

RAPORT FINAL ȘTIINȚIFIC ȘI TEHNIC

Denumire proiect (EN)	CULTiVA: Curriculum learning in text mining and visual question answering
Denumire proiect (RO)	CULTiVA: Învățarea pe bază de curriculum în text mining și visual question answering
Acronim	CULTiVA
Cod proiect	PN-III-P1-1.1-TE-2019-0235
Număr contract	72/2020
Contractor	UNIVERSITATEA BUCUREȘTI
Tip proiect	Proiecte de cercetare pentru stimularea tinerelor echipe independente (TE)
Autoritatea contractantă	Unitatea Executivă pentru Finanțarea Învățământului Superior, a Cercetării, Dezvoltării și Inovării
Perioada de raportare	07.09.2020 - 31.08.2022
Director de proiect	Radu Tudor Ionescu
Adresă web	https://cultiva-proj.herokuapp.com

OBIECTIVE PREVĂZUTE

Obiectivul 1 – Construirea unui predictor de dificultate a textului.

Obiectivul 2 – Antrenarea unor modele neuronale pentru clasificarea textelor și visual question answering, pentru a reproduce rezultatele de ultimă generație.

Obiectivul 3 – Antrenarea unor modele neuronale mai bune folosind învățarea bazată pe curriculum pentru clasificarea textelor.

Obiectivul 4 – Antrenarea unor modele neuronale mai bune folosind învățarea bazată pe curriculum pentru visual question answering.

Obiectivul 5 – Publicarea celor mai bune modele neuronale obținute pe un website public.

ETAPA 1/3

1. Etapa 1 – Rezumat:

În Etapa 1 – „Dezvoltarea și evaluarea unui predictor de dificultate a textelor (obiectiv 1)” am efectuat activitățile 1.1 și 1.2. Activitatea 1.1. are ca scop îndeplinirea primului obiectiv, acela de a concepe un model neuronal pentru estimarea dificultății textelor. Activitatea 1.2. are ca scop crearea și publicarea site-ului web asociat proiectului. Informațiile vor fi actualizate pe site pe parcursul desfășurării proiectului. Pe lângă cele două activități propuse, am demarat lucrul în vederea activităților 2.1 și 2.3, ce au ca scop îndeplinirea obiectivelor 2, respectiv 4.

2. Etapa 1 – Procentaj îndeplinire obiective la finalul etapei:

Obiectivul 1 – 100%

Obiectivul 2 – 20%

Obiectivul 3 – 0%
Obiectivul 4 – 20%
Obiectivul 5 – 20%

3. Etapa 1 – Descriere științifică și tehnică:

În conformitate cu activitățile prevăzute în Etapa 1 de raportare din Anexa II a contractului TE72/2020, am efectuat următoarele:

- **Activitatea 1.1. Task 1 – Dezvoltarea și evaluarea unui predictor de dificultate a textelor**

În vederea îndeplinirii primului nostru obiectiv, am decis să folosim un model neuronal state-of-the-art pentru estimarea dificultății textelor, anume un model tip BERT [Devlin et al., NAACL19]. Pentru a ne familiariza cu modelul BERT, am efectuat întâi experimente antrenând acest model pe problema geo-localizării textelor, aceasta fiind o problemă de regresie asemănătoare cu cea de estimare a complexității textelor. În urma experimentelor efectuate, am ajuns la concluzia că rezultatele modelului BERT sunt mai slabe în comparație cu modele precum metode nucleu pentru șiruri de caractere. Aceste rezultate au fost publicate în lucrarea [Găman et al., VarDial20]. Totuși, am decis continuarea studiului în direcția estimării dificultății textelor pe baza modelului BERT, aceasta fiind la rândul său tot o problemă de regresie. De această dată, rezultate obținute sunt apropiate de cele raportate în lucrarea [Butnaru et al., BEA18], conform Tabelului 1 de mai jos. Avantajul modelului BERT este extinderea simplă și directă de la estimarea complexității cuvintelor la estimarea complexității propozițiilor sau textelor.

Metodă	Acuratețe	Scor F1
[Butnaru et al., BEA18]	86.78%	85.94%
BERT	86.55%	82.96%

Tabel 1. Rezultatele modelului BERT în comparație cu rezultatele modelului propus în lucrarea [Butnaru et al., BEA18] pe problema estimării complexității la nivel de cuvânt pe setul de date English News.

- **Activitatea 1.2. Task 5 – Dezvoltarea unui website public pentru prezentarea proiectului și publicarea modelelor / codului aferent.**

Această activitate corespunde cu obiectivul 5 din cadrul propunerii de proiect, acela de a publica rezultatele obținute pe site-ul proiectului. În vederea îndeplinirii acestui obiectiv, am dezvoltat și publicat site-ul web al proiectului la: <https://cultiva-proj.herokuapp.com>. Conținutul publicat include un rezumat precum și informații despre membrii implicați în proiect și articolele publicate.

- **Activități premergătoare Activităților 2.1. și 2.3.**

În plus față de cele prevăzute în activitățile 1.1. și 1.2., am studiat posibilitatea aplicării învățării pe bază de curriculum pentru a antrena modele neuronale pentru problema detectării obiectelor din imagini. Rezultatele obținute demonstrează că învățarea pe bază de curriculum conduce către performanțe mai bune. Metoda propusă și rezultatele aferente au fost sintetizate în lucrarea [Soviany et al., CVIU21]. De menționat că în vederea dezvoltării unui model pentru visual question answering, una din etapele importante este tocmai detectarea obiectelor în

imagini. Astfel, studiul efectuat este relevant în îndeplinirea obiectivului 4, fiind un pas premergător activității 2.3.

Pe lângă cele enumerate mai sus, am studiat problema copierii unor modele state-of-the-art de tip black-box, folosind o tehnică de generare de imagini bazată pe calcul evoluționar. Replicarea rezultatelor unor modele din literatura recentă este obiectului 2 al proiectului nostru. Totuși, în anumit cazuri, aceste modele sunt expuse printr-un API ce permite doar propagarea unor exemple și aflarea predicțiilor corespunzătoare. În acest context, tehnica propusă este de interes, având o legătură importantă cu activitatea 2.1. Metoda propusă și rezultatele aferente au fost sintetizate într-o lucrare acceptată spre publicare [Bărbălău et al., NeurIPS20].

ETAPA 2/3

1. Etapa 2 – Rezumat:

În Etapa 2 – „Dezvoltarea și evaluarea unor modele neuronale deep cu rezultate state-of-the-art” am efectuat activitățile 2.1, 2.2 și 2.3. Activitatea 2.1. are ca scop îndeplinirea celui de-al doilea obiectiv, acela de a antrena rețele neuronale state-of-the-art pentru clasificarea textelor și visual question answering, utilizând seturi de date publice. Activitatea 2.2. are ca scop îndeplinirea parțială a celui de-al treilea obiectiv, anume dezvoltarea și evaluarea unui model neuronal deep pentru clasificarea textelor, folosind paradigma de învățare bazată pe curriculum. În mod asemănător, activitatea 2.3 are ca scop îndeplinirea parțială a celui de-al patrulea obiectiv, anume dezvoltarea și evaluarea unui model neuronal deep pentru visual question answering, folosind paradigma de învățare bazată pe curriculum.

2. Etapa 2 – Procentaj îndeplinire obiective la finalul etapei:

Obiectivul 1 – 100%

Obiectivul 2 – 100%

Obiectivul 3 – 50%

Obiectivul 4 – 40%

Obiectivul 5 – 66%

3. Etapa 2 – Descriere științifică și tehnică:

În conformitate cu activitățile prevăzute în Etapa 2 de raportare din Anexa II a contractului TE72/2020, am efectuat următoarele:

- **Activitatea 2.1. Task 2 – Dezvoltarea și evaluarea unor modele neuronale deep cu rezultate state-of-the-art.**

În vederea îndeplinirii celui de-al doilea obiectiv, am decis să folosim modele neuronale deep de tip transformer, cum ar fi BERT [Devlin et al., NAACL19], pentru clasificarea textelor, respectiv arhitecturi convoluționale, cum ar fi ResNet [He et al., CVPR16], pentru procesarea imaginilor. În zona procesării limbajului natural, am comparat modelul BERT cu o serie de abordări alternative în diverse probleme de clasificare a textelor, anume detecția polarității opiniei [Tache et al., EACL21], geo-localizarea textelor [Găman et al., VarDial21], detecția satirei [Ionescu et al., IJCNN21; Rogoz et al., ACL21] și recunoașterea sub-dialectelor limbii române [Găman et al., IJIS22].

Totodată, în zona procesării de imagini (naturale sau rezultate în urma aplicării unor transformări pe semnale audio), am comparat mai multe arhitecturi convoluționale de tip

ResNet [He et al., CVPR16] pe diverse probleme de clasificare sau detecție, anume recunoașterea și detecția obiectelor [Duță et al., ICCVW21], detecția speciilor de păsări sau a emoției din vorbire [Ristea et al., INTERSPEECH21]. Totodată, în problema de visual question answering, am reușit să reproducem rezultatele din lucrarea [Tan et al., EMNLP19]. Pentru a ajunge la rezultate de top în problemele studiate, am luat în calcul și crearea de ansamble de modele [Găman et al., IJIS22; Ristea et al., INTERSPEECH21]. Atingând performanțe state-of-the-art, putem concluzia că obiectivul 2 din propunerea de proiect a fost îndeplinit.

- **Activitatea 2.2. Task 3 – Dezvoltarea și evaluarea unui model neuronal deep pentru clasificarea textelor, bazat pe paradigma de învățare bazată pe curriculum.**

Această activitate corespunde cu obiectivul 3 din cadrul propunerii de proiect. În vederea îndeplinirii acestui obiectiv, am efectuat experimente pe o problemă de clasificare a textelor, anume detecția polarității opiniei, folosind un model de tip BERT. Pentru a introduce antrenarea graduală, bazată pe curriculum, am utilizat modelul de predicție a complexității dezvoltat în cadrul activității 1.1 din 2020. Pe lângă această abordare, am încercat să obținem o antrenare graduală utilizând dimensiunea vocabularului de cuvinte. Mai precis, vocabularul crește treptat, pe măsură ce modelul se optimizează. Rezultatele preliminare arată că învățarea bazată pe curriculum produce performanțe asemănătoare cu antrenarea standard, dar nu neapărat mai bune. În etapa următoare a proiectului, vom încerca să variem hiperparametrii implicați în învățarea bazată pe curriculum, cu scopul de a crește performanța peste metoda de antrenare standard.

- **Activitatea 2.3. Task 4 – Dezvoltarea și evaluarea unui model neuronal deep pentru visual question answering, bazat pe paradigma de învățare bazată pe curriculum.**

Am studiat posibilitatea aplicării învățării pe bază de curriculum pentru a antrena modele neuronale pentru diverse probleme din domeniul imaginilor: detectarea obiectelor, recunoașterea expresiilor faciale, estimarea vârstei persoanelor, clasificarea imaginilor obținute din semnale audio. Pentru problema detecției obiectelor, rezultatele obținute demonstrează că învățarea pe bază de curriculum conduce către performanțe mai bune. Metoda propusă și rezultatele aferente au fost sintetizate în lucrarea [Soviany et al., CVIU21]. De menționat că în vederea dezvoltării unui model pentru visual question answering, una din etapele importante este tocmai detectarea obiectelor în imagini. Astfel, studiul efectuat este relevant în îndeplinirea obiectivului 4. Pentru problemele de recunoașterea expresiilor faciale și estimarea vârstei persoanelor, am aplicat învățarea bazată pe curriculum folosind rețele neuronale de tip profesor-student. Metoda propusă și rezultatele aferente au fost sintetizate în lucrarea [Georgescu et al., MVA22]. Pentru clasificarea imaginilor obținute din semnale audio, am utilizat un algoritm de învățare care stabilește un curriculum pe baza etichetelor prezise de mai multe rețele neuronale convoluționale ce formează împreună un ansamblu. Metoda propusă și rezultatele aferente au fost sintetizate în lucrarea [Ristea et al., INTERSPEECH21]. Metodele de învățare bazată pe curriculum propuse în aceste lucrări [Georgescu et al., MVA22; Ristea et al., INTERSPEECH21; Soviany et al., CVIU21] au aplicabilitate imediată și în visual question answering. Astfel, în etapa următoare a proiectului, vom aplica abordările studiate pentru problema de visual question answering.

ETAPA 3/3

1. Etapa 3 – Rezumat:

În Etapa 3 – „Dezvoltarea și evaluarea unor modele neuronale deep pentru clasificarea textelor și visual question answering, bazate pe paradigma de învățare bazată pe curriculum” am efectuat activitățile 3.1, 3.2 și 3.3. Activitatea 3.1 are ca scop îndeplinirea completă a celui de-al treilea obiectiv, anume dezvoltarea și evaluarea unui model neuronal deep pentru clasificarea textelor, folosind paradigma de învățare bazată pe curriculum. Activitatea 3.2 are ca scop îndeplinirea completă a celui de-al patrulea obiectiv, anume dezvoltarea și evaluarea unui model neuronal deep pentru visual question answering, folosind paradigma de învățare bazată pe curriculum. Activitatea 3.3 are ca scop întocmirea prezentului raport.

2. Etapa 3 – Procentaj îndeplinire obiective la finalul etapei:

Obiectivul 1 – 100%
Obiectivul 2 – 100%
Obiectivul 3 – 100%
Obiectivul 4 – 100%
Obiectivul 5 – 100%

3. Etapa 3 – Descriere științifică și tehnică:

În conformitate cu activitățile prevăzute în Etapa 3 de raportare din Anexa II a contractului TE72/2020, am efectuat următoarele:

- **Activitatea 3.1. Task 3 – Dezvoltarea și evaluarea unui model neuronal deep pentru clasificarea textelor, bazat pe paradigma de învățare bazată pe curriculum.**

Această activitate corespunde cu obiectivul 3 din cadrul propunerii de proiect. În vederea îndeplinirii acestui obiectiv, am finalizat experimentele pe mai multe probleme de clasificare a textelor după diverse criterii, folosind 3 baze de date public disponibile: BoolQ [Clark et al., NAACL19], QNLI [Wang et al., ICLR19] și RTE [Wang et al., ICLR19]. În cadrul experimentelor, am folosit două modele state-of-the-art, anume BERT [Devlin et al., NAACL19] și LSTM [Hochreiter et al., NC97].

Deoarece abordarea bazată pe creșterea graduală a vocabularului încercată în cursul etapei 2.2 nu a condus la rezultate mai bune, am recurs la alte două strategii de învățare bazate pe curriculum. Prima dintre acestea, denumită „curriculum by smoothing” (CBS), a fost introdusă în lucrarea [Sinha et al., NeurIPS20] și a fost inițial aplicată doar pe imagini. CBS presupune netezirea imaginilor și activărilor straturilor convoluționale în iterațiile inițiale ale antrenării, netezirea fiind mai agresivă spre startul procesului, apoi redusă treptat pe măsură ce procesul de antrenare avansează. În acest fel, rețeaua va învăța în primele iterații informația de nivel grosier prezentă în imagini, detaliile fine devenind vizibile progresiv, odată cu deblurarea imaginilor. Noi am extins aplicabilitatea metodei CBS la texte, integrând metoda în modelele BERT și LSTM utilizate. A doua abordare reprezintă o strategie nouă introdusă în lucrarea [Croitoru et al., ArXiv22], fiind bazată pe ajustarea ratei de învățare pentru fiecare strat al rețelei. Cea de-a doua abordare se numește „learning rate curriculum” (LeRaC) și asignează fiecărui strat din rețea o rată de învățare astfel încât straturilor mai apropiate de intrare li se asociază o rată de învățare mai mare, iar, pe măsură ce adâncimea straturilor crește, rata de

învățare asignată este din ce în ce mai mică. Pe parcursul antrenării, ratele de învățare sunt actualizate astfel încât să convergă către aceeași valoare până la o anumită epocă. Ideea din spatele acestei metode este evitarea actualizării straturilor aflate mai aproape de ieșire la începutul antrenării pentru reducerea efectului produs de zgomotul cauzat de inițializarea rețelei.

Rezultatele comparative ale acestor metode sunt prezentate în Tabelul 2. După cum putem observa din acest tabel, performanțele obținute de abordarea noastră, numită LeRaC, sunt considerabil mai bune decât performanțele obținute aplicând modul convențional de antrenare. Prin urmare, considerăm că am reușit aplicarea cu succes a paradigmei de învățare bazată pe curriculum în domeniul clasificării textelor. Astfel, concluzionăm că obiectul 3 a fost îndeplinit cu succes.

Model	Paradigmă	BoolQ	RTE	QNLI
BERT	Convențională	74.12%	74.48%	92.13%
BERT	CBS	74.37%	74.97%	91.47%
BERT	LeRaC	75.55%	75.81%	92.45%
LSTM	Convențională	64.40%	54.12%	59.42%
LSTM	CBS	64.75%	54.03%	59.89%
LSTM	LeRaC	65.80%	55.71%	59.98%

Tabel 2. Rezultate obținute pe seturile de date BoolQ, QNLI și RTE folosind 3 strategii de antrenare aplicate modelelor BERT și LSTM, anume: convențională, CBS și LeRaC. Cele mai bune rezultate de pe fiecare coloană sunt marcate cu bold.

Observând succesul metodei LeRaC, am decis aplicarea acestei metode și în cazul recunoașterii obiectelor în imagini și recunoașterii vorbirii din semnale audio. În cazul din urmă, am aplicat LeRaC peste modelul SepTr, model ce a fost introdus în lucrarea [Ristea et al., INTERSPEECH22]. Rezultatele aferente sunt descrise în lucrarea [Croitoru et al., ArXiv22].

O altă contribuție în zona de procesare a textelor istorice constă în abordarea bazată pe ordonarea paginilor în funcție de numărul de rânduri adnotate și aplicarea acestei ordini pentru antrenarea unui model de detecție de rânduri. Paradigma de învățare bazată pe curriculum aplicată în acest caz a condus la rezultate semnificativ mai bune, așa cum rezultă din lucrarea [Găman et al., TiE22].

- **Activitatea 3.2. Task 4 – Dezvoltarea și evaluarea unui model neuronal deep pentru visual question answering, bazat pe paradigma de învățare bazată pe curriculum.**

Pentru problema visual question answering, am utilizat setul de date VQAv2 [Goyal et al., CVPR17], ce conține aproximativ 200.000 de imagini. Deoarece antrenarea unui model state-of-the-art, anume OFA [Wang et al., ICML22], pe întreg setul de date ar fi durat mai mult de 2 luni pe infrastructura de procesare disponibilă, am efectuat experimentele pe un scenariu de tip „few-shot learning”, păstrând din setul de date de antrenare inițial doar 100 de perechi formate din imagini și întrebări. Strategia de alegere a acestor perechi a fost următoarea: pentru fiecare din primele 10 cele mai frecvente tipuri de întrebări, am păstrat în mod aleator câte 10 exemple din setul de antrenare. În ceea ce privește setul de test, am păstrat doar exemplele care aparțin acestor cele mai frecvente 10 tipuri de întrebări, rezultând un set cu 1784 de exemple. Modelul OFA [Wang et al., ICML22] are 33 de milioane de parametrii, având la bază arhitectura ResNet-50 pentru procesarea imaginilor. Am antrenat acest model pentru 12000 de

iterații. Ca și în cazul activității 3.1, metodele de învățare bazată pe curriculum pe care le-am testat în experimentele noastre sunt „curriculum by smoothing” (CBS) [Sinha et al., NeurIPS20] și „learning rate curriculum” (LeRaC) [Croitoru et al., ArXiv22].

Model	Paradigmă	σ	k	Rată min	Rată max	Acuratețe
OFA	Convențională	-	-	-	-	44.51%
OFA	CBS	0-5	7	-	-	44.58%
OFA	CBS	0-2	5	-	-	44.64%
OFA	CBS	0-3	5	-	-	45.87%
OFA	LeRaC	-	-	5×10^{-8}	5×10^{-5}	44.09%

Tabel 3. Rezultate obținute pe setul de date VQAv2 (few-shot) folosind 3 strategii de antrenare a modelului OFA: convențională, CBS și LeRaC. Cele mai bune rezultate sunt marcate cu bold.

În Tabelul 3, sunt prezentate rezultatele obținute cu cele două metode de învățare bazată pe curriculum. Pentru CBS, am rulat 3 configurații diferite, în care am încercat diferite valori pentru deviația standard (σ) a filtrului gaussian și pentru dimensiunea acestuia (k). Deblurarea imaginilor are loc progresiv în primele 1000 de iterații, CBS fiind implementat prin schimbarea deviației standard a filtrelor. Se poate observa că toate experimentele cu CBS au depășit modelul baseline. Pentru LeRaC, încercările efectuate nu avut le fel de mult succes. Prezentăm un singur experiment în care rezultatul obținut este aproape de baseline. Pentru această metodă, este necesară stabilirea valorilor minime și maxime a ratei de învățare, valorile alese fiind menționate în tabel. În urma rezultatelor obținute cu metoda CBS, concluzionăm că am reușit aplicarea cu succes a paradigmei de învățare bazată pe curriculum în problema visual question answering. Prin urmare, susținem că obiectul 4 a fost îndeplinit cu succes.

- **Activități adiționale.**

Tot în cadrul proiectului, am realizat un articol de tip survey [Soviany et al., IJCV22] al paradigmei de învățare bazate pe curriculum, cuprinzând aici și metode aplicate în probleme de interes din cadrul proiectului, anume clasificarea textelor și visual question answering.

În vederea îndeplinirii celui de-al cincilea obiectiv, am publicat codul aferent diverselor modele și metode studiate pe platforma GitHub. Resursele publicate se regăsesc la următoarele adrese:

<https://github.com/adrianchifu/FreSaDa>

<https://github.com/ancatache/LaRoSeDa>

<https://github.com/iduta/coconv>

<https://github.com/antoniobarbalau/black-box-ripper>

<https://github.com/ristea/septr>

DISEMINAREA REZULTATELOR

În urma activităților de cercetare fundamentală efectuate în 2020, au rezultat 2 articole publicate în volume ale unor conferințe internaționale, dintre care 1 articol într-o conferință de categoria A* (NeurIPS 2020) și 1 articol într-un workshop de categoria B (VarDial 2020). Astfel, au fost îndeplinite cerințele minimale de diseminare a rezultatelor pe anul 2020, care prevedeau publicarea cel puțin a unui articol într-o conferință sau workshop de categoria B.

În urma activităților de cercetare fundamentală efectuate în 2021, au rezultat 7 articole publicate în volume ale unor jurnale sau conferințe internaționale, dintre care 1 articol într-un

jurnal din zona roșie (Q1), 1 articol într-o conferință de categoria A* (ACL 2021), 2 articole în conferințe de categoria A (INTERSPEECH 2021, EACL 2021), 1 articol într-o conferință de categoria B (IJCNN 2021) și două articole în workshop-uri (VarDial 2021, NeurArch 2021). Acestea se adaugă celor 2 articole publicate în 2020.

În urma activităților de cercetare fundamentală efectuate în 2022, au rezultat 6 articole publicate în volume ale unor jurnale sau conferințe internaționale, dintre care 2 articole în jurnale din zona roșie (Q1), 1 articol într-un jurnal din zona galbenă (Q2), 1 articol într-o conferință de categoria A* (NeurIPS 2022), 1 articol într-o conferință de categoria A (INTERSPEECH 2022), 1 articol într-un workshop (TiE 2022). Acestea se adaugă celor 9 articole publicate în anii anteriori, rezultând un număr total de 15 publicații finanțate din prezentul proiect.

Prin prisma rezultatelor obținute și a publicațiilor la conferințe și jurnale de top din domeniu, cum ar fi NeurIPS, ACL, INTERSPEECH, EACL, IJCV și CVIU, considerăm că impactul proiectului asupra domeniului de învățare automată va fi unul puternic. Conform Google Scholar, cele 15 articole finanțate din prezentul proiect însumează deja peste 120 de citări, dintre care cel mai citat (cu 47 de citări) este articolul de tip survey publicat în jurnalul IJCV. Totuși, considerăm că cel mai important rezultat al proiectului îl constituie lucrarea [Croitoru et al., ArXiv22], care a acumulat doar opinii pozitive din partea recenzorilor.

În concluzie, au fost îndeplinite cerințele minimale de diseminare a rezultatelor din proiect, care prevedeau publicarea a cel puțin 2 articole de categoria A sau A* (noi publicând 6 astfel de articole), cel puțin 2 articole de categoria B sau workshop-uri (noi publicând 5 astfel de articole) și cel puțin 1 articol de jurnal în zona roșie (noi publicând 3 astfel de articole). În plus, am publicat și 1 articol de jurnal în zona galbenă. Articolele finanțate prin proiectului de cercetare sunt listate în continuare:

- **2020:**

1. A. Bărbălău, A. Cosma, R.T. Ionescu, M. Popescu. Black-Box Ripper: Copying black-box models using generative evolutionary algorithms. In Proceedings of NeurIPS, pp. 20120–20129, 2020. **(Rank A* Conference)**

2. M. Găman, R.T. Ionescu. Combining Deep Learning and String Kernels for the Localization of Swiss German Tweets. In Proceedings of VarDial (COLING Workshop), pp. 242–253, 2020. **(Rank A/2 Workshop)**

- **2021:**

3. P. Soviany, R.T. Ionescu, P. Rota, N. Sebe. Curriculum Self-Paced Learning for Cross-Domain Object Detection. Computer Vision and Image Understanding, 204: 103166, 2021. **(Q1 journal)**

4. A. Tache, M. Găman, R.T. Ionescu. Clustering Word Embeddings with Self-Organizing Maps. Application on LaRoSeDa - A Large Romanian Sentiment Data Set. In Proceedings of EACL, pp. 949–956, 2021. **(Rank A Conference)**

5. M. Găman, S. Cojocariu, R.T. Ionescu. UnibucKernel: Geolocating Swiss German Jodels Using Ensemble Learning. In Proceedings of VarDial (EACL Workshop), pp. 84–95, 2021. **(Rank A/2 Workshop)**

6. R.T. Ionescu, A.G. Chifu. FreSaDa: A French Satire Data Set for Cross-Domain Satire Detection. In Proceedings of IJCNN, pp. 1–8, 2021. **(Rank B Conference)**

7. A.C. Rogoz, M. Găman, R.T. Ionescu. SaRoCo: Detecting Satire in a Novel Romanian Corpus of News Articles. In Proceedings of ACL, pp. 1073–1078, 2021. **(Rank A* Conference)**
8. N.C. Ristea, R.T. Ionescu. Self-paced ensemble learning for speech and audio classification. In Proceedings of INTERSPEECH, pp. 2836–2840, 2021. **(Rank A Conference)**
9. I.C. Duță, M.I. Georgescu, R.T. Ionescu. Contextual Convolutional Neural Networks. In Proceedings of NeurArch (ICCV Workshop), pp. 403–412, 2021. **(Rank A*/2 Workshop)**

- **2022:**

10. M. Găman, R.T. Ionescu. The Unreasonable Effectiveness of Machine Learning in Moldavian versus Romanian Dialect Identification. International Journal of Intelligent Systems, 37(8): 4928–4966, 2022. **(Q1 Journal)**
11. P. Soviany, R.T. Ionescu, P. Rota, N. Sebe. Curriculum Learning: A Survey. International Journal of Computer Vision, 130: 1526–1565, 2022. **(Q1 Journal)**
12. M.I. Georgescu, G. Duță, R.T. Ionescu. Teacher-Student Training and Triplet Loss to Reduce the Effect of Drastic Face Occlusion. Machine Vision and Applications, 33: 12, 2022. **(Q2 Journal)**
13. N.C. Ristea, R.T. Ionescu, F.S. Khan. SepTr: Separable Transformer for Audio Spectrogram Processing. In Proceedings of INTERSPEECH, 2022. **(Rank A Conference)**
14. M. Găman, L. Ghadamiyan, R.T. Ionescu, M. Popescu. Self-paced learning to improve text row detection in historical documents with missing labels. In Proceedings of TiE (ECCV Workshop), 2022. **(Rank A*/2 Workshop)**

REFERINȚE BIBLIOGRAFICE:

- [Bărbălău et al., NeurIPS20] A. Bărbălău, A. Cosma, R.T. Ionescu, M. Popescu. Black-Box Ripper: Copying black-box models using generative evolutionary algorithms. In Proceedings of NeurIPS, pp. 20120–20129, 2020.
- [Butnaru et al., BEA18] A. Butnaru, R.T. Ionescu. UnibucKernel: A kernel-based learning method for complex word identification. Proceedings of BEA-13, pp. 175–183, 2018.
- [Clark et al., NAACL19] C. Clark, K. Lee, M.W. Chang, T. Kwiatkowski, M. Collins, K. Toutanova. BoolQ: Exploring the Surprising Difficulty of Natural Yes/No Questions. In Proceedings of NAACL, pp. 2924–2936, 2019.
- [Devlin et al., NAACL19] J. Devlin, M.W. Chang, K. Lee, K. Toutanova. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of NAACL, pp. 4171–4186, 2019.
- [Croitoru et al., ArXiv22] F.A. Croitoru, N.C. Ristea, R.T. Ionescu, N. Sebe. LeRaC: Learning Rate Curriculum. In Proceedings of NeurIPS, 2022 (to appear).
- [Duță et al., ICCVW21] I.C. Duță, M.I. Georgescu, R.T. Ionescu. Contextual Convolutional Neural Networks. In Proceedings of NeurArch (ICCV Workshop), pp. 403–412, 2021.

- [Găman et al., VarDial21] M. Găman, S. Cojocariu, R.T. Ionescu. UnibucKernel: Geolocating Swiss German Jodels Using Ensemble Learning. In Proceedings of VarDial (EACL Workshop), pp. 84–95, 2021.
- [Găman et al., TiE22] M. Găman, L. Ghadamiyan, R.T. Ionescu, M. Popescu. Self-paced learning to improve text row detection in historical documents with missing labels. In Proceedings of TiE (ECCV Workshop), 2022.
- [Găman et al., VarDial20] M. Găman, R.T. Ionescu. Combining Deep Learning and String Kernels for the Localization of Swiss German Tweets. In Proceedings of VarDial (COLING Workshop), pp. 242–253, 2020.
- [Găman et al., IJIS22] M. Găman, R.T. Ionescu. The Unreasonable Effectiveness of Machine Learning in Moldavian versus Romanian Dialect Identification. *International Journal of Intelligent Systems*, 37(8): 4928–4966, 2022.
- [Georgescu et al., MVA22] M.I. Georgescu, G. Duță, R.T. Ionescu. Teacher-Student Training and Triplet Loss to Reduce the Effect of Drastic Face Occlusion. *Machine Vision and Applications*, 33: 12, 2022.
- [Goyal et al., CVPR17] Y. Goyal, T. Khot, Tejas, D. Summers-Stay, D. Batra, D. Parikh. Making the V in VQA Matter: Elevating the Role of Image Understanding in Visual Question Answering. In Proceedings of CVPR, pp. 6904–6913, 2017.
- [He et al., CVPR16] K. He, X. Zhang, S. Ren, J. Sun. Deep Residual Learning for Image Recognition. In Proceedings of CVPR, pp. 770–778, 2016.
- [Hochreiter et al., NC97] S. Hochreiter, J. Schmidhuber. Long Short-Term Memory. *Neural Computing*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [Ionescu et al., IJCNN21] R.T. Ionescu, A.G. Chifu. FreSaDa: A French Satire Data Set for Cross-Domain Satire Detection. In Proceedings of IJCNN, pp. 1–8, 2021.
- [Ristea et al., INTERSPEECH21] N.C. Ristea, R.T. Ionescu. Self-paced ensemble learning for speech and audio classification. In Proceedings of INTERSPEECH, pp. 2836–2840, 2021.
- [Ristea et al., INTERSPEECH22] N.C. Ristea, R.T. Ionescu, F.S. Khan. SepTr: Separable Transformer for Audio Spectrogram Processing. In Proceedings of INTERSPEECH, 2022.
- [Rogoz et al., ACL21] A.C. Rogoz, M. Găman, R.T. Ionescu. SaRoCo: Detecting Satire in a Novel Romanian Corpus of News Article.s. In Proceedings of ACL, pp. 1073–1078, 2021.
- [Sinha et al., NeurIPS20] S. Sinha, A. Garg, H. Larochelle. Curriculum By Smoothing. In Proceedings of NeurIPS, pp. 21653–21664, 2020.
- [Soviany et al., CVIU21] P. Soviany, R.T. Ionescu, P. Rota, N. Sebe. Curriculum Self-Paced Learning for Cross-Domain Object Detection. *Computer Vision and Image Understanding*, 2020, 204: 103166, 2021.

[Soviany et al., IJCV22] P. Soviany, R.T. Ionescu, P. Rota, N. Sebe. Curriculum Learning: A Survey. *International Journal of Computer Vision*, 130: 1526–1565, 2022.

[Tache et al., EACL21] A. Tache, M Găman, R.T. Ionescu. Clustering Word Embeddings with Self-Organizing Maps. Application on LaRoSeDa - A Large Romanian Sentiment Data Set. In *Proceedings of EACL*, pp. 949–956, 2021.

[Tan et al., EMNLP19] H. Tan, M. Bansal. LXMERT: Learning Cross-Modality Encoder Representations from Transformers. In *Proceedings of EMNLP*, pp. 5100–5111, 2019.

[Wang et al., ICLR19] A. Wang, A. Singh, J. Michael, F. Hill, O. Levy, S.R. Bowman. GLUE: A Multi-Task Benchmark and Analysis Platform for Natural Language Understanding. In *Proceedings of ICLR*, 2019.

[Wang et al., ICML22] P. Wang, A. Yang, R. Men, Rui, J. Lin, S. Bai, Z. Li, J. Ma, C. Zhou, Chang, J. Zhou, H. Yang. OFA: Unifying Architectures, Tasks, and Modalities Through a Simple Sequence-to-Sequence Learning Framework. In *Proceedings of ICML*, pp. 23318–23340, 2022.

Data,
31.08.2022

Director proiect,
Radu Tudor Ionescu