



ACADEMIA ROMÂNĂ

Școala de Studii Avansate a Academiei Române

Institutul de Matematică "Simion Stoilow"

REZUMATUL TEZEI DE DOCTORAT

Modele de învățare vizuală adâncă pentru imagistica și analiză medicală

CONDUCĂTOR DE DOCTORAT:

Prof. Dr. Marius Leordeanu

DOCTORAND:

Florin Condrea

2025

Cuprins

Abstract	2
1 INTRODUCERE	4
1.1 Învăţarea profundă în domeniul medical	4
1.2 Motivaţie	5
1.3 Provocări	5
1.4 Contribuţiile noastre principale	6
2 MODELE INFORMATE ANATOMIC	8
3 ÎNVĂŢAREA PROFUNDĂ DIN PERSPECTIVA CLINICĂ	11
4 ÎNVĂŢAREA SLAB SUPERVIZATA VIA CREŞTEREA GRANULARĂŢI ETI- CHETELOR	14
5 ÎNVĂŢAREA SARCINILOR TEMPORALE FĂRĂ ETICHETE UMANE	17
6 CONCLUZII	20
Bibliography	21

ABSTRACT

Weaving Anatomical Knowledge into Computer Vision for Data-Limited Medical Applications

“It’s still magic even if you know how it’s done.” – Terry Pratchett

Inteligența artificială (IA) medicală a apărut ca un domeniu transformator la intersecția dintre vedere computațională, asistență medicală și învățare automată, permițând îmbunătățirea acurateței diagnosticului și a eficienței în mediile clinice. Pe măsură ce tehnologiile de imagistică medicală continuă să avanseze, capacitatea de a analiza și interpreta cu precizie date radiologice complexe a devenit din ce în ce mai importantă pentru îmbunătățirea rezultatelor pacienților, reducerea erorilor de diagnostic și optimizarea fluxurilor de lucru clinice. Această teză avansează domeniul IA medical prin intermediul unei serii de abordări inovatoare de învățare profundă, concentrându-se pe detectarea emboliei pulmonare și monitorizarea semnelor vitale cu supraveghere umană limitată.

Începem prin abordarea sarcinii medicale de detectare a emboliei pulmonare cu un cadru dual-hop inovator, conștient anatomic, pentru detectarea emboliei pulmonare, care utilizează segmentarea organelor pentru a îmbunătăți precizia detectării. Această abordare îmbunătățește semnificativ metodele existente prin integrarea contextului anatomic în fluxul de detectare, obținând rezultate de ultimă generație pe un set de date public de detectare a emboliei pulmonare la scară largă. Pentru a aborda provocarea critică a granularității adnotării, introducem Label Up, un flux de lucru de învățare slab supervizată care transformă adnotările la nivel de imagine în măști de segmentare detaliate prin tehnici de explicabilitate a modelului, acoperind eficient decalajul dintre adnotările grosiere și segmentările fine fără a necesita efort suplimentar de adnotare manuală.

Teza abordează în continuare provocarea privind insuficiența datelor printr-o abordare inovatoare de generare a datelor sintetice, concepută special pentru aplicații de imagistică

medicală. Metoda noastră păstrează proprietățile statistice temporale ale datelor medicale din lumea reală, menținând în același timp constrângerile de confidențialitate, fără a necesita etichetare umană. Pentru monitorizarea semnelor vitale, dezvoltăm modele de învățare profundă capabile de evaluarea precisă și continuă a parametrilor fiziologici ai pacienților din senzori non-invazivi.

Cercetarea noastră avansează domeniul prin inovații tehnice care abordează provocări cheie în implementarea IA medicală: integrarea informațiilor anatomice pentru o înțelegere contextuală îmbunătățită, fluxuri de lucru de învățare slab supervizată pentru a maximiza utilitatea adnotărilor limitate și generarea de date sintetice pentru a depăși insuficiența datelor. Aceste contribuții îmbunătățesc colectiv performanța, interpretabilitatea și aplicabilitatea clinică a modelelor de învățare profundă în analiza imagisticii medicale, deschizând calea spre sisteme mai eficiente de diagnostic asistat de calculator care pot sprijini în mod semnificativ luarea deciziilor clinice și îmbunătăți îngrijirea pacienților.

Cuvinte cheie – *Învățare profundă, IA în sănătate, Detectare asistată de calculator, Inteligență artificială, Modele conștiente anatomic, Învățare slab supervizată, Generare de date sintetice, IA explicabilă, Imagistică medicală, Embolie pulmonară, Monitorizarea semnelor vitale, Suport pentru decizii clinice*

Capitolul 1

INTRODUCERE

1.1 Învățarea profundă în domeniul medical

Inteligența artificială a transformat domeniul medical, permițând analiza complexă a datelor medicale. Evoluția de la sistemele experte timpurii la aplicațiile moderne de învățare profundă reflectă progrese mai largi ale AI determinate de puterea de calcul, inovația algoritmică și disponibilitatea datelor[35]. Asistența medicală prezintă provocări unice de implementare a AI datorită datelor complexe, eterogene, sensibile la confidențialitate, care necesită abordări specializate[37], alături de cerințele de reglementare și provocările de integrare cu practicile clinice consacrate[13].

Evoluțiile recente ale inteligenței artificiale medicale s-au accelerat prin infrastructura de calcul îmbunătățită, implementarea dosarelor electronice de sănătate și progresele învățării automate[10]. Aceste fundații au îmbunătățit precizia analizei imagistice medicale și asistenței de diagnostic[20], domeniul îndreptându-se către instrumente de sprijinire a deciziilor bazate pe inteligență artificială și abordări medicale personalizate care abordează lacunele în furnizarea de asistență medicală, îmbunătățind în același timp rezultatele pacienților[9].

Învățarea profundă a avut un impact deosebit asupra aplicațiilor medicale, rețelele neuronale convoluționale revoluționând analiza imaginilor și arhitecturile recurente permițând procesarea sofisticată a textului clinic și a datelor în serii de timp[27]. Cu toate acestea, implementarea clinică rămâne limitată din cauza interpretabilității, calității datelor, generării și provocărilor legate de conformitatea reglementărilor[17].

Domeniului medical continuă să evolueze cu soluții de inteligență artificială care consolidează relațiile profesional-pacient, trecând de la telemedicină la sisteme inteligente integrate

care îmbunătățesc procesul decizional clinic, păstrând în același timp îngrijirea umană[35].

1.2 Motivație

Cercetarea medicală în AI reprezintă o intersecție unică de responsabilitate etică și inovație tehnică. Inteligența artificială în domeniul sănătății poate transforma rezultatele pacienților, poate îmbunătăți procesul decizional clinic și poate aborda disparitățile globale în domeniul sănătății[35]. Dintr-o perspectivă etică, asistența medicală este ideală pentru cercetarea AI, oferind beneficii tangibile prin detectarea precoce a bolii, acuratețe îmbunătățită a diagnosticului și planuri de tratament personalizate[13].

Inteligența artificială în domeniul sănătății abordează provocările de accesibilitate și echitate la nivel global. În regiunile cu acces limitat la medici specialiști, sistemele AI pot oferi asistență în diagnosticare și capacități de triaj altfel indisponibile[9]. Această dimensiune etică oferă atât o busolă morală, cât și un scop care transcende realizările pur tehnice.

Din punct de vedere tehnic, inteligența artificială medicală prezintă provocări solide de cercetare care implică date multimodale complexe, cu dimensiuni mari, care necesită abordări sofisticate de modelare[10]. Aplicațiile medicale necesită acuratețe, fiabilitate și interpretabilitate excepționale, împingând granițele învățării automate și necesitând inovații în proiectarea modelelor și protocoalele de validare[37].

Cercetarea medicală în AI beneficiază de obiective clar definite, cu criterii de evaluare obiective, fie că detectează embolii pulmonare, prezic progresia bolii sau generează exemple de antrenament sintetic pentru afecțiuni rare[20]. Această claritate permite cercetarea concentrată cu căi directe de traducere clinică.

1.3 Provocări

Mai multe provocări cheie definesc peisajul actual al cercetării medicale AI:

Deficitul datelor: Seturile de date medicale adnotate sunt limitate din cauza costurilor ridicate de etichetare și a cerințelor de expertiză, în special problematice pentru bolile rare[15]. Pentru afecțiunile cu prevalență sub 1 din 2.000 de persoane, datele imagistice insuficiente duc la o generalizare slabă a modelului[10].

Confidențialitatea datelor: Reglementările stricte privind confidențialitatea restricționează

partajarea datelor pacienților, necesitând învățare federată sau tehnici de generare de date sintetice[29]. GDPR și HIPAA impun cerințe stricte creând bariere de colaborare multi-instituțională[25], catalizând tehnici de păstrare a confidențialității, inclusiv confidențialitatea diferențială și criptarea homomorfă[16].

Interpretabilitate: Natura de cutie neagră a modelelor de învățare profundă împiedică adoptarea clinică acolo unde transparența este crucială[14]. Clinicienii au nevoie nu numai de predicții precise, ci și de explicații aliniate cu raționamentul medical[37]. Abordări precum mecanismele de atenție și explicațiile bazate pe concepte clinice urmăresc să pună legătura între complexitatea algoritmică și interpretabilitatea clinică[34].

Prejudecăți și corectitudine: Seturile de date dezechilibrate duc la predicții părtinitoare care afectează în mod disproporționat anumite grupuri demografice[11]. Prejudecățile istorice în practica medicală și accesul la asistență medicală codificate în datele de formare pot perpetua disparitățile existente[3], modelele instruite pe anumite populații având performanțe slabe pe grupuri subreprezentate[19].

1.4 Contribuțiile noastre principale

Această teză aduce contribuții semnificative la inteligența artificială medicală, abordând provocările aplicațiilor de asistență medicală, avansând în același timp metodologii pentru sisteme AI robuste, relevante din punct de vedere clinic:

Sisteme de diagnosticare end-to-end: Explorăm conducte de IA medicală de diagnosticare end-to-end utilizând informațiile medicale, utilizând informații anatomice pentru pre-procesarea datelor și preantrenarea modelelor din capitolul 2 și informații clinice pentru generarea de date sintetice în capitolul 5.

Combaterea deficitului de date: Am dezvoltat două cadre care permit formarea modelelor de învățare automată pentru sarcini complexe fără a necesita adnotări umane:

Cadrul nostru Label Up[7] îmbunătățește granularitatea adnotărilor prin valorificarea explicabilității modelului, aplicat la detectarea emboliei pulmonare prin actualizarea adnotărilor de la nivel de imagine la adnotări la nivel de voxel, permițând localizarea embolilor individuale.

În capitolul 5, prezentăm generarea de date sintetice care încorporează priorități medicale pentru monitorizarea la distanță a semnelor vitale. Videoclipurile noastre termice sintetice

folosesc primitive geometrice simple transformate pe baza unor frecvențe precise din punct de vedere fiziologic. Modelele antrenate exclusiv pe aceste date sintetice demonstrează o generalizare remarcabilă la videoclipurile termice reale, validând faptul că antrenamentul eficient necesită doar concepte temporale esențiale izolate de caracteristicile vizuale complexe.

Îmbunătățiri arhitecturale: Introducem o arhitectură care folosește atât informațiile globale, cât și cele locale în arhitectura noastră HopNet (Capitolul 2), utilizând mai multe treceri prin datele de intrare, timp în care agregăm caracteristicile de la trecerile anterioare și le folosim ca date de intrare pentru trecerile ulterioare.

Validare clinică: Evaluăm munca noastră în medii clinice (Capitolul 3), examinând valoarea medicală reală pentru pacienți, cu rezultate detaliate din perspectiva clinică care demonstrează impactul uman în lumea reală.

Capitolul 2

MODELE INFORMATE ANATOMIC

Embolia pulmonară (EP) rămâne o cauză principală a mortalității cardiovasculare, diagnosticul prompt și precis fiind crucial pentru rezultatele pacientului. În timp ce angiografia pulmonară tomografică computerizată (CTPA) servește drept standard de aur pentru diagnosticul EP, complexitatea și subtilitatea manifestărilor emboliei reprezintă provocări semnificative pentru radiologi. Pentru a aborda aceste provocări, în acest capitol prezentăm HopNet, o nouă abordare de învățare profundă în 3 etape care valorifică informația anatomică și o arhitectură dual-hop pentru detectarea EP îmbunătățită în scanările CTPA, obținând performanță care depășește stadiul actual al tehnicii pe un set de date publice la scară largă [5] pentru clasificarea al embolismului la nivel de imagine. Soluțiile noastre sunt construite peste un model de bază puternic bazat pe EfficientNetV2 [33], iar contribuțiile noastre sunt aplicate pe trei axe independente de dezvoltare, care se dovedesc a fi necesare pentru o performanță precisă. O prezentare generală a metodei noastre în 3 etape este afișată în Figura 2.1.

Prima etapă: segmentarea anatomică și izolarea regiunilor pulmonare și cardiace. Modulele neuronale profunde instruite pe informații fiziologice pentru segmentarea regiunilor plămânilor și inimii sunt utilizate pentru a segmenta doar informațiile relevante cu privire la detectarea EP. Prin valorificarea volumelor extrase, suntem capabili să mărim rezoluția structurilor anatomice relevante, păstrând în același timp dimensiunea inițială a datelor de intrare. Un avantaj suplimentar este eliminarea structurilor externe care pot fi confundate cu EP. Preprocesarea aduce o îmbunătățire cu 0,6 % a scorului F1 la nivel de studiu față de

Această secțiune se bazează pe lucrarea: Condrea, Florin, Marius Leordeanu et al. "Anatomically aware dual-hop learning for pulmonary embolism detection in CT pulmonary angiograms." *Computers in Biology and Medicine* 174 (2024): 108464.

utilizarea volumelor nedecupate în același timp de execuție.

A doua etapă: preantrenament anatomic. Rețeaua neuronală este antrenată în prealabil pe sarcina de a localiza repere anatomice specifice, înainte de a începe etapa de învățare a EP. Această etapă are rolul de a pregăti rețeaua neuronală atât cu cunoștințe anatomice, ajutând la diferențierea diferitelor structuri anatomice complexe, cât și la învățarea particularităților imagistice medicale. Aduce o îmbunătățire cu 0,4 % a scorului F1 la nivel de studiu față de pregătirea puternică Imagenet21K [28].

Etapa a treia: arhitectură dual-hop pentru detectarea EP. Arhitectura dual-hop realizează clasificarea în două faze. Prima fază realizează o evaluare inițială, iar a doua fază, având acces la datele de intrare inițiale, precum și la datele de ieșire primei faze, este capabilă să producă o predicție mai precisă și mai rafinată. Această îmbunătățire arhitecturală are avantajul cheie de a oferi caracteristici globale extrem de rafinate la începutul fazelor ulterioare. Această abordare simulează de la început un câmp receptiv foarte mare, care, combinat cu caracteristicile rafinate, obține un scor F1 de 1,6 % la îmbunătățirea nivelului de studiu față de arhitectura single hop.

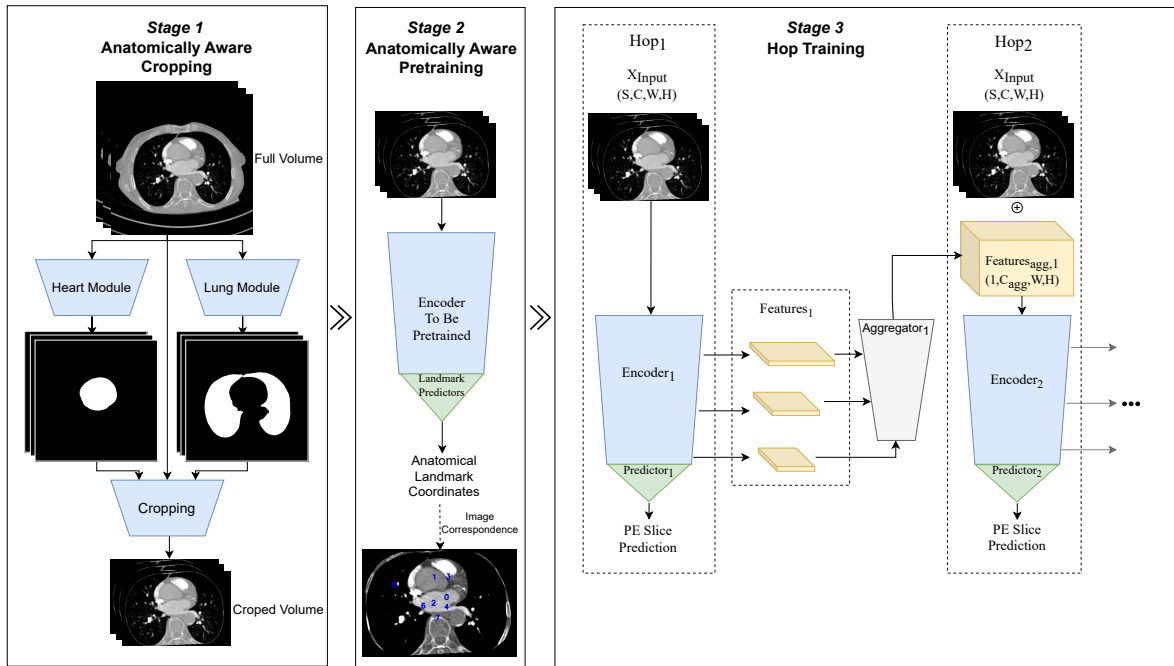


Figura 2.1: Flux de lucru propus: fiecare etapă reprezintă una dintre contribuții. **Stanga:** Segmentare și decupare anatomică, prin care datele sunt specializate pentru detectarea EP. **Centru:** Preantrenarea anatomică cu privire la sarcina aferentă de detectare a reperelor anatomice, prin care modelul este pregătit pentru sarcina noastră de detectare a emboliei pulmonare. **Dreapta** Antrenarea Hop, prin care predicțiile modelului sunt rafinate, peste două Hops de procesare neuronală.

Tabela 2.1: Compararea rezultatelor cu soluții de ultimă generație [1, 2, 39]. Modelul nostru, în comparație cu cele mai moderne [1, 2, 39], obține rezultate mai bune în ceea ce privește scorul F1. Modelul nostru funcționează puțin mai rău în comparație cu un radiolog uman.

Model	F1(↑)	AUC PR(↑)
Weiker [39] AI Model	86.0 %	-
Cheikh [2] AI Model	86.1 %	-
<i>Cheikh [2] Radiologist</i>	92.4 %	-
Buls [1] AI model	73.0 %	-
Modelul nostru de baza	88.4 %	63.87 %
+ Segmentare anatomica	89.0 %	64.29 %
+ Preantrenare	89.4 %	65.95 %
+ Hop (Modelul nostru final)	91.0 %	66.21 %

Modelul nostru final obține performanțe de ultimă generație, obținând un scor F1 la nivel de studiu de 91,0%, depășind alte metode puternice, cum ar fi scorul F1 al lui Weikert el al [38] 86,0% F1. Rezultatele complete sunt afișate în tabelul 2.1. Îmbunătățirile semnificative ale performanței față de standardul de bază puternic și metodele de ultimă generație demonstrează eficacitatea abordării noastre, chiar și în regimul bogat de date. Performanța modelului nostru, care se apropie de cea a radiologilor experți, sugerează potențialul său de aplicare clinică.

Mai mult, cadrul în trei etape introdus aici reprezintă o dovadă generală de concept care ar putea fi adaptată altor sarcini de analiză a imaginilor medicale. Prin descompunerea proceselor complexe de diagnosticare în pași intuitivi, secvențiali, am creat un model care nu numai că funcționează bine, ci și se aliniază cu raționamentul diagnosticului uman.

Capitolul 3

ÎNVĂȚAREA PROFUNDĂ DIN PERSPECTIVA CLINICĂ

Acest studiu a evaluat algoritmul nostru bazat pe rețeaua neuronală profundă (DNN) pentru detectarea automată a emboliei pulmonare (EP) pe scanările CT angiografice pulmonară (CTPA). EP reprezintă o afecțiune gravă clasată pe locul al treilea cel mai frecvent sindrom cardiovascular la nivel mondial, diagnosticul prompt fiind esențial pentru a preveni complicațiile care pun viața în pericol. Algoritmul a fost testat pe 903 scanări CTPA consecutive de la un singur spital universitar, cu o prevalență EP de 12,2% (n=110).

Din punct de vedere clinic, diagnosticul EP prezintă provocări semnificative. Radiologii se confruntă cu volume de lucru în creștere, scanările CTPA crescând cu 3-4% anual din cauza fricii de a rata un caz de EP, accesibilității crescute la CT și considerentelor financiare [36]. În situațiile de urgență, mai multe afecțiuni care pun viața în pericol trebuie excluse alături de EP, inclusiv sindromul coronarian acut, disecția aortică și pneumonia. Pacienții oncologici se confruntă cu o incidență crescută a EP din cauza trombozei asociate malignității, ceea ce face ca depistarea precoce să fie crucială pentru un management optim.

Algoritmul bazat pe DNN a demonstrat o performanță puternică, cu o sensibilitate de 84,6%, o specificitate de 95,1% și o precizie de 93,8% pentru detectarea EP. Această performanță se aliniază cu sensibilitatea tipică a radiologilor (67%-87%) și specificitatea (89%-99%) atunci când citirea consensului este considerată standardul de referință. Rezultatele deta-

Această secțiune se bazează pe munca noastră în: Tilman Emrich, Florin Condrea, et al "Accuracy of a Deep Neural Network for Automated Pulmonary Embolism Detection on Dedicated CT Pulmonary Angiograms"

Tabela 3.1: Analiza generală a performanței modelului AI. Intervalele de încredere de 95% sunt între paranteze pătrate. Toate cazurile $n = 903$.

Metrică	Valoare	95% CI
Adevărat pozitiv	93	-
Fals Negativ	17	-
Fals Pozitiv	39	-
Adevărat negativ	754	-
Sensibilitate	84,6%	[76,4%, 90,7%]
Specificitate	95,1%	[93,3%, 96,5%]
Valoare predictivă pozitivă	70,5%	[63,5%, 76,6%]
Valoare predictivă negativă	97,8%	[96,6%, 98,6%]
Precizie	93,8%	[92,0%, 95,3%]

liate sunt prezentate în tabelul 3.1

Analiza specifică locației a evidențiat o sensibilitate excelentă de 100% pentru EP central și de 96,7% pentru EP lobar, în timp ce detecția EP periferică a fost mai mică la 72,9%. Acest model reflectă experiența clinică, în care radiologii demonstrează, de asemenea, o precizie mai mare pentru EP centrală decât periferică (74%-86%) [18]. Fals negative ($n = 17$) au apărut în principal cu EP cronice mici în arterele subsegmentare și EP acute în ramurile segmentare sau subsegmentare. Fals pozitive ($n=39$) au fost cauzate în principal de fluxul turbulent, artefacte, consolidare și mase pulmonare. Rezultatele detaliate sunt prezentate în tabelul 3.2.

Din punct de vedere clinic, algoritmul a funcționat în mod similar la pacienții obezi și non-obezi, demonstrându-și robustețea în diferite populații de pacienți, în ciuda provocărilor legate de calitatea redusă a imaginii la pacienții cu indice de masă corporală mai mare din cauza zgomotului crescut, rezoluției spațiale reduse și artefactelor de întărire a fascicului. Populația studiată a reflectat în mod corespunzător scenariile clinice din lumea reală, pacienții cu EP pozitivi prezentând proporții mai mari de factori de risc cunoscuți, inclusiv mobilitate limitată, malignitate activă, intervenții chirurgicale recente și tulburări de coagulare.

Tabela 3.2: Analiza performanței bazată pe locație. Intervalele de încredere de 95% sunt între paranteze pătrate.

Locație	Adevărat pozitiv	Fals Negativ	Sensibilitate
Total	93	17	84,6% [76,4%, 90,7%]
Central	21	0	100,0% [83,9%, 100,0%]
Lobar	29	1	96,7% [82,8%, 99,9%]
Periferic	43	16	72,9% [59,7%, 83,6%]

Aceste rezultate sugerează utilitatea potențială a algoritmului în sprijinirea radiologilor prin prioritizarea examenului și ca al doilea cititor, potențial îmbunătățind eficiența fluxului de lucru și acuratețea diagnosticului în practica clinică. O astfel de integrare AI ar putea răspunde cererii tot mai mari de interpretare CTPA, menținând în același timp standarde înalte de diagnostic.

Capitolul 4

INVĂȚAREA SLAB SUPERVIZATA VIA CREȘTEREA GRANULARĂȚI ETICHETELOR

În acest capitol prezentăm o metodă de învățare slab supravegheată numită „Label Up” care îmbunătățește granularitatea adnotărilor de la nivel de imagine la nivel de pixel pentru detectarea emboliei pulmonare (EP). Metoda are trei etape, afișată în figura 4.1 antrenează mai întâi un clasificator la nivel de imagine, apoi utilizează un modul de explicabilitate bazat pe gradienti integrați [30] pentru a genera pseudo-etichete prin rafinare iterativă și, în final, antrenează un model de segmentare pe aceste pseudo-etichete.

Clasificatorul nostru la nivel de imagine, bazat pe munca noastră anterioară [6], atinge performanțe comparabile cu stadiul actual al cunoașterii cu 97,1% AUCROC și 75,1% scor F1 la performanța la nivel de imagine, oferind o bază solidă pentru etapa de explicabilitate. A doua etapă utilizează un proces inovator de rafinare iterativ în care regiunile EP prezise sunt mascate progresiv, forțând modelul să identifice embolii suplimentare. Această abordare îmbunătățește semnificativ sensibilitatea cu 10,1%, menținând în același timp o precizie rezonabilă, cu un vârf de performanță după aproximativ șase iterații de rafinare. Procesul se termină automat când nu sunt detectate alte EP, optimizând eficiența de calcul. Filtrele de post-procesare bazate pe dimensiunea și locația anatomică îmbunătățesc și mai

Această secțiune se bazează pe munca noastră în: Condrea, Florin, Saikiran Rapaka, and Marius Leordeanu. "Label up: Learning Pulmonary Embolism Segmentation from Image Level Annotation through Model Explainability." arXiv preprint arXiv:2412.07384 (2024).

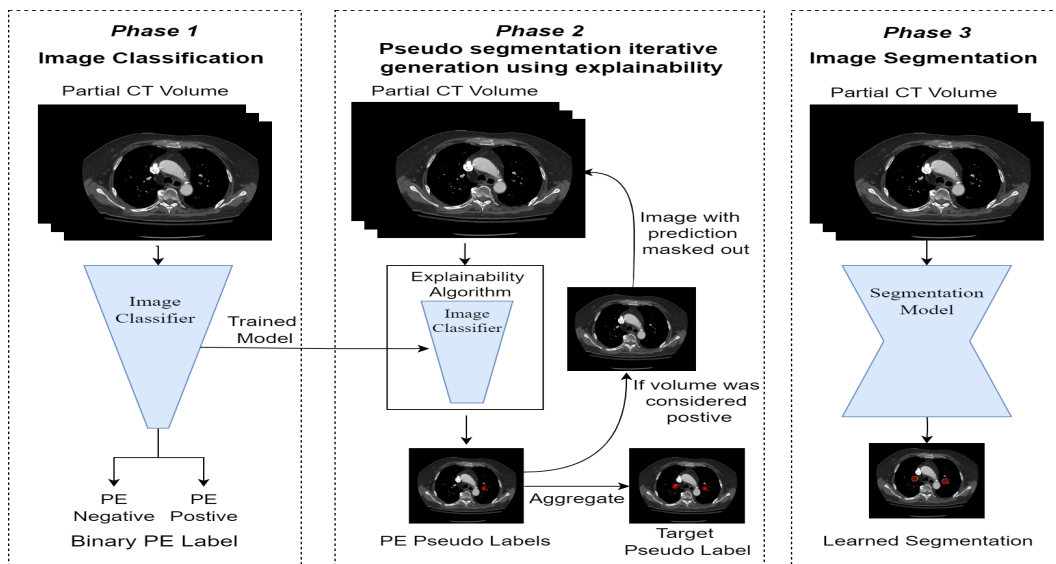


Figura 4.1: Metoda noastră pseudo-etichetare propusă în mai multe etape. În etapa 1, un clasificator la nivel de imagine este antrenat pentru a prezice dacă un teanc de imagini este pozitiv. În etapa 2, clasificatorul antrenat, împreună cu algoritmul de gradienti integrați, prezice locația EP. După fiecare iterație, zonele prezise sunt mascate și volumul introdus înapoi în rețea. Procesul se repetă până când EP nu mai sunt detectate și toate predicțiile sunt agregate în pseudo-etichete. În etapa 3, o rețea de segmentare a învățării profunde este antrenată pe pseudo-etichete, obținând performanțe superioare celei ale modelului profesorului bazat pe explicabilitate.

mult performanța prin eliminarea fals pozitive.

În comparație cu alte metode slab supravegheate privind localizarea EP, metoda noastră instruită pe un set de date de 6.824 de studii [5] atinge un scor F1 de 69,4% la localizarea EP pe un set de date publice de 442 de studii[?], care depășește puțin scorul F1 de 69.1 % al lui Pu et al. care este antrenat pe cantități similare de date[26] .Adăugând date suplimentare la setul nostru de date de antrenare, obținând un total de 13.329 de studii, performanța noastră se îmbunătățește la 71,6% scor F1, potrivit metode puternic supravegheate precum Zhu et al. [41] în ciuda faptului că folosește doar adnotări la nivel de imagine. Rezultatele complete sunt prezentate în tabelul 4.1.

În plus, antrenarea în continuare a modelului nostru slab supravegheat pe doar 111 studii adnotate uman produce un scor F1 de 75,5% la localizarea EP, depășind atât inițializarea aleatorie (71,0%), cât și metoda lui Zhu et al. (71,6%) [41] cu marje semnificative. Acest lucru demonstrează că conducta noastră slab supravegheată nu numai că generează pseudo-etichete de înaltă calitate, dar servește și ca o strategie eficientă de antrenament atunci când devin disponibile adnotări umane limitate. Rezultatele complete sunt prezentate în tabelul 4.2.

Tabela 4.1: Performanța modelului de învățare profundă. Raportăm performanța noastră pe două seturi de date de 6.824 și 13.329 de studii CTPA. Comparând cu alte metode de pseudo-etichetă [26] antrenate pe un set de date similar, obținem performanțe puțin mai bune. Adăugând mai multe date private la setul de antrenare, se obține performanțe similare cu metodele puternic supravegheate [41] evaluate folosind aceleași criterii de potrivire de detecție.

Model	Date	Pozitive	Recall	PPV	F1
Metode slab supervizate					
Pu et al [26]	6415	1990	61.8	78.2	69.1
Our Model	6824	1766	61.5	79.6	69.4
Our Model	13329	3389	66.9	77.0	71.6
Metode puternic supervizate					
Ozkan et al [24]	142	142	95.1	52.6	67.7
Tajbakhsh et al [31]	121	121	83.4	47.2	60.3
Tajbakhsh et al [32]	121	121	32.9	98.6	49.4
Xu et al [40]	113	113	93.2	51.2	66.1
Zhu et al [41]	142	142	86.0	61.3	71.6
Weikert et al [39]	30000	15858	82.2	86.8	85.8

Tabela 4.2: Rezultatele antrenării puternic supervizate. Ne antrenăm modelul pe adnotarea umană, pornind atât de la inițializarea aleatorie, cât și de la greutatea prea antrenate slab supervizate. Reglarea fină a modelului antrenat pe pseudo-etichete obține o creștere a performanței de 3,9 % scor F1, îmbunătățind atât sensibilitatea, cât și PPV. Modelul ajustat depășește, de asemenea, inițializarea aleatoare cu 4,5 % F1 și o altă metodă puternic supravegheată antrenată pe o cantitate similară de date [41]. Modelul nostru WSL cu caractere cursive este antrenat pe pseudo-etichete și adăugat pentru o vizualizare mai ușoară a îmbunătățirii obținute prin ajustare.

Model	Date	Pozitive	Sensitivitate	PPV	F1
Modelul nostru WSL de bază	13,329	3,389	66.9	77.0	71.6
Zhu et al [41]	142	142	86.0	61.3	71.6
Modelul nostru inițiat random	111	111	66.2	76.6	71.0
Modelul nostru WSL finetuned	111	111	73.9	77.5	75.5

Abordarea noastră abordează o provocare critică în IA medicală: deficitul de adnotări cu granulație fină. Prin valorificarea etichetelor la nivel de imagine ușor disponibile pentru a genera măști de segmentare detaliate, permitem dezvoltarea unor instrumente de diagnosticare mai precise fără costul prohibitiv al adnotării manuale la nivel de pixel. Modelul final de segmentare este, de asemenea, substanțial mai rapid (50x) decât modelul profesorul bazat pe explicabilitate, obținând în același timp o performanță mai bună, făcându-l potrivit pentru implementarea clinică. Această metodologie are aplicații potențiale largi dincolo de detectarea EP, oferind un cadru generalizabil pentru îmbunătățirea granularității adnotărilor în diferite sarcini de imagistică medicală.

Capitolul 5

ÎNVĂȚAREA SARCINILOR TEMPORALE FĂRĂ ETICHETE UMANE

În acest capitol, explorăm o abordare inovatoare a detectării semnelor vitale în videoclipurile termice prin valorificarea datelor sintetice construite din primitive de viziune computerizată de bază. Cercetarea demonstrează modul în care rețelele neuronale pot recunoaște semnale temporale specifice, în ciuda diferențelor semnificative dintre domeniile de antrenament și țintă. Metoda noastră, VSignNet, învață să prezică atât intensitatea semnelor vitale, cât și regiunile de interes corespunzătoare în videoclipurile termice fără supraveghere umană. Metoda obține rezultate de ultimă generație pe setul de date recent LCAS [8].

Abordarea noastră constă din patru etape: determinarea intervalului de frecvență de interes, generarea de date sintetice de antrenament, antrenamentul modelului și efectuarea inferenței cu post-procesare pentru a detecta vârfurile de semnal. Metodologia de generare a datelor sintetice se bazează în întregime pe principii euristice, creând mostre de antrenament care împărtășesc doar caracteristicile esențiale ale semnalului temporal țintă. Arhitectura noastră de învățare profundă surprinde dimensiunile temporale pe două niveluri - local și global - permițând detectarea eficientă a semnelor vitale. Ca prezentare generală a conduitei de ieșire ilustrată în figura 5.1,

Acest capitol se bazează pe munca noastră din: Condrea, Florin, Victor-Andrei Ivan, and Marius Leordeanu. "In search of life: Learning from synthetic data to detect vital signs in videos." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. 2020.

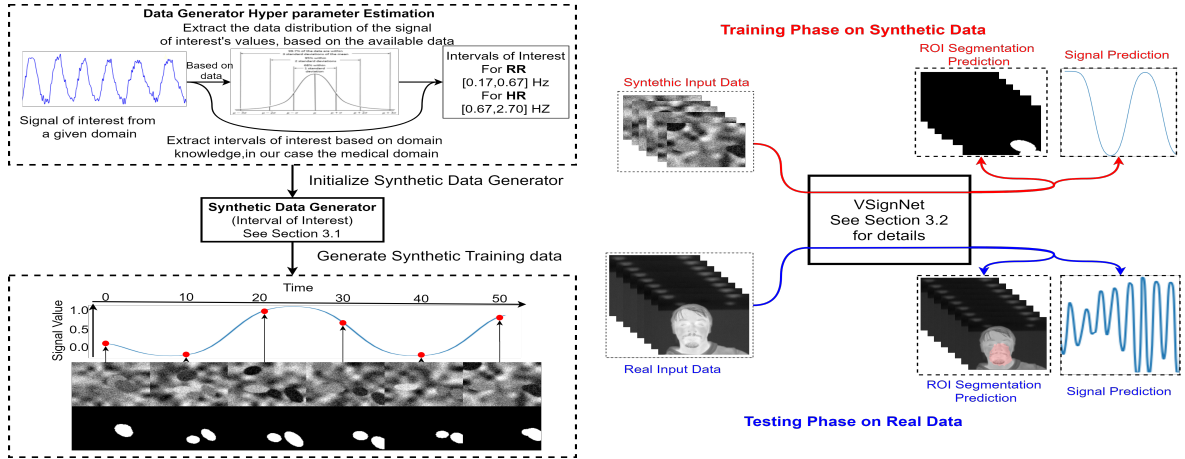


Figura 5.1: Cele patru etape ale abordării noastre: prima etapă constă în determinarea intervalului de frecvență de interes, fie prin utilizarea statisticilor setului de date, fie a cunoștințelor de domeniu. A doua etapă este generarea de mostre sintetice de date de antrenament. A treia etapă este formarea modelului. La etapa finală, de inferență, o etapă suplimentară de post-procesare detectează vârfurile semnalului, pe care ne bazăm rata medie a semnalului de ieșire.

Arhitectura VSigNet procesează cadrele de intrare printr-un encoder convoluțional temporal, urmat de LSTM-uri bidirecționale care agregă informații temporale globale. Modulele complet conectate transformă înglobările în valori numerice reprezentând mărimea semnelor vitale, în timp ce decodoarele convoluționale temporale generează hărți termice care codifică locațiile surselor de semnal. Această abordare cu două căi permite atât predicția intensității semnalului, cât și localizarea regiunii de interes.

Ritm cardiac (BPM)				
Experiment	Still		In miscare	
	MAE	STD	MAE	STD
LCAS [8]	29.68	± 15.76	18.96	± 22.51
VSigNet	15.51	± 9.93	14.91	± 7.99
Frecvența respirației (RPM)				
Experiment	Still		In miscare	
	MAE	STD	MAE	STD
LCAS [8]	3.72	± 0.78	5.87	± 2.18
M ROI	1.87	± 2.05	4.41	± 4.41
RF ROI	1.90	± 1.72	14.77	± 7.32
VSigNet	1.12	± 1.34	2.62	± 2.07

Tabela 5.1: Comparația performanței dintre rezultatele noastre (VSigNet) și rezultatele raportate de Cosar et al (LCAS [8]). De asemenea, raportăm rezultatele cele două standarde de baza, utilizând fie ROI adnotat manual (M ROI), fie ROI detectat cu RetinaFace (RF ROI).

VSigNet depășește metodele existente [4, 8, 22], obținând o eroare absolută medie de 1,12 ($\pm 1,34$) pentru rata de respirație pe minut în posturile nemișcate și 2,62 ($\pm 2,07$) în tim-

pul mișcării, semnificativ mai bună decât linia de bază LCAS și $2 \pm 5.1.87$. respectiv). Pentru detectarea ritmului cardiac, metoda noastră realizează o eroare medie absolută de 15,51 ($\pm 9,93$) de bătăi ale inimii pe minut în timpul pozițiilor nemișcate și 14,91 ($\pm 7,99$) în timpul mișcării, comparativ cu LCAS 29,68 ($\pm 15,76$) și 18,96 ($\pm 22,51$). Modelul demonstrează, de asemenea, robustețe în diferite modele de respirație și localizează în mod eficient regiunile de interes corespunzătoare sursei de semnal. Rezultatele complete sunt afișate în tabelul 5.1,

Eficacitatea remarcabilă a algoritmului nostru relativ simplu de generare de date sintetice în domeniul complex al imagisticii medicale merită investigații suplimentare. Aplicarea cu succes a unei astfel de metodologii ridică întrebări semnificative cu privire la potențialul paradigmelor de antrenament sintetic de a facilita învățarea nesupravegheată în sarcini de viziune spațio-temporală diverse și complexe [12, 21, 23]. Această descoperire sugerează căi promițătoare pentru abordarea provocărilor legate de deficitul de date în domeniile specializate în care seturile de date adnotate sunt deosebit de dificil de obținut.

Capitolul 6

CONCLUZII

Această teză a explorat soluții inovatoare pentru a aborda provocările cheie din IA medicală, concentrându-se pe îmbunătățirea performanței modelului în condiții de constrângeri de date, menținând în același timp relevanța clinică. Munca noastră a contribuit la analiza imaginilor medicale și la monitorizarea semnelor vitale prin design informat anatomic, generare de date sintetice și arhitecturi care folosesc atât caracteristici globale, cât și locale. Am demonstrat eficacitatea designului inspirat anatomic în cadrul nostru HopNet, care a folosit antecedente anatomice pentru a ghida preprocesarea și preformarea modelului, rezultând o învățare mai eficientă a caracteristicilor și o generalizare îmbunătățită.

Pentru a aborda deficitul de date, am propus două abordări complementare. Metodologia noastră Label-Up a crescut granularitatea adnotărilor fără etichetare manuală suplimentară, valorificând o supraveghere slabă și o rafinare iterativă. În activitatea noastră de monitorizare a semnelor vitale, am dezvoltat o metodă pentru generarea de date sintetice de antrenament bazate pe dinamica temporală țintă, creând primitive geometrice simplificate care au capturat modele fundamentale pentru detectarea precisă a ritmului respirator și cardiac. Ambele abordări au redus semnificativ sarcina de adnotare, menținând în același timp o acuratețe ridicată a diagnosticului.

Dincolo de inovațiile tehnice, am subliniat validarea clinică și interpretabilitatea. Colaborarea noastră cu profesioniștii medicali a oferit informații despre punctele forte și limitările modelului în scenarii realiste, în special în ceea ce privește anomaliile provocatoare. Această metodologie de evaluare informată clinic reprezintă un pas important către dezvoltarea sistemelor AI care pot sprijini în mod semnificativ luarea deciziilor medicale în practică.

Bibliografie

- [1] Buls, N., Watté, N., Nieboer, K., Ilse, B. and de Mey, J. [2021], ‘Performance of an artificial intelligence tool with real-time clinical workflow integration–detection of intracranial hemorrhage and pulmonary embolism’, *Physica Medica: European Journal of Medical Physics* **83**, 154–160.
- [2] Cheikh, A. B., Gorincour, G., Nivet, H., May, J., Seux, M., Calame, P., Thomson, V., Delabrousse, E. and Crombé, A. [2022], ‘How artificial intelligence improves radiological interpretation in suspected pulmonary embolism’, *European Radiology* pp. 1–12.
- [3] Chen, I. Y., Pierson, E., Rose, S., Joshi, S., Ferryman, K. and Ghassemi, M. [2021], ‘Ethical machine learning in healthcare’, *Annual Review of Biomedical Data Science* **4**, 123–144.
- [4] Cho, Y., Julier, S. J., Marquardt, N. and Bianchi-Berhouze, N. [2017], ‘Robust tracking of respiratory rate in high-dynamic range scenes using mobile thermal imaging’, *Biomedical Optics Express* **8**(10), 4480–4503.
- [5] Colak, E., Kitamura, F. C., Hobbs, S. B., Wu, C. C., Lungren, M. P., Prevedello, L. M., Kalpathy-Cramer, J., Ball, R. L., Shih, G., Stein, A., Halabi, S. S., Altinmakas, E., Law, M., Kumar, P., Manzalawi, K. A., Nelson Rubio, D. C., Sechrist, J. W., Germaine, P., Lopez, E. C., Amerio, T., Gupta, P., Jain, M., Kay, F. U., Lin, C. T., Sen, S., Revels, J. W., Brussaard, C. C., Mongan, J., Abdala, N., Bearce, B., Carrete, H., Dogan, H., Huang, S.-C., Crivellaro, P., Dincler, S., Kavnoudias, H., Lee, R., Lin, H.-M., Salehinejad, H., Samorodova, O., Rodrigues dos Santos, E., Seah, J., Zia, A., Arteaga, V. A., Batra, K., Castelli von Atzingen, A., Chacko, A., DiDomenico, P. B., Gill, R. R., Hafez, M. A., John, S., Karl, R. L., Kanne, J. P., Mathilakath Nair, R. V., McDermott, S., Mittal, P. K., Mumbower, A., Lee, C., Orausclio, P. J., Palacio, D., Pozzessere, C., Rajiah, P., Ramos, O. A., Rodriguez, S., Shaaban, M. N., Shah, P. N., Son, H., Sonavane,

-
- S. K., Spieler, B., Tsai, E., Vásquez, A., Vijayakumar, D., Wali, P. P., Wand, A. and Zamora Endara, G. E. [2021], ‘The rsna pulmonary embolism ct dataset’, *Radiology: Artificial Intelligence* **3**(2), e200254. PMID: 33937862.
URL: <https://doi.org/10.1148/ryai.2021200254>
- [6] Condrea, F., Rapaka, S., Itu, L., Sharma, P., Sperl, J., Ali, A. M. and Leordeanu, M. [2023], ‘Anatomically aware dual-hop learning for pulmonary embolism detection in ct pulmonary angiograms’, *arXiv preprint arXiv:2303.17593* .
- [7] Condrea, F., Rapaka, S. and Leordeanu, M. [2024], ‘Label up: Learning pulmonary embolism segmentation from image level annotation through model explainability’, *arXiv preprint arXiv:2412.07384* .
- [8] Cosar, S., Yan, Z., Zhao, F., Lambrou, T., Yue, S. and Bellotto, N. [2018], Thermal camera based physiological monitoring with an assistive robot, *in* ‘IEEE International Engineering in Medicine and Biology Conference’.
- [9] Davenport, T. and Kalakota, R. [2019], ‘The potential for artificial intelligence in healthcare’, *Future Healthcare Journal* **6**(2), 94–98.
- [10] Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., Corrado, G., Thrun, S. and Dean, J. [2019], ‘A guide to deep learning in healthcare’, *Nature Medicine* **25**(1), 24–29.
- [11] Gichoya, J. W., Banerjee, I., Bhimireddy, A. R., Burns, J. L., Celi, L. A., Chen, L.-C., Correa, R., Dullerud, N., Ghassemi, M., Huang, S.-C., Kuo, P.-C., Lungren, M., Palmer, L., Price, B. J., Purkayastha, S., Pyrros, A., Oakden-Rayner, L., Okechukwu, C., Seyyed-Kalantari, L., Trivedi, H., Wang, R., Zaiman, Z. and Zhang, H. [2022], ‘Ai recognition of patient race in medical imaging: a modelling study’, *The Lancet Digital Health* **4**(6), e406–e414.
- [12] Gupta, A., Vedaldi, A. and Zisserman, A. [2016], Synthetic data for text localisation in natural images, *in* ‘Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)’.
- [13] He, J., Baxter, S. L., Xu, J., Xu, J., Zhou, X. and Zhang, K. [2019], ‘The practical

-
- implementation of artificial intelligence technologies in medicine’, *Nature Medicine* **25**(1), 30–36.
- [14] Holzinger, A., Langs, G., Denk, H., Zatloukal, K. and M"uller, H. [2019], ‘Causability and explainability of artificial intelligence in medicine’, *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* **9**(4), e1312.
- [15] Johnson, A. E. W., Pollard, T. J. and Mark, R. G. [2018], ‘Reproducibility in critical care: a mortality prediction case study’, *Proceedings of Machine Learning Research* **85**, 361–376.
- [16] Kaissis, G. A., Makowski, M. R., R"uckert, D. and Braren, R. F. [2020], ‘Secure, privacy-preserving and federated machine learning in medical imaging’, *Nature Machine Intelligence* **2**(6), 305–311.
- [17] Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G. and King, D. [2019], ‘Key challenges for delivering clinical impact with artificial intelligence’, *BMC Medicine* **17**(1), 195.
- [18] Kligerman, S. J., Mitchell, J. W., Sechrist, J. W., Meeks, A. K., Galvin, J. R. and White, C. S. [2018], ‘Radiologist performance in the detection of pulmonary embolism: features that favor correct interpretation and risk factors for errors’, *Journal of thoracic imaging* **33**(6), 350–357.
- [19] Larrazabal, A. J., Nieto, N., Peterson, V., Milone, D. H. and Ferrante, E. [2020], ‘Gender imbalance in medical imaging datasets produces biased classifiers for computer-aided diagnosis’, *Proceedings of the National Academy of Sciences* **117**(23), 12592–12594.
- [20] Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A., van Ginneken, B. and Sánchez, C. I. [2017], ‘A survey on deep learning in medical image analysis’, *Medical Image Analysis* **42**, 60–88.
- [21] Mayer, N., Ilg, E., Fischer, P. et al. [2018], ‘What makes good synthetic training data for learning disparity and optical flow estimation?’, *International Journal of Computer Vision* **126**, 942–960.

-
- [22] Mutlu, K., Esquivelzeta Rabell, J., Martin del Olmo, P. and Haesler, S. [2018], ‘Infrared thermography-based monitoring of respiration phase without image segmentation’, *Journal of Neuroscience Methods* **301**, 1–8.
- [23] Oh, T., Jaroensri, R., Kim, C., Elgharib, M., Durand, F., Freeman, W. and Matusik, W. [2018], ‘Learning-based video motion magnification’, *arXiv preprint arXiv:1804.02684*.
- [24] Özkan, H., Osman, O., Şahin, S. and Boz, A. F. [2014], ‘A novel method for pulmonary embolism detection in cta images’, *Computer methods and programs in biomedicine* **113**(3), 757–766.
- [25] Price, W. N. and Cohen, I. G. [2019], ‘Privacy in the age of medical big data’, *Nature Medicine* **25**(1), 37–43.
- [26] Pu, J., Gezer, N. S., Ren, S., Alpaydin, A. O., Avci, E. R., Risbano, M. G., Rivera-Lebron, B., Chan, S. Y.-W. and Leader, J. K. [2023], ‘Automated detection and segmentation of pulmonary embolisms on computed tomography pulmonary angiography (ctpa) using deep learning but without manual outlining’, *Medical Image Analysis* **89**, 102882.
- [27] Rajpurkar, P., Chen, E., Banerjee, O. and Topol, E. J. [2022], ‘Ai in health and medicine’, *Nature Medicine* **28**(1), 31–38.
- [28] Ridnik, T., Ben-Baruch, E., Noy, A. and Zelnik-Manor, L. [2021], ‘Imagenet-21k pre-training for the masses’, *arXiv preprint arXiv:2104.10972*.
- [29] Rieke, N., Hancox, J., Li, W., Milletari, F., Roth, H. R., Albarqouni, S., Bakas, S., Galtier, M. N., Landman, B. A., Maier-Hein, K., Ourselin, S., Sheller, M., Summers, R. M., Trask, A., Xu, D., Baust, M. and Cardoso, M. J. [2020], ‘The future of digital health with federated learning’, *npj Digital Medicine* **3**(1), 119.
- [30] Sundararajan, M., Taly, A. and Yan, Q. [2017], Axiomatic attribution for deep networks, in ‘International conference on machine learning’, PMLR, pp. 3319–3328.
- [31] Tajbakhsh, N., Gotway, M. B. and Liang, J. [2015], Computer-aided pulmonary embolism detection using a novel vessel-aligned multi-planar image representation and

-
- convolutional neural networks, in ‘Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part II 18’, Springer, pp. 62–69.
- [32] Tajbakhsh, N. et al. [2019], ‘Computer-aided detection and visualization of pulmonary embolism using a novel, compact, and discriminative image representation’, *Medical Image Analysis* **58**, 101541.
- [33] Tan, M. and Le, Q. [2021], Efficientnetv2: Smaller models and faster training, in ‘International Conference on Machine Learning’, PMLR, pp. 10096–10106.
- [34] Tjoa, E. and Guan, C. [2020], ‘A survey on explainable artificial intelligence (xai)’, *IEEE Access* .
- [35] Topol, E. J. [2019], ‘High-performance medicine: the convergence of human and artificial intelligence’, *Nature Medicine* **25**(1), 44–56.
- [36] Wang, R. C., Miglioretti, D. L., Marlow, E. C., Kwan, M. L., Theis, M. K., Bowles, E. J., Greenlee, R. T., Rahm, A. K., Stout, N. K., Weinmann, S. et al. [2020], ‘Trends in imaging for suspected pulmonary embolism across us health care systems, 2004 to 2016’, *JAMA network open* **3**(11), e2026930–e2026930.
- [37] Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M. and Summers, R. M. [2020], ‘Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases’, *IEEE Transactions on Medical Imaging* **38**(7), 1588–1601.
- [38] Weikert, T., Winkel, D. J., Bremerich, J., Stieltjes, B., Parmar, V., Sauter, A. W. and Sommer, G. [2020a], ‘Automated detection of pulmonary embolism in ct pulmonary angiograms using an ai-powered algorithm’, *European radiology* **30**, 6545–6553.
- [39] Weikert, T., Winkel, D. J., Bremerich, J., Stieltjes, B., Parmar, V., Sauter, A. W. and Sommer, G. [2020b], ‘Automated detection of pulmonary embolism in ct pulmonary angiograms using an ai-powered algorithm’, *European radiology* **30**, 6545–6553.
- URL:** <https://pubmed.ncbi.nlm.nih.gov/32621243/>

-
- [40] Xu, H., Li, H., Xu, Q., Zhang, Z., Wang, P., Li, D. and Guo, L. [2023], ‘Automatic detection of pulmonary embolism in computed tomography pulmonary angiography using scaled-yolov4’, *Medical Physics* **50**(7), 4340–4350.
- [41] Zhu, H., Tao, G., Jiang, Y., Sun, L., Chen, J., Guo, J., Wang, N., Wei, H., Liu, X., Chen, Y. et al. [2024], ‘Automatic detection of pulmonary embolism on computed tomography pulmonary angiogram scan using a three-dimensional convolutional neural network’, *European Journal of Radiology* p. 111586.