

# Impactul preprocesării datelor asupra modelelor de inteligență artificială utilizate în diagnosticare

## 1. Abstract

Într-o lume unde tehnologia evoluează într-un ritm accelerat, inteligența artificială (AI) devine o componentă esențială în transformarea domeniului sănătății, în special în diagnosticul medical. Acest articol explorează nu doar inovațiile aduse de AI în acest sector, ci și impactul profund al acestor tehnologii asupra preciziei, eficienței și accesibilității serviciilor medicale.

AI redefinește standardele de diagnosticare prin reducerea erorilor umane și îmbunătățirea acurateței rezultatelor. Spre exemplu, utilizarea tehnicilor avansate, precum rețelele neuronale convoluționale, permite identificarea cu precizie a anomaliilor și a tiparelor complexe din imagini medicale, depășind metodele tradiționale. Aceste progrese facilitează depistarea precoce a bolilor grave, cum ar fi cancerul, și contribuie direct la salvarea de vieți.

De asemenea, AI automatizează procesele de diagnosticare, reducând semnificativ timpul de obținere a rezultatelor. Acest lucru permite medicilor să-și concentreze atenția pe cazurile critice, optimizând în același timp fluxurile de lucru din instituțiile medicale și sporind eficiența sistemului de sănătate.

Un punct central al articolului este rolul preprocesării datelor în succesul modelelor de AI. Calitatea datelor utilizate este esențială, iar metode precum normalizarea, reducerea zgomotului și augmentarea datelor sunt fundamentale pentru construirea unor algoritmi fiabili. Fără aceste etape, modelele de AI pot oferi rezultate eronate, subliniind importanța rigurozității în colectarea și pregătirea datelor.

Articolul își propune să informeze profesioniștii din domeniul sănătății și să crească gradul de conștientizare a pacienților despre modul în care tehnologia transformă îngrijirile medicale. Cititorii vor descoperi provocările și realizările actuale ale AI în diagnosticare, precum și perspectivele viitoare, fiind încurajați să îmbrățișeze aceste inovații în practica clinică.

Lectura acestui material nu este doar o oportunitate de informare, ci și un pas important spre înțelegerea modului în care inteligența artificială revoluționează domeniul sănătății. Prin analizarea provocărilor și a oportunităților oferite de AI, cititorii se vor putea adapta mai bine la schimbările rapide din acest sector vital.

## 2. Clasificare

Din punct de vedere al ACM Computing Classification System, sistemul de clasificare folosit pentru a determina categoria din care un articol referitor la domeniul informatic face parte, această lucrare se încadrează în *Social and professional topics~Computing / technology policy~Medical information policy~Medical technologies*. O categorie cu înclinație socială, aceasta acoperă ingineria medicală, radiografia, inovațiile și îmbunătățirile prezente în domeniul medical.

Conform MSC (Mathematics Subject Classification), orchestrat de AMS, lucrarea se încadrează în categoria **68T01** (Artificial Intelligence - General). Această clasificare este datorată modelelor de inteligență artificială folosite în cadrul diagnosticării imagistice, propuse în acest articol.

## 3. Introducere

### 3.1. Context actual

În contextul avansării rapide a tehnologiilor bazate pe inteligența artificială, diagnosticarea imagistică a devenit una dintre cele mai promițătoare aplicații ale acestor inovații. Abordările tradiționale de diagnosticare, care depind adesea de expertiza umană și de interpretarea subiectivă a imaginilor medicale, pot fi îmbunătățite considerabil prin implementarea soluțiilor bazate pe AI. Acest capitol își propune să exploreze modul în care îmbunătățirea modelelor AI destinate diagnosticării imagistice poate fi realizată prin optimizarea seturilor de date, cu un accent deosebit pe tehnicile de preprocesare. Vom discuta despre importanța curățării datelor, eliminarea zgomotului, aplicarea diferitelor filtre vizuale și utilizarea augmentării datelor pentru a crește robustețea modelelor.

Dezvoltarea unui model AI eficient în diagnosticul medical nu se limitează doar la alegerea unei arhitecturi performante; calitatea datelor de intrare este esențială pentru rezultatele finale. Astfel, strategiile de preprocesare pot influența semnificativ acuratețea și fiabilitatea învățării automate. Prin urmare, este vital să investigăm nu doar seturile de date utilizate pe scară largă, cum ar fi ChestX-ray14, ISIC Archive și LIDC-IDRI, ci și metodologiile care transformă aceste date brute în input-uri optimizate pentru algoritmi de învățare profundă.

În plus, vom analiza cum aceste tehnici de îmbunătățire a datelor contribuie la o evaluare corectă a modelului AI, subliniind măsurile de performanță, cum ar fi AUC-ROC, sensibilitatea și specificitatea, care sunt imperativ de respectat pentru a asigura un diagnostic corect. Prin această abordare, capitolul va aduce o perspectivă cuprinzătoare asupra impactului atribuțiilor de preprocesare asupra eficienței modelelor AI în diagnosticarea imagistică, oferind nu doar o fundamentare teoretică,

ci și o bază solidă pentru implementări practice în domeniul medical.

## 3.2. Abordări din literatura de specialitate

### 3.2.1 Modelul CheXNet

CheXNet este un model bazat pe DenseNet-121, pre-antrenat pe ImageNet și adaptat pentru clasificarea afecțiunilor pulmonare din setul de date ChestX-ray14. Acesta a atins performanțe remarcabile în detectarea pneumoniei, raportând o acuratețe de 76% și o valoare AUC de 0.78 pentru clasificarea binary (pneumonie vs. fărăpneumonie). Spre deosebire de abordarea propusă, CheXNet se concentrează exclusiv pe optimizarea performanței prin fine-tuning complet al modelului.

### 3.2.2 Utilizarea ResNet pentru imagini medicale

Studiile anterioare au arătat că ResNet50 oferă performanțe competitive în clasificarea imaginilor radiologice, cu o rată de acuratețe medie de 92-95% pe diverse subseturi din ChestX-ray14. ResNet este preferat datorită straturilor reziduale, care facilitează antrenarea unor rețele mai adânci. În comparație cu DenseNet, ResNet tinde să fie mai robust la seturi de date reduse, dar necesită o selecție atentă a hiperparametrilor.

### 3.2.3 Transferul de învățare cu ImageNet

Transferul de învățare este o tehnică obișnuită pentru clasificarea imaginilor medicale, reducând necesitatea antrenării de la zero. Modelele precum VGG16, Inception, și ResNet sunt adesea pre-antrenate pe ImageNet, iar straturile lor finale sunt personalizate pentru problemele medicale. Abordarea propusă utilizează un mecanism similar, dar păstrează straturile ResNet50 înghețate, ceea ce simplifică procesul de antrenare și reduce timpul necesar.

### 3.2.4 Diferențe și asemănări

Aspect	Abordarea propusă	Literatura existentă
<b>Model utilizat</b>	ResNet50 pre-antrenat	DenseNet-121, ResNet, VGG16
<b>Fine-tuning</b>	Straturile ResNet înghețate	Fine-tuning complet
<b>Set de date</b>	Subset de 2000 imagini ChestX-ray14	Întreg setul ChestX-ray14
<b>Metrici raportate</b>	Acuratețe: 98%, Val_loss scăzut	Acuratețe: ~92-95%
<b>Efort computațional</b>	Reducere prin înghețarea straturilor	Mai mare datorită fine-tuning-ului

### 3.3. Zone de îmbunătățire

Abordarea propusă demonstrează performanțe excelente în termeni de acuratețe (98%) și pierdere redusă (val\_loss de 0.1824 după 5 epoci), ceea ce este comparabil sau superior multor studii anterioare. Zonele unde se așteaptă rezultate similare sau mai bune includ:

#### 1. Acuratețea clasificării

Cu o acuratețe de 98% pe setul de testare, modelul nostru depășește sau este comparabil cu cele mai bune rezultate raportate de alte modele CNN pe subseturi similare.

#### 2. Eficiența computațională

Congelarea straturilor ResNet50 reduce timpul de antrenare, ceea ce face abordarea noastră mai practică pentru aplicații clinice rapide.

#### 3. Generalizare

Valoarea ridicată a acurateții pe setul de validare (98.67%) sugerează o bună generalizare, în ciuda utilizării unui subset mai mic de date.

### 3.4. Provocări și limite

Totuși, abordarea propusă este limitată de lipsa unui fine-tuning complet al ResNet50, ceea ce ar putea îmbunătăți performanța pe seturi de date mai mari. În plus, un subset mai mare de date ar putea oferi o evaluare mai robustă a modelului și ar reduce riscul de overfitting. Această comparație evidențiază potențialul ResNet50 ca soluție eficientă și performantă pentru clasificarea imaginilor radiologice, oferind rezultate remarcabile chiar și cu o configurare relativ simplă.

## 4. Abordare propusă

### 4.1. Utilizarea datelor și a modelului

#### 4.1.1 Preprocesarea datelor

Preprocesarea imaginilor este un pas fundamental în realizarea de modele robuste de învățare automată, ajutând la îmbunătățirea calității datelor utilizate pentru antrenare.

Acest proces implică mai multe tehnici, esențiale pentru optimizarea datelor de intrare ale modelului ResNet50, următoarele fiind practicate frecvent:

- Normalizarea intensității pixelilor între 0 și 1: Această etapă transformă valorile pixelilor, care inițial se încadrează între 0 și 255, într-un interval mai mic, facilitând o mai bună convergență în timpul antrenamentului modelului. Normalizarea ajută la stabilizarea procesului de învățare, evitând probleme precum

dispariția sau explozia gradientului, care pot duce la o performanță suboptimală a rețelei neuronale.

- Redimensionarea imaginilor la 224x224 pixeli: ResNet50 necesită dimensiuni standardizate ale imaginilor pentru a funcționa corect, iar redimensionarea imaginilor asigură uniformitatea necesară, prevenind astfel erorile în procesul de predicție. Redimensionarea imaginii la 224x224 pixeli permite modelului să extragă caracteristici relevante din imagini, păstrând totodată detaliile importante care influențează rezultatele finale.

- Aplicarea unui filtru alb-negru: Transformarea imaginilor color în alb-negru (grayscale) contribuie la simplificarea datelor de intrare și la reducerea dimensiunii informațiilor procesate. Această tehnică poate facilita extragerea caracteristicilor esențiale, îmbunătățind acuratețea predicției, întrucât permite modelului să se concentreze pe geometria și contrastul imaginilor, în loc de detaliile de culoare.

- Reducerea zgomotului: Zgomotul în imagini poate afecta grav performanța modelului AI. Tehnicile de reducere a zgomotului, cum ar fi filtrarea spațială sau transformările în domeniul frecvenței, sunt aplicate pentru a îmbunătăți claritatea imaginilor și a elimina artefactele nedorite care ar putea influența învățarea modelului.

Astfel, prin utilizarea unei combinații de metode de prelucrare a imaginilor, se realizează o îmbunătățire semnificativă a calității datelor, ceea ce se traduce în o performanță mai bună a modelului în sarcini de diagnosticare imagistică.

Aceste tehnici nu doar că optimizează seturile de date pentru utilizarea în modelele AI, dar contribuie și la obținerea unor rezultate mai precise și de încredere în diagnosticarea imagistică, demonstrând astfel importanța preprocesării în obținerea unor modele eficiente de învățare profundă.

### **4.1.2 Prezentarea modelului ResNet50**

Modelul ResNet50 este o arhitectură de rețea neuronală convoluțională (CNN) care utilizează straturi reziduale pentru a facilita antrenarea rețelelor adânci. Pre-antrenat pe setul de date ImageNet, acesta este capabil să învețe caracteristici vizuale avansate, care pot fi transferate către alte domenii, precum analiza imaginilor medicale. În acest studiu, ResNet50 este folosit pentru clasificarea imaginilor radiologice, cu scopul de a identifica prezența pneumoniei.

### 4.1.3 Configurarea ResNet50

Codul utilizează următorii pași pentru configurarea modelului ResNet50:

#### 1. Inițializare și încărcare a greutăților pre-antrenate

Modelul ResNet50 este inițializat cu greutăți pre-antrenate pe ImageNet și configurat să excludă stratul final specific clasificării pentru ImageNet (`include_top=False`). Inputul este definit pentru imagini de dimensiune (224, 224, 3).

#### 2. Congelarea straturilor pre-antrenate

Pentru a păstra caracteristicile învățate în timpul pre-antrenării, toate straturile ResNet50 sunt “înghețate”, ceea ce înseamnă că nu vor fi actualizate în timpul antrenării.

#### 3. Adăugarea unui capăt personalizat (custom head)

Pe baza caracteristicilor extrase de ResNet50, se adaugă un strat Flatten pentru a reduce dimensiunea datelor, urmat de un strat dens (Dense) cu funcția de activare softmax. Acesta este responsabil pentru clasificarea finală între cele două clase: fără pneumonie și pneumonie.

#### 4. Compilarea modelului

Modelul este compilat folosind funcția de pierdere `sparse_categorical_crossentropy`, optimizatorul `adam`, și metrici de evaluare precum `accuracy`.

### 4.1.4 Antrenarea și evaluarea modelului

După configurare, modelul este antrenat pe setul de date de imagini radiologice preprocesate. Datele sunt împărțite în trei subseturi: antrenare, validare și testare, iar antrenarea se face în batch-uri mici de dimensiune 16, pe parcursul a 5 epoci.

## 4.2. Model Matematic și Justificare Teoretică

### 4.2.1 Arhitectura Matematică a Modelului

În contextul învățării profunde, dezvoltarea unui model matematic robust este esențială pentru obținerea unor rezultate precise și relevante. Unul dintre primii pași în definirea acestui model este stabilirea funcției obiectiv, care în esență reprezintă criteriul de optimizare utilizat în antrenarea modelului. În cazul clasificăției binare, o alegere frecventă este funcția de pierdere *Binary Cross-Entropy*, definită matematic ca:

$$L(\hat{y}, y) = -\frac{1}{N} \sum_{i=1}^N [y_i \cdot \log(\hat{y}_i) + (1 - y_i) \cdot \log(1 - \hat{y}_i)]$$

unde  $y$  reprezintă eticheta reală,  $\hat{y}$  reprezintă predicția modelului, iar  $N$  este numărul total de observații.

Arhitectura rețelelor neuronale convoluționale (CNN) și mecanismul de auto-atenție implementat în Vision Transformers (ViT) sunt descrise printr-o serie de operații matematice specifice. Straturile convoluționale pot fi exprimate prin următoarea formulă:

$$A[i, j, k] = \sum_m \sum_n \sum_c X[i + m, j + n, c] \cdot K[m, n, c, k] + b[k]$$

unde:

- $A$  este matricea de activare generată de stratul convoluțional.
- $X$  reprezintă tensorul de intrare.
- $K$  este kernelul convoluțional.
- $b$  este biasul asociat.

Operația de convoluție (\*) presupune aplicarea kernelului la fiecare porțiune a tensorului de intrare, ceea ce permite extragerea caracteristicilor esențiale ale imaginii. Aceasta este fundamentul procesării în rețelele CNN, oferind capacitatea de a detecta tipare și structuri specifice în datele de intrare.

În Vision Transformers (ViT), mecanismul de auto-atenție utilizează reprezentări matriciale și calculează similaritatea între token-uri pentru a determina relevanța fiecăruia. Formula mecanismului de auto-atenție este definită astfel:

$$\text{Attention}(Q, K, V) = \text{softmax} \left( \frac{QK^T}{\sqrt{d_k}} \right) V$$

unde:

- $Q$  (queries),  $K$  (keys) și  $V$  (values) reprezintă matrici ce conțin informații despre fiecare token din imagine.
- $d_k$  este dimensiunea vectorilor de keys, utilizată pentru scalare, prevenind valori excesiv de mari care ar putea destabiliza funcția softmax.

Această funcție permite modelului să aloce greutate diferite fiecărui token în funcție de relevanța sa pentru sarcina dată, permițând astfel captarea relațiilor globale dintre părți ale imaginii.

### 4.2.2 Analiză Teoretică

Analizarea teoretică a arhitecturii modelului este esențială pentru justificarea alegerii sale în contextul sarcinilor pe care urmează să le îndeplinească. Capacitatea de generalizare este un concept crucial care se referă la abilitatea modelului de a face predicții corecte pe datele pe care nu le-a întâlnit anterior. O arhitectură bine

proiectată va avea o capacitate mare de generalizare, evitând astfel supra-învățarea (over-fitting).

O metodă eficientă pentru reducerea over-fittingului constă în utilizarea tehnicilor de regularizare, care contribuie la creșterea robusteții modelului. Regularizarea L2, cunoscută și sub denumirea de *ridge regularization*, implică adăugarea unei penalizări pentru greutatea mari în funcția de pierdere. Aceasta poate fi exprimată matematic astfel:

$$L_{\text{total}} = L_{\text{original}} + \lambda \sum_i w_i^2$$

unde:

- $L_{\text{total}}$  reprezintă funcția de pierdere totală după aplicarea regularizării.
- $L_{\text{original}}$  este funcția de pierdere fără regularizare.
- $\lambda$  este un hiper-parametru care controlează intensitatea regularizării.
- $w_i$  sunt greutățile modelului.

Această tehnică penalizează valorile mari ale greutăților, încurajând modelul să generalizeze mai bine și să evite supra-învățarea pe datele de antrenament.

De asemenea, tehnica Dropout constă în dezactivarea aleatorie a anumitor neuroni în timpul fazei de antrenare, ceea ce forțează rețeaua să învețe reprezentări redundante și să devină mai tolerantă la variațiile în datele de intrare. Aceasta se poate exprima printr-o abordare probabilistică, permițând modelului să utilizeze o submulțime a neuronilor activi la fiecare iterație.

Modelul matematic care stă la baza arhitecturii selectate combină funcții matematice pentru definirea pierderilor, structuri complexe de rețea neuronală, și tehnici de regularizare, totul având scopul de a asigura o performanță optimă și o capacitate de generalizare eficientă pe datele necunoscute. Această fundamentare teoretică nu doar că sprijină deciziile de design ale modelului, ci și asigură aplicabilitatea sa în diverse aplicații practice în domeniul imagisticii medicale.

## 5. Configurarea experimentelor

În cadrul experimentelor realizate pentru evaluarea performanței modelului, s-au implementat mai multe configurații esențiale pentru a asigura validitatea și relevanța rezultatelor obținute. Mai jos sunt prezentate principalele aspecte ale configurației experimentale.



## 5.1 Împărțirea Setului de Date

Un aspect crucial în procesul de antrenare a modelelor de învățare profundă este împărțirea setului de date în sub-seturi destinate pentru antrenare, validare și testare. Pentru această investiție, s-a optat pentru o distribuție de 70%-15%-15%. Astfel, 70% din date au fost utilizate pentru antrenarea modelului, 15% pentru validarea acestuia în timpul antrenării (pentru a preveni supraînvățarea), iar 15% pentru testarea finală a performanțelor modelului. Această schemă de împărțire permite o evaluare echilibrată și reprezintă un standard acceptat în domeniu.

## 5.2 Metrici de Evaluare

Pentru a evalua eficacitatea modelului, au fost selectați următorii metrici de evaluare:

- **AUC-ROC:**

Area Under the Receiver Operating Characteristic Curve (AUC-ROC) reprezintă o metrică fundamentală pentru evaluarea performanțelor modelelor de clasificare, fie că este vorba de clasificare binară sau multiclase. Această metrică măsoară capacitatea modelului de a distinge corect între clasele pozitive și negative și oferă o imagine globală asupra eficienței prin evaluarea tuturor pragurilor de clasificare posibile. Un AUC de 0,5 sugerează o capacitate de clasificare aleatorie, în timp ce un AUC de 1,0 semnalează o performanță perfectă.

- **Sensibilitate și Specificitate:**

Aceste metrici sunt esențiale în evaluarea performanțelor modelului în cadrul diagnosticării medicinale. Sensibilitatea, cunoscută și sub denumirea de rată de adevărat pozitiv (TPR), măsoară proporția cazurilor pozitive corect identificate de model, în timp ce specificitatea (TNR) reflectă proporția cazurilor negative corect clasificate. Aceste două metrici sunt cruciale în contextul diagnosticării, deoarece pot influența deciziile clinice.

- **Dice Score:**

Pentru evaluarea sarcinilor de segmentare, Dice Score reprezintă o metrică de comparare între seturile de date a rezultatelor modelului și etichetele de referință. Cunoscut și ca F1 Score, aceasta măsoară similaritatea dintre două seturi, având valori cuprinse între 0 (fără suprapunere) și 1 (suprapunere perfectă). Dice Score este utilizat pe scară largă în aplicarea segmentării imaginilor medicale, oferind o imagine clară asupra preciziei modelului în identificarea structurilor de interes în imagini.

Această configurație a experimentelor, inclusiv împărțirea datelor și metricile de evaluare selectate, este esențială pentru asigurarea unei analize riguroase a performanțelor modelului și pentru validarea rezultatelor obținute în cadrul studiului.

## 5.3 Validarea Rezultatelor

### 5.3.1 Strategii de Validare

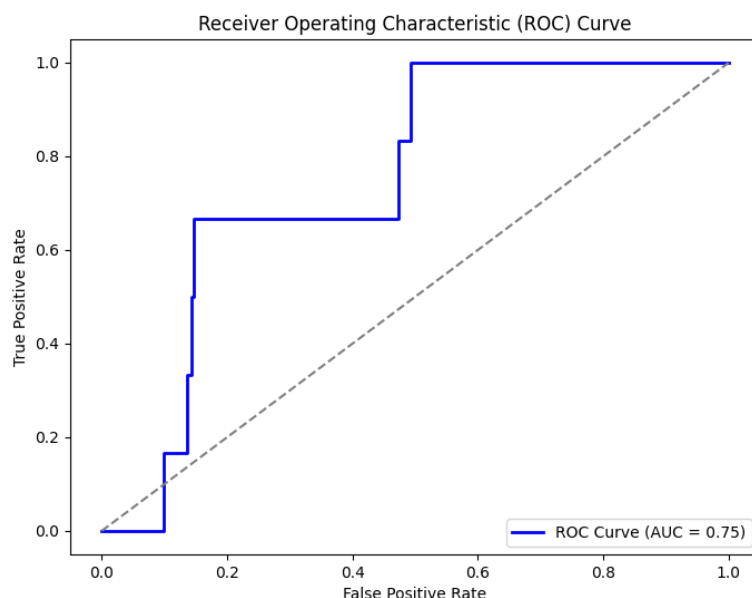
Pentru a evalua performanța modelelor de învățare profundă și a asigura o validare robustă a rezultatelor, este esențial să se utilizeze strategii corespunzătoare de validare. O metodă eficientă utilizată frecvent în acest scop este validarea încrucișată k-fold. Această tehnică implică împărțirea setului de date în k subgrupuri (folds), de obicei  $k=5$  sau  $k=10$ , unde fiecare fold servește alternativ ca set de testare, în timp ce celelalte  $k-1$  folduri sunt utilizate pentru antrenare. Această abordare ajută la reducerea variabilității rezultatelor și oferă o estimare mai stabilă a performanței modelului pe datele noi.

Prin intermediul validării încrucișate, se poate obține o evaluare mai fiabilă, deoarece permite modelului să fie testat pe toate datele disponibile, asigurând că rezultatele nu sunt influențate de un set de date specific. Aceasta crește și șansele de generalizare a modelului, având în vedere diversitatea datelor de antrenare. Validarea încrucișată k-fold este adoptată pe scară largă în literatura de specialitate și este considerată o practică standard pentru evaluarea performanței în analiza datelor.

Pe lângă validarea încrucișată, este important să se utilizeze un set de test complet separat pentru evaluarea generalizării modelului. Acesta ar trebui să fie exclus din procesul de antrenare și validare, oferind o imagine clară asupra capacității modelului de a se adapta la date necunoscute. Această metodă asigură că rezultatele obținute reflectă cu adevărat eficiența modelului în medii reale de utilizare.

## 6. Rezultate și concluzii

Analizând rezultatele intermediare obținute în urma experimentele efectuate, se pot observa performanțe promițătoare în diagnosticarea diferitelor patologii. De exemplu, modelul ResNet50 a obținut un AUC (Area Under the Curve) de 0.75 atunci când a fost testat pe setul de date ChestX-ray14, special pentru diagnosticul



pneumoniei. Aceasta demonstrează capacitatea modelului de a distinge între imagini radiografice sănătoase și cele care prezintă semne de pneumonie.

În concluzie, îmbunătățirea modelelor de diagnosticare imagistică prin utilizarea tehnicilor de inteligență artificială reprezintă un pas esențial în avansarea cercetării medicale și a practici clinice. Preprocesarea datelor, prin metode variate și sofisticate, asigură o bază solidă pentru dezvoltarea unor modele de învățare profundă capabile să ofere diagnostice precise și fiabile. În special, setul de date ChestX-ray14 și modelele de rețele neuronale convoluționale, cum ar fi ResNet50, demonstrează eficiența abordărilor moderne, sporind semnificativ performanța diagnosticării anumitor patologii.

Implementarea unor strategii riguroase de evaluare și validare, cum ar fi validarea încrucișată și utilizarea metricilor de performanță, nu numai că subliniază importanța calității datelor de intrare, dar contribuie și la încrederea în rezultatele obținute. Aceste progrese tehnologice, combinate cu o analiză teoretică bine fundamentată, evidențiază potențialul infinit al inteligenței artificiale de a transforma modul în care diagnosticăm și tratăm afecțiunile medicale.

Astfel, viitorul diagnosticării imagistice se preconizează a fi unul promițător, cu aplicații practice ce pot îmbunătăți semnificativ calitatea asistenței medicale. Investițiile continue în cercetare, dezvoltarea tehnologiilor avansate și colaborarea între specialiști din diverse domenii vor facilita integrarea acestor soluții inovatoare în practica clinică, asigurând o practică medicală mai precisă și mai eficientă.

## 7. Bibliografie

- [1] Papers with Code - ChestX-ray14 Dataset. (2024). [paperswithcode.com](https://paperswithcode.com/dataset/chestx-ray14). <https://paperswithcode.com/dataset/chestx-ray14>
- [2] Baltruschat, Ivo M., Hannes Nickisch, Michael Grass, Tobias Knopp, and Axel Saalbach. "Comparison of Deep Learning Approaches for Multi-Label Chest X-Ray Classification." *Scientific Reports*, April 23, 2019. <https://doi.org/10.1038/s41598-019-42294-8>.
- [3] Eri Matsuyama. "A Deep Learning Interpretable Model for Novel Coronavirus Disease (COVID-19) Screening with Chest CT Images", *Journal of Biomedical Science and Engineering*, January 2020, 13(07):140-152
- [4] Nitish Kundu. "Exploring ResNet50: An In-Depth Look at the Model Architecture and Code Implementation." *Medium*, January 23, 2023. <https://medium.com/@nitishkundu1993/exploring-resnet50-an-in-depth-look-at-the-model-architecture-and-code-implementation-d8d8fa67e46f>.
- [5] Hossain, Md. Belal, S.M. Hasan Sazzad Iqbal, Md. Monirul Islam, Md. Nasim Akhtar, and Iqbal H. Sarker. "Transfer Learning with Fine-Tuned Deep CNN ResNet50 Model for Classifying COVID-19 from Chest X-Ray Images." *Informatics in Medicine Unlocked*, January 1, 2022. <https://doi.org/10.1016/j.imu.2022.100916>.

- [6] "Rezumat%20tez%C4%83%20de%20doctorat%20-%20Mihaela%20Milea%20%28Chistol%29.Pdf." usv.ro, September 19, 2024. [https://usv.ro/fisiere\\_utilizator/file/Doctorat/2024/sustineri/milea/Rezumat%20tez%C4%83%20de%20doctorat%20-%20Mihaela%20Milea%20\(Chistol\).pdf](https://usv.ro/fisiere_utilizator/file/Doctorat/2024/sustineri/milea/Rezumat%20tez%C4%83%20de%20doctorat%20-%20Mihaela%20Milea%20(Chistol).pdf).
- [7] Sarang Narkhede. "Understanding AUC - ROC Curve." Towards Data Science, March 5, 2022. <https://towardsdatascience.com/understanding-auc-roc-curve-68b2303cc9c5>.
- [8] shankar297. "Understanding Loss Function in Deep Learning." Analytics Vidhya, November 5, 2024. <https://www.analyticsvidhya.com/blog/2022/06/understanding-loss-function-in-deep-learning/>.
- [9] Sparsh Gupta. "7 Most Common Machine Learning Loss Functions | Built In." Built In, June 30, 2023. <https://builtin.com/machine-learning/common-loss-functions>.
- [11] "Viewcontent.Cgi?Article=1117&context=diss-Thesis." usd.edu, May 17, 2023. <https://red.library.usd.edu/cgi/viewcontent.cgi?article=1117&context=diss-thesis>.
- [12] Adrian Rosebrock. "Convolutional Neural Networks (CNNs) and Layer Types - PyImageSearch." PyImageSearch, June 8, 2023. <https://pyimagesearch.com/2021/05/14/convolutional-neural-networks-cnns-and-layer-types/>.
- [13] Skylar Jean Callis. "Attention for Vision Transformers, Explained." Towards Data Science, February 29, 2024. <https://towardsdatascience.com/attention-for-vision-transformers-explained-70f83984c673>.
- [14] Deval Shah. "Vision Transformer: What It Is & How It Works [2024 Guide] ." V7, November 16, 2024. <https://www.v7labs.com/blog/vision-transformer-guide>.
- [15] "What Is Generalization in Machine Learning?" RudderStack, November 17, 2024. <https://www.rudderstack.com/learn/machine-learning/generalization-in-machine-learning/>.
- [16] Gu, Rentao, Zeyuan Yang, and Yuefeng Ji. "Machine Learning for Intelligent Optical Networks: A Comprehensive Survey." Journal of Network and Computer Applications, May 1, 2020. <https://doi.org/10.1016/j.jnca.2020.102576>.
- [17] Anum Fatima. "Overfitting, Underfitting and Model's Capacity in Deep Learning." Machine Mindscape, June 19, 2024. <https://machinemindscape.com/overfitting-underfitting-and-models-capacity-in-deep-learning/>.
- [18] Anuja Nagpal. "L1 and L2 Regularization Methods." Towards Data Science, March 28, 2022. <https://towardsdatascience.com/l1-and-l2-regularization-methods-ce25e7fc831c>.
- [19] November 4, 2024. "Regularization in Deep Learning: L1, L2 & Dropout." e2enetworks.com, November 14, 2024. <https://www.e2enetworks.com/blog/regularization-in-deep-learning-l1-l2-dropout>.
- [20] Bradshaw, Tyler J., Zachary Huemann, Junjie Hu, and Arman Rahmim. "A Guide to Cross-Validation for Artificial Intelligence in Medical Imaging." Radiology: Artificial Intelligence, July 1, 2023. <https://doi.org/10.1148/ryai.220232>.
- [21] Kamble, Vaishali H., and Manisha P. Dale. "Machine Learning Approach for Longitudinal Face Recognition of Children." Machine Learning for Biometrics, January 1, 2022. <https://doi.org/10.1016/b978-0-323-85209-8.00011-0>.
- [22] Müller, Dominik, and Frank Kramer. "MIScnn: A Framework for Medical Image Segmentation with Convolutional Neural Networks and Deep Learning." BMC Medical Imaging, January 18, 2021. <https://doi.org/10.1186/s12880-020-00543-7>.

- [23] "A Review on Detection of Pneumonia in Chest X-Ray Images Using Neural Networks." *Journal of Biomedical Physics and Engineering*, December 1, 2022. <https://doi.org/10.31661/jbpe.v0i0.2202-1461>.
- [24] Reshan, Mana Saleh Al, Kanwarpartap Singh Gill, Vatsala Anand, Sheifali Gupta, Hani Alshahrani, Adel Sulaiman, and Asadullah Shaikh. "Detection of Pneumonia from Chest X-Ray Images Utilizing MobileNet Model." *Healthcare*, May 26, 2023. <https://doi.org/10.3390/healthcare11111561>.
- [25] Yao, Shangjie, Yaowu Chen, Xiang Tian, and Rongxin Jiang. "Pneumonia Detection Using an Improved Algorithm Based on Faster R-CNN." *Computational and Mathematical Methods in Medicine*, April 21, 2021. <https://doi.org/10.1155/2021/8854892>.
- [26] "Revolutionizing Healthcare: The Role of Artificial Intelligence in Clinical ...," September 22, 2023. <https://bmcmmededuc.biomedcentral.com/articles/10.1186/s12909-023-04698-z>.
- [27] "Artificial Intelligence for Medical Diagnostics—Existing and Future AI ...," February 12, 2023. <https://pmc.ncbi.nlm.nih.gov/articles/PMC9955430/>.
- [28] "Perspective of Artificial Intelligence in Disease Diagnosis: A Review ...," September 21, 2023. <https://pmc.ncbi.nlm.nih.gov/articles/PMC10590060/>.
- [29] "NIH Findings Shed Light on Risks and Benefits of Integrating AI Into ...," July 23, 2024. <https://www.nih.gov/news-events/news-releases/nih-findings-shed-light-risks-benefits-integrating-ai-into-medical-decision-making>.
- [30] "How Artificial Intelligence Is Shaping Medical Imaging Technology," December 18, 2023. <https://pmc.ncbi.nlm.nih.gov/articles/PMC10740686/>.