

# Clasificarea bolilor dermatologice

Clivet Matei-Constantin  
Ciobanu Mircea-Razvan

# 1. Context & Motivație

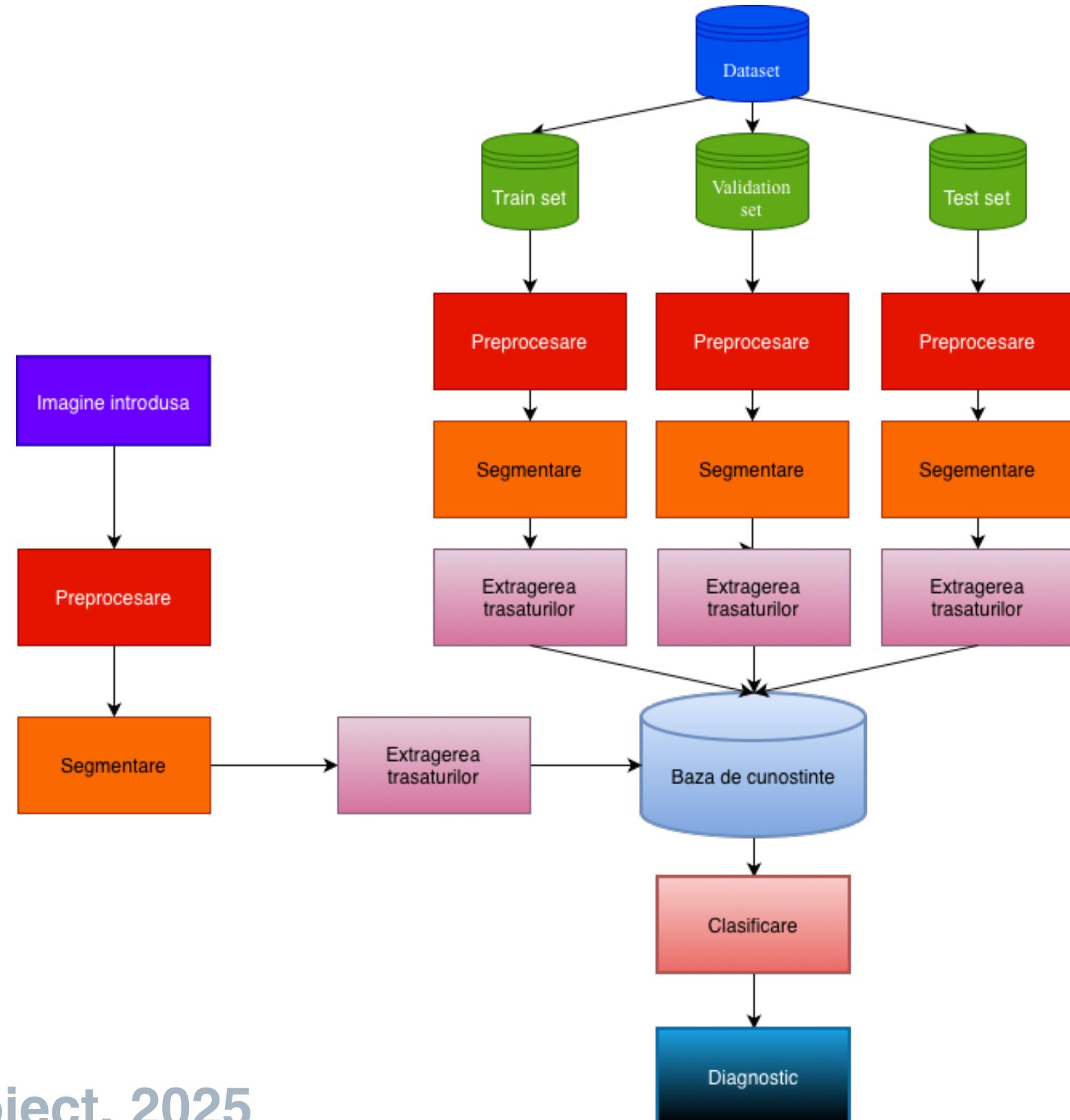
- **Context:** În ultimii ani, bolile dermatologice au devenit o problemă tot mai frecventă la nivel global. Diagnosticul corect al acestora depinde în mare măsură de experiența medicilor dermatologi și de aparatul la care acestia au acces. Astfel, nevoia de dezvoltare a unor metode inovative și eficiente de identificare a bolilor dermatologice este una ce crește în fiecare an.
- **Motivație:** Sistemul de detectare automata a bolilor dermatologice contribuie la oferirea timpurie a unui diagnostic, reducerea erorilor de diagnostic, o accesibilitate sporita pentru centrele cu resurse medicale limitate și la reducerea costurilor prin eliminarea consulturilor suplimentare.
- **Obiectivul proiectului:** Problema pe care o adresează proiectul este identificarea și clasificarea automata a bolilor dermatologice folosind tehnici de procesare a imaginilor și machine learning pentru sprijinul unui diagnostic medical rapid și precis.

# Clasificarea bolilor dermatologice

## 2. Arhitectura preliminară a soluției

- **Descrierea componentelor:**
- **Dataset:** Setul de date de la care pornim, imaginile pe care le folosim in training-ul bazei de cunoștințe.
- **Train set:** subset-ul de date folosit pentru a invata modelul. (70% din total)
- **Validation set:** subset-ul folosit pentru verificarea modului în care modelul invata. (15% din total)
- **Test set:** subset-ul folosit pentru testarea reală a modelului. (15% din total)

- **Schema arhitecturii:**



## 2. Arhitectura preliminară a soluției

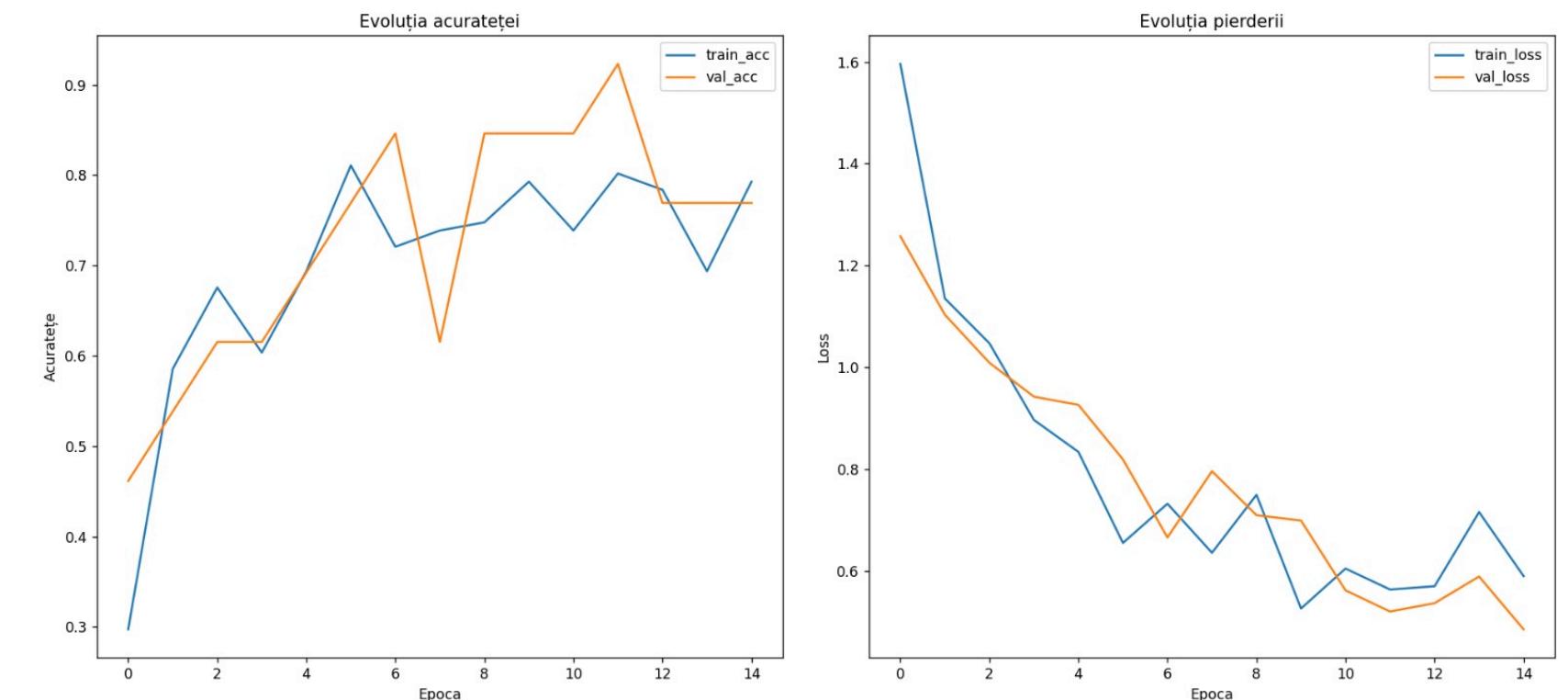
- **Preprocesare:** Realizam toate procesele de preprocesare ale imaginilor, cum ar fi redimensionarea, conversia la grayscale, eliminarea zgomotului.
- **Segmentare:** Detectarea marginilor, thresholding.
- **Extragerea trasaturilor:** Detectarea trasaturilor complexe, cum ar fi formele sau regiunile tipice leziunilor sau a bolilor.
- **Baza de cunoștințe:** Reteaua neuronală convolutională în sine, adică totalitatea operațiilor de convoluție aplicate imaginilor.
- **Clasificare:** Pasul final de clasificare în care modelul decide din ce categorie aparține imaginea.
- **Diagnostic:** Ofere diagnosticul final și acuratetea.

## 2. Arhitectura preliminară a soluției

- Fluxul de date:** Imaginele din dataset sunt impartite în 4 categorii, ca apoi să fie folosite în număr majoritar de către train set pentru antrenarea modelului. Validation set asigură învățarea corectă, prin testarea în timpul antrenamentului. În test set se află imagini care nu sunt văzute de către model în timpul antrenamentului, folosite pentru testarea să după ce etapa de învățare a avut loc. Cele 3 realizează rețea una convolutională, folosită apoi pentru detectarea bolii din imaginea introdusa.

### 3. Evaluarea Preliminară a Soluției

- **Metodologia de evaluare:** Solutia a fost testata pe un set de aproximativ 156 de imagini luate din diferite dataset-uri pentru 4 boli: acnee, psoriazis, melanom si eczeme. Metrica principala folosita a fost acuratețea, adica procentul de imagini clasificate corect pentru train si validation set.
  - **Setul de date:** Setul de date folosit este un cumul de 156 de imagini din diferite dataset-uri stiintifice si medicale care este impartit in 4 categotii. unul pentru fiecare boala.
  - **Exemple de cazuri de test:**



# 4. Rezultate Preliminare

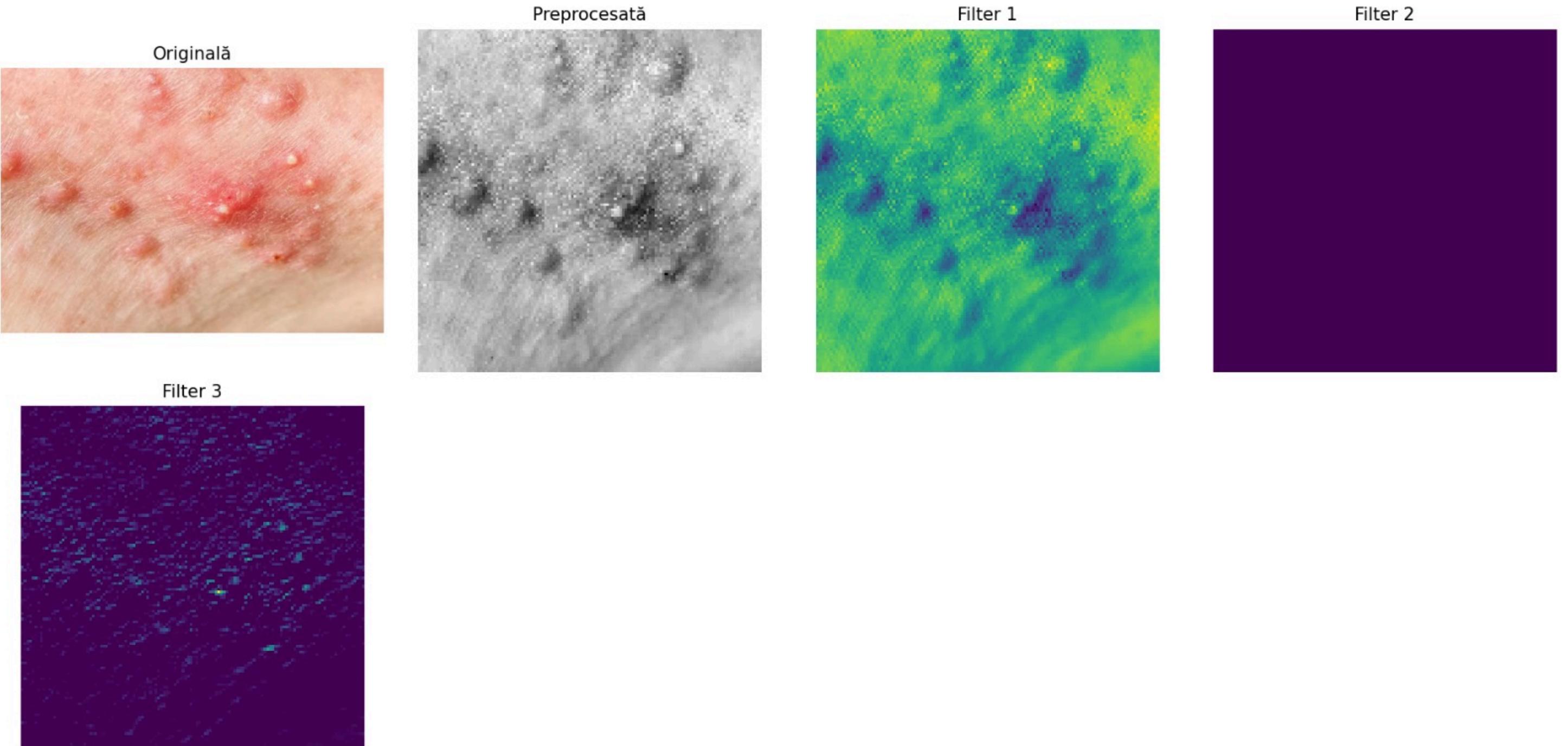
- **Rezultate obținute:**

```
Epoch 1/15
28/28 4s 86ms/step - accuracy: 0.4234 - loss: 1.3789 - val_accuracy: 0.3077 - val_loss: 1.2286
Epoch 2/15
28/28 2s 75ms/step - accuracy: 0.5856 - loss: 0.9964 - val_accuracy: 0.5385 - val_loss: 1.1404
Epoch 3/15
28/28 2s 74ms/step - accuracy: 0.6126 - loss: 0.9220 - val_accuracy: 0.4615 - val_loss: 1.0645
Epoch 4/15
28/28 2s 75ms/step - accuracy: 0.7027 - loss: 0.7991 - val_accuracy: 0.5385 - val_loss: 0.7839
Epoch 5/15
28/28 2s 76ms/step - accuracy: 0.7207 - loss: 0.8094 - val_accuracy: 0.8462 - val_loss: 0.7352
Epoch 6/15
28/28 2s 74ms/step - accuracy: 0.6847 - loss: 0.7404 - val_accuracy: 0.9231 - val_loss: 0.5817
Epoch 7/15
28/28 2s 75ms/step - accuracy: 0.7297 - loss: 0.6876 - val_accuracy: 0.7692 - val_loss: 0.6188
Epoch 8/15
28/28 2s 75ms/step - accuracy: 0.7027 - loss: 0.6562 - val_accuracy: 0.7692 - val_loss: 0.5637
Epoch 9/15
28/28 2s 77ms/step - accuracy: 0.7117 - loss: 0.7187 - val_accuracy: 0.6923 - val_loss: 0.6638
Epoch 10/15
28/28 2s 76ms/step - accuracy: 0.7207 - loss: 0.6564 - val_accuracy: 0.8462 - val_loss: 0.4548
Epoch 11/15
28/28 2s 74ms/step - accuracy: 0.7838 - loss: 0.6100 - val_accuracy: 0.7692 - val_loss: 0.5174
Epoch 12/15
28/28 2s 73ms/step - accuracy: 0.7658 - loss: 0.5446 - val_accuracy: 0.9231 - val_loss: 0.3944
Epoch 13/15
28/28 2s 75ms/step - accuracy: 0.7477 - loss: 0.6090 - val_accuracy: 0.7692 - val_loss: 0.4247
Epoch 14/15
28/28 2s 74ms/step - accuracy: 0.7568 - loss: 0.5467 - val_accuracy: 0.7692 - val_loss: 0.6410
```

# Clasificarea bolilor dermatologice

## 4. Rezultate Preliminare

- Rezultate obținute:



## 4. Rezultate Preliminare

- Interpretarea rezultatelor:** De la prima pana la a 15-a epoca avem o evolutie a acuratetei si o scadere a pierderilor, iar filtrele functioneaza. Totusi, modelul nu are suficiente imagini cat sa se antreneze corespunzator in identificarea categoriilor. In faza aceasta , modelul reuseste sa identifice, cu aproximatie, 2 din 5 imagini introduse.

## 5. Concluzii Preliminare

- **Rezumatul progresului:** Pana in prezent am realizat pregatirea datelor, etapa de preprocessare, antrenarea modelului si introducerea unui sistem de evaluare a rezultatelor.
- **Limitările soluției actuale:** Clase inegale, dataset mic, lipsa etapei explicite de segmentare.
- **Potențiale îmbunătățiri:**
  - Adaugarea unei etape de segmentare explicite.
  - Im bunatatiarea dataset-ului cu un numar mai mare de date.(Ex: HAM10000)

# Clasificarea bolilor dermatologice

## 6. Directii Viitoare

- **Pasi următori:** Introducerea explicita a segmentarii si crearea unui fisier main care sa permita rularea intregului program dintr-un singur fisier.
- **Plan de implementare:** Crearea unui fisier numit download\_dataset.py care sa permita instalarea la nivel local a imaginilor. Adaugarea pasilor de segmentare in train\_model.py. Imbunatatirea generala a structurii proiectului.
- **Obiectivele finale:** O aplicatie care primeste imagini aleatorii din cele 4 categorii de boli si le clasifica cu o acuratete de macar 80%.
- Folosirea telefonului ca element de achizitie a imaginilor.