

Rezumat

Lucrarea de față reprezintă un studiu al problemei recunoașterii optice de caractere, utilizând rețele neuronale feedforward și oferă o soluție în implementarea acestui model, urmând structura perceptronului multi-strat.

Rețeaua implementează, pentru antrenare și detecție a caracterelor o rețea neuronală de tip perceptron multi-strat datorită capabilității sale de a generaliza și de a răspunde la tipare necunoscute/impredictibile, caracterizate prin nonlinear-separabilitate.

Rețeaua este antrenată cu algoritmul backpropagation și este configurabilă, funcționalități ca numărul de straturi de neuroni acunși, parametri de antrenare cât și neuronul de deplasare (bias) fiind disponibile utilizatorului în construirea structurii rețelei. De asemenea, diferite funcții precum Sigmoidă Bipolară, Sigmoidă Unipolară sau funcția Identitate pot fi folosite în activarea neuronilor rețelei. Programul poate fi extins în scopul de a oferi posibilitatea construirii de modele diferite ce pot fi folosite cu succes și în alte aproximări de funcții, predicții, clasificări sau recunoaștere de tipare.

În aprofundarea problemei recunoașterii optice de caractere, proiectul implementează un modul de detecție a caracterelor, abordând dificultăți ca:

- extragerea caracterelor/entităților definite de caractere dintr-o imagine;
- transformarea acestor caractere în date de intrare specifice pentru rețeaua neuronală identificând cele mai importante trăsături definitorii pentru simbolurile de detectat;
- antrenarea rețelei cu datele extrase;
- recunoașterea simbolurilor și a entităților din document și construirea documentului aferent imaginii.

Pentru segmentarea imaginii de detectat, sunt folosiți diferiți algoritmi personalizați, bazați pe histograme, cu parametri pre-ajustați în mod empiric. Algoritmii bazați pe histograme au fost aleși datorită simplității în procesare cât și a rapidității efectuării deciziilor de segmentare.

Extragerea atributelor sau descriptorilor de imagine în definirea trăsăturilor pe baza căreia se face clasificarea, se face cu *Transformata cosinus discretă* (DCT), un algoritm de compresie a imaginilor, ce permite minimizarea intrărilor în clasificator cât și un grad ridicat de generalizare.

Modulul de recunoaștere a caracterelor este expus printr-o interfață grafică, utilizând o rețea neuronală pre-antrenată, iar modulul de antrenare al rețelei neuronale este disponibil prin intermediul unui API.

Ca limbaj de programare, proiectul utilizează limbajul Java, iar ca interfață, se folosește JavaFX, o platformă de programare bazată pe java, ce permite dezvoltarea aplicațiilor grafice. Având în vedere limbajul folosit, platforma pe care rulează aplicația este Mașina Virtuală Java fapt care conferă aplicației un grad de portabilitate ridicat.