

Klasifikacija i operacije sa pokretnim ciframa

Vukašin Jović RA131/2015

Asistent: Ivan Perić

Uvod

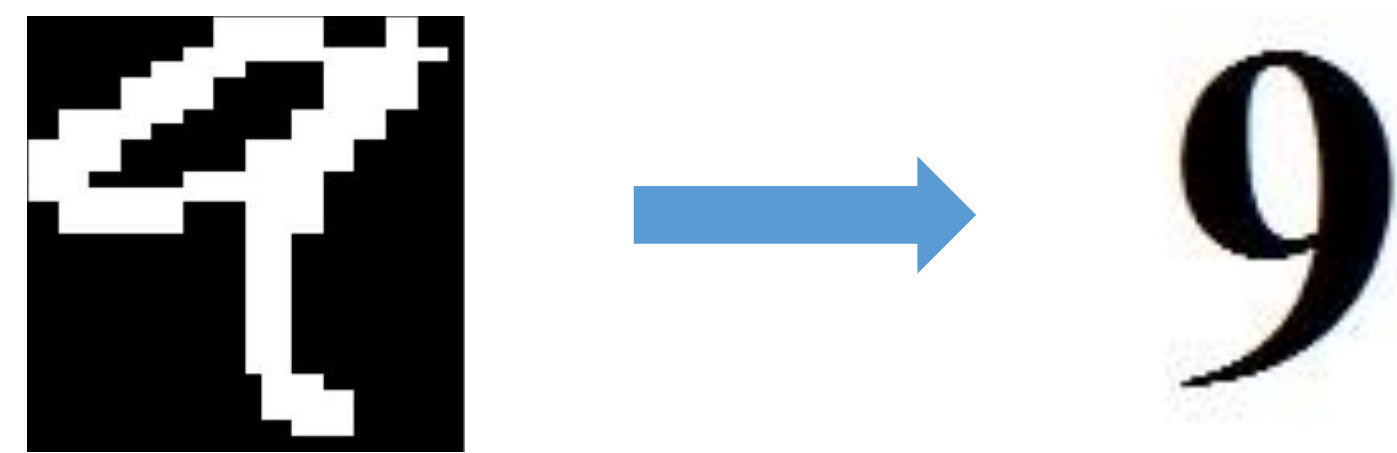
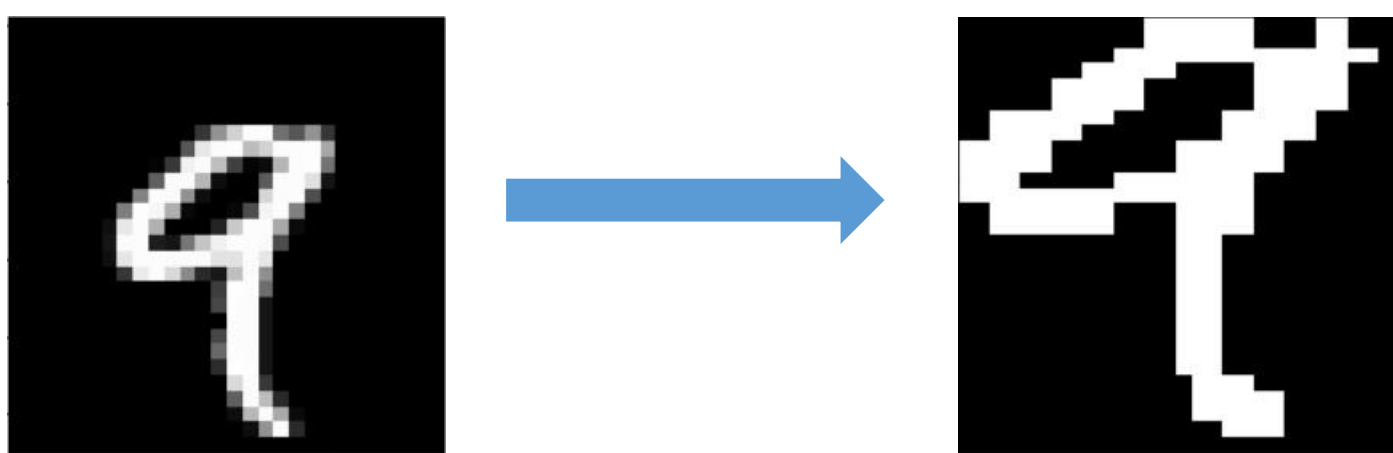
Cilj zadatka jeste detektovati položaj dve iscrtane linije na svakom videu nakon čega je potrebno locirati konture ručno napisanih brojeva i klasifikovati ih. Brojevi se kreću po videu i prelaze ili ne prelaze preko linije. Oni brojevi koji pređu prvu(plavu) liniju treba da se saberu na ukupnu sumu, dok oni koji pređu preko druge(zelene) linije treba da se oduzmu od ukupne sume.

Za rešavanje ovog problema korišćeni su:

1. Hough transformacija iz OpenCV – putem koje se lociraju linije
2. findContours za prepoznavanje kontura
3. Neuronska mreža za obučavanje i prepoznavanje brojeva

Priprema podataka za treniranje

Za treniranje neuronske mreže iskorišćen je MNIST dataset koji se sastoji iz 60000 slika za treniranje. S obzirom da brojevi na slikama nemaju isto rastojanje od svih ivica vršilo se otklanjanje tih praznina do ivica.

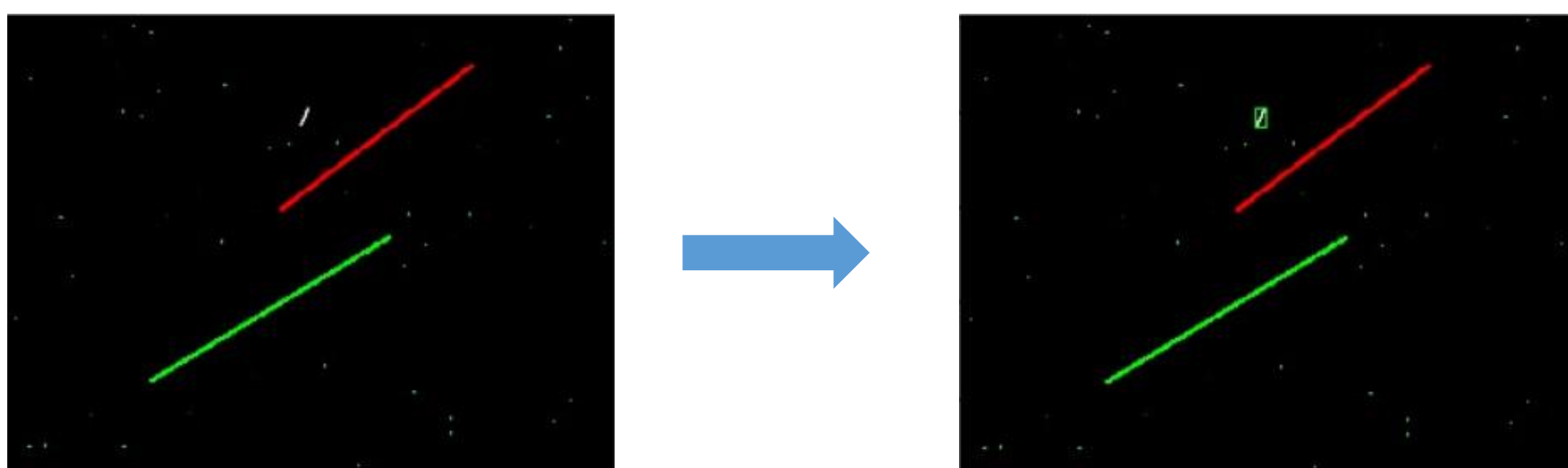


Detekcija linija Hough transformacijom

S obzirom da su linije fiksne na svakom videu, dovoljno je naći njihov položaj na prvom frejmu videa. Najpre se korisni CannyEdge algoritam za prepoznavanje ivica, nakon čega se primenjuje Hough transformacija sa parametrima kao što su minimalna dužina linije i maksimalni razmak između dve linije da bi se gledale kao jedna. Putem dobijenih koordinata linija računa se jednačina prave za obe linije.

Detekcija kontura cifara i priprema za klasifikaciju

Detekcija konture se vršila tako što se pronađu sve konture na fotografiji – frejmu putem cv2.findContours(). Nakon toga se kontura oiviči pravougaonikom te se proveravaju dimenzije pravougaonika kako bi se utvrdilo da li je uhvaćen objekat od interesa ili ne. Kao rezultat dobijamo niz fotografija na kojima se nalazi po jedan pronađen broj koji je spreman za klasifikaciju. Sledi slika pronađene konture.



Trenutak klasifikacije

U svakom frejmu se gleda donji desni ćošak pravougaonika kojim je broj uokviren. S obzirom da se brojevi kreću po dijagonali videa ukoliko donji desni ćošak dodirne liniju broj će sigurno proći kroz liniju. Proverava se da li njegove koordinate pripadaju jednačini prave linije s razlikom od 2 i čim upadne u taj opseg, detektovana kontura se priprema za klasifikaciju i prosleđuje neuronskoj mreži radi predikcije.

Sprečavanje ponavljanja

Kako broj u jednom frejmu može a ne mora da se pomeri za određen korak, potrebno je utvrditi da li je broj koji je na liniji već detektovan i računat. To se radilo tako što se koordinate donjeg desnog ćoška broja koji se dodirne s linijom pamte te se proverava da li su koordinate trenutno obrađivanog ćoška jako blizu nekim ranije prepoznatim koordinatama. Da bi izbegli situaciju gde dva različita broja u potpuno dva različita frejma dodirnu liniju u (približno) istoj tački, lista ranijih koordinata se anulira na svakih 15 frejmova.

Neuronska mreža

Neuronska mreža se sastoji iz tri sloja. Ulazni i međusloj imaju sigmoid, dok izlazni sloj ima softmax aktivacionu funkciju. Izlazni sloj se sastoji od 10 neurona jer postoji 10 klasa (brojevi od 0 do 9). Od 10 neurona prilikom jedne klasifikacije aktivira se jedan koji označava određenu klasu. Što znači da ako se pri klasifikaciji aktivira peti neuron, neuronska mreža je prepoznala broj 5.

Zaključak

Obrada svakog videa se završava uspešno. Tačnost obrade zavisi od kompleksnosti videa. Ukoliko je pojava preklapajućih cifara veoma česta, konture sakrivenih brojeva ne budu pronađene te se rezultat u odnosu na njih ne menja. Problem koji rezultuje lošijom tačnošću takođe nastaje ukoliko se pogrešno prepozna cifra veće vrednosti. Na taj način ukoliko neuronska mreža npr. umesto broja 9 klasifikuje broj 0, nakon sabiranja ili oduzimanja dolazi do velikog odstupanja u tačnosti.