

Prepoznavanje i praćenje cifara u video zapisu

Aleksandar Jeremić | RA198-2014

Problem

Video zapis poseduje 2 linije ispod kojih prolaze cifre (rukom pisane). Cifre koje prođu ispod plave linije treba sabrati, dok one koje prođu ispod zelene linije treba oduzeti od konačne sume. Jasno, da bi se ovaj zadatak izveo, neophodno je pronaći način da prepoznamo linije, cifre, kao i samo kretanje cifara u snimku, i da ove pojave predstavimo putem promenljivih u programu.

Rešenje

S obzirom da linije na snimku ne menjaju poziciju, i da su uvek u pitanju tačno 2 linije, iz jednog frejma snimka je moguće **prepoznati linije** na njemu putem Hough-ove transformacije.

Što se tiče **prepoznavanja cifara**, situacija je malo komplikovanija. Prvo je potrebno pronaći cifre na frejmu snimka, a potom ih i prepoznati.

U ovom projektu, za **pronalaženje cifara** je korišćena sledeća metoda: frejm snimka je prvo proveden kroz više filtera koji su uklonili šum i već prepoznate linije iz slike, a potom su uklonjene i boje na slici, i uradjena posterizacija slike tako da smo dobili crno-belu sliku bez glatkih prelaza nijansi. Iz ovakve slike je potom moguće izvuci konture objekata na slici (u našem slučaju, to će biti čisto beli objekti – cifre, na crnoj pozadini). Iz pozicije ovako dobijenih kontura dobijamo i pozicije cifara na slici.

Samo **prepoznavanje cifara** je u nadležnosti SVM modela. Ovaj model se prvo obučava priloženim data setom, nakon čega se koristi za prepoznavanje cifara sa malih isečaka tih cifara iz frejma (slike).

Praćenje cifara (dok putuju po snimku) je u osnovi implementirano veoma jednostavno – za svaku prepoznatu cifru beleži putanja kojom se kretala tokom snimka (u listi (zapravo, dictionary-ju) pod imenom 'strings' u kodu; svaki 'string' u ovoj listi je kao nit kojom se kretala cifra, predstavljena nizom tačaka), a cifre povezujemo na njihove niti tako što upoređujemo poziciju cifre u ovom frejmu sa pozicijom cifara zabeleženih u prethodnom frejmu. Ukoliko je udaljenost dovoljno mala, pretpostavljamo da se radi o istoj cifri koja se samo malo pomerila od prethodnih frejma. Očigledno je da ovako primitivan način praćenja cifata nije idelan, te može doći do gubljenja cifara ukoliko one budu

zaklonjene drugim ciframa ili čak linijama ispod kojih treba da prođu. Za ovu svrhu se vodi evidencija i o ciframa čije su putanje prekinute, a postoji i mehanizam koji može da poveže novo-pronađene cifre na njihove ranije prekinute putanje. Sam mehanizam povezivanja na prekinute putanje funkcioniše tako što, pri pronalasku nove cifre (za koju pretpostavlja da je zapravo cifra koja se ranije pojavila, ali kojoj smo izgubili trag), predviđa gde bi se u ovom frejmu nalazile cifre iz liste prekinutih putanja (lista (dictionary) 'izgubljene' u kodu), i onda povezuje novo-pronađenu cifru na putanju čija je predviđena pozicija najbliža poziciji novo-pronađene cifre.

S obzirom da su nam poznate putanje svih cifata kao i tačke koje određuju linije na slici, **provera da li je cifra prošla ispod linije** se svodi na jednostavnu proveru preseka putanje cifre i linije. Ukoliko presek postoji, cifra je prošla ispod linije.

Treba pomenuti da sama implementacija sadrži i niz manjih mehanizama koji su tu da obezbede veću tačnost programa.

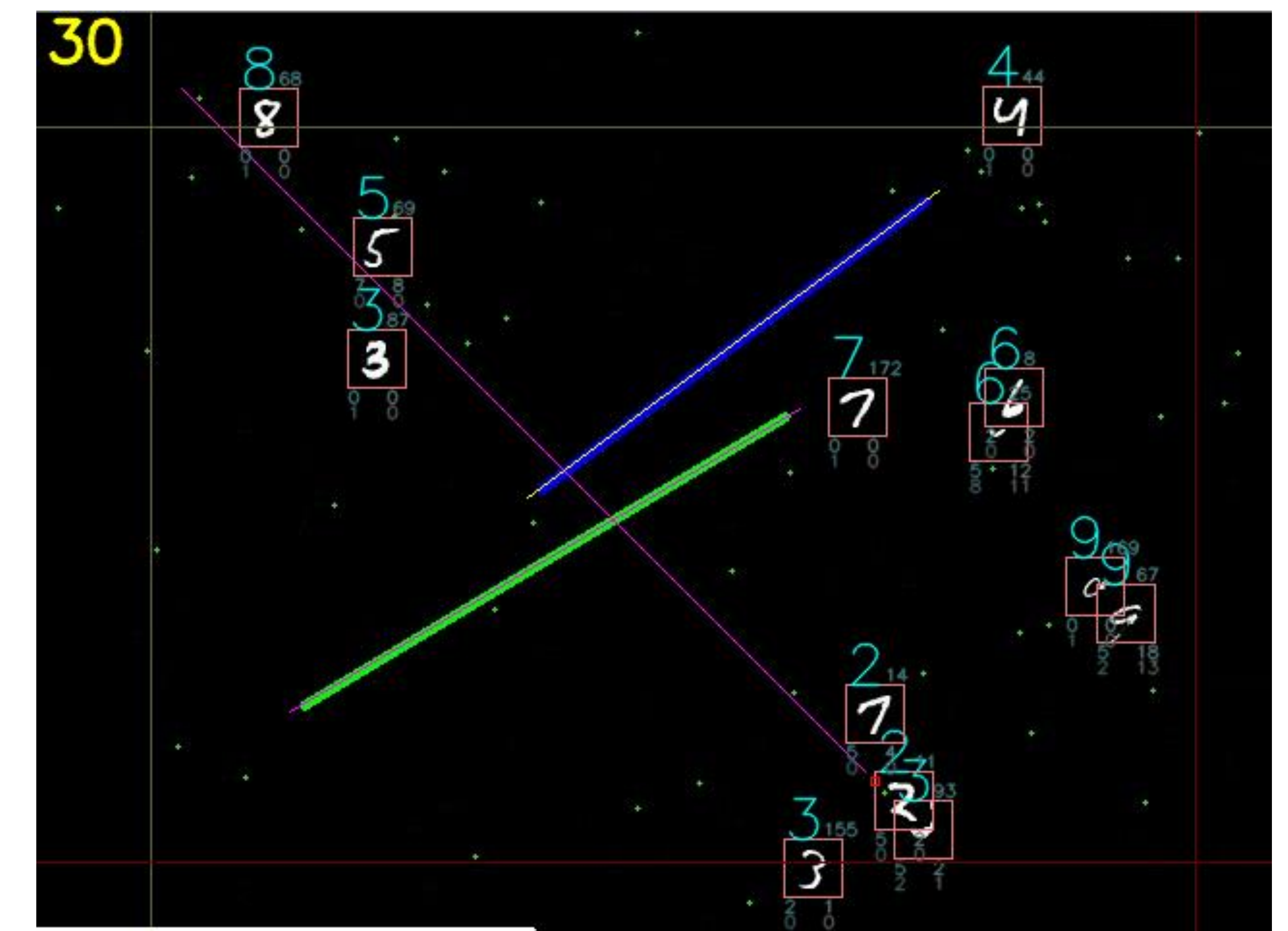
Na primer, predikcije SVM modela o cifri na osnovu isečka iz frejma, rade se više puta u toku samog životnog veka cifre na snimku, i cuvaju se u listi (dict. 'votes') unutar svake putanje cifre (dict. 'strings'). Kada je potrebno dobiti vrednost neke cifre, uzima se najviše puta 'izglasana' vrednost (iz te liste 'votes').

Već sam napomenuo da se predviđanje cifre radi više puta u toku životnog veka cifre na snimku. Doduše predviđanje se ne radi u svakom frejmu snimka. Program će raditi predikciju cifre (i dodati glas u 'votes' listu za tu cifru) samo kada su dimenzije prepoznate cifre (to jest' dimenzije konture cifre) veće od nekih očekivanih vrednosti. Same očekivane vrednosti dimenzija cifre se postavljaju i menjaju u toku životnog veka cifre u snimku. Ovo će sprečiti prepoznavanje cifre u slučajevima kad su njene dimenzije premale, jer je u tom slučaju cifra najverovatnije zaklonjena nekom drugom cifrom (te se ni 'votes' lista neće puniti pogrešnim vrednostima, ili bar ne u tolikoj meri).

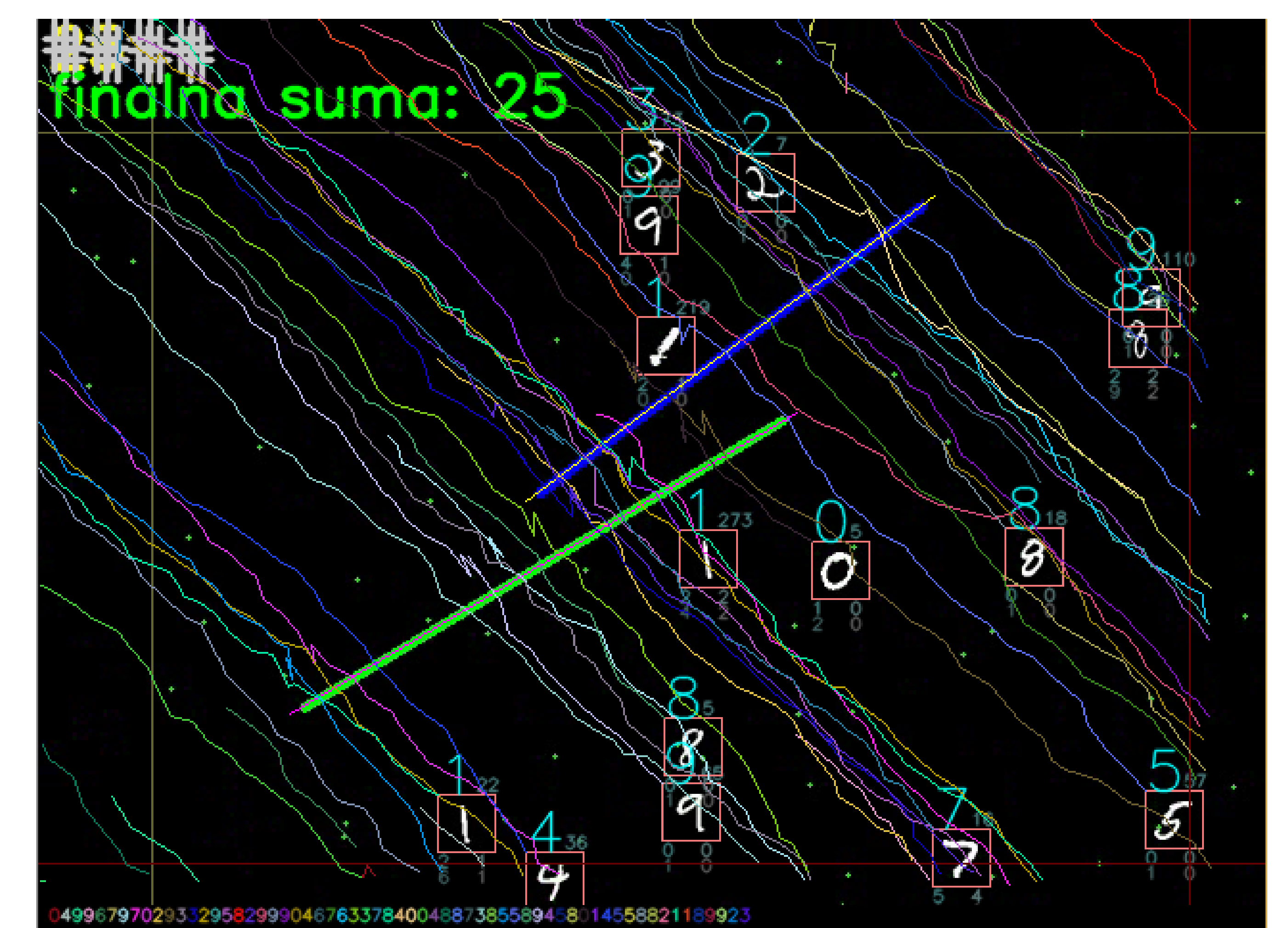
Korišćeni alati

Python v3.7, sa bibilotekama:
numpy-1.14.0+mkl, opencv_contrib_python-3.3.0.10,
scikit_learn-0.19.1, scipy-1.0.0

Program u radu



- Pronađena cifra se povezuje na prekinutu nit. Ljubičasta linija prestavlja aproksimaciju kretanja ove cifre, dobijenu linearnom regresijom nad tačkama koje predstavljaju putanju izgubljene cifre. Crvenu kvadratic je pretpostavka gde bi se ta cifra, na toj putanji, nalazila u ovom frejmu.



- Po završetku, program prikazuje putanje svih cifara koje je zabeležio u toku svog rada.