Sprint4 Discurs

**Arhitectura**

Arhitectura sistemului a ramas aceeasi si anume OCR + Keras si Tensorflow.

A. Achizitie de imagini

Este cea mai simpla parte, se scaneaza o imagine si se incarca in script ca data de intrare pentru modelul antrenat. Imaginile vor fi negru pe alb.

B. Preprocesarea

Pentru acest pas nu am facut nimic pe partea de antrenare deoarece seturile de date sunt deja curatate, insa pentru propriile poze a trebuit sa le convertim la grayscale, am aplicam un blur Gaussian (to reduce noise) si le-am binarizat.

C. Segmentare

Sistemul nostru a utilizat metode precum detectarea marginilor, contururi și filtrarea conturului pentru a segmenta caracterele dintr-o imagine.

D. Extragerea caracteristicilor

Textul selectat este asociat cu baza de date preîncărcată în sistem și în setul de date, cea mai mare corelație este selectată și declarată ca un caracter. Odată ce caracterul este recunoscut pe baza clasificării, acesta este convertit în text.

**Alte abordari**

Data trecuta am mentionat Long Short Term Memory care este o arhitectura de rețele neuronale recurente. In acest timp am incercat sa-l implementam insa am facut-o la scala mica si am reusit sa obtinem rezultate doar pe numere.

Rezultatul este un scor RSME de 0.26 pentru 3 layere, cu 30 de epochs si 100 de bachuri.

Pentru rezultate state of the art care implica LSTM am gasit un experiment care foloseste curbe Bezier si care foloseste baza de date IAM On-Line pentru a antrena modelul.

Rata de eroare a caracterelor este:

**Tabel rezultate performanta**

Revenind la setul de date MNIST, pentru metoda implementata de noi, rezultate la antrenament sunt de 96% acuratete si cu o eroare de caractere de 20-30%

Rezultate state of the art arata o precizie ridicată de 99,30%, care a fost a fost obtinuta cu o rețea neuronală convoluțională (CNN) antrenată cu date distorsionate. Simard et al. a îmbunătățit atât generarea distorsionată a eșantionului, cât și implementarea CNN și a dus la o precizie a testului de 99,60 În loc de extractorii de caracteristici antrenabili din CNN, extragând caracteristici discriminatorii euristic poate duce la precizii ridicate. Fără antrenament cu probe distorsionate, alti 2 cercetatori au obținut o precizie de 99,57% prin extragerea structurilor locale de caracteristici și clasificari folosind SVM-uri liniare **triowise**.

Intr-un alt articol s-au comparat diverse clasificări statistice și neuronale pe caracteristici extrase cu PCA din imagini normalizate. Cu toate acestea, caracteristicile PCA nu funcționează satisfăcător.

Clasificatorul **SVM** cu kernel **RBF** oferă cea mai mare precizie. Cel mai bun clasificator neuronal s-a dovedit a fi cel polinomial (**PC**), care este mult mai puțin complex în stocare și execuție decât SVM-urile. Și rețeaua RBF depășește în mare parte **MLP** atunci când își antrenează în mod discriminator toți parametrii.

Tabelul include rata erorilor unde: „\4-grad" si „\8-grad" inseamna orientare in 4 si 8 directii ale gradientului, de asemenea; \SVC-poly" si \SVC-rbf" indica suport pentru clasificatori vectoriali de tipul unul versus toti cu kernel polinomial si kernel RBF.

In acest tabel, RBF pare ca este inferior in fata MLP pe setul de date MNIST, dar pe alte seturi de date, reteaua RBF domina MLP.

~The idea of PCA(PRINCIPAL COMPONENT ANALYSIS) este o metoda de extragere a caracteristicilor — reducerea numarului de variabile din setul de date, in timp ce pastreaza cat mai multe informatii.

~Svm

~RBF

~PC

~MLP

**Concluzie si Idei** **viitoare**

In concluzie, am implementat cu succes un model de recunoastere a literelor si numerelor discrete cu o rata de acuratete la antrenare de 96% iar pe imagini proprii de 80%.

O posibila solutie la cresterea acuratetii artificial ar fi integrarea unui alt algoritm de machine learning care sa verifice rezultatul solutiei initiale, astfel incat daca cuvantul nu are o acuratete de 100% pe fiecare caracter, acest algortim intra in actiune si verifica daca cuvantul scris de utilizator se gaseste in dictionar sau sa dea o sugestie inapoi care are o acuratete mai mare.