

Detekcija knjiga na slici i detekcija naslova, autora i izdavača knjige

Autori: Milica Đumić i Tatjana Gavrilović

Fakultet tehničkih nauka u Novom Sadu

DEFINICIJA PROBLEMA I MOTIVACIJA

"Onaj ko ne voli čitati, nije pronašao pravu knjigu za sebe" - J. K. Rowling.

Problem koji se rešava je detektovanje knjiga na slici, potom prepoznavanje njenog naslova, pisca i izdavačke kuće.

Knjige su oduvek bile odličan način da se oslobodimo dosade i istrgnemo iz monotonije. Ljubiteljima knjiga bi ovakva aplikacija olakšala vršenje popisa knjiga, proveru brojnog stanja ili skratila vreme pri online kupovini.

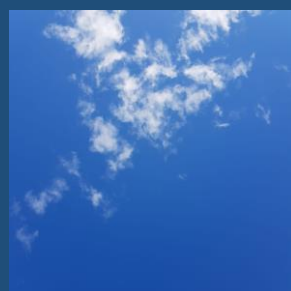
Široku primenu bi softver imao u bibliotekama i izdavačkim kućama.

SKUP PODATAKA

Skup podataka se sastoji iz dve celine, trening i test skupa, kao što je prikazano u tabeli. Svaka slika ima dimenzije 1600x1600 piksela.

Skup	Sa knjigama	Bez knjiga
Trening	210	210
Test	100	50

Primeri:



METODOLOGIJA

DETEKCIJA KNJIGA:

Prvi korak je da se slike pretprocesiraju (svesti na nijanse sive). Nakon toga potrebno je iskoristiti, nad svakom slikom, funkciju `skimage.feature.hog` čiji rezultat je vektor obeležja. Model („`joblib.load("pedestrian.pkl")`“) treniramo nad tim vektorima (funkcija `fit`). Za svaku sliku iz testnog skupa, računamo vektor obeležja (`hog-om`). Procenu da li na slici ima ili nema knjiga dobijamo koristeći funkciju `predict` nad modelom (sa dobijenim vektorom kao parametrom).

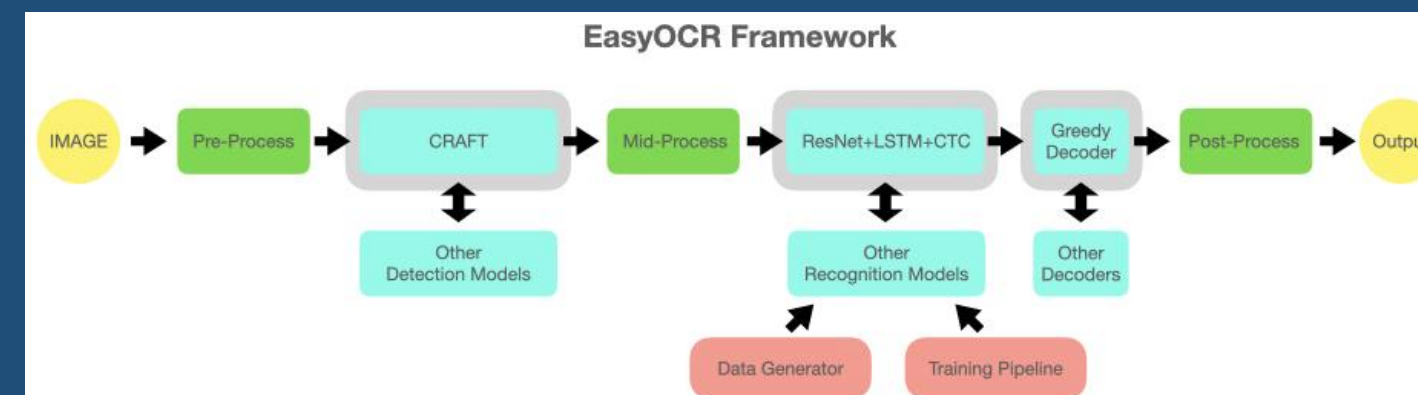
DETEKCIJA REČI:

Za detekciju reči smo koristili biblioteku `easyocr`.

Prvo smo definisali čitača (srpsku latinicu), pozivajući funkciju "`easyocr.Reader(['rs_latin'])`". Ovaj čitač radi u smeru s leva na desno, prepoznavajući između ostalog i specijalne karaktere Š, Č, Ć, Ž i Đ.

Funkciji "`reader.readtext(image)`" se prosledi slika, te kao izlaznu vrednost dobijamo niz od tri elementa:

1. koordinate uočene reči
2. prepoznatu reč
3. procentualnu sigurnost da je to ta reč.



REZULTATI

U svrhu detekcije knjige isprobane su tehnologije: detekcija kontura, *Canny* algoritam, *Hog* deskriptor, *non-maximum suppression*, biblioteka *joblib* i linijska detekcija. Svi pristupi, izuzev biblioteke *joblib*, detektuju na kom mestu se naleze knjige na slici, ali sa velikom greškom. Pristup baziran na upotrebi biblioteke *joblib* detektuje isključivo da li se na slici nalaze ili ne knjige, sa velikom uspešnošću.

Kada je u pitanju detekcija reči, isprobane su dve mogućnosti. Prvi pokušaj je bio klasični *OCR*, dati algoritam je rezultovao veoma lošom detekcijom i regije u kojoj se karakter nalazi i samih karaktera. Drugi pokušaj je korišćenje biblioteke *easyocr* koja daje veoma dobro detektovane i prepoznate karaktere.

ZAKLJUČAK

Za detekciju knjiga je odabrana biblioteka *joblib*. Tačnost dobijena na testnom skupu iznosi 88.67%. Za detekciju reči je odabrana biblioteka *easyocr*. Tačnost dobijena na testnom skupu iznosi 81.2%.

Dalji rad bi uključivao ubacivanje neuronskih mreža (da bi se zaista detektovale knjige) i obradu video snimka.