

Soft computing – predefinisani projekat

Darko Tačić, RA190-2013

Asistent: Miroslav Kondić

Uvod i motivacija

Predefinisani projekat iz predmeta soft computing. Glavni problem zadatka je obrada video snimka pomoću kojeg možemo da saznamo neke informacije. Iako je ovo školski zadatak, on ipak pokrtiva osnovne aspetke detekcije i obrade slika i videa.

Pristup problemu i opis problema

Postoji više načina na osnovu kojeg se pristupa rešavanju ovakvih problema. Potrebno je strpljenje, sistematičnost i želja. Prilikom obrade videa, prvenstveno treba razmišljati o trenutnom stanju problema. Da li su podaci dobri? Da li se mogu izvršiti neke operacije radi dobijanja nekih odgovora? Da li je neko već rešavao sličan problem? Svaki video snimak sadrži jednu pokretnu liniju koja je iste boje. Bilo je potrebno realizovati detekciju te linije pomoću Hough-ove transformacije. Takođe se na snimku nalaze brojevi koji se nasumično kreću na ekranu.

Metod i algoritam za rešavanje problema

Projekat je relizovan u python programkom jeziku i mnogim pomoćnim bibliotekama sa kojima on dolazi. Neke od poznatih biblioteka su OpenCV,Numpy,MNIST...Prvi zadatak je bio realizovati detekciju pokretne linije. S obzirom da je video niz pokretnih slika(framova) koji se smenjuju u realnom vremenu, potrebno je učitati video i posebno posmatrati stanje u svakom frame-u. Na početku slika se konvertuje u sivu boju, radi uprošćavanja daljoj manipulaciji slike. Nakon toga primenjen je globali trashold radi odvajanja bitnog od nebitnog. S obzirom da je linija tamno-plave boje, za detekciju iste je korišćen drugačiji trashold nego za detekciju brojeva. Dalje primenjujemo kombinaciju morfoloških operacija(Dilacija i erozija) radi uklanjanja nepotrebnih šumova. Traženi brojevi su podaci iz poznatog MNIST dataseta. Ovaj dataset obuhvata oko 70000 ručno pisanih brojeva dimenzija 28x28 piksela. S obzirom da su datu brojevi proizvoljne veličine, potrebno je izdvojiti regione, ukloniti šumove i još dodatno izvršiti transfomacije nad testnim podacima. Radi poboljšanja tačnosti svaki region biva konvertovan u sličicu 28x28. Takođe bilo koji region koji je manji od 8 piksela se ignoriše, jer najverovatnije ne predstavlja cifru iz datog skupa. Nakon detekcije broja i linija, sledi razmatranje uslova prelaska broja preko linije. Zbog sporog izvršavanja zadatka ova logika je izostavljena, ali bi se mogla implementirati na osnovu dimenzija regiona u datom frame-u.

Rezultati

Nakon uspešnog pokretanja koda, detekcija linije,regiona i brojeva daje rezultate u realnom vremenu. Međutim zbog sporog izdršavanja frameova svakog videa, nije istetirana potpuna funkcionalnost. Postoji mogućnost proširenja logike radi postizanja potpune funkcionalnosti zadatka.

Zaključak

Ovaj zadatak, iako deluje bez svrhe i kao da nema neku težinu. Ipak on otvara vrata ka mnogim drugim problemima i može da posluži kao odskočna daska ka rešavanju ozbiljnih problema u praksi. Pokazano je kako se jednostavnom programskom logikom protumačiti slika ili video na veoma zanimljiv način. Primer problema gde bi detekcija linija bila korisna je u fudbali prilikom detekcije ofsajd pozicije igrača. Takođe primenljivo je prilikom provere servisa u tenisu(Da li se loptila nalazi ispred ili iza dozvoljene linije). Što se detekcije brojeva tiče, ova operacija bi se mogla koristiti u bilo kakvom softveru koji ima ulogu u interakciji sa klijentom. Veštačka ineligencija koja bi sama mogla da tumači tekst napisan od strane korisnika i da verno daje smisleno mišljenje i odgovor.