MINISTERUL EDUCAȚIEI ȘI CERCETĂRII
UNIVERSITATEA NAȚIONALĂ DE ȘTIINȚĂ ȘI TEHNOLOGIE
POLITEHNICA BUCUREȘTI

Şcoala Doctorală de Electronică, Telecomunicații și Tehnologia Informației

REZUMAT TEZĂ DE DOCTORAT

Alexandra-Georgiana ANDREI

SISTEME DE INTELIGENȚĂ ARTIFICIALĂ PENTRU DIAGNOSTIC ȘI MONITORIZARE

ARTIFICIAL INTELLIGENCE SYSTEMS FOR DIAGNOSIS AND MONITORING OF MEDICAL CONDITIONS

Prof. Dr. Ing. Bogdan IONESCU
POLITEHNICA Bucharest

Conducător de doctorat

BUCUREȘTI 2026

Cuprins

1 Introducere **1**
1.1 Prezentarea domeniului de doctorat 1
1.2 Scopul tezei 2
1.3 Conținutul tezei 2

2 Concepte Teoretice **3**
2.1 Necesitatea imaginilor medicale sintetice 3
2.2 Confidențialitatea datelor în imagistica biomedicală 3
2.3 Modalități de imagistică medicală 4
2.4 Modele generative 4

3 Identificarea Tendințelor în Rețelele Generative Adversariale **5**
3.1 Domeniu și motivație 5
3.2 Metodologie 5
3.3 Rezultate 5
3.4 Concluzii 6

4 Metode de învățare automată pentru analiza tulburărilor neurologice **7**
4.1 Domeniu și motivație 7
4.2 Recenzie sistematică a epilepsiei pe baza EEG 7
4.3 Detectarea Blocajului de Mers 8
4.4 Concluzii 8

5 Augmentarea Imaginilor Histopatologice cu Protejarea Confidențialității **9**
5.1 Domeniu și motivație 9
5.2 Metodologie 9
5.3 Rezultate 11
5.4 Concluzii 11

6 Investigarea Impactului Modelelor Generative asupra Imaginilor Biomedicale Sintetice **12**
6.1 Introducere și motivație 12
6.2 Metodologie 12

6.3 Rezultate 13
6.4 Concluzii inter-ediții 14

7 Compromisuri Utilitate-Confidențialitate în Generarea de Imagini Medicale Sintetice **15**
7.1 Introducere și motivație 15
7.2 Metodologie și cadru de evaluare 15
7.3 Rezultate 16
7.4 Concluzii 16

8 Concluzii **17**
8.1 Rezultate obținute 17
8.2 Contribuții originale 18
8.3 Lista publicațiilor originale 19
8.4 Perspective pentru dezvoltări ulterioare 22

Bibliografie **26**

Capitolul 1

Introducere

Inteligența Artificială (IA) și Învățarea Automată (ÎA) au devenit componente esențiale atât în cercetarea biomedicală, cât și în practica clinică, facilitând abordări bazate pe date pentru diagnosticarea bolilor, monitorizarea pacienților și planificarea tratamentelor într-o gamă largă de specialități medicale. Capacitatea acestor tehnologii de a identifica tipare relevante în date complexe și de înaltă dimensionalitate, precum imagini medicale, biosemnale sau date provenite de la senzori, a deschis noi perspective pentru dezvoltarea unor instrumente automatizate care să sprijine procesul decizional clinic. Cu toate acestea, implementarea unor sisteme de IA fiabile în domeniul sănătății rămâne dificilă. Datele medicale sunt adesea limitate, sensibile, costisitor de adnotat și supuse unor reglementări stricte, precum Regulamentul General privind Protecția Datelor, ceea ce complică dezvoltarea unor modele care să fie în același timp eficiente și de încredere.

Această teză abordează o parte dintre aceste provocări prin contribuții care acoperă mai multe domenii medicale și tipuri de date. Teza este unificată de preocuparea pentru dezvoltarea unor metode de IA care nu doar ating performanțe ridicate, ci sunt și responsabile, oferind rezultate relevante din punct de vedere clinic, funcționând robust în condiții de date limitate și asigurând gestionarea adecvată a informațiilor pacienților, cu respectarea cerințelor de confidențialitate.

1.1 Prezentarea domeniului de doctorat

Imagistica medicală este un instrument esențial pentru diagnostic și planificarea tratamentului, iar aplicarea IA în acest domeniu a demonstrat un potențial considerabil. Totuși, progresul este adesea limitat de lipsa unor seturi de date mari, diverse și bine adnotate. Modelele generative, în special Rețelele Generative Adversariale (GAN) [6] și modelele de difuzie [10], oferă o soluție promițătoare prin generarea de imagini sintetice realiste, care reproduc proprietățile datelor clinice. Cu toate acestea, utilizarea lor ridică întrebări legate de utilitatea clinică și de protecția confidențialității, deoarece pot reține informații identificabile din datele de antrenament. Provocări similare apar și în analiza

datelor fiziologice și a celor provenite de la senzori purtabili. Tulburări neurologice precum epilepsia (peste 50 de milioane de cazuri la nivel global [26]) și boala Parkinson (peste 10 milioane [20]) implică scenarii dificile: date EEG cu variabilitate ridicată și adnotări limitate, respectiv necesitatea unor clasificatori robuști pentru monitorizarea simptomelor motorii în condiții reale, pe baza unor seturi reduse de date etichetate.

1.2 Scopul tezei

Scopul acestei teze este de a investiga aplicarea Inteligenței Artificiale și a Învățării Automate asupra datelor medicale din multiple modalități și contexte clinice, urmărind trei obiective interdependente: dezvoltarea metodelor de generare a datelor sintetice, evaluarea implicațiilor asupra confidențialității ale modelelor generative și elaborarea unor metode de învățare automată pentru aplicații clinice cu date limitate. În domeniul generării de imagini, teza analizează arhitecturi generative aplicate în histopatologie și imagistica CT, propune cadre de augmentare a datelor cu protecție a confidențialității și realizează o evaluare comparativă a modelelor, evidențiind compromisurile dintre calitatea imaginilor, utilitatea lor în sarcini ulterioare și riscurile privind confidențialitatea. O întrebare centrală este dacă imaginile sintetice pot fi simultan utile clinic și sigure sau dacă păstrează urme identificabile din datele reale utilizate la antrenare. Dincolo de generarea de imagini, teza abordează și scenarii cu deficit de date prin două contribuții în analiza tulburărilor neurologice: o recenzie sistematică a metodelor de învățare nesupervizată și auto-supervizată pentru epilepsie pe baza datelor EEG și un studiu experimental privind detectarea automată a blocajului mersului în boala Parkinson, folosind date de la senzori accelerometrici.

1.3 Conținutul tezei

Teza este structurată în două părți. Partea I (Capitolele 1-2) stabilește cadrul teoretic și conceptual al lucrării. Partea a II-a (Capitolele 3-7) prezintă contribuțiile originale: o recenzie sistematică a metodelor de sinteză a imaginilor medicale bazate pe GAN (Capitolul 3), învățarea nesupervizată pentru tulburări neurologice (Capitolul 4),dezvoltarea unor tehnici de augmentare a imaginilor histopatologice cu mecanisme de protecție a confidențialității (Capitolul 5), o campanie multianuală de evaluare comparativă a confidențialității modelelor generative (Capitolul 6) și un studiu comparativ al compromisului dintre utilitate și confidențialitate în sinteza imaginilor obtinute prin Tomografie Computerizată (CT) (Capitolul 7). Capitolul 8 sintetizează concluziile, evidențiază contribuțiile originale, prezintă lista publicațiilor și conturează direcțiile de cercetare viitoare.

Capitolul 2

Concepte Teoretice

2.1 Necesitatea imaginilor medicale sintetice

Sarcini precum detectarea bolilor, segmentarea și predicția evoluției clinice necesită acces la seturi de date extinse, diverse și bine etichetate. În practică, acest acces este limitat constant de trei bariere majore. În primul rând, constrângerile legale și instituționale împiedică, de cele mai multe ori, partajarea datelor brute ale pacienților între spitale. În al doilea rând, procesul de adnotare necesită expertiza radiologilor sau patologilor, fiind astfel costisitor și influențat de variabilitatea între observatori. În al treilea rând, bolile rare și cazurile atipice sunt, în mod inerent, subreprezentate, ceea ce duce la un dezechilibru de clase în seturile de date disponibile. În acest context, imaginile medicale sintetice oferă o posibilă soluție pentru depășirea acestor limitări.

2.2 Confidențialitatea datelor în imagistica biomedicală

În ciuda potențialului lor, modelele generative ridică probleme importante legate de confidențialitate și securitate. Studii recente arată că atât GAN-urile, cât și modelele de difuzie pot memora sau reproduce caracteristici identificabile din imaginile reale utilizate la antrenare. În context biomedical, acest aspect este deosebit de sensibil, deoarece imaginile pot conține informații biometrice sau diagnostice despre pacienți.

Literatura descrie trei tipuri principale de atacuri: atacuri de legătură (re-identificarea pacienților prin corelarea datelor sintetice cu cele reale), inferența atributelor (deducerea unor caracteristici sensibile) și inferența apartenenței (determinarea dacă o anumită imagine a fost inclusă în setul de antrenament) [23, 16, 15, 25]. În acest context, asigurarea faptului că imaginile sintetice sunt atât utile clinic, cât și sigure din punctul de vedere al confidențialității necesită metode riguroase și standardizate de evaluare, un aspect care motivează direct contribuțiile acestei teze.

2.3 Modalități de imagistică medicală

Teza se concentrează pe două modalități principale de imagistică. Histopatologia presupune examinarea microscopică a secțiunilor de țesut colorate cu Hematoxilină și Eozină (HâE), oferind detalii esențiale despre morfologia celulară, necesare pentru diagnosticul și stadializarea cancerului. Imaginile de tip whole-slide (WSI) sunt, de regulă, împărțite în fragmente mai mici pentru a permite analiza automată [7]. TC utilizează raze X pentru a genera imagini în secțiune transversală ale corpului, facilitând analiza volumetrică a structurilor anatomice. În cadrul acestei teze, imagistica tuberculozei pulmonare este utilizată ca domeniu de aplicație pentru experimentele din Capitolele 6 și 7.

2.4 Modele generative

Rețelele Generative Adversariale (GAN) constau dintr-un generator G și un discriminator D antrenați în mod adversarial. G mapează un vector latent $z \sim p_z(z)$ la un eșantion sintetic, în timp ce D clasifică fiecare imagine ca reală sau falsă [6]. Antrenamentul este guvernat de obiectivul minimax:

$$\min_G \max_D V(D,G) = \mathbb{E}_{x\sim p_{\text{data}}}[\log D(x)] + \mathbb{E}_{z\sim p_z}[\log(1-D(G(z)))]. \tag{2.1}$$

Variantele GAN utilizate în această teză includ DCGAN [21], cGAN [17], BigGAN [5] și StyleGAN [13].

Modelele Probabilistice de Difuzie prin Denoising (DDPM) generează imagini prin învățarea inversării unui proces gradual de adăugare a zgomotului pe parcursul a T pași [10]. Procesul direct $q(x_t!\mid!x_{t-1}) = \mathcal{N}(x_t;\sqrt{1-\beta_t},x_{t-1},,\beta_t I)$ degradează progresiv datele până la zgomot gaussian izotrop, iar distribuția inițială este reconstruită cu ajutorul unei rețele neuronale care modelează procesul invers $p_\theta(x_{t-1}!\mid!x_t) = \mathcal{N}(x_{t-1};\mu_\theta(x_t,t),\Sigma_\theta(x_t,t))$. În teză sunt analizate atât variantele necondiționate, cât și cele condiționate (cDiffusion).

Metrici de evaluare. Calitatea imaginilor și fidelitatea distribuțională sunt evaluate folosind Distanța Fréchet de Inception (FID) [9], Scorul Inception (IS) [22] și metrici bazate pe precizie/recall (precizie-α, recall-β, densitate, acoperire). Riscurile privind confidențialitatea sunt cuantificate prin ratele de succes ale atacurilor de inferență a apartenenței.

Capitolul 3

Identificarea Tendințelor în Rețelele Generative Adversariale

3.1 Domeniu și motivație

GAN-urile au devenit printre cele mai utilizate modele pentru sinteza realistă a imaginilor medicale, cu aplicații care acoperă radiologia, histopatologia, oftalmologia și dermatologia. Expansiunea rapidă a metodelor bazate pe GAN a generat o literatură extinsă, însă fragmentată. În acest context, acest capitol prezintă o revizuire sistematică a metodelor de sinteză a imaginilor medicale bazate pe GAN, realizată în conformitate cu ghidurile PRISMA [19], oferind astfel un cadru metodologic unitar pentru contribuțiile experimentale din capitolele următoare.

3.2 Metodologie

A fost utilizat un șir de căutare structurat, care a combinat termeni specifici Rețelelor Generative Adversariale cu diferite modalități de imagistică medicală, aplicat în bazele de date PubMed, IEEE Xplore și Scopus. Criteriile de includere au impus ca studiile să: (i) utilizeze metode bazate pe GAN pentru generarea sau translatarea imaginilor medicale; (ii) să se concentreze pe cel puțin o modalitate de imagistică medicală; (iii) să includă o evaluare cantitativă a rezultatelor. În total au fost identificate 11.592 de înregistrări. După eliminarea duplicatelor, filtrarea titlurilor și rezumatelor, precum și evaluarea textului integral, au fost selectate 286 de studii pentru analiza aprofundată.

3.3 Rezultate

Tendințe de publicare și seturi de date. Dintre cele 286 de studii incluse, 227 (79,2%) au fost publicate în ultimii cinci ani. În ceea ce privește tipul seturilor de date utilizate, 137 de articole s-au bazat exclusiv pe seturi de date publice, 127 au utilizat date private,

iar 22 au combinat ambele tipuri. RMN-ul reprezintă cea mai frecvent utilizată modalitate de imagistică, fiind urmat de radiografiile pe raze X și de imagistica fundului de ochi (retinografie). De asemenea, CT-ul și histopatologia sunt bine reprezentate în literatură. Predominanța neuroimagisticii și a imagisticii toracice reflectă existența unor seturi de date publice de tip benchmark, de mari dimensiuni, disponibile în aceste domenii.

Pre-procesare și arhitecturi. Etapele de pre-procesare diferă în funcție de modalitatea de imagistică utilizată. În cazul datelor CT, acestea sunt supuse ferestrelor HU și normalizării intensității. Datele RMN necesită, de regulă, eliminarea craniului și corecția câmpului de bias. În cazul lamelor histopatologice, sunt aplicate tehnici de normalizare a colorației (precum Reinhard și Macenko), precum și fragmentarea în patch-uri. Din perspectiva paradigmelor de generare, traducerea imagine-la-imagine este cea mai frecvent întâlnită, fiind utilizate pe scară largă modele precum Pix2Pix (59 de studii, 20,3%), CycleGAN (59 de studii, 20,6%) și cGAN (58 de studii, 20,3%). Sinteza necondiționată este dominată de variante ale DCGAN și PGGAN, în timp ce sinteza condiționată utilizează tot mai frecvent familii de modele de tip StyleGAN pentru obținerea unor rezultate la rezoluții înalte.

Lacune în evaluare. Metricile FID și IS sunt cele mai frecvent raportate, fiind utilizate în peste 80% dintre studii. În schimb, metricile care reflectă utilitatea clinică, cum ar fi performanța în sarcini clinice, apar în mai puțin de 40% dintre articole. Evaluarea confidențialității, adică verificarea faptului că imaginile sintetice nu pot reproduce sau dezvălui datele de antrenare, este raportată în doar 5 studii, în ciuda afirmațiilor frecvente privind beneficiile acestor metode în protejarea confidențialității.

3.4 Concluzii

Această recenzie evidențiază adoptarea rapidă a GAN-urilor pentru generarea de imagini medicale în diverse modalități și contexte clinice. Aceste modele au demonstrat un potențial semnificativ în aplicații precum augmentarea datelor, traducerea între modalități și simularea patologiilor, reușind să genereze imagini anatomic realiste și relevante din punct de vedere clinic. În același timp, persistă o serie de limitări importante. Multe metode se bazează încă pe seturi de date pereche, practicile de evaluare rămân neuniforme și adesea nu reflectă adecvat utilitatea clinică, iar aspectele legate de confidențialitate sunt rareori abordate în mod riguros. Aceste probleme evidențiază necesitatea unor cadre de lucru mai robuste, mai relevante clinic și care să asigure protecția datelor.

Deși GAN-urile reprezintă un instrument promițător pentru sinteza imaginilor medicale, depășirea provocărilor discutate în această recenzie este esențială pentru tranziția acestor metode din zona cercetării experimentale către aplicații sigure, fiabile și utilizate pe scară largă.

Capitolul 4

Metode de învățare automată pentru analiza tulburărilor neurologice

4.1 Domeniu și motivație

Acest capitol prezintă două contribuții complementare care abordează tulburările neurologice prin metode de IA, eficiente din punct de vedere al utilizării etichetelor: (i) o recenzie sistematică a literaturii privind învățarea nesupervizată și auto-supervizată aplicată cercetării epilepsiei pe bază de EEG; (ii) un studiu experimental privind detectarea automată a blocajului de mers (Freezing of Gait - FoG) în boala Parkinson, utilizând date provenite de la accelerometre tri-axiale.

4.2 Recenzie sistematică a epilepsiei pe baza EEG

Metodologie. Recenzia a fost realizată în conformitate cu ghidurile PRISMA. Au fost consultate cinci baze de date: Embase, Cochrane, Web of Science, PubMed și Google Scholar. Criteriile de includere au impus ca studiile să utilizeze algoritmi nesupervizați sau auto-supervizați aplicați înregistrărilor EEG provenite de la subiecți umani diagnosticați cu epilepsie, publicate după anul 2013 și redactate în limba engleză. În total au fost identificate 11.592 de înregistrări. După procesul de filtrare, au fost reținute 108 studii pentru analiză.

Rezultate. Dintre cele 108 studii incluse, 82 (75,9%) au utilizat exclusiv seturi de date publice, cele mai frecvent utilizate fiind CHB-MIT [24] și setul de date al Universității din Bonn [3]. Detecția crizelor epileptice a fost cea mai frecventă sarcină (57% dintre studii), urmată de predicția crizelor (25%), clasificarea tipului de criză (22%), localizarea crizelor (4%) și caracterizarea pacienților (2%). Au fost identificate două paradigme principale de modelare: modele specifice pacientului (42 de studii), care oferă performanțe ridicate dar generalizare limitată, și modele multi-pacient generalizate (63 de studii), cu performanțe mai stabile între subiecți. Din punct de vedere metodologic, autoencodoarele și metodele

de clustering sunt predominante, în timp ce abordările contrastive și auto-supervizate reprezintă o direcție emergentă. Totodată, 73% dintre studii nu raportează proceduri de eliminare a artefactelor, ceea ce ridică probleme de reproductibilitate. Transformata Fourier pe termen scurt este cea mai utilizată reprezentare timp-frecvență. Lipsa unor protocoale standardizate de evaluare este identificată ca principal factor limitativ al domeniului.

4.3 Detectarea Blocajului de Mers

Metodologie. Blocajul de mers reprezintă o incapacitate tranzitorie de inițiere sau menținere a mersului și constituie unul dintre cele mai invalidante simptome motorii ale bolii Parkinson. Detectarea automată prin senzori purtabili este relevantă clinic pentru monitorizarea evoluției bolii și activarea intervențiilor asistive [8].

Au fost evaluate cinci cadre de învățare profundă: două variante AlexNet (AlexNet-A, cu un strat ascuns suplimentar, și AlexNet-B, fără acesta), LeNet, VGG-16 și un model bazat pe autoencoder pentru extragerea caracteristicilor, urmat de un clasificator SVM. Toate modelele au utilizat ca intrare segmente de 2, 3, 4 și 5 secunde, cu suprapunere de 50%, din semnalele accelerometrului tri-axial. Studiul a analizat sistematic influența poziționării senzorului (gleznă, coapsă, trunchi), a axelor semnalului (X, Y, Z și magnitudine), precum și a strategiilor de echilibrare a datelor (sub-eșantionare versus lipsa echilibrării) asupra performanței de detecție. A fost utilizat setul de date Daphnet FoG [4], care conține 8 ore și 20 de minute de înregistrări provenite de la 10 pacienți.
Rezultate. Toate cele cinci modele evaluate au atins o acuratețe maximă de până la 90,92%. Modelul bazat pe autoencoder și SVM, împreună cu LeNet, au oferit cele mai consistente performanțe în funcție de poziția senzorului. Senzorii amplasați pe coapsă și trunchi au generat în mod constant cele mai discriminative semnale, în timp ce senzorii de la gleznă au prezentat o variabilitate mai mare a performanței. Intrările bazate pe magnitudine au depășit consistent utilizarea axelor individuale în toate arhitecturile testate. Cea mai bună configurație (autoencoder, coapsă, magnitudine) a atins o acuratețe de 90,92%, depășind abordările CNN comparabile din literatură cu 0,8–7%.

4.4 Concluzii

Recenzia privind epilepsia bazată pe EEG evidențiază o literatură extinsă, dar eterogenă, și identifică drept lacune majore lipsa standardizării evaluării și diversitatea redusă a seturilor de date. Studiul privind detectareablocajului de mers oferă recomandări concrete privind plasarea senzorilor și alegerea arhitecturilor, obținând performanțe superioare față de lucrările anterioare. În ansamblu, ambele contribuții subliniază importanța disponibilității datelor și a standardizării evaluării ca direcții esențiale pentru dezvoltările viitoare în domeniu.

Capitolul 5

Augmentarea Imaginilor Histopatologice cu Protejarea Confidențialității

5.1 Domeniu și motivație

Examinarea histopatologică a secțiunilor de țesut colorate H&E este esențială pentru diagnosticarea cancerului colorectal (CRC), al treilea cel mai frecvent tip de cancer la nivel global, precum și a cancerului de sân [18]. Dezvoltarea unor modele computaționale robuste este însă limitată de disponibilitatea unor seturi de date mari și bine adnotate.

În acest context, capitolul de față propune un GAN Condițional (cDCGAN) pentru sinteza multi-clasă a imaginilor histopatologice și evaluează în mod riguros atât calitatea imaginilor generate din perspectivă clinică, cât și proprietățile de confidențialitate ale acestora.

5.2 Metodologie

Arhitectura cGAN propusă. Modelul cDCGAN propus se bazează pe arhitecturile DCGAN [21] și cGAN [17], fiind extins prin modificări arhitecturale adaptate sintezei histopatologice multi-clasă. Etichetele de clasă sunt utilizate ca semnale de condiționare atât pentru generator, cât și pentru discriminator. Generatorul condițional este compus din cinci straturi de convoluție transpusă (kernel 4, 128 filtre, stride 2), care procesează un vector latent de dimensiune 1500 concatenat cu etichetele clasei. Fiecare strat deconvoluțional este urmat de activarea LeakyReLU cu pantă 0,2, iar ieșirea finală este obținută prin funcția Tanh. Discriminatorul este format din cinci blocuri convoluționale cu kernel 4×4 și stride 2, fiind încheiat cu un strat sigmoid pentru clasificarea real/sintetic. Modelul este antrenat cu un batch size de 125, utilizând optimizatorul Adam (learning rate 0,0002, $\beta_1 = 0,5$) și funcția de pierdere entropie încrucișată, pe un

set de 79.026 de imagini de antrenament din datasetul NCT-CRC-HE-100K [14]. Sunt incluse șapte clase de țesut: DEB, LYM, MUC, MUS, STR, TUM și NORM.

Evaluarea confidențialității. Analiza confidențialității urmează protocolul ImageCLEFmedical GANs [2]. Pornind de la un set de imagini generate G și un set mixt de imagini reale (care include atât imagini utilizate în antrenament, cât și imagini neutilizate), sarcina este de a identifica dacă se poate determina care imagini reale au contribuit la procesul de antrenare.

Metoda propusă are două etape: (i) extragerea de caracteristici folosind rețele neuronale convoluționale pre-antrenate (ResNet50, EfficientNet și DenseNet121); (ii) gruparea nesupervizată a caracteristicilor imaginilor generate, utilizând k-means și clustering ierarhic. Ipoteza de bază este că imaginile utilizate în antrenament tind să fie mai apropiate în spațiul de reprezentare de grupurile formate de imaginile generate, comparativ cu imaginile nevăzute.

Evaluarea utilității. Utilitatea imaginilor sintetice este evaluată prin performanța în sarcini de clasificare. Sunt utilizate patru arhitecturi: DenseNet, ResNet-50, EfficientNet și MobileNetV2, antrenate pe combinații de date reale și sintetice și comparate cu modele antrenate exclusiv pe date reale.

În plus, este realizat un test de tip Turing, în care patru specialiști în histopatologie evaluează realismul perceptual al imaginilor generate.

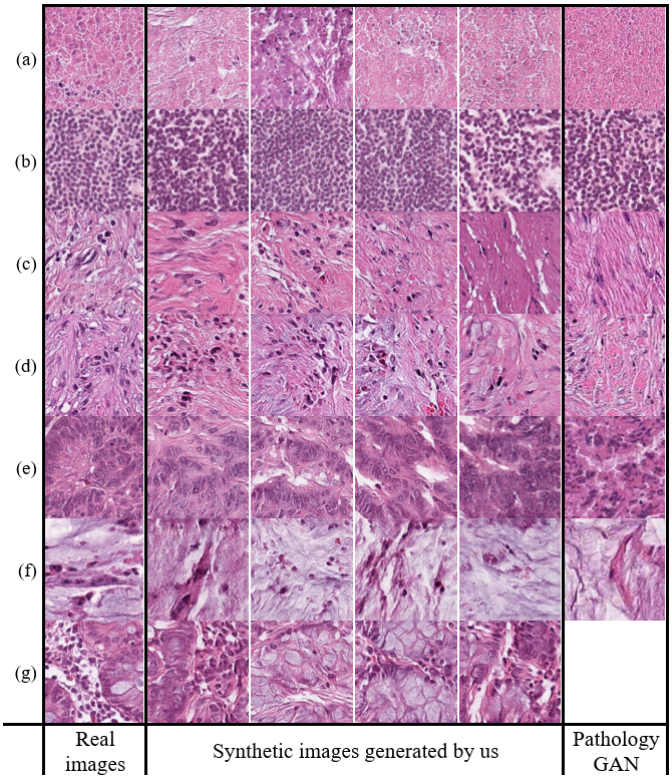


Fig. 5.3 Exemple de imagini reale și sintetice din setul NCT-CRC-HE-100K pentru cancer colorectal: (a) DEB, (b) LYM, (c) MUS, (d) STR, (e) TUM, (f) MUC, (g) NORM.

Tabelul 5.2 Compararea scorului FID cu metode de ultimă generație pe setul CRC (R = imagini reale, S = imagini sintetice).

Metodă	**5k R vs 5k S**	**10k R vs 10k S**
PathologyGAN [1]	83,52	56,03
Difuzie Condiționată [11]	292,86	286,25
StyleGAN [13]	202,42	194,01
Propus	**50,14**	**41,79**

5.3 Rezultate

Calitatea imaginii. Figura 5.3 ilustrează exemple de imagini reale și generate pentru toate cele șapte clase de țesuturi. Imaginile sintetice reușesc să surprindă în mod fidel caracteristicile morfologice relevante, inclusiv textura, tiparele de colorare și organizarea celulară, atingând un nivel de realism care le face dificil de distins față de imaginile reale. Evaluarea cantitativă indică un FID mediu de 64,09 pe toate clasele, cu cea mai bună performanță înregistrată pe clasa LYM (21,2). După cum este prezentat în Tabelul 5.2, metoda propusă depășește semnificativ baseline-urile existente.

Evaluare perceptuală și confidențialitate. Rezultatele testului de tip Turing arată că specialiștii nu au reușit să distingă în mod consistent imaginile reale de cele sintetice, acuratețea de detecție situându-se aproape de nivelul aleator pentru majoritatea tipurilor de țesut.

În ceea ce privește confidențialitatea, acuratețea identificării amprentelor datelor de antrenament rămâne în jur de 50%, indiferent de extractorul de caracteristici, metrica de distanță sau algoritmul de clustering utilizat. Acest rezultat sugerează că imaginile generate nu păstrează informații identificabile despre datele de antrenament.

Utilitatea imaginilor generate. Augmentarea seturilor de antrenament cu imagini sintetice conduce la îmbunătățiri constante ale performanței clasificatorilor, cu câștiguri între 1,83% și 35,36%, în funcție de arhitectură. În plus, testarea pe un set independent de date pentru cancer de sân confirmă generalizabilitatea și utilitatea metodei dincolo de domeniul inițial de antrenare.

5.4 Concluzii

Acest capitol demonstrează că sinteza de imagini histopatologice bazată pe GAN condițional poate asigura simultan calitate ridicată a imaginilor, utilitate în sarcini clinice și protejarea confidențialității datelor.

Modelul cDCGAN propus depășește atât PathologyGAN, cât și baseline-urile mai recente precum modelele de difuzie și StyleGAN în ceea ce privește scorul FID, generând în același timp imagini perceptual indistinctibile de țesuturile reale și care nu expun amprentele datelor de antrenament.

Capitolul 6

Investigarea Impactului Modelelor Generative asupra Imaginilor Biomediale Sintetice

6.1 Introducere și motivație

În ciuda utilizării tot mai frecvente a modelelor generative în imagistica medicală, evaluările sistematice privind scurgerea de informații private în imaginile sintetice rămân limitate. Pentru a aborda acest gol, am lansat seria de evaluare comparativă ImageCLEFmedical GANs în cadrul ImageCLEF, oferind un cadru structurat pentru analiza modului în care imaginile sintetice codifică, păstrează sau dezvăluie informații despre datele de antrenament.

Sunt investigate trei întrebări principale de cercetare: (1) sunt detectabile amprentele datelor de antrenament în imaginile sintetice? (2) poate fi identificată arhitectura generativă care a produs o imagine sintetică? (3) pot fi atribuite imaginile sintetice unor subseturi specifice de date de antrenament?

6.2 Metodologie

Fiecare ediție a pus la dispoziția participanților seturi de imagini CT biomedicale, atât reale, cât și sintetice, generate de diferite arhitecturi, împreună cu un protocol standardizat de evaluare. Modelele generative incluse acoperă cGAN, BigGAN, DCGAN și modele de difuzie. Metodologiile dezvoltate de Laboratorul AI Multimedia au evoluat de-a lungul edițiilor: în 2023 s-au utilizat extragerea de caracteristici și grupare nesupervizată; în 2024 s-a trecut la clasificare cu CNN-uri profunde; iar în 2025 a fost introdusă o arhitectură bazată pe rețele siameze. Pentru subtask-ul 2 (detectarea amprentelor modelului), caracteristicile sunt extrase folosind MobileNetV2, EfficientNet și DenseNet-121. PCA este utilizată pentru reducerea dimensionalității (95% sau 90% variață

păstrată), iar apoi imaginile sunt grupate prin k-means sau clustering ierarhic aglomerativ în funcție de arhitectura generatoare.

6.3 Rezultate

Subtask 1: Identificarea amprentelor datelor de antrenament Aceast subtask este formulat ca o clasificare binară („utilizat pentru antrenament” și „neutilizat”) și este evaluată prin scorul F1 (2023–2024) și coeficientul kappa Cohen (2025).

Abordările Laboratorului AI Multimedia. În 2023, a fost antrenat un DCGAN pe imagini reale, iar rezultatele sale au fost utilizate ca referință. Pentru clasificare, au fost extrase caracteristici de textură LBP și caracteristici profunde VGG-16, care au fost introduse într-un SVM binar. Rezultatele au arătat că LBP depășește caracteristicile profunde (F1 = 0,626 față de 0,585), sugerând că textura surprinde mai bine semnăturile de tip „amprentă”. În 2024 au fost propuse două abordări: (i) segmentarea regiunii pulmonare cu U-Net, urmată de extragerea caracteristicilor (ResNet-50, DenseNet-121) și clustering ierarhic; (ii) detecție de anomalii cu autoencoder, folosind eroarea de reconstrucție (MSE sau SSIM). Cea mai bună performanță a fost obținută de autoencoderul bazat pe SSIM (F1 = 0,594 pe setul cGAN). În 2025, a fost utilizată o rețea siameză antrenată cu pierdere contrastivă pe perechi de imagini reale și sintetice, obținând $\kappa = 0,036$ pe setul de test.

Constatări. În 2023–2024, mai multe echipe au obținut performanțe semnificativ peste nivelul aleator (F1 până la 0,80), confirmând că modelele de difuzie și cGAN lasă amprente detectabile la nivel de imagine. În schimb, ediția 2025, bazată pe StyleGAN, a fost mult mai dificilă: toți cei 33 de participanți au obținut valori ale kappa aproape de zero, iar cea mai mare acuratețe raportată a fost 0,574. Acest rezultat sugerează că StyleGAN este considerabil mai rezistent la inferența apartenenței la nivel de imagine. Rămâne deschisă întrebarea dacă acest lucru reflectă o îmbunătățire reală a confidențialității sau limitări ale metodelor actuale de analiză.

Subtask 2: Detectarea amprentelor modelului generativ Introdus în 2024, aceast subtask investighează dacă imaginile sintetice pot fi grupate după modelul generativ care le-a produs, fără supraveghere. Setul de antrenament conține imagini CT sintetice generate de trei modele, iar setul de test include imagini generate de patru modele diferite. Evaluarea se face prin Indicele Rand Ajustat (ARI).

Abordarea Laboratorului AI Multimedia utilizează caracteristici extrase din cinci CNN-uri pre-antrenate pe ImageNet (VGG-16, ResNet-50, MobileNetV2, EfficientNet, DenseNet-121), precum și LBP. Aceste caracteristici sunt apoi grupate prin k-means și clustering ierarhic. Pe setul de antrenare, DenseNet-121 a obținut ARI perfect (1,0). Pe setul de test oficial, MobileNetV2 cu clustering ierarhic a atins ARI = 0,9971, iar DenseNet-121 ARI = 0,9965, rezultând locul 2 din 46 de echipe (cea mai bună echipă: ARI = 1,0 folosind t-SNE + k-means).

Constatări. Rezultatul central constă în contrastul puternic față de subtask-ul 1. În timp ce identificarea apartenenței la datele de antrenament este dificilă, gruparea imaginilor după modelul generativ este aproape trivială. Diferite arhitecturi imprimă semnături stabile și distincte în textură, conținut frecvențial și artefacte structurale, ușor de detectat chiar și în prezența modelelor necunoscute la testare.

Subtask 3: Identificarea subseturilor de date de antrenament Introdus în 2025, acest subtask mută accentul de la confidențialitatea la nivel de imagine către confidențialitatea la nivel de set de date: fiecare imagine sintetică trebuie atribuită subsetului de date reale care a fost utilizat la antrenarea modelului generativ.

Sunt definite cinci subseturi ($t1$–$t5$), utilizate pentru antrenarea unui model de difuzie. Setul de test conține imagini sintetice din regiuni CT toracice, cervicale și abdominale, introducând variabilitate anatomică ridicată. Metrica oficială este acuratețea clasificării.

Abordarea Laboratorului AI Multimedia. Au fost propuse două metode. Metoda 1 combină caracteristici CNN (MobileNetV2, ResNet-50, EfficientNet, DenseNet-121) cu k-means, clustering ierarhic sau SVM. Cel mai bun rezultat este obținut cu ResNet-50 + SVM (52,36%).

Metoda 2 utilizează o rețea siameză cu patru blocuri convoluționale, embedding L2-normalizat de dimensiune 256 și pierdere contrastivă, atribuind fiecare imagine celui mai apropiat centroid de subset. Pe setul de dezvoltare, modelul atinge 100% acuratețe, iar pe setul oficial de test obține **99,04%**, clasându-se pe locul 1 din 11 echipe.

Constatări. Rezultatele arată că imaginile generate prin difuzie păstrează informații statistice puternice despre subsetul de antrenament, inclusiv tipare anatomice, distribuția leziunilor și caracteristici de achiziție. Acest lucru permite o atribuire aproape perfectă la nivel de set de date. Acest aspect reprezintă un risc important și adesea subestimat pentru confidențialitate: chiar dacă identificarea imaginilor individuale este dificilă, informațiile agregate la nivel de cohortă pot fi recuperate cu acuratețe foarte ridicată.

6.4 Concluzii inter-ediții

Din ansamblul rezultatelor se desprind trei concluzii principale: (i) modelele de difuzie prezintă amprente mai ușor detectabile ale datelor de antrenament decât GAN-urile clasice; (ii) variantele StyleGAN sunt cele mai rezistente atât la detectarea amprentelor, cât și la atribuirea modelului; (iii) atribuirea imaginilor sintetice la subseturi de date este fezabilă, indicând că riscurile de confidențialitate depășesc simpla inferență a apartenenței.

Aceste observații susțin necesitatea dezvoltării unor cadre generative care integrează confidențialitatea încă din faza de proiectare și arată că protecția datelor nu poate fi presupusă, ci trebuie verificată empiric.

Capitolul 7

Compromisuri Utilitate-Confidențialitate în Generarea de Imagini Medicale Sintetice

7.1 Introducere și motivație

Acest capitol prezintă un studiu experimental comparativ amplu asupra a șase arhitecturi reprezentative pentru sinteza imaginilor CT: DCGAN, cGAN, StyleGAN1, StyleGAN2, difuzie necondiționată și difuzie condiționată (cDiffusion). Întrebarea centrală de cercetare este dacă există corelații între calitatea imaginii, utilitatea în sarcini clinice și riscurile de confidențialitate la nivel de arhitectură sau dacă aceste dimensiuni sunt independente și necesită evaluare separată.

7.2 Metodologie și cadru de evaluare

Toate cele șase arhitecturi au fost antrenate pe același set de date CT toracic. Cadrul de evaluare integrează șase direcții complementare: (i) **Metrici de fidelitate și diversitate**: FID, IS, α-precizie, β-recall, densitate și acoperire; (ii) **Plauzibilitate anatomică**: un model UNet de segmentare pulmonară este utilizat pentru a verifica prezența, localizarea și proporțiile structurilor anatomice din imaginile sintetice; (iii) **Evaluare perceptuală umană (test de tip Turing)**: radiologii analizează imagini sintetice în sarcini cu imagine unică și cu imagini multiple, clasificând fiecare caz ca real, sintetic sau incert; (iv) **Explicabilitate (Grad-CAM)**: hărțile de activare bazate pe gradient evidențiază regiunile imaginii utilizate de rețelele discriminator pentru a diferenția imaginile reale de cele sintetice; (v) **Evaluarea confidențialității**: atacurile de inferență a apartenenței sunt utilizate pentru a cuantifica riscul ca imaginile sintetice să dezvăluie informații despre datele de antrenament; (vi) **Suprapunerea distribuțională**: proiecțiile t-SNE ale

caracteristicilor extrase din rețele profunde sunt utilizate pentru a analiza suprapunerea distribuțională și acoperirea modurilor între datele reale și cele sintetice.

7.3 Rezultate

Calitatea imaginii și fidelitatea distribuțională. Modelul de difuzie condiționat obține cea mai echilibrată performanță, cu FID scăzut (56,4) și JSD (0,014), precum și valori ridicate pentru precizie (0,712), recall (0,375), densitate (0,576) și acoperire (0,636).

În contrast, DCGAN prezintă valori foarte reduse pentru toate aceste metrici, indicând un colaps al modelului. StyleGAN2 obține un FID moderat (57,9), dar un recall foarte scăzut (0,0035) și o acoperire limitată (0,182), sugerând o calitate vizuală bună, dar concentrată într-o zonă restrânsă a distribuției.

Evaluarea perceptuală umană. În sarcina cu imagine unică, imaginile cGAN sunt cel mai ușor de identificat ca sintetice (93,3%), în timp ce StyleGAN1 este cel mai convingător (39,4% detectare corectă, 59,2% clasificate greșit ca reale).

Imaginile CT reale sunt recunoscute corect în doar 38,1% din cazuri, ceea ce evidențiază atât dificultatea sarcinii, cât și realismul ridicat al modelelor generative.

În sarcina cu imagini multiple, performanța de discriminare crește pentru toate modelele, deoarece comparația directă scoate în evidență inconsistențe subtile.

Evaluarea confidențialității. Riscul de confidențialitate variază semnificativ în funcție de arhitectură. Modelele de difuzie sunt cele mai vulnerabile la inferența apartenenței, în acord cu Capitolul 6.

În schimb, StyleGAN prezintă cea mai mică susceptibilitate, în ciuda realismului ridicat, fiind astfel cea mai sigură familie din perspectiva confidențialității.

Analiza Grad-CAM arată diferențe clare în mecanismele de detecție: GAN-urile sunt identificate prin artefacte locale de textură, în timp ce modelele de difuzie prin inconsistențe structurale globale.

7.4 Concluzii

Nicio arhitectură nu domină toate dimensiunile evaluate. Modelul cDiffusion oferă cel mai bun echilibru general, dar prezintă riscuri mai mari de confidențialitate. Modelele StyleGAN sunt cele mai realiste și mai robuste din punct de vedere al confidențialității, dar au o diversitate mai redusă. DCGAN suferă în mod clar de colaps de moduri. Aceste rezultate evidențiază necesitatea utilizării unor cadre de evaluare multidimensionale pentru validarea imaginilor medicale sintetice și arată că metodele cu confidențialitate integrată în design sunt esențiale pentru aplicații clinice reale.

Capitolul 8

Concluzii

Această teză investighează aplicarea IA și ÎA pe date medicale din multiple modalități și contexte clinice, contribuind la dezvoltarea unor sisteme de încredere și eficiente pentru domeniul sănătății. Contribuțiile vizează generarea de imagini medicale sintetice și evaluarea confidențialității, învățarea nesupervizată ți pentru analiza tulburărilor neurologice, precum și învățarea automată pentru detectarea simptomelor motorii pe baza senzorilor purtabili. Împreună, aceste direcții avansează înțelegerea modului în care metodele de IA pot fi proiectate, evaluate și implementate responsabil în contexte clinice reale.

8.1 Rezultate obținute

Partea I introduce fundamentele tezei. Capitolul 2 stabilește cadrul conceptual, acoperind motivația utilizării datelor medicale sintetice, provocările de confidențialitate în imagistica biomedicală și arhitecturile modelelor generative analizate în lucrare.

Partea a II-a prezintă contribuțiile originale. Este propusă o recenzie sistematică a metodelor bazate pe GAN pentru sinteza imaginilor medicale, care oferă o privire de ansamblu asupra arhitecturilor, seturilor de date și practicilor de evaluare, evidențiind în special lacunele legate de confidențialitate și utilitate în sarcini clinice.

Un GAN condițional convoluțional profund (cDCGAN) este dezoltat pentru generarea de imagini histopatologice pentru șapte clase de țesuturi. Rezultatele arată că augmentarea cu date sintetice îmbunătățește performanța clasificării și contribuie la protejarea confidențialității pacienților. Evaluarea realizată de experți confirmă realismul imaginilor generate, iar analizele de confidențialitate indică absența scurgerii de informații identificabile din datele de antrenament. Modelul demonstrează și capacitate de generalizare pe un set independent de date pentru cancer de sân.

O campanie de evaluare comparativă desfășurată pe parcursul a trei ediții consecutive (2023–2025) oferă o analiză sistematică a comportamentului modelelor generative din perspectiva confidențialității. Rezultatele arată că modelele de difuzie lasă în mod

constant urme detectabile ale datelor de antrenament, în timp ce arhitecturile StyleGAN prezintă cea mai redusă susceptibilitate la inferența apartenenței. În plus, studiile evidențiază că diferitele arhitecturi imprimă semnături specifice și că atribuirea imaginilor sintetice la subseturi de date este posibilă cu acuratețe ridicată.

Un studiu comparativ a șase arhitecturi de sinteză CT (DCGAN, cGAN, StyleGAN1, StyleGAN2, difuzie și difuzie condiționată) confirmă faptul că nu există o corelație directă între calitate, utilitate și confidențialitate. Aceste dimensiuni trebuie evaluate independent, motivând utilizarea unor cadre de evaluare multidimensionale. Difuzia condiționată oferă cel mai bun echilibru între fidelitate, diversitate și acoperire, în timp ce StyleGAN obține cel mai ridicat realism perceptual.

O recenzie sistematică a metodelor de învățare nesupervizată și auto-supervizată aplicate datelor EEG în epilepsie analizează utilizarea clusteringului, autoencodoarelor și învățării contrastive în condițiile unor date etichetate limitate. Studiul evidențiază lipsa standardizării evaluării ca principală limitare a domeniului.

Un studiu experimental privind detectarea automată a blocajului de mers în boala Parkinson utilizează semnale accelerometrice 3D pentru a evalua cinci cadre de învățare profundă, analizând influența plasării senzorilor, a axelor semnalului și a strategiilor de echilibrare a datelor. Toate modelele ating acurateți de până la 90,92%, iar autoencoderul și LeNet oferă cele mai stabile rezultate. Senzorii plasați pe coapsă și trunchi sunt cei mai informativi, iar reprezentarea pe magnitudine depășește consistent abordările pe axe individuale.

8.2 Contribuții originale

- [J1] propune un GAN condițional pentru generarea de imagini histopatologice de cancer colorectal a șapte clase de țesuturi. Contribuția mea constă în proiectarea arhitecturii, validarea imaginilor prin evaluare de specialitate, compararea cu metode de ultimă generație și analiza confidențialității prin verificarea absenței scurgerii de date identificabile. De asemenea, demonstrez îmbunătățirea performanței în sarcini de clasificare și capacitatea de generalizare pe seturi de date multiple.

- [J2] prezintă o recenzie sistematică a metodelor de învățare nesupervizată și auto-supervizată aplicate EEG în epilepsie. Lucrarea sintetizează metodele existente pentru detecție, predicție și caracterizare a pacienților fără dependență de etichete, analizând arhitecturi, seturi de date și protocoale de evaluare. Sunt evidențiate limitările legate de lipsa datelor, variabilitatea inter-subiect și absența benchmark-urilor standardizate.

- [J3] propune un cadru bazat pe Vision Transformers pentru detectarea și stadializarea glomerulonefritei membranoase din imagini de microscopie electronică.

Contribuțiile mele includ participarea la proiectarea și evaluarea experimentală a modelului, precum și analiza performanței pe seturi de date interne și externe.

- [J4] oferă o recenzie sistematică a GAN-urilor pentru imagini medicale, analizând tendințele din domeniu, seturile de date, preprocesarea, arhitecturile și metricile de evaluare. Lucrarea evidențiază aplicațiile în sinteză necondiționată, traducere imagine-la-imagine și sinteză inter-modalitate, stabilind baza conceptuală a tezei.

- [J5] prezintă un studiu comparativ al modelelor generative pentru imagini medicale, analizând compromisul dintre calitate, utilitate și confidențialitate. Lucrarea evaluează șase arhitecturi într-un cadru unificat, utilizând metrici cantitative, embedding-uri și evaluări umane, evidențiind limitele evaluării bazate pe o singură metrică.

- [C1, C4, C6, C7, C8, C9, B1, B2, B3] investighează impactul modelelor generative asupra imaginilor biomedicale sintetice printr-un cadru de evaluare a confidențialității. Contribuția mea constă în organizarea și coordonarea unei campanii internaționale de evaluare pe parcursul a trei ani, proiectarea task-urilor și curarea seturilor de date. Analizele realizate explorează modul în care arhitecturile generative influențează confidențialitatea și proprietățile datelor sintetice. În paralel, [J4] completează această direcție printr-o analiză comparativă a comportamentului modelelor generative din perspectiva utilității și confidențialității.

- [C2, C3, C5] propun metode pentru evaluarea confidențialității în generarea de imagini biomedicale, vizând identificarea amprentelor datelor de antrenament și a scurgerilor de informații din date sintetice. Contribuțiile mele includ dezvoltarea de metode bazate pe învățare profundă și evaluarea lor sistematică pe multiple arhitecturi și seturi de date. Rezultatele obținute se situează constant printre cele mai performante abordări din competiții.

- [C10, C11] investighează detectarea automată a blocajului de mers în boala Parkinson pe baza semnalelor de la senzori purtabili. Contribuțiile mele includ proiectarea și evaluarea modelelor CNN pentru analiza semnalelor accelerometrice, studiul influenței configurațiilor de senzori și al reprezentărilor de semnal, precum și comparația cu metode existente pentru analiza mersului.

8.3 Lista publicațiilor originale

În această secțiune este prezentată lista tuturor lucrărilor publicate și a lucrărilor trimise spre publicare. Fig. 8.1 prezintă o listă a lucrărilor publicate și disponibile pe Google Scholar la data de 3 mai 2026.

Articole de jurnal:

(J1) **A.-G. Andrei**, M.G. Constantin, M. Graziani, H. Müller, B. Ionescu, "Privacy Preserving Histopathological Image Augmentation with Conditional Generative Adversarial Networks", Elsevier Pattern Recognition Letters, vol. 188, pp. 185-192, https://doi.org/10.1016/j.patrec.2024.12.014, ISSN: 0167-8655, 2025 (Q2 journal, Impact Factor: 3.3, WOS: 001407367700001).

(J2) A.-M. Tautan, **A.-G. Andrei**, C.L. Smeralda, G. Vatti, S. Rossi, B. Ionescu, "Unsupervised Learning from EEG Data for Epilepsy: A Systematic Literature Review", Elsevier Artificial Intelligence in Medicine, vol. 162,ISSN: 0933-3657, DOI:https://doi.org/10.1016/j.artmed.2025.103095, 2025 (Q1 journal, Impact Factor: 6.2, WOS:001437200900001)

(J3) (Accepted) Mihai Gabriel Constantin, George Terinte-Balcan, Ioana Maria Lambrescu, Tudor Emanuel Fertig, Oana Bianca Lazăr, **Alexandra-Georgiana Andrei**, Nicoleta Petre, Bogdan Ionescu, Mihaela Gherghiceanu. Two-stage deep learning networks for the diagnosis and staging of membranous glomerulonephritis from electron microscopy images. Laboratory Investigation. ISSN: 0023-6837 (Q2 journal, Impact Factor: 4.2)

(J4) (Under review) **Andrei AG**, Ionescu B. Generative Adversarial Networks for Medical Image Generation: A Systematic Review. IEEE Journal of Biomedical and Health Informatics. ISSN: 2168-2208 (Q1 journal, Impact Factor: 6.7).

(J5) (In preparation) **Andrei AG**, M.G. Constantin, M. Dogariu, B. Ionescu, "Comparative Study of Generative Models: Utility and Privacy Implications".

Articole de conferințe:

(C1) **A.-G. Andrei**, A. Radzhabov, I. Coman, V. Kovalev, B. Ionescu, H. Müller, "Overview of the ImageCLEFmedical GANs 2023 Task: Identifying Training Data "Fingerprints" in Synthetic Biomedical Images Generated by GANs for Medical Image Security", ImageCLEF 2023, CLEF Conference and Labs of the Evaluation Forum, ISSN: 1613-0073, September 18–21, 2023, Thessaloniki, Greece (indexed in DBLP Computer Science Bibliography);

(C2) **A.-G. Andrei**, B. Ionescu, "AIMultimediaLab at ImageCLEFmedical GANs 2023: Determining "Fingerprints" of Training Data in Generated Synthetic Images", ImageCLEF 2023, CLEF Conference and Labs of the Evaluation Forum, ISSN: 1613-0073, September 18–21, 2023, Thessaloniki, Greece (indexed in DBLP Computer Science Bibliography);

(C3) **A.-G. Andrei**, M.G. Constantin, M. Dogariu, B. Ionescu, "AI Multimedia Lab at ImageCLEFmedical GANs 2024: Deep Learning Approaches for Analyzing Synthetic Medical Images", ImageCLEF 2024, CLEF Conference and Labs of the Evaluation Forum, CEUR-WS, 3740, ISSN: 1613-0073, September 9–12, 2024, Grenoble, France (indexed in DBLP Computer Science Bibliography);

(C4) **A.-G. Andrei**, A. Radzhabov, D. Karpenka, Y. Prokopchuk, V. Kovalev, B. Ionescu, H. Müller, "Overview of 2024 ImageCLEFmedical GANs Task–investigating Generative Models' Impact on Biomedical Synthetic Images", ImageCLEF 2024, CLEF Conference and Labs of the Evaluation Forum, CEUR-WS, 3740, ISSN: 1613-0073, September 9–12, 2024, Grenoble, France (indexed in DBLP Computer Science Bibliography);

(C5) **A.-G. Andrei**, M.G. Constantin, M. Dogariu, L. Ştefan, B. Ionescu, "AI Multimedia Lab at ImageCLEFmedical GANs 2025: Identifying Real-image Usage in Generated Medical Images", ImageCLEF 2025, CLEF Conference and Labs of the Evaluation Forum, CEUR-WS, 4038, ISSN: 1613-0073, September 9–12, 2025, Madrid, Spain (indexed in DBLP Computer Science Bibliography);

(C6) **A.-G. Andrei**, M.G. Constantin, M. Dogariu, A. Radzhabov, L.D. Ştefan, Y. Prokopchuk, V. Kovalev, H. Müller, B. Ionescu, "Overview of ImageCLEFmedical 2025 GANs Task: Training Data Analysis and Fingerprint Detection", ImageCLEF 2025, CLEF Conference and Labs of the Evaluation Forum, CEUR-WS, 4038, ISSN: 1613-0073, September 9–12, 2025, Madrid, Spain (indexed in DBLP Computer Science Bibliography);

(C7) B. Ionescu, H. Müller, A.M. Drăgulinescu, A. Idrissi-Yaghir, A. Radzhabov, A.G.S. de Herrera, **A. Andrei**, et al. "Advancing Multimedia Retrieval in Medical, Social Media and Content Recommendation Applications with ImageCLEF 2024", in Advances in Information Retrieval: 46th European Conference on Information Retrieval, ECIR 2024 Proceedings, March 24–28, Glasgow, UK, 2024 (indexed in Web of Science);

(C8) B. Ionescu, H. Müller, A.-M. Drăgulinescu, A. Popescu, A. Idrissi-Yaghir, A.G. Seco de Herrera, **A. Andrei**, et al. "ImageCLEF 2023 Highlight: Multimedia Retrieval In Medical, Social Media And Content Recommendation Applications", in Springer Lecture Notes in Computer Science, 13982, pp. 557–567, ISBN: 978-3-031-28240-9, DOI: https://doi.org/10.1007/978-3-031-28241-6_64, ECIR 2023 Proceedings, April 2 – 6, Dublin, Ireland, 2023 (indexed in Web of Science);

(C9) B. Ionescu, H. Müller, D.-C. Stanciu, A. Idrissi-Yaghir, A. Radzhabov, A.G. Seco de Herrera, **A. Andrei**, et al. "ImageCLEF 2025: Multimedia Retrieval in Medical, Social Media and Content Recommendation Applications", in Advances

in Information Retrieval: 47th European Conference on Information Retrieval, ECIR 2025 Proceedings, April 6–10, Lucca, Italy, 2025 (indexed in Web of Science);

(C10) **A.-G. Andrei**, A.M. Tăuțan, B. Ionescu, "Freezing of Gait Detection in Parkinson's Disease: Comparison of Deep Learning Frameworks", 18th IEEE International Symposium on Medical Measurements and Applications - MeMeA 2024, 26–28 June, Eindhoven, Netherlands, 2024, DOI: 10.1109/MeMeA60663.2024.10596827, ISSN: 2837-5882 (indexed in Web of Science);

(C11) A.-M. Tăuțan,**A.-G. Andrei**, B. Ionescu, "Freezing of Gait Detection for Parkinson's Disease Patients using Accelerometer data: Case study", IEEE International Conference on E-Health and Bioengineering - EHB, WOS: 000646194100095, October 29-30, Iași, Romania, 2020, (indexed in Web of Science).

Book chapters:

(B1) B. Ionescu, H. Müller, ..., **A.-G. Andrei**, ..., A. Popescu, "Overview of the ImageCLEF 2023: Multimedia Retrieval in Medical, Social Media and Internet Applications", in Springer Lecture Notes in Computer Science, 14163, pp. 370–396, ISBN: 978-3-031-42447-2, DOI: https://doi.org/10.1007/978-3-031-42448-9_25, CLEF 2023 Proceedings, 18-21 September, Thessaloniki - Greece, 2023;

(B2) B. Ionescu, H. Müller, ..., **A.-G. Andrei**, Y. Prokopchuk, ..., B. Stein. "Overview of the ImageCLEF 2024: Multimedia Retrieval in Medical Applications", in Springer Lecture Notes in Computer Science, 14959, pp. 140–164, ISBN: 978-3-031-71907-3, DOI: https://doi.org/10.1007/978-3-031-71908-0_7, CLEF 2024 Proceedings, 9-12 September, Grenoble - France, 2024;

(B3) B. Ionescu, H. Müller, D.-C. Stanciu, **A.-G. Andrei**, ... B. Stein, Overview of ImageCLEF 2025: Multimedia Retrieval in Medical, Social Media and Content Recommendation Applications", in Springer Lecture Notes in Computer Science, 16089, pp. 290–314, ISBN: 978-3-032-04353-5, DOI: https://doi.org/10.1007/978-3-032-04354-2_17, CLEF 2025 Proceedings, 9-12 September, Madrid - Spain, 2025.

8.4 Perspective pentru dezvoltări ulterioare

Contribuțiile acestei teze deschid mai multe direcții de cercetare viitoare, care acoperă generarea de imagini medicale sintetice, proiectarea de modele cu protecția confidențialității, analiza clinică și integrarea mai amplă a IA în practica medicală. În domeniul generării de imagini medicale sintetice, rezultatele arată că riscul de confidențialitate depinde

puternic de arhitectura utilizată și nu poate fi dedus doar pe baza metricilor de calitate a imaginii. O direcție naturală pentru cercetările viitoare este integrarea mecanismelor de protecție a confidențialității, precum confidențialitatea diferențială, regularizarea adversarială sau obiective de antrenare care țin cont de inferența apartenenței, direct în procesul de antrenare al modelelor generative, și nu doar evaluarea lor ulterior. Această tranziție de la evaluare post-hoc la confidențialitate integrată în design este esențială pentru utilizarea sigură a datelor medicale sintetice în contexte de partajare și reutilizare clinică.

În paralel, modelele de evaluare pentru modelele generative necesită dezvoltări suplimentare. Rezultatele tezei evidențiază faptul că realismul vizual, fidelitatea distribuțională, plauzibilitatea anatomică și rezistența la atacuri de confidențialitate sunt dimensiuni complementare, care pot varia semnificativ între arhitecturi. Lucrările viitoare ar trebui să se concentreze pe dezvoltarea unor benchmark-uri standardizate și multidimensionale, care să integreze metrici bazate pe sarcini, evaluări anatomice și teste de rezistență la scurgeri de date. Astfel de cadre ar permite comparații mai consistente între modele și ar oferi recomandări mai clare pentru utilizarea lor în aplicații sensibile din punct de vedere al confidențialității.

În acest context, inițiativele de evaluare comparativă coordonate de comunitate, precum ImageCLEFmedical GANs, reprezintă o platformă importantă pentru studierea aspectelor de securitate și confidențialitate ale imaginilor medicale sintetice. Ediția 2026 a acestei provocări introduce sarcini care vizează identificarea utilizării imaginilor reale în antrenarea modelelor generative, asocierea imaginilor generate cu reprezentări latente și generarea de secțiuni CT cu protecția confidențialității. Aceste direcții investighează explicit amprentele datelor de antrenament, memorarea în spațiul latent și posibila reconstrucție neintenționată a anatomiei pacienților. Extinderea și participarea la astfel de cadre reprezintă o direcție importantă, deoarece permit analiza sistematică a mecanismelor de scurgere a informațiilor și susțin dezvoltarea de modele care echilibrează realismul cu garanții solide de confidențialitate.

Extinderea acestor metode la alte modalități imagistice precum RMN, ultrasunete sau imagistica retiniană, precum și la sarcini mai complexe, cum ar fi sinteza volumetrică 3D și traducerea multi-modală, reprezintă o altă direcție relevantă. În domeniul analizei tulburărilor neurologice, contribuțiile din Capitolul 4 sugerează mai multe direcții viitoare. Pentru cercetarea epilepsiei bazate pe EEG, este necesară dezvoltarea unor benchmark-uri standardizate și compararea sistematică a metodelor de învățare nesupervizată pe seturi de date diverse, pentru a consolida o literatură în prezent fragmentată. În cazul detectării blocajului de mers, provocările legate de generalizare, inclusiv numărul redus de subiecți și dezechilibrul sever între clase, trebuie abordate în lucrările viitoare. Direcțiile posibile includ evaluarea leave-one-subject-out pentru o analiză mai robustă, integrarea arhitecturilor convoluționale și recurente în autoencodoare pentru captarea

mai bună a dinamicii temporale și utilizarea modelelor generative pentru augmentarea episoadelor rare de FoG.

Mai larg, integrarea datelor sintetice și a instrumentelor bazate pe IA în contexte clinice reale necesită o aliniere mai strânsă cu expertiza medicală și cu cadrele de reglementare. Cercetările viitoare ar trebui să analizeze modul în care seturile de date sintetice pot fi validate, documentate și partajate în mod responsabil, precum și modul în care reglementări precum RGPD pot fi integrate în fluxurile de lucru ale modelelor generative. Colaborarea interdisciplinară între cercetători în IA, clinicieni și autorități de reglementare va fi esențială pentru transformarea acestor metode în instrumente de încredere, utilizabile efectiv în cercetare și în practica medicală.

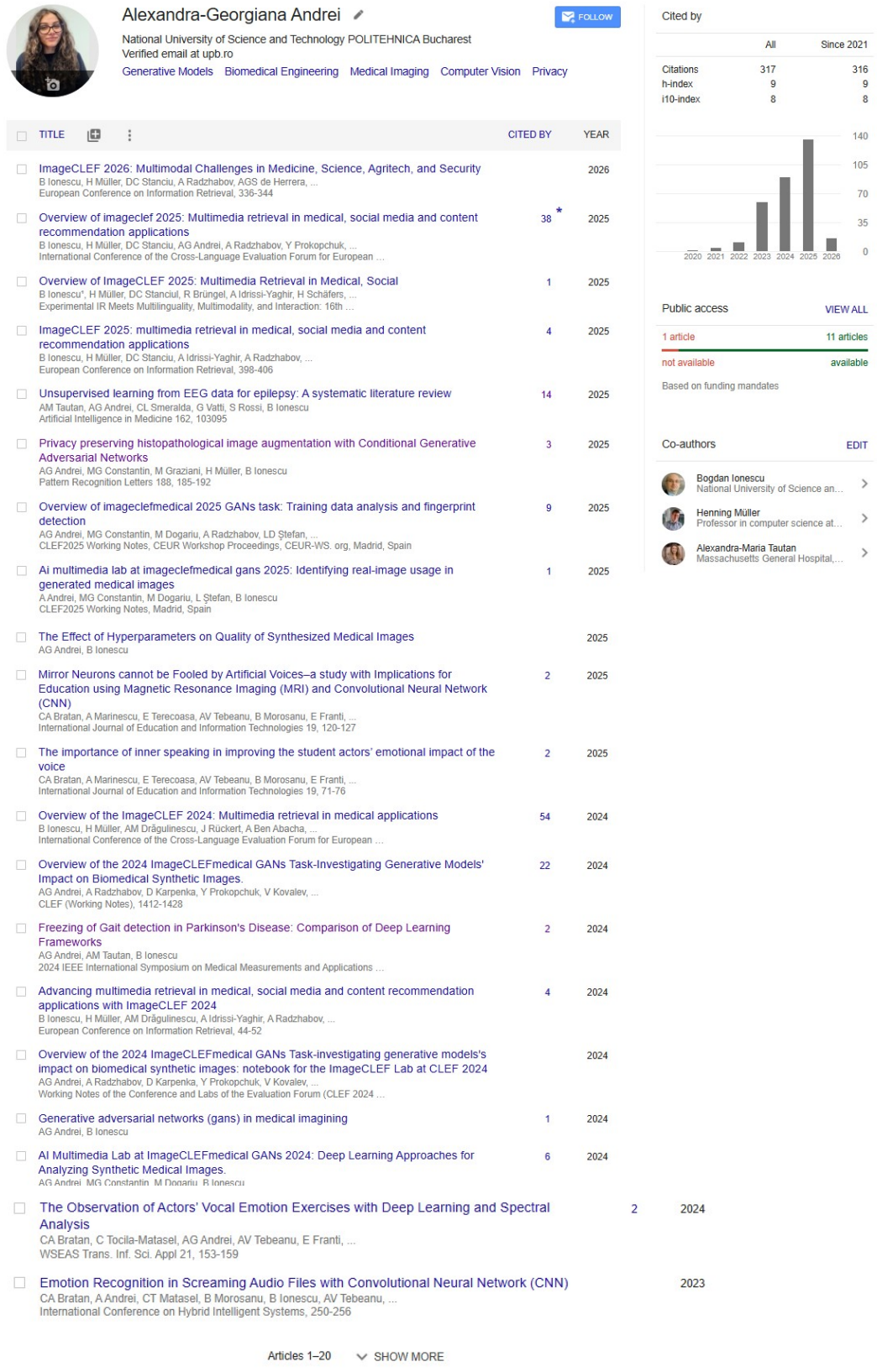


Figura 8.1 Lista lucrărilor publicate pe Google Scholar cu numărul de citări (captură de ecran realizată la 3 mai 2026).

Bibliografie

[1] A. C. Quiros, R. Murray-Smith, K. Y. (2019). Pathologygan: Learning deep representations of cancer tissue. *arXiv preprint arXiv:1907.02644*.

[2] Andrei, A.-G., Constantin, M. G., Dogariu, M., and Ionescu, B. (2024). Ai multimedia lab at imageclefmedical gans 2024: Deep learning approaches for analyzing synthetic medical images. In *CLEF2024 Working Notes*, CEUR Workshop Proceedings, Grenoble, France. CEUR-WS.org.

[3] Andrzejak, R. G., Lehnertz, K., Mormann, F., Rieke, C., David, P., and Elger, C. E. (2001). Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. *Physical Review E*, 64(6):061907.

[4] Bachlin, M., Plotnik, M., Roggen, D., Maidan, I., Hausdorff, J. M., Giladi, N., and Troster, G. (2009). Wearable assistant for parkinson's disease patients with the freezing of gait symptom. *IEEE Transactions on Information Technology in Biomedicine*, 14(2):436–446.

[5] Brock, A., Donahue, J., and Simonyan, K. (2018). Large scale gan training for high fidelity natural image synthesis. *arXiv preprint arXiv:1809.11096*.

[6] Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. (2014). Generative adversarial nets. *Advances in neural information processing systems*, 27.

[7] Gurina, T. S. and Simms, L. (2020). Histology, staining.

[8] Heremans, E., Nieuwboer, A., and Vercruysse, S. (2013). Freezing of gait in parkinson's disease: where are we now? *Current neurology and neuroscience reports*, 13(6):350.

[9] Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., and Hochreiter, S. (2017). Gans trained by a two time-scale update rule converge to a local nash equilibrium. *Advances in neural information processing systems*, 30.

[10] Ho, J., Jain, A., and Abbeel, P. (2020). Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851.

[11] Ho, J. and Salimans, T. (2021). Classifier-free diffusion guidance. In *NeurIPS 2021 Workshop on Deep Generative Models and Downstream Applications*.

[12] Kaabachi, B., Despraz, J., Meurers, T., Otte, K., Halilovic, M., Kulynych, B., Prasser, F., and Raisaro, J. L. (2025). A scoping review of privacy and utility metrics in medical synthetic data. *NPJ digital medicine*, 8(1):60.

[13] Karras, T., Laine, S., and Aila, T. (2019). A style-based generator architecture for generative adversarial networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 4401–4410.

[14] Kather, J. N., Halama, N., and Marx, A. (2018). 100,000 histological images of human colorectal cancer and healthy tissue. Zenodo.

[15] Liu, Q., Shakya, R., Jovanovic, J., Khalil, M., and de la Hoz-Ruiz, J. (2025). Ensuring privacy through synthetic data generation in education. *British Journal of Educational Technology*, 56(3):1053–1073.

[16] Min, G. and Oh, J. (2025). Can synthetic data protect privacy? *IEEE Access*.

[17] Mirza, M. and Osindero, S. (2014). Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784*.

[18] Organization, W. H. (2023). Colorectal cancer. https://www.who.int/news-room/fact-sheets/detail/colorectal-cancer. Accessed: 10 November 2025.

[19] Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., et al. (2021). The prisma 2020 statement: an updated guideline for reporting systematic reviews. *bmj*, 372.

[20] Parkinson's Foundation (n.d.). Statistics. Accessed: 2025-03-09.

[21] Radford, A., Metz, L., and Chintala, S. (2015). Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*.

[22] Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., and Chen, X. (2016). Improved techniques for training gans. *Advances in neural information processing systems*, 29.

[23] Sanchez-Serrano, P., Rios, R., and Agudo, I. (2025). A decision framework for privacy-preserving synthetic data generation. *Computers and Electrical Engineering*, 126:110468.

[24] Shoeb, A. H. (2009). *Application of machine learning to epileptic seizure onset detection and treatment*. PhD thesis, Massachusetts Institute of Technology.

[25] Stadler, T., Oprisanu, B., and Troncoso, C. (2020). Synthetic data-a privacy mirage. *arXiv preprint arXiv:2011.07018*.

[26] World Health Organization (2019). Epilepsy: A public health imperative: Summary. https://apps.who.int/iris/handle/10665/325440. Accessed: Jan. 30, 2023.

[27] Zhu, J.-Y., Park, T., Isola, P., and Efros, A. A. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, pages 2223–2232.