



Universitatea
Transilvania
din Braşov

ŞCOALA DOCTORALĂ INTERDISCIPLINARĂ

Facultatea: Inginerie Electrică și Știința Calculatoarelor

Ing. Anamaria VIZITIU

**Învățarea automată pentru generarea de
imagini medicale, diagnosticare neinvazivă
și asigurarea confidențialității datelor**

**Deep Learning for medical image
generation, non-invasive diagnosis and
privacy preservation**

REZUMAT / ABSTRACT

Conducător științific

Prof.dr.ing. Florin MOLDOVEANU

BRAȘOV, 2020



Universitatea
Transilvania
din Braşov

D-lui (D-nei)

**Componenta
Comisiei de doctorat**

Numită prin ordinul Rectorului Universităţii Transilvania din Braşov

Nr. din

- PREŞEDINTE: - Prof. dr. ing. MORARU Sorin-Aurel
Director de departament
Universitatea Transilvania din Braşov
- CONDUCĂTOR ŞTIINŢIFIC: - Prof. dr. ing. MOLDOVEANU Florin Dumitru
Universitatea Transilvania din Braşov
- REFERENŢI: - Prof. dr. ing. MICLEA Liviu
Universitatea Tehnică din Cluj Napoca
- Prof. dr. ing. POPESCU Dan
Universitatea Politehnică din Bucureşti
- Prof. dr. ing. SUCIU Constantin
Universitatea Transilvania din Braşov

Data, ora şi locul susţinerii publice a tezei de doctorat:, ora, sala

Eventualele aprecieri sau observaţii asupra conţinutului lucrării vor fi transmise electronic, în timp util, pe adresa anamaria.vizitiu@unitbv.ro

Totodată, vă invităm să luaţi parte la şedinţa publică de susţinere a tezei de doctorat.

Vă mulţumim.

Cuprins

	Pg. rezumat	Pg. teză
1 Introducere	3	1
1.1 Către medicina asistată de calculator: prezentare generală, provocări și viitor	3	1
1.2 Obiectivele tezei	4	3
1.3 Structura și conținutul tezei	5	4
2 Deep Learning în medicină	9	7
2.1 Rețele neurale	9	7
2.2 Deep Learning	10	10
2.3 Deep Learning in imagistica medicală	10	13
3 Reconstrucția automată a tomografiei computerizată	11	27
3.1 Introducere	11	27
3.2 Formularea matematică a scanării computer tomografice	12	28
3.3 Reconstrucția imaginilor bazată pe învățare	13	30
3.3.1 Antrenarea modelului Deep Learning de reconstrucție tomografică	13	30
3.3.2 Arhitectura modelului Deep Learning	15	33
3.4 Experimente	15	35
3.5 Rezultate	15	36
3.5.1 Evaluarea calitativă	16	36
3.5.2 Evaluarea cantitativă	17	37
3.6 Discuții generale și concluzii	19	39
4 Detecția automată a leziunilor mamare în mamografiile digitale cu tomosinteză	21	41
4.1 Introducere	21	50
4.2 Detecția leziunilor mamare asistată de calculator	22	52
4.2.1 Formularea problemei	22	54
4.3 Detecția leziunilor mamare cu modele Deep Learning	23	55
4.3.1 Setul de date clinic	23	55
4.3.2 Arhitectura rețelei	23	55
4.3.3 Detaliile de antrenare	24	58
4.3.4 Rezultate	25	60
4.4 Corelarea leziunilor în proiecțiile mamografice	27	62
4.4.1 Depistarea corespondențelor dintre proiecțiile mamografice	28	64
4.5 Rezultate	31	69
4.5.1 Corelarea leziunilor	31	69
4.5.2 Fuziunea proiecțiilor mamografice pentru detectarea leziunilor	32	69
4.6 Discuții generale și concluzii	32	72
5 Protecția datelor în aplicațiile medicale bazate pe învățarea automată	35	77
5.1 Introducere	35	77
5.2 Stadiul actual	35	78
5.2.1 Protecția datelor în inteligența artificială	35	78

5.2.2	Criptarea homomorfică	35	79
5.3	Sistemul criptografic MORE	36	81
5.3.1	Efectuarea operațiilor pe date criptate	36	82
5.4	Aplicarea modelelor Deep Learning pe date criptate homomorfic	38	84
5.4.1	Metoda	38	84
5.5	Partea experimentală	40	86
5.5.1	Formularea problemelor	40	87
5.5.1.1	MNIST: Setul de date tipic pentru rețelele neuronale	40	87
5.5.1.2	Modelul de ordin redus a circulației arteriale	40	88
5.5.1.3	Angiografiile coronariene cu raze X	41	90
5.5.2	Pregătirea bazei de date criptate	42	93
5.5.3	Arhitectura modelelor Deep Learning	43	93
5.5.3.1	Rețeaua neurală pentru clasificarea imaginilor ce conțin cifre scrise de mână	43	94
5.5.3.2	Rețeaua neurală pentru analiza hemodinamică în timp real	43	94
5.5.3.3	Rețeaua neurală pentru analiza angiografiilor coronariene cu raze X	44	95
5.6	Rezultate	45	97
5.6.1	Performanța	45	97
5.6.1.1	Clasificarea imaginilor MNIST	45	98
5.6.1.2	Analiza hemodinamică	45	100
5.6.1.3	Clasificarea angiografiilor coronariene	45	100
5.6.2	Timpul de rulare	47	102
5.7	Discuții și concluzii	47	106
6	Concluzii finale	51	109
6.1	Concluzii	51	109
6.2	Contribuții originale	52	110
6.2.1	Reconstrucția imaginilor medicale bazată pe învățarea automată	52	110
6.2.2	Diagnosticarea medicală bazată pe învățarea automată	52	110
6.2.3	Asigurarea confidențialității datelor în procesul de învățare automată	53	110
6.3	Diseminarea rezultatelor	54	112
	Bibliografie	56	114
	Abstract	60	

1. Introducere

1.1 Către medicina asistată de calculator: prezentare generală, provocări și viitor

Imagistica medicală este o procedură medicală standard, neinvazivă, utilizată pe scară largă în procesul de examinare a corpului uman. Tehnicile imagistice, cum ar fi radiografia, ecografia, tomografia computerizată, permit obținerea de imagini medicale detaliate ale structurilor interne ale corpului uman în scopuri clinice, cum ar fi prevenirea bolilor, diagnosticarea, planificarea tratamentului și monitorizarea.

În prezent, imagistica medicală a devenit o parte din ce în ce mai importantă a sectorului medical deoarece facilitează detecția timpurie, îmbunătățind sănătatea și bunăstarea pacienților. Precizia diagnosticului depinde în mare măsură de calitatea imaginii achiziționate dar, mai important, de interpretarea imaginilor medicale de către radiologi. Procesul de analiză și interpretare a imaginilor medicale este unul extrem de complex, fiind frecvent predispus la erori umane, ca urmare a experienței practice, a subiectivității și a nivelului de oboseală acumulat. La acestea se adaugă variațiile mari care există între pacienți, complexitatea imaginilor și volumul mare de muncă.

În ultimii ani, datorită progreselor tehnologice, un efort considerabil a fost investit în furnizarea de noi soluții și instrumente pentru automatizarea proceselor excesiv de complexe sau de rutină care necesită constant intervenția umană. Deși în practică, interpretarea imagisticii medicale se bazează în mare măsură pe experiența medicilor, în timp, odată cu experiența și anii de practică, acest proces devine automatizat într-o anumită măsură. Astfel, analiza imaginilor medicale poate fi foarte bine formulată ca o soluție software ce oferă instrumentele necesare diagnosticării asistate de calculator. Combinarea puterii de calcul cu experiența radiologilor, cu scopul de a facilita crearea de soluții software în medicină, este un pas crucial în îmbunătățirea procesului decizional.

Astfel de soluții pot aduce beneficii substanțiale, reducând costurile și îmbunătățind calitatea serviciilor. O mare parte din aceste soluții se bazează pe inteligența artificială. Mai exact, pe abilitatea acestor instrumente de a procesa cantități masive de date și de a oferi predicții și estimări prin recunoașterea diverselor caracteristici.

Ca urmare a rezultatelor promițătoare obținute în automatizarea unor probleme complexe din domeniul procesării de imagini și computer vision, inteligența artificială a fost integrată în multe domenii ale industriei, printre care și cel medical. Cu toate că inteligența artificială presupune o componentă cognitivă, tehnologiile actuale sunt departe de a atinge inteligența ideală. În cel mai optimist caz, ele oferă mijloacele algoritmice de a detecta și recunoaște anumite modele sau regularități în seturi mari de date. Machine Learning (ML) este o ramură a inteligenței artificiale care se bazează pe extragerea de cunoștințe și statistici pentru a permite calculatoarelor să învețe din cantități mari și diverse de date fără a fi programate în mod explicit [1].

Abilitatea de a extrage informații semnificative din seturi mari de date face ca soluțiile ML să fie potrivite pentru rezolvarea unor probleme complexe de analiză a datelor, inclusiv analiza medicală bazată pe imagini. Cu toate acestea, crearea sistemelor inteligente de asistență medicală vine cu multe provocări și obstacole. Unul dintre cele mai mari obstacole în calea integrării soluțiilor ML în industria medicală este dată de natura datelor. Nu numai că manipularea datelor cu caracter personal ridică anumite dificultăți, ci și faptul că, odată colectate, acestea trebuie să fie adnotate manual de

către experți medicali.

În pofida faptului că este unul dintre sectoarele cu cea mai rapidă creștere din economia mondială, piața fiind estimată să ajungă la 6.6 miliarde de \$ până în 2021 [2], industria biomedicală adoptă lent soluțiile și practicile bazate pe inteligență artificială. Decalajul se produce într-o zonă care produce o colecție extrem de abundentă de date valoroase. În timp ce datele biomedicale sunt abundente, ele sunt greu de circulat și manipulat din cauza constrângerilor etice și juridice. Prin reglementările adoptate în prezent, cum ar fi regulamentul general privind protecția datelor (GDPR), au fost ridicate mai multe preocupări cu privire la protecția, securitatea și partajarea datelor cu caracter personal. Astfel de reglementări restricționează utilizarea sau divulgarea datelor personale, lucru care afectează în mod clar aplicarea și dezvoltarea algoritmilor bazați pe învățarea automată din date. Având în vedere că accesul la datele pacienților este necesar în aplicațiile medicale inteligente, în prezent un efort considerabil este alocat în găsirea unor soluții care să trateze problemele legate de securitatea și protecția datelor într-o manieră care să permită accesul la datele medicale, asigurând în același timp progresul și utilitatea tehnicilor de ML în sectorul medical.

În consecință, în ultimii ani, s-a depus un efort continuu pentru a oferi soluții bazate pe inteligență artificială pentru prelucrarea rapidă, precisă și sigură a datelor medicale pentru screening, prevenire, diagnostic și tratamente personalizate. Cu toate acestea, anumite provocări rămân a fi rezolvate pentru a putea integra pe deplin aceste soluții în practica medicală.

1.2 Obiectivele tezei

Fluxul de lucru clinic tradițional al interpretării imagisticii medicale constă din trei etape: (i) achiziția de date, (ii) reconstrucția structurii interne a corpului uman pe baza datelor obținute și (iii) interpretarea clinică a imaginii reconstruite. Cu toate acestea, în combinație cu soluțiile de detecție asistată de calculator, este necesară introducerea unei etape suplimentare, responsabilă pentru confidențialitatea, integritatea și securitatea datelor pacientului, după cum se arată în Figura 1.1.

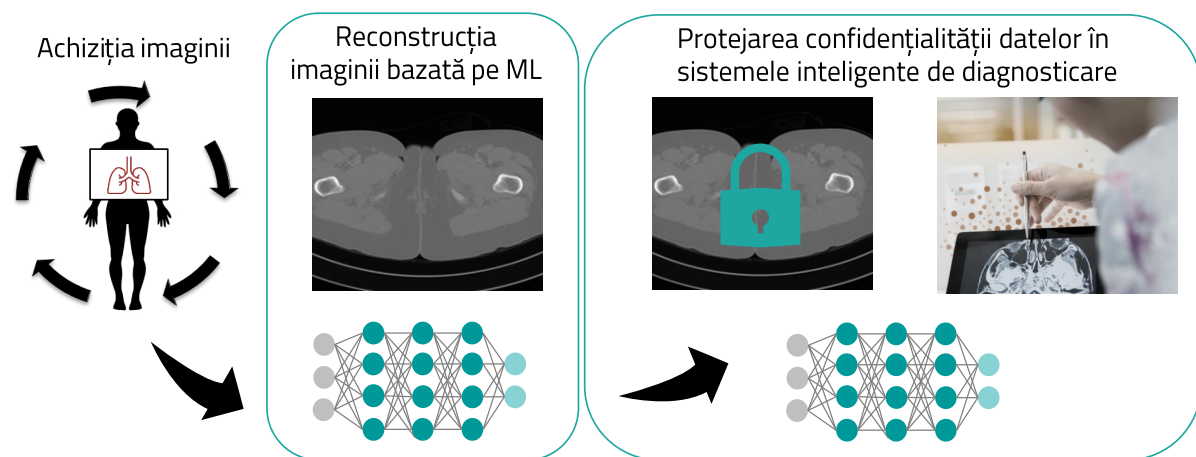


Figura 1.1: Procesul de interpretare al imaginilor medicale bazat pe inteligență artificială în cadrul unui flux clinic de rutină, de la achiziția de date până la diagnostic.

Având în vedere că radiologia joacă un rol major în diagnosticarea și tratarea bolilor și știind că timpul de analiză imagistică este relativ lung iar interpretarea este susceptibilă de a fi afectată de erori umane, accentul se pune pe investigarea beneficiilor pe care inteligența artificială le poate aduce domeniului medical, cu accent pe imagistică.

Prin urmare, teza actuală contribuie la dezvoltarea și aplicarea rețelelor neurale în fluxul de lucru al interpretării imagisticii medicale. Mai precis, se introduc soluții inovative bazate pe modele Deep Learning în domeniile reconstrucției și analizei imaginii dar și în asigurarea confidențialității datelor medicale. Cercetarea actuală a avut ca scop studierea, implementarea, testarea și validarea utilizării

modelelor Deep Learning în vederea furnizării unei analize complete a datelor medicale bazate pe învățare.

Astfel, pot fi identificate următoarele obiective principale de cercetare:

- Dezvoltarea, implementarea, și testarea unui algoritm bazat pe Deep Learning pentru reconstrucția rapidă a tomografiilor computerizate de înaltă rezoluție pornind de la date achiziționate cu o doză mult redusă de radiații;
- Dezvoltarea unui instrument de evaluare realist din punct de vedere clinic pentru stabilirea calității reconstrucțiilor;
- Dezvoltarea, implementarea, și testarea unui sistem de diagnosticare asistată de calculator bazat pe Deep Learning care evidențiază zonele suspecte sau cu risc în mamografii;
- Dezvoltarea unei metodologii de validare și evaluare a performanțelor sistemului de diagnosticare asistată de calculator;
- Extinderea metodologiei existente pentru analiza imagisticii pe baza proiecțiilor mamografice;
- Dezvoltarea unui sistem securizat de criptare homomorfică;
- Dezvoltarea unei biblioteci generice de Deep Learning care înglobează criptarea homomorfică pentru a permite realizarea calculului din rețeaua neurală direct pe date criptate fără a dezvălui informațiile cu caracter personal ale pacienților;
- Dezvoltarea unei metodologii de testare și evaluare a bibliotecii de Deep Learning pe date criptate homomorfic;
- Dezvoltarea, implementarea, și evaluarea unui sistem securizat de analiză a angiografiilor coronariene criptate homomorfic;
- Dezvoltarea, implementarea, și evaluarea unui sistem securizat de personalizare a simulărilor hemodinamice ale circulației sângelui;
- Identificarea și examinarea atât al nivelului de securitate, cât și al utilității sistemului criptografic;
- Întărirea securității sistemului de criptare, menținând în același timp performanța și utilitatea în aplicațiile reale;
- Dezvoltarea și implementarea unui sistem securizat pentru estimarea riscului pentru un eveniment cardiovascular.

1.3 Structura și conținutul tezei

Teza de doctorat conține 6 capitole și este organizată după cum urmează:

Capitolul întâi, Introducere (eng. Introduction), conține o prezentare generală a domeniului de cercetare, definindu-se contextul și justificându-se relevanța științifică. Sunt evidențiate problemele identificate spre a fi soluționate în vederea îmbunătățirii procesului de diagnosticare medicală, împreună cu activitățile desfășurate și metodologiile utilizate pe parcursul cercetărilor. De asemenea, sunt definite principalele obiective ale tezei de doctorat. În finalul capitolului se prezintă succint structura și conținutul tezei.

Capitolul al doilea, Învățarea automată în medicină (eng. Deep Learning in Medicine), se axează pe o prezentare generală dar și aplicată a domeniului învățării automate. Pentru început, sunt introduse și explicate noțiunile fundamentale ale algoritmilor de învățare automată bazați pe rețele neurale artificiale, cu accent pe procesul de învățare și evoluție spre rețele adânci (eng. Deep Learning). În

continuare, sunt trecute în revistă tendințele actuale al cercetărilor științifice în domeniul proiectării arhitecturii rețelelor, fiind vizate și formulate principalele sarcini care pot fi abordate în domeniul medical cu ajutorul modelelor bazate pe învățare. Astfel, sunt discutate și prezentate în detaliu cele patru probleme clasice din sfera inteligenței artificiale (clasificarea, segmentarea, localizarea și generarea de imagini), precum și tehnicile și metricile adecvate de evaluare a performanțelor.

În **capitolul al treilea**, Reconstrucția automată a tomografiei computerizate (eng. Towards Data-Driven CT Imaging Reconstruction), se introduce un model bazat pe rețele neurale pentru reconstrucția imaginilor medicale. Pentru început, se analizează nivelul actual al cercetărilor științifice în domeniul reconstrucției imaginilor medicale, fiind vizate cu precădere tehnicile de reconstrucție bazate pe inteligență artificială. În continuare, se prezintă în detaliu soluția inovativă propusă. Algoritmul propus combină tehnicile de învățare automată cu noțiuni de bază din fizică și geometrie ce permit descrierea sistemului de achiziție al tomografiei computerizate și care este îmbunătățit prin utilizarea unor tehnici avansate de antrenare ce au rolul să integreze percepția umană în contextul generării de imagini medicale realiste. Sunt descrise modalitățile de evaluare luate în considerare și utilizate pentru a se demonstra și valida calitatea imaginilor obținute prin intermediul soluției propuse. În continuarea capitolului se prezintă rezultatele reconstrucțiilor efectuate pentru o serie de imagini provenite de la pacienți. Aceste rezultate sunt validate prin intermediul unor indicatori de performanță, precum și prin realizarea unui studiu orb randomizat cu radiologi experimentați. Capitolul se încheie cu un studiu comparativ, menit să compare performanțele algoritmilor clasici din literatura de specialitate cu metoda propusă, demonstrându-se faptul că soluția dezvoltată conduce la obținerea unor performanțe superioare.

Capitolul al patrulea, Detecția automată a leziunilor mamare în mamografiile digitale cu tomosinteză (eng. Towards Computer-aided Detection System in Digital Breast Tomosynthesis), descrie metodologia utilizată pentru dezvoltarea unui sistem bazat pe rețele neurale pentru detecția leziunilor la nivelul sânului din mamografiile digitale tridimensionale. Prima parte a capitolului este dedicată prezentării contextului medical, introducerii tehnicii de imagistică medicală și a setului de date considerat pentru realizarea cercetărilor, precum și a noțiunilor medicale necesare în analiza imaginilor. După o scurtă introducere în domeniu și în literatura de specialitate, se descrie în detaliu problema tratată, și metodologiile propuse pentru dezvoltarea unui sistem complet autonom de analiză a imaginilor medicale. În continuare, se prezintă sistemul de detecție propus și metodologia de validare și evaluare a performanțelor. Capitolul continuă cu analiza performanțelor și este evidențiat un prim set de concluzii. Rezultatele obținute sunt promițătoare, demonstrând capacitățile modelelor neurale de a fi integrate în procesul de diagnosticare în cadrul unui flux clinic de rutină. În pofida rezultatelor promițătoare obținute, s-a identificat posibilitatea îmbunătățirii acurateței sistemului de detecție, propunându-se în acest sens un sistem optimizat bazat pe o metodă inovativă de corelare geometrică a informațiilor din cele două imagini standard achiziționate pentru fiecare sân. Astfel, detecțiile obținute de modelul neural pentru fiecare imagine sunt ulterior filtrate prin metoda de corelație geometrică. Următoarea etapă a cercetării este reprezentată de evaluarea rezultatelor obținute prin combinarea algoritmului de detecție cu cel responsabil de calculul corespondențelor dintre leziuni (eng. registration). În urma unei noi evaluări s-a dovedit că această strategie este viabilă, obținându-se o reducere a numărului de detecții false în raport cu metoda inițială.

Capitolul al cincilea, Protecția datelor în aplicațiile medicale bazate pe învățarea automată (eng. Towards Privacy-Preserving Deep Learning-based Medical Applications), descrie metodologia utilizată pentru dezvoltarea de soluții medicale personalizate prin protejarea integrității datelor pacienților. Soluția propusă folosește o combinație inovativă de algoritmi de procesare a datelor bazați pe învățare automată și algoritmi de criptare pentru a permite manipularea datelor cu caracter personal. Mai exact, analiza datelor bazată pe rețele neurale adânci realizând-se direct pe date criptate, fără a fi necesar accesul la datele originale. Astfel, în prima parte a capitolului se prezintă principalele tehnologii utilizate precum și lucrările de referință din literatura de specialitate în care se prezintă stadiul actual în securitatea datelor în inteligența artificială. În continuare, se introduce sistemul de criptare considerat pentru realizarea cercetărilor (schema de criptare homomorfică MORE) și soluția propusă

pentru a se utiliza rețelele neurale pe date criptate. Pentru a se valida algoritmul propus se consideră două probleme de actualitate din sfera medicală: analiza automată a imaginilor angiografice achiziționate cu raze X criptate și estimarea automată a parametrilor pentru personalizarea modelului hemodinamicii sistemului cardiovascular. Suplimentar, o a treia problemă tipică din sfera inteligenței artificiale, detecția automată a cifrelor din imagini, este considerată ca metodă de analiză comparativă. Astfel, în a doua parte a capitolului se descriu în detaliu cele trei probleme tratate, seturile de date, precum și modelele neurale considerate pentru realizarea cercetărilor. În final, sunt descrise modalitățile de evaluare dezvoltate și utilizate pentru a analiza abilitățile sistemului propus. În continuare, se prezintă rezultatele studiului de fezabilitate, menit să compare performanțele modelelor care operează direct pe date criptate cu cele convenționale în care procesarea se face direct pe datele necriptate. Experimentele au arătat că utilizarea criptării homomorfe ca mecanism de protecție a integrității datelor permite aplicarea tehnicilor de inteligență artificială pe date criptate fără a compromite precizia. În urma analizei securității sistemului criptografic s-au identificat o serie de limitări și vulnerabilități și s-a propus o nouă metodă de criptare (Hybird MORE). În urma analizei și evaluării noului sistem criptografic pe o problemă de estimare a riscul unui atac de cord, s-a dovedit că această nouă metodă este promițătoare, obținându-se o îmbunătățire a securității însă cu un cost de precizie mai mic.

Teza de doctorat se încheie cu **capitolul al șaselea**, Concluzii finale (eng. Final Conclusions), în care se prezintă concluziile finale evidențiate în urma activității de cercetare întreprinse. Suplimentar, sunt enumerate contribuțiile personale din cadrul tezei de doctorat, precum și lucrările științifice publicate de-a lungul programului de doctorat. Capitolul se încheie cu prezentarea unor recomandări și enunțarea posibilelor direcții viitoare de cercetare.

Porțiuni din această teză au fost publicate anterior în [3–6].

2. Deep Learning în medicină

De la prima lor apariție în 1943 [7], funcționalitatea rețelelor neurale a fost asociată continuu cu modul în care oamenii învață și prelucrează informația. Mai precis, ele au fost concepute pentru a imita conexiunile sinaptice dintre neuronii creierului, iar mai târziu, au devenit baza modelelor Deep Learning.

2.1 Rețele neurale

La nivel înalt, o rețea neurală poate fi definită ca un model matematic care conectează o serie de intrări la ieșiri printr-o compoziție de straturi cu blocuri de procesare interconectate (transformări și funcții de activare). Arhitectura unei rețele neurale clasice este descrisă în Figura 2.1. Pentru a permite o formulare matematică mai complexă, la fiecare bloc de prelucrare se adaugă, de obicei, funcții de activare neliniare. Acestea filtrează informațiile care trec prin rețea, determinând semnalul de intrare relevant pentru a fi transmis la nivelul următor. Astfel, se decide dacă un anumit neuron ar trebui activat sau nu, iar fără acestea, rețeaua neurală devine un model liniar simplu.

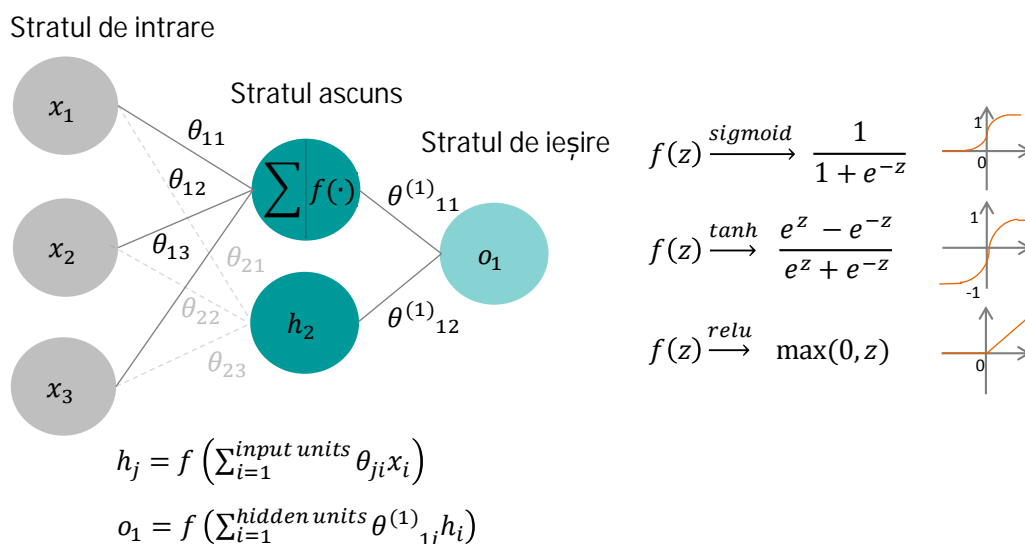


Figura 2.1: Arhitectura unei rețele neurale simple descrisă prin trei straturi de neuroni: intrare, ieșire și unul ascuns. Informația parcurge toate straturile, începând cu stratul de intrare și până la cel de ieșire. În această configurație, fiecare neuron primește informații de la toți neuronii din stratul anterior (în literatura de specialitate acest strat poartă denumirea de strat complet conectat). Conexiunile θ dintre neuroni sunt parametrii care trebuie ajustați în funcție de date și de problema formulată. Fiecare bloc de procesare efectuează o transformare (aici o sumă ponderată a parametrilor de intrare), iar rezultatul este transmis către o funcție de activare f care va fi utilizată pentru a adăuga proprietăți neliniare în rețea. Funcțiile de activare sunt selectate dintr-un set de funcții limitate cu anumite proprietăți matematice.

Ca orice altă metodă de ML, modelele bazate pe rețele neurale au ca scop să învețe din experiența anterioară pentru a face predicții pe baza unor observații noi. În procesul de învățare supervizată, modelul învață automat parametrii modelului, și deci funcția de mapare a spațiului datelor de intrare de cel de ieșire, pe baza exemplurilor de antrenare adnotate, într-un mod iterativ, prin efectuarea treptată a unor ajustări cu intenția de a minimiza eroarea dintre valorile prezise și cele dorite.

După antrenare, rețeaua este capabilă să furnizeze rezultate care sunt din punct de vedere statistic similare cu cele preconizate, chiar și atunci când datele de intrare sunt unele pe care rețeaua nu le-a întâlnit niciodată în timpul antrenării. În consecință, rețelele neurale pot fi utilizate pentru prezicerea unor ieșiri pe baza anumitor caracteristici de intrare, clasificarea datelor și chiar localizarea unor obiecte în imagini.

2.2 Deep Learning

În esență, o rețea Deep Learning nu reprezintă altceva decât o rețea neurală mai complexă, compusă din mai multe straturi de neuroni și organizată ca un strat de intrare, urmat de multiple straturi ascunse și un strat de ieșire. De-a lungul anilor s-a demonstrat că o astfel de arhitectură facilitează modelarea funcțiilor extrem de complexe, permițând învățarea unor reprezentări intermediare mai bogate. Prin urmare, diferența esențială dintre rețelele convenționale și cele de adâncime este dată de adâncimea modelelor, deși nestandardizată, de obicei o rețea cu o adâncime mai mare de două straturi, se încadrează în categoria rețelelor adânci.

2.3 Deep Learning in imagistica medicală

Tendința actuală în domeniul medicinei este de a stabili un diagnostic personalizat, de a planifica tratamentul, precum și de a preveni bolile prin identificarea și corelarea unor cantități masive de date provenite de la pacienți (de exemplu simptome, diagnostic, tratamente etc.). În analiza imagisticii medicale, modelele Deep Learning au fost introduse ca o modalitate de a înțelege într-o manieră automatizată conținutul semantic al imaginilor pacientului pentru diagnosticarea, detectarea și segmentarea structurilor anatomice sau al patologiilor. În consecință, în ultimii ani, Deep Learning a oferit multe soluții concepute pentru a aborda diferite sarcini de imagistică (Figura 2.2), inclusiv segmentarea, detectarea obiectelor, clasificare și, mai recent, generarea de imagini (de exemplu, îmbunătățire, densificare, reconstrucție, etc.).

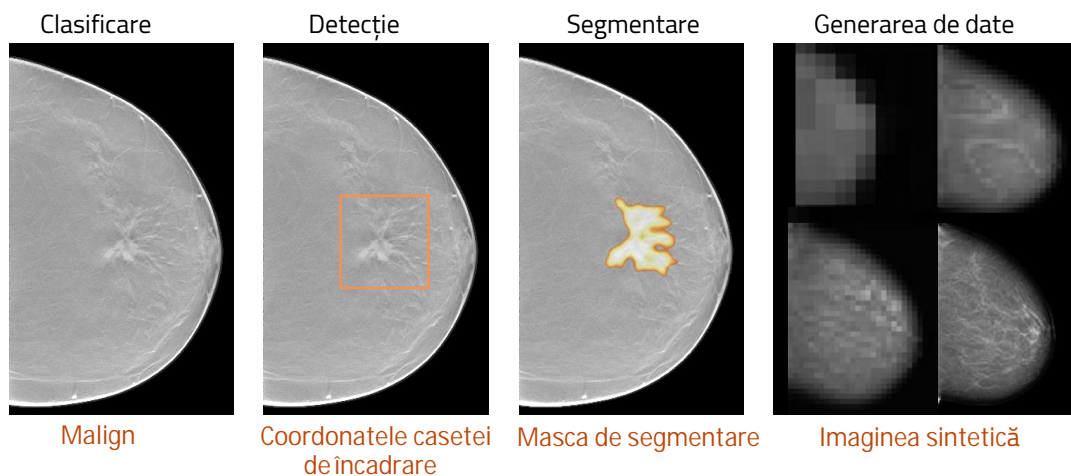


Figura 2.2: O prezentare generală a problemelor tratate în imagistica medicală prin intermediul modelelor Deep Learning.

3. Reconstrucția automată a tomografiei computerizată

3.1 Introducere

Tomografia computerizată (CT) este o tehnică de investigare imagistică des utilizată în rutinele clinice care, prin folosirea razelor X, permite studiul organelor și țesuturilor corpului uman în scopuri diagnostice în numeroase specialități medicale. Mai exact, prin utilizarea fascicolului de raze X se obține imaginea unei secțiuni transversale a obiectului observat. Această metodă de investigare implică măsurarea gradului de atenuare a radiației electromagnetice care traversează obiectul și ulterior reconstrucția imaginii obiectului investigat folosind proiecțiile obținute din secțiunile sale transversale.

Există o dependență clară între calitatea reconstrucției tomografice și precizia de diagnosticare [8]. Pentru a realiza o reconstrucție tomografică de calitate este necesară utilizarea unei doze crescute de radiații, ceea ce sporește riscul pacientului de a dezvolta patologii legate de radiații. O reducere a dozei de iradiere conduce însă la o deteriorare a calității imaginii, introducând artefacte și zgomot, și astfel compromițând precizia de diagnosticare.

Mai recent, un interes sporit a fost acordat integrării tehnicilor de inteligență artificială în procesul de reconstrucție pentru a crea imagini de calitate din proiecții achiziționate cu doze reduse de radiații. Aceste tehnici au avantajul de a crea imagini cu rezoluție îmbunătățită, cu un nivel de zgomot redus, permițând reducerea substanțială a expunerii la radiații ionizante în timpul unei scanări.

Considerând procedeul de achiziție al tomografiilor computerizate, soluțiile de reconstrucție bazate pe modele Deep Learning se împart astfel în două mari categorii: (i) reconstrucția bazată pe post-procesarea în spațiul imaginii și (ii) reconstrucția optimizată în spațiul proiecției. Metodele din prima categorie formulează reconstrucția mai mult ca o problemă de post-procesare responsabilă cu reducerea zgomotului din imagini. Aceasta este una dintre cele mai simple metode de a introduce inteligența artificială în procesul de reconstrucție tomografică. În acest context, reconstrucția constă din două etape: (i) reconstrucția propriu-zisă obținută prin aplicarea unui operator simplu, cum ar fi transformarea inversă a Radonului, asupra proiecțiilor achiziționate cu o doză redusă de radiații și (ii) post-procesarea imaginilor obținute de către o rețea neurală pentru a reduce artefactele și zgomotul în vederea obținerii unei reprezentări asociate cu o doză ridicată de radiații.

În pofida rezultatelor impresionante obținute prin reconstrucția bazată pe post-procesarea în spațiul imaginii, anumite detalii pot fi pierdute în timpul etapei de post-procesare cu rețele neurale, deoarece cunoștințele legate de procesul de achiziție (modelul intern al scannerului și legile fizice ale interacțiunilor razelor X), sunt de obicei neglijate. Prin urmare, metodele neurale care încorporează anumite cunoștințe legate de procesul de achiziție au potențialul de a suprima efectele nedorite, păstrând în același timp detaliile importante. Recent, ideea de a învăța reconstrucția direct prin procesarea proiecțiilor prin intermediul rețelelor neurale a fost investigată în mai multe lucrări științifice [9].

În consecință, metodele din a doua categorie vizează învățarea întregii operațiuni de reconstrucție direct din datele măsurate (proiecții), într-o manieră complet automatizată. În acest capitol se propune un model de reconstrucție bazat pe modele Deep Learning pentru a rezolva problema reconstrucției tomografiilor computerizate cu doză redusă. Deși algoritmul propus în [10] (Learned Primal-Dual) conduce la obținerea unor rezultate impresionante, imaginile obținute prezintă un efect global

de supra-netezire, cauzat de procedura de optimizare adoptată în antrenarea rețelei (minimizarea erorii pătratică medii dintre două imagini). Pentru a depăși această limitare, se propune o strategie de optimizare a metodei Learned Primal-Dual prin introducerea unei rețele adversariale generative, încurajând, de asemenea, integrarea percepției umane, care îmbunătățește detaliile structurale și texturale.

3.2 Formularea matematică a scanării computer tomografice

CT este o procedură imagistică medicală utilizată pentru a produce reprezentări anatomice volumetrice detaliate ale corpului uman. Procesul de obținere al unei tomografii computerizate include doi pași: achiziția datelor și reconstrucția imaginii. Procesul de obținere al tomografiei computerizate este descris în Figura 3.1.

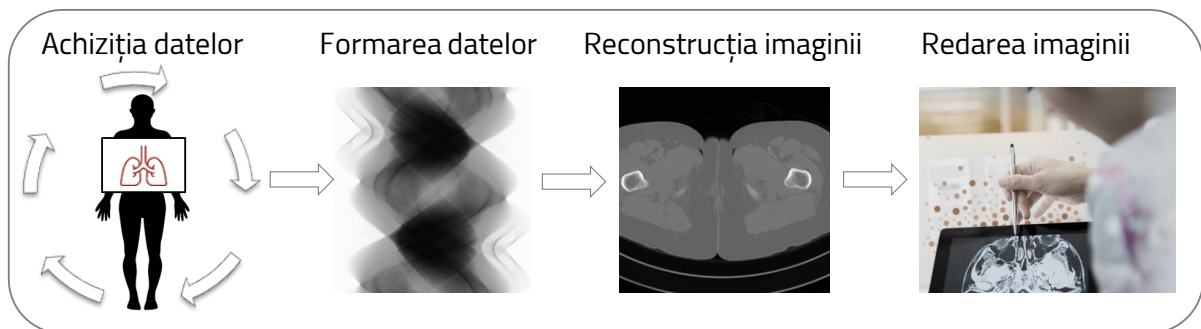


Figura 3.1: Procesul de examinare prin tomografie computerizată, de la achiziția datelor la randarea volumului.

În timpul etapei de achiziție, un obiect este plasat între sursa de raze X care emite un fascicul conic cu raze X și un detector cu ecran plat care captează absorbția de raze X a diferitelor țesuturi (materiale) de-a lungul unei linii drepte L . Informațiile obținute sunt cunoscute sub denumirea de proiecții. Pentru a observa obiectul din diferite direcții, perechea detector-sursă este rotită în jurul obiectului, capturând sute de proiecții. Proiecțiile generate la distanțe unghiulare echidistante sunt suprapuse pentru a forma o așa-numită sinogramă. Sinograma este apoi convertită în imaginea unei secțiuni transversale în timpul procesului de reconstrucție. În final, secțiunile transversale generate sunt grupate pentru a forma imaginea tridimensională a obiectului, pe baza căruia ulterior se pot identifica și localiza țesuturi, organe, anomalii etc.

Astfel, în termeni matematici, obiectivul tomografiei computerizate poate fi formulat ca recuperarea hărții valorilor coeficientului de atenuare a obiectului observat $f \in \mathcal{X}$ din măsurătorile $g \in \mathcal{Y}$, unde \mathcal{X} denotă spațiul de reconstrucție și \mathcal{Y} reprezintă spațiul de măsurare. CT este astfel obținut prin intermediul transformatei Radon (operatorul de proiecție) care creează măsurătorile g din imaginea f :

$$g = H(f), \quad (3.1)$$

unde $H : \mathcal{X} \rightarrow \mathcal{Y}$ este matricea sistemului.

Soluția tipică pentru obținerea reconstrucției se bazează pe o transformare inversă directă, descrisă în imagistica CT de transformata Radon inversă (operatorul de proiecție inversă), $H^* : \mathcal{Y} \rightarrow \mathcal{X}$, astfel încât:

$$f = H^*(g). \quad (3.2)$$

Reconstrucția obținută pe baza operatorului convențional de inversiune suferă de o puternică neclaritate, detaliile fiind estompate.

Recent, reconstrucția iterativă a devenit o soluție alternativă. Pornind de la o imagine inițială și încorporând ipotezele cu privire la sistemul de achiziție al datelor, algoritmul actualizează în mod iterativ reconstrucția, utilizând atât operatorul de proiecție, cât și cel de proiecție inversă, pentru a minimiza o funcție obiectiv. Prin urmare, tomografia cu raze X poate fi formulată ca o problemă de optimizare după cum urmează [11]:

$$\arg \min_{f \in \mathcal{X}} \|(H(f) - g)\|_2 + \gamma R(f), \quad (3.3)$$

unde termenul din stânga egalității reprezintă termenul de fidelizare a datelor și măsoară discrepanța dintre datele de proiecție inițiale măsurate g și datele de proiecție sintetizate $H(f)$. Termenul din dreapta egalității, $R : \mathcal{X} \rightarrow \mathbb{R}$, este termenul de regularizare și încorporează cunoștințele anterioare legate de f , introdus pentru a transforma problema într-una bine definită, hiperparametrul γ controlând echilibrul dintre cunoștințele anterioare și fidelitatea datelor.

Problema de optimizare formulată în 3.3 este de obicei rezolvată utilizând metodele bazate pe gradient. Principalul dezavantaj al unui algoritm de reconstrucție iterativ este timpul mare de calcul.

3.3 Reconstrucția imaginilor bazată pe învățare

Astfel, sistemele iterative bazate pe învățare reprezintă o alternativă promițătoare pentru a reduce limitările impuse de complexitatea înaltă de calcul a metodelor clasice iterative. Algoritmii de reconstrucție bazați pe modelele Deep Learning au ca și scop găsirea parametrului $\theta \in \mathbb{Z}$ al operatorului parametric H_θ^+ pe bazat setului de date de antrenare (perechi de imagini ale reconstrucțiilor reale și măsurătorile corespunzătoare ale acestora) [12], prin minimizarea unei funcții obiectiv dintre reconstrucțiile reale și cele estimate de modelul neural, astfel încât $H_\theta^+(g) \approx f$.

În reconstrucția primal-duală bazată pe învățarea automată pe baza exemplelor (Learned Primal-Dual) o schemă iterativă este desfășurată de-a lungul unei rețele neurale. Modelul obținut realizează reconstrucția, pornind de la proiecțiile achiziționate, prin actualizarea în mod alternativ atât al proiecțiilor, cât și al reconstrucției.

Rețeaua neurală combină informațiile furnizate de operatorii de proiecție și proiecție inversă (transformata Radon și inversa) cu tehnicile de filtrare specifice imaginii, obținute prin intermediul blocurile convoluționale din rețea, pentru a genera o reconstrucție intermediară îmbunătățită la fiecare iterație. Încapsularea cunoștințelor despre legile fizice de obținere a tomografiilor computerizate are multiple avantaje: permite păstrarea detaliilor și reducerea numărului de necunoscute (ceea ce conduce la o estompare a cerințelor modelelor Deep Learning referitoare la un volum mare de date de antrenare).

Prin urmare, în cadrul fiecărei iterații rețeaua învață funcția care actualizează atât imaginea, cât și proiecțiile :

$$g_i = G_{\theta_i}(g_{i-1}, H(f_{i-1}), g), \quad (3.4)$$

$$f_i = F_{\theta_i}(f_{i-1}, H^*(g_i)), \quad (3.5)$$

unde G_{θ_i} și F_{θ_i} descriu rețelele neurale convoluționale (Figura 3.2), $i = \overline{1, N}$ reprezintă iterația curentă, rezultatul ultimei iterații fiind operatorului parametric învățat (pseudo-inversa).

3.3.1 Antrenarea modelului Deep Learning de reconstrucție tomografică

Problema descrisă anterior este formulată în mod tradițional ca o problemă de regresie [10], unde rețeaua neurală este instruită pentru a furniza imagini reconstruite similar celor originale prin minimizarea unei funcții obiectiv la nivel de pixeli [de exemplu, eroare medie pătratică [eng. mean squared error (MSE)] și eroare medie absolută [eng. mean absolute error (MAE)], între reconstrucțiile reale și cele generate. Atât MSE, cât și MAE tind să fie suboptime pentru sarcina de generare de imagini, încurajând formarea de imagini neclare și descurajând obținerea texturilor și detaliilor structurale [13].

În consecință, pentru a genera tomografiile computerizate, este propusă o soluție bazată pe rețele neurale, PD-WGAN Perceptual, care operează direct pe proiecțiile achiziționate. Pentru a permite obținerea unui nivel mai ridicat de detalii în imaginile reconstruite, modelul este antrenat nu pentru a minimiza o eroare de tipul MSE, ci o funcție obiectiv de percepție (eng. perceptual loss) [14].

Pentru a genera imagini sintetice realiste, statistic indistinctibile de reconstrucțiile reale, se utilizează o rețea neurală generativă. Un astfel de model Deep Learning este format din două rețele neurale cuplate: prima rețea (generatorul) este responsabilă cu generarea de imagini CT realiste pornind de la proiecțiile achiziționate, a doua rețea (discriminatorul) trebuie să facă distincția dintre imaginile reale și cele sintetice. Antrenarea modelului se realizează antrenând simultan atât generatorul cât și discriminatorul. Astfel, cu cât discriminatorul devine mai bun în a distinge imaginile false de cele adevărate cu atât generatorul va genera imagini mai realiste.

Pentru a asigura menținerea stabilității în timpul antrenării, se utilizează o rețea de tip WGAN (Wasserstein Generative Adversarial Network) cu penalizarea gradientului [15], al cărei obiectiv general este de a învăța distribuția datelor reale prin minimizarea distanței Wasserstein dintre distribuția datelor reale și distribuția generatorului. Problema este definită astfel: fie un exemplu f extras din distribuția reală a imaginilor reconstruite \mathbb{P}_f , și g din distribuția datelor de proiecție \mathbb{P}_g , generatorul G învață să mapeze datele dintr-o distribuție \mathbb{P}_g la spațiul datelor \mathbb{P}_f , în timp ce discriminatorul D estimează probabilitatea ca un exemplu generat să facă parte din distribuția reală \mathbb{P}_f . În final, obiectivul învățării pentru modelul WGAN poate fi formulat astfel ca un joc minmax în care discriminatorul este antrenat pentru a maximiza acuratețea clasificării, iar generatorul pentru a minimiza diferența dintre cele două distribuții:

$$\min_G \max_D L_{adversarial} = L_{critic} + \lambda L_{penalty}, \quad (3.6)$$

$$L_{critic} = \mathbb{E}_{\tilde{f} \sim \mathbb{P}_g} [D(\tilde{f})] - \mathbb{E}_{f \sim \mathbb{P}_f} [D(f)], \quad (3.7)$$

$$L_{penalty} = \mathbb{E}_{\hat{f} \sim \mathbb{P}_{\hat{f}}} [(\|\Delta_{\hat{f}} D(\hat{f})\|_2 - 1)^2], \quad (3.8)$$

unde L_{critic} este responsabil cu estimarea distanței Wasserstein și $L_{penalty}$ este termenul de penalizare al gradientului, a cărui contribuție este echilibrată prin intermediul parametrului λ . În ecuațiile anterioare $\tilde{f} = G(g)$ reprezintă o imagine generată de generatorul G din proiecțiile g , în timp ce \hat{f} descrie un exemplu extras dintr-o distribuție uniformă $\mathbb{P}_{\hat{f}}$ (eșantionarea este efectuată de-a lungul liniilor drepte între perechile de exemple eșantionate din \mathbb{P}_f și \mathbb{P}_g).

În consecință, obiectivul generatorului este de a face discriminatorul să creadă că un exemplu generat este real, minimizând astfel funcția:

$$L_{generator} = - \mathbb{E}_{\tilde{f} \sim \mathbb{P}_g} [D(\tilde{f})]. \quad (3.9)$$

Obiectivul discriminatorului este de a face o mai bună distincție între exemplele reale și cele generate, astfel încât să se maximizeze probabilitatea de asociere corectă a datelor:

$$L_{discriminator} = \mathbb{E}_{\tilde{f} \sim \mathbb{P}_g} [D(\tilde{f})] - \mathbb{E}_{f \sim \mathbb{P}_f} [D(f)] + \mathbb{E}_{\hat{f} \sim \mathbb{P}_{\hat{f}}} [(\|\Delta_{\hat{f}} D(\hat{f})\|_2 - 1)^2]. \quad (3.10)$$

În timp ce versiunea originală a modelului WGAN utilizează un discriminator global, în cadrul acestui model se adoptă un discriminator local [16], care analizează regiunile individuale din imaginea generată și furnizează o hartă a scorului discriminativ. Scorul discriminativ oferă informații constructive reconstrucției, aplicând constrângeri suplimentare generatorului, ceea ce duce la obținerea unor imagini de o calitate mai înaltă.

În plus, pentru a încuraja imaginile generate să fie cât mai realiste, generatorul are asociată o funcție obiectiv de percepție, bazată pe similitudinea a două imagini (distanța dintre caracteristicile

de nivel înalt extrase din imagini dintr-o rețea pre-antrenată):

$$L_{content} = \mathbb{E}_{g \sim \mathbb{P}_g, x \sim \mathbb{P}_f} [\|\Phi(G(g)) - \Phi(f)\|_2^2], \quad (3.11)$$

unde Φ este o hartă de trăsături (eng. feature map) extrasă de către rețea, $G(g)$ reprezintă reconstrucția generată și f este imaginea tomografică reală. Ultimul strat convoluțional din rețeaua VGG19 [17] operează ca un modul de extragere al caracteristicilor din imagini (eng. feature extractor).

Considerând funcția obiectiv adversarială $L_{adversarial}$ și cea de percepție $L_{content}$, o funcție obiectiv finală este definită ca o sumă ponderată astfel:

$$\min_G \max_D L_{content} + \lambda_1 L_{adversarial} + \lambda_2 L_{supervision}, \quad (3.12)$$

A se reține că o funcție obiectiv de supervizare pentru o comparație directă a imaginilor din spațiul reconstrucției, poate fi, de asemenea, încorporată pentru o constrângere suplimentară a generatorului:

$$L_{supervision} = \mathbb{E}_{g \sim \mathbb{P}_g, f \sim \mathbb{P}_f} [\|G(g) - f\|_1], \quad (3.13)$$

Procedura de antrenare pentru modelul PD-WGAN propus este descrisă în Algoritmul 3.1.

3.3.2 Arhitectura modelului Deep Learning

Arhitectura modelului de reconstrucție propus, Perceptual PD-WGAN, pentru formularea problemei reconstrucției tomografice iterative, este descrisă în Figura 3.2.

3.4 Experimente

Pentru a compara și evalua algoritmul propus, s-a considerat următoarea aplicație clinică realistă: reconstrucția tomografiei computerizate cu doză de radiație completă pe baza datelor de proiecție simulate cu doză redusă.

Datele de antrenare conțin scanări abdominale CT dintr-un set de date pus la dispoziția participanților la competiția NIH, AAPM and Mayo Clinic Low Dose CT Grand Challenge [18]. Setul de date include volume CT provenite de la 10 pacienți. Pentru experimentele propuse, setul de date a fost împărțit la nivel de pacienți, rezultând 2198 512×512 secțiuni CT bidimensionale pentru faza de antrenare și 210 de secțiuni CT pentru validare. Astfel, problema este formulată ca reconstrucția secțiunilor transversale bidimensionale din mai multe proiecții unidimensionale (integrala de linie) ale obiectului observat.

Deoarece rețeaua propusă încorporează cunoștințele privind geometria sistemului CT sub forma operatorilor de proiecție și proiecție inversă, modelul CT a fost utilizat pentru a genera în mod sintetic datele de proiecție aferente unei doze reduse de radiații, pornind de la reconstrucțiile cu doză completă de radiații.

3.5 Rezultate

Rețeaua Learned Primal-Dual antrenată cu funcția obiectiv MSE, PD-MSE, a fost implementată ca metodă principală de referință. Structura rețelei a fost identică cu cea din lucrarea [10]. Soluția propusă (Perceptual PD-WGAN) a fost astfel comparată cu trei algoritmi de reconstrucție: algoritmul clasic de reconstrucție bazat pe filtrarea proiecțiilor [Filtered Back-Projection (FBP)], reconstrucția cu reducerea variației totale [Total Variation (TV)] și reconstrucția bazată pe învățare PD-MSE.

Algorithm 3.1 Algoritmul de antrenare al modelului PD-WGAN.

Input: Generatorul G cu parametrii θ , Discriminatorul D cu parametrii ω , rețeaua VGG-19 pre-antrenată.

Input: Operatorul de proiecție și proiecție inversă H, H^* , numărul de iterații N , spațiul primal și dual N_{primal}, N_{dual} .

Input: Setul de antrenare $\{\mathbf{g}^{(j)}, \mathbf{f}^{(j)}\}_{j=1}^M$, numărul de epoci de antrenare N_{epochs} , dimensiunea unui batch b , numărul de iterații pentru discriminator și generator N_D, N_G , factorii de ponderare $\lambda, \lambda_1, \lambda_2$, hiperparametrii optimizatorului Adam α, β_1, β_2 .

Output: Actualizarea parametrilor θ și ω .

```
1: procedure Antrenare( $\mathbf{g}, \mathbf{f}$ )
2:   Inițializarea parametrilor modelului:  $\theta, \omega$ 
3:   for  $epoch = 1, \dots, N_{epochs}$  do
4:     for  $critic = 1, \dots, N_D$  do
5:       Extragerea unui batch de  $b$  exemple de antrenare:  $\{\mathbf{g}^{(i)}, \mathbf{f}^{(i)}\}_{i=1}^b \sim \{\mathbf{g}^{(j)}, \mathbf{f}^{(j)}\}_{j=1}^M$ 
6:       for  $i = 1, \dots, m$  do
7:         Generarea unui număr aleator:  $\epsilon \sim Uniform[0, 1]$ 
8:         Generarea reconstrucției:  $\tilde{\mathbf{f}}^{(i)} \leftarrow G(\mathbf{g}^{(i)}, H, H^*, N, N_{primal}, N_{dual})$ 
9:         Calculul:  $\hat{\mathbf{f}}^{(i)} \leftarrow \epsilon \mathbf{f}^{(i)} + (1 - \epsilon) \tilde{\mathbf{f}}^{(i)}$ 
10:        Calculul funcției obiectiv:  $L_D^{(i)} \leftarrow D(\tilde{\mathbf{f}}^{(i)}) - D(\mathbf{f}^{(i)}) + \lambda(\|\Delta \hat{\mathbf{f}}^{(i)} D(\hat{\mathbf{f}}^{(i)})\|_2 - 1)^2$ 
11:      end for
12:    end for
13:    Actualizarea discriminatorului  $D$ :  $\omega \leftarrow Adam(\omega, L_D, \alpha, \beta_1, \beta_2)$ 
14:    for  $generator = 1, \dots, N_G$  do
15:      Extragerea unui batch de  $b$  exemple de antrenare:  $\{\mathbf{g}^{(i)}, \mathbf{f}^{(i)}\}_{i=1}^b \sim \{\mathbf{g}^{(j)}, \mathbf{f}^{(j)}\}_{j=1}^M$ 
16:      for  $i = 1, \dots, m$  do
17:        Generarea reconstrucției:  $\tilde{\mathbf{f}}^{(i)} \leftarrow G(\mathbf{g}^{(i)}, H, H^*, N, N_{primal}, N_{dual})$ 
18:        Calculul funcției obiectiv perceptuală:  $L_{content}^{(i)} \leftarrow \|VGG(\tilde{\mathbf{f}}^{(i)}) - VGG(\mathbf{f}^{(i)})\|_2^2$ 
19:        Calculul funcției obiectiv de supervizare:  $L_{supervision}^i \leftarrow \|G(\tilde{\mathbf{g}}^{(i)}) - \mathbf{f}^{(i)}\|_1$ 
20:        Calculul funcției obiectiv:  $L_G^{(i)} \leftarrow \lambda_1 L_{content}^i + \lambda_2 L_{supervision}^i + D(\tilde{\mathbf{f}}^{(i)})$ 
21:      end for
22:    end for
23:    Actualizarea generatorului  $G$ :  $\theta \leftarrow Adam(\theta, L_G, \alpha, \beta_1, \beta_2)$ 
24:  end for
25: end procedure
```

3.5.1 Evaluarea calitativă

Pentru a evalua calitatea imaginilor tomografice generate, au fost selectate secțiuni reprezentative din setul de date de testare, iar rezultatele vizuale obținute sunt prezentate în Figura 3.3. Se poate observa că metoda de reconstrucție propusă, PD-WGAN, poate atenua, într-o anumită măsură, efectul de supra-netezire. Cu toate acestea, în comparație cu imaginea de referință, există încă unele detalii care par a fi zgomotoase, și structuri care tind să fie ușor distorsionate.

Calitatea reconstrucțiilor a fost în continuare validată, cu accent pe utilizarea clinică, prin intermediul unui studiu orb randomizat cu trei radiologi experimentați. Imaginile au fost evaluate pe scala Likert cu cinci valori, din trei perspective: zgomot, artefacte și diagnosticare.

Media calculată a valorilor alese de către radiologi, alături de punctajele individuale, este prezentată în Figura 3.4. Pe scara comparativă în cinci puncte, metoda propusă de reconstrucție, PD-WGAN, a obținut cele mai mari scoruri pe toate cele trei criterii. Toți evaluatorii au fost de acord că imaginile reconstruite pe baza algoritmului PD-WGAN dispun de un nivel redus de zgomot și artefacte în

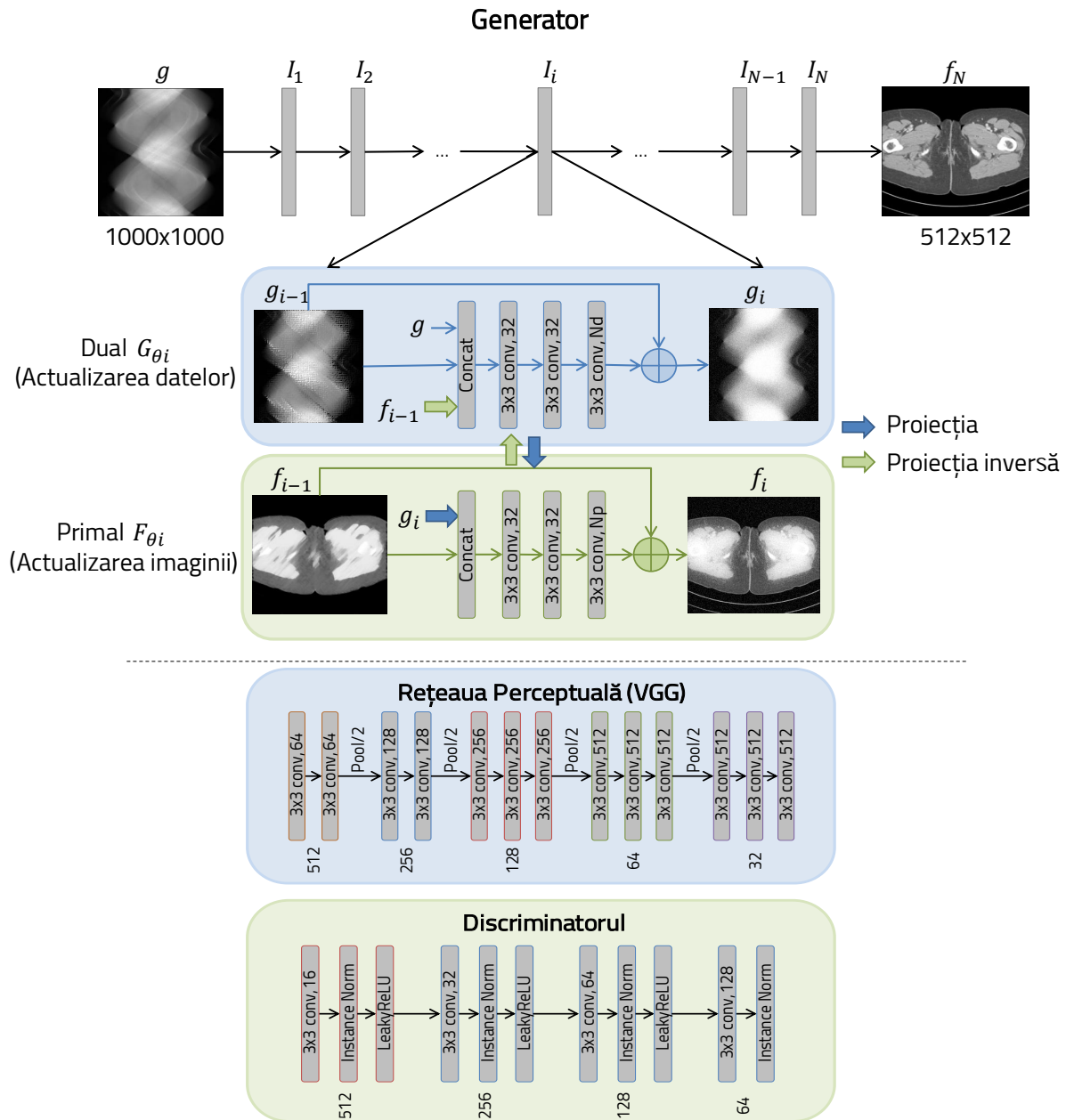


Figura 3.2: Structura generală a rețelei propuse, Perceptual PD-WGAN, pentru rezolvarea unei probleme tomografice. Schema iterativă este desfășurată de-a lungul a $N = 8$ iterații, cu parametrii de antrenare independenți. Pentru a forța persistența informațiilor între iterații, 5 hărți de trăsături au fost adoptate pentru cele două ramuri, primal și dual, $Nd = Np = 5$.

comparație cu nivelurile percepute pentru imaginile reconstruite cu modelul PD-MSE. De asemenea, evaluatorii au identificat metoda de reconstrucție propusă PD-WGAN ca oferind imagini de calitate superioară acceptabile în vederea diagnosticării.

3.5.2 Evaluarea cantitativă

Pentru o evaluare cantitativă a calității reconstrucției, Tabela 3.1 afișează valorile medii și abaterile standard pentru două măsurători de calitate a imaginii: raportul semnal zgomot de vârf [eng. Peak Signal to Noise Ratio (PSNR)] și indicele de similaritate structurală [eng. Structural Similarity Index Measure (SSIM)]. De asemenea, timpii necesari realizării unei reconstrucții, pentru fiecare al-

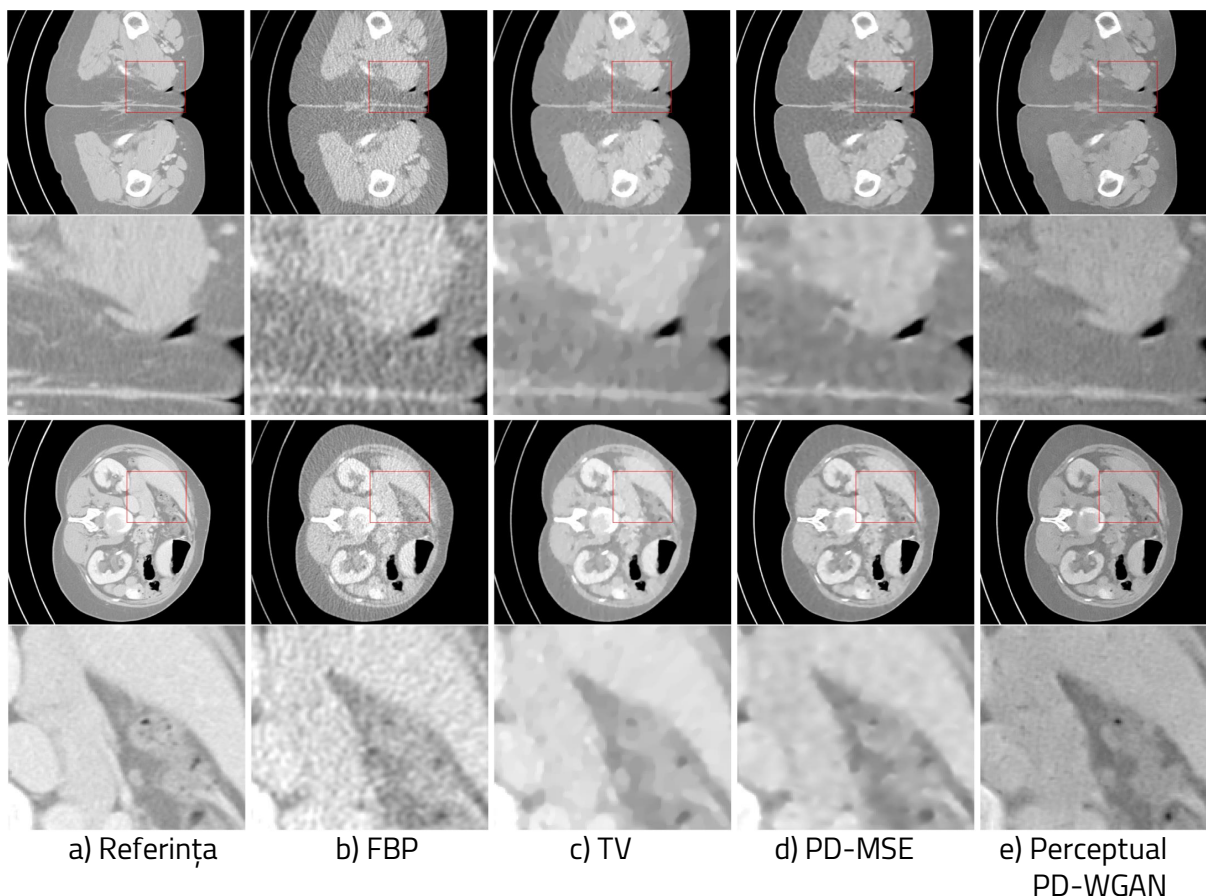


Figura 3.3: Reconstrucții tomografice cu doză completă: Imagini CT cu secțiunea transversală abdominală, alături de o regiune mărită delimitată de dreptunghiul roșu. Fereastra de afișare este de [-500, 200] HU.

goritm considerat, sunt afișați în Tabela 3.1.

Tabela 3.1: Compararea cantitativă a algoritmilor de reconstrucție.

Metoda	Metrica de evaluare		Timpul de rulare (s)
	SSIM	PSNR	
FBP	0.827±0.018	34.533±1.451	0.452±0.084
TV	0.940±0.002	37.887±1.432	59.581±0.829
PD-MSE	0.954±0.004	42.598±1.312	0.537±0.137
Perceptual PD-WGAN	0.911±0.006	31.375±2.218	0.520±0.176

Deși rezultatele obținute de modelul propus sunt din punct de vedere vizual mai asemănătoare cu imaginile de referință, acest lucru nu este reflectat de măsurătorile PSNR și SSIM. Totuși, acest lucru nu este surprinzător, deoarece rețeaua tomografică propusă a fost antrenată să minimizeze o funcție obiectiv bazată pe caracteristici într-o manieră adversarială, nu cu intenția de a reduce diferențele la nivel de pixeli, ci pentru a încuraja generarea de imagini realiste.

În timp ce modelului PD-WGAN necesită un timp mai îndelungat de antrenare în comparație cu modelul PD-MSE, timpul necesar pentru reconstrucție este relativ identic, structura adversarială a rețelei este neglijată în timpul procesului de inferență. Mai mult decât atât, se poate constata că reconstrucțiile bazate pe modele Deep Learning sunt semnificativ mai rapide, două ordini de mărime,

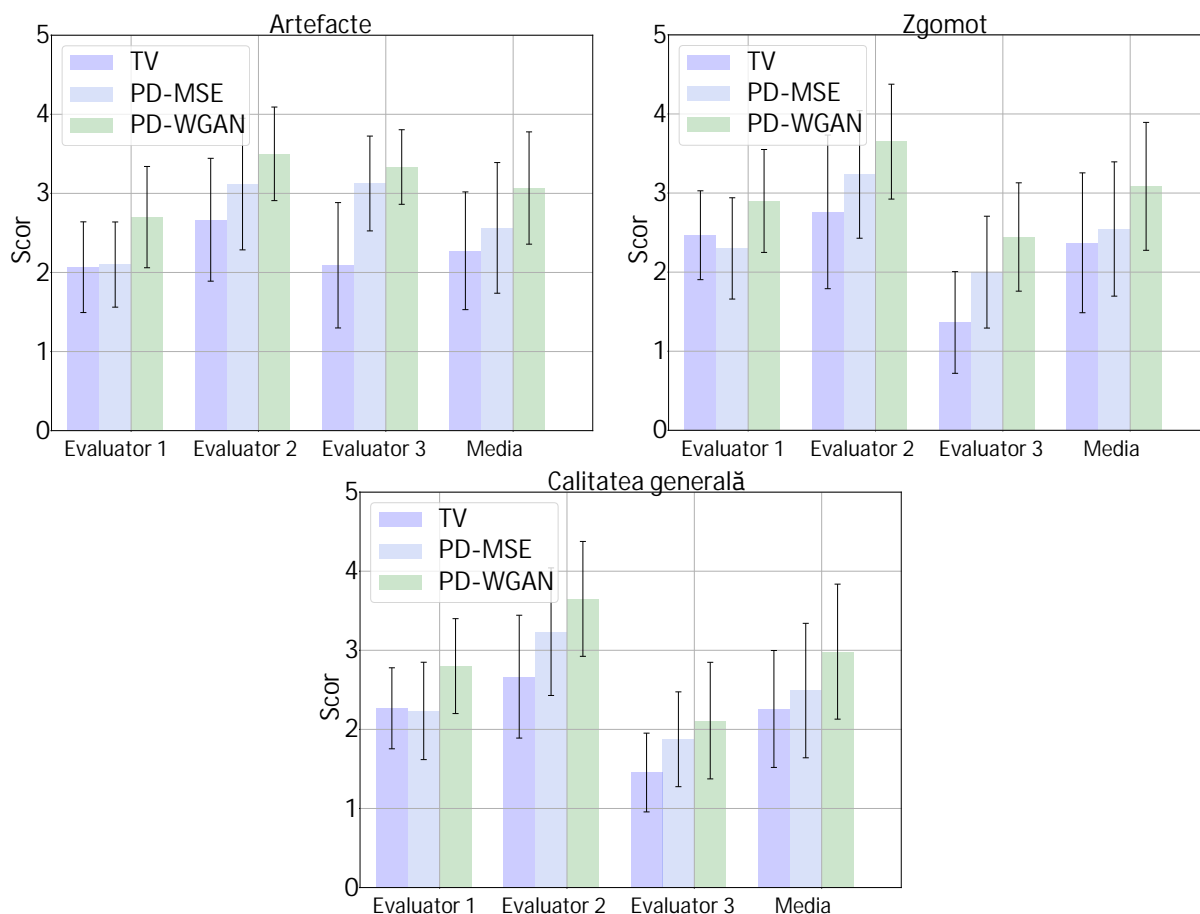


Figura 3.4: Media scorurilor calitative provenite de la cei trei evaluatori pe o scară de tip Likert cu cinci puncte (1 = inacceptabil, 2 = slabă, 3 = acceptabilă, 4 = bună și 5 = excelentă). Punctajele medii cu privire la calitatea generală a imaginii pentru diagnosticare, reducerea zgomotul de imagine și a artefactelor sunt afișate pentru fiecare examinator în parte. În final, sunt descrise și punctajele medii globale.

față de metoda iterativă TV.

3.6 Discuții generale și concluzii

S-a demonstrat într-un scenariu realist din punct de vedere clinic că operatorul parametric de reconstrucție tomografică poate fi învățat direct din perechi de date de antrenare, proiecție-reconstrucție, fără a fii necesară o inițializare a imaginii în prealabil sau efectuarea unei inversiuni directe. Spre deosebire de soluțiile Deep Learning bazate pe ideea de post-procesare, în care reconstrucția tomografică este formulată pentru a îmbunătăți o reconstrucție existentă, algoritmul propus se bazează pe proprietățile interne ale scannerului și pe legile fizice ale interacțiunilor razelor X permițând obținerea de reconstrucții direct de la datele achiziționate.

Deși în literatura de specialitate au fost raportate rezultate promițătoare în ceea ce privesc metodele de post-procesare bazate pe modele Deep Learning în tomografia computerizată, acestea sunt în general limitate datorită informațiilor incomplete de care dispun. Astfel, soluția propusă, nu numai că include cunoștințele utile legate de obținerea imagisticii CT, ci permite efectuarea corecțiilor atât în spațiul reconstrucției, cât și în cel al proiecției.

Soluția propusă se bazează pe ideea desfășurării unui algoritm iterativ de reconstrucție a imaginii într-un număr finit de iterații reprezentat prin intermediul unui model Deep Learning. Astfel, s-a adoptat o arhitectură de tipul WGAN în care blocul generator este responsabil de generarea recon-

strucției finale prin învățarea filtrelor și ajustând datele corespunzător. În plus, modelul neural integrează informațiile existente despre obținerea imagisticii CT și îmbunătățește strategia de antrenare prin impunerea percepției umane asupra calității reconstrucției generate.

În cadrul experimentelor propuse, modelul propus a condus la obținerea de rezultate promițătoare, încurajând atât conservarea detaliilor cât și al texturilor. Algoritmul depășește din punct de vedere calitativ metoda propusă anterior în literatura de specialitate (PD-MSE), reducând tendința de supra-netezire și astfel sporind calitatea imaginii în ceea ce privește capacitatea generală de diagnosticare, eliminarea artefactelor și suprimarea zgomotului.

4. Detecția automată a leziunilor mamare în mamografiile digitale cu tomosinteză

4.1 Introducere

Deși cancerul de sân este cunoscut ca fiind unul dintre principalele cauze ale deceselor cauzate de cancer în rândul femeilor, diagnosticul precoce și tratamentul pot reduce drastic rata mortalității [19]. Explorarea mamografică (examinarea cu raze X) a sânelor este principala metodă de investigație imagistică utilizată pentru depistarea cancerului de sân în stadii incipiente. În timp ce utilizarea mamografiei s-a dovedit a reduce semnificativ rata mortalității, această tehnică suferă de o rată mai mare a rechemărilor (număr mare de rezultate fals-pozitive) și a erorilor de depistare cancerului la sân [20].

În prezent, tehnologia de vârf în screeningul cancerului de sân este mamografia digitală cu tomosinteză (mamografia 3D). Această metodă imagistică este o formă avansată a mamografiei convenționale, având potențialul de a îmbunătăți atât sensibilitatea, cât și specificitatea în localizarea și caracterizarea leziunilor de la nivelul sânelor [21]. Mamografia cu tomosinteză permite o redare volumetrică a întregului sân, reducând astfel efectele suprapunerii țesuturilor.

În timpul examinării, două proiecții mamografice standard sunt achiziționate pentru fiecare sân, corespunzând planurilor cranio-caudal (CC) și mediolateral oblic (MLO). Astfel, fiecare sân este comprimat și examinat din două direcții non-ortogonale, obținându-se o vedere în spațiu a structurilor sânelor. Prin urmare, examinarea mamografică pe baza celor 4 imagini, după cum se arată în Figura 4.1, poate oferi o mai bună evidențiere a diferitelor structuri mamare care se suprapun, sau a anumiților patologii și, implicit, poate îmbunătăți detecția și diagnosticul tumorilor [22].

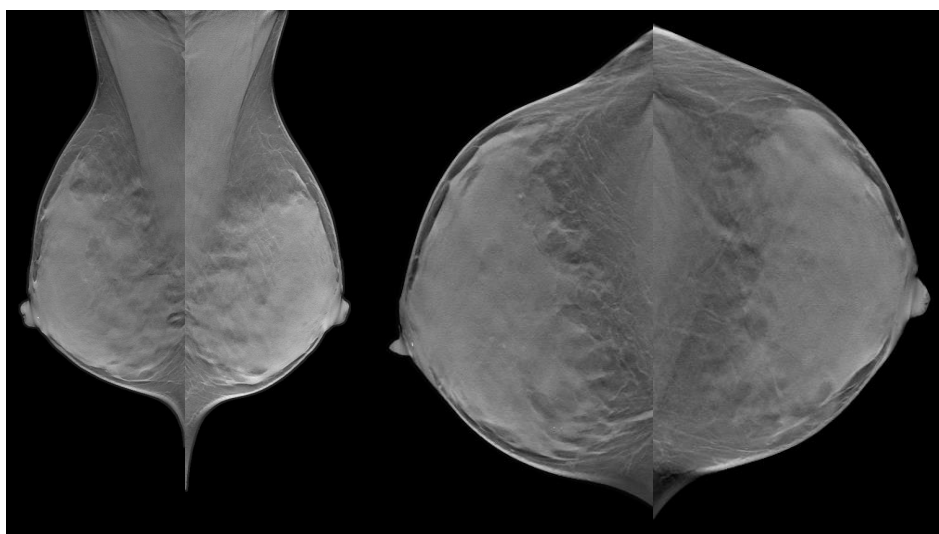


Figura 4.1: Secțiuni din mamografiile 3D ce descriu proiecțiile mediolaterale oblice dreaptă și stângă și proiecțiile cranio-caudale dreaptă și stângă ale sânelor pentru diagnostic.

Chiar dacă mamografia 3D este o tehnică de screening din ce în ce mai des folosită datorită avan-

tajelor clare aduse procesului de screening al cancerului de sân, în rutinele clinice, mamografia clasică rămâne încă metoda preferată. Acest lucru se datorează în principal faptului că mamografia cu tomosinteză este încă o tehnică relativ nouă, care vine cu un cost de achiziție mai mare și care, datorită volumului mai mare de informații, necesită un timp mai lung pentru interpretare.

Având în vedere că examinarea și interpretarea mamografiilor este o problemă clinică complexă, se propune o soluție bazată pe modele Deep Learning pentru identificarea și marcarea zonelor de interes din mamografiile cu tomosinteză în vederea facilitării detecției cancerului de sân. Soluția dezvoltată utilizează o rețea neurală complet convoluțională capabilă să identifice trăsăturile ce semnalează prezența unei leziuni și să recunoască într-o manieră complet automată modele-tip în vederea localizării leziunilor mamare la nivelul secțiunilor mamografiilor 3D. Prin urmare, modelul neural este proiectat să marcheze zonele suspecte, și astfel asista radiologul la examinarea acelor zone care altfel ar fi rămas nedetectate. În consecință, soluția dezvoltată poate fi utilizată ca un instrument de sprijin în procesul de detecție, însă responsabilitatea de a lua decizia finală, pe baza raționamentelor clinice și experienței, revine radiologului.

Pentru a valorifica potențialul modelelor bazate pe învățarea automată din date în interpretarea imaginii tridimensionale a sânelui, se urmează trei direcții principale: (i) reducerea influenței lipsei unui set masiv de date mamografice cu tomosinteză, (ii) adresarea problemei referitoare la natura subtilă a delimitării leziunilor de țesutul mamar și (iii) reducerea numărului de detecții fals-pozitive.

Astfel, într-o primă instanță, se adoptă o strategie de pre-antrenare a modelului neural în două etape cu scopul de a reduce, într-o anumită măsură, problemele cauzate de lipsa unui set de date suficient de mare pentru modelele Deep Learning. Mai precis, soluția propusă ia în considerare ideea de a reutiliza filtrele convoluționale, și deci caracteristicile, învățate de rețea pe un set de date diferit care dispune de milioane de imagini de antrenare. În acest caz, în loc de a inițializa în mod aleatoriu modelul și antrena de la zero pe date mamografice cu tomosinteză, o serie de straturi ale rețelei sunt pre-inițializate cu ponderile identificate anterior. Astfel, modelul este pre-antrenat, într-o primă instanță, pe imagini din sfera computer vision. Ulterior, modelul este adaptat pentru a funcționa pe imagini mamografice convenționale, și în final antrenat pe mamografiile cu tomosinteză. În plus, pentru a trata problema tranziției inexacte dintre leziuni și țesuturile adiacente, algoritmul este proiectat special pentru a facilita identificarea leziunilor iar strategia de antrenare împreună cu funcția obiectiv sunt adaptate în consecință. În final, pentru a crește precizia modelului în detecția leziunilor, reducând numărul rezultatelor fals-pozitive sau fals-negative, se propune un algoritm de identificare a corespondențelor dintre leziunile marcate de model pe cele două proiecții mamografice standard ale sânelui (CC și MLO).

4.2 Detecția leziunilor mamare asistată de calculator

4.2.1 Formularea problemei

În cadrul sarcinilor de imagistică medicală orientate pe identificarea obiectelor de interes, scopul este de a localiza și evidenția regiunile în care se află obiectele de interes. Atunci când se antrenează un model Deep Learning într-o manieră supervizată, rețeaua neurală învață să extragă automat astfel de regiuni prin minimizarea unei funcții obiectiv dintre referința adnotată manual și predicție.

Pentru a reduce complexitatea problemei, în loc de a viza regresia directă a coordonatelor dreptunghiului de încadrare (Figura 4.2a) sau de a clasifica fiecare pixel din imagine (Figura 4.2b), problema de detecție poate fi formulată ca generarea hărții probabilistice (Figura 4.2c) pentru fiecare leziune. Prin urmare, referința (ieșirea așteptată) este codificată prin intermediul densității de probabilitate a repartiției normale astfel:

$$f(x, y, z) = e^{-\frac{(x-\mu_x)^2}{0.1\sigma_x^2}} \cdot e^{-\frac{(y-\mu_y)^2}{0.1\sigma_y^2}} \cdot e^{-\frac{(z-\mu_z)^2}{0.1\sigma_z^2}}, \quad (4.1)$$

unde x, y, z sunt coordonatele voxelilor imaginii tridimensionale, μ_x, μ_y, μ_z sunt coordonatele ce descriu centrul leziunii (extras din dreptunghiului de încadrare marcat de către radiolog) și $\sigma_x, \sigma_y, \sigma_z$ re-

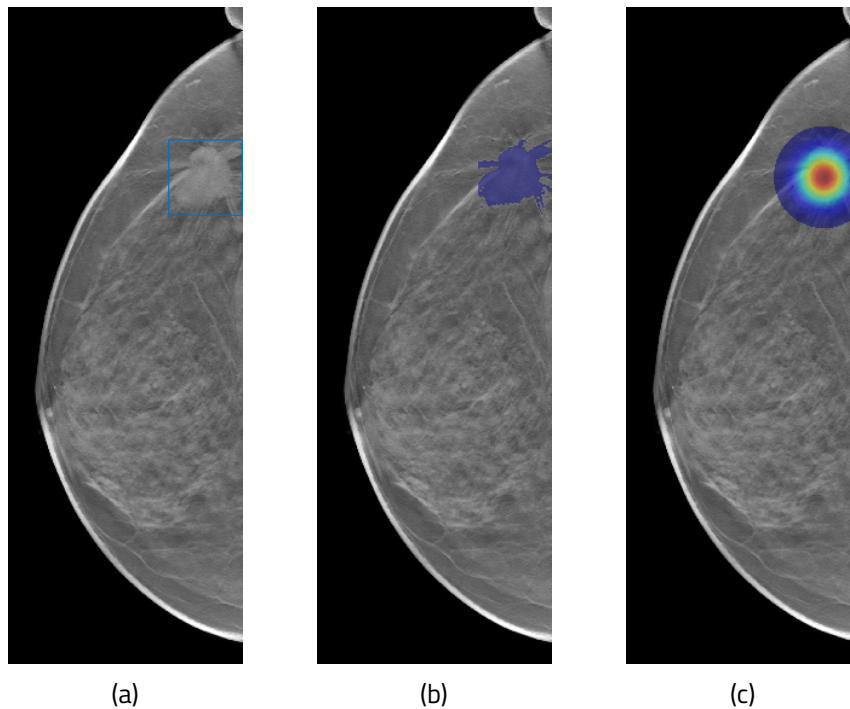


Figura 4.2: Exemplu de secțiune a mamografiei cu tomosinteză ce prezintă o leziune marcată prin (a) coordonatele dreptunghiului de încadrare a leziunii, (b) harta segmentării și (c) harta probabilistică.

prezintă lățimea, înălțimea și adâncimea dreptunghiului. Prin urmare, în procesul de antrenare, o rețea neurală încearcă să minimizeze distanța dintre harta probabilistică de referință și cea generată.

Prin codificarea informațiilor la nivel de pixel, modelul poate identifica un număr variabil de obiecte într-o imagine, amplasate în diferite locații spațiale, fără a necesita cunoștințe a priori despre numărul total de obiecte așteptate. Deoarece harta probabilistică oferă o încredere mai mică pixelilor aflați la limitele obiectului, aceasta introduce o modalitate de a reduce implicațiile introduse de natura subtilă a delimitării leziunilor de țesutul mamar.

4.3 Detecția leziunilor mamare cu modele Deep Learning

4.3.1 Setul de date clinic

Pentru a efectua experimente bazate pe învățare, se utilizează o bază de date formată din mamografii cu tomosinteză. Datele sunt grupate la nivel de pacient în seturi de antrenare, validare și testare. Partiționarea datelor se realizează astfel încât toate proprietățile statistice să fie bine păstrate în seturile de date.

Tabela 4.1 descrie seturile de date utilizate în acest studiu.

4.3.2 Arhitectura rețelei

În acest document, se propune o arhitectură de tipul encoder-decoder pentru a aborda problema detecției leziunilor în tomografiile cu tomosinteză. Rețeaua (Figura 4.3) este formată din două părți: un encoder responsabil cu extragerea caracteristicilor contextuale prin reducerea continuă a dimensiunii hărților de caracteristici și un decoder utilizat pentru expandarea hărților de caracteristici și recuperarea informațiilor spațiale. În plus, se adaugă conexiuni (skip) între straturile corespunzătoare ale celor două module pentru a facilita recuperarea informațiilor spațiale pierdute, dar și pentru a captura detaliile fine. Procedând astfel, informațiile contextuale capturate de encoder sunt combinate

Tabela 4.1: Împărțirea datelor mamografice cu tomosinteză în seturi de: antrenare, validare și testare.

Set de date	Categorie	Nr. de pacienți	Nr. de volume	Nr. de secțiuni
Antrenare	Pozitiv	271	345	5375
	Negativ	871	1318	74011
Validare	Pozitiv	130	156	2500
	Negativ	384	531	29213
Testare	Pozitiv	172	228	3528
	Negativ	573	805	45241

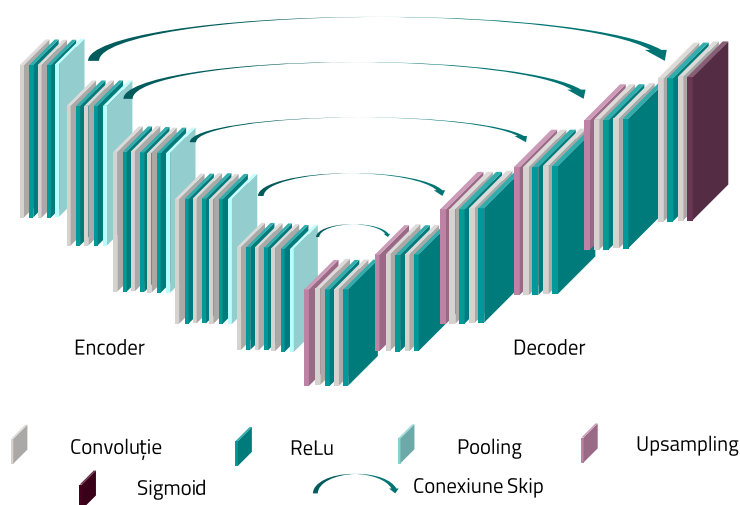


Figura 4.3: Arhitectura modelului Deep Learning propus pentru detecția automată a leziunilor mamare în tomografiile cu tomosinteză.

cu informațiile spațiale ce provin de la decoder. Rețeaua este formată integral din straturi convoluționale (complet convoluțională) cu filtre de dimensiuni mici (3×3), iar pentru reducerea spațială a hărților de caracteristici se utilizează straturi de pooling. Tabela 4.2 prezintă în detaliu topologia modelului propus.

4.3.3 Detaliile de antrenare

Analiza bazată pe modele Deep Learning direct pe mamografiile 3D cu rezoluție înaltă este însă dificil de realizat, deoarece atât complexitatea rețelei, cât și dimensionalitatea datelor au un impact semnificativ asupra utilizării memoriei procesoarelor grafice. În activitatea de cercetare curentă, accentul se pune pe facilitarea dezvoltării unui sistem de detecție asistată de calculator. Mai precis, problema de identificare a leziunilor mamare este abordată prin analiza, pe baza modelelor Deep Learning, a mamografiilor cu tomosinteză la nivelul secțiunilor bidimensionale.

Datorită rezoluției spațiale înalte, utilizarea directă a secțiunilor bidimensionale limitează atât alegerea arhitecturii cât și complexitatea rețelei. Astfel, pentru o mai mare flexibilitate, s-a optat pentru antrenarea modelului pe subregiuni extrase din secțiunile bidimensionale. Modelul propus este optimizat prin maximizarea unei variante continue și diferențiabile a coeficientului de similaritate Jaccard

Tabela 4.2: Detaliile de configurare a modelului Deep Learning propus pentru detecția leziunilor mamare (DB = downsampling block, UB = upsampling block).

Blocuri	Straturi	Nr. filtre	Dimensiune
Intrare	-	-	(1024,1024,3)
DB + MaxPooling	2 × (3×3 Conv-ReLU)	64	(1024,1024,64)
DB + MaxPooling	2 × (3×3 Conv-ReLU)	128	(512,512,128)
DB + MaxPooling	3 × (3×3 Conv-ReLU)	256	(256,256,256)
DB + MaxPooling	3 × (3×3 Conv-ReLU)	512	(128,128,512)
DB + MaxPooling	3 × (3×3 Conv-ReLU)	512	(64,64,512)
Upsampling + UB	3×3 Conv-ReLU	512	(64,64,512)
	3×3 Conv-ReLU	256	(64,64,256)
Upsampling + UB	3×3 Conv-ReLU	512	(128,128,512)
	3×3 Conv-ReLU	256	(128,128,256)
Upsampling + UB	3×3 Conv-ReLU	512	(256,256,512)
	3×3 Conv-ReLU	128	(256,256,128)
Upsampling + UB	3×3 Conv-ReLU	256	(512,512,256)
	3×3 Conv-ReLU	64	(512,512,64)
Upsampling + UB	3×3 Conv-ReLU	128	(1024,1024,128)
	3×3 Conv-ReLU	64	(1024,1024,64)
Ieșire	3×3 Conv-ReLU	32	(1024,1024,32)
	3×3 Conv-Sigmoid	1	(1024,1024,1)
Nr. de parametrii	17M		

[23, 24] dintre imaginea de referință T și imaginea cu harta de probabilități estimată de rețea P :

$$L_{Jaccard} = 1 - \frac{\sum_i T_i \cdot P_i}{\sum_i T_i^2 + \sum_i P_i^2 - \sum_i T_i \cdot P_i} \quad (4.2)$$

Trebuie remarcat faptul că numai exemplele de antrenare ce conțin leziuni (exemple pozitive) contribuie la calculul funcției eroare.

Ca și strategie de antrenare, se adoptă tehnica transferului de cunoștințe în două etape, după cum se arată în Figura 4.4. Într-o primă instanță, pornind de la un decoder inițializat aleatoriu și de la un encoder pre-antrenat (parametrii modelului VGG19 [17] pe setului de date ImageNet[25]), parametrii rețelei sunt ajustați pentru a rezolva o problemă de segmentare a leziunilor mamelor în mamografiile convenționale. Antrenarea este realizată pe subregiuni bidimensionale cu dimensiunea de 1024×1024 pixeli. Funcția eroare Dice [26] este adoptată ca funcție de optimizare a modelului de segmentare:

$$L_{Dice} = 1 - \frac{2 \sum_i T_i \cdot P_i}{\sum_i T_i^2 + \sum_i P_i^2} \quad (4.3)$$

În continuare, modelul rezultat este ajustat pentru a rezolva problema de detecție a leziunilor mamare pe mamografiile cu tomosinteză, prin optimizarea funcției Jaccard.

4.3.4 Rezultate

Câteva exemple ale estimărilor rețelei (harta de probabilități), obținute de modelul propus pe imagini din setul de date de testare, sunt prezentate în Figura 4.5a. Pentru a determina locația 3D finală

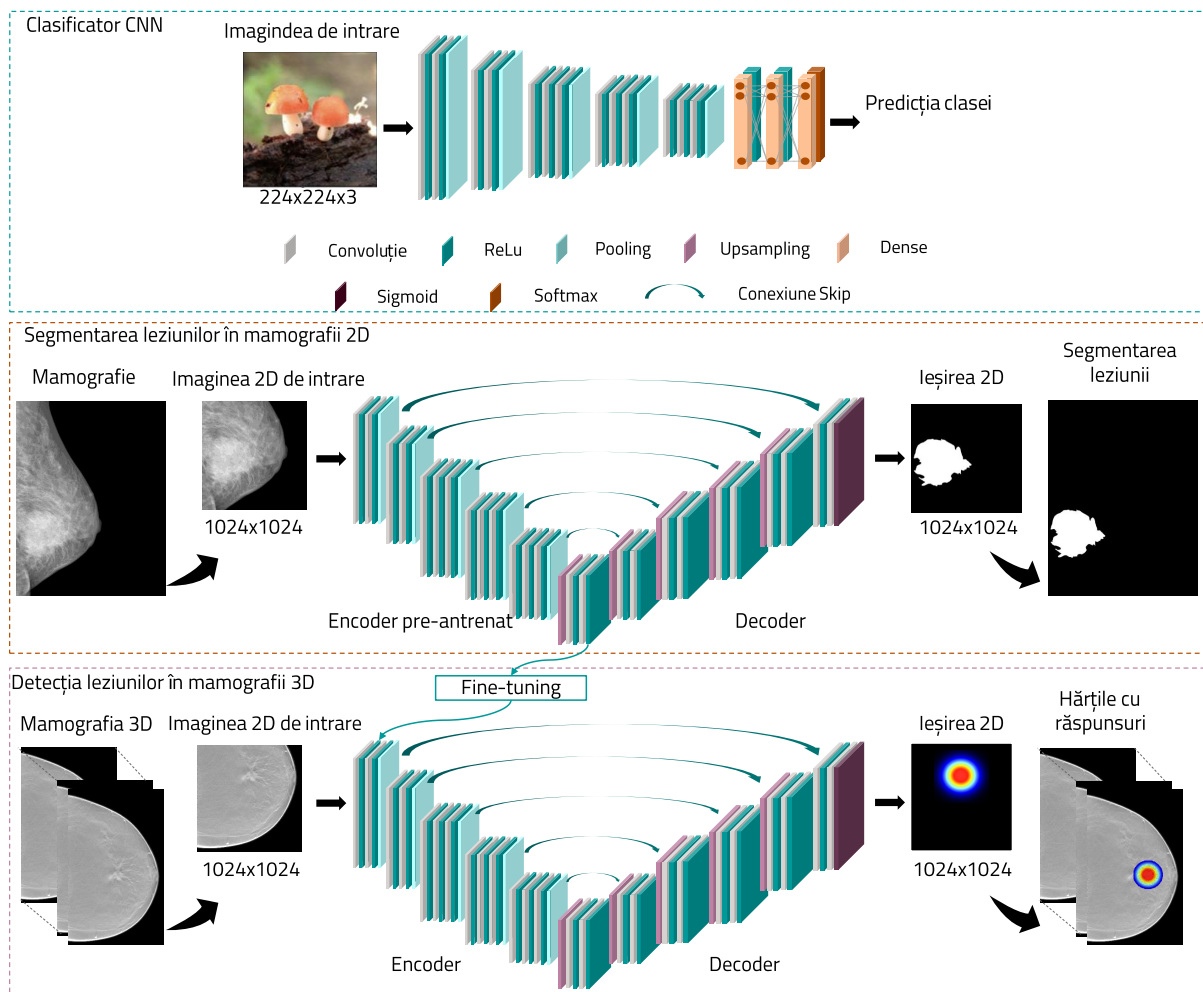
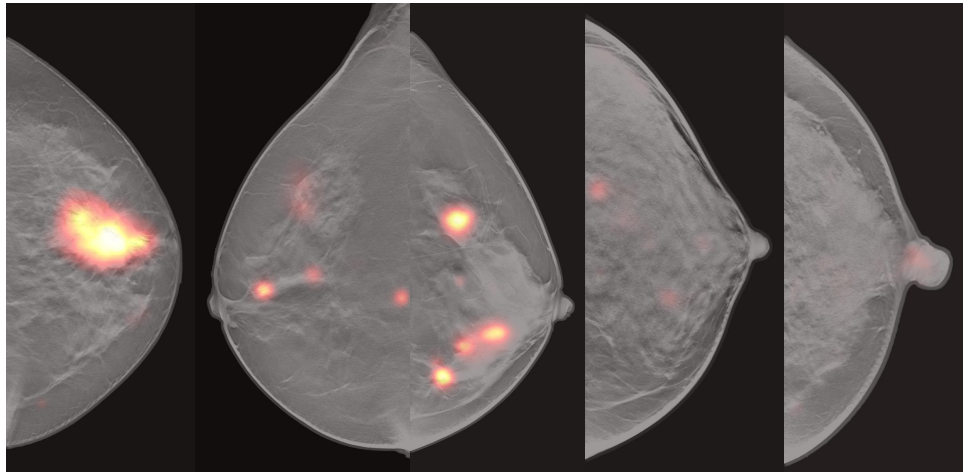


Figura 4.4: Prezentarea generală a procesului de antrenare al modelului Deep Learning propus. Parametrii straturilor de convoluție ale unui clasificador antrenat pe imagini din setul de date ImageNet sunt utilizați pentru a inițializa enoderul modelului propus. Ulterior, rețeaua este adaptată pentru a rezolva segmentarea leziunilor în mamografiile 2D convenționale. În final, rețeaua este ajustată pentru problema detectării leziunilor în mamografiile cu tomosinteză.

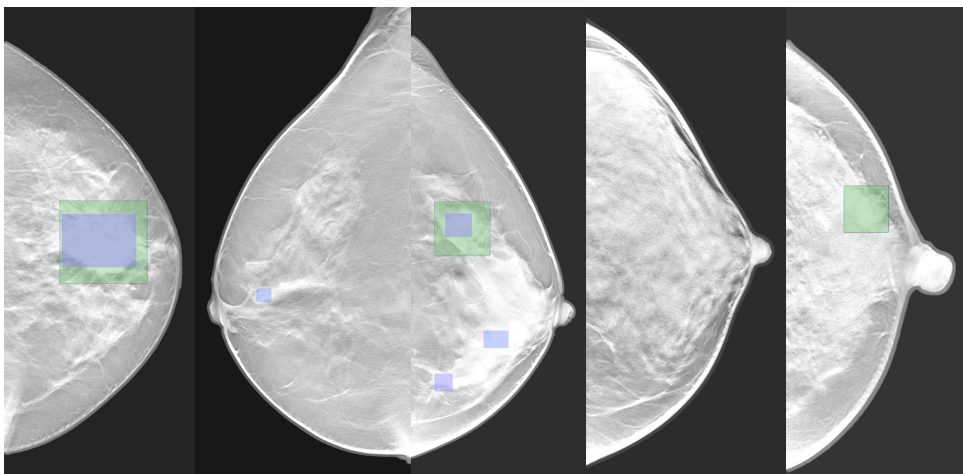
corespunzătoare unei leziuni, răspunsurile oferite de rețea pentru fiecare secțiune din mamografia cu tomosinteză sunt suprapuse. Pe volumul, astfel obținut, se aplică o metodă de post-procesare pentru localizarea maximelor funcției de densitate de probabilitate (mean-shift). În această etapă se obțin atât parametrii ce descriu dreptunghiurile de încadrare, cât și un scor de încredere al regiunii evidențiate ca fiind o leziune. Figura 4.5b arată exemple de rezultate ale detecției leziunilor după etapa de post-procesare. După cum se arată în aceste imagini, modelul propus de detectare a leziunilor este capabil să identifice leziuni de diferite dimensiuni și aspecte. Cu toate acestea, există și cazuri în care modelul nu a putut identifica corect leziunile, în special atunci când leziunile apar în zone cu țesut mamar dens.

Performanța modelului propus de detectare a leziunilor mamare este evaluată în raport cu metrica FROC (Free Response Operating Characteristic) la nivel de volum (mamografie 3D). FROC măsoară rata de pozitivi adevărați (True Positive Rate - TPR) în raport cu numărul de rezultate fals- pozitive per volum. În mod convențional, TPR descrie procentul leziunilor adnotate care au fost identificate corect de către model.

Curba FROC obținută pe setul de date de test, pentru un prag de detectare $T = 0.05$, este prezentată în Figura 4.6. Graficul indică o rată de detecție de 80% la 0.7 FP în medie pentru fiecare volum. Cu toate acestea, echilibrul dintre cele două cantități poate fi controlat prin modificarea pragului de



(a)



(b)

Figura 4.5: Exemple de hărți de răspuns precise de modelul Deep Learning (a) și (b) dreptunghiurile de încadrare corespunzătoare generate în etapa de post-procesare. Dreptunghiurile marcate cu culoare verde reprezintă dreptunghiurile adnotate de radiolog și cu albastru sunt descrise cele estimate. Fiecare imagine descrie proiecția de intensitate maximă în planul xy al mamografiei 3D.

decizie.

4.4 Corelarea leziunilor în proiecțiile mamografice

Pentru a spori fiabilitatea detecției leziunilor, radiologii combină informațiile obținute din proiecțiile CC și MLO ale unui sân. O evaluare cu două vizualizări mărește șansele ca o leziune să fie văzută în cel puțin una din proiecții, îmbunătățind de asemenea și rata de rechemare prin eliminarea detecțiilor false [27]. Cu toate acestea, din cauza compresiei și a informațiilor suplimentare aduse de natura volumetrică a mamografiilor cu tomosinteză, găsirea regiunilor corespunzătoare în cele două proiecții standard amplifică timpul necesar pentru examinare. În plus, există cazuri în care, din cauza geometriei achiziției, structurile văzute într-o proiecție pot fi vizibile doar parțial în cealaltă, după cum se arată și în Figura 4.7.

Deși o analiză bazată pe cele două proiecții poate mări performanța unui sistem de detecție asistată de calculator, după cum se arată în studiile realizate pe mamografii 2D [22], datorită lipsei de date complete, majoritatea sistemelor bazate pe învățare pentru detectarea leziunilor în mamografii 3D se bazează pe procesarea unei singure proiecții. O soluție alternativă care atenuază problema

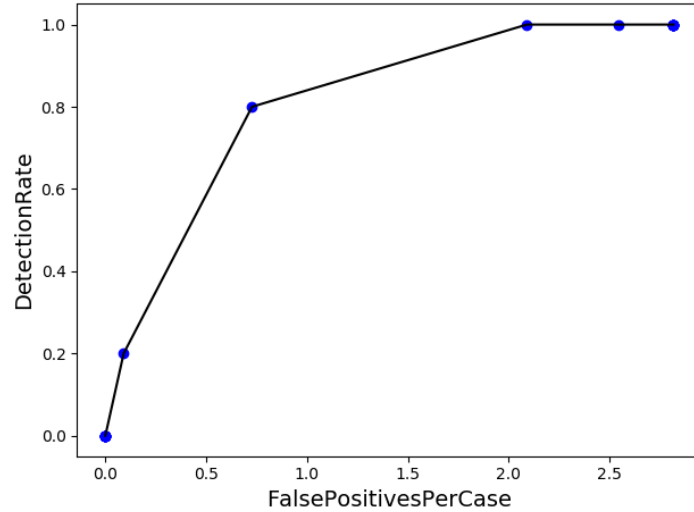


Figura 4.6: Curba FROC obținută pentru detecția leziunilor în mamografiile cu tomosinteză.

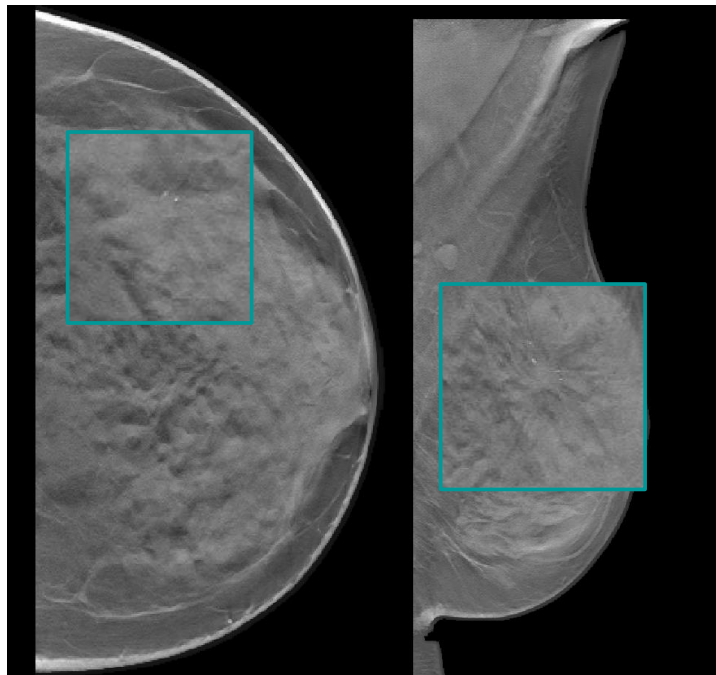


Figura 4.7: Exemplu de leziune evidențiată într-o mamografie cu tomosinteză: o secțiune extrasă din volumele CC și MLO care conține o leziune, marcată aici prin intermediul chenarului rectangular. Leziunea este clar vizibilă în vederea MLO (dreapta), dar numai parțial în CC (stânga). Regiunea este mărită pentru o mai bună vizualizare.

disponibilității datelor, menținând în același timp beneficiile aduse de existența acestor proiecții, este corelarea la nivel regional a celor două imagini.

4.4.1 Depistarea corespondențelor dintre proiecțiile mamografice

În acest sens, se propune o soluție de găsimă a corespondențelor dintre proiecțiile CC și MLO. Mai precis, prin combinarea informațiilor preluate din geometria mamografiilor cu tomosinteză cu distanța dintre mamelon și leziune, un model intermediar poate fi asamblat sub forma unui sistem de tip cadran de ceas (o'clock), așa cum se arată în Figura 4.8. Un astfel de sistem poate fi utilizat pentru

a găsi corespondențele dintre regiunile corespunzătoare din volumele CC și MLO fără a lua în considerare mecanismul de compresie din timpul achiziției sau proprietățile materiale ale sânului. Astfel, metoda este rapidă și intuitivă, având potențialul de a îmbunătăți detectabilitatea leziunilor.

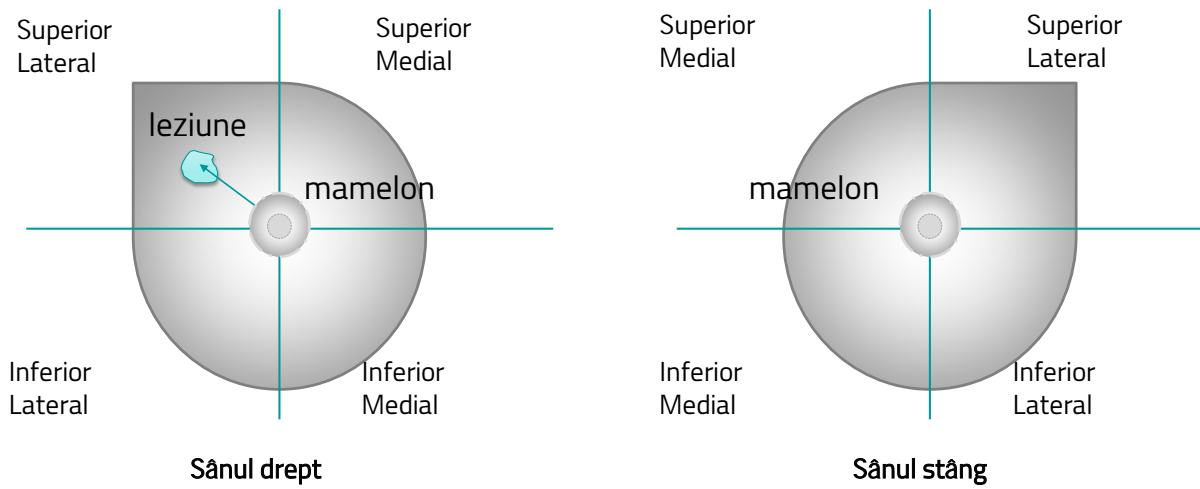


Figura 4.8: Împărțirea virtuală a sânelui stâng și drept în patru cadrane.

Dată fiind natura imagisticii mamografiei cu tomosinteză, simpla examinare a mamografiilor 3D achiziționate oferă o intuiție asupra poziției leziunii în sân. În combinație cu informațiile de „scrolling” prin volum, se poate stabili dacă o leziune proiectată în imaginea MLO este mediană sau laterală în sistemul o’clock sau dacă este superioară sau inferioară în cazul proiecției CC. De exemplu, știind că, în cazul imaginii MLO, fasciculul de raze X traversează sânul din partea interior-superioară către partea exterior-inferioară și măsurând distanța dintre poziția leziunii (adică, secțiunea bidimensională în care leziunea este vizibilă) și marginea sânelui, se poate estima poziția leziunii în planul vertical (lateralitatea). Astfel, prin examinarea secțiunii bidimensionale dintr-o proiecție MLO se poate determina locația leziunii în cadranul superior/inferior, în timp ce glisând prin volum, locația în cadranul medio/lateral poate fi dedusă. Aceeași ipoteză este valabilă și pentru proiecțiile CC, însă vice-versa. Figura 4.9 ilustrează procedura propusă de gărire a corespondențelor.

Astfel, dispunând de locația 3D a unei structuri într-una din proiecții, locația 3D corespunzătoare în proiecția adiacentă este determinată în două etape, după cum este prezentat în Figura 4.10. Pentru o locație dată într-o imagine, se determină locația 2D corespunzătoare (în planul xy) în a doua proiecție prin proiectarea informațiilor în sistemul o’clock și înapoi în imaginea țintă (Figura 4.9). Pentru a putea determina însă poziția 3D, sunt necesare informații suplimentare referitoare la secțiunea (axa z) pe care este localizată leziunea. Pe lângă informațiile 2D, sistemul o’clock oferă, de asemenea, o intuiție asupra cadranelor în care este localizată leziunea. Prin urmare, în a doua etapă, secțiunea exactă este dedusă printr-o analiză de similaritate structurală dintre subregiunea bidimensională extrasă din locația leziunii cunoscute și regiunile potențiale din volumul țintă.

Similaritatea între imagini este calculată prin metoda potrivirii cu un șablon (eng. template matching). Coeficientul de auto-corelație normalizată dintre regiunea leziunii localizată într-o proiecție (I_s) și o regiune propusă din proiecția adiacentă țintă (I_t) se determină după cum urmează:

$$R = \frac{\sum_{i,j} (I_s(i,j) - \bar{I}_s)(I_t(i,j) - \bar{I}_t)}{\sqrt{\sum_{i,j} (I_s(i,j) - \bar{I}_s)^2} \sqrt{\sum_{i,j} (I_t(i,j) - \bar{I}_t)^2}} \quad (4.4)$$

unde $\bar{I}_s = \frac{1}{NM} \sum_{i,j} I_s(i,j)$ și $\bar{I}_t = \frac{1}{NM} \sum_{i,j} I_t(i,j)$. Subregiunile I_s și I_t au aceeași dimensiune $N \times M$, cu (i,j) coordonatele pixelilor.

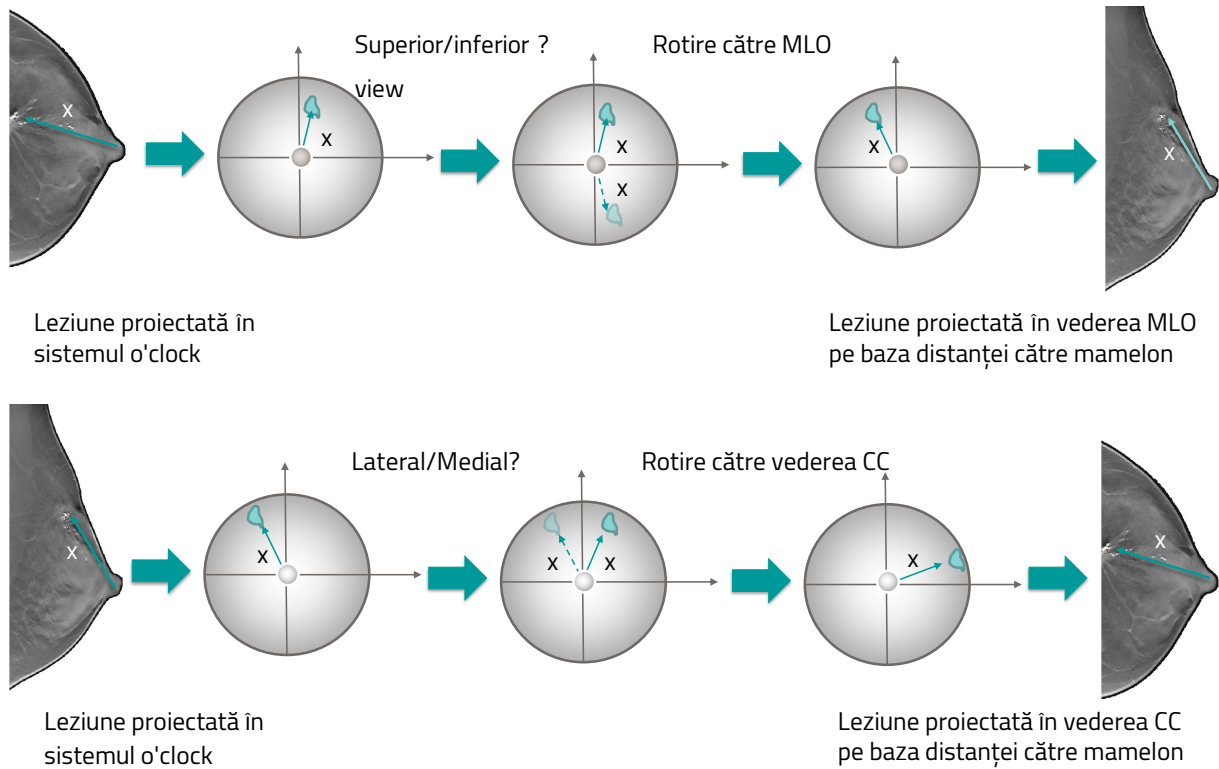


Figura 4.9: O prezentare generală a algoritmului de depistare a locației leziunilor în planul xy dintr-o proiecție mamografică în cealaltă.

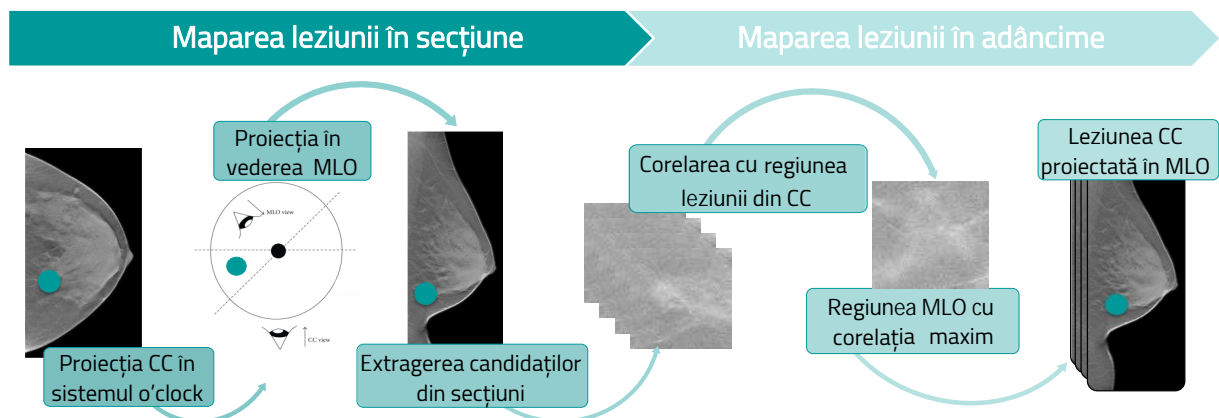
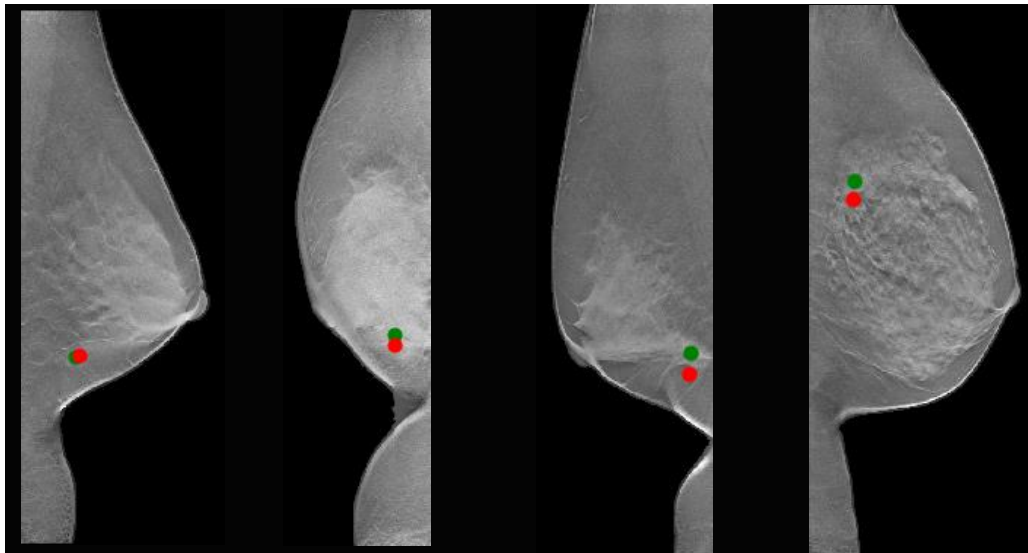
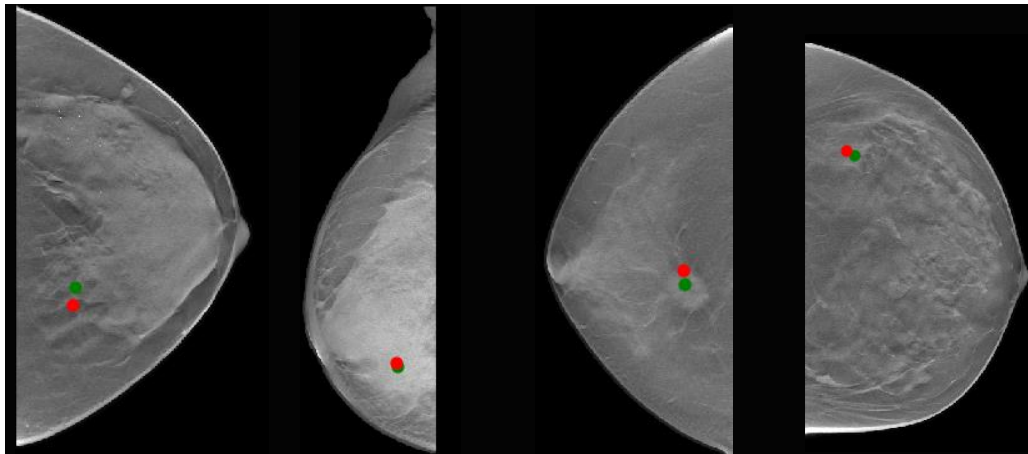


Figura 4.10: O prezentare generală a algoritmului propus pentru determinarea locației 3D a unei leziuni într-o proiecție pornind de la locația acesteia în proiecția adiacentă. În prima etapă, leziunea din vederea CC este proiectată în planul xy al vederii MLO, utilizând un criteriu de potrivire geometrică și sistemul de coordonate o'clock intermediar. În continuare, informația de adâncime este recuperată prin determinarea regiunii cu valoarea maximă a corelației dintre regiunile propuse și șablonul (regiunea leziunii extrasă din vederea CC).

Regiunea identificată ca având valoarea maximă a corelației determină informația de adâncime necesară pentru a determina poziția finală a leziunii.



(a)



(b)

Figura 4.11: Rezultatele corelării poziției leziunilor de la (a) CC la MLO și (b) MLO la CC pentru patru pacienți diferiți (reprezentați pe coloane). Locațiile leziunilor marcate cu culoarea verde reprezintă reperele definite de radiolog și cu roșu sunt descrise pozițiile determinate cu metoda propusă.

4.5 Rezultate

Soluția propusă pentru depistarea corespondențelor în proiecțiile mamografiilor cu tomosinteză a fost aplicată doar pe acele cazuri din setul de date pentru care au fost disponibile ambele proiecții (CC și MLO). Algoritmul a fost evaluat atât din punct de vedere al performanței în depistarea corespondențelor, cât și din cel al implicațiilor aduse integrării soluției în sistemul Deep Learning de detecție al leziunilor.

4.5.1 Corelarea leziunilor

Având poziția leziunii într-o proiecție mamografică cu tomosinteză, poziția corespunzătoare în a doua proiecție a fost estimată și comparată cu poziția cunoscută, adnotată. Pentru fiecare caz, maparea leziunilor a fost efectuată și evaluată în ambele direcții: CC la MLO și MLO la CC. Figura 4.11 cuprinde o comparație vizuală a rezultatelor obținute.

Considerând mapările leziunilor din ambele direcții, s-a constatat că eroarea medie pătratică a diferențelor dintre pozițiile 3D estimate a leziunilor și cele reale este de aproximativ 21.9 mm. Distanța

medie absolută de-a lungul fiecărei axe (Figura 4.12) ilustrează faptul că diferența dintre leziunile reale și cele corelate din planul axial a fost relativ mai mare în direcția y, comparativ cu direcția x.

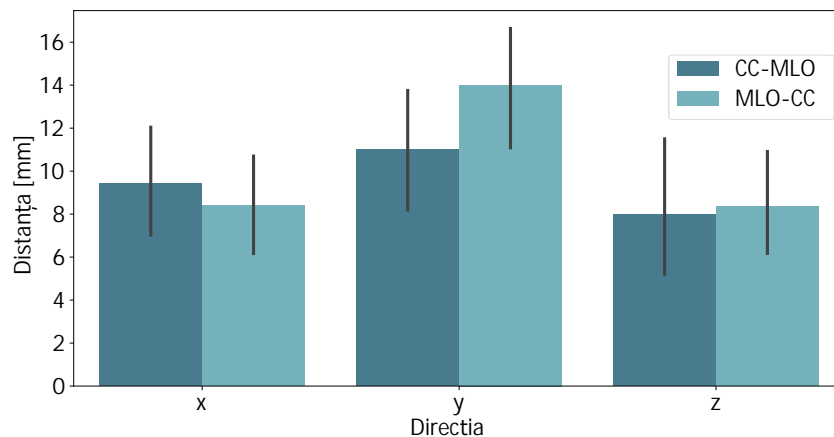


Figura 4.12: Graficul cu bare al distanței medii absolute dintre leziunile 3D reale și cele corelate de-a lungul axei X, Y și respectiv Z. Pentru fiecare caz, maparea a fost realizată din ambele direcții: CC-MLO și MLO-CC. Mamografiile cu tomosinteză analizate dispun de o rezoluție de $(0.34 \times 0.34 \times 1)$ mm.

4.5.2 Fuziunea proiecțiilor mamografice pentru detectarea leziunilor

Pentru a detecta leziunile mamare prin analiza concomitentă a celor două proiecții standard mamografice, modelul Deep Learning propus în secțiunea 4.2 a fost inițial aplicat independent pe volumele CC și MLO. În continuare, răspunsurile generate au fost utilizate pentru a prezice locațiile corespunzătoare în a doua proiecție, utilizând algoritmul propus pentru depistarea corelațiilor în două mamografii cu tomosinteză.

Pentru cazurile aflate în setul de testare în care atât vizualizările CC, cât și cele MLO au fost disponibile, performanța algoritmului de detectare al leziunii a fost ușor îmbunătățită prin fuziunea răspunsurilor rețelei neurale în cele două proiecții. Mai precis, rezultatele au arătat o creștere a specificității cu 13% în detrimentul reducerii cu 4% a sensibilității. Cu toate acestea, echilibrul dintre rata fals-positivă și rata real-positivă poate fi controlat prin modificarea pragului la care o constatare este considerată suspectă, după cum se arată în Figura 4.13.

4.6 Discuții generale și concluzii

În acest capitol s-a investigat problema detectării automate a leziunilor în mamografiile cu tomosinteză. În vederea proiectării unui sistem eficient, au fost investigate trei aspecte: (i) lipsa datelor, (ii) adnotările inexacte și (iii) reducerea rezultatelor fals-pozitive. Pentru a combate problemele introduse de utilizarea unui set redus de date pentru formularea soluției Deep Learning, a fost adoptată o strategie cunoscută sub numele de fine-tuning. Astfel, pentru a evita producerea fenomenului de overfitting, cunoștințele extrase dintr-o problemă de computer vision au fost încorporate în modelul neural. Ulterior, modelul a fost adaptat pentru problema segmentării de leziuni în mamografiile computerizate și în final pentru detecția leziunilor în mamografiile cu tomosinteză. Pentru a îmbunătăți și facilita identificarea leziunilor, problema a fost formulată într-o manieră care să reducă implicațiile introduse de adnotările inexacte. Cu toate acestea, datorită aspectului variabil al leziunilor, s-a identificat o rată ridicată de detecții fals-pozitive, problemă des întâlnită în rândul sistemelor de detecție a leziunilor mamare asistată de calculator. Pentru a reduce numărul de detecții false determinate de modelul Deep Learning, a fost propusă o metodă de post-procesare ce utilizează și fuzionează informațiile furnizate de rețea pentru cele două proiecții standard, și anume CC și MLO.

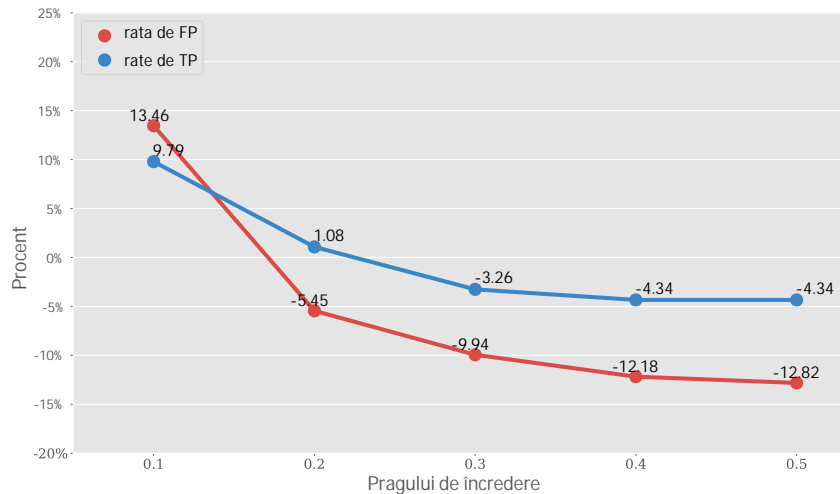


Figura 4.13: Influența pragului de încredere asupra ratei de TP și FP utilizat în corelarea locației între vizualizările standard mamografice. Pragul controlează valoarea de la care o constatare este considerată suspectă (certitudine ridicată). Modificarea acestor cantități este exprimată în raport cu performanțele obținute de sistemul de detecție aplicat independent per volum.

În concluzie, sistemul propus bazat pe Deep Learning a obținut rezultate promițătoare pentru problema dificilă a detectării leziunilor în mamografia 3D. În plus, analiza realizată pe două vederi a arătat potențialul, dar și necesitatea de a include informații suplimentare în procesul de detectare al leziunilor. Mai precis, această etapă de post-procesare a îmbunătățit rata de rechemare prin eliminarea detecțiilor false prin intermediul depistării corespondențelor între imagini.

5. Protecția datelor în aplicațiile medicale bazate pe învățarea automată

5.1 Introducere

Soluțiile de asistență medicală asistate de calculator care încorporează tehnici de învățare automată din date se bazează în primul rând pe datele existente, respectiv viitoare, ale pacientului pentru a furniza rezultate precise și fiabile [28]. Deoarece aplicațiile bazate pe modele Deep Learning necesită acces la datele personale ale pacienților, au fost ridicate o serie de probleme privind confidențialitatea și securitatea informațiilor personale [29]. În plus, reglementările adoptate în prezent cu privire la garanțiile de confidențialitate pentru manipularea datelor cu caracter personal (de ex. GDPR în UE, HIPAA în SUA) impun adoptarea unor tehnici mai eficiente de protecție a datelor.

Datorită dificultăților care apar în practică atunci când se utilizează modelele Deep Learning pe date criptate, precum și ineficienței soluțiilor actuale, se propune o soluție care să crească eficiența modelelor criptate pentru a asigura confidențialitatea datelor în aplicații reale. Metoda propusă permite: (i) calcule folosind numere raționale, (ii) operații mai rapide și (iii) rezultate comparabile cu cele obținute cu modele necriptate. Algoritmii de învățare bazați pe modele Deep Learning care păstrează confidențialitatea datelor sunt aplicați inițial pe o problemă clasică de recunoaștere a cifrelor din imagini pentru a valida metoda de criptare și, ulterior, pe două aplicații medicale personalizate. În timpul evaluării soluției propuse, s-a demonstrat că nu există pierderi de performanță atunci când rețelele funcționează pe date criptate folosind schema homomorfică de criptare MORE.

5.2 Stadiul actual

5.2.1 Protecția datelor în inteligența artificială

Au fost făcute mai multe încercări de a aborda problema păstrării confidențialității datelor în analiza bazată pe tehnici de inteligență artificială prin intermediul criptării homomorfe. Acest tip special de criptare permite manipularea directă a datelor criptate. Prin urmare, sistemul criptografic se concentrează pe păstrarea datelor private, permițând unei terțe părți să proceseze datele într-o formă criptată, fără a fi nevoie să dezvăluie informațiile reale. Păstrând structurile matematice care caracterizează datele, criptarea homomorfică reprezintă o soluție promițătoare pentru garantarea confidențialității, menținând totodată utilitatea deplină.

5.2.2 Criptarea homomorfică

Odată cu prima introducere a schemei de criptare homomorfică [30], în literatura de specialitate au fost propuse numeroase variații ale schemei inițiale [31]. Cele mai multe dintre aceste scheme criptografice sunt cunoscute pentru eficiența lor în ceea ce privește securitatea, dar necesită calcule intense și doar un număr limitat de operațiuni pot fi efectuate înainte ca decriptarea să nu mai fie posibilă. Acest lucru limitează în mod clar posibilitatea de utilizare în aplicații reale. Deși progresele recente în domeniul criptării homomorfe au dus la numeroase variante de sisteme de criptare, niciun sistem disponibil în prezent nu poate manipula direct numere raționale.

În consecință, o variantă a unei metode de criptare homomorfică matriceală, numită MORE (eng. Matrix Operation for Randomization or Encryption) [32] a fost adaptată pentru a sprijini direct aritmetica cu virgulă mobilă.

5.3 Sistemul criptografic MORE

Folosind schema de criptare MORE, o valoare numerică este criptată ca o matrice și prin calcul matriceal se obține comportamentul homomorfic, schema satisfăcând atât proprietățile de însumare, cât și de multiplicare. În consecință, toate operațiile efectuate pe date criptate (eng. ciphertext) devin matriceale, de exemplu, adunarea a două numere presupune, în spațiul criptat, adunarea celor două matrici rezultate după criptarea numerelor. Ordinul matricii de criptare reprezintă un parametru care controlează compensarea dintre securitate și eficiență: prin creșterea complexității schemei (și anume, ordinul matricii utilizate pentru criptarea unui mesaj) securitatea este îmbunătățită cu un cost de procesare mai lung. Schema de criptare poate fi direct generalizată la o matrice $n \times n$, cu toate acestea, pentru simplitate, în Tabela 5.1 este prezentată schema configurată pentru a funcționa cu matrici 2×2 .

Tabela 5.1: Configurarea schemei de criptare MORE pentru a funcționa cu numere raționale.

Mesajul	Marime scalară $m \in \mathbb{R}$
Generarea cheii secrete	Matrice inversabilă $S \in \mathbb{R}^{2 \times 2}$
Construcția matricii	$M = \begin{pmatrix} m & 0 \\ 0 & r \end{pmatrix}$, unde $r \in \mathbb{R}$ este o valoare aleasă arbitrar
Operația de criptare	$Criptare(m) = C = SMS^{-1}$
Operația de decriptare	$Decriptare(C) = K = (S^{-1}CS)$
Recuperarea mesajului	$m = K_{(1,1)}$

5.3.1 Efectuarea operațiilor pe date criptate

Schema de criptare MORE permite ca operațiile algebrice să fie efectuate pe matrici criptate. Astfel, date fiind două matrici criptate $C_1 = SM_1S^{-1}$ și $C_2 = SM_2S^{-1}$ pentru operația de înmulțire

$$C_1C_2 = SM_1S^{-1}SM_2S^{-1} = SM_1M_2S^{-1}, \quad (5.1)$$

care este echivalentul criptării M_1M_2 și pentru adunare

$$C_1 + C_2 = SM_1S^{-1} + SM_2S^{-1} = S(M_1 + M_2)S^{-1}. \quad (5.2)$$

În mod similar, această proprietate se aplică și pentru operațiile de scădere și împărțire, dar și pentru operațiile mixte cu valori necriptate, făcând schema complet homomorfică în raport cu operațiile algebrice. Un exemplu de aplicare al operațiilor matematice pe mesaje criptate este prezentat în Figura 5.1.

În aplicațiile din lumea reală, inclusiv în cele bazate pe învățarea automată din date, funcțiile neliniare (de exemplu, radical, exponențială, logaritmică etc.) sunt nelipsite. Atunci când un sistem de criptare este limitat la utilizarea exclusivă a operațiilor algebrice, abordarea tipică pentru a susține un spectru mai larg de operații implică un procedeu de aproximare a funcției neliniare printr-o funcție polinomială (de exemplu, seriile Taylor trunchiate). Schema MORE permite o abordare mai simplă pentru efectuarea unor astfel de operații. Un exemplu care demonstrează aplicarea funcției logaritmice pe date criptate cu schema MORE este prezentat în Figura 5.2. Algoritmul 5.1 arată cum, având un mesaj criptat $C \in \mathbb{R}^{2 \times 2}$, două metode pot fi utilizate pentru a formula funcția $f(x) = \frac{1}{1 + e^{-x}}$ definită pe $x \in \mathbb{R}$ în cazul utilizării schemei MORE. Această funcție este cunoscută sub numele de

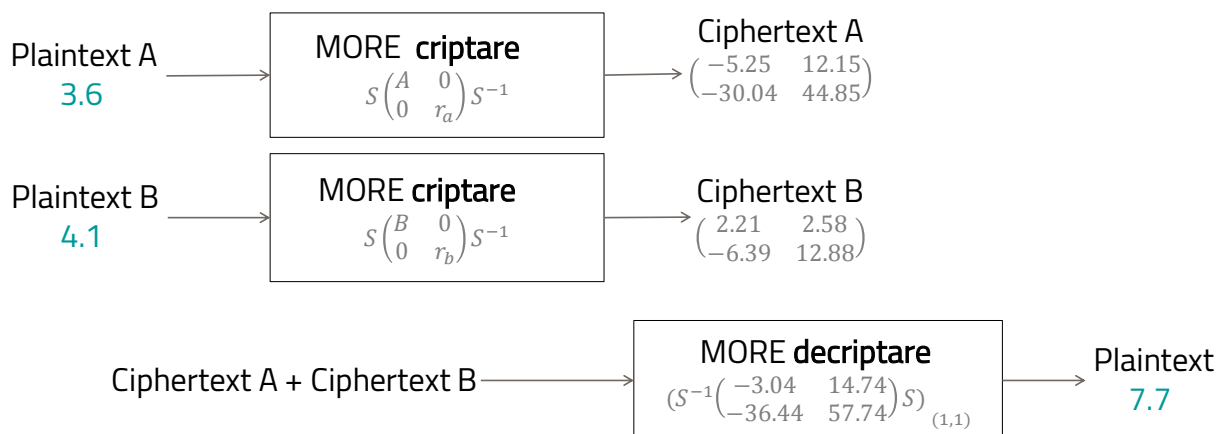


Figura 5.1: Operația de adunare în cazul schemei de criptare MORE.

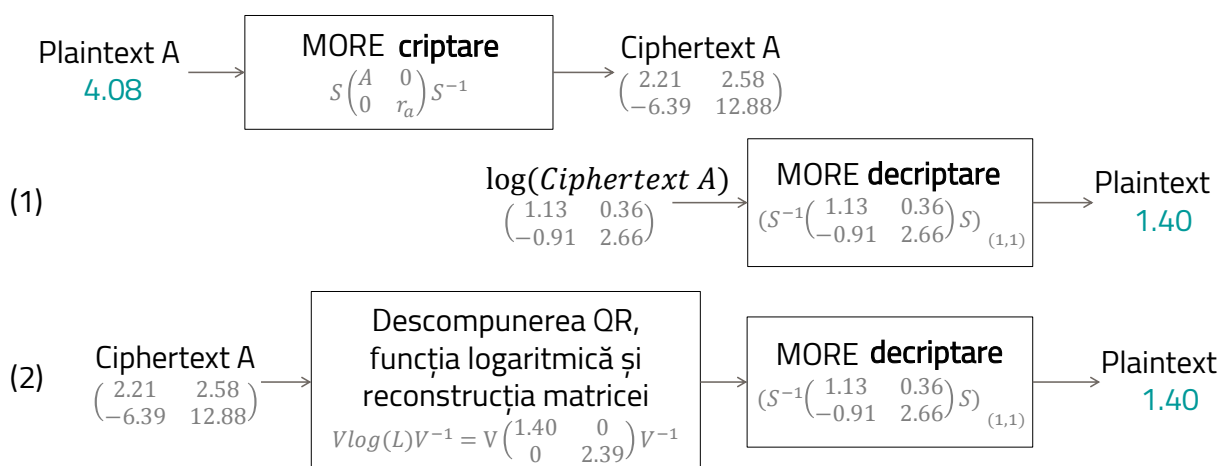


Figura 5.2: Exemplu de aplicare al funcției logaritmice în cazul schemei de criptare MORE. Două metode sunt propuse: (1) abordarea matriceală clasică și (2) bazată pe descompunerea QR.

funcția sigmoidă și este utilizată pe scară largă în rețelele neurale datorită proprietăților sale neliniare.

Algorithm 5.1 Implementarea funcției sigmoidă în cazul criptării MORE.

Input: Mesajul criptat $C \in \mathbb{R}^{2 \times 2}$

Output: Rezultatul criptat $R \in \mathbb{R}^{2 \times 2}$

```

1: function Sigmoid( $C$ )                                     // Utilizând direct operații matriceale
2:    $R \leftarrow I_2 \times (I_2 + \text{MatrixExp}(-C))^{-1}$       //  $I_2$  reprezintă matricea identitate
3:   return  $R$ 
4: end function
5: function Sigmoid( $C$ )                                     // Utilizând descompunerea QR
6:    $L, V \leftarrow \text{EigenDecomposition}(-C)$ 
7:    $L_f \leftarrow \text{Diag}(\text{Exp}(L))$ 
8:    $C_{exp} \leftarrow V \times L_f \times V^{-1}$ 
9:    $R \leftarrow I_2 \times (I_2 + C_{exp})^{-1}$ 
10:  return  $R$ 
11: end function

```

5.4 Aplicarea modelelor Deep Learning pe date criptate homomorfic

În această secțiune sunt descrise aspecte ale rețelelor neuronale care păstrează confidențialitatea. Metoda propusă se bazează pe schema de criptare MORE și permite atât antrenarea cât și exploatarea rețelelor neuronale direct pe datele criptate homomorfic.

5.4.1 Metoda

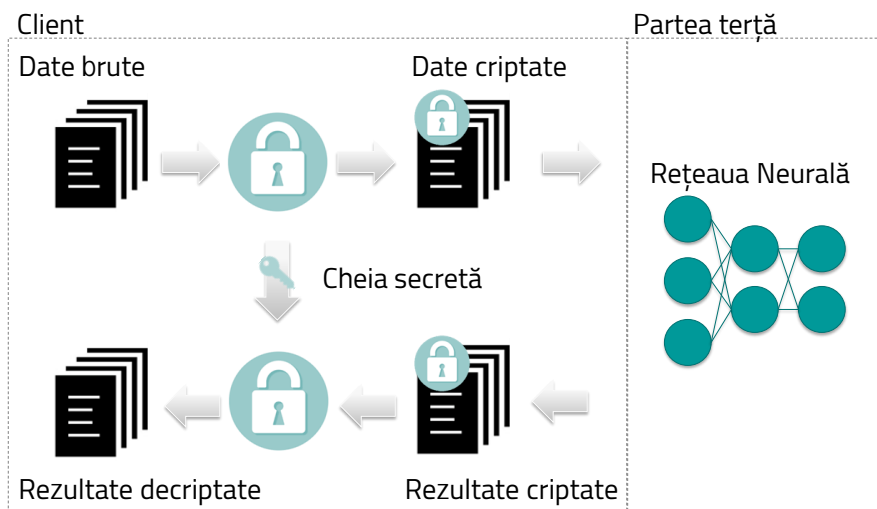


Figura 5.3: Fluxul de lucru al aplicației propuse bazate pe modele Deep Learning care păstrează confidențialitatea datelor prin intermediul criptării homomorfice.

Fluxul de lucru propus bazat pe criptarea homomorfică și Deep Learning este prezentat în Figura 5.3. Înainte de a fi procesate, datele de antrenare sunt criptate cu o cheie secretă simetrică care nu este niciodată partajată (Algoritmul 5.3). Modelul Deep Learning va avea acces numai la versiunea criptată a datelor, în timp ce datele reale sunt detașate de unitatea de procesare și rămân private la furnizorul de date. În cele din urmă, datorită proprietății homomorfe care stă la baza schemei de criptare MORE, a suportul direct pentru aritmetica în virgulă mobilă, și cu toate calculele efectuate în interiorul rețelei formulate pentru a fi aplicate pe date criptate, o rețea neurală poate fi antrenată direct pe date criptate folosind algoritmi clasici de antrenare. În consecință, rezultă un model criptat care furnizează predicții criptate, descifrabile doar de proprietarul cheii secrete în urma aplicării algoritmului 5.4 de decriptare. Odată ce faza de antrenare este finalizată, forma criptată a unui model poate fi utilizată pentru a realiza noi predicții pe date criptate (faza de inferență), unde datele de intrare sunt criptate cu aceeași cheie folosită în faza de antrenare. Criptosistemul MORE se bazează pe o cheie simetrică. Prin urmare, o cheie secretă, generată utilizând Algoritmul 5.2, este folosită atât pentru criptarea datelor cât și pentru decriptarea rezultatelor.

Soluția propusă de analiză a datelor criptate pe baza modelelor Deep Learning este descrisă în Algoritmul 5.6. Cu scopul de a compara și valida soluția propusă, varianta clasică, necriptată este furnizată în Algoritmul 5.5. A se reține că în Algoritmul 5.6 toate operațiile efectuate în timpul procesului de antrenare și predicție sunt formulate în conformitate cu Secțiunile 5.3 și 5.3.1.

În urma acestei abordări, confidențialitatea datelor este păstrată la trei nivele: (i) în timpul antrenării, când partea externă procesează direct datele criptate, (ii) în timpul inferenței, când datele pacientului rămân confidențiale (algoritmul primește date criptate ca intrare și generează rezultate criptate) și (iii) modelul neural al părții externe rămâne confidențial. În consecință, manipularea în siguranță a datelor medicale este realizată astfel încât partea externă să nu acceseze direct datele personale ale pacienților, iar utilizatorul să nu poată obține informații cu privire la modelul neural.

Algorithm 5.2 Generarea cheii secrete

Output: Cheia secretă $S \in \mathbb{R}^{2 \times 2}$

```
1: function KeyGeneration()
2:   while True do
3:      $S \leftarrow \text{RandomUniform}(size = (2, 2), minval, maxval)$ 
4:     if  $\det(S) \neq 0$  then // Asigură inversabilitatea matricială
5:       break
6:     end if
7:   end while
8:   return  $S$ 
9: end function
```

Algorithm 5.3 Criptarea MORE

Input: Mesajul $m \in \mathbb{R}$ **Input:** Cheia secretă $S \in \mathbb{R}^{2 \times 2}$ **Output:** Mesajul criptat $C \in \mathbb{R}^{2 \times 2}$

```
1: function Encryption( $m, S$ )
2:    $M \in \mathbb{R}^{2 \times 2} \leftarrow$  matrice nulă
3:    $M(0, 0) \leftarrow m$ 
4:    $M(1, 1) \leftarrow \text{RandomUniform}(minval, maxval)$ 
5:    $C \leftarrow S \times M \times S^{-1}$ 
6:   return  $C$ 
7: end function
```

Algorithm 5.4 Decriptarea MORE

Input: Mesajul criptat $C \in \mathbb{R}^{2 \times 2}$ **Input:** Cheia secretă $S \in \mathbb{R}^{2 \times 2}$ **Output:** Mesajul $m \in \mathbb{R}$

```
1: function Decryption( $C, S$ )
2:    $K \leftarrow S^{-1} \times C \times S$ 
3:    $m \leftarrow K(0, 0)$ 
4:   return  $m$ 
5: end function
```

Algorithm 5.5 Analiza pe date necriptate bazată pe modele Deep Learning.

```
1: function TrainOnPlaintext()
2:    $X_{train}, Y_{train} \leftarrow \text{LoadDataset}()$ 
3:    $X_{train} \leftarrow \text{Normalize}(X_{train})$ 
4:   BuildModel()
5:   Train( $X_{train}, Y_{train}$ )
6:   return model
7: end function

8: function PredictOnPlaintext()
9:    $X_{test} \leftarrow \text{LoadSamples}()$ 
10:   $X_{test} \leftarrow \text{Normalize}(X_{test})$ 
11:  LoadModel()
12:   $\tilde{Y}_{test} \leftarrow \text{Predict}(X_{test})$ 
13:  return  $\tilde{Y}_{test}$ 
14: end function
```

Algorithm 5.6 Analiza pe date criptate bazată pe modele deep.

```
1: function TrainOnCiphertext()
2:    $X_{train}, Y_{train} \leftarrow \text{LoadDataset}()$ 
3:    $X_{train} \leftarrow \text{Normalize}(X_{train})$ 
4:    $S \leftarrow \text{KeyGeneration}()$ 
5:    $X_{train_{enc}} \leftarrow \text{Encryption}(X_{train}, S)$ 
6:    $Y_{train_{enc}} \leftarrow \text{Encryption}(Y_{train}, S)$ 
7:   BuildModel()
8:   Train( $X_{train_{enc}}, Y_{train_{enc}}$ )
9:   return  $model_{enc}$ 
10: end function

11: function PredictOnCiphertext()
12:   $X_{test} \leftarrow \text{LoadSamples}()$ 
13:   $X_{test} \leftarrow \text{Normalize}(X_{test})$ 
14:   $S \leftarrow \text{LoadKey}()$ 
15:   $X_{test_{enc}} \leftarrow \text{Encryption}(X_{test}, S)$ 
16:  LoadModel()
17:   $\tilde{Y}_{test_{enc}} \leftarrow \text{Predict}(X_{test_{enc}})$ 
18:   $\tilde{Y}_{test} \leftarrow \text{Decryption}(\tilde{Y}_{test_{enc}}, S)$ 
19:  return  $\tilde{Y}_{test}$ 
20: end function
```

5.5 Partea experimentală

Pentru validarea metodei propuse au fost analizate trei tipuri de aplicații Deep Learning: regresia, clasificarea binară și multiplă. În primul rând s-a fost considerat problema clasică de recunoaștere a cifrelor în imagini iar mai apoi confidențialitatea datelor a fost analizată în două aplicații de asistență medicală, prin antrenarea modelelor Deep Learning pe datele criptate, (i) pentru a evalua hemodinamica întregului corp și (ii) pentru a distinge între arterele coronare (stângă și dreaptă) în angiografii.

5.5.1 Formularea problemelor

5.5.1.1 MNIST: Setul de date tipic pentru rețelele neuronale

O problemă tipică studiată în contextul rețelelor neuronale este cea a clasificării. Mai precis, problema clasificării imaginilor în funcție de informațiile conținute. Baza de date MNIST (Modified National Institute of Standards and Technology) [33] conține imagini reprezentând cifre scrise de mână și este de obicei folosită ca referință pentru algoritmi de clasificare a imaginilor (Figura 5.4). Problema de recunoaștere a cifrelor este formulată ca prezicerea probabilității unei imagini să aparțină fiecăreia dintre cele 10 clase (0-9 cifre).

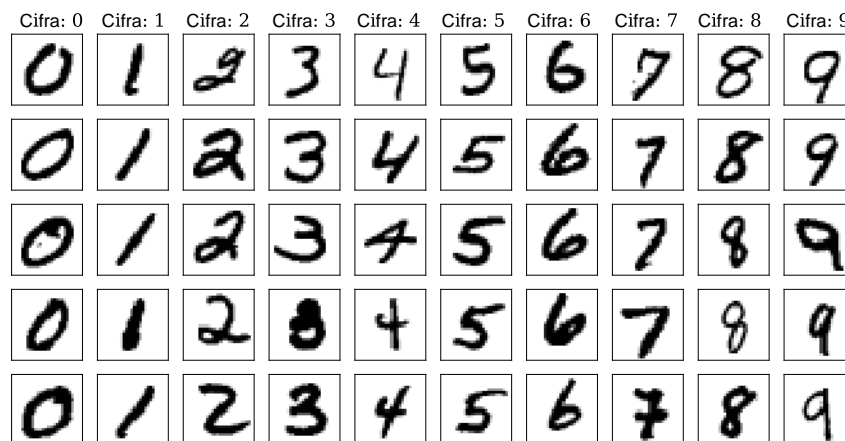


Figura 5.4: Exemple de imagini din setul de date MNIST.

5.5.1.2 Modelul de ordin redus a circulației arteriale

Pentru a demonstra fezabilitatea soluției propuse în cadrul unei aplicații de medicină personalizată, a fost ales un model hemodinamic al sistemului cardiovascular. Mai exact, un model de ordin redus a circulației arteriale (eng. whole body circulation - WBC). Modelul WBC folosit se bazează pe analogia dintre hidraulică și electricitate și este prezentat în Figura 5.5. Modelul cu buclă închisă al sistemului cardiovascular considerat conține modelul inimii (ventriculul stâng (LV) și atrium, ventriculul drept și atrium, valvele), circulația sistemică (artere, capilare, vene) și circulația pulmonară (artere, capilare, vene) [34].

Modelul WBC poate fi rulat în condiții specifice pacientului pentru a determina diverse mărimi de interes relevante clinic: rezistența arterială, complianța arterială, volumul mort al ventriculului stânga/drept, elastanța ventriculară/atrială/arterială, etc. Cu toate acestea, parametrii modelului trebuie personalizați pentru a se potrivi condițiilor și stării specifice pacientului.

Metoda de personalizare adoptată a fost descrisă anterior în detaliu [35] și constă din două etape secvențiale. În primul rând, o serie de parametri sunt calculați direct, iar apoi, o metodă de calibrare complet automată bazată pe algoritmi de optimizare estimează valorile parametrilor rămași, asigurându-se corespondența dintre calculele personalizate și măsurători.

Parametrii de intrare specifici pacientului sunt:

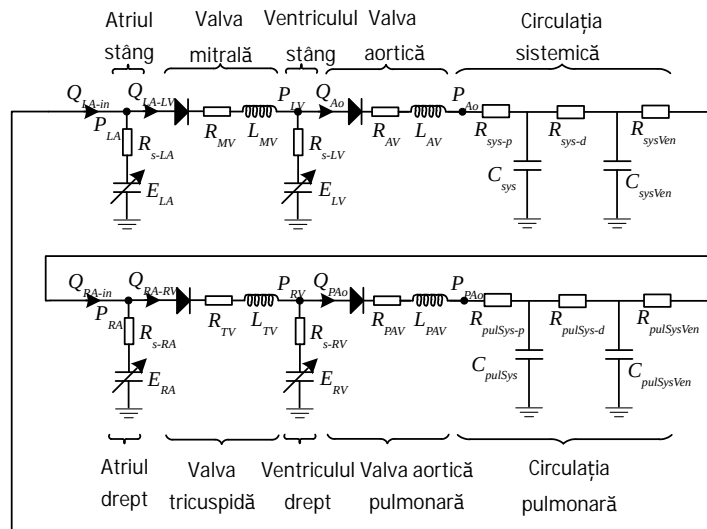


Figura 5.5: Modelul cu buclă închisă al sistemului cardiovascular.

- Circulația sistemică: presiunea sistolică aortică de vârf, presiune aortică la finalul diastolei, volumele ventriculului stâng la finalul sistolei și la finalul diastolei, volumul diastolic final, timpul de ejecție a ventriculului stâng
- Circulația pulmonară: presiunea sistolică de vârf a arterei pulmonare, presiunea arterială pulmonară la final de diastolă, volumele ventriculului drept la finalul sistolei și diastolei, timpul de ejecție a ventriculului drept

Mărimile personalizate de interes determinate după executarea personalizării sunt:

- Circulația sistemică: volumul mort al ventriculului stâng, timpul la elanța ventriculară stângă maximă, rezistența sistemică, complianța sistemică, raportul dintre rezistența proximală și distală a circulației sistemice
- Circulația pulmonară: volumul mort al ventriculului drept, timpul la elanța ventriculară dreaptă maximă, rezistența pulmonară, complianța pulmonară, raportul dintre rezistența proximală și distală a circulației pulmonare

În timp ce modelul parametric este foarte eficient, procesul de personalizare necesită un timp de calcul general de 30 - 60 de secunde pentru a determina mărimile de interes specifice pacientului. Astfel, un model capabil să transmită în timp real parametrii, care altfel ar fi determinați prin simulări ale modelului WBC, ar fi un instrument util.

În contextul rețelelor neuronale, această problemă este formulată ca prezicerea unor valori numerice reale în virgulă mobilă dintr-un set de parametri de intrare.

5.5.1.3 Angiografiile coronariene cu raze X

Angiografia coronariană cu raze X invaziv (ICA) este o procedură imagistică de diagnostic care oferă informații importante despre structura și funcția inimii și constituie baza imagisticii afecțiunilor coronariene [36]. ICA permite evaluarea severității anatomice a stenozelor coronariene, fie vizual, fie prin intermediul unei angiografii coronariene cantitative asistate de calculator [37]. Angiografiile coronariene sunt înregistrate separat și secvențial pentru artera coronară dreaptă (RCA) și cea stângă (LCA) (Figura 5.6).

Un domeniu important de cercetare în afecțiunile coronariene este post-procesarea complet automatizată a angiografiilor coronariene [38], cu următoarele obiective:

- Evaluare anatomică: aprecierea automată a severității stenozelor.
- Evaluare funcțională non-invazivă: calcularea automată a indicilor de diagnosticare funcțională [39], [40].
- Raportare: redactarea automată a rapoartelor medicale pe baza constatărilor din angiografiile coronariene.

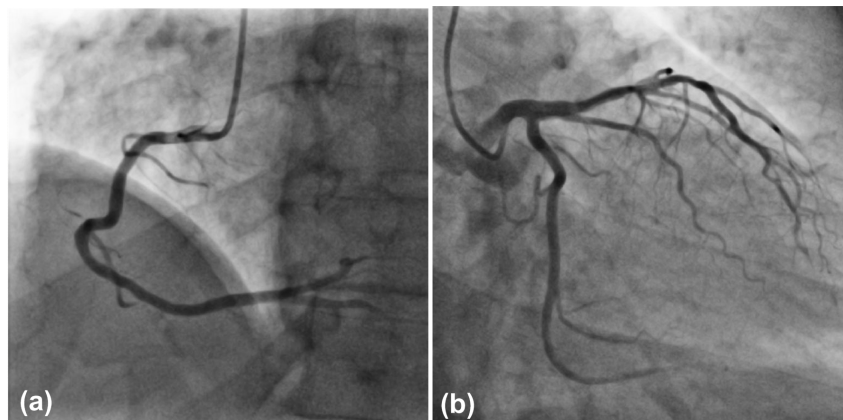


Figura 5.6: (a) Artera coronară dreaptă (b) Artera coronară stângă.

În multe aplicații clinice bazate pe utilizarea ICA, clasificarea automată a arterelor LCA / RCA reprezintă o etapă importantă de pre-procesare.

Recunoașterea automată a arterelor în angiografiile coronariene cu raze X poate fi formulată ca o problemă de clasificare binară, unde o rețea neurală învățată să prezică probabilitatea unei imagini să aparțină clasei pozitive (reprezentată de valoarea 1).

5.5.2 Pregătirea bazei de date criptate

Pentru a evalua performanța metodei propuse, au fost utilizate trei baze de date: imagini cu cifre scrise de mână (MNIST), angiografiile coronariene cu raze X și datele generate sintetic pentru modelul WBC. O prezentare generală a acestor baze de date este realizată în Tabela 5.2.

Tabela 5.2: Prezentarea generală a bazelor de date utilizate pentru evaluarea experimentală a rețelelor neuronale care păstrează confidențialitatea datelor.

	MNIST	WBC	Angio
Numărul de instanțe din setul de antrenare	50000	7000	1996 (7984*)
Numărul de instanțe din setul de validare	10000	1000	680
Numărul de instanțe din setul de testare	10000	2000	702

*după augmentare

Pentru a adresa provocările aplicării calculului pe date criptate homomorfic și pentru a evalua utilizarea modelelor Deep Learning pe date criptate, pentru fiecare set de date, mărimile de intrare (imagini sau vectori de caracteristici), au fost criptate utilizând strategia MORE, așa cum este descris în Algoritmul 5.3. Reprezentarea criptată a unei imagini extrase din baza de date MNIST este prezentată în Figura 5.7. În mod similar, valorile țintă, adică etichetele de clasă sau valorile reale, au fost, de asemenea, criptate.

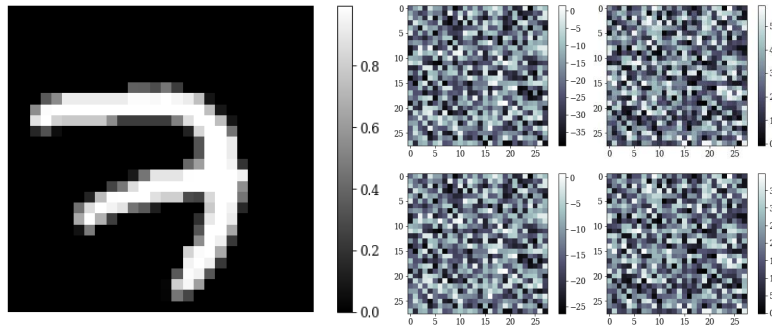


Figura 5.7: Imaginea unei cifre cu pixeli în interval $p \in [0, 1]$ afișată alături de componentele mesajului criptat.

5.5.3 Arhitectura modelelor Deep Learning

Pentru a evalua fezabilitatea și eficiența rețelelor neurale de a opera direct pe date criptate homomorfic, au fost realizate trei experimente prin antrenarea: (i) unei rețele convoluționale pentru recunoașterea cifrelor în imagini criptate, (ii) unei rețele neurale pentru analiza hemodinamică în timp real și (iii) unei rețele neurale pentru analiza angiografiilor coronare cu raze X criptate. Pentru o comparație a performanței și a convergenței modelelor, antrenarea s-a realizat, de asemenea, pe seturile de date echivalente necriptate.

5.5.3.1 Rețeaua neurală pentru clasificarea imaginilor ce conțin cifre scrise de mână

Pornind de la cele mai recente rezultate obținute de modelele convoluționale pentru recunoașterea cifrelor MNIST, un model convoluțional a fost utilizat pe perechi de date intrare-ieșire criptate. Topologia modelului propus pentru păstrarea confidențialității datelor este descrisă în Tabela 5.3.

Tabela 5.3: CNN-MNIST: Topologia rețelei convoluționale propuse pentru clasificarea cifrelor scrise de mână.

Straturi	Parametrii	Dimensiuni
Intrare	-	(1,28,28)
Convoluție	(8,3,3)	(8,28,28)
Activare (Sigmoid)	-	-
Average Pooling	(2,2)	(8,14,14)
Convoluție	(16,3,3)	(16,14,14)
Activare (Sigmoid)	-	-
Average Pooling	(2,2)	(16,7,7)
Flatten	-	(784,)
Fully Connected	100	(100,)
Activare (Sigmoid)	-	-
Fully Connected	10	(10,)
Activare (Softmax)	-	-

5.5.3.2 Rețeaua neurală pentru analiza hemodinamică în timp real

Dată fiind natura datelor de intrare (informațiile reprezentate ca un vector de caracteristici) și nevoia de a modela decizia rețelei pe baza unei dependențe globale între caracteristicile de intrare,

a fost utilizată o rețea neurală complet conectată cu 3 straturi ascunse. Topologia rețelei neurale propuse este descrisă în Tabela 5.4.

Tabela 5.4: FCNN-WBC: Topologia rețelei neurale propuse pentru analiza hemodinamică.

Straturi	Parametrii	Dimensiuni
Intrare	-	(9,)
Fully connected	40	(40,)
Activare (Tanh)	-	-
Fully connected	40	(40,)
Activare (Tanh)	-	-
Fully connected	40	(40,)
Activare (Sigmoid)	-	-
Fully connected	12	(12,)
Activare (Linear)	-	-

5.5.3.3 Rețeaua neurală pentru analiza angiografiilor coronariene cu raze X

Motivat de cele mai recente rezultate în analiza imaginilor pe baza modelelor neurale, a fost adoptată o rețea convoluțională pentru a rezolva problema de recunoaștere a angiografiei coronariene. Topologia rețelei convoluționale propuse este descrisă în Tabela 5.5.

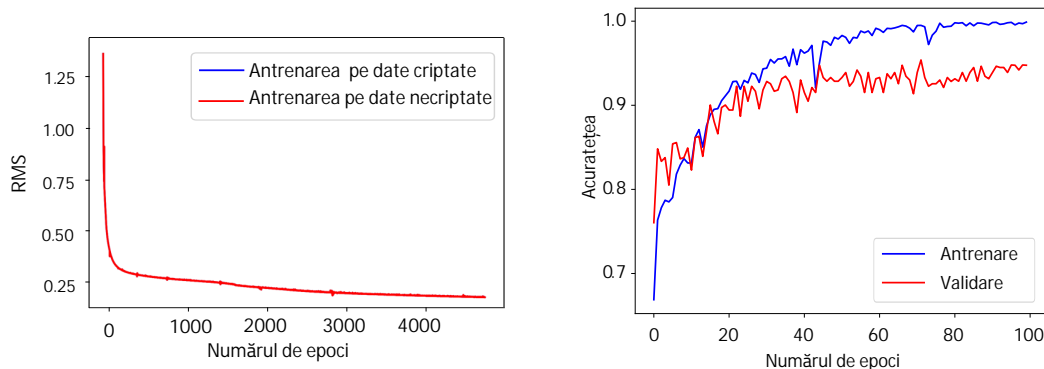
Tabela 5.5: CNN-Angio: Topologia rețelei convoluționale propuse clasificarea angiografiilor coronariene.

Straturi	Parametrii	Dimensiuni
Intrare	-	(1,256,256)
Convoluție	(4,3,3)	(4,256,256)
Activare (Sigmoid)	-	-
Average Pooling	(2,2)	(4,128,128)
Convoluție	(8,3,3)	(8,128,128)
Activare (Tanh)	-	-
Average Pooling	(2,2)	(8,64,64)
Convoluție	(16,3,3)	(16,64,64)
Activare (Tanh)	-	-
Average Pooling	(2,2)	(16,32,32)
Convoluție	(32,3,3)	(32,32,32)
Activare (Tanh)	-	-
Average Pooling	(2,2)	(32,16,16)
Flatten	-	(8192,)
Fully connected	64	(64,)
Activation (Tanh)	-	-
Dropout	25%	-
Fully connected	-	(1,)
Activare (Sigmoid)	-	-

5.6 Rezultate

5.6.1 Performanța

Pentru a analiza potențialul rețelelor de a învăța direct din datele criptate, evoluția erorii în timpul antrenării pentru problema analizei hemodinamice, după decriptare, este prezentată în Figura 5.8a. În mod similar, evoluția performanței de clasificare pe parcursul antrenării și validării modelului convoluțional pe angiografii coronare cu raze X criptate, obținută după decriptare, este prezentată în Figura 5.8b.



(a) Modele neurale pentru estimarea parametrilor modelului WBC (b) Modelul neural pentru clasificarea angiografiilor coronariene

Figura 5.8: (a) Evoluția erorii de antrenare pentru rețeaua criptată și necriptată: diferențele dintre curbele de învățare, cauzate de aritmetica cu virgulă mobilă, sunt insesizabile. (b) Evoluția performanței de clasificare pe parcursul antrenării cu date criptate.

Evoluția antrenării demonstrează capacitatea metodei propuse de a păstra corectitudinea calculelor. Mai mult, după decriptare, parametrii învățați de un model atunci când a fost antrenat pe date criptate s-au dovedit a fi identici, până la precizia mașinii, cu cei învățați de modelul necriptat.

5.6.1.1 Clasificarea imaginilor MNIST

Metrica considerată pentru evaluarea performanței unui clasificator pe setul de date MNIST este acuratețea absolută, și anume procentul de imagini clasificate corect. Rețeaua necriptată a condus la o acuratețe de clasificare de 98,2 % pe setul de date de testare, similar cu rețeaua criptată.

Precizia, rechemarea și scorul F1 pentru fiecare cifră sunt raportate în Tabela 5.6. De asemenea, matricea de confuzie a fost calculată și este afișată în Figura 5.9.

5.6.1.2 Analiza hemodinamică

Pentru a evalua capacitatea rețelei neurale de a estima ieșirile modelului de ordin redus a circulației arteriale, a fost calculată eroarea relativă medie absolută și corelația Pearson, iar rezultatele sunt afișate în Tabela 5.7. Graficele de dispersie ale parametrilor măsurați comparativ cu cei estimați, având cel mai mare și cel mai mic coeficient de corelație, sunt prezentate în Figura 5.10. Prima diagramă afișează rezultatele rețelei neurale obținute pentru estimarea raportului dintre rezistența proximală și distală în circulația sistemică. A doua diagramă prezintă rezultatele prezicerii rezistenței sistemice.

5.6.1.3 Clasificarea angiografiilor coronariene

Pentru a evalua precizia modelului de recunoaștere a angiografiilor coronariene, curba ROC (eng. receiver operator characteristic) este afișată în Figura 5.12. Tabela 5.8 conține precizia, rechemarea și

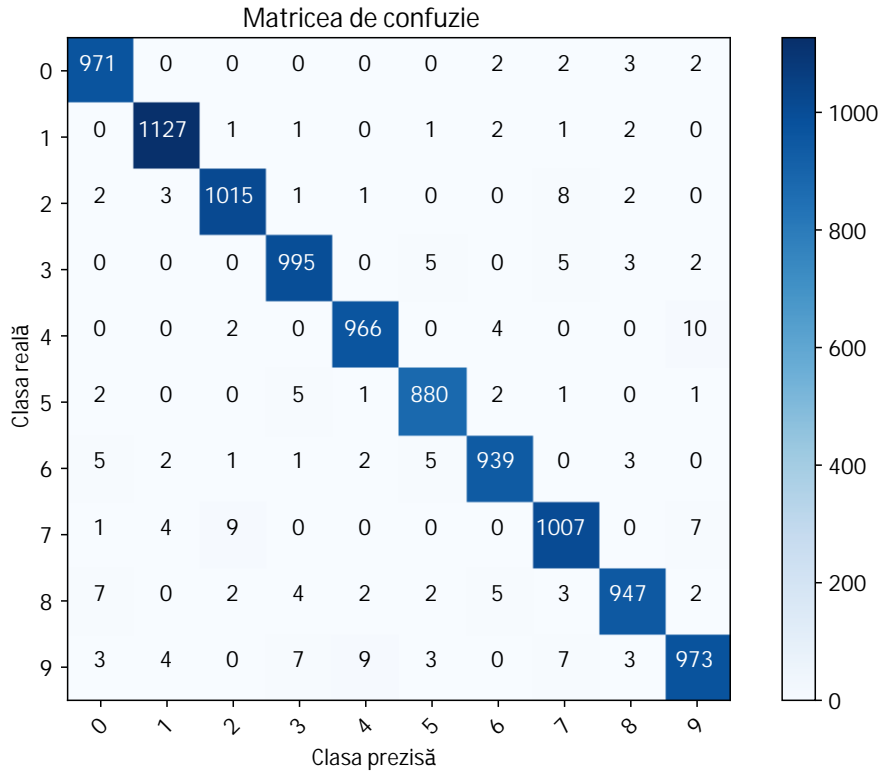


Figura 5.9: Matricea de confuzie obținută de clasificatorul propus pe imaginile MNIST din setul de teste. Numărul de pe diagonală indică numărul de imagini clasificate corect, în timp ce restul reprezintă cele clasificate greșit.

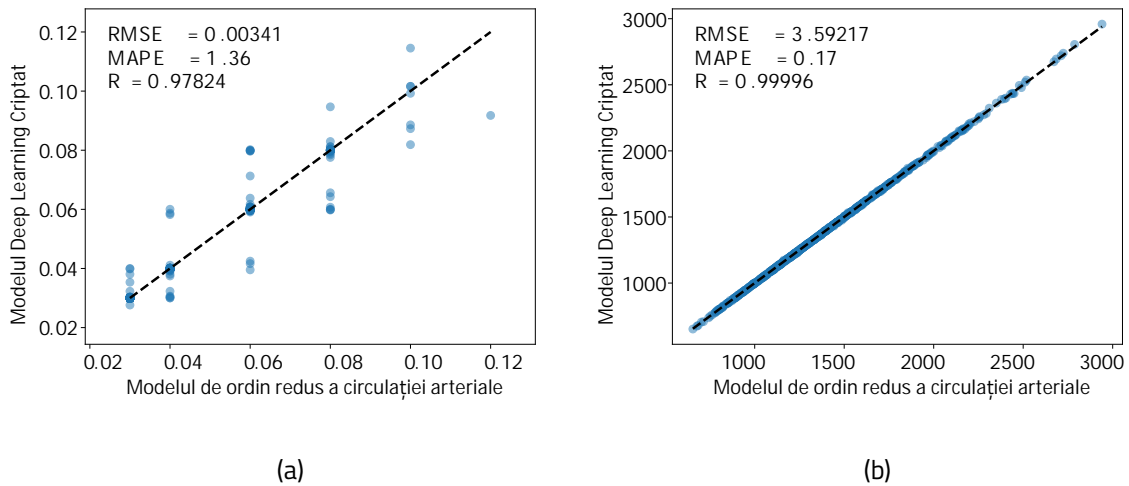


Figura 5.10: Predicții versus valorile țintă (a) raportul dintre rezistența proximală și distală în circulația sistemică și (b) rezistența sistemică.

scorul F1 pentru cele două clase, LCA și RCA. Figura 5.11 afișează matricea de confuzie, portretizând măsurile de asociere între etichetele adevărate și predicțiile făcute de rețeaua neurală pe setul de date de testare.

Rețeaua antrenată pe date criptate a clasificat corect 96.2% din angiografiile din setul de date de testare. În comparație cu modelul necriptat, precizia a fost identică.

Tabela 5.6: Precizia, rechemarea și scorul F1 obținute de modelul neural pentru clasificarea imaginilor MNIST.

Cifra	Precizia (%)	Rechemarea (%)	Scorul F1 (%)
0	97.9	99.0	98.5
1	98.8	99.2	99.0
2	98.5	98.3	98.4
3	98.1	98.5	98.3
4	98.4	98.3	98.4
5	98.2	98.6	98.4
6	98.4	98.0	98.2
7	97.3	97.9	97.6
8	98.3	97.2	97.7
9	97.5	96.4	97.0
Media	98.1	98.1	98.1

Tabela 5.7: Rezultatele rețelei neurale pentru analiza hemodinamică în timp real pe setul de date de testare.

Circulația	Parametrii	MAPE (%)	Corelația Pearson (%)
Sistemică	Volumul mort	7.03	0.9997
	Timpul la elastația max.	0.13	0.9995
	Rezistența	0.17	0.9999
	Complianța	2.45	0.9867
Pulmonară	Volumul mort	9.88	0.9991
	Timpul la elastația max.	0.10	0.9994
	Rezistența	0.32	0.9998
	Complianța	0.67	0.9983

5.6.2 Timpul de rulare

O comparație detaliată a timpului de rulare pentru fiecare aplicație medicală este prezentată în Tabela 5.9. Deși modelele neurale care operează direct pe date criptate homomorfic sunt semnificativ mai lente (până la un ordin de mărime) atât în faza de antrenare, cât și în cea de inferență, sistemul de criptare utilizat este mult mai rapid în comparație cu sistemele clasice homomorfe, unde diferența este de aproximativ 6 până la 7 ordine de mărime.

5.7 Discuții și concluzii

În ultimii ani, accentul tot mai puternic pus pe protejarea confidențialității datelor medicale și promovarea furnizării de soluții medicale personalizate a încurajat dezvoltarea de instrumente care să permită prelucrarea datelor sensibile prin intermediul rețelelor neuronale.

Soluția propusă urmărește să asigure confidențialitatea datelor în analiza bazată pe inteligența artificială prin integrarea unui mecanism de criptare a datelor și oferirea de rezultate fiabile, care să fie utilizate în fluxurile de lucru clinice. Aplicabilitatea încorporării schemei de criptare MORE în modelele Deep Learning a fost evidențiată prin abordarea a trei probleme diferite: recunoașterea cifrelor,

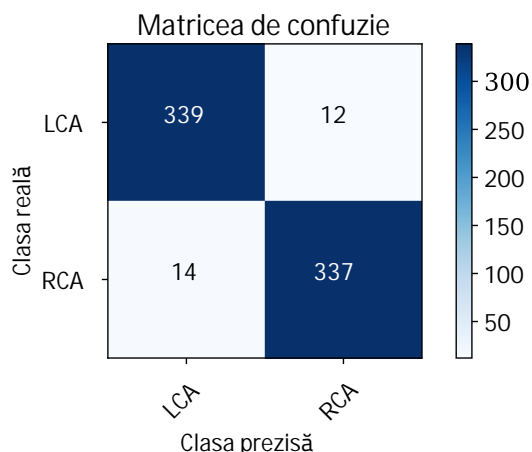


Figura 5.11: Matricea de confuzie obținută de clasificatorul propus pentru identificarea angiografiilor coronariene.

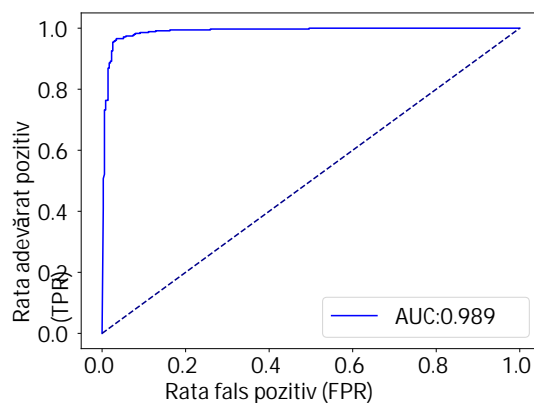


Figura 5.12: Curba ROC obținută de clasificatorul propus pentru identificarea angiografiilor coronariene.

Tabela 5.8: Precizia, rechemarea și scorul F1 obținute de modelul neural pentru analiza hemodinamică.

Label	Precizia (%)	Rechemarea (%)	Scorul F1 (%)
LCA	96.0	96.5	96.3
RCA	96.5	96.0	96.2
Media	96.2	96.2	96.2

Tabela 5.9: Analiza timpului de rulare: valorile medii și abaterea standard a rețelelor criptate și necriptate în cazul celor două aplicații medicale.

Sarcină	Operația	Timpul de rulare (s) pe date criptate	Timpul de rulare (s) pe date necriptate	Raportul criptat/necriptat
Clasificarea angiografiilor	Antrenare (1 epocă)	1075.47±45.54	34.48±1.12	31.19
	Predicție (702 imagini)	26.36±1.98	0.8±0.06	32.95
Analiza hemodinamică	Antrenare (1 epocă)	0.66±0.09	0.021±0.001	31.4
	Predicție (2000 date)	0.102±0.01	0.006±0.0009	17

analiza hemodinamică și clasificarea angiografiei coronariene. S-au făcut cercetări atât pe etapele de antrenare, cât și pe cele de inferență, iar rezultatele demonstrează că ambele pot fi efectuate pe date criptate. S-a demonstrat că acuratețea modelului criptat nu diferă statistic de cea a modelului necriptat și că, urmând strategia propusă, calculele pe date criptate sunt doar ușor mai costisitoare decât cele efectuate pe datele necriptate.

În concluzie, experimentele au arătat că utilizarea schemei de criptare homomorfică MORE ca mecanism de protejarea a confidențialității datelor a permis aplicarea modelelor Deep Learning pe date criptate, fără a compromite precizia. Deși timpul de rulare a crescut cu mai mult de un ordin de mărime, modelele criptate produc rezultate într-un timp rezonabil. Cu sprijinul său direct pentru calcule pe numere raționale și capacitatea de a efectua operații fără a adăuga zgomot, schema devine eligibilă pentru aplicarea unor modele Deep Learning mai complexe. A se reține că schema MORE

permite un compromis între securitate și eficiență: prin creșterea complexității schemei (adică, ordinul matricei utilizată pentru criptarea unui mesaj) securitatea este îmbunătățită cu un cost de rulare ceva mai lung.

6. Concluzii finale

6.1 Concluzii

Atunci când sunt utilizate de către cadrele medicale în cadrul unui flux clinic de rutină, soluțiile bazate pe inteligență artificială și, în special, pe modele Deep Learning, pot aduce progrese majore în furnizarea de asistență medicală prin îmbunătățirea sănătății (de exemplu, diagnosticare timpurie sau îmbunătățită, planuri de tratamente personalizate, prevenirea bolilor etc.), îmbunătățirea procesului de examinare (de exemplu, reducerea timpului de expunere, diagnosticarea mai rapidă, proceduri neinvazive etc.) și reducerea costurilor asistenței medicale. Datorită tranziției către fișele medicale electronice și a fluxurilor curente de lucru clinice, informațiile din fișa de sănătate a unui pacient pot include sute de date clinice, genetice și imagistice, colectate de-a lungul anilor, puternic interconectate și interdependente. Pentru a atenua dificultățile de analiză, interpretare și sintetizare a datelor medicale complexe, activitățile întreprinse de-a lungul programului de cercetare s-au axat pe propunerea de noi soluții de analiză a datelor bazate pe învățarea automată, ca și mijloc de a permite extragerea de informații suplimentare din cantități masive de date.

Scopul cercetării descrise în prezentul document a fost acela de a investiga potențialul modelelor Deep Learning (rețele neurale avansate) de a îmbunătăți procesul de asistență medicală. În mod specific, atenția s-a îndreptat spre a dezvolta, implementa, testa și verifica o serie de soluții bazate pe învățarea automată din date pentru analiza imagisticii medicale. Deși tehnicile de inteligență artificială sunt integrate în principal în dezvoltarea sistemelor de detectare și diagnosticare asistate de calculator, obiectivul final al tezei a fost de a introduce modelele Deep Learning în procesul de analiză a imagisticii medicale, începând de la achiziția imaginii, continuând cu accesul securizat la datele privind sănătatea pacientului și până la etapa finală a interpretării imaginii. Prin urmare, teza stabilește bazele pentru înlocuirea algoritmilor clasici cu soluții complet automatizate bazate pe învățarea automată în vederea obținerii unui timp de diagnostic redus și o acuratețe îmbunătățită.

Considerând sistemul bazat pe modele Deep Learning pentru rezolvarea problemelor de imagistică medicală dezvoltat, implementat și testat în Capitolul 3 s-a dovedit că soluția de reconstrucție tomografică cu doză completă, pornind de la măsurătorile achiziționate cu doză redusă, produce rezultate promițătoare, validate de radiologi experimentați. În plus, reconstrucția imaginii tomografice computerizate este obținută într-un timp adecvat pentru o setare clinică de rutină. În mod similar, în Capitolul 4 se dezvoltă și se validează un sistem complet automat de detectare a leziunilor la nivelul sânilor. Soluția bazată pe modele Deep Learning încearcă să reducă volumul de lucru și să îmbunătățească eficiența radiologului în interpretarea mamografiilor tridimensionale, despre care se știe că ridică mari probleme în rutinele clinice.

Deși de-a lungul anilor s-au obținut rezultate promițătoare în ceea ce privește analiza imagisticii medicale asistată de calculator, mai este încă un drum lung de parcurs pentru a realiza sisteme complet fiabile. Un obstacol clar în calea dezvoltării unor astfel de sisteme este dat de reglementările actuale de protecție a datelor. În pofida faptului că o cantitate abundentă de date este colectată zilnic în centrele medicale din toată lumea, acestea rămân blocate în interiorul spitalelor. Prin urmare, una dintre cele mai mari provocări din industria biomedicală este dezvoltarea și furnizarea de soluții medicale personalizate fără a dezvălui informațiile sensibile privind sănătatea pacienților. Cu toate acestea, aceste două cerințe se exclud reciproc. Pentru a oferi soluții medicale fiabile, personalizate, cercetătorii trebuie să utilizeze datele existente ale pacienți, lucru dificil de realizat. În plus,

există o necesitate clară pentru un efort comun de agregare a datelor din diferite centre medicale, pentru adaptarea și îmbunătățirea continuă a soluțiilor bazate pe învățarea automată din date. Acest lucru implică însă schimbul de date cu caracter personal. Capitolul 5 oferă o soluție posibilă pentru garantarea confidențialității datelor cu caracter personal în cazul utilizării modelelor Deep Learning. Rezultatele raportate indică faptul că soluția propusă are un potențial mare: (i) performanțele nu diferă de cele obținute cu variantele necriptate ale aplicațiilor bazate pe învățarea automată a datelor și (ii) timpul de execuție crește doar marginal. Această soluție a fost introdusă ca un instrument pentru a facilita colaborarea între cercetători în vederea dezvoltării mai rapide a sistemelor de asistență pentru luarea deciziilor medicale.

6.2 Contribuții originale

Contribuțiile aduse în această lucrare se încadrează în următoarele trei arii de cercetare:

- reconstrucția imaginilor medicale bazată pe învățarea automată;
- diagnosticarea medicală bazată pe învățarea automată;
- asigurarea confidențialității datelor cu caracter personal în procesul de învățare automată.

6.2.1 Reconstrucția imaginilor medicale bazată pe învățarea automată

Unul dintre cele mai active domenii de cercetare în tomografia computerizată (CT) este cel responsabil de reducerea expunerii la radiații și de menținere a calității necesare pentru o diagnosticare precisă. Progresele recente oferite de modelele Deep Learning pentru rezolvarea problemelor inverse din imagistica biomedicală au condus la dezvoltarea de metode alternative de obținere a unor imagini medicale de înaltă calitate din datele achiziționate cu doză redusă de radiații. În timp ce majoritatea soluțiilor de reconstrucție a imaginilor bazate pe învățarea automată abordează problema dintr-o perspectivă de post-procesare, soluția dezvoltată s-a concentrat pe reconstrucția directă a imaginilor tomografice cu doză de radiație completă de la măsurătorile obținute cu doză redusă.

În continuare sunt evidențiate principalele contribuții aduse:

- proiectarea și dezvoltarea unui sistem complet automatizat pentru reconstrucția tomografiei computerizate prin integrarea tehnicilor de învățare automată, precum și al conceptelor de bază ale fizicii și geometriei care descriu sistemul de achiziție;
- proiectarea și dezvoltarea unui sistem îmbunătățit de reconstrucție iterativă, folosind modele Deep Learning și tehnici avansate de antrenare, cu rolul de a integra percepția umană în contextul generării de imagini medicale realiste;
- proiectarea, validarea și aplicarea sistemului într-un scenariu clinic realist;
- proiectarea și dezvoltarea unui instrument de evaluare medicală a calității reconstrucției imaginilor CT.

6.2.2 Diagnosticarea medicală bazată pe învățarea automată

Motivat de progresele recente înregistrate de modelele Deep Learning într-o gamă largă de domenii ale medicinei și de nevoia de sisteme avansate de diagnosticare asistată de calculator, se introduce un algoritm de analiză a mamografiei mamare bazat pe învățarea automată din imagini. Mai precis, o rețea neurală convoluțională este antrenată într-o manieră supervizată pentru a evidenția regiunile suspecte din mamografiile digitale tridimensionale. Pentru a reduce implicațiile introduse de variațiile de formă, mărime și aspect ale leziunilor, precum și de incapacitatea de a delimita exact leziunile de țesutul mamar, problema localizării este formulată ca o problemă de identificare a unei

hărți probabilistice. Pentru a atenua dificultățile care apar în antrenarea rețelelor neurale pe seturi de date mici, o bază de date publică cu mamografiile convenționale este utilizată pentru a pre-antrena modelul și a îmbunătăți capacitatea de generalizare. Suplimentar, se propune un algoritm de depistare a corespondențelor dintre cele două mamografiile tipice achiziționate din unghiuri diferite pentru fiecare sân pentru a îmbunătăți detecția și a reduce numărul de rezultate fals pozitive.

În continuare sunt evidențiate principalele contribuții aduse:

- proiectarea și dezvoltarea unui sistem de detecție automată a maselor de țesut tumoral la nivelul sânilor din mamografiile tridimensionale;
- exploatarea formulării detecției leziunilor pe baza hărții probabilistice;
- proiectarea și implementarea unei strategii de antrenare a modelului neural pentru exploatarea naturii tridimensionale a datelor;
- proiectarea și implementarea unei strategii de pre-antrenare a modelului neural pentru îmbunătățirea generalizării;
- proiectarea și implementarea unui algoritm de post-procesare bazat pe detectarea corespondențelor dintre leziunile identificate în proiecțiile mamografiilor cu tomosinteză pentru a îmbunătăți precizia;
- proiectarea, aplicarea și validarea algoritmului prin utilizarea unei baze de date cu mamografiile tridimensionale;
- studiul experimental privind evaluarea performanței.

6.2.3 Asigurarea confidențialității datelor în procesul de învățare automată

În pofida potențialului uriaș al tehnicilor de învățare automată de a facilita dezvoltarea de aplicații medicale personalizate, adoptarea acestora în fluxurile de lucru clinice a fost îngreunată în multe cazuri de reglementările stricte cu privire la confidențialitatea datelor pacienților. O soluție bazată pe criptarea homomorfică, în deosebi pe schema MORE (eng. Matrix Operation for Randomization or Encryption), este propusă ca un mecanism care să faciliteze manipularea datelor cu caracter personal, fără a necesita acces direct la datele originale. Sistemul criptografic utilizat permite aplicarea operațiilor matematice în cadrul rețelelor neurale direct pe numere cu virgulă mobilă criptate, în detrimentul creșterii ușoare a a gradului de complexitate și implicit a timpului de execuție. Pentru a evalua fezabilitatea soluției propuse și pentru a arăta că aplicarea rețelelor neuronale pe datele criptate nu afectează performanțele, s-a propus dezvoltarea unui model pentru recunoașterea automată a cifrelor din imagini.

Pentru a evalua potențialul metodei propuse în cadrul aplicațiilor medicale, s-au considerat două probleme clinice. Astfel, s-a propus un model pentru a estima ieșirile unui model hemodinamic de circulație a întregului corp din date criptate, respectiv un model pentru a clasifica angiografiile coronariene criptate. Rezultatele experimentale obținute subliniază potențialul soluției propuse de a depăși soluțiile actuale și de a oferi rezultate comparabile, într-un interval de timp rezonabil, cu cele ale modelelor neurale necriptate.

După analizarea aspectelor de securitate ale sistemului de criptare, s-a arătat că sistemul criptografic promovează performanța și utilitatea în detrimentul unei securități mai slabe. Deși soluția poate fi integrată în unele aplicații practice, anumite restricții trebuie să fie luate în considerare în viitoarele activități de cercetare. A fost, de asemenea, examinată posibilitatea reducerii vulnerabilității sistemului propus și s-a propus un sistem criptografic hibrid care încorporează un strat suplimentar de securitate. Rezultatele indică faptul că securitatea sistemului poate fi îmbunătățită, însă sunt necesare modificări suplimentare pentru a limita creșterea nivelului de zgomot introdus.

În continuare sunt evidențiate principalele contribuții aduse:

- proiectarea, dezvoltarea și evaluarea unui sistem securizat de criptare homomorfică care permite efectuarea operațiilor matematice direct pe numere reale criptate;
- proiectarea și dezvoltarea unei platforme software generice care să permită aplicarea modelelor neurale direct pe date criptate homomorfic asigurând confidențialitatea datelor;
- validarea și integrarea într-un sistem de estimare automată a parametrilor necesari personalizării modelului sistemului cardiovascular;
- validarea și integrarea într-un sistem de recunoaștere automată a tipului arterei coronare (dreapta sau stânga) din imagini angiografice criptate;
- realizarea unui studiu experimental pentru evaluarea performanței și securității;
- proiectarea, dezvoltarea și evaluarea sistemului hibrid de criptare homomorfică care adaugă un strat suplimentar de protecție pentru o securitate îmbunătățită.

6.3 Diseminarea rezultatelor

Cercetările întreprinse în perioada de studii doctorale au permis elaborarea și publicarea mai multor articole științifice. Astfel, s-au publicat 5 articole în calitate de prim autor și 4 în calitate de coautor, după cum urmează:

- **Vizitiu, A.**, Nita, C., Puiu, A., Suciu, C., Itu, L., Applying Deep Neural Networks over Homomorphic Encrypted Medical Data, Computational and Mathematical Methods in Medicine, 2020;
- **Vizitiu, A.**, Nita, C., Puiu, A., Suciu, C., Itu, L., Towards Privacy-Preserving Deep Learning based Medical Imaging Applications, IEEE International Symposium on Medical Measurements and Applications (MeMeA), 2019;
- **Vizitiu, A.**, Nita, C., Puiu, A., Suciu, C., Itu, L., Privacy-Preserving Artificial Intelligence: Application to Precision Medicine, 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2019;
- **Vizitiu, A.**, Puiu, A., Reaungamornrat, S., Itu, L., Data-Driven Adversarial Learning for Sinogram-Based Iterative Low-Dose CT Image Reconstruction, 23rd International Conference on System Theory, Control and Computing (ICSTCC), 2019;
- Danu, M., Nita, C., **Vizitiu, A.**, Suciu, C., Itu, L., Deep Learning-based Generation of Synthetic Blood Vessel Surfaces, 23rd International Conference on System Theory, Control and Computing (ICSTCC), 2019;
- **Vizitiu, A.**, Nita, C., Itu, L., Homomorphic Encryption in Deep Learning-based Applications for Healthcare Data Analysis, Transylvanian Machine Learning Summer School (TMLSS), 2018;
- Ciusdel, C., **Vizitiu, A.**, Moldoveanu, F., Suciu, C., Itu, L., Towards Real Time Machine Learning-based Estimation of Fracture Risk in Osteoporosis Patients, International Conference on Optimization of Electrical and Electronic Equipment (OPTIM), 2017 and International Aegean Conference on Electrical Machines and Power Electronics (ACEMP), 2017;
- Suciu, C., Itu, L., Nita, C., **Vizitiu, A.**, Stroia, I., Lazăr, L., Gîrbea, A., Foerster, U., Mihalef, V., Patient-specific Hemodynamic Computations: Application to Personalized Diagnosis of Cardiovascular Pathologies, pp. 177-227, 2017;
- Ciusdel, C., **Vizitiu, A.**, Moldoveanu, F., Suciu, C., Itu, L., Towards Deep Learning-based Estimation of Fracture Risk in Osteoporosis Patients, 40th International Conference on Telecommunications and Signal Processing (TSP), 2017.

Rezultatele cercetărilor au fost prezentate la o serie de conferințe și au atras multă atenție și discuții. În consecință, au fost câștigate 3 premii:

- **“Best Poster Award”** pentru activitatea desfășurată pentru asigurarea confidențialității datelor în rețelele neurale în cadrul unei conferințe a școlii de vară (Transylvanian Machine Learning Summer School) organizată de Google DeepMind;
- **“Innovation Radar Prize 2019”** la categoria “Industrial & Enabling Tech” unde a fost recunoscut potențialul de inovare al soluției de păstrare a confidențialității datelor dezvoltat în cadrul proiectului de cercetare H2020 “My Health – My Data” finanțat de UE;
- **“Best Paper Award for Ph.D. students”** pentru lucrarea de reconstrucție a tomografiei computerizate folosind modele Deep Learning prezentată la conferința IEEE ICSTCC.

Aceste premii recunosc valoarea tezei, precum și performanța cercetării și impactul potențial al activităților întreprinse asupra viitorului industriei medicale.

Bibliografie

- [1] Arthur L. Samuel. "Some Studies in Machine Learning Using the Game of Checkers". In: IBM Journal of Research and Development 3 (1959), pp. 210–229.
- [2] Frost and Sullivan. From \$600M to \$6 billion, artificial intelligence systems poised for dramatic market expansion in healthcare. 2016. url: <https://ww2.frost.com/news/press-releases/%20600-m-6-billion-artificial-intelligence-systems-poised-dramaticmarket-expansion-healthcare>.
- [3] Anamaria Vizitiu, Andrei Puiu, Sureerat Reangamornrat, and Lucian M Itu. "Data-Driven Adversarial Learning for Sinogram-Based Iterative Low-Dose CT Image Reconstruction". In: 2019 23rd International Conference on System Theory, Control and Computing (ICSTCC) (2019), pp. 668–674.
- [4] Anamaria Vizitiu, Cosmin Ioan Nita, Andrei Puiu, Constantin Suciu, and Lucian Mihai Itu. "Privacy Preserving Artificial Intelligence: Application to Precision Medicine". In: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (2019), pp. 6498–6504.
- [5] Anamaria Vizitiu, Cosmin Ioan Nita, Andrei Puiu, Constantin Suciu, and Lucian Mihai Itu. "Towards Privacy Preserving Deep Learning based Medical Imaging Applications". In: 2019 IEEE International Symposium on Medical Measurements and Applications (MeMeA) (2019), pp. 1–6.
- [6] Anamaria Vizitiu, Cosmin Ioan Nita, Andrei Puiu, Constantin Suciu, and Lucian Mihai Itu. "Applying Deep Neural Networks over Homomorphic Encrypted Medical Data". In: Computational and Mathematical Methods in Medicine (2020).
- [7] Warren Sturgis McCulloch and Walter Pitts. "A logical calculus of the ideas immanent in nervous activity". In: Bulletin of Mathematical Biology 52 (1988), pp. 99–115.
- [8] Hao Yan, Laura Cervino, Xun Jia, and Steve B. Jiang. "A comprehensive study on the relationship between the image quality and imaging dose in low-dose cone beam CT." In: Physics in medicine and biology 57 7 (2012), pp. 2063–80.
- [9] Ge Wang, Jong Chul Ye, Klaus Mueller, and Jeffrey A. Fessler. "Image Reconstruction is a New Frontier of Machine Learning". In: IEEE Transactions on Medical Imaging 37 (2018), pp. 1289–1296.
- [10] Jonas Adler and Ozan Oktem. "Learned Primal-Dual Reconstruction". In: IEEE Transactions on Medical Imaging 37 (2018), pp. 1322–1332.
- [11] Michael T. McCann, Kyong Hwan Jin, and Michael Unser. "A Review of Convolutional Neural Networks for Inverse Problems in Imaging". In: ArXiv abs/1710.04011 (2017).
- [12] Jonas Adler and Ozan Öktem. "Solving ill-posed inverse problems using iterative deep neural networks". In: CoRR abs/1704.04058 (2017).
- [13] Hang Zhao, Orazio Gallo, Iuri Frosio, and Jan Kautz. "Loss Functions for Image Restoration With Neural Networks". In: IEEE Transactions on Computational Imaging 3 (2017), pp. 47–57.

- [14] Justin Johnson, Alexandre Alahi, and Li Fei-Fei. "Perceptual Losses for Real-Time Style Transfer and Super-Resolution". In: ECCV. 2016.
- [15] Ishaan Gulrajani, Faruk Ahmed, Martín Arjovsky, Vincent Dumoulin, and Aaron C. Courville. "Improved Training of Wasserstein GANs". In: Conference on Neural Information Processing Systems (NIPS). 2017.
- [16] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. "Image-to-Image Translation with Conditional Adversarial Networks". In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2017), pp. 5967–5976.
- [17] Karen Simonyan and Andrew Zisserman. "Very Deep Convolutional Networks for Large-Scale Image Recognition". In: International Conference on Learning Representations (ICLR) (2015).
- [18] Cynthia H. McCollough et al. "Low-dose CT for the detection and classification of metastatic liver lesions: Results of the 2016 Low Dose CT Grand Challenge." In: Medical physics 44 10 (2017), e339–e352.
- [19] Independent UK Panel on Breast Cancer Screening. "The benefits and harms of breast cancer screening: an independent review". In: The Lancet 380 (2012), pp. 1778–1786.
- [20] Robyn Gartner Roth, Andrew D. A. Maidment, Susan P. Weinstein, Susan Orel Roth, and Emily F. Conant. "Digital breast tomosynthesis: lessons learned from early clinical implementation." In: Radiographics : a review publication of the Radiological Society of North America, Inc 34 4 (2014), E89–102.
- [21] Junqiang Lei, Pin Fan Yang, Li Zhang, Yinzong Wang, and Kehu Yang. "Diagnostic accuracy of digital breast tomosynthesis versus digital mammography for benign and malignant lesions in breasts: a meta-analysis". In: European Radiology 24 (2013), pp. 595–602.
- [22] Jun Wei, Heang-Ping Chan, Chuan Zhou, Y. Wu, Berkman Sahiner, Lubomir M. Hadjiiski, Marilyn A. Roubidoux, and Mark A. Helvie. "Computer-aided detection of breast masses: four-view strategy for screening mammography." In: Medical physics 38 4 (2011), pp. 1867–76.
- [23] Md Atiqur Rahman and Yuhuai Wang. "Optimizing Intersection-Over-Union in Deep Neural Networks for Image Segmentation". In: International Symposium on Visual Computing (ISVC). 2016.
- [24] Jeroen Bertels, Tom Eelbode, Maxim Berman, Dirk Vandermeulen, Frederik Maes, Raf Bisschops, and Matthew B. Blaschko. "Optimizing the Dice Score and Jaccard Index for Medical Image Segmentation: Theory and Practice". In: ArXiv abs/1911.01685 (2019).
- [25] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. "ImageNet Classification with Deep Convolutional Neural Networks". In: Conference on Neural Information Processing Systems (NIPS). 2012.
- [26] Fausto Milletari, Nassir Navab, and Seyed-Ahmad Ahmadi. "V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation". In: 2016 Fourth International Conference on 3D Vision (3DV) (2016), pp. 565–571.
- [27] Edward A. Sickles, Willi Weber, Hazel Galvin, Steven H. Ominsky, and Richard A. Sollitto. "Baseline screening mammography: one vs two views per breast." In: AJR. American journal of roentgenology 147 6 (1986), pp. 1149–53.
- [28] Ziad Obermeyer and Ezekiel J. Emanuel. "Predicting the Future - Big Data, Machine Learning, and Clinical Medicine." In: The New England journal of medicine 375 13 (2016), pp. 1216–9.
- [29] Reza Shokri and Vitaly Shmatikov. "Privacy-Preserving Deep Learning". In: Proceedings of the 22Nd ACM SIGSAC Conference on Computer and Communications Security. CCS '15. Denver, Colorado, USA: ACM, 2015, pp. 1310–1321. isbn: 978-1-4503-3832-5. doi: 10 . 1145 / 2810103 . 2813687. url: <http://doi.acm.org/10.1145/2810103.2813687>.

- [30] Craig Gentry and Shai Halevi. "Implementing Gentry's fully-homomorphic encryption scheme". In: *Advances in Cryptology (EUROCRYPT)*. 2010.
- [31] Ahmed El-Yahyaoui and Mohamed Dafir Elkettani. "Fully homomorphic encryption: state of art and comparison". In: *International Journal of Computer Science and Information Security* 14.4 (2016), p. 159.
- [32] Aviad Kipnis and Eliphaz Hibshoosh. "Efficient Methods for Practical Fully Homomorphic Symmetric-key Encryption, Randomization and Verification". In: *IACR Cryptology ePrint Archive* (2012), p. 637.
- [33] Yann Lecun, Leon Bottou, Y. Bengio, and Patrick Haffner. "Gradient-Based Learning Applied to Document Recognition". In: *Proceedings of the IEEE* 86 (Dec. 1998), pp. 2278–2324. doi: 10.1109/5.726791.
- [34] Viorel Mihalef, Lucian Itu, Tommaso Mansi, and Puneet Sharma. "Lumped Parameter Whole Body Circulation Modelling". In: *Patient-specific Hemodynamic Computations: Application to Personalized Diagnosis of Cardiovascular Pathologies*. Ed. by Lucian Mihai Itu, Puneet Sharma, and Constantin Suci. Cham: Springer International Publishing, 2017, pp. 111–152.
- [35] Lucian Mihai Itu, Puneet Sharma, Bogdan Georgescu, Ali Kamen, Constantin Suci, and Dorin Comaniciu. "Model based non-invasive estimation of PV loop from echocardiography". In: *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (2014), pp. 6774–6777.
- [36] Thomas James Ryan. "The coronary angiogram and its seminal contributions to cardiovascular medicine over five decades." In: *Circulation* 106 6 (2002), pp. 752–6.
- [37] Vivian G. Ng and Alexandra J Lansky. "Novel QCA methodologies and angiographic scores". In: *The International Journal of Cardiovascular Imaging* 27 (2010), pp. 157–165.
- [38] Armin Arbab-Zadeh. "What Will it Take to Retire Invasive Coronary Angiography". In: *JACC. Cardiovascular imaging* 9 5 (2016), pp. 565–7.
- [39] Monique Tröbs et al. "Comparison of Fractional Flow Reserve Based on Computational Fluid Dynamics Modeling Using Coronary Angiographic Vessel Morphology Versus Invasively Measured Fractional Flow Reserve." In: *The American journal of cardiology* 117 1 (2016), pp. 29–35.
- [40] Lucian Mihai Itu, Saikiran Rapaka, Tiziano Passerini, Bogdan Georgescu, Chris Schwemmer, Max Schoebinger, Thomas G. Flohr, Puneet Sharma, and Dorin Comaniciu. "A machine-learning approach for computation of fractional flow reserve from coronary computed tomography." In: *Journal of applied physiology* 121 1 (2016), pp. 42–52.

Abstract

In recent years, powered by state-of-the-art achievements in a broad range of areas, machine learning, with emphasis on deep neural networks, has received considerable attention from the healthcare sector.

The present work focuses on the exploration, development, and evaluation of deep learning-based solutions for automatic medical data analysis. The final goal of the current thesis is to incorporate learning-based solutions in the medical imaging analysis pipeline, starting from image acquisition, continuing with image interpretation and up to secure patient health data manipulation.

To exploit the potential of deep learning-based methods for medical imaging analysis, an advanced type of breast imaging, three-dimensional mammography has been considered. Moreover, the reconstruction, which represents the crucial component in producing images of the internal structure of the human body, has been tackled by integrating deep-learning with physics and acquisition geometry of computed tomography (CT). To allow for the deep-learning-based analysis to be performed on medical data without disclosing patient-related health information, privacy-preserving deep learning solutions that operate directly on homomorphically encrypted data have been proposed.

While still far from being deemed trustworthy solutions for practical medical image analysis, the results hence obtained reflect the potential of learning-based approaches to shaping the future of the healthcare industry.

În ultimii ani, ca urmare a avansului tehnologic, popularitatea inteligenței artificiale a explodat afectând un spectru larg al domeniilor de activitate, printre care și sectorul medical.

Lucrarea de față se concentrează pe exploatarea, dezvoltarea și evaluarea soluțiilor bazate pe rețele neurale adânci (eng. Deep Learning) pentru analiza automată a datelor medicale. Astfel, scopul final al prezentei teze este de a include soluții bazate pe învățare în procesul de analiză și prelucrare a imaginilor medicale, pornind de la achiziția imaginii, continuând cu interpretarea acesteia și până la asigurarea confidențialității datelor cu caracter personal în vederea manipulării acestora.

În acest context, pentru a exploata potențialul modelelor neurale adânci în analiza imagisticii medicale, s-a propus o metodă de detecție automată a maselor de țesut tumoral la nivelul sânului din mamografiile digitale cu tomosinteză (mamografia 3D). Pentru a demonstra avantajele pe care rețelele neurale adânci le pot aduce în procesul de formare al imaginilor medicale, reconstrucția, care reprezintă componenta esențială în reprezentarea structurii interne a corpului uman, este adresată prin combinarea rețelelor neurale adânci cu noțiunile de bază ale fizicii și geometriei ce descriu sistemul de achiziție al tomografiei computerizate (CT). În finalul tezei, s-a propus o soluție care permite realizarea unor aplicații de medicină personalizată bazate pe inteligență artificială și care, în același timp, protejează datele personale ale pacienților. Soluția propusă se bazează pe rețele neurale adânci și pe criptarea homomorfică, un tip special de criptare care permite realizarea de operații, de exemplu aritmetice, asupra informațiilor criptate.

Deși încă nu au ajuns la nivelul de încredere necesar pentru a putea fi utilizate în practica medicală, rezultatele obținute reflectă potențialul soluțiilor bazate pe rețele neurale adânci în conturarea viitorului sistemului medical.