

Capitolul 4. Îmbunătățirea modelelor de diagnosticare imagistică

4.1 Introducere

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.

4.2 Îmbunătățirea Diagnosticului cu ChestX-ray14 și ResNet50

4.2.1 Setul de date ChestX-ray14 și utilizarea acestuia cu ResNet50

ChestX-ray14 este un set de date extins care cuprinde 112.120 imagini de radiografii toracice, provenite de la 30.805 pacienți unici, colectate între anii 1992 și 2015. Fiecare imagine este etichetată cu una dintre cele 14 patologii comune, inclusiv

pneumonie, emfizem și cancer pulmonar, fiind un instrument valoros în cercetarea în domeniul diagnosticării imagistice. Acest set de date este utilizat pentru antrenarea și evaluarea performanței modelelor de învățare profundă, cum ar fi ResNet50, un model de rețea neuronală convoluțională dezvoltat de Microsoft care a revoluționat domeniul clasificării imaginilor.

În cadrul utilizării ChestX-ray14 cu ResNet50, modelul este adaptat printr-un proces de fine-tuning, care implică ajustarea greutateilor și optimizarea rețelei prezente pe ImageNet pentru a se potrivi cerințelor specifice ale datelor de radiografie. Acest proces include modificarea ultimei straturi dens de ieșire a modelului pentru a corespunde numărului de patologii din setul de date (14 clase diferite). Fine-tuning permite modelului să învețe din caracteristicile particulare ale radiografiilor și să îmbunătățească acuratețea diagnosticării imaginilor, demonstrând o performanță semnificativ mai bună comparativ cu modele antrenate de la zero (2019).

4.2.2 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.

4.3 Configurația 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.

4.3.1 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.

4.3.2 Metrice de Evaluare

Pentru a evalua eficacitatea modelului, au fost selectați următorii metrice 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 metrice 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.

4.4 Validarea Rezultatelor

4.4.1 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.

4.4.2 Rezultate Intermediare

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 0.88 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.

Un alt exemplu semnificativ provine de la modelul EfficientNet, care a furnizat un AUC de 0.94 în clasificarea melanomului pe setul de date ISIC. Aceasta arată nu doar acuratețea modelului în identificarea melanomului, dar și eficiența abordărilor bazate pe rețele neuronale convoluționale avansate în analiza imaginilor dermatologice.

Aceste rezultate sunt în conformitate cu literatura de specialitate, care susține eficiența acestor modele în aplicarea practică. De exemplu, studiile anterioare au evidențiat că modelele ResNet50 și EfficientNet se află printre cele mai eficiente arhitecturi pentru clasificarea imaginilor medicale, demonstrând realizări similare în evaluările AUC.

Prin implementarea unei strategii solide de validare, precum validarea încrucișată k-fold și utilizarea unui set de test separat, se pot obține rezultate intermediare relevante care reflectă eficiența modelelor de învățare profundă. Exemplele de AUC obținute subliniază nu doar performanța acestor modele în diverse domenii de aplicație, ci și importanța adoptării unor metode riguroase de evaluare pentru asigurarea generalizării și aplicabilității lor în medii clinice reale.

4.5 Model Matematic și Justificare Teoretică

4.5.1 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 clasificății 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.

4.5.2 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_{total} reprezintă funcția de pierdere totală după aplicarea regularizării.
- L_{original} este funcția de pierdere fără regularizare.
- λ 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.

4.6 Concluzii

În concluzie, utilizarea tehnicilor de inteligență artificială pentru îmbunătățirea modelelor de diagnosticare imagistică reprezintă un pas crucial în avansarea cercetării medicale și a practicii clinice. Preprocesarea riguroasă a datelor, prin metode avansate, oferă o fundație solidă pentru dezvoltarea modelelor de învățare profundă capabile să furnizeze diagnostice precise și de încredere. Setul de date ChestX-ray14, împreună cu modele de rețele neuronale convoluționale precum ResNet50, evidențiază eficiența abordărilor moderne, îmbunătățind semnificativ performanța diagnosticării anumitor patologii.

Adoptarea unor strategii riguroase de evaluare, cum ar fi validarea încrucișată și utilizarea unor metrici relevante, nu doar că subliniază importanța calității datelor de intrare, ci și întărește încrederea în acuratețea rezultatelor. Aceste progrese tehnologice, susținute de o analiză teoretică bine fundamentată, confirmă potențialul imens al inteligenței artificiale de a revoluționa diagnosticul și tratamentul medical.