Izvješće

**Uvod**

Izvješće pokriva opis aktivnosti projektnog zadatka, dobivene rezultate, naše komentare, zaključke te popis izvora.

**Opis projektnog zadatka**

Projektni zadatak je bio usporediti točnost ResNet-18 arhitekture naspram klasične CNN arhitekture pri treniranju nad CIFAR i MNIST skupom podataka. Glavna razlika između ovih mrežnih arhitektura je postojanje takozvanih shortcut connectiona (residual connection) između slojeva. Njihova uloga je da u sljedeći sloj prenose input trenutnog sloja, koji nije bio transformiran prolaskom kroz sloj. Proveli smo treniranje ResNet-18 mreže s različitim optimizerima (adam, stohastic gradient descent) te različitim batch size-om (128, 512) da saznamo kako ta dva koncepta utječu na njenu točnost. Treniranje je provedeno kroz 10 epoha.

**Dobiveni rezultati**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Točnost | | | |
| optimizer | stohastic gradient descent | | Adam (learning rate = 3e-4) | |
| batch size | 128 | 512 | 128 | 512 |
| CIFAR | 88,594% | 86,788% | 89,350% | 88,074% |
| MNIST | 98,915% | 98,938% | 98,772% | 98,938% |

**Komentari rezultata**

Uočena je razlika u točnosti mreže ovisno o skupu podataka. Točnost je veća nad skupom MNIST, nego nad skupom CIFAR. Smatramo da svaki skup podataka daje drugačiji n-dimenzionalni oblik koji utječe na traženje globalnog minimuma funckije i na djelovanje optimizera. Pretpostavljamo da skup podataka MNIST daje „pogodniji oblik“ za mrežu ResNet-18.

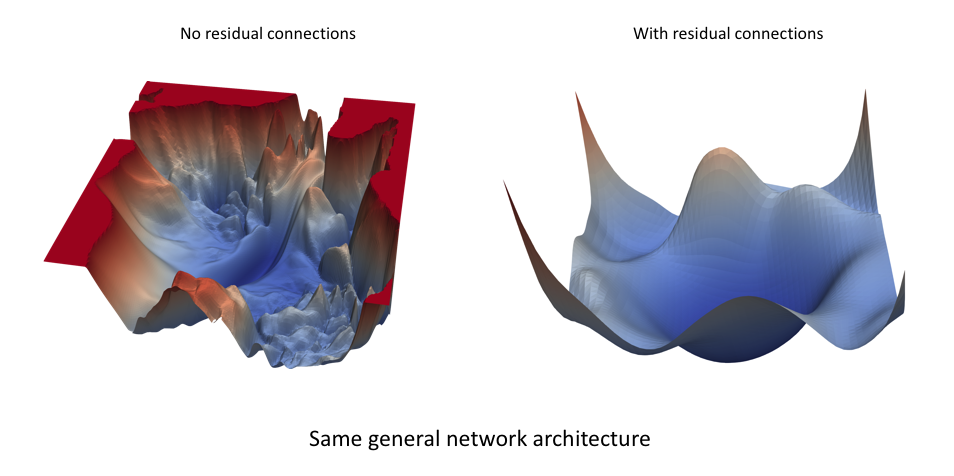
Također, za skup CIFAR (vrijedi i za sgd i za adam) primjećujemo malo veću točnost kad koristimo batch size 128 (+1,806%, +1,276). Dok za skup MNIST primjećujemo bolju točnost korištenjem batch size 512, no te razlike su u izuzetno male (+0,023%, +0,166%). Smatramo da se iz ovih vrijednosti ne mogu donijeti zaključci u skladu sa već poznatim efektima različitih vrijednosti batch size-a.

**Usporedba točnosti ResNet-18 modela i osnovnog CNN modela**

--usporediti sa CNN i plottati--

**Zaključak**

Bolja performansa ResNet-18 mreže nad klasičnom CNN mrežom je posljedica postojanja shortcut connectiona. Njihov učinak na loss landscape neuronske mreže je vidljiv na priloženoj slici. Traženje minimuma i optimizacija funkcije gubitka je manje zahtjevna nad „ljepšim“ loss landscape-om i manja je vjerojatnost da završimo u nekom lokalnom minimumu.



**Izvori:**

--linkati samo znanstvene radove--