	91.88%	89.66%	94.13%	86.83%
4(Zebra)	83.00%	77.81%	75.31%	66.97%



(a)



(b)



(c)



(d)

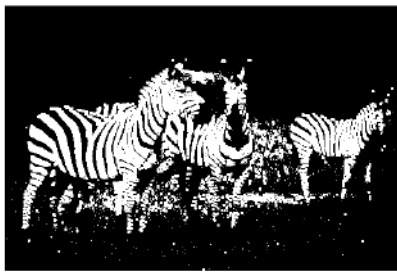


(e)

Fig. 8.1 (3.20) Segmentarea imaginii echipajului de pe Apollo, din lotul de antrenare. (a) Imagine originală; (b) PCNN pentru segmentarea imaginilor în tonuri de gri; (c) PCNN pentru segmentarea imaginilor color; (d) PCNN propus pentru segmentarea imaginilor color; (e) referința.



(a)



(b)



(c)



(d)



(e)

Fig. 8.2 (3.10) Segmentarea imaginii care conține zebre, din lotul de validare. (a) Imagine originală; (b) PCNN pentru segmentarea imaginilor în tonuri de gri; (c) PCNN pentru segmentarea imaginilor color; (d) PCNN propus pentru segmentarea imaginilor color; (e) referința.

8.1.2 Rezultate si concluzii referitoare la identificarea si recunoașterea plăcuțelor de înmatriculare

Structura de intrare clasică a neuronului PCNN este compusă din valoarea pixelului cu aceleași coordonate ca cele ale neuronului aferent și valorile pixelilor aparținând vecinătății acestuia.

În versiunea propusă de segmentare a imaginii PCNN se modifica structura de intrare a neuronului, astfel încât să conțină valorile pixelilor, nu numai dintr-un singur plan spectral, ci din toate planurile spectrale ale imaginii. In cazul acestei aplicatii, imaginile sunt formate din 3 plane spectrale, fiind obtinute folosind o camera speciala pentru sisteme LPR. Extragerea informatiei din cele trei plane spectrale este realizata prin alegerea unui plan spectral de referință, care furnizeaza valoarea variabilei lui S_{ij} din ecuația (8.1), și în loc de valorile vecinătăților vom furniza neuronului informații de la celelalte plane spectrale p .

$$\mathbf{S}_{ij} = \left\{ \begin{array}{ccc} I_{1_{i-1,j}}, & I_{2_{i-1,j}}, & I_{3_{i-1,j}}, \\ I_{1_{i,j}}, & I_{2_{i,j}}, & I_{3_{i,j}}, \\ I_{1_{i+1,j}}, & I_{2_{i+1,j}}, & I_{3_{i+1,j}} \end{array} \right\} \quad (8.1)(4.1)$$

unde S_{ij} reprezintă variabila de intrare a neuronului PCNN, $I_{q_{i,j}}$ reprezintă valoarea pixelului de la poziția i, j din planul spectral q . în acest caz S_{ij} este $I_{2_{i,j}}$ [21].

Faptul ca modelul PCNN standard nu are un mecanism intern de oprire, poate deveni o problemă importantă, deoarece este imposibil să se decidă la care iterație rezultatul modelului este favorabil.

În această situație, datorita faptului că unele informații sunt cunoscute in prealabil, referitor la imagine, putem folosi aceste cunoștințe pentru a deduce o condiție de oprire. Luând în considerare că dimensiunea unei litere este cunoscută, pot fi deduse următoarele ecuații:

$$D = \begin{cases} D + 1 & \text{daca } Y_K[n] = 1 \\ D & \text{daca } Y_K[n] = 0 \end{cases} \quad (8.2)(4.2)$$

$$F = \begin{cases} 1 & \text{daca } D > r \text{ si } D < R \\ 0 & \text{daca } D < r \text{ sau } D > R \end{cases} \quad (8.3)(4.3)$$

unde D este numărul de pixeli activați, r și R sunt dimensiunile minime si maxime a literelor. F reprezinta semaforul de oprire, unde $F=1$ semnifica stop.

In concluzie, se aduc două completări principale ale modelului PCNN clasic: un nou vector de alimentare pentru modelul PCNN și o condiție robustă de oprire. Aceste două

adaptări permit modelului PCNN să ajungă la performanțe foarte bune în segmentarea imaginilor, fără nici o intervenție umană.



Fig. 8.3 (4.6) Număr de înmatriculare murdar pe timp de zi;
 (a) Imaginea originală; (b) Rezultat PCNN; (c) Rezultat SOM;
 (d) Rezultat Mean Shift;



Fig. 8.4 (4.7) Număr de înmatriculare segmentat;
 (a) Imaginea originală; (b) Rezultat final segmentare;

Tab. 8.3 (4.1) Precizia recunoașterii caracterelor
 cu modelul SVM in faza de antrenare

	Kernel: liniar; Dimensiunea imaginii: 20x28	Kernel: liniar; Dimensiunea imaginii: 15x21	Kernel: liniar; Dimensiunea imaginii: 10x14
Rezultat	92.30%	95.24%	96.08%
	Kernel: polinomial de gradul 2; Dimensiunea imaginii: 20x28	Kernel: polinomial de gradul 2; Dimensiunea imaginii: 15x21	Kernel: polinomial de gradul 2; Dimensiunea imaginii: 10x14
Rezultat	88.27%	93.87%	95.42%

Tab. 8.4 (4.2) Precizia modelului LPR

10000 evenimente	PCNN	SOM	Mean Shift
Număr de plăcute de înmatriculare complet nerecunoscute	117	322	497
Număr de plăcute de înmatriculare nerecunoscute	530	1791	3381
Procentul de plăcute de înmatriculare recunoscute corect	93.60%	78.90%	61.30%

8.1.3 Rezultate si concluzii referitor la recunoașterea vehiculelor militare în imagini SAR

Se prezinta un sistem de recunoaștere a vehiculelor militare terestre (Automatic Target Recognition – ATR) în imagini SAR. Acest algoritm consta dintr-o cascadă de prelucrare a imaginilor cu mai multe etape. Prima etapa, detectia obiectelor, se realizează cu ajutorul unui modul PCNN. Urmează doua module de selecție a caracteristicilor: filtrarea Gabor (GF) și analiza componentelor principale (PCA). Ultima etapa consta din clasificatorul Support Vector Machine (SVM) îmbunătățit prin utilizarea de date generate artificial (Virtual Training Data Generation - VTDG) cu Concurrent Self-Organization Map (CSOM). Pentru experimentarea modelului s-a utilizat baza de date Moving and Stationary Target Acquisition and Recognition (MSTAR).

Pentru a obține o performanță îmbunătățită de segmentare, ne-am propus să rulăm în paralel două modele de segmentare PCNN, una pentru n iterații și în al doilea pentru m iterații.

Rezultatele experimentale sunt ilustrate în figura 8.5 și tabelul 8.5.

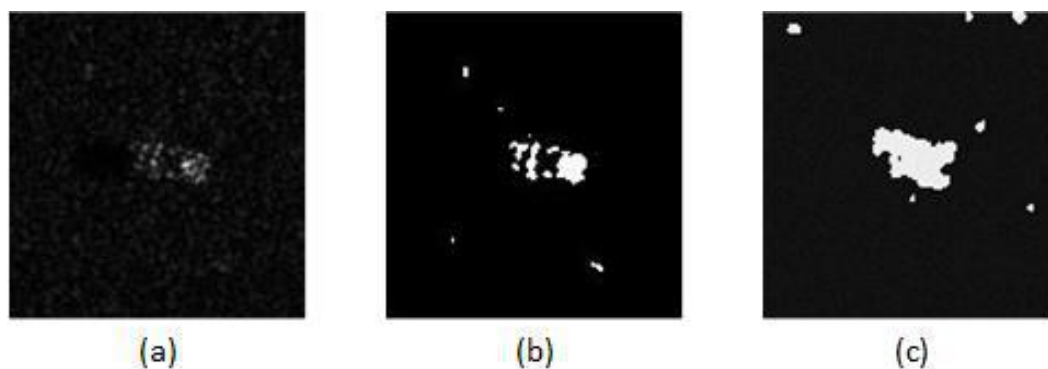


Fig. 8.5 (5.7) Ilustrarea segmentării cu modelele folosite:
 (a) Imaginea originală SAR (BMP 2);
 (b) Imaginea obținută prin segmentarea de referință;
 (c) Imaginea obținută prin segmentarea PCNN ($n = 21, m = 24$).

Tab. 8.5 (5.1) Scorul maxim de recunoaștere pentru setul de date MSTAR.
 $R = (\text{numărul de date de antrenare generate virtual folosind CSOM}) / (\text{număr de date de antrenare inițială})$.

Algoritmi	Scor de recunoaștere	Parametrii
PCNN-GF-PCA- {VTDG-CSOM}-SVM	97.36%	R=3.25; modele CSOM rectangulare de dimensiuni: (48x47), (28x27), (48x47); SVM cu nucleu RBF; gamma=0.0001
PCNN-GF-PCA- SVM	94.06%	SVM cu nucleu RBF; gamma=0.0001
RS-GF-PCA-SVM	86.15%	SVM cu nucleu RBF; gamma=0.01

8.1.4 Rezultate si concluzii referitoare la recunoașterea stării de ebrietate a subiecților

Algoritmul propus pentru diagnosticarea stării de ebrietate este compus din următoarele etape: i) captarea imaginilor termale; ii) segmentarea imaginilor cu ajutorul modelului PCNN; iii) substituirea pixelilor activi din masca de segmentare cu valorile acestora din imaginea originală; iv) selecția caracteristicilor cu metoda PCA; v) selecția caracteristicilor cu metoda LDA; vi) clasificarea pe baza metodei SVM.

Pentru evaluarea algoritmului propus s-a realizat o baza de date continand 400 de imagini faciale in spectrul termal. Rezultatele experimentale sunt ilustrate in figurile 8.6 si 8.7 si in tabelele 8.6 - 8.9.

In figura 8.6 se ilustrează diferența între imaginea unui subiect in stare normala si imaginea aceluiasi subiect in stare de ebrietate, după segmentarea cu modelul PCNN înlocuirea pixelilor activi din masca obținută în urma segmentării cu valorile acestora din imaginea originală.

In figura 8.7 se ilustrează întreg procesul de procesare a imaginilor, pana la pașii de extragerea caracteristicilor, pentru un subiect, atât in stare normala cat si in stare de ebrietate.

Un alt element de noutate, pe lângă dezvoltarea algoritmului, este faptul ca pentru realizarea acestei aplicații s-a creat o baza de date mentionata mai sus.



Fig. 8.6 (6.8) Imaginile finale prelucrate pentru a fi clasificata starea de ebrietate:
(a) normalitate; (b) ebrietate.

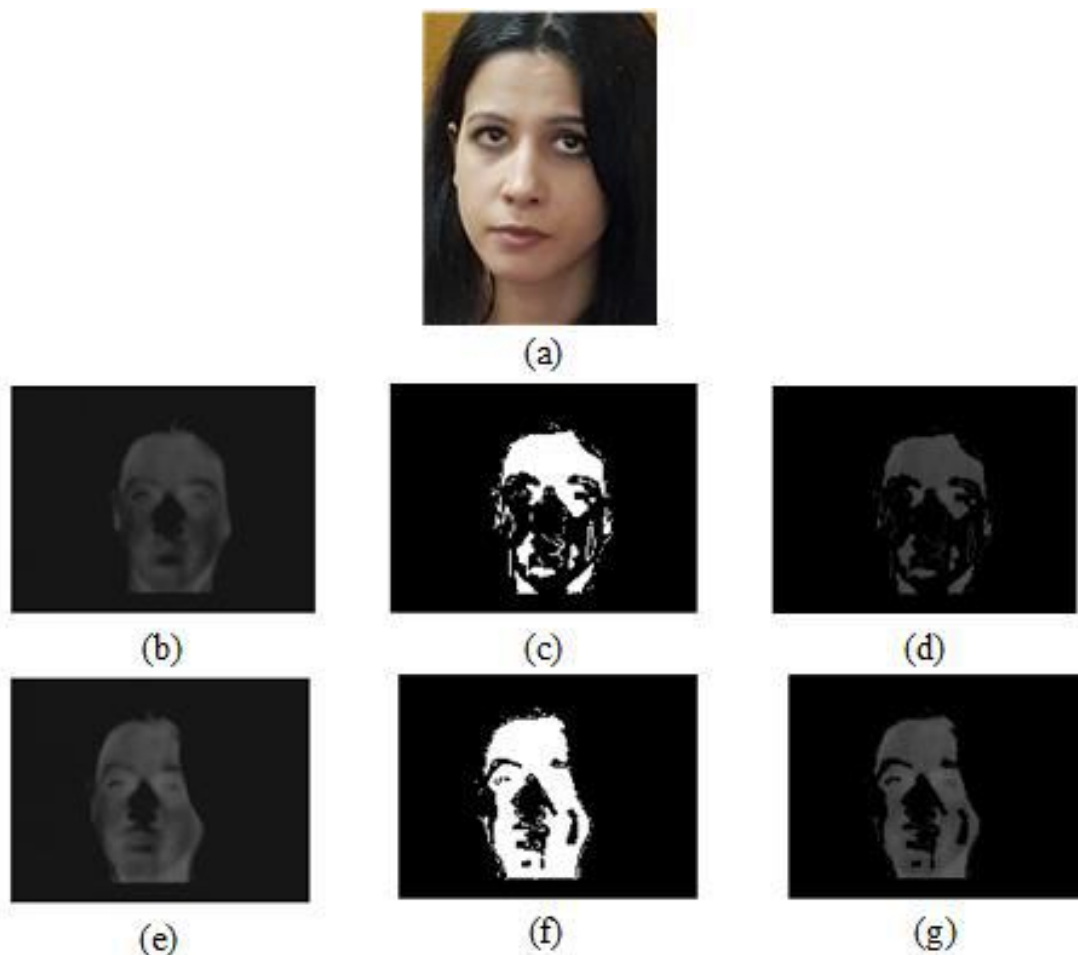


Fig. 8.7 (6.7) *Cristina: (a) imaginea în spectrul vizibil; (b) imaginea de intrare, în domeniul infraroșu, normalitate; (c) imaginea segmentată PCNN (alb corespunde segmentelor), normalitate; (d) pixelii din segmentele extrase de PCNN înlocuiți cu nivelurile de gri corespunzătoare, normalitate; (e) imaginea de intrare, în domeniul infraroșu, ebrietate; (f) imaginea segmentată PCNN (alb corespunde segmentelor), ebrietate; (g) pixelii din segmentele extrase de PCNN înlocuiți cu nivelurile de gri corespunzătoare, ebrietate;*

Tab. 8.6 (6.1) Evaluarea performanțelor, independenta de subiect, pentru identificarea stării de ebrietate prin utilizarea modelului propus,
(m PCA=150; m' LDA=1)

Nucleul si parametrii clasificatorului SVM	Dimensiunea mareicilor M si W	n (Nr. de iterații PCNN)	$V_L = V_F$	V_E	$\alpha_L = \alpha_F$	α_E	β	OA (%)	MAR (%)	FAR (%)
Liniar	4	3	0.1	0.31	2.8	16.6	0.4	97.5	1	4
Polinomial de grad 2	4	3	0.1	0.31	2.8	16.6	0.4	97.5	1	4
Polinomial de grad 3	4	3	0.1	0.31	2.8	16.6	0.4	97.5	1	4
Polinomial de grad 4	4	3	0.1	0.31	2.8	16.6	0.4	97.5	1	4
Polinomial de grad 5	4	3	0.1	0.31	2.8	16.6	0.4	97.5	1	4
RBF sigma = 0.571	6	5	0.28	0.63	36.03	2.05	3.62	97.5	1	4
RBF sigma = 1	6	5	0.28	0.63	36.03	2.05	3.62	97.5	1	4

Tab. 8.7 (6.2) Evaluarea performanțelor independenta de subiect pentru identificarea stării de ebrietate prin modelul în care nu intervine segmentarea PCNN ci imaginea este data în totalitatea sa spre etapa de extragere de trăsături.
(m PCA=150; m' LDA=1)

Nucleul si parametrii clasificatorului SVM	OA (%)	MAR (%)	FAR (%)
Liniar	77.5	24%	14%
Polinomial de grad 2	76.5	24%	16%
Polinomial de grad 3	78	20%	17%
Polinomial de grad 4	76.5	16%	23%
Polinomial de grad 5	80	8%	23%
RBF sigma=0.411	77	24%	0%
RBF sigma=1	77.5	24%	21%

Tab. 8.8 (6.3) Matricea de confuzie a subiectului "i" pentru identificarea stării de ebrietate folosind modelul propus ($m\ PCA=150$; $m' LDA=1$)

Subiectul de index "i"	1		2		3		4		5	
	normalitate	ebrietate	normalitate	ebrietate	normalitate	ebrietate	normalitate	ebrietate	normalitate	ebrietate
Starea reala → Starea dedusa ↓										
Normalitate	100%	0%	100%	0%	100%	0%	90%	10%	100%	30%
Beat	0%	100%	0%	100%	0%	100%	10%	90%	0%	70%
Subiectul de index "i"	6		7		8		9		10	
	normalitate	ebrietate	normalitate	ebrietate	normalitate	ebrietate	normalitate	ebrietate	normalitate	ebrietate
Starea reala → Starea dedusa ↓										
Normalitate	100%	0%	100%	0%	100%	0%	100%	0%	100%	0%
Beat	0%	100%	0%	100%	0%	100%	0%	100%	0%	100%

Tab. 8.9 (6.4) Matricea de confuzie globala pentru identificarea stării de ebrietate folosind modelul propus ($m\ PCA=150$; $m' LDA=1$)

Starea reala → Starea dedusa ↓	normalitate	ebrietate
	99%	4%
normalitate	99%	4%
ebrietate	1%	96%

8.1.5 Rezultate si concluzii referitor la detecția schimbărilor în serii de imagini de observație terestra

În acest capitol s-a propus un model inovativ, nesupervizat pe baza ideii prof. Neagoe, de detecție a schimbărilor bazat pe rețele neuronale, constând din următoarele etape: i) calculul diferenței dintre imagini (DI); ii) evaluarea numărului total de activări pe fiecare neuron PCNN pentru un anumit interval de timp; iii) clustering cu algoritmul K-Means.

Rezultatele experimentale pentru setul de date Mexico (Landsat-5) a condus la harta schimbărilor reprezentată în figura 8.8, corespunzând unei precizii globale de 97.66%.



(a) (b)
Fig. 8.8 (4.3) Vizualizarea rezultatelor experimentale. a) Harta schimbărilor obținută folosind modelul CVA-PCNN cu acumularea activărilor - clustering (OA = 97.86%). (b) harta de referință. (Pixeli negri sunt "schimbate"; pixelii albi sunt "neschimbate").

8.2 Contribuții originale

Contribuțiile originale aduse de autor in aceasta teza vor fi prezentate in ordinea capitolelor din lucrare.

8.2.1 Contribuții la segmentarea imaginilor in tonuri de gri si a imaginilor color (Capitolul 3):

- Un model de segmentare a imaginilor in tonuri de gri prin combinarea a doua modele de segmentare PCNN si optimizarea parametrilor cu ajutorul unui algoritim genetic;
- Un model de segmentare a imaginilor color cu modelul PCNN prin rearanjarea pixelilor din imaginea color, astfel incat toata informatia continuta initial in cele trei plane de culori fundamentale sa poata fi reprezentata intr-un singur plan de culoare;
- Implementarea si testarea algoritmului propus; interpretarea rezultatelor experimentale pentru evaluarea modelului propus.

Aceste contributii au fost prezentate in lucrarea

- S. V. Carata, V. E. Neagoe, „An Innovative Pulse-Coupled Neural Network Approach to Image Segmentation” *Proc. 6th European Conference of Computer Science (ECCS '15)*, Rome, Italy, pp. 137-141, November 2015, (BDI)

8.2.2 Contribuții la recunoașterea plăcuțelor de înmatriculare (Capitolul 4):

- Un mecanism de segmentare a imaginilor multispectrale folosind modelul PCNN;
- Un mecanism robust de stop a sistemului de segmentare PCNN;
- Un algoritm complet de recunoaștere a plăcuțelor de înmatriculare;
- Implementarea și testarea algoritmului propus pentru o largă bază de date (10 000 evenimente a 20 de imagini fiecare); interpretarea rezultatelor experimentale pentru evaluarea modelului propus.

Aceste contribuții au fost prezentate în lucrarea

- S. V. Carata, V. E. Neagoe „A Pulse-Coupled Neural Network Approach for Image Segmentation and Its Pattern Recognition Application”. *Proc. 11th International Conference on Communications (COMM)*, pp. 61-61, June 2016, (IEEE Xplore, ISI Thomson)

8.2.3 Contribuții la identificarea și recunoașterea vehiculelor militare în imagini SAR (Capitolul 5):

- Un algoritm complet pentru identificarea și recunoașterea vehiculelor militare în imagini SAR incluzând segmentarea imaginilor cu modelul PCNN;
- Implementarea și testarea algoritmului propus utilizând baza de date MSTAR, care conține 2987 imagini; interpretarea rezultatelor experimentale pentru evaluarea modelului propus.

Aceste contribuții au fost prezentate în lucrările:

- V.E. Neagoe, S.V. Carata, A. D. Cioteș, „Automatic Target Recognition in SAR Imagery Using Pulse-Coupled Neural Network Segmentation Cascaded with Virtual Training Data Generation CSOM-Based Classifier”, *Proc. International Geoscience and Remote Sensing Conference (IGARSS 2015)*, Milano, Italy, pp. 3274-3277, July 2015, (IEEE Xplore, ISI Thomson)
- V.E. Neagoe, S.V. Carata, A. D. Cioteș „An Advanced Neural Network-Based Approach for Military Ground Vehicle Recognition in SAR Aerial Imagery” *Scientific Research and Education in the Air Force-AFASES*, Brasov, pp.41-47, May 2016, (BDI)

8.2.4 Contribuții la diagnosticarea stării de ebrietate (Capitolul 6):

- Construirea unei baze de date indexate pentru testarea metodelor de detecție a stării de ebrietate, conținând 400 de imagini faciale în spectrul infraroșu îndepărtat (termal), corespunzând stărilor de normalitate / ebrietate ale subiecților;
- Un algoritm complet pentru diagnosticarea stării de ebrietate utilizând un modul PCNN pentru segmentarea imaginilor faciale din spectrul termal;
- Implementarea și testarea algoritmului propus; interpretarea rezultatelor experimentale pentru evaluarea modelului propus.

Aceste contributii au fost prezentate in lucrarea

- V.E. Neagoe, S.V. Carata „Subject Independent Drunkenness Detection Using Pulse-Coupled Neural Network Segmentation of Thermal Infrared Facial Imagery” *International Journal of Mathematical and Computational Methods*, pp. 305-312, August 2016, (BDI)

8.2.5 Contribuții la identificarea schimbărilor in serii temporale de imagini satelitare (Capitolul 7):

- O metoda inovativa, pentru detecția nesupervizata a schimbărilor in imagini satelitare observatie terestra (Earth Observation – EO) utilizand modelul PCNN, bazat pe ideea prof. Neagoe;
- Implementarea si testarea algoritmului propus; interpretarea rezultatelor experimentale pentru evaluarea modelului propus.

Aceste contributii au fost prezentate in lucrarea:

- V.E. Neagoe, A. D. Ciotec, S.V. **Carata** „A new multispectral pixel change detection approach using pulse-coupled neural networks for change vector analysis”, *Proc. International Geoscience and Remote Sensing Conference (IGARSS 2016)*, Beijing, China, pp. 3386-3389, July 2016, (IEEE Xplore, ISI Thomson)

8.3 Lista lucrărilor originale

- [1] **S. V. Carata**, V. E. Neagoe, „An Innovative Pulse-Coupled Neural Network Approach to Image Segmentation”, Proc. 6th European Conference of Computer Science (ECCS '15) Rome, Italy, pp. 137-141, November 2015, (BDI)

- [2] **S. V. Carata**, V. E. Neagoe, „A Pulse-Coupled Neural Network Approach for Image Segmentation and Its Pattern Recognition Application”, Proc. 11th International Conference on Communications (COMM), pp. 61-61, June 2016, (IEEE Xplore, ISI Thomson)

- [3] V.E. Neagoe, **S.V. Carata**, A. D. Ciotec, „Automatic Target Recognition in SAR Imagery Using Pulse-Coupled Neural Network Segmentation Cascaded with Virtual Training Data Generation CSOM-Based Classifier”, Proc. International Geoscience and Remote Sensing Conference (IGARSS 2015), Milano, Italy, pp. 3274-3277, July 2015, (IEEE Xplore, ISI Thomson)

- [4] V.E. Neagoe, **S.V. Carata**, A. D. Ciotec, „An Advanced Neural Network-Based Approach for Military Ground Vehicle Recognition in SAR Aerial Imagery”, Proc. Scientific Research and Education in the Air Force-AFASES, Brasov, pp.41-47, May 2016, (BDI)

- [5] V.E. Neagoe, **S.V. Carata**, „Subject Independent Drunkenness Detection Using Pulse-Coupled Neural Network Segmentation of Thermal Infrared Facial Imagery”, International Journal of Mathematical and Computational Methods, pp. 305-312, August 2016, (BDI)

- [6] V.E. Neagoe, A. D. Ciotec, **S.V. Carata**, „A new multispectral pixel change detection approach using pulse-coupled neural networks for change vector analysis”, Proc. International Geoscience and Remote Sensing Conference (IGARSS 2016), Beijing, China, pp. 3386-3389, July 2016, (IEEE Xplore, ISI Thomson)

8.4 Perspective de dezvoltare ulterioară

Vom prezenta mai departe perspectivele de dezvoltare pentru metodele propuse in aceasta lucrare.

In cazul aplicatiei de segmentare a imaginilor (capitolul 3), se pot testa alte metode de rafinare a parametrilor modelului PCNN in vederea obtinerii unor performante imbunatatite. O alta directie este sa se testeze utilizarea unei metode de adaptare automata a parametrilor in functie de imaginea de segmentat.

Pentru aplicatia de recunoastere a placutelor de inmatriculare, prezentata in capitolul 4, o potentiala directie de dezvoltare este combinarea metodei de segmentare a imaginilor prezentata in capitolul 3 cu metoda de segmentare prezentata in aplicatia din capitolul 4.

Algoritmul de recunoastere a vehiculelor militare terestre in imagini SAR (capitolul 5) ar putea beneficia de adaugarea unui algoritm genetic pentru rafinarea parametrilor metodei de segmentare PCNN.

Algoritmul propus de diagnosticare a starii de ebrietate (capitolul 5) ar putea beneficia de extinderea bazei de date; o baza de date mai larga pentru antrenare poate conduce la performante superioare ale clasificatorului.

Performantele metodei bazate pe model PCNN pentru detectia schimbarilor in imagini satelitare multitemporale de observatie terestra (capitolul 6), se poate evalua mai complet folosind pentru experimente alaturi de baza de date de imagini multispectrale Mexico utilizata in teza si alte baze de date satelitare multitemporale : imagini multispectrale sau imagini SAR (Synthetic Aperture Imagery).

Bibliografie

- [1] Eckhorn R, Reitboeck HJ, Arndt M et al „Feature linking via synchronization among distributed assemblies: simulation of results from cat cortex.” *Neural Computation* 2(3), pp. 293 – 307, 1990
- [2] Johnson JL „Pulse-coupled neural nets: translation, rotation, scale, distortion, and intensity signal invariance for images”. *Applied Optics* 33(26), pp. 6239 – 6253, 1994
- [3] Johnson JL, Padgett ML, „PCNN models and applications” *IEEE Transactions on Neural Networks* 10(3), pp. 480 – 498, 1999
- [4] Y. Ma, K. Zhan, Z. Wang, „Applications of Pulse-Coupled Neural Networks”, Springer, 2010.
- [5] Dunn, J. C., "A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters". *Journal of Cybernetics* 3 (3): 32–57. 1973-01-01
- [6] Ahmed, Mohamed N., Yamany, Sameh M., Mohamed, Nevin, Farag, Aly A., Moriarty, Thomas. "A Modified Fuzzy C-Means Algorithm for Bias Field Estimation and Segmentation of MRI Data" (PDF). *IEEE Transactions on Medical Imaging* 21 (3), pp: 193–199, 2002
- [7] Kohonen, Teuvo; Honkela, Timo (2011). "Kohonen network". *Scholarpedia*. Retrieved 2012-09-24.
- [8] K. F. Man, K. S. Tang, and S. Kwong, „Genetic Algorithms: Concepts and Applications”, *IEEE Transactions on Industrial Electronics*, Vol. 43, No. 5, pp. 519 – 534, October 1996
- [9] Mitchell, Melanie, „An Introduction to Genetic Algorithms”. Cambridge, MA: MIT Press, pp. 2, 1996
- [10] Whitley, Darrell, "A genetic algorithm tutorial". *Statistics and Computing* 4 (2), pp. 65–85, 1994
- [11] G. Kuntimad and H. S. Ranganath, „Perfect Image Segmentation Using Pulse Coupled Neural Networks”; *IEEE Transactions on Neural Networks*, Vol. 10, No. 3, pp. 591 – 598 , May 1999
- [12] S. S. Adagale, S. S. Pawar, „Image segmentation using PCNN and template matching for blood cell counting”; *IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, pp. 1-5, 2013

- [13] T. Lindblad and J. M. Kinser, "Image Processing Using Pulse Coupled Neural Networks", Springer, 2010
- [14] S. V. **Carata**, V. E. Neagoe, „An Innovative Pulse-Coupled Neural Network Approach to Image Segmentation”, 6th European Conference of Computer Science (ECCS '15) Rome, Italy, pp. 137-141, November 7-9, 2015
- [15] H. Zhuang et al, "Multichannel Pulse-Coupled-Neural-Network-Based Color Image Segmentation for Object Detection", IEEE Trans. On Industrial Electronics, vol 59, nr. 8, Aug 2012, pp. 3299-3308.
- [16] Cortes, C.; Vapnik, V., "Support-vector networks". Machine Learning 20 (3), pp. 273–297, 1995.
- [17] Cristianini, Nello; and Shawe-Taylor, John, „An Introduction to Support Vector Machines and other kernel-based learning methods”, Cambridge University Press, 2000.
- [18] R. C. Gonzalez, R. E. Woods, "Digital image processing", Prentice Hall, 2008
- [19] J. Shi, J. Malik, „Normalized Cuts and Image Segmentation”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 8, pp. 888 – 905, August 2000.
- [20] R. Szeliski, "Computer Vision: Algorithms and Applications", Springer, 2010
- [21] S. V. **Carata**, V. E. Neagoe , „A Pulse-Coupled Neural Network Approach for Image Segmentation and Its Pattern Recognition Application”, 11th International Conference on Communications (COMM), pp. 61-61, June 2016, (IEEE Xplore, ISI Thomson)
- [22] A. Streicher, E-Learning for radar image interpreters, in: New Security Learning, <http://www.newsecuritylearning.com/index.php/feature/105-e-learning-for-radar-image-interpreters>.
- [23] Aach, T.; Kaup, A.; Mester, R.: „On texture analysis: Local energy transforms versus quadrature filters." In Signal Processing, vol. 45, pp. 173-181, 1995.
- [24] Fleet, D. J.; Jepson, A. D.: „Stability of Phase Information." In IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 15, no. 12, pp. 1253-1268, December 1993.
- [25] Fleet, D. J.; Jepson, A. D.; Jenkin, M. R. M.: „Phase-Based Disparity Measurement." In Computer Vision, Graphics and Image Processing: Image Understanding, vol. 53, no. 2, pp. 198-210, March 1991.
- [26] Jenkin, M. R. M.; Jepson, A. D.: „The Fast Computation of Disparity from Phase Differences." In Proceedings of the IEEE Computer Vision and Pattern Recognition, pp. 398-403, 1989.
- [27] Jenkin, M. R. M.; Jepson, A. D.: „Recovering Local Surface Structure through Local Phase Difference Measurements." In Computer Vision, Graphics and Image Processing: Image Understanding, vol. 59, no. 1, pp. 72-93, January, 1994.
- [28] Jenkin, M. R. M.; Jepson, A. D.; Tsotsos, J. K.: „Techniques for Disparity Measurement." In Computer Vision, Graphics and Image Processing: Image Understanding, vol. 53, no. 1, pp. 14-30, January, 1991.
- [29] Maki, A.; Bretzner, L.; Eklundh, J.: „Local Fourier Phase and Disparity Estimates: An Analytical Study." In Hlavac, Sara (eds.): Proceedings of the 6th International Conference on Computer Analysis of Images and Patterns, Springer-Verlag, Berlin, Heidelberg, New York, pp. 868-873, September 1995.
- [30] Sanger, T. D.: „Stereo Disparity Computation Using Gabor Filters." In Biological Cybernetics, vol. 59, pp. 405-418, 1988.

- [31] Daugman, J. G.: „Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters." J. Optical Society of America A, vol. 2, no. 7, pp. 1160-1169, July 1985.
- [32] Daugman, J. G.: „Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters." J. Optical Society of America A, vol. 2, no. 7, pp. 1160-1169, July 1985
- [33] inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/TRAPP1/filter.html
- [34] Pearson, K. , "On Lines and Planes of Closest Fit to Systems of Points in Space" .Philosophical Magazine. pp. 559–572. 1911.
- [35] Hotelling, H., „Analysis of a complex of statistical variables into principal components”, Journal of Educational Psychology, pp. 417–441, and 498–520, 1933.
- [36] Hotelling, H., „Relations between two sets of variates”, Biometrika, pp. 321–77, 1936
- [37] V.E. Neagoe, A.D. Ciotec, New Approach for Accurate Classification of Hyperspectral Images Using Virtual Sample Generation by Concurrent Self-Organizing Maps, Proc. IEEE Internat. Geoscience and Remote Sensing Conf. (IGARSS), Melbourne (Australia), pp. 1031-1034, July 21-26 2013
- [38] V.E. Neagoe and A. Ropot, Concurrent Self-Organizing Maps - A Powerful Artificial Neural Tool for Biometric Technology, pp. 291-298, in: Harbour Protection Through Data Fusion Technologies, Edited by: Shahbazian, E; Rogova, G; DeWeert, MJ, NATO Science for Peace and Security Series-C: Environmental Security, Springer, 2009.
- [39] V.E. Neagoe, R.M. Stoica, A.I. Ciurea, L. Bruzzone, F. Bovolo, Concurrent Self-Organizing Maps for Supervised/Unsupervised Change Detection in Remote Sensing Images, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, Vol. 7, No. 8, pp. 3525-3533, August 2014.
- [40] M. Sezgin, B. Sankur, Survey over Image Thresholding Techniques and Quantitative Performance Evaluation, Journal of Electronic Imaging, Vol. 13, no 1, pp. 146–165, Jan. 2004.
- [41] V.E. Neagoe, S.V. **Carata**, A. D. Ciotec, “Automatic Target Recognition in SAR Imagery Using Pulse-Coupled Neural Network Segmentation Cascaded with Virtual Training Data Generation CSOM-Based Classifier”, Proceedings of International Geoscience and Remote Sensing Conference (IGARSS 2015), Milano, Italy, pp. 3274-3277, July 2015
- [42] Y. Adini, Y. Moses, and S. Ullman, Face recognition: The problem of compensating for changes in illumination direction, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 19, No. 7, July 1997, pp. 721-732.\
- [43] D.A. Socolinsky and A. Selinger, A Comparative Analysis of Face Recognition Performance with Visible and Thermal Infrared Imagery, Proc. 16th Int'l Conf. Pattern Recognition, Vol. 4, pp. 217-222, 2002.
- [44] D.A. Socolinsky, A. Selinger, and J.D. Neuheisel, Face recognition with visible and thermal infrared imagery, Computer Vision and Image Understanding, Vol. 91, pp. 72-114, 2003.
- [45] J. Heo, S.G. Kong, B.R. Abidi, and M.A. Abidi, Fusion of Visual and Thermal Signatures with Eyeglass Removal for Robust Face Recognition, IEEE Workshop on ObjectTracking and Classification Beyond the Visible Spectrum in conjunction with CVPR 2004, July 2004, Washington DC, pp. 94-99.

- [46] P. Buddharaju, I.T. Pavlidis, P. Tsiamyrtzis, M. Bazakos, Physiology-Based Face Recognition in the Thermal Infrared Spectrum, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 29, No. 4, April 2007, pp. 613-626.
- [47] V.E. Neagoe, A. Mugioiu and C. Tudoran, Concurrent Self-Organizing Maps for Multispectral Facial Image Recognition, *Proc. of the 2007 IEEE Symposium on Computational Intelligence in Image and Signal Processing (CIISP 2007)*, April 1-5, 2007, Honolulu, Hawaii, USA, pp. 330-335, ISBN:1-4244-0707-9.
- [48] V.E. Neagoe, A. Ropot, and A. Mugioiu, Real Time Face Recognition Using Decision Fusion of Neural Classifiers in the Visible and Thermal Infrared Spectrum, *Proc. of the 2007 IEEE International Conference on Advanced Video and Signal based Surveillance (AVSS 2007)*, London (United Kingdom), 5-7 September 2007, pp. 1-6, ISBN:978-1-4244-1696-7
- [49] V.E. Neagoe, A.P. Barar, N. Sebe, P. Robitu, A Deep Learning Approach for Subject Independent Emotion Recognition from Facial Expressions, *Recent Advances in Image, Audio, and Signal Processing, Proc. 1st International Conference on Image Processing and Pattern Recognition (IPPR13)*, Budapest (Hungary), Dec. 10-12, 2013, pp. 93-98, ISBN: 978-960-474-350-6, ISSN:1790-5117.
- [50] UN, Road Safety report, 2013. [Online]. Available: <http://www.un.org/ar/roadsafety/pdf/roadsafetyreport.pdf>.
- [51] D. Isa, G.C. Chieh, R. Arelhi, Abnormal Driver Behavior Detection Using Parallel CPU and GPU Algorithm through Facial Expression, Thermal Imaging and Heart Rate Data Fusion, *Proc. International Journal of Engineering and Innovative Technology (IJEIT)*, Vol 2, Issue 3, September 2012, pp. 53-64, ISSN: 2277-3754.
- [52] B. Hunicka, H. Laurell, and H. Bergman, Psychosocial Characteristics of Drunk Drivers Assessed by Addiction Severity Index, Prediction of Relapse, *Scandinavian Journal of Public Health*, February, Vol. 38, No. 1, 2010, pp.71-77.
- [53] Y.C. Wu, Y.Q. Xia, P. Xie, and X.W. Ji, (2009) The Design of an Automotive Anti-Drunk Driving System to Guarantee the Uniqueness of Driver, *Proc. International Conference on Information Engineering and Computer Science (ICIECS 2009)*, December 2009, pp.1-4.
- [54] G. Koukiou and V. Anastassopoulos, Drunk Person Identification Using Thermal Infrared Images, *Int. J. Electronic Security and Digital Forensics*, Vol. 4, No. 4, 2012, pp. 229-243.
- [55] G. Koukiou and V. Anastassopoulos, Face Locations Suitable Drunk Persons Identification, *2013 International Workshop on Biometrics and Forensics (IWBF)*, Lisbon, 4-5 April 2013, pp. 1-4, ISBN 978-1-4673-4987-1.
- [56] G. Koukiou and V. Anastassopoulos, Eye Temperature Distribution in Drunk Persons using Thermal Imagery, *Proc. 2013 International Conference of the Biometrics Special Interest Group (BIOSIG)*, Darmstadt, 5-6 Sept. 2013, pp. 233-240.
- [57] G. Koukiou and V. Anastassopoulos, Neural Networks for Identifying Drunk Persons using Thermal Infrared Imagery, *Forensic Science International*, Vol. 252, July 2015, pp. 69-76.
- [58] R.L. Haupt, S.E. Haupt, *Practical Genetic Algorithms*, (2nd Ed.), Wiley, New York, 2004.
- [59] M. Bishop, *Pattern Recognition and Machine Learning*, Springer, New York, 2006.
- [60] V.E. Neagoe, **S.V. Carata**, „Subject Independent Drunkenness Detection Using Pulse-Coupled Neural Network Segmentation of Thermal Infrared Facial Imagery”,

International Journal of Mathematical and Computational Methods, pp. 305-312, August 2016, (BDI)

- [61] L. Bruzzone and D.F. Pretto, "Automatic analysis of the difference image for unsupervised change detection," IEEE Trans. Geoscience and Remote Sensing, vol. 38, no. 3, pp. 1171-1182, May 2000.
- [62] A. Singh, "Digital change detection techniques using remotely-sensed data," Int. J. Remote Sensing, vol. 10, no. 6, 1989.
- [63] V. Neagoe, A. Ciurea, L. Bruzzone, F. Bovolo, „A novel neural approach for unsupervised change detection using SOM clustering for pseudo-training set selection followed by CSOM classifier”, Proc. 2014 IEEE Int. Geoscience and Remote Sensing Symposium, July 13-18, 2014, Quebec-city, Canada, pp. 1437-1440.
- [64] V.E. Neagoe, S.V. **Carata**, „A new multispectral pixel change detection approach using pulse-coupled neural networks for change vector analysis”,
- [65] V. Neagoe, Inteligenta computationala, capitol al cartii Enciclopedia Matematica , (coordonatorii volumului M. Iosifescu si O. Stanasila), Ed. AGIR, Bucuresti, 2010.
- [66] Francesca Bovolo, Lorenzo Bruzzone, „A Theoretical Framework for Unsupervised Change Detection Based on Change Vector Analysis in the Polar Domain”, Ieee Transactions on Geoscience and Remote Sensing, Vol. 45, No. 1, January 2007
- [67] A. Streicher, E-Learning for radar image interpreters, in: *New Security Learning*, <http://www.newsecuritylearning.com/index.php/feature/105-e-learning-for-radar-image-interpreters>.
- [68] F. Del Frate, F. Pacifici, G. Schiavon, and C. Solimini, "Use of neural networks for automatic classification from high-resolution images," IEEE Trans. Geosci. Remote Sens., vol. 45, no. 4, pp. 800–809, Apr. 2007.
- [69] F. Pacifici, F. del Frate, C. Solimini, and W. Emery, "An innovative neural-net method to detect temporal changes in high resolution optical satellite imagery," in IEEE Trans. Geosci. Remote Sens., vol. 45, pp. 2940-2952, 2007.
- [70] V.E. Neagoe, S.V. **Carata**, A. D. Ciotec „An Advanced Neural Network-Based Approach for Military Ground Vehicle Recognition in SAR Aerial Imagery" Scientific Research and Education in the Air Force-AFASES, Brasov, pp.41-47, May 2016
- [71] O. Tifrea, „Recunoasterea umorului in texte” , Teza de doctorat (Conducator Prof. Dan Cristea), Universitatea A.I. Cuza, Iasi, 2008