

Klasifikacija saobraćajnih znakova pomoću konvolucionih neuronskih mreža

Darko Tica

Uvod

Problem klasifikacije saobraćajnih znakova na osnovu slike predstavlja jedan od bitnih problema za rešavanje u cilju postizanja autonomne vožnje. Saobraćajni znaci signaliziraju opasnost, naredbe, upozorenja i obaveštenja vozačima i predstavljaju glavni način njihovog informisanja o pravilima bezbednog ponašanja na putu. Ta pravila moraju biti poštovana od strane svih učesnika u saobraćaju, te je za agenta koji samostalno vozi potrebno da iz spoljašnje sredine prikupi i istumači informacije sa takvih znakova. To je moguće postići detekcijom saobraćajnih znakova u agentovom okruženju, a zatim i njihovom klasifikacijom, odnosno identifikacijom. U ovom radu je istražen drugi deo problema, to jest klasifikacija znakova na osnovu prosleđene slike, pri čemu je problem klasifikacije rešen pomoću konvolucionih neuronskih mreža.

Podaci za treniranje

Kao skup za treniranje korišćen je **GTSRB (German Traffic Sign Recognition Benchmark) dataset**, koji se sastoji od oko 40000 slika koje su podeljene u 43 klase, pri čemu svaka klasa predstavlja po jedan saobraćajni znak. Radi pojednostavljivanja projekta i procesa treniranja, broj klasa je smanjen na 12, pri čemu te klase obuhvataju 16,470 slika. Odabrani tipovi znakova su prikazani na sledećoj slici:



okviru *dataset*-a. Ovakva izmena veličine je nužna stoga što konvolutivna neuronska mreža kao ulaz očekuje konstantan broj parametara (ulaznih piksela u ovom slučaju).

Arhitektura neuronske mreže

Kao što je već napomenuto, kao rešenje ovog problema korišćena je konvolutivna neuronska mreža. Kao uzor je korišćena *LeCun* neuronska mreža. U ovom primeru, ona se sastoji od ulaznog sloja koji je vidu trodimenzionalne matrice dimenzija 55x55x3: prve dve dimenzije predstavljaju dimenzije ulazne slike, dok treća dimenzija predstavlja standardan broj kanala za neuronske mreže koje obrađuju slike (po jedan kanal za crvenu [R], zelenu [G] i plavu [B] boju - RGB). U unutrašnjem delu mreže se naizmenično smenjuju konvolucioni (filteri veličine 3x3) i *pooling* (*pool* veličine 2x2) slojevi, kao što se vidi na slici ispod. Pored toga, na izlazu iz konvolucionih slojeva, kao i na pretposlednjem, potpuno povezanom sloju korišćena je ReLU funkcija aktivacije, dok je na izlaznom sloju korišćena *Softmax* funkcija, s obzirom da se radi o klasifikacionom problemu. Takođe, u toku treniranja korišćena je *droupout* tehnika. Ona je primenjena nakon drugog i trećeg konvolucionog sloja, i glavni cilj njene primene je da se smanji mrežni *overfitting*.

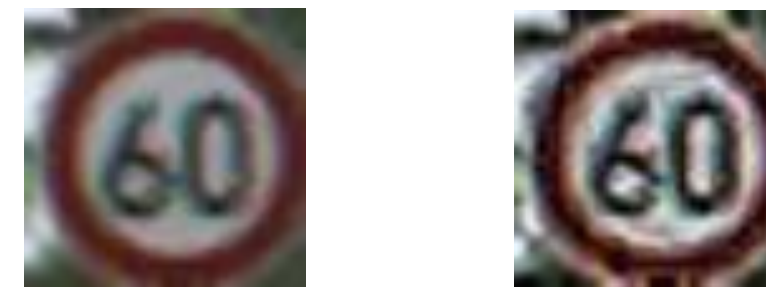
Kao funkcija greške za mrežu je korišćena *categorical cross-entropy* funkcija, pri čemu je kao optimizaciona tehnika odabrana *Adam* funkcija. Kako je u pitanju neuronska mreža čiji je cilj klasifikacija znakova, pri čemu jedan znak može pripadati samo jednoj klasi, izbor ove funkcije kao funkcije greške predstavlja očigledan izbor.

Dobijeni *dataset* sastavljen od pomenutih 12 klasa znakova je dalje podeljen na trenažni skup podataka (11863 slika), validacioni skup podataka (2960 slika) i testni skup podataka (1647).

Slike znakova koji se nalaze u *dataset*-u su različitih dimenzija, od 15x15 pa do 250x250 piksela. Pored toga, svaka slika je labelirana sa koordinatama pravougaonika (*bounding box*) koji obuhvata granice samog znaka u okviru slike. Usled toga, za svaku sliku je urađeno isecanje (*crop*), a nakon toga i promena veličine na dimenzije 55x55 piksela, što predstavlja prosečnu veličinu slike u

Pretprocesiranje

Kao i u većini drugih slučajeva neuronskih mreža čiji je ulaz slika, i kod ovog problema se performanse mogu povećati pretprocesiranjem ulaznih slika. U ovom primeru, korišćena je metoda **CLAHE** (*contrast limited adaptive histogram equalization*). Glavna prednost ovog metoda jeste u tome što se povećava kontrast slike, ali se takođe naglašavaju ivice objekata na njoj, čime se mreži daje precizniji prikaz nekog znaka, a samim time i povećava šansa za uspešnije treniranje i obradu.



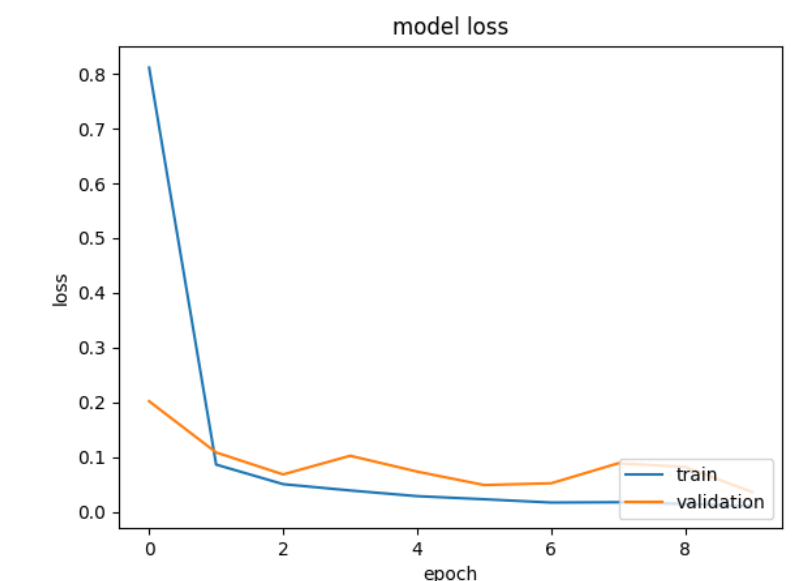
Slika: znak pre i posle primene CLAHE metode

Rezultati

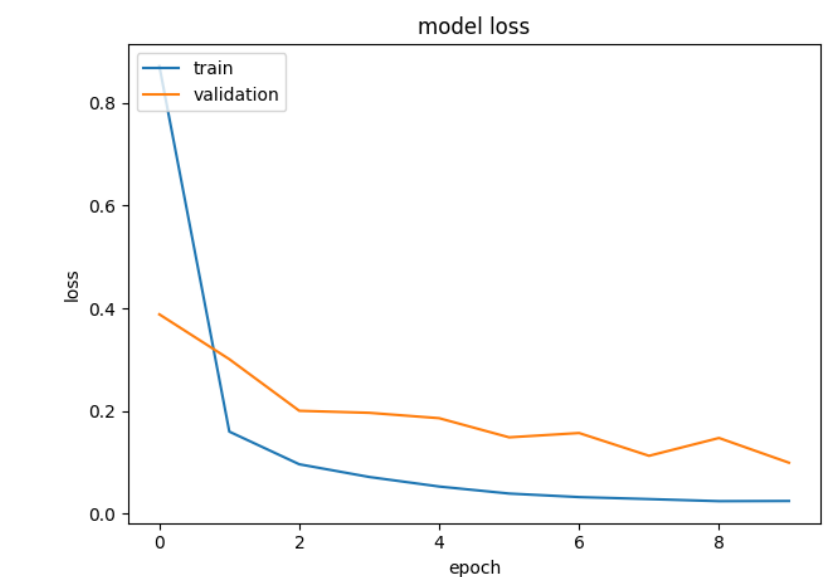
Kao glavna mera za obradu rezultata odnosno mere performansi mreže koristi se *accuracy*, odnosno procenat tačno klasifikovanih znakova. Samo testiranje je vršeno na već pomenutom testnom skupu, koji je izdvojen pre treniranja, i čiji podaci nisu bili korišćenu u trenažnom procesu.

Performanse, odnosno preciznost nad datim testnim skupom, uz korišćenje pretprocesiranja u vidu CLAHE metode jeste **99.5%**. Sa druge strane, preciznost mreže koja je bila trenirana bez pretprocesiranja bila je **96.3%**, što predstavlja napredak i opravdava korišćenje pomenute pretprocesne metode. Pored veće krajnje preciznosti, takođe se može primetiti i brža konvergencija funkcije greške po epohama, što takođe govori o bržem procetu treniranja funkcije sa korišćenjem CLAHE metode u

odnosu na metod koji nije koristio nikakvo pretprocesiranje.



Slika: greška sa CLAHE pretprocesiranjem, po epohama



Slika: greška bez CLAHE pretprocesiranja, po epohama

Zaključak

Kao što je prikazano, prednost u rešavanju ovog problema ima konvoluciona funkcija koja koristi CLAHE metod u odnosu na onu koja ne koristi nijedan metod pretprocesiranja.

Iako je rezultujuća preciznost mreže velika, primenom mreže na neki drugi *dataset* može se primetiti pad preciznosti na 91% (primenom na **rMASTIF dataset**). Uzrok ovome predstavlja ograničenost primenjenog skupa podataka, kao i potencijalan *overfitting* koji je nastao usled treniranja podataka. Jedno od rešenja ovog problema bi stoga mogla biti primena nekih drugih metoda regularizacije, ali i treniranje na većem skupu podataka. Takođe, unapređenje mogu doneti i neke druge vrste pretprocesiranja slike, kao što su normalizacija, rotacija, skaliranje slike, drugačije vrste kontrasta itd.