



Detekcija cifara i linija koristeći Hough transformaciju

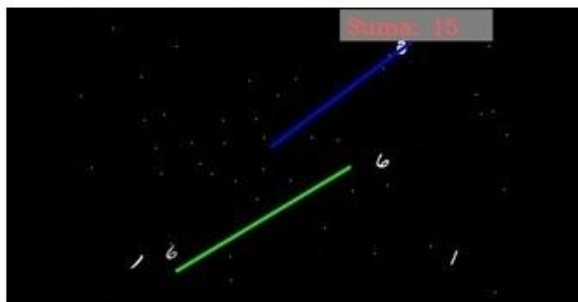


FTN-AI-LAB

Problem

Prepoznavanje objekata na digitalnim formatima se u poslednje vreme izdvojilo kao značajna oblast obrade slike koristeći veštačku inteligenciju.

Tema projekta jeste prepoznavanje cifara i linija na video zapisu. Problem obuhvata detektovanje prelaska cifre ispod linije. Linije se kreću, a u podacima postoji određen nivo šuma.



Podaci

Predmet prepoznavanja jesu dve pokretne linije, jedna plave i jedna zelene boje, kao i decimalne ručno pisane cifre bele boje koji se nalaze na 10 različitih video zapisa.

Obučavanje neuronske mreže je izvršeno na data setu predstavljenom slikom veličine 2000x1000 piksela, koji sadrži 5000 ručno pisanih brojeva, po 500 za svaku decimalnu cifru. Pošto je slika podeljena tako da svaki red zauzima 20 piksela, i svaka grupa istih cifara zauzima 5 redova, traženi izlaz neuronske mreže je izračunat kao količnik celobrojnog deljenja y koordinate gornjeg levog piksela segmenta slike sa 100.

Dataset korišćen u projektu se može naći na sledećoj adresi:
<https://www.learnopencv.com/wp-content/uploads/2016/12/digits.png>

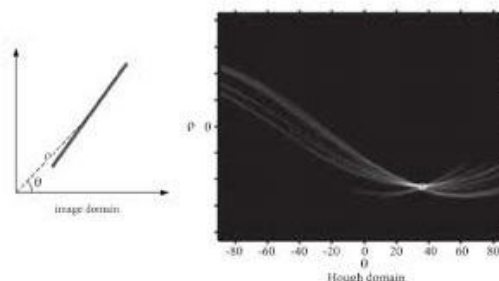
Metodologija

Kao algoritam za detekciju linija, u ovom projektu je korišćena probabilistička Hough transformacija. Za klasifikaciju cifara je korišćena neuronska mreža sa jednim skrivenim slojem, 784 ulaza i 10 izlaza.

Hough transformacija

Hough transformacija predstavlja metod za detekciju pravih linija na rasterskoj slici. Temelji se na pretpostavki da kroz svaki piksel na slici možemo provući beskonačno mnogo pravih linija, po jednu za svaki mogući ugao.

Algoritam uvodi Houghov prostor, koji koristi parametre eksplicitne jednačine prave, $Y = k * X + n$, kao koordinate novog prostora, a tretira promenljive X i Y kao konstante. Transformacija pronalazi pravce u Houghovom prostoru (koji odgovaraju pikselima u prostornom domenu slike) koji se seku u jednoj tački (k,n), što znači da pikseli predstavljeni tim pravcima pripadaju istoj liniji u prostornom domenu slike.



Pošto problematika projekta obuhvata prepoznavanje duži, a ne pravih, korišćen je probabilistički oblik Hough transformacije, koji pronalazi duži koristeći funkciju gustine verovatnoće prilikom određivanja piksela koji pripadaju istoj liniji.

Parametri

Nedostaci podataka jesu promenljiva priroda ručno pisanih objekata, čija se nepravilnost ne može tačno predvideti, kao i stohastičnost probabilističke transformacije, koja ne vraća uvek iste vrednosti kao parametre duži. Promenljiva priroda ručno pisanih podataka je zaobiđena pažljivim korišćenjem erozije i dilatacije crno belih objekata, kojima se postiže pravilnije konturisanje cifara, ali koji takođe dovode do nepravilnijeg oblika nekih drugih cifara (tipičan primer je prevelika dilatacija na ciframa osmice, koja eliminiše unutrašnje praznine cifre i dovodi do pogrešnog prepoznavanja od strane neuronske mreže). Problemi takođe nastaju prelaskom cifara preko linije, kada usled promenljive veličine cifara, manji delovi bivaju prepoznati kao cifre. Radi rešavanja ovog problema, prepoznavanje cifre se vrši u trenutku kada se gornji levi ugao konture nađe na dovoljno maloj udaljenosti ispod tražene linije. Ovo dovodi do nešto kasnijeg prepoznavanja cifre, ali takođe uvodi veću tačnost prilikom prepoznavanja.

Zaključak

Postignuta tačnost od 66% zadovoljava potrebe problema, ali takođe ostavlja dovoljno prostora za poboljšanje pristupa. Korišćenjem metode praćenja objekata dobijaju se bolji rezultati prilikom prepoznavanja objekata koji prelaze jedni preko drugih u trenutku prelaska linije, ali se javlja problem prilikom sudaranja objekata tokom njihovog kretanja i greške koju uvodi algoritam prilikom određivanja koja je kontura u pitanju.

Aspekt projekta koji se takođe može poboljšati jeste prepoznavanje cifara koje prolaze preko ivica linije, gde se određen broj cifara koje prelaze preko gornje ivice ne prepozna usled prevelike udaljenosti.