

Detekcija i prepoznavanje rukopisa sa slike

Milica Jovović, SW-15/2018
Fakultet tehničkih nauka, Novi Sad

Motivacija

Čitanje ručno napisanog teksta predstavlja radnju koju redovno vršimo, kako u profesionalne svrhe, tako i u svakodnevnom životu. Softver za detekciju i prepoznavanje rukopisa sa slike omogućava konvertovanje ručno napisanih karaktera u sebi odgovarajuće kompjuterske karaktere, korišćenjem kompjuterske vizije. Implementacija ovog alata bi znatno doprinela digitalizaciji sveta i olakšanoj obradi teksta.

Skup podataka

Podaci korišćeni za rešavanje opisanog problema su preuzeti sa www.kaggle.com. Dataset se sastoji od skoro 415 hiljada slika ručno napisanih imena, kao i .csv datoteka koje sadrže nazive slika uz odgovarajuće ime sa slike. Od ukupnog broja slika, 80% čine slike nad kojima se model trenira, a preostalih 20% je jednako podeljeno na validacioni i test skup podataka. Zbog slabijih performansi računara na kom je implementiran softver, korišćeno je 15% podataka iz svakog od skupova.

▲ FILENAME	▲ IDENTITY
URL of handwritten image	Transcribed handwritten name
330961 unique values	100540 unique values
TRAIN_00001 . jpg	BALTHAZAR
TRAIN_00002 . jpg	SIMON

ARGITXU

ELISA

Opis problema

Problem prepoznavanja rukopisa sa slike se vrši u nekoliko faza, od kojih su osnovne: dobavljanje slike, pretprocesiranje slike i klasifikacija detektovanih alfanumeričkih karaktera. Slika koju sistem dobavlja i obrađuje je slika ručno napisanog teksta, tj. imena u konkretno opisanom sistemu. Da bi izlaz iz sistema bio korektan, mora se vršiti pretprocesiranje, odnosno priprema slike. Ovaj korak podrazumeva izolovanje područja slike od značaja, izmena dimenzija slike, kao i uklanjanje smetnji i nejasnoća. Učitavanje slika se vrši u „grayscale” modu, što omogućava računaru bolju vidljivost detalja, a nakon toga se menja njihova dimenzija na 256x64. Slike koje sadrže nečitljiva imena nisu uzimane u obzir. Nakon pretprocesiranja, klasifikacija ima za zadatak da prepozna o kom konkretnom karakteru je reč. Kao klasifikatori se najčešće koriste konvolucione neuronske mreže (CNN), što je i implementirano pomoću Python jezika i Keras biblioteke.

Rešenje i rezultati

Algoritam korišćen za implementaciju softvera za detekciju i prepoznavanje rukopisa sa slike je konvoluciona neuronska mreža (CNN), delimično u kombinaciji sa rekurentnom neuronskom mrežom (RNN). Implementirana neuronska mreža se sastoji od ulaznog sloja, čiji ulazi su pikseli slike, i izlaznog sloja, čiji izlaz daje predviđene vrednosti prepoznatih karaktera. Između ulaznog i izlaznog sloja se nalazi nekoliko skrivenih slojeva, kao što su: Conv2D, MaxPooling2D, Dropout, Reshape, Dense, Bidirectional i LSTM. U konvolucionim slojevima se kao aktivaciona funkcija primenjuje ReLU, a na samom izlazu se primenjuje Softmax. Kao funkcija greške koristi se Connectionist Temporal Classification (CTC). Metrika koja je upotrebljena za merenje performansi je procenat uspešnosti klasifikovanih slika, tj. „accuracy”, gde je sve preko 85% tačnosti uspešno klasifikovano, a ispod neuspešno. Konkretni rezultati koje implementirani sistem daje su od 74% do 87% uspešnosti. Procenat uspešnosti varira upravo zbog slabijih performansi mašine na kojoj se model trenira, kompajlira i validira, odnosno uticaj na uspešnost ima trenutno opterećenje mašine. Pretpostavka je da bi sistem poboljšano radio na mašinama sa boljom hardverskom podrškom.

Budući rad

Za unapređenje softvera i povećanje njegove tačnosti, u budućnosti bi se mogla koristiti mašina sa boljim performansama. Ovo bi omogućilo upotrebu većeg broja podataka, odnosno bolje treniranje neuronske mreže. Samim tim, sistem bi imao bolje znanje pri prepoznavanju različitih vrsta rukopisa. Još jedan vid unapređenja sistema bi se mogao ostvariti ukoliko bi pre treniranja modela na detekciju kompletnih reči, istrenirali model da detektuje pojedinačne karaktere. Ovim bi se sistem bio spreman da uvidi razlike između konkretnih karaktera i lakše razazna od čega je sačinjena prepoznata reč. Takođe, znatno šira primena alata za detekciju i prepoznavanje rukopisa bi se mogla ostvariti ukoliko bi se model trenirao i na detekciju većeg teksta, odnosno rečenica i pasusa, a ne samo reči. Za korektno prepoznavanje teksta bi se morala vršiti segmentacija linija teksta, reči i karaktera, u fazi pretprocesiranja slike, i korišćenje nekih mehanizama za proveru smislenosti prepoznatog teksta, kao što su rekurentne neuronske mreže (RNN), Naivni Bayes i slično.

