



**UNIVERSITATEA „POLITEHNICA” din BUCUREȘTI
ȘCOALA DOCTORALĂ ETTI-B**

Rezumat TEZĂ DE DOCTORAT

**METODE ȘI ALGORITMI DE EXTRACȚIE ȘI
CĂUTARE DE INFORMAȚII ÎN SERII TEMPORALE
DE IMAGINI SATELITARE**

**Doctorand: Cpt. Ing. Alexandru-Cosmin Grivei
Conducător de doctorat: Prof. Dr. Ing. Mihai Dateu**

BUCUREȘTI 2021

Cuprins

Capitolul 1	Introducere	1
1.1	Prezentarea domeniului tezei de doctorat.....	1
1.2	Scopul tezei de doctorat	1
1.3	Conținutul tezei de doctorat	2
Capitolul 2	Noțiuni privind achizițiile de teledetecție utilizate	3
2.1	Principii de teledetecție	3
2.2	Instrumente de teledetecție.....	3
2.3	Serii temporale de imagini satelitare.....	4
2.4	Concluzii	4
Capitolul 3	Metode și algoritmi utilizați pentru extracția de informație din date de teledetecție	5
3.1	Concepte existente pentru extragerea de informație din date de teledetecție satelitară	5
3.2	Metode de extragere de informații din date de teledetecție satelitară utilizate.....	5
3.3	Metode de clasificare a datelor obținute din teledetecție satelitară.....	6
3.4	Metrice utilizate pentru determinarea performanțelor metodelor de extragere de informație	6
3.5	Concluzii	6
Capitolul 4	Metode și sisteme software dezvoltate pentru extragerea de informații cantitative din serii temporale de imagini satelitare	7
4.1	Metodă propusă pentru analiza automată a seriilor temporale de imagini satelitare	7
4.2	Sistem software pentru extragerea de informații din date Sentinel-2	9
4.3	Sistem software pentru îmbunătățirea rezultatelor de clasificare a datelor de teledetecție.....	10
4.4	Sistem software pentru extragerea de informații din volume foarte mari de date de teledetecție	12
4.5	Concluzii de capitol.....	13
Capitolul 5	Metode și sisteme software dezvoltate pentru identificarea și cuantificarea schimbărilor din serii temporale de imagini satelitare	15
5.1	Metodă de indexare și extracție de informații bazată pe algoritmi de hashing	15

5.2	Sistem software pentru construirea și analiza temporală a seriilor temporale de imagini satelitare	16
5.3	Sistem software pentru extragerea de informații din serii temporale de imagini satelitare	18
5.4	Platformă web pentru prezentarea informațiilor extrase din serii temporale de imagini satelitare	19
5.5	Concluzii de capitol.....	21
Capitolul 6	Concluzii.....	22
6.1	Contribuții originale	22
6.2	Lista lucrărilor originale.....	24
6.2.1	Articole de jurnal	24
6.2.2	Articole de conferință	24
6.3	Perspective de dezvoltare ulterioară.....	26
Anexe.....		27
A.1	Contribuții în proiecte finanțate de ESA	27
Bibliografie		29

Capitolul 1 Introducere

Primele misiuni de observare a Pământului bazate pe sateliți datează de la sfârșitul anilor '50, odată cu lansarea primului satelit artificial, *Sputnik-1*, ce avea scopul de a studia ionosfera [1]. O dată cu trecerea timpului a crescut și interesul mai multor țări și organizații pentru observarea Pământului. Prin urmare, cantitatea de date colectate de-a lungul timpului a crescut considerabil, de multe ori depășind capacitatea de procesare.

1.1 Prezentarea domeniului tezei de doctorat

Așa cum se întâmplă și în cazul altor domenii din sfera *big data* [2], cei care au ca obiectiv procesarea volumelor mari de date provenite de la sateliți se vor confrunta cu probleme legate de stocarea și organizarea a datelor, manipularea, extragerea de informații utile și prezentarea acestora.

Datele rezultate în urma observațiilor satelitare trec printr-o serie de etape pentru putea fi valorificate de către utilizatorul final. Aceste etape, sunt în ordine, următoarele: creare, stocare, procesare și prezentare. În cazul datelor produse de misiunile de teledetecție, cei care doresc să le proceseze se lovesc problema celor 5V (Volum, Viteză, Varietate, Veridicitate și Valoare) [3].

1.2 Scopul tezei de doctorat

Achizițiile ce acoperă același suprafață geografică formează serii temporale. În cadrul acestora se pot observa diferite tendințe sau se pot cuantifica schimbările survenite asupra unei zone într-un interval de timp dat.

Teza tratează problemele ridicate de procesarea seriilor temporale de imagini satelitare (STIS), atât din punct de vedere al extragerii de informații, cât și al manipulării lor în vederea atingerii acestui scop.

O primă direcție de cercetare este reprezentată de dezvoltarea de algoritmi, metode și sisteme software capabile să proceseze volume mari de date în vederea extragerii de informații utile, într-un timp relativ scurt. Luând în considerare tipurile de informații latente în datele de teledetecție s-au tratat două abordări:

- Prima constă în analiza individuală a achizițiilor ce formează o serie temporală, și prezentarea rezultatelor sub forma unor grafice ce conțin evoluții temporale cantitative sau statistice.
- Cea de-a doua abordare constă în analiza STIS ca o macrostructură de date în care se urmărește identificarea zonelor în care se produc schimbări (cunoscute sau nu) ale suprafeței Pământului.

Pentru a permite analiza facilă a volumelor mari de date teza tratează și metode de indexare a achizițiilor de teledetecție în vederea generării rapide de serii temporale pentru o zonă geografică de interes, cât și pentru a permite funcționalități de tip căutări bazate pe conținut.

1.3 Conținutul tezei de doctorat

Teza prezintă o serie de implementări software pentru extragerea de informație din serii temporale de imagini satelitare, optimizarea procesului de căutare în arhive de date de teledetecție, și prezentarea rezultatelor, fiind structurată pe șase capitole.

Capitolul 2 prezintă caracteristicile datelor de teledetecție utilizate în cadrul sistemelor software dezvoltate. Informațiile prezentate sunt grupate, în funcție de tipul de achiziție utilizat de senzorii ambarcați pe satelit, multi-spectrală, respectiv radar.

În capitolul 3 sunt descrise metodele și algoritmi dezvoltați și utilizați pentru extragerea de informații din datele de teledetecție, ce au fost integrați în sistemele software dezvoltate.

Capitolul 4 prezintă patru sisteme software ce facilitează căutarea și extragerea rapidă și precisă a informațiilor latente din achiziții individuale de teledetecție.

În cadrul capitolului 5 sunt prezentate metodele și sistemele software dezvoltate ce permit manipularea volumelor mari de date, indexarea datelor de teledetecție, generarea rapidă de serii temporale, extragerea de informație, și prezentarea rezultatelor într-o manieră interactivă cu ajutorul unui sistem bazat pe tehnologii de tip web.

În ultimul capitol, 6, sunt prezentate concluziile pentru sistemele software dezvoltate pe parcursul tezei, contribuțiile originale obținute, lista articolelor publicate, cât și perspective pentru dezvoltări ulterioare.

Capitolul 2 Noțiuni privind achizițiile de teledetecție utilizate

În cadrul acestui capitol sunt prezentate caracteristicile fiecărui senzor ale cărui achiziții au fost utilizate pe parcursul tezei pentru demonstrarea funcționalităților și performanțelor sistemelor software dezvoltate pentru extragerea de informații.

2.1 Principii de teledetecție

În funcție de modul în care se realizează achiziția imaginilor satelitare, senzorii utilizați se împart în două categorii: multi-spectrali (pasivi), și radar (activi). Cei multi-spectrali înregistrează radiația spectrului electromagnetic provenită de la sol ca urmare a iluminării realizate de Soare. Metodele active, de tip RAS presupun iluminarea solului de către instrumentele amplasate la bordul sateliților, înregistrând ulterior semnalele reflectate de suprafața Pământului.

În continuare, sunt prezentate caracteristicile senzorilor, pasivi și activi, ce au achiziționat datele utilizate în sistemele software dezvoltate și prezentate în cadrul tezei.

2.2 Instrumente de teledetecție

Cea mai mică structură de date dintr-o imagine satelitară este pixelul. Informația stocată în fiecare pixel depinde de rezoluția spațială, spectrală și radiometrică a senzorului ce face achiziția.

În cazul instrumentelor multi-spectrale, rezoluția spațială definește care sunt cele mai mici structuri observabile pe Pământ cu ajutorul unui anumit senzor. Acestea se încadrează în următoarele categorii [4]: rezoluție grosieră (> 1000 m), rezoluție medie (100-1000 m), rezoluție fină (5-100 m), rezoluție foarte fină (< 5 m).

Rezoluția spectrală este definită prin numărul și lățimea benzilor spectrale folosite pentru achiziția datelor. În funcție de numărul benzilor fiecare senzor se încadrează în una dintre categoriile: pancromatic (o bandă), multi-spectral (3 benzi sau mai multe), sau hiper-spectral (zeci sau sute de benzi).

Rezoluția radiometrică face referire la gama dinamică a senzorului, sau numărul de nivele a radiației electromagnetice recepționate pe fiecare bandă spectrală.

În format digital aceste nivele sunt date de numărul de biți folosiți pentru reprezentare.

Rezoluția temporală este definită prin intervalul de timp scurs între două vizități succesive ale aceleiași suprafețe a Pământului.

2.3 Serii temporale de imagini satelitare

Seriile Temporale de Imagini Satelitare reprezintă un set de imagini achiziționate pe aceeași zonă geografică la momente diferite de timp. Pentru a obține serii mai dense, pentru un interval relativ scurt de timp, STIS pot fi formate și din achiziții provenite de la misiuni satelitare diferite. Senzorii cu rezoluții spațiale și temporale ridicate permit observarea mai precisă a structurilor și tiparelor existente în cadrul scenelor dinamice.

2.4 Concluzii

Din punct de vedere al eșantionării temporale, achizițiile de tip RAS nu sunt afectate de condițiile meteorologice sau de nivelul de iluminare al scenelor făcând mult mai fezabile aceste tipuri de date pentru monitorizarea schimbărilor rapide ce pot surveni în cazul unor calamități naturale precum uragane, inundații, incendii de vegetație ș.a. Din păcate acest tip de date necesită un nivel de procesare mult mai ridicat față de achizițiile de multi-spectrale, având nevoie de metode avansate pentru reducerea zgomotului din imagini.

Achizițiile de tip multi-spectral permit o mai bună identificare a claselor de acoperire de la nivelul solului, în cazul unor achiziții realizate în condiții optime de iluminare și în absența fenomenelor meteorologice și a perturbațiilor atmosferice. De asemenea achiziționarea datelor cu o rezoluție spectrală variată permite o procesare mai rapidă a datelor pentru diferite aplicații.

Capitolul 3 Metode și algoritmi utilizați pentru extracția de informație din date de teledetecție

În cadrul acestui capitol sunt prezentați algoritmi de extragere de trăsături și clasificare implementați în cadrul instrumentelor software dezvoltate, cu scopul de a oferi soluții pentru problemele legate de manipularea datelor rezultate din observarea Pământului și extragerea de informații conținute în acestea.

3.1 Concepte existente pentru extragerea de informație din date de teledetecție satelitară

Un sistem software capabil să automatizeze procesul de adnotare a datelor de teledetecție și să genereze informații exacte pentru orice tip de date, pentru orice scenariu posibil va avea mereu un nivel ridicat de complexitate. În dezvoltarea unui astfel de sistem este foarte importantă optimizarea componentei software, flexibilitatea și scalabilitatea fiind puncte cheie pentru datele de teledetecție datorită problemelor asociate celor 5V. Din acest motiv, este preferabilă utilizarea unei arhitecturi modulare [5], capabile să integreze mult mai ușor noi funcționalități în eventualitatea utilizării unor noi tipuri de date. De asemenea, este recomandată utilizarea arhitecturilor software stratificate, așa cum se sugerează în [6] și [7], cu funcționalități și logici bine definite.

3.2 Metode de extragere de informații din date de teledetecție satelitară utilizate

Un prim pas în procesul de extragere de informații din imagini satelitare constă în descrierea conținutului existent cu ajutorul reprezentărilor matematice bazate pe trăsăturile spectrale, textură sau formă. Aceste reprezentări sunt utilizate în procesul de clasificare pentru crearea unor modele ce pot asocia etichete semantice reprezentărilor matematice. Algoritmi de clasificare se împart în supervizați, în care utilizatorul trebuie să ofere exemple de elemente ce trebuie asociate unei clase

semantice, și nesupervizați, în care asocierea elementelor se realizează automat în funcție de reprezentarea matematică a acestora și de funcția de optimizare utilizată.

3.3 Metode de clasificare a datelor obținute din teledetecție satelitară

Algoritmii de clasificare permit identificarea de structuri similare în cadrul datelor de teledetecție, reprezentând un mijloc important pentru extragerea de informații. Aceștia se împart în algoritmi nesupervizați, și algoritmi supervizați. Primii încearcă identificarea, în mod automat, a unor elemente reprezentative (centroizi) ce pot fi asociate unor clase semantice. Numărul centroizilor este dictat de către utilizator și are un rol important în rezultatele generate pentru majoritatea algoritmilor nesupervizați. Algoritmii supervizați au nevoie de exemple pentru a stabili modul de distribuire a datelor în clase semantice.

3.4 Metrice utilizate pentru determinarea performanțelor metodelor de extragere de informație

În procesul de testare a performanțelor algoritmilor de clasificare implementați în cadrul instrumentelor software pentru căutarea informației în date de teledetecție, pe parcursul tezei am utilizat o serie de metrice pentru evaluarea performanțelor acestora: acuratețe, precizie, recall, F-score, și specificitate. Aceste metrice sunt adesea utilizate pentru stabilirea nivelului de performanță a algoritmilor și metodelor pentru căutarea informației, clasificare, învățare, recunoaștere de tipare ș.a..

3.5 Concluzii

În cadrul acestui capitol am prezentat o serie de concepte existente pentru extragerea de informații din date de teledetecție. Acestea rezolvă diverse probleme legate de automatizarea procesului de extragere de informații în domeniul observării Pământului.

Extragerea de informații din datele de teledetecție se realizează cu ajutorul algoritmilor de extragere de trăsături și clasificare. Acești algoritmi au fost integrați în cadrul sistemelor software dezvoltate, dovedind performanțe bune în cadrul diferitelor scenarii prezentate în capitolele 4 și 5. Performanțele obținute au fost evaluate cu ajutorul diferitelor metrice, precum: acuratețe, precizie, recall, specificitate și F-score.

Capitolul 4 Metode și sisteme software dezvoltate pentru extragerea de informații cantitative din serii temporale de imagini satelitare

În cadrul acestui capitol sunt prezentate o serie de instrumente software și metode dedicate procesării și extragerii de informații din date rezultate din observarea Pământului.

Prima metodă, prezentată, utilizează reprezentări bazate pe indici spectrali pentru a extrage în mod automat și rapid, informații cantitative, privind acoperirea diferitelor clase semantice dintr-o anumită zonă geografică de interes.

În partea doua a capitolului, este prezentat un instrument software de tip plug-in ce facilitează extragerea de informații din date Sentinel-2, fiind disponibil către comunitatea dedicată observării Pământului. Acesta integrează algoritmi de extragere de trăsături din date de teledetecție, algoritmi de clasificare, și metode de vizualizare.

Pentru a accelera și îmbunătăți procesul de extragere de informație din date de teledetecție, am dezvoltat un sistem software ce permite salvarea și reutilizarea automată a parametrilor utilizați în cadrul scenariilor de clasificare.

La finalul capitolului este prezentat un sistem software complex optimizat pentru procesarea și extragerea de informații din date TerraSAR-X. Acesta conține o serie de funcționalități pentru gestionarea bazelor foarte mari de date, vizualizarea datelor, căutarea pe baza conținutului informațional și adnotarea semantică a datelor.

4.1 Metodă propusă pentru analiza automată a seriilor temporale de imagini satelitare

Pentru automatizarea procesului de extragere de informații cantitative privind suprafețele ocupate de clasele de acoperire existente în seriile temporale de imagini satelitare, am dezvoltat o metodă rapidă și exactă bazată pe indici spectrali și ce referințe geografice pentru reducerea diferențelor existente între datele provenite de la mai mulți senzori. Pentru demonstrarea metodei am prezentat evoluția claselor semantice în zona geografică București, România, pe parcursul a 33 de ani [8].

Metoda utilizată pentru extragere automată a claselor semantice de acoperire din STIS este reprezentată în *Figura 4.1*. Primul pas constă în interogarea bazei de date formată din imagini satelitare pentru observarea Pământului. Datele returnate, trec printr-un proces de corecție atmosferică, după care se păstrează doar benzile necesare calculării indicilor spectrali. Datele procesate sunt ulterior secționare în funcție de zona de interes.

În contextul în care nu există date măsurate la nivelul solului, corecțiile atmosferice sunt realizate cu metoda Dark Object Subtraction 1 (DOS1) [9]. Această metodă presupune că, în absența perturbațiilor atmosferice, zonele care de obicei sunt mai întunecate în imaginile de teledetecție (zone acoperite de apă) trebuie să fie cât mai aproape de negru. Pentru aplicarea corecției se identifică valoarea cea mai mică la nivel de fiecare bandă și se scade această valoare din fiecare pixel al benzii.

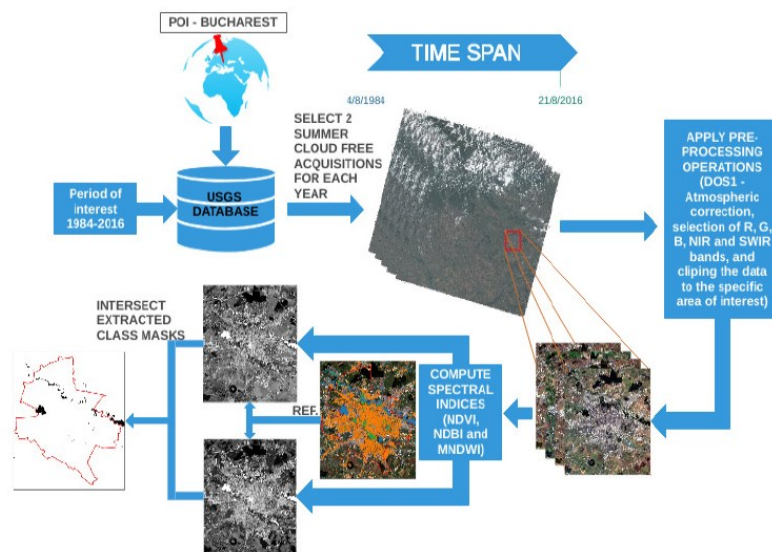


Figura 4.1 Schemă logică pentru extragerea automată a claselor semantice din achiziții de teledetecție.

Deoarece datele provin de la mai mulți senzori, pentru a elimina diferențele dintre valorile obținute pentru indicii spectrali obținuți din acestea, în cadrul metodei se utilizează mai multe zone de referință pentru fiecare clasă semantică.

Ultimul pas al metodei constă în identificarea tuturor pixelilor aferenți măștilor obținute pentru fiecare clasă semantică de acoperire. Aceste valori sunt multiplicare cu suprafața acoperită de fiecare pixel (900 m^2 pentru datele Landsat, respectiv 100 m^2 pentru datele Sentinel-2).

Metoda utilizată pentru analiza seriei temporale s-a dovedit a fi viabilă pentru seturi formate din date provenite de la mai multe misiuni satelitare și senzori, atât timp cât există o unitate de referință pentru clasele semantice monitorizate.

4.2 Sistem software pentru extragerea de informații din date Sentinel-2

În cadrul acestui capitol prezint un instrument software pentru extragerea de informații din date Sentinel-2, sub forma unui plug-in pentru Sentinel Application Platform (SNAP) [10]. SNAP este o platformă software open-source dezvoltată de ESA pentru exploatarea datelor ERS-ENVISAT, Sentinel-1, Sentinel-2, Sentinel-3 și o serie de alte misiuni satelitare.

Sistemul integrează două lanțuri de procesare distincte: analiza exploratorie vizuală a datelor, și reprezentarea semantică a conținutului datelor satelitare. Acestea pot fi folosite atât separat, cât și înlănțuit, procesul de adnotare semantică putând succeda cel de analiză exploratorie, prin care se pot antrena optim algoritmi de clasificare, integrați. În *Figura 4.2* săgețile verzi sunt utilizate pentru a indica fluxul computațional și dependențele, în timp ce săgețile roșii sunt utilizate pentru a evidenția scopul utilizatorului și modul de exploatare al celor două lanțuri de procesare.

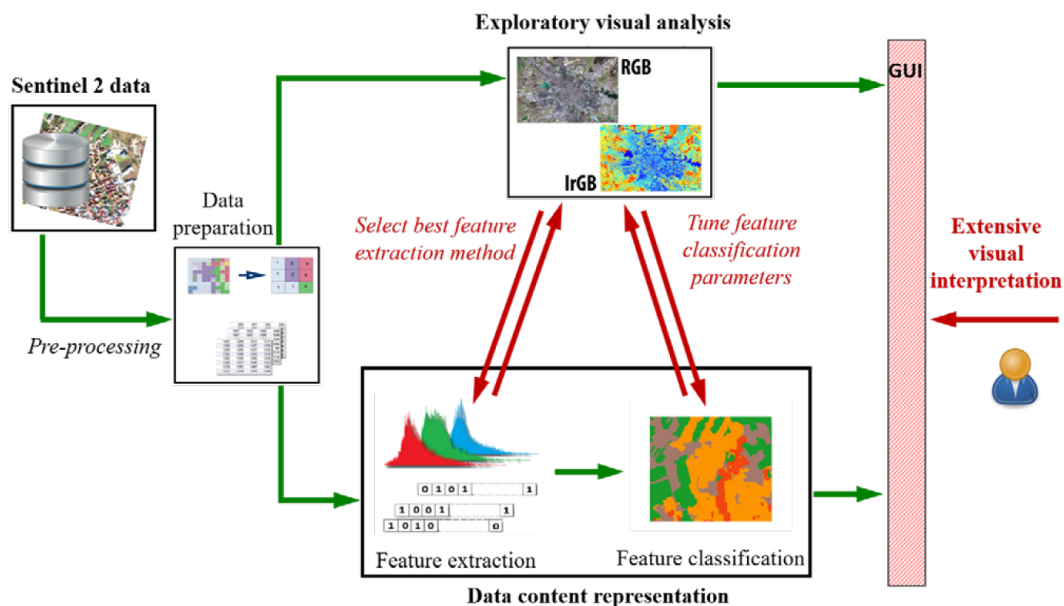


Figura 4.2 Metodologie pentru exploatarea și analiza datelor Sentinel-2.

Sistemul a fost dezvoltat folosind o logică modulară fiind integrabil în platforma SNAP sub forma unui plug-in. Acesta utilizează multitudinea de

capabilități puse la dispoziție de platformă, precum funcționalitățile de vizualizare și exploatare a imaginilor satelitare, managementul straturilor, operații aritmetice cu benzile spectrale, ș.a.. Toate acestea fac din SNAP o platformă flexibilă ce poate deservi mai multe tipuri de date și utilizatori.

Arhitectura plug-in-ului urmărește metodologia prezentată în *Figura 4.2* și este formată din două module distincte: *Exploratory Visual Analysis* și *Semantic Annotation*. Cel din urmă este împărțit în două sub-module: *Feature extraction* și *Feature classification*.

Metodologia prezentată poate acomoda o gamă largă de algoritmi pentru vizualizarea datelor și reprezentarea conținutului, dar din experiențe anterioare [11], am identificat o serie de algoritmi ce au fost integrați în plugin-ul dezvoltat.

Pentru a prezenta capabilitățile sistemului, am ales două scenarii ce utilizează achiziții Sentinel-2. În primul scenariu am realizat o analiză a performanțelor tuturor algoritmilor integrați în plug-in, folosind o scenă pentru care există o hartă de referință a claselor semantice de acoperire, în timp ce în al doilea scenariu, obiectivul a constat în identificarea acoperirii unei clase semantice pentru care nu există hartă de referință.

Sistemul propus [10], facilitează analiza și exploatarea datelor Sentinel-2 prin intermediul unei platforme cunoscute de comunitatea dedicată observării Pământului, SNAP. Sistemul integrează metode pentru căutare de informații cu ajutorul metodelor de augmentare vizuală, respectiv de adnotare semantică bazată pe metode și algoritmi pentru extragere de trăsături și clasificare.

4.3 Sistem software pentru îmbunătățirea rezultatelor de clasificare a datelor de teledetecție

Pentru a putea obține rezultate optime în procesul de clasificare supervizată, s-a demonstrat că utilizatorul trebuie să ofere atât exemple pozitive pentru clasa căutată, cât și exemple negative [12]. Am dezvoltat un sistem software cu scopul de a permite utilizatorului să realizeze clasificări cu acuratețe ridicată pentru seturile de date rezultate din misiunile de teledetecție [13], și să reutilizeze modele de clasificare construite, pe alte seturi de date.

Funcționalitățile sistemului pot fi împărțite în cinci componente reprezentate în *Figura 4.3* sub forma unui flux informațional, de la stânga către dreapta, ce cuprinde

următoarele componente: adnotarea datelor, extragerea trăsăturilor imagine, clasificarea, bucla pentru evaluarea relevanței, și manipularea rezultatelor.

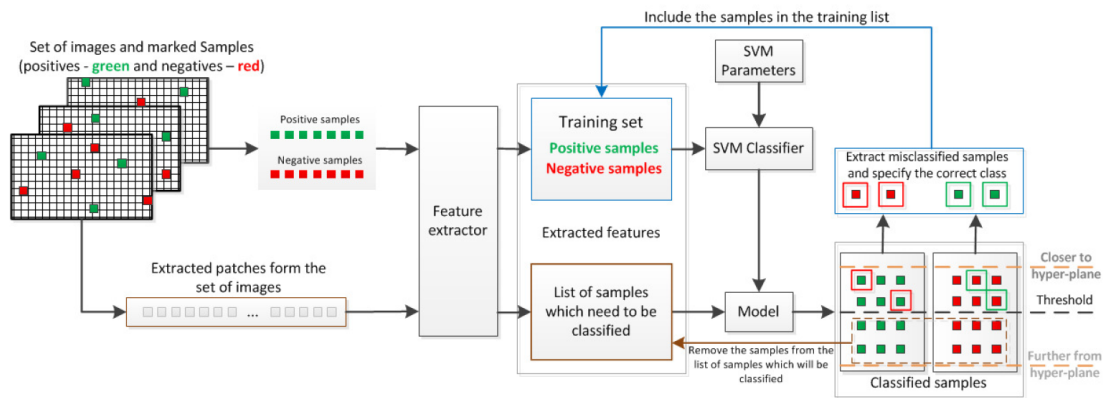


Figura 4.3 Schema funcțională a sistemului folosită pentru îmbunătățirea rezultatelor clasificării.

Sistemul conține un set de componente funcționale, ce deservește roluri bine definite, precum: manipularea grafică a imaginilor, extragerea de trăsături imagine, clasificarea vectorilor de date, optimizarea setului de antrenare, sau salvarea diferitelor componente rezultate în urma clasificării.

În dezvoltarea sistemului am folosit paradigma de programare de tip Model View Controller (MVC) [14] ce separă conținutul software în trei componente mari: model, view și controller. Acestea integrează cinci pachete software utilizate pentru: reprezentarea datelor de teledetecție, procesarea datelor de tip imagine, clasificare, antrenare interactivă și interfață grafică.

Pentru a demonstra funcționalitățile sistemului și compatibilitatea acestuia cu mai multe tipuri de date de teledetecție am construit trei scenarii în care am analizat performanțele ce pot fi obținute cu sistemul dezvoltat. În cadrul primului scenariu sunt comparate două seturi de date formate din achiziții de teledetecție de rezoluție medie, Landsat 8 OLI/TIRS (rezoluție spațială de 30 m) și Sentinel-2 MSI (rezoluție spațială de 10 m). În cadrul scenariilor 2 și 3 am analizat seturi de date provenite din achiziții realizate de misiunile GeoEye și din achiziții de imagini aeriene.

Modul de învățare interactivă bazat pe un număr redus de exemple este eficient atât din punct de vedere al utilizatorului (acesta poate indica ce anume caută într-o scenă) cât și al efortului computațional. Sistemul este capabil, datorită multiplelor tipuri de algoritmi MSV, să realizeze atât clasificări binare (one-versus-all) cât și clasificări multi-clasă (se pot realiza hărți semantice pentru scene complete).

Sistemul este capabil să salveze automat, după fiecare iterație, rezultatele clasificării, modelul MSV, și parametrii scenariului (exemplele folosite pentru antrenare, dimensiunea grătarului, trăsăturile, ș.a.). Ultimele două componente pot fi reîncărcate în sistem pentru a continua procesul de antrenare pentru același set de date sau pe un alt set de date, ce prezintă caracteristici similare.

Pentru a ajuta comunitatea dedicată observării Pământului, codul sursă al aplicației este disponibil pe platforma *github.com* [15].

4.4 Sistem software pentru extragerea de informații din volume foarte mari de date de teledetecție

În cadrul acestui capitol prezint un sistem software, dezvoltat pe o arhitectură modulară, ce oferă capabilități inovative pentru extragerea de informații din date TerraSAR-X [16] și de teledetecție. Sistemul este format din mai multe module ce oferă funcționalități precum: adăugarea datelor, extragerea trăsăturilor din date RAS, extragerea de meta-date, definirea de etichete semantice pentru conținutul arhivelor de imagini, metode avansate de interogare a bazelor de date pe baza conținutului, meta-datelor și categoriilor semantice, precum și diverse funcționalități de vizualizarea a datelor.

Pentru testarea platformei, din arhiva de date TerraSAR-X, am construit un set de date, ce acoperă zone urbane și industriale, din diferite părți ale Globului. Acesta este format din 1100 de scene selectate în funcție de disponibilitate, conținut, diversitatea claselor semantice, și de parametrii de achiziție. În vederea stocării datelor, fiecare achiziție a fost segmentată în secțiuni de 200 x 200 și 160 x 160 pixeli. Vectori de trăsături au fost extrași din fiecare dintre secțiuni și stocați într-o bază de date. Baza de date conține aproximativ 9.24 milioane vectori pentru fiecare tip de descriptor.

Sistemul dezvoltat oferă funcționalități multiple precum: generarea unui model de date, căutarea de informații folosind metode de vizualizare 3D a datelor, interogării diverse și crearea de dicționare de clase semantice. Sistemul este capabil să interpreteze datele TerraSAR-X și meta-datele ce le însoțesc pentru a îmbunătății întregul lanț de procesare ce conține calibrarea datelor, extragerea de trăsături, adnotarea semantică, și crearea de cataloage digitale cu produse.

Figura 4.4 prezintă principalele module ale platformei, împreună cu interfețele existente în cadrul PGS, aplicației utilizatorului, precum și fluxurile de date dintre module. Principalele componente ale platformei sunt următoarele: Data Model Generation, Data-Mining Database, Query Engine, Content-based Image Retrieval, Visual Data Mining, Knowledge Discovery in Databases, Epitome Generation, și System Evaluation.

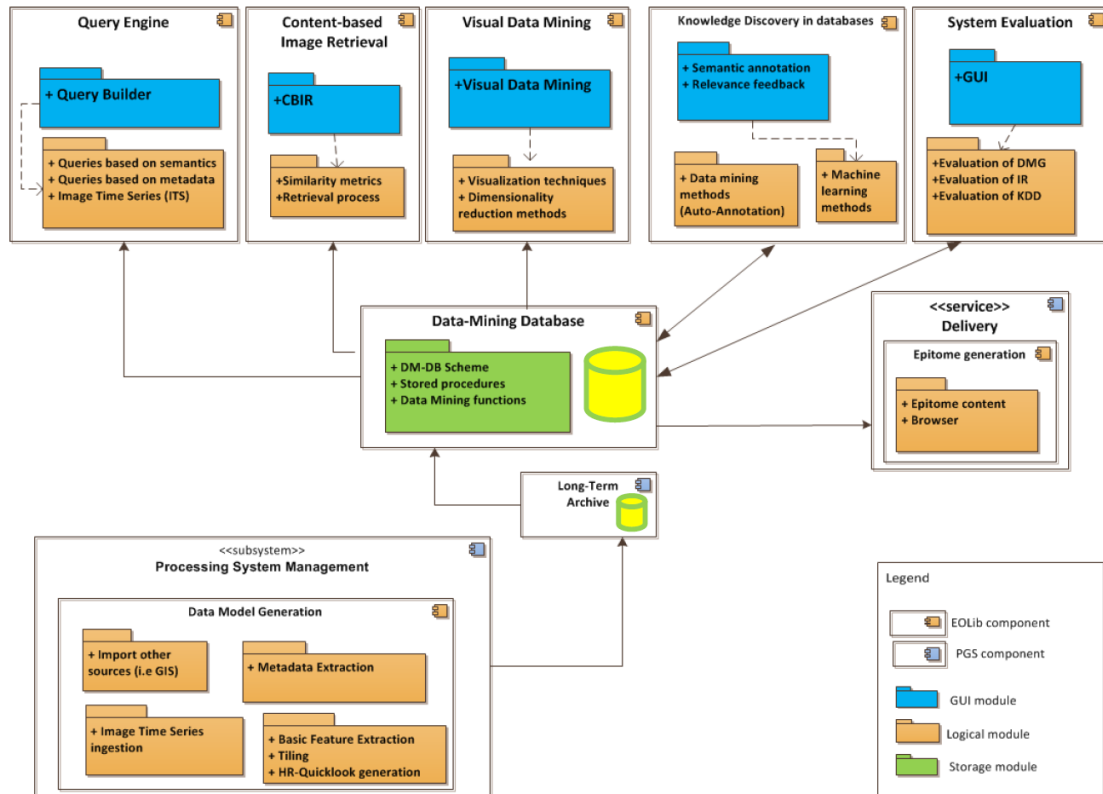


Figura 4.4 Componentele principale ale sistemului, funcționalități generale și dependențe.

După adăugarea datelor în platformă, utilizatorii pot realiza următoarele operații: căutarea de informații, crearea de noi categorii/clase semantice prin explorarea bazei de date, căutarea de imagini, etichetarea și introducerea noilor taxonomii în baza de date, căutarea de imagini sub forma interogărilor pe bază de imagini date drept exemplu, meta-date, clase sau categorii semantice.

4.5 Concluzii de capitol

În cadrul acestui capitol au fost prezentate o serie de metode și aplicații software pentru extragere de informații din achiziții de teledetecție. Toate metodele propuse integrează algoritmi de inteligență artificială și extragere de trăsături ce sunt adaptați

pentru procesarea imaginilor de teledetecție. Toate sistemele propuse au fost validate cu ajutorul scenariilor de utilizare, dovedind rezultate notabile pentru domeniul observațiilor satelitare.

Capitolul 5 Metode și sisteme software dezvoltate pentru identificarea și cuantificarea schimbărilor din serii temporale de imagini satelitare

În cadrul acestui capitol, sunt prezentate o serie de metode și sisteme software propuse ca soluții pentru rezolvarea următoarelor probleme . indexarea eficientă a bazelor de date cu achiziții de teledetecție, generarea seriilor temporale pe baza unor parametri definiți de utilizator, identificarea zonelor în care au loc schimbări semantice, și prezentarea interactivă și selectivă a informațiilor generate din analiza seriilor temporale.

5.1 Metodă de indexare și extracție de informații bazată pe algoritmi de hashing

Pentru optimizarea modului în care se face indexarea datelor de teledetecție am implementat o metodă bazată pe coduri hash [17]. Aceasta facilitează căutarea informațiilor bazată pe conținut în baze de date foarte mari cu imagini satelitare rezultate din observarea Pământului. Scopul principal al implementării constă în reducerea dimensionalității vectorilor de trăsături extrase din imagini, în vederea îmbunătățirii timpilor de extragere de informații din baze de date foarte mari. Logica din spatele sistemului conține două etape prezentate în *Figura 5.1*.

Metoda se bazează pe o combinație de extragere de trăsături, algoritmi de hashing și metode de indexare. Cu ajutorul acestui sistem complexitatea și dimensiunea bazei de date este mult redusă. De asemenea viteza de răspuns pentru un număr de 20 milioane de valori stocate (coduri binare) este de doar 0.4 s.

Pentru a testa performanțele metodei propuse am construit o baza de date formată din cinci imagini Sentinel-2 din diferite părți ale Europei. Setul de date este format din 4 imagini, ce acoperă fiecare o suprafață de 109.8 km x 109.8 km și o imagine cu o suprafață de 40 km x 78 km. Pentru efectuarea testelor, fiecare imagine a fost divizată în secțiuni de 5 x 5 pixeli (sau 50 m x 50 m). În total au rezultat 20.537.664 secțiuni, ale căror coduri hash au fost introduse în baza de date.

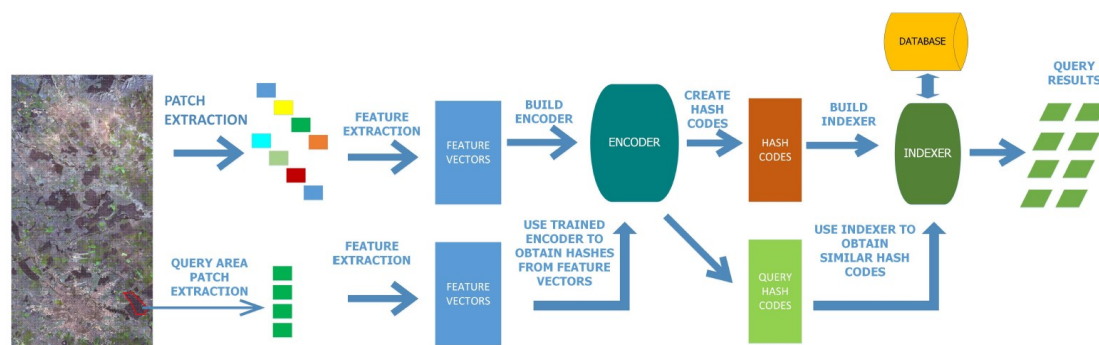


Figura 5.1 Schemă logică reprezentând fluxul informațional de realizare a encoder-ului utilizat pentru indexarea bazei de date (în partea de sus), respectiv, fluxul informațional de căutare în baza de date pe baza unor imagini date drept exemplu.

Sistemul prezintă scalabilitate foarte bună și permit generarea de rezultate într-un timp foarte scurt, în detrimentul timpilor de calcul mai mari determinați de calcularea codurilor hash.

5.2 Sistem software pentru construirea și analiza temporală a seriilor temporale de imagini satelitare

Pentru a facilita manipularea volumelor foarte mari de date de teledetecție, am dezvoltat o platformă software ce oferă utilizatorului posibilitatea să își creeze propriile serii temporale asupra cărora poate să aplice diferite metode de vizualizare pentru a facilita observarea diferitelor evoluții prezente în STIS, sau să descarce datele în vederea unor potențiale procesări ulterioare [18].

Sistemul software a fost dezvoltat folosind limbajul de programare Java și are o structură stratificată și modulară. Funcționalitățile implementate de platformă sunt următoarele: vizualizarea datelor geo-referențiate, indexare în baza de date și stocare non-perisabilă, căutare și extragere a datelor, extragere de trăsături și clasificare. Toate modulele sunt împărțite pe niveluri de abstractizare: stocare, modelare/reprezentare, control/procesare și vizualizare. Arhitectura creată este modulară, flexibilă, permițând dezvoltări ulterioare ale sistemului.

Modulul asociat bazei de date este dezvoltat pe baza sistemului de gestiune MySQL și folosește mecanismul de indexare de tip arbore binar (B-Tree) aplicat asupra tabelor. Indexarea a fost realizată pentru câmpurile satelit, senzor, bandă și dată pentru tabela “images” în vederea facilitării extragerii informațiilor din baza de date cu 830.000 de înregistrări.

Interfețele grafice dezvoltate sunt componente ale nivelului vizualizare și reprezintă principala modalitate de interacțiune a utilizatorilor cu funcționalitățile sistemului. Acestea sunt formate dintr-o serie de ferestre, precum cea principală, prezentată în *Figura 5.2*.

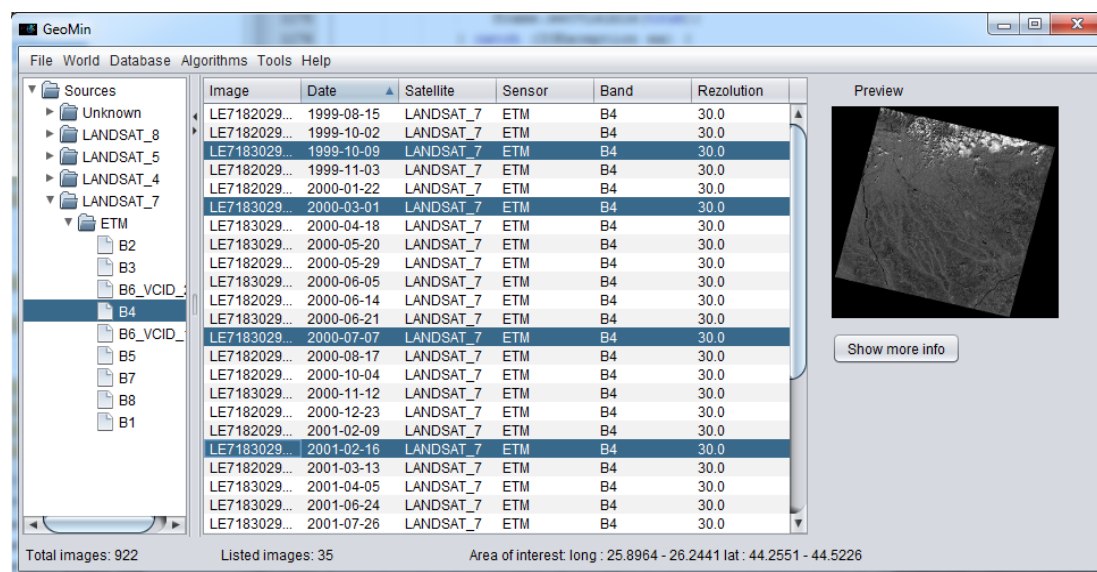


Figura 5.2 Fereastra principală a aplicației.

Interacțiunea cu datele existente în baza de date se poate face direct din fereastra principală prezentată în *Figura 5.2* folosind meniul din partea stângă, în care datele sunt filtrate printr-o structură de tip arborescentă. Aceasta conține următoarele filtre: misiunea satelitară, senzorul utilizat și benzile disponibile.

Principala sursă pentru datele folosite în cadrul aplicației este reprezentată de către platforma USGS. Acestea cumulează 121 de achiziții, realizate de către misiunile satelitare: Landsat 4, 5, 6, și 8. Datele acoperă partea de sud-est a României și au fost realizate în perioada 1989-2015 [19].

Sistemul dezvoltat și-a demonstrat utilitatea în manipularea datelor spațio-temporale și facilitează căutarea de informații în seturile de date. Aceasta permite crearea de STIS în vederea procesărilor ulterioare de tip extragere de trăsături și clasificare.

5.3 Sistem software pentru extragerea de informații din serii temporale de imagini satelitare

În cazul sistemelor bazate pe algoritmi de clasificare supervizați, sunt necesare funcționalități pentru încărcarea și vizualizarea datelor, adnotarea acestora, aplicarea algoritmilor de extragere de trăsături și vizualizarea rezultatelor.

Toate funcționalitățile menționate mai sus au fost integrate în cadrul sistemului software dezvoltat pentru procesarea și extragerea de informații din STIS, prezentat în cadrul acestui capitol.

Pentru validarea sistemului dezvoltat am utilizat o serie temporală formată din 9 achiziții distribuite relativ uniform pe o perioadă de 9 ani, pe care am efectuat o serie de teste în vederea identificării schimbărilor survenite asupra suprafețelor de acoperire.

Pentru a putea observa STIS folosite ca seturi de antrenare, am modificat formatul listei (din partea centrală jos), din orizontal în vertical (vezi *Figura 5.3*). Acest lucru permite o mai bună utilizare a spațiului disponibil și facilitează observarea schimbărilor survenite la nivelul fiecărei serii temporale.

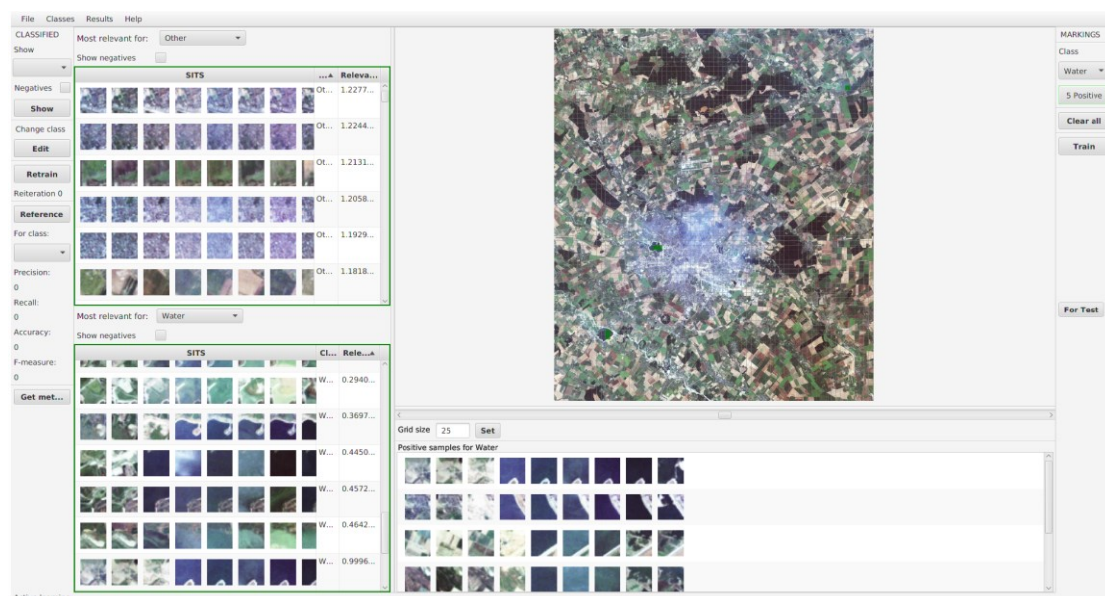


Figura 5.3 Interfața grafică a sistemului capabil să proceseze STIS.

Afișarea rezultatelor, în cele două tabele aflate în partea stângă a fost păstrată în format vertical, dar de a ceastă dată fiecare element din listă nu conține o singură sub-imagine, ci toate sub-imaginile ce compun seriile temporale. În acest fel utilizatorul poate ajusta seturile de antrenare pentru a putea ajuta clasificatorul în a înțelege categoriile semantice de schimbare căutate. Prin creșterea complexității

datelor analizate, crește și probabilitatea erorilor de procedură, iar astfel utilizatorul are control mai bun asupra întregului proces.

Setul de date utilizat în cadrul acestui scenariu este format din trei achiziții Landsat 4 TM și șase achiziții Landsat 5 TM, peste București, România, în perioada 1984-1993. Pentru a analiza schimbările survenite în STIS analizată, am creat o hartă de referință la nivel de pixel pentru trei clase de evoluție. Acestea cuprind *amenajarea de lacuri, remodelarea zonelor urbane și fragmentarea terenurilor agricole*.

În cadrul testelor realizate, am analizat rezultatele obținute prin segmentarea setului de date în secțiuni de 10 x 10 pixeli (rezultând un total de 33490 secțiuni), și 25 x 25 pixeli (rezultând un total de 5372 secțiuni).

În cazul utilizării sistemului pentru analiza STIS, am demonstrat că, acesta, este capabil să obțină rezultate foarte bune, utilizând un număr foarte mic de exemple de antrenare (5-25), în raport cu numărul total de elemente analizate (5372, respectiv 33490 secțiuni). Prin utilizarea mecanismului de învățare asistată, rezultatele s-au îmbunătățit, în medie, cu 5-35%, în funcție de descriptorul utilizat pentru reprezentarea datelor și dimensiunea secțiunilor ce compun STIS.

5.4 Platformă web pentru prezentarea informațiilor extrase din serii temporale de imagini satelitare

Prin monitorizarea zonelor urbane, se pot identifica și descrie procesele prin care au loc schimbările suprafețelor de teren. În această idee am dezvoltat o platformă web pentru monitorizarea zonelor vulnerabile la mișcări ale solului. Aceasta folosește o serie de date de teledetecție satelitară, împreună cu măsurători efectuate la sol pentru a oferi informații legate de diferențele de nivel ce pot apărea în zonele urbane, ca urmare a scufundării pământului, sau ca urmare a mișcărilor seismice.

Sistemul software dezvoltat implementează o arhitectură de web, construită cu ajutorul tehnologiilor tip client-server, Java. Arhitectura este de tip modular, interfața grafică a aplicației fiind dezvoltată cu ajutorul tehnologiei Java Server Faces (JSF) [20], în timp ce partea de back-end a aplicației utilizează o bază de date MySQL controlată cu ajutorul tehnologiei JPA. Datele geo-referențiate (de tip raster și vector) sunt stocate și administrate cu ajutorul GeoServer [21]. Acesta furnizează datele într-un format optimizat pentru aplicații web. Acestea sunt afișate în interfața clientului stratificat cu ajutorul librăriei Javascript, OpenLayers3 [22].

Sursele de informație afișate ca produse pe platformă sunt următoarele: interferometrie PS (sub formă de hărți de persistent scatterers, și profile de deformări punctuale), detecție de schimbări (sub formă de hărți binare de schimbări, hărți multi-clasă de schimbări), și analize geologice (sub formă de hărți geologice, și descrieri textuale a unor observații realizate la nivelul solului).

Avantajul utilizării unei arhitecturi de tip web constă în faptul că utilizatorul nu trebuie să instaleze nimic, permițând un acces rapid și facil pentru orice utilizator, indiferent de gradul de expertiză, locația geografică sau resursele disponibile pe dispozitivul deținut.

Pentru produsele de tip PS Layer, ce afișează informații de tip deformări ale suprafeței terenului se utilizează date RAS. Aceste date provin din misiunile ERS1, ERS2, ASAR și Sentinel-1. Produsele disponibile pe platforma se încadrează în următoarele categorii: *hărți de deformare*, *evoluții de deformare* (vezi Figura 5.4), *hărți geologice*, *zone cu risc geologic*, *schimbări ale claselor de acoperire*, *statistici*, și *alte produse*.

Pe lângă produsele disponibile în interfața grafică de tip hartă, pe platforma web, mai sunt disponibile și alte tipuri de informații ce se regăsesc în pagini secundare. Acestea cuprind pagina *Info* în care sunt descrise diverse terminologii legate de straturile cu risc geologic, sau litologie. De asemenea în cazul în care utilizatorul este autentificat, prin accesarea paginii *Download*, acesta are la dispoziție o serie de documente cu diverse informații suplimentare legate de zonele de interes.

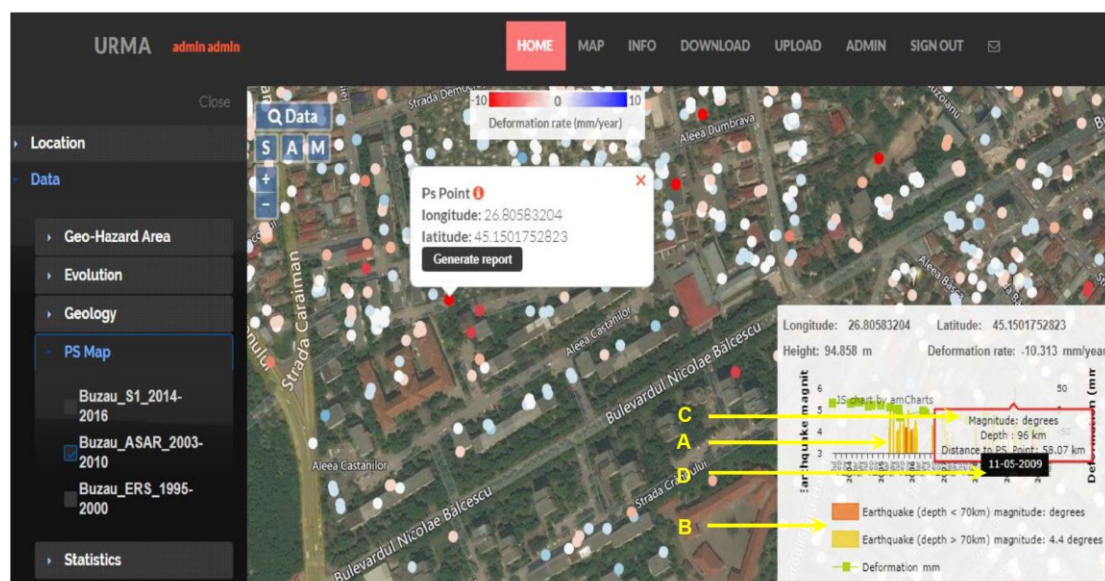


Figura 5.4 Reprezentarea grafică a informațiilor legate de cutremure.

Sistemul software a fost dezvoltat pentru a facilita integrarea mai multor nivele informaționale, ce furnizează următoarele servicii: vizualizarea ratelor de deformare pentru diverse puncte din zonele urbane (componente de clădire), vizualizarea dinamicilor urbane sub forma evoluțiilor claselor de acoperire, identificarea, măsurarea și vizualizarea zonelor cu risc geologic, afișarea unor informații geologice (tip de sol, plăci tectonice, cutremure), și vizualizarea diferitelor statistici pentru diverse domenii (agricultură, demografie, rezidență, ș.a.). Versiunea curentă a platformei poate fi accesată la adresa <http://urma.ceospacetech.pub.ro>.

5.5 Concluzii de capitol

Extragerea de informații utile din seturile de date de teledetecție ridică probleme legate de stocarea și indexarea acestora, căutarea și generarea seriilor temporale, procesarea datelor cu ajutorul algoritmilor de inteligență artificială și prezentarea rezultatelor obținute. În cadrul acestui capitol au fost prezentate o serie de soluții ce tratează una sau mai multe dintre problemele enumerate anterior și întâlnite în cazul procesării seriilor temporale de imagini satelitare.

Capitolul 6 Concluzii

În cadrul acestei teze am încercat să identific potențiale soluții pentru problemele ridicate de procesarea datelor de teledetecție. Cele două direcții urmate s-au axat pe tipurile de informații ce pot fi extrase din datele de teledetecție, cantitative (suprafețe totale la nivel de achiziții individuale) sau evolutive (hărți de schimbare). Prin soluțiile prezentate și demonstrate, s-au tratat toate problemele legate de procesarea datelor de teledetecție.

Pentru a facilita extragerea de informații cantitative, rezultate prin procesarea individuală a fiecărei achiziții dintr-o serie temporală, am identificat seturi de algoritmi de extragere de trăsături și clasificare, ce pot fi aplicați datelor multi-spectrale, respectiv datelor RAS. Aceștia au fost integrați în sisteme software dedicate procesării imaginilor satelitare, iar conceptele propuse sunt puse la dispoziția comunității pentru observarea Pământului.

Cea de-a doua direcție a generat soluții privind manipularea datelor din perspectiva big data. S-a prezentat o metodă optimizată pentru stocarea eficientă a datelor de teledetecție și căutarea rapidă în baze de date cu volum foarte mare de intrări, împreună cu un sistem software pentru generarea rapidă a seriilor temporale cu anumite caracteristici cerute de utilizator. Generarea hărților de evoluție a fost tratată în cadrul unui concept software ce facilitează dezvoltarea de clasificatori optimizați pentru identificarea schimbărilor în date de teledetecție.

Prezentarea informațiilor extrase din seriile temporale de imagini satelitare a fost realizată în cadrul conceptului software de tip client server cu interfață web. Acesta integrează o serie de funcționalități pentru agregarea mai multor moduri de prezentare a informațiilor pentru o mai bună înțelegere a fenomenelor observate de către datele satelitare.

6.1 Contribuții originale

Contribuțiile originale dezvoltate pe parcursul tezei, corelațiile cu articolele în care au fost prezentate (enumerate în capitolul 6.2), și, respectiv, proiectele în care au fost implementate, sunt după cum urmează:

1. O metodă bazată pe indici spectrali ce poate fi utilizată pentru analiza

- seriilor temporale formate din achiziții multi-spectrale provenite de la mai mulți senzori. Soluția propusă a fost aplicată pe o serie temporală de imagini multi-spectrale, ce acoperă o perioadă de 33 de ani [C7].
2. O arhitectură de tip plug-in ce facilitează procesarea datelor de teledetecție de către comunitatea implicată în observarea Pământului. Aceasta integrează mulți algoritmi de extragere de trăsături și clasificare optimizați pentru procesarea datelor Sentinel-2 cu ajutorul platformei SNAP. Mai multe scenarii au fost testate și prezentate în cadrul [J1], [C1], [C3], și [C5].
 3. Un concept pentru optimizarea clasificării volumelor mari de date utilizând un număr mic de exemple, bazat pe o buclă pentru evaluarea relevanței. Acesta integrează și mecanisme de aplicare a clasificatorului antrenat pe un set distinct de date față de cel pe care s-a făcut antrenarea. Soluția propusă a fost validată prin mai multe scenarii prezentate în articolele [C2], și [C12].
 4. O optimizare a conceptului de la punctul 3 pentru analiza seriilor temporale ca o structură individuală de date. Conceptul dezvoltat a fost validat prin articolele [C9] și [C10].
 5. Un concept software pentru extragerea de informații din date TerraSAR-X. Acesta integrează soluții de baze de date și indexări optimizate pentru tipul de date folosit, algoritmi de extragere de trăsături și clasificare, și metode pentru căutări bazate pe conținut [J2], [C6].
 6. O soluție de indexare a volumelor foarte mari de date bazată pe algoritmi de hashing, ce permite căutări bazate pe conținut [C4].
 7. Un sistem pentru manipularea volumelor foarte mari de date de teledetecție provenite de la mai mulți senzori, ce poate fi folosit pentru generarea seriilor temporale, și afișarea modelelor de evoluție prezente în cadrul acestora [C14].
 8. Un concept software pentru prezentarea informațiilor obținute din date de teledetecție, îmbunătățirea perspectivei asupra fenomenelor fizice observate și generarea automată a situațiilor pentru puncte geografice de interes. Acest sistem integrează soluții web pentru afișarea evoluțiilor [C8].
 9. S-au analizat o serie de concepte de tip state-of-the-art pentru a determina potențiale direcții viitoare de cercetare [C11], [C13], [C15].

6.2 Lista lucrărilor originale

Pe parcursul tezei de doctorat conceptele propuse au fost validate de către comunitatea științifică prin 2 articole de jurnal și 15 articole de conferință. De asemenea, majoritatea conceptelor software dezvoltate au fost implementate în 4 proiecte de cercetare, din domeniul observării Pământului, finanțate de ESA (vezi Anexa A.1).

6.2.1 Articole de jurnal

- J1. **A.-C. Grivei**, I. Neagoe, F.A. georgescu, A. Griparis, C. Vaduva, Z. Bartalis, and M. Datcu, *Multispectral data analysis for semantic assessment – A SNAP framework for Sentinel-2 use case scenarios*, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 13, pp. 4429-4442, 2020, DOI: 10.1109/JSTARS.2020.3013091.
- J2. M. Datcu, **A.-C. Grivei**, D. Espinoza-Molina, C. O. Dumitru, C. Reck, V. Manilici, and G. Schwarz. *The Digital Earth Observation Librarian: A Data Mining Approach for Large Satellite Images Archives*, Journal: Big Earth Data (TBED), 2020, DOI: 10.1080/20964471.2020.1738196.

6.2.2 Articole de conferință

- C1. C. Vaduva, F.A. Georgescu, **A.-C. Grivei**, I. Coca (Neagoe), A. Griparis, Z. Bartalis, and M. Datcu, Sentinel-2 data analysis based on explainable features and exploratory visual analysis, ESA EO Φ -WEEK 2020, 28 September – 02 October 2020, Virtual Event.
- C2. **A.-C. Grivei**, *SVMIRE - An Open Source Svm Image Retrieval With Relevance Feedback System For Earth Observation Data Classification*, 2020 13th International Conference on Communications (COMM), Bucharest, 18-20 June 2020, pp. 171-176, DOI: 10.1109/COMM48946.2020.9141975.
- C3. **A.-C. Grivei**, C. Văduva, and M. Datcu, *Assessment Of Burned Area Mapping Methods For Smoke Covered Sentinel-2 Data*, 2020 13th International Conference on Communications (COMM), Bucharest, 18-20 June 2020, pp. 189-192, DOI: 10.1109/COMM48946.2020.9141999.
- C4. **A.-C. Grivei**, C. Văduva, and M. Datcu, *Improved Earth Observation Data Retrieval Through Hashing Algorithms*, IGARSS 2019 - 2019 IEEE

- International Geoscience and Remote Sensing Symposium, Yokohama, 28 July - 2 August 2019, pp. 5909-5912, doi: 10.1109/IGARSS.2019.8898911.
- C5.C. Văduva, F. A. Georgescu, A. Griparis, I. Neagoe, **A.-C. Grivei**, and Mihai Datcu, *Exploratory Search Methodology for Sentinel 2 Data: A Prospect of Both Visual and Latent Characteristics*, IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium, Yokohama, 28 July - 2 August 2019, pp. 10067-10070, DOI: 10.1109/IGARSS.2019.8900349.
- C6.C. O. Dumitru, G. Schwarz, **A. Grivei**, R. Hugues, J. Lorenzo, M. Datcu, *Artificial Intelligence for EO Sensor Fusion*, Living Planet Symposium, Milan, 13-17 May 2019.
- C7.**A.-C. Grivei**, and M. Datcu, *Analysis of Bucharest's Land Cover Evolution Over a Period of 33 Years Using Multi-Sensor Data*, IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, 23-27 July 2018, pp. 1680-1683, DOI: 10.1109/IGARSS.2018.8518615.
- C8.**A.-C. Grivei**, D. D. Taralunga , A. Popescu, V. Poncos, S.-A. Toma, D. Teleaga, A.-M. Vîjdea, A. Baltres and I. Cociuba, *Multi-Layer Geohazards Information System Concept for Urban Areas – URMA Urban Mapping*, EURAS 2018; 12th European Conference on Synthetic Aperture Radar, Aachen, 4-7 June 2018, ISBN: 978-3-8007-4636-1.
- C9.**A.-C. Grivei**, A. Rădoi and M. Datcu, *Land Cover Change Detection in Satellite Image Time Series Using an Active Learning Method*, 2017 9th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp), Bruges, 27-29 June 2017, pp. 1-4, DOI: 10.1109/Multi-Temp.2017.8035213.
- C10. **A.-C. Grivei**, A. Rădoi, C. Văduva and M. Datcu, *Land Cover Change Detection Using a Relevance Feedback System*, WorldCover 2017 Conference, Frascati, 14-16 March 2017.
- C11. **A.-C. Grivei** and M. Datcu, *Data Analytics for Spatio-Temporal Patterns in Satellite Image Time Series: Methods and Architectures*, 2nd BigSkyEarth Workshop – Big Data processing and management concepts for new platforms, Sopron, 23-24 February 2017.
- C12. **A.-C. Grivei**, A. Rădoi, C. Văduva and M. Datcu, *An Active-Learning Approach to Query by Example Retrieval in Remote Sensing Images*, 2016

- International Conference on Communications (COMM), Bucharest, 9-11 June 2016, pp. 377-380, DOI: 10.1109/ICComm.2016.7528339.
- C13. **A.-C. Grivei**, and A.-M. Ghimeș , *Modern Considerations for Data Mining and Big Data*, 2016 International Conference on Communications (COMM), Bucharest, 9-11 June 2016, pp. 385-390, doi: 10.1109/ICComm.2016.7528265.
- C14. **A.-C. Grivei**, and M. Datcu, *HiperMINE – An Earh Observation Spatio-Temporal Data Mining System*, 2015 23rd Telecommunications Forum TELFOR (TELFOR), Belgrade, 24-26 November 2015, pp. 906-909, DOI: 10.1109/TELFOR.2015.7377612.
- C15. C. Văduva, A. Rădoi, **A.-C. Grivei**, G. Schwarz, and M. Datcu, *Urban Mapping using Satellite Time Series*, Mapping Urban Areas from Space – MUAS 2015, Frascati, 4-5 November 2015.

6.3 Perspective de dezvoltare ulterioară

Funcționalitățile instrumentelor software prezentate pot fi extinse cu ajutorul algoritmilor bazați pe rețele neuronale, ce produc rezultate cu o nivel de precizie ridicat.

Integrarea tuturor conceptelor realizate în cadrul tezei, pentru indexarea și afișarea datelor, extragerea de informații din seriile temporale de imagini satelitare și afișarea acestora, într-un sistem complet, bazat pe servicii de tip cloud, la care utilizatorii au acces pe baza unei interfețe web, reprezintă o altă direcție pentru dezvoltări ulterioare.

O altă direcție de dezvoltare ulterioară, prin care se poate crește viteza de procesare a datelor de teledetecție constă în implementarea algoritmilor pe plăci grafice cu număr mare de nuclee de procesare.

Anexe

A.1 Contribuții în proiecte finanțate de ESA

Soluțiile software dezvoltate pe parcursul tezei au fost implementate în o serie de platforme software, de tip desktop și web ce au fost rezultatul a 4 proiecte finanțate de ESA.

P1. Open Source Image Retrieval – Integration of Developed Tools (OSIRIDE)

OSIRIDE a fost câștigat în competiția "AO/1-7966/14/I/CS", fiind finanțat prin European Space Agency, General Support Technology Program (GSTP 6 El. 1), și desfășurat în perioada 2015-2017. Obiectivul proiectului a constat în dezvoltarea, și implementarea de instrumente pentru căutarea de imagini de teledetecție pe baza conceptului CIBC. Algoritmii dezvoltați au fost integrați într-o platformă de tip open-source, prin care s-au extins capacitățile de procesare a volumelor foarte mari de date obținute din observarea Pământului. Arhitectura modulară include o serie de algoritmi și metode pentru căutarea și extragerea de informații latente, vizualizarea datelor, învățarea supervizată a corespondențelor dintre caracteristicile fizice ale datelor și etichetele semantice asociate, prin interacțiunea utilizatorului cu platforma web ce integrează toate aceste funcționalități.

Contribuția personală la acest proiect a constat în dezvoltarea și demonstrarea conceptului de învățare supervizată bazată pe o evaluarea, de către utilizator, a relevanței datelor returnate de sistem și optimizarea algoritmilor de clasificare. De asemenea, am realizat o serie de optimizări pentru reducerea efortului computațional din procesul de extragere de trăsături și clasificare.

P2. Data Mining For Analysis and Exploitation of Next Generation of Time Series (DAMATS)

Proiectul DAMATS a fost câștigat prin competiția "AO/1-7967/14/I/CS", fiind finanțat prin European Space Agency, General Support Technology Program (GSTP 6 El.1), având perioada de desfășurare: 2015-2017. Scopul proiectului a fost de a dezvolta un sistem software ce integrează diferite metode analitice pentru exploatarea informației conținute de STIS. Informația extrasă este prezentată sub forma de categorii de evoluție a claselor de acoperire din datele rezultate din observarea

Pământului. Interacțiunea cu utilizatorul este realizată cu ajutorul unei platforme web ce permite accesul acestuia la bazele de date cu date de teledetecție, și la resurse de calcul pentru utilizarea algoritmilor pentru extragerea claselor de evoluție.

În cadrul acestui proiect am implementat și demonstrat performanțele unor algoritmi state-of-the-art ce au fost integrați în sistem și am realizat un demonstrator pentru o metodă bazată pe MSV și evaluarea relevanței ce poate fi aplicată pentru extragerea de clase semantice de evoluție din serii temporale de imagini satelitare.

P3. The Multi-layer Geohazards Information System Concept for Urban Areas (URMA - urban risk mapping)

Proiectul a fost câștigat printr-o competiție ESA de tip "Call for Outline Proposals under the Romanian Industry Incentive Scheme", fiind desfășurat în perioada: 2016-2018. Acesta a urmărit dezvoltarea unei arhitecturi și metodologii pentru evaluarea și cartografierea zonelor vulnerabile la mișcările solului.

În cadrul acestui proiect am dezvoltat un sistem software de tip platformă web pentru exploatarea interactivă a informațiilor privind dinamica zonelor urbane și care permite realizarea corelațiilor dintre acestea și datele geologice, seismice și PSI disponibile. Platforma pune la dispoziția utilizatorilor funcționalități pentru generarea automată a rapoartelor ce conțin informații privind evoluția unui anumit punct de interes.

P4. Multispectral Data Analysis Toolbox for SNAP – ESA's SentiNel Application Platform (DAS-Tool)

Proiectul a fost câștigat în urma competiției "AO/1-8545/16/NL/Cbi Romanian Industry Incentive Scheme", fiind finanțat prin European Space Agency, și desfășurat în perioada: 2017-2020. Obiectivele proiectului au constat în elaborarea de algoritmi dedicați pentru descrierea și analiza semantică a datelor Sentinel-2. Pentru atingerea obiectivului a fost dezvoltată o arhitectură software de tip plug-in ce se interfațează cu platforma SNAP a ESA.

Contribuțiile mele în acest proiect au constat în implementarea și testarea algoritmilor de extragere de trăsături și clasificare utilizați pentru adnotarea semantică pe regiuni, a datelor Sentinel-2. O altă contribuție este reprezentată de dezvoltarea arhitecturii software și interfețelor grafice pentru extragerea de informației din datele de teledetecție, la nivel de regiune.

Bibliografie

- [1] V. Kuznetsova, V. Sinelnikova și S. Alpertb, „Yakov Alpert: Sputnik-1 and the first satellite ionospheric experiment,” *Advances in Space Research*, vol. 55, nr. 12, pp. 2833-2839, 2015.
- [2] A.-C. Grivei și A.-M. Ghimeș, „Modem Considerations for Data Mining and Big Data,” în *11th International Conference on Communications - COMM2016*, București, 2016.
- [3] Y. Demchenko, P. Grosso, C. de Laat și P. Membrey, „Addressing big data issues in Scientific Data Infrastructure,” în *2013 International Conference on Collaboration Technologies and Systems*, 2013.
- [4] S. Liang, X. Li și W. Jindi, *Advanced Remote Sensing*, Elsevier, 2012.
- [5] D. Espinoza-Molina și M. Datcu, „Architecture Concept for Earth,” în *IGARSS 2013*, 2013.
- [6] M. M. U. Rathore, A. Paul, A. Ahmad, B.-W. Chen, B. Huang și W. Ji, „Real-Time Big Data Analytical Architecture for Remote Sensing Application,” *IEEE Journal of Selected Topics in Applied*, 2015.
- [7] P. Mazzetti și S. Nativi, „Multidisciplinary Interoperability for Earth Observation: Some Architectural Issues,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, nr. 3, pp. 1054-1059, 2012.
- [8] A.-C. Grivei și M. Datcu, „ANALYSIS OF BUCHAREST’S LAND COVER EVOLUTION OVER A PERIOD OF 33 YEARS USING MULTI-SENSOR DATA,” în *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, Valencia, 2018.
- [9] M. Moran, R. Jackson, P. Slater și P. Teillet, „Evaluation of simplified procedures for retrieval of land surface reflectance factors from satellite sensor output,” *Remote Sensing of Environment*, vol. 41, nr. 5, p. 69–184, 1992.
- [10] A. Grivei, I. Neagoe, F. Georgescu, A. Griparis, C. Vaduva, Z. Bartalis și M. Datcu, „Multispectral data analysis for semantic assessment – A SNAP framework for Sentinel 2 use case scenarios,” *J-STARs*, 2020.
- [11] C. Vaduva, F. Georgescu, A. Griparis, I. Neagoe, A. Grivei și M. Datcu, „EXPLORATORY SEARCH METHODOLOGY FOR SENTINEL 2 DATA: A PROSPECT OF BOTH VISUAL AND LATENT CHARACTERISTICS,” în *IGARSS 2019*, Yokohama, 2019.
- [12] M. Babae, S. Tsoukalas, G. Rigoll și M. Datcu, „Visualization-based active learning for the annotation of sar images,” *IEEE Journal of Selected Topics in Applied Earth*

Observations and Remote Sensing, vol. 8, nr. 10, p. 4687–4698, 2015.

- [13] A.-C. Grivei, „SVMIRE - AN OPEN SOURCE SVM IMAGE RETRIEVAL WITH RELEVANCE FEEDBACK SYSTEM FOR EARTH OBSERVATION,” în *The 13th International Conference on Communications*, Bucharest, 2020.
- [14] F. Buschmann, R. Meunier, H. Rohnert, P. Sommerlad și M. Stal, *Pattern-Oriented Software Architecture - Volume 1: A System of Patterns*, Wiley Publishing, 1996.
- [15] A. Grivei, „GitHub,” [Interactiv]. Available: <https://github.com/acgrivei/svmire>. [Accesat 22 07 2020].
- [16] M. Datcu, A.-C. Grivei, D. Espinoza-Molina, C. Dumitru, C. Reck, V. Manilici și G. Schwarz, „The Digital Earth Observation Librarian: A Data Mining Approach for Large Satellite Images Archives,” *Big Earth Data*, pp. 1-30, 2020.
- [17] A.-C. Grivei, C. Văduva și M. Datcu, „Improved Earth Observation Data Retrieval through Hashing Algorithms,” în *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, Yokohama, 2019.
- [18] A.-C. Grivei și M. Datcu, „HyperMINE – An Earth Observation Spatio-Temporal Data Mining System,” în *23rd Telecommunications forum TELFOR 2015*, Serbia, 2015.
- [19] C. Văduva, A. Rădoi, A.-C. Grivei, G. Schwarz și M. Datcu, „Urban Mapping using Satellite Time Series,” în *Mapping Urban Areas from Space Conference*, Frascati, 2015.
- [20] „Java Server Faces,” [Interactiv]. Available: <https://javaee.github.io/javaxserverfaces-spec/>.
- [21] „GeoServer,” 14 06 2020. [Interactiv]. Available: <http://geoserver.org/>.
- [22] „Open Layers,” 14 06 2020. [Interactiv]. Available: <https://openlayers.org/>.
- [23] T. Costăchioiu, R. Constantinescu și M. Datcu, „Multitemporal Satellite Image Time Series analysis of urban development in Bucharest and Ilfov areas,” în *10th International Conference on Communications (COMM)*, Bucharest, 2014.
- [24] C. Dumitru, G. Schwarz, A. Grivei, R. Hugues, J. Lorenzo și M. Datcu, „Artificial Intelligence for EO Sensor Fusion,” în *Living Planet Symposium*, Milan, 2019.
- [25] USGS, „USGS Earth Explorer,” USGS, [Interactiv]. Available: <http://earthexplorer>.