

Universitatea „1 Decembrie 1918” din Alba Iulia

Anul universitar 2017-2018,

DETECȚIA AUTOMATĂ A ANOMALIILOR ÎN MAMOGRAFIILE DIGITALE FOLOSIND TEHNICI DE ÎNVĂȚARE PROFUNDĂ

RAPORT de CERCETARE

Autor: masterand Onița Daniela Marcela, facultatea de Științe Exacte și Inginerești, specializarea Programare Avansată și Baze de Date, anul II de studiu

Profesor coordonator: Lect. univ. dr. Adriana Birluțiu

CUPRINS

INTRODUCERE	2
METODE ȘI INSTRUMENTE DE CERCETARE UTILIZATE	4
Învățarea profundă (en. Deep Learning)	4
Învățarea activă prin transfer (en. Active Transfer Learning)	4
Criteriul de eșantionare a incertitudinii (en. Uncertainty Sampling Criterion)	5
Active Transfer Criterion	5
DESCRIEREA SETULUI DE DATE	6
Preprocesare	7
PROTOCOLUL EXPERIMENTAL	8
REZULTATE OBȚINUTE	10
CONCLUZII ȘI DIRECȚII DE CONTINUARE A CERCETĂRII	13
Comentarii asupra rezultatelor obținute	13
BIBLIOGRAFIE	14

I. INTRODUCERE

Obiectivul general al acestui proiect a fost de a dezvolta un sistem automat de analiză a mamografiilor digitale folosind tehnici de învățare automată.

Proiectul de cercetare științifică propus a fost realizat pe o perioadă de 6 luni în care ne-am propus studiul unei teme de cercetare actuale și anume învățarea profundă. În acest domeniu al învățării profunde rezultatele sunt influențate de cantitatea și calitatea imaginilor. Datorită numărului mic de imagini ale setului de date, algoritmi bazați pe învățarea profundă au o performanță scăzută față de algoritmi de învățare automată. Acest fapt conduce la cercetarea unor noi tehnici aplicate în domeniul învățării automate: învățarea activă și învățarea prin transfer.

Propunerea acestui proiect științific s-a bazat pe cercetarea anterioară pe care am desfășurat-o și care a fost prezentată la conferința SYNASC (*19th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing*) din Timișoara, 21-24 Septembrie 2017, realizată împreună cu Adriana Birluțiu, Adrian Burlacu și Manuella Kadar, cu titlul: ***Defect Detection in Porcelain Industry based on Deep Learning Techniques*** [3].

Rezultatele proiectului de cercetare au fost prezentate la simpozionul național *Convergențe și provocări în domeniul Științelor Exacte și Inginerești*, Ediția a VII-a, din 18 noiembrie 2017, organizat în cadrul Universității “1 Decembrie 1918” din Alba Iulia, și publicate în volumul simpozionului în lucrarea cu titlul ***Învățarea automată activă și învățarea automată prin transfer aplicate în analiza imaginilor medicale*** [1].

Pe baza experienței dobândite în cadrul acestui proiect am continuat cercetarea în domeniu și am realizat un articol cu titlul ***Active Learning based on Transfer Learning*** [2], care a fost trimis spre publicare la conferința internațională *European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, din Belgia.

În viitor ne propunem realizarea unui articol pe o temă conexă proiectului de cercetare care să fie trimis la conferința internațională ***YEF-ECE 2018 - 2nd International Young Engineers Forum on Electrical and Computer Engineering***, pentru masteranzi și doctoranzi, organizată în Portugalia.

Învățarea automată este o ramură a inteligenței artificiale care se ocupă cu proiectarea și dezvoltarea de algoritmi și metode care permit sistemelor informatice să învețe. Pe baza

observațiilor și a experienței, sistemele informatice pot învăța să facă previziuni exacte și să ia decizii optime. În multe probleme de imagistică biomedicală, numai un expert poate analiza datele. Deși experții în domeniu sunt disponibili, munca lor este costisitoare și este de dorit să evităm să punem întrebări atunci când nu este neapărat nevoie. Sarcina noastră este să folosim cât mai eficient resursele existente pentru a facilita învățarea în cazul problemelor cu date de dimensiune redusă.

Învățarea automată aplicată în medicină și în special în analiza imaginilor medicale, este un subiect de cercetare foarte actual și cu rezultate promițătoare. Algoritmii de învățare automată, printr-un proces numit antrenare - adică învățare din exemple - sunt capabili să identifice, să clasifice și să eticheteze conținutul unei imagini. De exemplu, procesul de reconstrucție a imaginilor medicale care furnizează vederi bi și tri-dimensionale ale organelor interne ale pacienților folosește algoritmi de învățare automată pentru analiza și procesarea imaginilor. Acești algoritmi pot contribui astfel la întreg actul medical, începând cu diagnostic și până la tratament și intervenții.

Performanța algoritmilor de învățare automată folosiți în analiza imaginilor medicale depinde de datele de antrenare. Datele folosite pentru antrenare sunt însă limitate pentru că necesită etichetarea lor de către cadrele medicale. Cu cât există mai multe date care pot fi folosite pentru antrenarea algoritmilor, cu atât crește performanța și implicit calitatea imaginilor medicale. Învățarea automată prin transfer a apărut ca urmare a acestei nevoi de a include informații și date adiționale în procesul de învățare automată.

Obiectivul cercetării noastre a fost de a îmbunătăți tehnicile de învățare automată prin punerea împreună a două tehnici de învățare: învățarea activă și învățarea prin transfer. Primul pas este crearea unui model de clasificare pentru setul de date ales, folosind tehnici de învățare automată. Se dorește demonstrarea faptului că învățarea activă folosind una din strategiile de învățare se pretează mai bine în comparație cu alegerea aleatorie a punctelor care urmează a fi adăugate setului de antrenare. Al doilea pas este combinarea învățării active și cea prin transfer pentru construirea unui algoritm eficient care să poată clasifica imaginile medicale. Criteriile de învățare activă pe care le propunem sunt similare cu cele introduse în [5], [14].

II. METODE ȘI INSTRUMENTE DE CERCETARE UTILIZATE

1. **Învățarea profundă (en. Deep Learning)**

Învățarea profundă (en. Deep Learning) este o metodă relativ nouă în domeniul tehnologiilor de inteligență artificială și de învățare automată [4, 16]. Deep learning este considerată o ramură a învățării automate (en. Machine Learning), care are ca obiectiv apropierea dintre aceasta și inteligența artificială, prin învățarea computerelor a unor lucruri de care în trecut doar oamenii erau capabili (recunoașterea imaginilor, a vocii, etc.). Această tehnică a înregistrat succese considerabile în aplicații diverse, în special în cercetarea medicală. Rețelele neuronale profunde sunt utilizate în aplicații pentru clasificarea imaginilor medicale, pentru a determina specificitățile ADN-ului și ARN-ului, pentru clasificarea unor bacterii sau pentru diagnosticarea diverselor boli, cum ar fi Alzheimer, etc. Doar câteva lucrări au explorat metode profunde de învățare pentru a aborda clasificarea automată a leziunilor identificate pe mamografie, de exemplu în [3] s-au folosit rețele neuronale convoluționale pentru diagnosticarea cancerului mamar.

2. **Învățarea activă prin transfer (en. Active Transfer Learning)**

Învățarea automată prin transfer (Pan, Yang, 2010) s-a dezvoltat în cadrul învățării automate și a inteligenței artificiale ca o nouă paradigmă ce investighează modalitățile de a recunoaște și a aplica cunoștințe și abilități dobândite anterior pentru a rezolva sarcini noi dar similare cu cele anterior rezolvate. Ideea de bază din spatele algoritmilor de învățare automată prin transfer este că domeniile implicate conțin informații comune, dar ascunse care pot fi descoperite și exploatate folosind diferite tehnici și care reprezintă o punte de legătură pentru transferul cunoștințelor dobândite anterior. Învățarea automată prin transfer se inspiră din învățarea prin transfer din psihologie, mai precis din dependența învățării umane de experiența anterioară. Pe baza analogiei, se poate realiza transferul de la o problemă anterior rezolvată la o problemă țintă. Acest tip de învățare ridică următoarele întrebări: ce este analogia, cum sunt recunoscute analogiile, cum se transferă cunoștințele relevante? De exemplu, un sistem informatic conținând cunoștințe despre diagnosticarea defectelor din calculatoare poate învăța prin analogie diagnosticarea bolilor umane prin corelarea disfuncțiilor calculatorului cu simptomele pacienților, respectiv a unei metode de depanare cu tratamentul unei boli.

În ultimii ani algoritmi de învățare automată prin transfer au început să fie investigați pentru analiza imaginilor medicale pentru că s-a demonstrat prin studii științifice că aceștia duc la creșterea performanței actului medical (Heimann et al., 2014; van Opbroek et al., 2015; Zheng, 2015).

Învățarea activă, cunoscută și în literatura de specialitate ca *optimal experimental design*, este potrivită pentru situațiile în care etichetarea datelor este o activitate dificilă, consumatoare de timp și costisitoare. Ideea din spatele învățării active este că prin selectarea optimă a datelor de antrenare, poate fi obținută o performanță mai bună față de selecția aleatoare a datelor ce vor fi folosite în antrenare. Scenariile în care învățarea activă poate fi aplicată fac parte din următoarele trei categorii: (i) generarea de date noi care urmează a fi etichetate; (ii) învățarea activă în fluxul de date, unde la fiecare moment se decide dacă să se solicite sau nu eticheta unei instanțe de date; (iii) învățarea activă bazată pe selectarea punctului dintr-un set separat de date care conține și date neetichetate. În acest proiect considerăm al treilea scenariu pentru realizarea învățării active.

3. Criteriul de eșantionare a incertitudinii (en. **Uncertainty Sampling Criterion**)

Criteriul incertitudinii este o strategie de învățare activă în care este selectat pentru etichetare exemplul pentru care predicțiile modelului sunt cele mai incerte. Incertitudinea predicțiilor poate fi măsurată, de exemplu, folosind entropia Shannon:

$$Uncertainty(a) = - \sum_c p(c|a) \log p(c|a),$$

unde a reprezintă punctul care urmează a fi etichetat, iar c reprezintă posibila etichetă. Pentru un clasificator binar această strategie se reduce la interogarea punctelor ale căror predicție este aproape de 0.5. Această strategie este cunoscută datorită modului rapid de creștere a performanței algoritmului.

4. Active Transfer Criterion

Criteriul propus se bazează pe Transfer Learning (TL). Învățarea de transfer (Pan, Yang, 2010) este o tehnică folosită pentru a transfera cunoștințele de la o sarcină de învățare sursă la o sarcină de învățare țintă. În contextul algoritmilor de învățare automată, învățarea de transfer poate fi implementată prin transferarea diferitelor elemente ale modelului învățat pe sarcina de învățare sursă către sarcina de învățare țintă. Învățarea prin transfer se utilizează

atunci când datele de antrenare ale sarcinii țintă sunt de dimensiuni mai mici, similare, dar nu identice cu datele sarcinii sursă. [14]

Ideea principală din spatele criteriului Active Transfer Learning propusă în acest proiect este de a exploata învățarea cu mai multe seturi de date și de a folosi modelele învățate pe alte seturi de date M_1, \dots, M_M la determinarea cunoștințelor dobândite cu un nou punct de date.

Vom folosi următoarea notație pentru probabilitatea corespunzătoare la alte modele $m = 1, \dots, M$,

$$p_m(c|a) \equiv p(c|a, M_m).$$

Inspirat de [5], se măsoară dezacordul ținând seama de predicția medie a întregului comitet și se calculează media Kullback-Leibler (KL) divergența previziunilor individuale față de media:

$$Committee(a) = \sum_{m=1}^M \frac{1}{M} KL[\bar{p}(\cdot | a) || p_m(\cdot | a)],$$

cu $p(\cdot | a)$ probabilitatea medie predictivă a întregului comitet.

Diferența KL pentru probabilități discrete este definită ca

$$KL[p_1(\cdot | a) || p_2(\cdot | a)] = \sum_c p_1(c|a) \log\left(\frac{p_1(c|a)}{p_2(c|a)}\right).$$

Diferența KL poate fi văzută ca o distanță între probabilități, distanță folosit într-un sens abuziv, deoarece divergența KL nu este simetrică, adică, $KL[p_1 || p_2] \neq KL[p_2 || p_1]$. Acest dezavantaj al divergenței KL poate fi depășit luând în considerare o măsură simetrică, de exemplu $KL[p_1 || p_2] + KL[p_2 || p_1]$.

II. DESCRIEREA SETULUI DE DATE

Pentru clasificarea imaginilor s-a folosit setul Breast Cancer Histopathological Images Classification (BreakHis) [6]. Acest set de date este format din imagini microscopice ale țesutului tumoral mamar, colectate de la 82 de pacienți. Imaginile utilizează diferiți factori de mărire (40X, 100X, 200X, 400X), așa cum este ilustrat în Fig. nr. 1.

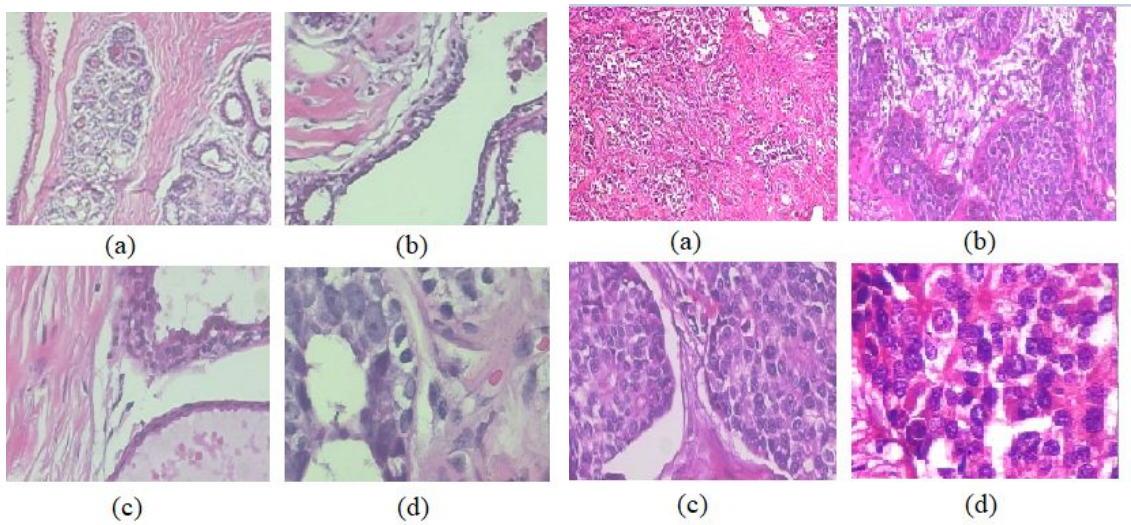


Fig. nr. 1 – Diferite tipuri de tumori de sân observate în factori de mărire diferiți: (a) 40x, (b) 100x, (c) 200x, (d) 400x. Imaginile din partea stânga prezintă tumori benigne, iar cele din dreapta prezintă tumori maligne.

Imaginile sunt structurate după cum urmează:

Tabel nr. 1 - Structura setului de date BreakHis

Factor de mărire	Benign	Malign	Total
40X	640	1354	1994
100X	644	1437	2081
200X	623	1390	2013
400X	588	1095	1683

Preprocesare

Fiecare imagine a fost convertită la scară gri și a fost redimensionată la 28x28, astfel o imagine este descrisă de un vector caracteristic 784-dimensional.

Imaginile au fost preprocesate folosind Principal Component Analysis (PCA). În prima fază, s-a efectuat centrarea datelor în jurul valorii zero: pentru fiecare porțiune a imaginii, valoarea medie a pixelilor a fost calculată și scăzută din date. În a doua fază, dispersia datelor s-a adus în jurul valorii 1, în modul următor: i) s-a calculat matricea de covarianță a datelor și s-a realizat factorizarea SVD a matricei; ii) s-a realizat decorelarea datelor prin rotirea și reducerea dimensiunii; iii) datele decorelate au fost împărțite la valorile proprii.

III. PROTOCOLUL EXPERIMENTAL

Pentru realizarea experimentelor bazate pe tehnici de învățare profundă s-au construit modele bazate pe rețele neuronale convoluționale (en. Convolutional Neural Networks). S-a dorit obținerea unui model a cărui performanță să fie îmbunătățită comparativ cu algoritmi de învățare automată. Pentru aceasta s-au testat mai multe tipuri de arhitecturi a rețelei neuronale, cu diferiți parametri și cu un număr diferit de straturi ascunse. Am observat faptul că incrementarea numărului de straturi ascunse nu îmbunătățește performanța algoritmului. Modelul este o rețea neurală, a cărei structură este prezentată în Figura nr. 2. A fost utilizată o funcție rectilinie de activare pentru neuronii din stratul ascuns. Pentru a activa stratul de ieșire a fost folosit o funcție sigmoidă. S-a folosit entropia încrucișată binară ca funcția de pierdere. Algoritmul “stochastic gradient descent”: rata de învățare = 0.1, impuls = 0.7 și dezintegrarea în greutate = 0,14 a fost utilizată pentru a învăța ponderile. Modelul a fost antrenat pe 50 de epoci cu o dimensiune a lotului de 28. Performanța modelului, 0.64 ± 0.62 (acuratețe \pm deviația standard), nu a fost îmbunătățită comparativ cu algoritmi de învățare automată.

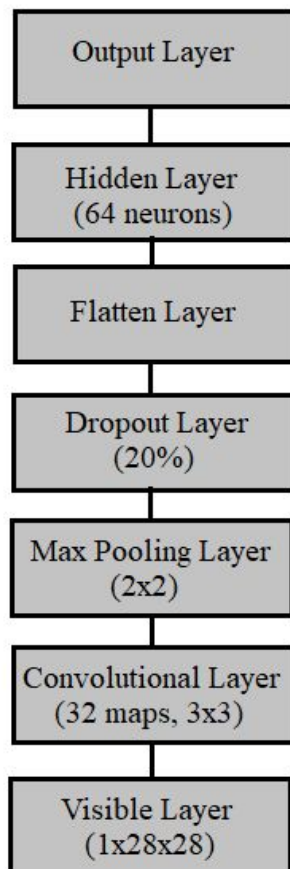


Figura nr. 2 - Structura rețelei neuronale convoluționale

Astfel, s-au comparat mai mulți algoritmi de învățare automată: Regresie logistică, Analiză discriminatorie liniară, Arbori de decizie, Algoritm Naive Bayes, Algoritm Random Forest și Mașini cu suport vectorial (SVM). Performanța algoritmilor este prezentată în Tabel nr. 2 din capitolul *Rezultate obținute*.

În continuare, am testat rezultatele obținute cu învățarea activă în comparație cu cele ale selecției aleatorii. Pentru a evidenția experimentele de învățare activă, am creat un model care să fie instruit pe un set de date folosind algoritmul SVM. Datele de antrenament au fost utilizate ca un grup de puncte din care au fost selectate puncte pentru etichetare, fie aleator, fie activ. În primul rând, un punct aleatoriu a fost selectat dintre datele neetichetate și a fost adăugat la datele de antrenare. După selectare, acest punct a fost șters din datele neetichetate. În al doilea rând, punctul a fost selectat activ. Pentru eșantionarea incertitudinii, deoarece avem un clasificator binar, punctul care a fost ales a fost unul pentru care probabilitățile de predicție (obținute pe date neetichetate cu SVM model) sunt mai apropiate de 0,5. După fiecare actualizare au fost făcute previziuni pe setul de validare. Am folosit acuratețea (medie \pm deviația standard) ca măsură a performanței. Rezultatele au fost medii pe parcursul a 20 de divizări a setul de antrenare, setul de date neetichetate și setul de date de validare. Toți algoritmi au fost instruiți pe baza selecției aleatorii, respectiv active a 50 de puncte.

Pentru experimentele active de transfer, alegem unul din seturile de date ca țintă. Acesta va fi instruit folosind un model pre-instruit într-un set de date format de restul seturilor de date care corespund diferiților factori de mărire. Punctul care a fost selectat a fost cel pentru care celelalte modele au obținut cea mai slabă performanță. Punctele selectate au fost adăugate la setul de antrenament și au fost șterse din seturile de date neetichetate.

Un alt experiment pentru a evidenția comparația dintre învățarea activă folosind sau nu un model pre-instruit: se alege unul din factorii de mărire ca set de date. Acesta va fi antrenat folosind un model pre-instruit pe un set de date format din restul factorilor de mărire. Fiecare punct a fost selectat folosind strategia activă a strategiei de eșantionare a incertitudinilor. Pașii de adăugare și ștergere ai punctelor selectate sunt similari cu cei de la experimentele descrise anterior. Rezultate obținute sunt prezentate în graficul din Figura nr. 5.

În Figura nr. 6 este prezentat următorul experiment realizat pentru a compara de asemenea învățarea activă folosind modele pre-instruite și învățarea activă folosind algoritmul SVM. Au fost definite patru modele SVM și fiecare dintre ele a fost instruit după cum urmează: unul pentru fiecare factor de mărire. Aceste modele vor fi folosite ca modele

pre-instruite. De această dată, punctele au fost selectate activ folosind transferul de învățare după cum urmează. A fost ales un factor de mărire ca set de date și acesta a fost antrenat folosind fiecare dintre modelele pre-instruite rămase. De exemplu, dacă alegem factorul de mărire 200X ca set de date, va fi antrenat folosind modele pre-instruite pe seturile de date 40X, 100X și 400X. Pentru fiecare din cele 50 iterații, punctul a fost selectat utilizând strategia activă de eșantionare a incertitudinii, dar de această dată ne referim la probabilitățile de predicție obținute folosind modele pre-instruite.

Rezultatele experimentelor descrise sunt prezentate în graficele din capitolul următor.

IV. REZULTATE OBȚINUTE

Pentru realizarea experimentelor descrise anterior s-a folosit limbajul de programare Python și mediul de programare Spyder. Principalele biblioteci utilizate au fost Keras, Theano și sklearn.

Rezultatele comparației algoritmilor de învățare sunt prezentate în Tabelul nr. 2. SVM funcționează cel mai bine, astfel se justifică folosirea lui în restul experimentelor.

Tabel nr. 2 - Comparația performanței algoritmilor de învățare

Algoritm de învățare	Acuratețea (media ± deviația standard)
SVM	0.66 ± 0.0004
Regresie liniară	0.66 ± 0.0004
Analiză discriminatorie liniară	0.57 ± 0.002
Arbori de decizie	0.59 ± 0.04
Naive Bayes	0.56 ± 0.04
Random Forest	0.59 ± 0.01

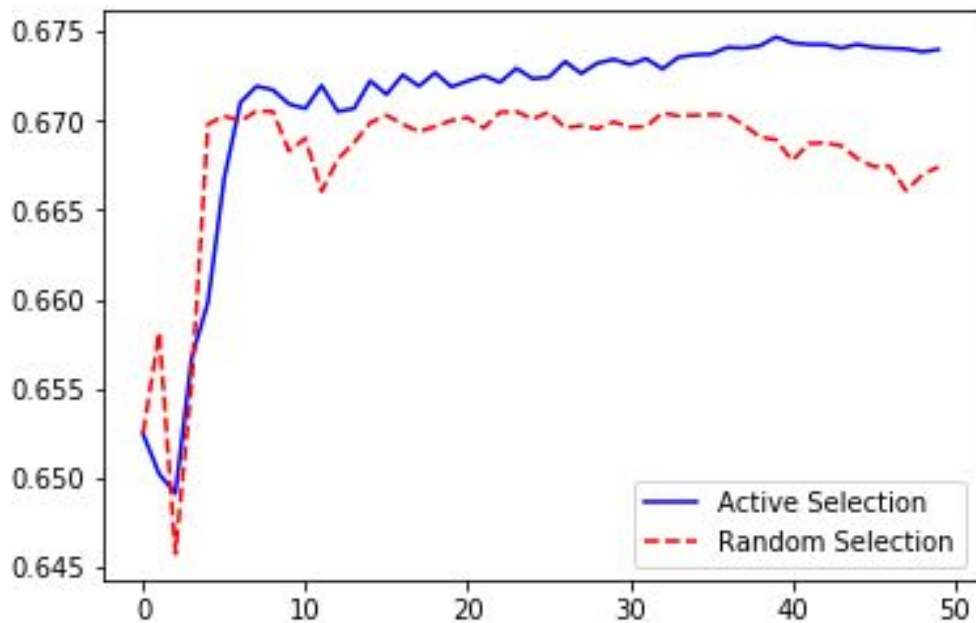


Fig. nr. 3 - Acuratețea obținută după selecția activă vs selecția aleatorie a punctelor de antrenament.

Figura nr. 3 compară performanța obținută cu selectarea aleatorie și activă pentru setul de date. Rezultatele obținute utilizând o selecție activă sunt mai bune decât rezultatele obținute folosind selecția aleatorie a punctelor de antrenament.

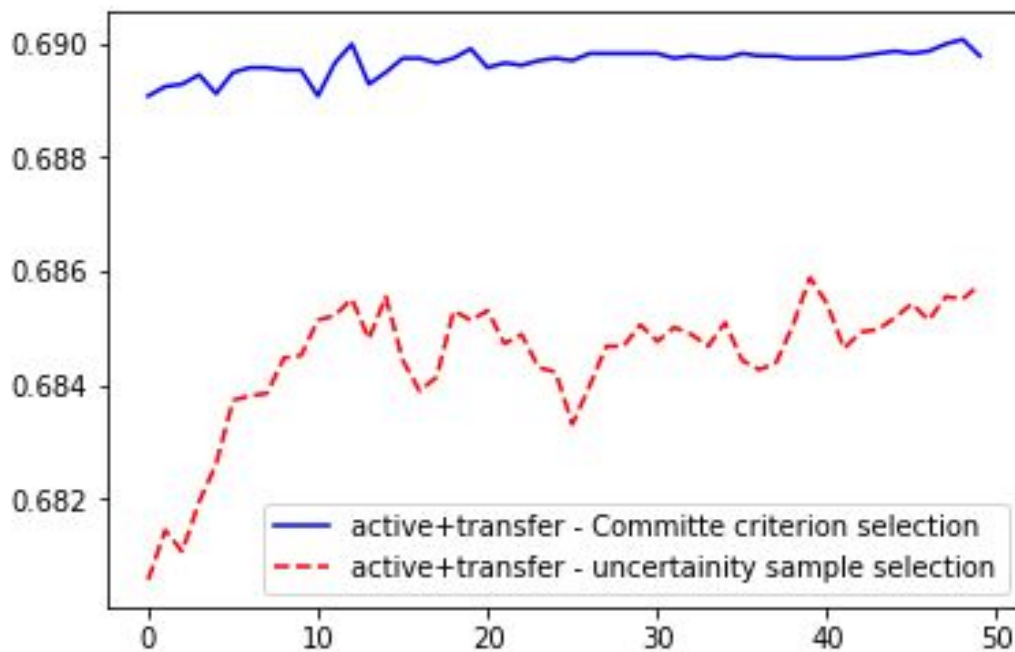


Fig. nr. 4 - Precizia obținută după selecția activă de transfer utilizând criteriul de transfer activ vs. selecția activă de transfer utilizând criteriul de eșantionare a incertitudinii a punctelor de antrenament pentru setul BreakHis.

Figura nr. 4 compară acuratețea obținută cu ajutorul învățarea activă de transfer folosind diferite tipuri de strategii active: criteriul de eșantionare a incertitudinii și criteriu de transfer activ. Acest grafic arată că criteriul transferului activ îmbunătățește performanța în unele cazuri.

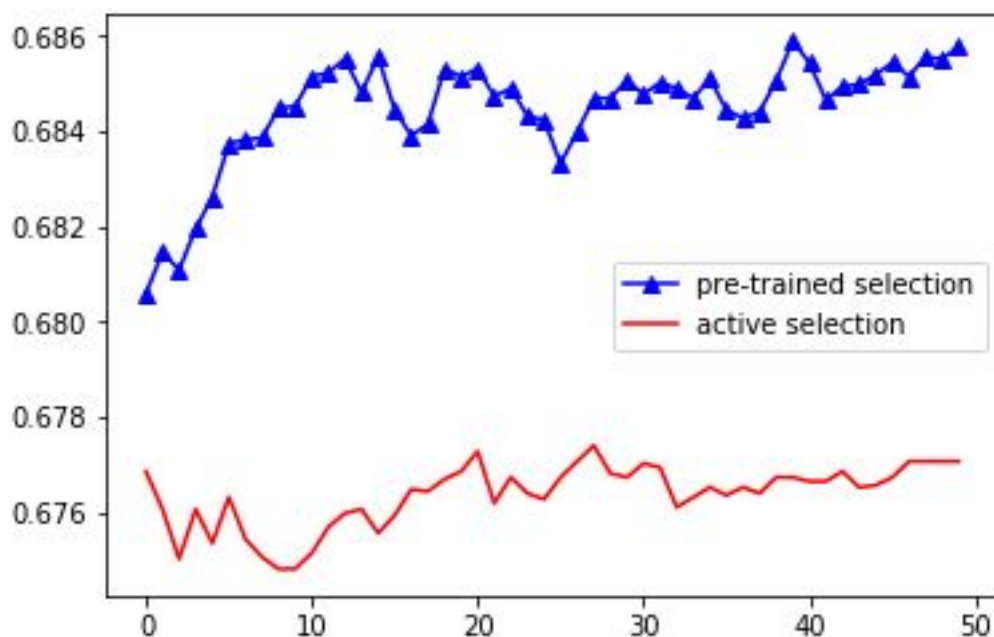


Figura nr. 5 - Acuratețea obținută folosind selecția activă de transfer și un model pre-instruit (pre-trained selection) vs. selecția activă de transfer fără un model pre-instruit utilizând eșantionul de incertitudine a punctelor de antrenament pentru setul BreakHis.

În Figura nr. 5 sunt evidențiate experimentele în care s-au comparat rezultate obținute folosind selecția activă a punctelor folosind criteriul incertitudinii. Diferența o face faptul că în primul caz selecția s-a făcut folosind algoritmul SVM ca algoritm de antrenare, iar în al doilea caz s-a folosit pentru antrenare un model pre-instruit descris anterior în capitolul *Protocolul experimental*. Din grafic rezultă faptul că folosirea unui model pre-instruit este benefică performanței algoritmului.

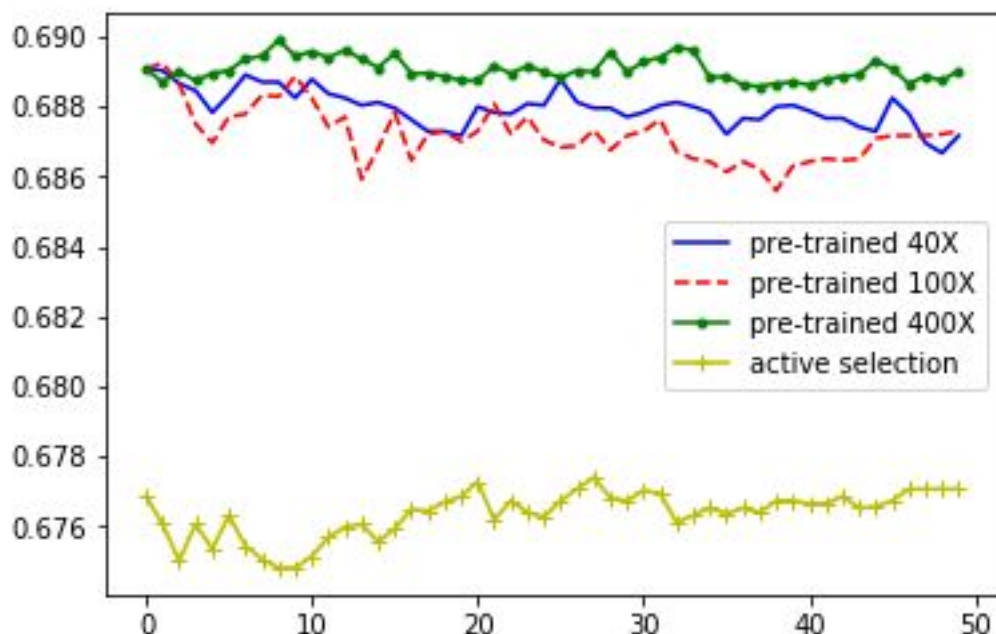


Figura nr. 6 - Acuratețea obținută folosind selecția activă de transfer și criteriul de eșantionare a incertitudinii folosind sau nu modele pre-instruite.

În graficul din Figura nr. 6 este comparată de asemenea performanța algoritmilor folosind sau nu modele pre-instruite. Comparativ cu graficul din Figura nr. 5 unde setul sursă a fost obținut prin concatenarea a trei seturi, în graficul din Figura nr. 6 au fost folosite 3 modele pre-instruite ca seturi sursă. Asemănarea reiese din faptul că în ambele cazuri modele pre-instruite au o performanță mai ridicată.

IV. CONCLUZII ȘI DIRECȚII DE CONTINUARE A CERCETĂRII

În acest raport s-a demonstrat faptul că performanța algoritmilor de învățare automată pentru clasificarea imaginilor din domeniul biomedical poate fi crescută prin combinarea învățării active cu învățarea de transfer. În viitor ne propunem extinderea cercetărilor pe aceeași temă propusă în proiectul de cercetare, prin dezvoltarea unor modele asemănătoare cu cele descrise la nivelul acestui proiect și prin testarea altor criterii din domeniul inteligenței artificiale.

Comentarii asupra rezultatelor obținute

Experimentele tratate la nivelul acestui proiect de cercetare îmbină învățarea automată cu cea activă și cea de transfer, cu scopul creșterii performanței algoritmilor. Rezultatele

obținute au fost publicate în cadrul simpozionului național “Convergențe și provocări în domeniul științelor exacte și ingineresti”, organizată de Universitatea “1 Decembrie 1918” din Alba Iulia. De asemenea o lucrare bazată pe tehnici asemănătoare cu cele folosite în acest proiect de cercetare a fost trimisă spre recenzie la simpozionul internațional din Belgia, “European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning”.

V. BIBLIOGRAFIE

1. Onița D., Birlutiu A., Furdui A., Învățarea automată activă și învățarea automată prin transfer aplicate în analiza imaginilor medicale, Ed. Aeternitas, Alba Iulia, 2017, ISSN: 2285-9500.
2. Onița D., Birluțiu A., Active Learning based on Transfer Learning - trimis spre recenzie la conferința European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, Belgia.
3. A. Birlutiu, A. Burlacu, M. Kadar, D. Onita. Defect Detection in Porcelain Industry based on Deep Learning Techniques. 19th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, Timisoara, September 21- 24, 2017. (accepted)
4. Daniela Onița, Golban Igor, Detecția defectelor în industria porțelanului folosind rețele neuronale adânci, TOM II Științe Exacte și Inginerești, IN-EXTENSO Nr. 4/2017, Ed. Aeternitas, Alba Iulia, 2017, ISSN 2360 - 0675, ISSN-L 2360 - 0675.
5. Birlutiu A., Groot P., Heskes T. (2013) Efficiently learning the preferences of people. Machine Learning Journal, 90 (1), pp.1-28, Springer, ISSN: 0885-6125.
6. Spanhol, F., Oliveira, L. S., Petitjean, C., Heutte, L., A Dataset for Breast Cancer Histopathological Image Classification, IEEE Transactions on Biomedical Engineering (TBME), 63(7):1455-1462, 2016.
7. Lewis, D., Gale, W. A sequential algorithm for training text classifiers. In Proceedings of the 17th annual international ACM SIGIR conference on research and development in information retrieval (pp. 3^a12), 1994.
8. Bishop, C.: Pattern Recognition and Machine Learning. Springer, (2006).
9. Mnih,V.: Machine learning for aerial image labeling, Ph.D. thesis, University of Toronto, (2013)

10. McCallum, A., Nigam, K. Employing EM and pool-based active learning for text classification. In Proceedings of the 15th international conference on machine learning (pp.350â358), 1998.
11. Diethe, T., Twomey, N., and Flach, P. Active transfer learning for activity recognition. ESANN 2016 proceedings, European Symposium on Artificial Neural Network, Computational Intelligence and Machine Learning.
12. Zhao, L., Jialin Pan, S., Wei Xiang, E., Zhong,E., Lu, Z., and Yang, Q. Active transfer learning for cross-system recommendation. In Proceedings of the 27th AAAI Conference on Artificial Intelligence, 2013.
13. Bengio, Y. (2009), 'Learning deep architectures for AI', Foundations and Trends in Machine Learning 2(1), 1–127.
14. Wang, X. Active Transfer Learning, PhD Thesis. CMU, 2016.
15. Burr Settles and Mark Craven. 2008. An analysis of active learning strategies for sequence labeling tasks. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP '08)*. Association for Computational Linguistics, Stroudsburg, PA, USA, 1070-1079.
16. Arevalo, J., González, F. A. & Ramos-Pollán, R. Representation learning for mammography mass lesion classification with convolutional neural networks. *Computer Methods and Programs in Biomedicine* 127, 248–257 (2016).