Recunoasterea scrisului de mana

**Recapitulare**

Am antrenat un model de recunoastere a scrisului de mana bazat pe: MNIST 0-9 , Kaggle A-Z folosind: Keras, Tensorflow, Deep Learning

**Ce s-a adaugat de data trecuta**

Am folosit modelul anterior pentru a recunoaste propriile noastre imagini

**Care este arhitectura preliminara si functionalitatile aplicatiei demonstrator**

Proiectul pana acum este format din:

Un modul pentru funcții I / O de ajutor si pentru modelul de recunoastere (implementarea modelului de învățare ResNet)

Un fișier CSV care conține setul de date Kaggle A-Z

Modelul antrenat de recunoastere

Un graf in care afisam rezultatele cele mai recente si aplicatia principala.

De asemenea folosim un pachet numit imutlis care face functii precum translatia, rotația, redimensionarea, scheletizarea mai convenabile precum și afișarea mai ușoară a imaginilor Matplotlib cu OpenCV si Python.

Dupa ce importam tot ce ne trebuie, convertim imaginea in grayscale apoi aplicam un blur gaussian pentru a reduce zgomotul. De aici, detectam marginile imaginii blurate urmand apoi sa localizam conturul caracterelor si initializam lista ce le va tine.

Fiecare contur trece prin 4 pasi:

1. Selectam contururile de dimensiuni rezonabile si le extragem aria de interes (luam informatiile bounding boxului si punem noi delimitari cat de mari sau mici sa fie contururile si le luam pe cele care se potrivesc parametrilor)
2. Curatam imaginea folosind un algoritm de prag (Otsu) pentru a avea o imagine cu caractere albe pe un fundal negru
3. Redimensionarea caracterelor la o imagine de 32x32 pixeli si apoi centralizam conturul
4. Scalam intensitatea pixelilor intre [0,1] si setam o dimensiune pentru gramada apoi impachetam conturul curatat impreuna cu gramada intr-un tuplu dublu si il adaugam la lista de caractere.

Dupa ce trecut prin toate caracterele si am terminat de facut lista, o dam modelului pentru a face predictii, urmand apoi sa ne uitam cu o buclă la fiecare predicție și bounding boxul acesteia si sa luam probabilitatea cea mai mare.

**Teste preliminare**

Pentru poza cu “Hello world” modelul nostru de recunoaștere a funcționat bine, dar a făcut două greșeli.

La a doua imagine avem doar o singură cifră greșită (1 confundat cu 7)

În timp ce modelul nostru de recunoaștere a scrisului de mână a obținut o precizie de 96-97% pe setul nostru de testare, precizia noastră de recunoaștere a scrisului de mână pe propriile noastre imagini personalizate este puțin mai mică decât aceasta.

De exemplu in urmatoarea imagine avem cel mai prost rezultat, chiar daca modelul are o predictie buna, datorita felului in care scriu, uneori vede o singura litera ca pe mai multe.

**Concluzii preliminare**

Avand in vedere rezultatele, problema este mult mai complicata decat ne asteptam mai ales ca modelul nostrum de recunoastere nici nu a luat in condierare cazul in care caracterele pot fi conectate ceea ce inseamna ca mai multe caractere unite vor fi tratate ca unul singur, iar asta rezulta in predictii gresite din partea modelului create.

Tratarea caracterelor conectate este încă o zonă deschisă de cercetare în domeniul computer vision; cu toate acestea, exista modelele care arata a fi promitatoare pentru a imbunatati precizia recunoașterii scrisului de mână cum ar fi : LSTM (Long Short-Term Memory). Insa noi nu vom promite ca putem reusi sa implementam un astfel de model in timpul ramas.

LSTM: Conceptul de bază al LSTM este starea celulei și este o poartă diferită. Starea celulei acționează ca o autostradă de transport care transferă informații relative pe tot parcursul lanțului de secvență. Vă puteți gândi la aceasta ca la „memoria” rețelei. Starea celulei, în teorie, poate transporta informații relevante pe parcursul procesării secvenței. Deci, chiar și informațiile din pașii de timp anteriori pot face ca aceasta să fie pași de timp mai târziu, reducând efectele memoriei pe termen scurt. Pe măsură ce starea celulei își continuă călătoria, informațiile sunt adăugate sau eliminate în starea celulei prin porți. Porțile sunt rețele neuronale diferite care decid ce informații sunt permise în starea celulei. Porțile pot afla ce informații sunt relevante de păstrat sau uitate în timpul antrenamentului.