



KONVERTOVANJE SLIKE U TEKST

Smisljana Tedić RA 94/2014
Maja Gulović RA 85/2014

Soft kompjuting

Neuronske mreže

Veštačke neuronske mreže (ANNs) su oblik implementacije sistema veštačke inteligencije inspirisane biološkim neuronskim mrežama. Takvi sistemi uče zadatke zasnovane na primerima bez generalnog programiranja samog zadatka. Njihov cilj je da rešavaju probleme na isti način kao i čovekov mozak.

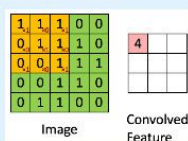
Konvolutivne mreže (CNNs) su višeslojna konvolucija sa nelinearnom aktivacionom funkcijom. Razlikuju se od ANNs jer se ne spajaju svaki ulazni sa svakim izlaznim neuronom u sloju (potpuni povezani sloj), već koristimo konvoluciju na ulazni da izračunamo izlazni sloj. Kao rezultat dobijamo lokalnu vezu u kojoj je svaki region ulaza povezan sa neuronom na izlazu. CNNs su projektovane tako da prednost postižu u radu sa 2D strukturama, kao što su slike ili ulazi poput govornog signala, a čine i jezgro većine sistem veštačke inteligencije danas.

Filteri

Konvolucionni filter (kernel) u bazičnim CNNs je generalizovan linearni model za region slike na koji se primenjuje. Filter se predstavlja dvodimenzionalnom matricom, malih dimenzija, u poređenju sa slikom na koju se primenjuje i sastoji se od realnih vrednosti. Filteri se odnose na konvolucionni sloj arhitekture CNN, u kome se primenjuje na sliku koja se obrađuje u datom sloju. Sam postupak primene filtera na sliku, zove se **konvolucija**.



Sa leve strane je originalna slika, u sredini je filter označen crvenom bojom, primenjen na receptivno polje. Sa desne strane je rezultujuća slika nakon primene filtera.

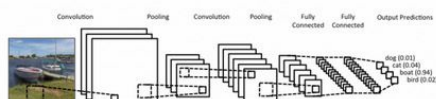


Princip rada CNNs

Ideja CNNs jeste da se postavi veći broj slojeva za otkrivanje bitnih osobina ulaznih podataka. U skladu sa tom idejom se konvolucionni filteri primenjuju na sliku kako bi se izvukle korisne karakteristike i kreirale njihove mape.

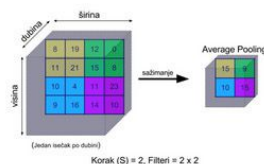
Svaki filter je skup karakteristika nižeg nivoa i on gradi karakter višeg.

U klasifikaciji slika CNN uči da detektuje **ivice** "sirovog" (raw) **piksela** u prvom sloju, zatim koristi **ivice** da detektuje jednostavne **oblike** u sledećem sloju. Zatim te **oblike** koristi u detekciji **kompleksnih karakteristika** (različitih slova, cifara, znakova).



Sažimanje usrednjavanjem

Saživanje usrednjavanjem (average pooling) je forma nelinearnog smanjivanja broja piksela gde se bira piksel sa prosečnom vrednošću. Sloj sažimanja (pooling layer) koristi se sa ciljem progresivnog smanjenja veličine slike, samim tim i broja karakteristika. Izvršava tako sto se računa prosečan piksel u odnosu na vrednosti njegovih suseda, a razlikovajući piksel od ivice suseda. Omogućava fiksnu veličinu izlazne matrice što je potrebno za klasifikaciju. Rezultat sažimanja je "zamagljena" slika.



Računanje prosečnih vrednosti iz matrice oblika 4x4, primenom operacije sažimanja usrednjavanjem (veličinom filtera 2x2 i korakom 2), dobijamo matricu oblika 2x2, koja sadrži prosečne vrednosti.

Treniranje mreže

PREDMET RADA

Testiranje mreže

- Pronalazi sve moguće varijacije istih slova, cifara i specijalnih znakova upoređujući njihove različite ulazne slike.
- Obrada podataka na osnovu kojih se obučava mreža: konvertovanje baferovanih slika (RGB -> Hex). Potrebno je da se dobiju vrednosti za crvenu, zelenu i plavu boju u rasponu od 0 do 255 iz slike, a zatim se one pretvaraju u heksadecimalni niz koji se može koristiti za određivanje boje u html/css kodu.
- Obezbeđuje da sve ulazne slike budu istih dimenzija: menja veličinu matrice tako što prvo konvertuje matricu u baferovanu sliku (Hex -> RGB), zatim se promeni veličina na odgovarajuću i naposljetku se opet vrši obrnut proces pretvaranja skalirane baferovane slike u matricu.
- Vrš se sažimanje usrednjavanjem (average pooling).
- Obučava se mreža.
- Obučene podatke čuvamo u fajlu NaucenoAlfabet.txt, a slike u folderu naucenoSlike.

- Učitavanje u prethodno obučene mreže iz fajla NaucenoAlfabet.txt.
- Detekcija granica ulazne slike: razlikovanje piksela od ivica suseda i omogućavanje kreiranja i kretanja filtera (submatrice).
- Proces poredjenja matrice ulazne slike sa matricama naučenih slika u cilju prepoznavanja pojedinačnih slova, cifara ili specijalnih znakova sa ulazne slike.
- Ispis izlazne vrednosti - svih prepoznatih znakova tj. konvertovanog teksta.

