

# **Recunoașterea caracterelor inscripționate pe obiecte ceramice folosind tehnici de vedere artificială și învățare automată profundă**

**Autor: Oancea Maria Nicoleta**

**Profesori coordonatori: Conf. univ. dr. Bîrluțiu Adriana, Conf. univ. dr. ing. Kadar Manuella**

**Instituție: Universitatea "1 Decembrie 1918" din Alba Iulia**

*This application is developed to help the companies from the porcelain industry to identify the engraved characters on plates. In the first phase, the application identifies the part of the object where the characters are engraved and then it starts to section the image to extract the characters or the group of characters. After all groups of characters were identified, it uses a convolutional neural network to recognize the characters from that image.*

## **1. Introducere**

Tema acestei lucrări este dezvoltarea unei aplicații care este capabilă să recunoască caracterele inscripționate pe obiecte ceramice. Pentru aceasta este important să se identifice zonele în care sunt ștanțate caracterele, să se extragă aceste caractere din imagine și să se folosească o rețea neuronală care va returna valoarea caracterelor.

Această automatizare răspunde cerințelor din mediul industrial, universitatea noastră colaborând cu o companie din industria porțelanului, iar această cercetare vine în sprijinul acestei colaborări. Prin implementarea acestui proiect se dorește înlocuirea personalului care se ocupă cu verificarea calității produselor, astfel reducându-se costurile și timpii de producție.

Există mai multe lucrări științifice care analizează problema recunoașterii caracterelor folosind tehnici de vedere artificială și învățare profundă. Această lucrare este o extindere a lucrării "Automatic Character Recognition In Porcelain Ware"[1] care se concentrează pe recunoașterea caracterelor inscripționate pe farfuriile rotunde. În prezenta lucrare se analizează un alt tip de obiect ceramic, farfuria dreptunghiulară, și se dorește recunoașterea caracterelor inscripționate pe acestea.

Această cercetare s-a desfășurat în cadrul bursei de performanță științifică acordată de către Universitatea "1 Decembrie 1918" din Alba Iulia pe perioada ianuarie – iunie 2019.

## **2. Tehnologii folosite**

### **2.1. Tehnici de vedere artificială**

Vederea artificială este un domeniu științific care se află la intersecția dintre informatică, matematică, inginerie, biologie și psihologie, având câte puțin din fiecare [2].

Pentru a putea recunoaște caracterele dintr-o imagine este important ca imaginea să fie de o calitate cât mai bună. Dacă însă nu dispunem de astfel de imagini ele trebuie procesate, aplicând diferite filtre asupra lor. Procesarea presupune eliminarea zgomotului, accentuarea zonelor de interes, renunțarea la detaliile ne semnificative, etc.

Segmentarea imaginii este procesul prin care o imagine se împarte în mai multe zone cu scopul de a simplifica și de a scoate în evidență doar zonele semnificative ce urmează a fi analizate [3]. Prin analogie în cazul farfuriilor ne interesează doar zonele în care sunt ștanțate caracterele ce urmează a fi recunoscute.

## 2.2. Rețele neuronale adânci

Rețelele neuronale adânci sunt o ramură a învățării automate cu foarte multe aplicații de succes. După cum îi spun și numele, o rețea neuronală adâncă este compusă din mai multe straturi ascunse, ceea ce duce la o performanță mai bună în comparație cu rețelele simple. Straturile ascunse multiple îi permit rețelei să învețe prin tehnica ierarhiei de caracteristici, adică fiecare strat se ocupă de anumite caracteristici, care se combină de la un strat la altul pentru a produce un rezultat mai bun [4].

În prezent, algoritmi bazați pe rețele neuronale adânci au stabilit noi recorduri în ceea ce privește rezultatele obținute în diverse probleme practice. Acest tip de rețele excelează în recunoașterea imaginilor și a sunetului, dar și în sistemele de recomandări. Ca și exemplu, rețelele profunde fac parte din algoritmul dezvoltat de *DeepMind*, *AlphaGo*, care a bătut campionul mondial, Lee Sedol, la Go din 2016 și pe noul campion mondial din 2017, Ke Jie [5].

## 2.3. Rețele neuronale convoluționale

Rețelele neuronale artificiale reprezintă un mod de recreare a rețelelor neuronale biologice. În mod asemănător creierului uman acestea folosesc exemple pentru a învăța informații noi. Rețelele neuronale se pot clasifica după două criterii fundamentale: după modul de propagare a semnalului se împart în două subcategorii: rețele feed-forward și rețele recurente (feed-back); după modul de învățare se împart în următoarele subcategorii: rețele cu învățare supervizată și rețele cu autoorganizare [6].

Rețelele neuronale convoluționale (en. *Convolutional Neural Network (CNN)*) au fost menționate pentru prima dată în 1940, dar au început să fie utilizate doar în ultima perioadă deoarece aveau nevoie de resurse hardware ce în anii 1940 nu erau disponibile. Aceste rețele neuronale sunt compuse din mai multe straturi prin care datele sunt filtrate în categorii. CNN-urile sunt utilizate cel mai adesea în recunoașterea imaginilor și prelucrarea textului [7].

Straturile convolutive sunt denumite așa după operația matematică de convoluție. Convoluția este o operație în care două funcții produc o a treia funcție care arată cum forma uneia este modificată de cealaltă. Convoluția este similară corelației încrucișate [8]. Rețelele neuronale convoluționale sunt printre cele mai comune și robuste rețele neuronale. Cercetătorii le folosesc într-o mulțime de domenii inclusiv în medicină, unde s-au dovedit a fi eficiente în a da diagnostice corecte, asemănătoare cu cele ale unui medic cu mulți ani de experiență [9].

Unul din principalele obstacole în calea răspândirii CNN-urilor o constituie necesitatea unui spațiu mare de stocare a datelor.

## 3. Analiza experimentală a problemei

### 3.1. Limbajul și pachetele utilizate

Pentru implementarea problemei abordate în această lucrare s-a folosit Matlab. Pachetele utilizate sunt: Deep Learning Toolbox, Image Processing Toolbox și Statistics and Machine Learning Toolbox [10].

**Matlab** este un limbaj de programare bazat pe matrici care poate fi integrat cu alte limbaje de programare. Ingineri și oameni de știință din întreaga lume îl folosesc pentru a analiza și proiecta sisteme și produse care ne influențează viața.

**Deep Learning Toolbox** oferă un framework pentru proiectarea și implementarea rețelelor neuronale profunde. Acest pachet vine cu un set de algoritmi, modele precompilate și aplicații care ușurează munca de dezvoltare a unui algoritm de recunoaștere a formelor, obiectelor sau caracterelor, sau a oricărui alt algoritm de *deep learning*.

**Image Processing Toolbox** oferă o gamă variată de algoritmi de procesare, analiză, vizualizare și dezvoltare de imagini. Acest pachet poate fi folosit pentru segmentarea și îmbunătățirea imaginilor, reducerea zgomotului și pentru diverse transformări geometrice.

**Statistics and Machine Learning Toolbox** oferă funcții și aplicații pentru a descrie, analiza și modela datele. Setul de instrumente conține algoritmi de învățare automată supervizată și nesupervizată.

### 3.2. Setul de date utilizat

Setul de date utilizat conține 50 de imagini cu farfurii fabricate de o companie specializată în industria porțelanului. Imaginile folosite pentru această aplicație conțin farfurii dreptunghiulare care au caracterele inscripționate atât pe partea dreaptă cât și pe partea stângă. În Figura 1 sunt reprezentate două farfurii de acest tip, prima are caracterele inscripționate pe partea dreaptă iar a doua pe partea stângă.



**Figura 1: Imagini analizate pentru recunoașterea caracterelor ștanțate pe farfurie.**

### 3.3. Descrierea aplicației

Pentru a putea identifica caracterele inscripționate pe farfurii este important să se identifice zona în care se afla acestea, să se stabilească dacă este un singur caracter sau dacă este un grup de caractere. Pentru a rezolva această problemă s-a folosit procedeul de segmentare a imaginii [1].

Primul pas este găsirea conturului farfuriei și decuparea zonei. Pentru aceasta, imaginea este convertită în alb negru, după care se elimină toate obiectele conectate care au un număr de pixeli mai mic decât valoarea specificată folosind funcția *bwareaopen*. Identificarea zonei în care se află farfuria se face folosind funcția *regionprops* care identifică regiunile conectate și măsoară proprietățile acestora. Secvența de cod care face cele menționate dar și alte operațiuni de procesare a imaginii este reprezentată în Figura 2, iar rezultatul obținut se regăsește în Figura 3.

```
% to gray
Ig = rgb2gray(I);
% to black and white
Ibw = im2bw(Ig,graythresh(Ig));

% erode holes
Ibw2 = bwareaopen(Ibw, 80, 8);
SE = strel('disk',4);
Ibw2 = imerode(Ibw2,SE);

% Find conected region and their properties
stats = regionprops(Ibw2,'boundingbox','Area','Perimeter');
```

Figura 2: Procesarea imaginii și identificarea zonei în care se găsește farfuria.

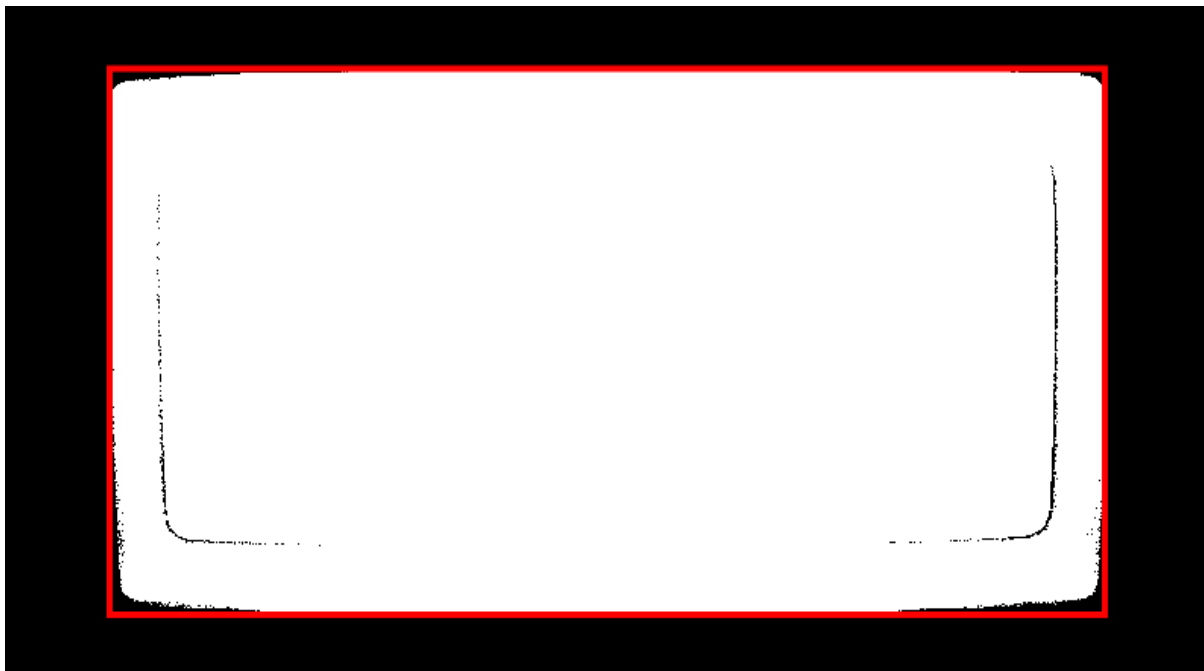


Figura 3: Încadrarea zonei în care se găsește farfuria.

Următorul pas este identificarea zonei în care sunt inscripționate caracterele, în partea dreaptă sau în partea stângă a farfuriei. Pentru aceasta se folosește funcția *edge* care identifică contururile din imagine, cu cât valoarea pragului este mai mică cu atât identifică mai multe margini. Imaginea

obținută se dilatează folosind *imdilate*, accentuându-se zonele în care se găsesc caracterele. După aceasta se aplică din nou funcția de identificare a zonelor conectate. Se împarte imaginea în două părți și se numără zonele conexe. În funcție de numărul zonelor conectate predominante în una din părți, se stabilește dacă caracterele sunt ștanțate pe partea stângă sau pe partea dreaptă. Secvența de cod descrisă se regăsește în Figura 4, iar rezultatele sunt reprezentate în Figura 5.

```
Ig = rgb2gray(I1);  
Ig_ad = imadjust(Ig);  
% Edge detection  
I_edge = edge(Ig_ad, 'Canny', 0.15, 2.0);  
% Dilate edge  
SE = strel('Disk', 18, 4);  
I_edge = imdilate(I_edge, SE);  
% Find connected region and their properties  
stats = regionprops(I_edge, 'Area', 'BoundingBox', 'Orientation');
```

Figura 4: Secvența de cod prin care se identifică zona în care sunt inscripționate caracterele.

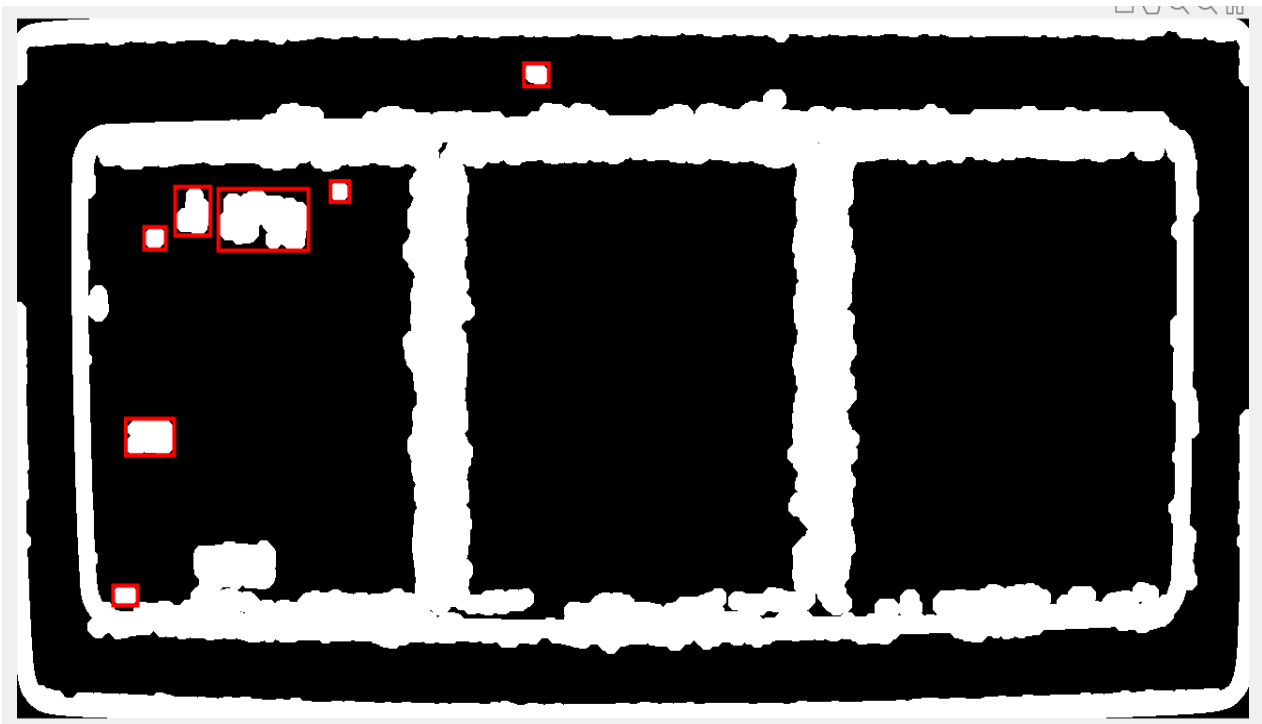
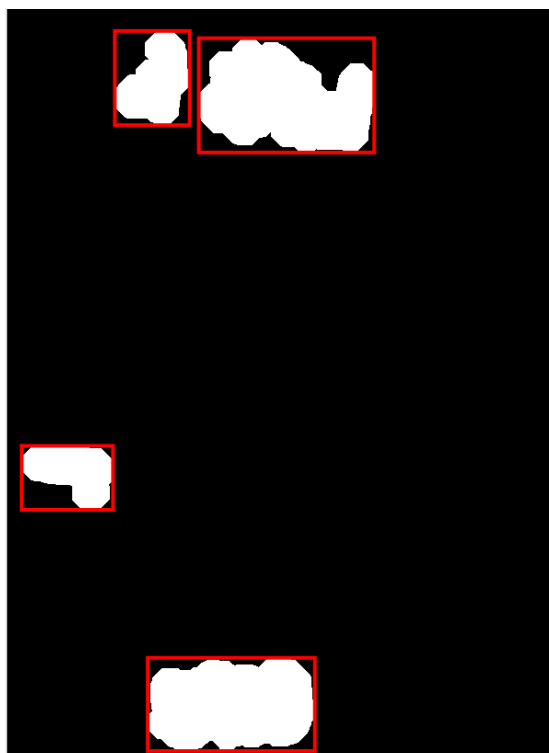


Figura 5: Zona în care sunt ștanțate caracterele.

După identificarea părții (stânga sau dreapta) pe care se află caracterele se trece la identificarea zonei în care sunt inscripționate caracterele. Pentru aceasta se folosesc tehnicile menționate mai sus. Rezultatul obținut este reprezentat în Figura 6.



**Figura 6: Marcarea zonelor în care sunt inscripționate caracterele.**

În pasul următor se trece la segmentarea imaginii în patru imagini care conțin zonele unde sunt ștanțate caracterele. Aceste imagini sunt reprezentate în Figura 7.



**Figura 7: Caracterele și grupurile de caractere după extragerea din imagine.**

Pentru fiecare imagine în parte se trece la recunoașterea caracterului sau a caracterelor inscripționate. Ca prim pas algoritmul s-a concentrat doar pe recunoașterea caracterului "E". Pentru recunoașterea acestui caracter se folosește o rețea neuronală deja antrenată pe recunoașterea cifrelor și a literelor. Secvența de cod care clasifică caracterele este reprezentată în Figura 8.

```
332 - grayIm = imresize(grayIm, [53 26]);
333 - [label, score] = classify(convnet, grayIm(:, :, 1));
334 - labels = [labels label];
335 - scores = [scores max(score)];
336
```

---

Command Window

```
ans =
    'E'
```

**Figura 8: Secvența de cod care clasifică caracterele din imagine.**

Pentru identificarea caracterelor se folosește funcția *classify* care primește ca input rețeaua neuronală și imaginea, iar ca rezultat returnează valoarea caracterului inscripționat în imagine.

### 3.4. Rezultate obținute

Din cele 50 de imagini analizate, a fost identificat corect caracterul "E" în 10 (20%) dintre acestea, în 30 (60%) dintre ele s-a obținut ca și rezultat caracterul "9" (care este apropiat de E), iar în celelalte 10 (20%) imagini s-au clasificat ca fiind alte caractere.

Aceste rezultate s-au obținut din cauză că s-a folosit o rețea neuronală care a fost antrenată pe un alt tip de farfurii, în care caracterul "9" seamănă mult cu caracterul "E" din tipul curent de farfurii. Un alt motiv al obținerii acestor rezultate este calitatea mult mai scăzută a imaginilor. În Figura 9 sunt reprezentate cele două caractere, în partea stângă este caracterul "9" inscripționat pe primul tip de farfurii, iar în partea dreaptă este caracterul "E" ștanțat pe actualul tip de farfurii.



**Figura 9: Reprezentarea caracterului "9" și "E".**

## 4. Concluzii și direcții de cercetare viitoare

În concluzie această aplicație vine în sprijinul companiilor din industria porțelanului care au nevoie de identificarea caracterelor inscripționate pe farfurii.

Pe viitor această aplicație va fi extinsă pentru a recunoaște caracterele ștanțate și pe alte tipuri de farfurii. Se poate trece și la un pas mult mai avansat acela de a utiliza această aplicație pentru recunoașterea caracterelor marcate și pe alte obiecte.

Tot ca direcție viitoare se dorește re-antrenarea rețelei neuronale cu o gamă mult mai variată de caractere, ceea ce va duce la îmbunătățirea semnificativă a rezultatelor obținute în recunoașterea caracterelor.

## Mulțumiri

Această lucrare a fost susținută de Universitatea "1 Decembrie 1918" din Alba Iulia prin intermediul bursei de performanță științifică din perioada ianuarie - iunie 2019.

## Bibliografie

- [1] D.A. Sanchez, S.G. Bulon, L. Moreno, A. Birlutiu, M. Kadar. Automatic Character Recognition in Porcelain Ware. ACTA TECHNICA NAPOCENSIS – Electronica - telecomunicatii (Electronics and Telecommunications) ISSN 1221 – 6542, vol. 59, nr.3/2018-seria electronica.
- [2] James Le, The 5 Computer Vision Techniques That Will Change How You See The World, <https://heartbeat.fritz.ai/the-5-computer-vision-techniques-that-will-change-how-you-see-the-world-1ee19334354b>
- [3] Linda G. Shapiro and George C. Stockman, Computer Vision, New Jersey, Prentice Hall, 2001
- [4] Artificial Intelligence (AI) vs. Machine Learning vs. Deep Learning, <https://skymind.ai/wiki/ai-vs-machine-learning-vs-deep-learning>
- [6] Ileana Ioan, Rotar Corina, Muntean Maria, Inteligență artificială, Aeternitas, 2009.
- [7] Tristan Greene, A beginner's guide to AI: Neural networks, <https://thenextweb.com/artificial-intelligence/2018/07/03/a-beginners-guide-to-ai-neural-networks/>
- [8] Convolution, <http://mathworld.wolfram.com/Convolution.html>

- [9] Nicolas Loerbroks, State of the Art in Compressing Deep Convolutional Neural Networks, [https://medium.com/@nicolas\\_19145/state-of-the-art-in-compressing-deep-convolutional-neural-networks-cfd8c5404f22](https://medium.com/@nicolas_19145/state-of-the-art-in-compressing-deep-convolutional-neural-networks-cfd8c5404f22)
- [10] Matlab, <https://www.mathworks.com/help/matlab/>