

Predefinisani projekat Soft kompjuting 2018

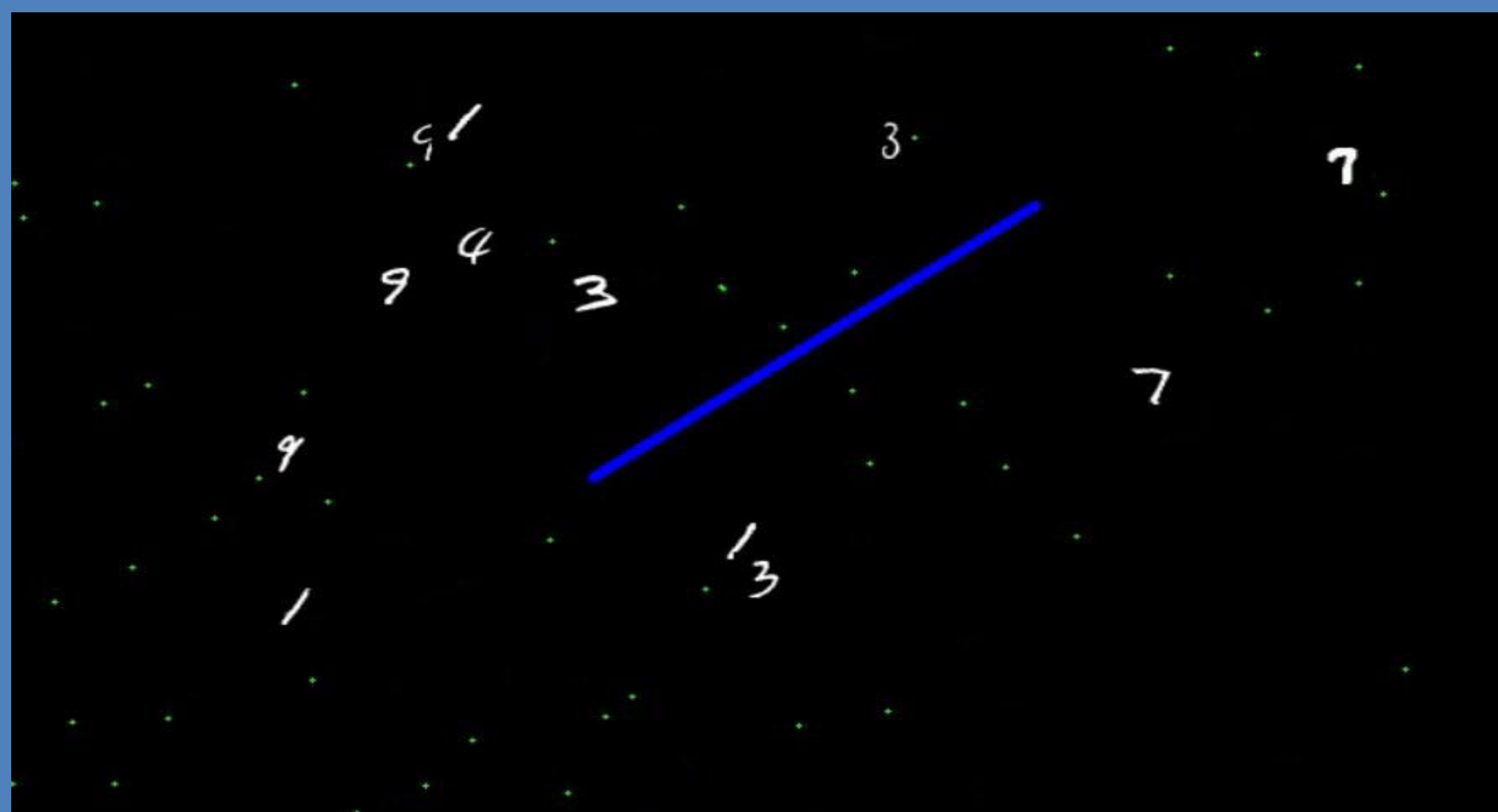
Profesor: Jelena Slivka

Asistent: Stefan Anđelić

Student: Uroš Jakovljević RA 176/2015

Opis problema

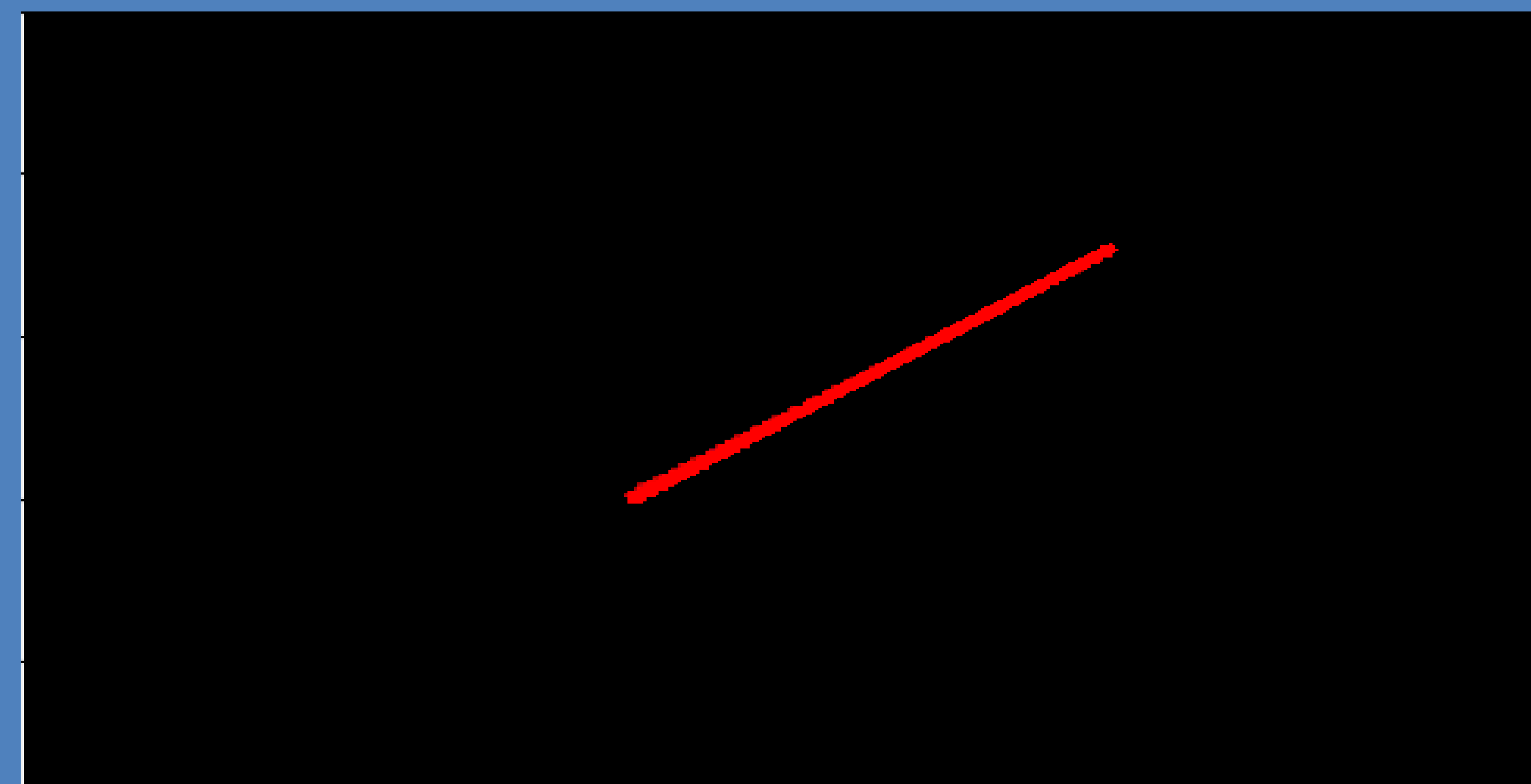
Video zapis poseduje jednu liniju plave boje iza koje prolaze ručno napisane cifre (cifre iz MNIST skupa podataka). Cifre se u video zapisu kreću za slučajni broj koraka i bele su boje. Takođe, čitav video zapis poseduje šum u različitim intenzitetima različite boje, koji otežava izdvajanje i prepoznavanje cifara (slika 1). Potrebno je prepoznati i sabrati sve cifre koje prođu ispod plave linije sa tačnošću prepoznavanja preko 90%. Pokušati rešiti slučaj preklapajućih cifara koje istovremeno prelaze preko linije.



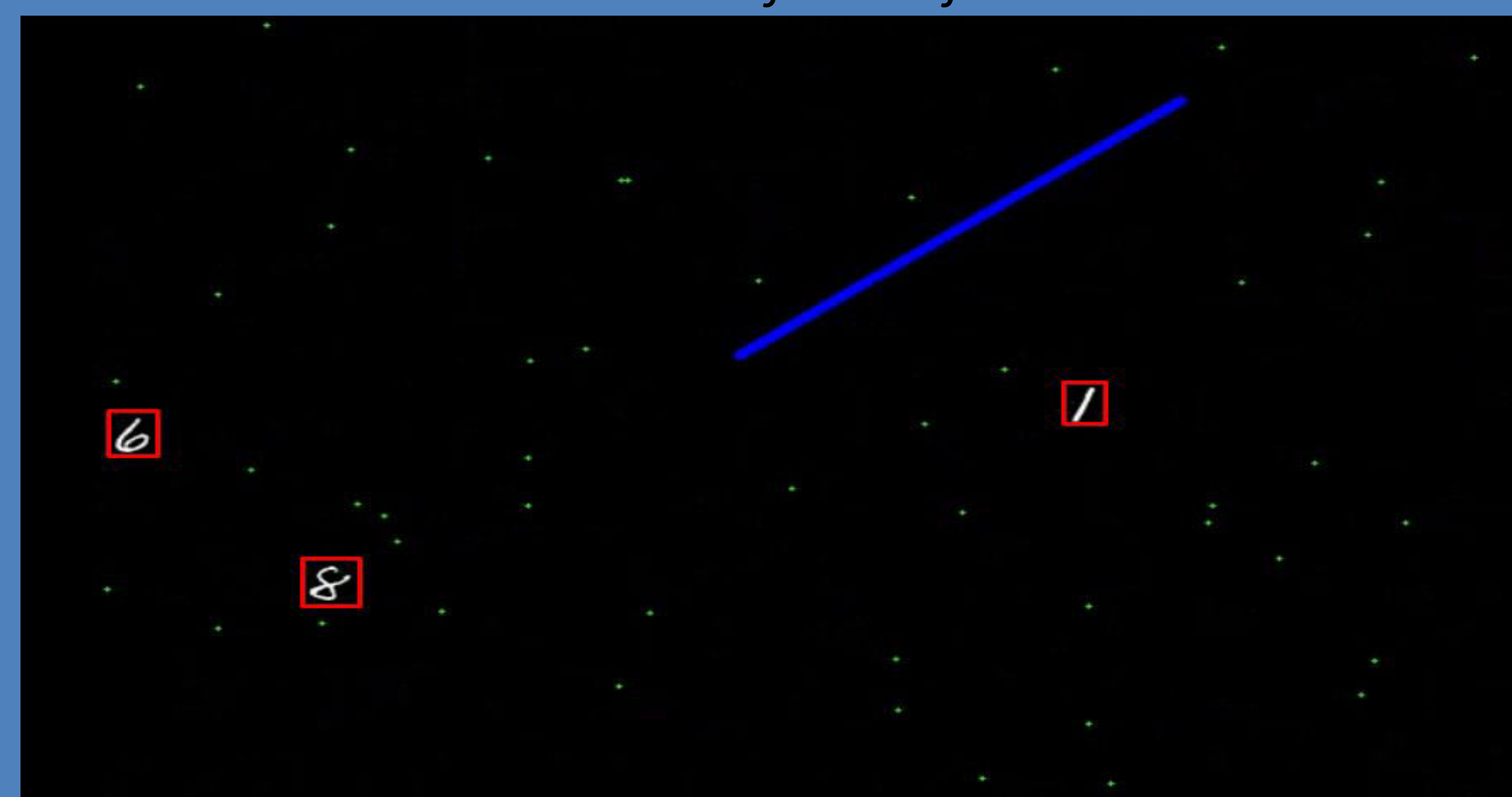
Slika 1 – početni problem

Postupak rešavanja problema

1. **Detekcija linije** – za detekciju linije se koristi *Hough* transformacija (probabilistička). Prilikom učitavanja video zapisa, uzima se prvi frejm iz video zapisa i on se filtrira na plavu boju, a zatim se prebacuje u crno belu sliku (*gray*). Nad tako filtriranom slikom se primenjuje *Canny Edge Detection* algoritam za pronalaženje ivica linije na slici. Nad tom slikom ivica se zatim primeni *Hough* transformacija za pronalaženje linija na slici (slika 2). Iz skupa pronađenih linija se izdvajaju dve najudaljenije tačke koje predstavljaju temena linije.



Slika 2 – detektovana linija pomoću Hough transformacije



Slika 3 – selektovani regioni od interesa u kojima se nalaze cifre

2. **Prepoznavanje cifara** – kada se pronađe linija, dalje se video obrađuje frejm po frejm radi pronalaska i prepoznavanja cifara. Frejm se prvo filtrira na belu boju, radi uklanjanja šuma (cifre su bele boje, dok šum nije), a zatim se vrši transformacija frejma u binarnu sliku. Takva slika se prosledi algoritmu za traženje regiona od interesa (delovi slike u kojima se nalaze cifre) koji vraća niz regiona skaliranih na veličinu 28x28 u kojima se nalaze slike cifara (slika 3). Za svaki region/cifru se prvo utvrdi da li je prvi put prešla preko linije (korak 3) i ukoliko jeste, prosleđuje se K-nearest Neighbours klasifikatoru koji je prepoznaje (klasifikuje) kao jednu od 10 cifara. Klasifikator je prethodno obučen na 60 hiljada primeraka slika ručno napisanih cifara iz MNIST skupa podataka.
3. **Praćenje da li je cifra prešla preko linije** – na osnovu prethodno pronađenih koordinata linije se zaključuje da li je cifra prešla preko nje. Za svaku cifru se pamte koordinate regiona koji je predstavlja i podatak o tome da li je prošla liniju. Pri obradi svakog pojedinačnog frejma, pronađena cifra se poredi sa ciframa iz prethodnog frejma i zaključuje se da li je u pitanju nova cifra ili već postojeća koja se pomerila i da li je prešla preko linije.

Zaključak i analiza

Ovakvim postupkom detekcije i prepoznavanja cifara, ostvari se tačnost prepoznavanja cifara preko 90%. Korišćenjem drugog klasifikatora kao što je na primer konvoluciona neuronska mreža, može se postići veća preciznost ali uz cenu dužeg obučavanja klasifikatora. Bolja preciznost bi se mogla postići i korišćenjem dodatnih metoda za eliminaciju šuma i sofisticiranijim postupkom izdvajanja i prepoznavanja regiona od interesa. Takođe, problem preklapajućih cifara koje prelaze istovremeno preko linije nije u potpunosti rešen i njegovim rešavanjem bi se dosta povećala preciznost rezultata.