

Prepoznavanje saobraćajnih znakova pomoću CNN

Jana Jovičić, Jovana Pejkić

Prezentacija seminarskog rada
u okviru kursa
Računarska inteligencija
Matematički fakultet

jana.jovicic755@gmail.com, jov4ana@gmail.com

Sadržaj

- 1 Cilj rada
- 2 Informacije o korišćenom skupu podataka
- 3 Modeli
 - Model 1
 - Model 1 nastavak
 - Model 2
 - Model 2 nastavak
 - Unutrašnja struktura CNN
 - Filtriranje i proširivanje
 - Aktivaciona funkcija
 - Agregacija metodom Maxpool
 - Model 3: AlexNet
 - Model 3: AlexNet - nastavak
 - Model 3: AlexNet - statistike
 - Model 3: AlexNet - statistike - nastavak
 - AlexNet i LeNet-5

Cilj rada

- Za bazu podataka kineskih saobraćajnih znakova izvršiti što precizniju klasifikaciju
- Implementirati CNN u programskom jeziku Python uz korišćenje Keras biblioteke
- Isprobati nekoliko različitih arhitektura mreže
- Uporediti dobijene rezultate i izvesti zaključke

Informacije o korišćenom skupu podataka

- Baza sadrži **6164 slika** saobraćajnih znakova
 - podeljenih u **58 klasa**
 - pri čemu **trening skup** sadrži **4170 slika**
 - a **test skup** **1994 slika**
- Zbog nejednakog broja slika (negde 5, negde 400) po klasama, korišćen je deo baze
- Izdvojeno je **10 klasa** koje su imale približno jednak broj slika
- Dobijen je **trening skup** od **1693 slika** i **test skup** od **764 slika**

Unutrašnja struktura CNN

■ Konvolutivni sloj

- Operacija konvolucije:

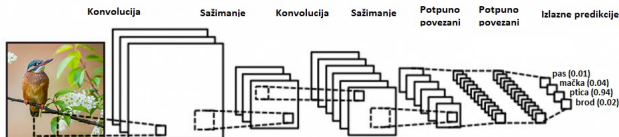
$$(f * g)_{ij} = \sum_{k=0}^{p-1} \sum_{l=0}^{q-1} f_{i-k, j-l} * g_{k,l}$$

■ Sloj agregacije

- Maxpooling / Averagepooling

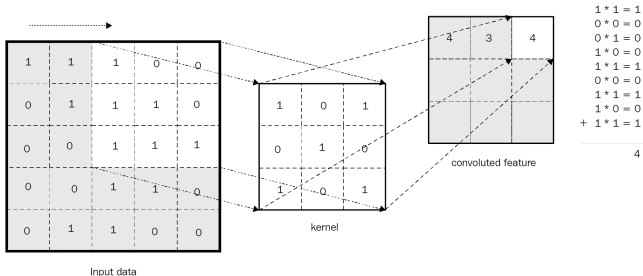
■ Potpuno povezani sloj

- Softmax funkcija
- Categorical crossentropy



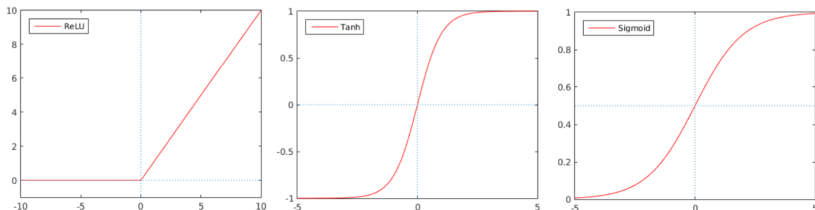
Filtriranje, proširivanje i korak

- Vrš se na konvolutivnom sloju
- **Filtriranje** je „ekstraktovanje“ karakteristika ulaza (slike)
 - tako što se izvršava operacija konvolucije
- Dimenzija mape nakon filtriranja je manja od dimenzije ulaza
- Vrš se **proširivanje** ulazne matrice
 - nulama / vrednostima koje su već na obodu
- **Korak** - za koliko piksela se filter pomera duž slike



Aktivaciona funkcija

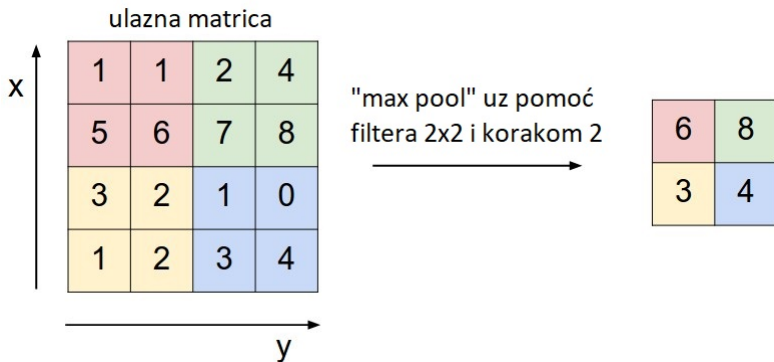
- Funkcija kojom se ograničavaju vrednosti izlaza neurona
 - opseg izlaza neurona obično je u intervalu $[0,1]$ ili $[-1,1]$
- Više vrsta aktivacionih funkcija:
 - **ReLU** (Rectified Linear Unit): $f(z) = \max(0, z)$
 - **tanh**: $\tanh(x) = \frac{e^{2x} - 1}{e^{2x} + 1}$
 - **sigmoid**: $\sigma(x) = \frac{1}{1 + e^{-x}}$



Slika: Grafici funkcija: ReLU, tanh i sigmoid

Agregacija metodom Maxpool

- Jedan od metoda koji se koristi na sloju za agregaciju, najzastupljeniji
- Vraća **maksimum** dela slike prekrivene filterom
- Ideja je da se informacije o slici što više „ukrupne“



Model 1

- Jedan od prvih modela koji je imao uspeha nad test podacima
- Sastoji se iz:
 - 4 konvolutivna sloja
 - 2 agregirajuća sloja
 - 2 potpuno povezana sloja
- U svim konvolutivnim slojevima:
 - velicina jezgra je 3x3
 - broj iteracija na izlazu iz konvolucije je 32
- U svakom sloju se koristi **ReLU** aktivaciona funkcija
- Agregacija se vrši biranjem **maksimalne vrednosti** dela mape karakteristika koji je prekriven iterom

Model 1 nastavak

- Funkcijom **Dropout()** je isključivan određen broj nasumično odabranih neurona (da bi se sprečilo preprilagodjavanje)
- Nakon agregacija je isključeno 20% neurona, pre FC sloja 50%
- Poslednji potpuno povezani (FC) sloj
 - ima onoliko neurona koliko ima klasa
 - koristi softmax aktivacionu funkciju
- Učenje modela je sprovedeno u 30 epoha
- Batch size je postavljen na 32
 - sto znaci da u svakoj iteraciji uzima 32 primerka iz trening skupa koja ce biti propagirana kroz mrežu
- Optimizacija modela je izvršena pomocu **gradijentnog spusta**

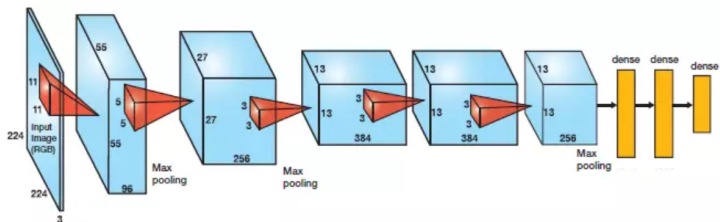
Model 2

Model 2 nastavak



Model 3: AlexNet

- AlexNet arhitektura je jedna od prvih dubokih mreža
- Sastoji se iz:
 - 5 konvolutivna sloja
 - 3 potpuno povezana sloja
- Kao aktivaciona funkcija koristi se ReLu



Slika: AlexNet arhitektura

Model 3: AlexNet - nastavak

- Mreža koju smo implementirale svaki izlaz iz konvolutivnog sloja normalizuje pre nego što ga prosledi narednom sloju
- Pre normalizacije (nakon 1., 2. i 5. konv. sloja) - agregacija
 - sa parametrom padding = 'valid' (nema proširenja)
- Poravnavajući sloj ("ispravlja" mapu karakteristika u vektor)
- 3 Dense sloja + Dropout()
 - Dropout() sprečava prilagodavanje modela
- Na kraju je izlazni sloj koji preslikava ulaz u zadati broj klasa
 - kao funkciju aktivacije koristi Softmax.

Model 3: AlexNet - statistike

- Kod AlexNet modela, vreme izvršavanja je oko 10 puta veće za 100 nego za 10 epoha
- Najmanja preciznost za AlexNet model se postiže za funkciju sigmoid, 10 epoha i Batch size 256, a iznosi 0.696
 - to je jedini slučaj da je preciznost ispod 0.71
 - uglavnom su preciznosti u intervalu [0.801, 0.934]
 - ne osciluje mnogo izvan pomenutog intervala
- Prosečna preciznost za 10 epoha je 82%, a za 100 epoha 90%
- Iz ovoga se može zaključiti da je za AlexNet model i dati skup podataka bolje trenirati model u što više epoha
- Ovo potvrđuje i činjenica da je u 1000 epoha ovaj model dostigao najveću preciznost od 95% (veću od sva tri modela)

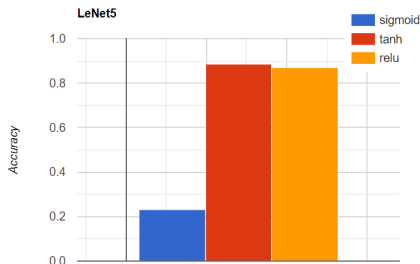
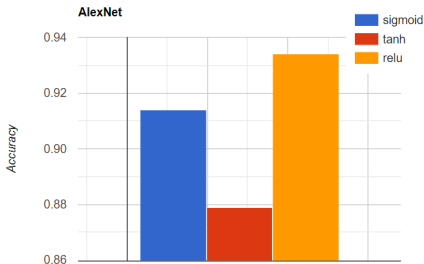
Model 3: AlexNet - statistike - nastavak

AlexNet					
Br. epoha	Funkcija aktivacije	Optimizator	Batch size	Preciznost na test skupu	Vreme izvršavanja (sekunde)
10	relu	adam	32	0.838	44.356
10	relu	adam	64	0.848	37.813
10	relu	adam	256	0.715	33.786
10	sigmoid	adam	32	0.919	43.755
10	sigmoid	adam	64	0.845	38.291
10	sigmoid	adam	256	0.696	36.811
10	tanh	adam	32	0.801	43.664
10	tanh	adam	64	0.906	39.264
10	tanh	adam	256	0.851	34.403
100	relu	adam	32	0.934	372.984
100	relu	adam	64	0.934	334.667
100	relu	adam	256	0.906	295.523
100	sigmoid	adam	32	0.914	372.971
100	sigmoid	adam	64	0.914	334.225
100	sigmoid	adam	256	0.856	292.743
100	tanh	adam	32	0.879	373.153
100	tanh	adam	64	0.921	339.918
100	tanh	adam	256	0.906	296.921

Slika: Rezultati za AlexNet model za različite vrednosti broja epoha (10 i 100), funkcije aktivacije (relu, tanh i sigmoid) i Batch size (32, 64 i 256)

AlexNet i LeNet-5

- Primititi da se vrednosti na y-osi ova dva grafika razlikuju
- To je zato što LeNet-5 daje malu preciznost za f-ju sigmoid
- Maksimalna preciznost AlexNet mreže je 0.94, a LeNet-5 0.9



Slika: Grački prikaz preciznosti za mreže AlexNet i LeNet-5 za 100 epoha, za funkcije sigmoid, tanh i ReLu

Zaključak

-
-
-

Literatura



Ime Prezime pisca (godina)

Ime knjige Mesto, tekst, godina.



Ime Prezime pisca (od-do)

Naziv dela, on-line at: <http://adresa.org/>.