

**FACULTATEA DE AUTOMATICĂ ȘI CALCULATOARE**
DEPARTAMENTUL CALCULATOARE

SINTEZA

proiectului de diplomă cu titlul:

**REDUCEREA DIMENSIONALITĂȚII REPREZENTĂRII VIZUALE A IMAGINILOR OCT PENTRU
ANALIZA AFECȚIUNILOR OFTALMOLOGICE FOLOSIND CONTRASTIVE LEARNING**Autor: **Bianca VESA**Coordonator: **Conf. dr. ing. Anca Nicoleta MĂRGINEAN****1. Cerințele temei:**

Tomografia în coerență optică (*engl.* Optical Coherence Tomography - OCT) reprezintă o investigație modernă și non-invazivă, care se bazează pe lumină pentru a genera imagini în secțiune transversală ale retinei. Măsurătorile pe care le pune la dispoziție OCT sunt esențiale în identificarea și urmărirea răspunsurilor la tratament pentru anumite boli serioase, cum ar fi degenerescența maculară legată de vârstă (DMLV), care reprezintă cea mai răspândită cauză a pierderii vederii, întâlnită la populația cu vârstă de peste 60 de ani.

Această lucrare propune un model bazat pe metoda de învățare contrastive learning, care să obțină o reprezentare vizuală în spațiu dimensional redus a imaginilor de tip OCT, cu scopul de a analiza anumite afecțiuni oftalmologice, precum DMLV. Reprezentările au fost evaluate prin tehnici de clasificare și regresie, pentru a corela rezultatele procedurii OCT cu acuitatea vizuală a pacientului.

2. Soluții alese:

Specificul acestei lucrări este dat de faptul că:

1. include procesarea imaginilor OCT, care prezintă forme avansate ale afecțiunii DMLV, obținute de la 94 de pacienți consultați la Spitalul Clinic Județean de Urgență din Cluj-Napoca
2. modificările retinei, relevante monitorizării afecțiunii, sunt surprinse de un volum de imagini OCT (în număr de 25 - pot fi referite ca B-scans), care are asociată o valoare a acuității vizuale măsurată de medicii oftalmologi

Lucrarea prezintă detaliază un sistem care include procesarea datelor din setul DMLV și utilizarea lor în analiza la nivel de volum a imaginilor OCT. O abordare obișnuită de tip clasificare supervizată nu va aduce rezultate utile, în special datorită numărului redus de volume, motiv pentru care s-a abordat metoda contrastive learning pentru a învăța o reprezentare redusă, ce include principalele trăsături din cadrul unui B-scan. Sistemul dezvoltat este ilustrat în **Figura 1**.

Așadar, soluția propusă include:

- A. Dezvoltarea unui model bazat pe contrastive learning, care include un modul de tip encoder, antrenat pentru a extrage trăsăturile cele mai importante din imaginile OCT. Antrenarea urmează atât învățarea de tip self-supervised, conform framework-ului SimCLR, propus de *Chen et al.* în articolul [2], cât și cea de tip supervised, propusă de *Khosla et al.* în lucrarea [3], sub numele de SupCon.

- B. Evaluarea celor două tipuri de antrenare a encoder-ului, prin includerea modelului într-un task de clasificare a setului propus de *Kermany et al.* în articolul [1], care cuprinde 83484 de imagini de tip B-scan individuale, grupate în 4 clase, în funcție de afecțiunile oftalmologice pe care le prezintă.
- C. Selectarea encoder-ului care a adus rezultatele cele mai bune, pentru a extrage informațiile din imaginile OCT din setul DMLV, sub forma unor vectori de trăsături.
- D. Concatenarea vectorilor de trăsături care corespund unui volum de B-scans, într-o singură imagine volumetrică și asocierea ei cu acuitatea vizuală aferentă.
- E. Alcătuirea unui nou set de date format din imagini volumetrice, cu clasele reprezentate de valorile acuității vizuale.
- F. Evaluarea calității reprezentărilor volumetrice rezultate de encoder, cu ajutorul unui model de clasificare și a unui model de regresie, care să prezică acuitatea vizuală asociată unei instanțe de imagine OCT volumetrică.

3. Rezultate obținute:

Pentru modelul de tip contrastive learning self-supervised în task-ul de clasificare, acuratețea maximă pe setul de validare este 88%, iar acuratețea maximă aferentă pe setul de test Kermany este 95%, având 19 predicții fals-pozitive (instanțe de imagini care prezentau afecțiuni ale retinei, însă au fost clasificate ca fiind normale) din 1000 de predicții. Modelul de tip contrastive learning supervised a obținut o acuratețe maximă de 95% pe setul de validare și 99% pe setul de test, cu 0 predicții fals-pozitive.

Calitatea reprezentărilor encoder-ului din modelul supervised contrastive a fost evaluată cu ajutorul unei rețele neuronale convoluționale de clasificare a acuității vizuale, care a adus o acuratețe maximă pe setul de validare de 45%, la antrenarea cu setul de date grupat în 12 clase de acuitate. Pentru 2 clase de acuitate (valori ≤ 0.1 , respectiv > 0.1), acuratețea maximă pe setul de validare a fost de 88%, însă cu apariția fenomenului de overfit. Modelul de tip regresie, a oferit rezultate mai promițătoare, rezultând în valorile următoare la evaluarea pe setul de test: root mean squared error = 0.18, mean absolute error = 0.12, unde domeniul de valori pentru acuitate este $[0, 1]$.

4. Testări și verificări:

Modelul contrastive learning a fost antrenat în task-ul de clasificare în cele 4 clase de afecțiuni oftalmologice, pe baza setului de date Kermany-train, atât folosind funcția supervised loss, care include și etichetele setului de date, cât și folosind funcția self-supervised loss, fără a include etichetele. Setul a fost împărțit în 80% imagini pentru antrenare și 20% imagini pentru validare. În urma testelor pe setul Kermany-test de 1000 de imagini, sistemul final include doar encoder-ul de tip supervised, care a dat rezultate mult mai bune. S-au efectuat teste care variau dimensiunea vectorului de trăsături construit cu ajutorul encoder-ului, cu valorile 64, 128, 512, 1024, iar valoarea care a adus rezultatele cele mai bune este 128.

Componenta encoder din modelul contrastive a fost utilizată pentru a transforma imaginile din setul DMLV, în reprezentări de dimensiune 128, care au fost concatenate. Un volum cuprinde 25 de B-scans, rezultând o reprezentare finală într-o imagine de dimensiune 25×128 , care a fost asociată valorii acuității vizuale corespunzătoare. Clasificatorul pentru predicția acuității a fost o simplă rețea neuronală convoluțională, iar modelul de regresie ales este Support Vector Regression.

5. Contribuții personale:

- Utilizarea paradigmei contrastive learning în procesul de învățare al encoder-ului și analiza empirică a arhitecturii optime a modelului de clasificare pe setul Kermany
- Extragerea trăsăturilor din imaginile DMLV cu ajutorul encoder-ului și concatenarea vectorilor pentru a obține reprezentări de dimensiune redusă ale volumelor de B-scans

- Dezvoltarea unui clasificator și antrenarea lui pentru predicția acuității vizuale
- Utilizarea unui algoritm machine learning de regresie pentru predicția acuității vizuale

6. Surse de documentare:

- [1] D. Kermany et al., “Identifying medical diagnoses and treatable diseases by image-based deep learning,” vol. 172, no. 5, 2018, pp. 1122—1131. [Online]. Available: <https://doi.org/10.1016/j.cell.2018.02.010>
- [2] T. Chen et al., “A simple framework for contrastive learning of visual representations,” CoRR, vol. abs/2002.05709, 2020. [Online]. Available: <https://arxiv.org/abs/2002.05709>
- [3] P. Khosla et al., “Supervised contrastive learning,” in Advances in Neural Information Processing Systems, H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin, Eds., vol. 33. Curran Associates, Inc., 2020, pp. 18 661–18 673.

Data: 05.07.2022

Autor *Bianca*

Coordonator _____

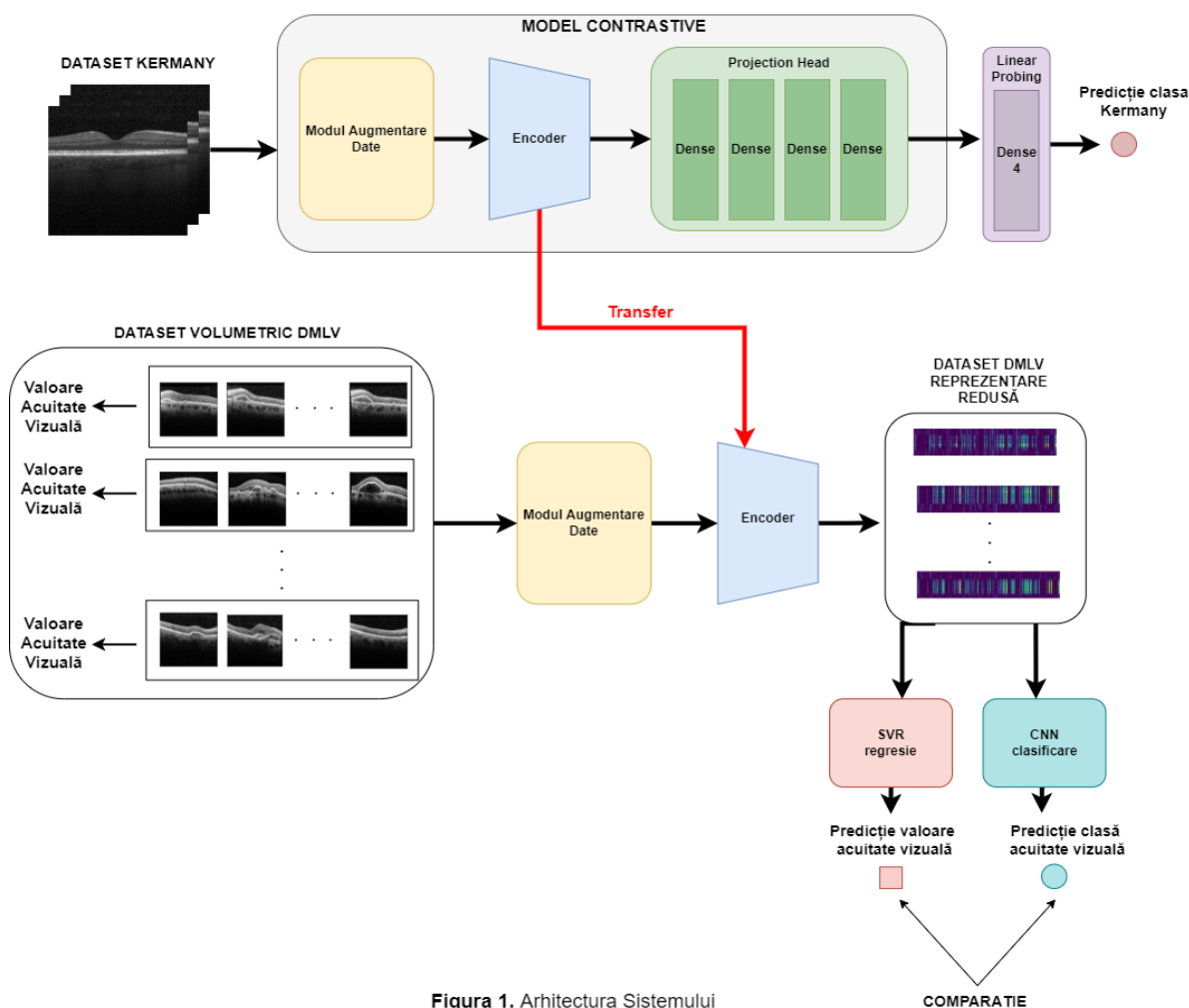


Figura 1. Arhitectura Sistemului