

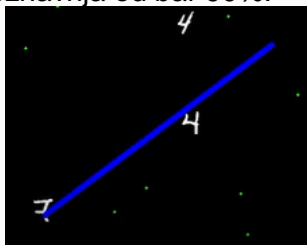
Predefinisani projekat za ocene 7 i 8

Tanja Provči
FTN, Soft kompjuting

Problematika

Svaki video zapis sadrži jednu liniju čija je boja konstanta bez obzira na snimak koji se obrađuje. Cifre se kreću za nasumičan broj koraka, te je potrebno sabrati sve one koje su prošle ispod plave linije ili su je velikim delom zakačile.

Cilj zadatka je postići tačnost prepoznavnja od bar 90%.



Slika 1.1 Cifre na snimku koje je potrebno detektovati

Podaci

MNIST baza podataka sadrži 70 000 ručno pisanih cifara. Ova baza se često koristi za obučavanje i testiranje u mašinskom učenju.

Slika 1.2 Primer ručno pisanih cifara koje se nalaze u MNIST data set-u



Metodologija

Zadatak je rešavan u Python programskom jeziku uz oslonac na pojedine biblioteke kao što su OpenCV, NumPY i Scikit-image.

Za početak je potrebno svaki drugi frejm prebaciti u nijansu sive. Iako se ovakvim konvertovanjem gubi informacija o boji piksela na slici, sama slika postaje mnogo lakša za dalju obradu. Nakon konverzije potrebno je dobiti binarnu sliku korišćenjem određenog praga tolerancije. U ovom zadatku je primenjen globalni trešhold kod kojeg se isti prag primenjuje na sve piksele te se izdvajaju informacije od značaja.

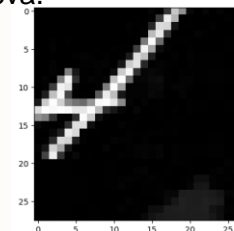
Linija sa video snimka detektovana je pomoću *skeletonize* i *HoughLinesP* funkcija. *Skeletonize* redukuje binarne objekte na širinu od 1 piksela identifikujući i uklanjajući granične piksele, te se takva slika prosleđuje funkciji *HoughLinesP* koja vraća početne i krajnje koordinate detektovane linije.

Nakon toga, potrebno je ponovo primeniti trešhold kako bi se izdvojile cifre od pozadine, kao i implementirati zatvaranje (dilacija + erozija, zatvaranje sitnih otvora među belim pikselima).

Napokon, vrši se određivanje regiona u kojima se nalaze cifre pomoću funkcije koja vraća listu regiona sa njihovim osobinama-*regionprops* (prosleđuje joj se slika sa obeleženim regionima dobijena funkcijom *label*).

Ukoliko se za cifre utvrdi da se seku sa pravom, u tom slučaju ih je potrebno proslediti *KNeighborsClassifier* algoritmu koji vrši prepoznavanje cifre te ju je kao takvu potrebno dodati na ukupnu sumu brojeva.

Slika 2.1 Detektovanja cifra



Rezultati i analiza

Tačnost i poklapanje rezultata rešenja dobijenog prethodno opisanom implementacijom je 96,64%.

Ova preciznost je određena na osnovu predefinisano testu. Metodologija je primenjena na 10 video snimaka.

Kod pojedinih snimaka tačnost prepoznavanja cifara i njihovo sumiranje su 100%, dok kod drugih dolazi do blagih odstupanja.

Naime, ukoliko jedna cifra u potpunosti preklapa neku drugu u trenutku preseka sa pravom, program nije u stanju da detektuje zaklonjenu cifru.

Takođe, u pojedinim slučajevima jedna te ista cifra biva detektovana kao dve različite jer u dva susedna frejma dolazi do poklapanja sa linijom te algoritam to prepoznaje kao istu cifru.