

Generisanje lica korišćenjem GAN neuronskih mreža

Seminarski rad u okviru kursa
Računarska inteligencija
Matematički fakultet

Antić Dimitrije, Novaković Andrija
mi16128@alas.matf.bg.ac.rs, mi16068@alas.matf.bg.ac.rs

27. januar 2020.

Sažetak

U ovom tekstu je ukratko opisan rad na GAN neuronskim mrežama primenjenih na problem generisanja ljudskih lica i postignuti rezultati uz osvrt na originalne prateće radove.

Sadržaj

1	Uvod	2
2	Osnovni pojmovi	2
2.1	Konvolutivne neuronske mreže	2
2.2	Generativne adversarijalne neuronske mreže	2
2.3	Nešov ekvilibrijum	3
3	Adversarijalne mreže	3
4	Prednosti i mane adversarijalnih mreža	4
5	Eksperimentalni rezultati	5
5.1	Konfiguracija modela	5
5.2	Postignuti rezultati	6
6	Zaključak	7
7	Literatura	7

1 Uvod

Korišćenjem generativnih adversarijalnih mreža, predložene u radu [1], cilj je bio generisanje ljudskih lica, i usput se upoznamo sa matematičkim principima koji stoje iza GAN mreža.

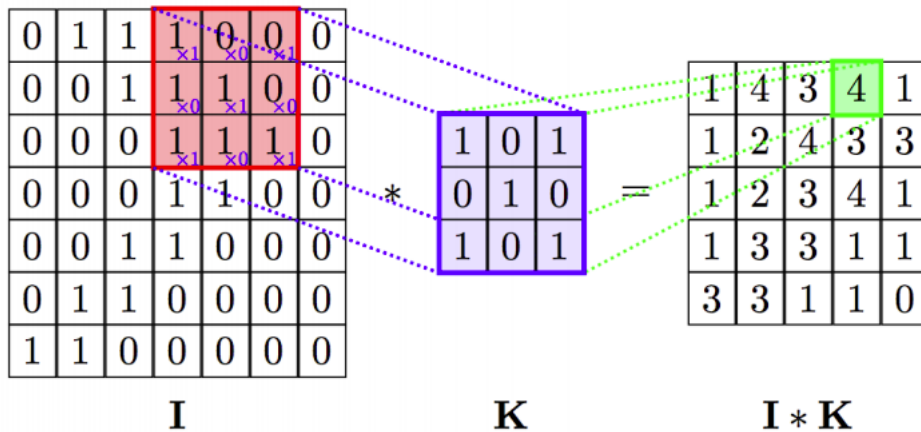
Korišćen je skp podataka **CelebFaces Attributes Datasets (CelebA)** koji sadrži više od 200.000 slika poznatih ličnosti. Svaka slika je isečena na veličinu od približno 28px kako bi na slici bilo samo lice osobe.

2 Osnovni pojmovi

2.1 Konvolutivne neuronske mreže

Konvolutivne neuronske mreže se intenzivno koriste u obradi signala poput slika i zvuka, s toga su i generator i diskriminator upravo konvolutivne neuronske mreže. Zasnivaju se na sposobnosti da konstruišu atribute, ali ne iz već datih, nego iz sirovih signla. Uče tzv filtere, čijom konvolutivnom primenom detektuju određena svojstva signala. Njihov značaj je u tome što ne zahtevaju ljudski angažman u definisanju relevalntnih svojstava signala, već u zavisnosti od problema mreža sama ustanovi koja su svojstva bitna kroz učenje adekvatnih filtera.

Uobičajna struktura konvolutivne mreže podrazumeva smenjivanje konvolutivnih slojeva i slojeva agregacije. Izlazi konvolutivnog sloja se tranf-sormišu nelinearnom aktivacionom funkcijom. Na izlaze poslednjeg od tih slojeva se obično nadovezuje potpuno povezana mreža, koja uči nad atributima koje prethodni slojevi konstruišu.



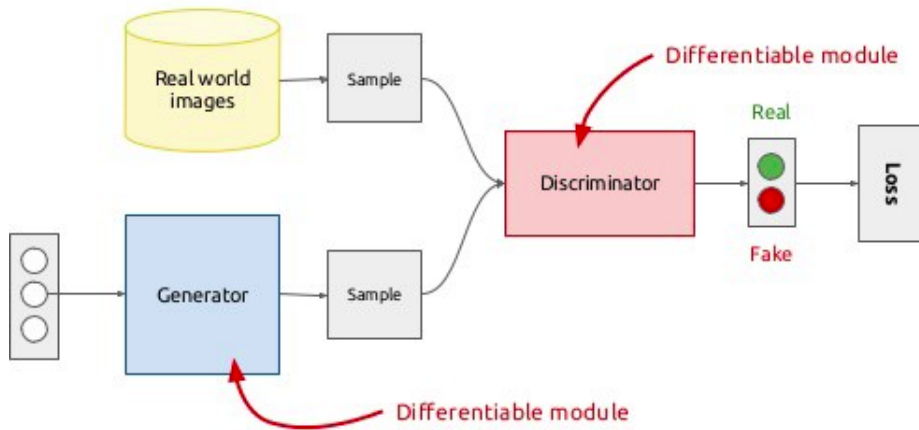
Slika 1: Slika 1: Konvolucija ulazne slike sa filterom

2.2 Generativne adversarijalne neuronske mreže

Generativne adversarijalne mreže se sastoje iz dva modela. Generativni model G koji treba da zaključi raspodelu podataka i diskriminator D koji procenjuje verovatnoću da je uzorak došao iz trening podataka a ne od generatora G. Ovakav način rada podseća na min-max igru dva

igrača. Kada se oba modela definišu kao višeslojni perceptroni, ceo sistem može biti treniran propagacijom unazad. Proces treniranja traje dok se ne postigne Nešov ekvilibrijum.

U predloženoj arhitekturi adversarijalnih mreža, generativni model se stavlja pred protivnika diskriminativni model koji ima uči da prepozna da li je uzorak iz raspodele modela ili iz raspodele podataka. Generativni model može da se posmatra kao falsifikator koji želi da generiše falsifikat i da prođe "nekažnjeno", dok diskriminator ima zadatak da otkrije kada generator vara.



2.3 Nešov ekvilibrijum

U teoriji igara, Nešov ekvilibrijum je koncept rešenja igre koja uključuje dva ili više igrača, kod kog se podrazumeva da svaki igrač zna strategije ekvilibrijuma ostalih igrača, i nijedan igrač ništa ne može da dobije tako što samo on promeni svoju strategiju (jednostrano je promeni).

Ako je svaki igrač izabrao strategiju, i nijedan igrač ne može da profitira promenom svoje strategije pod pretpostavkom da ostali igrači ne promene svoje strategije, onda trenutni skup izabranih strategija, i odgovarajućih dobitaka predstavlja Nešov ekvilibrijum. Drugim rečima, da bi postojao Nešov ekvilibrijum, svaki igrač mora odrično da odgovori na pitanje: Znajući strategije ostalih igrača, i pretpostavljajući da ih oni neće promeniti, da li ja mogu da profitiram promenom svoje strategije? Prosto rečeno, dva igrača su u Nešovom ekvilibrijumu ako je svaki doneo najbolju moguću odluku, uzevši u obzir odluku protivnika.

3 Adversarijalne mreže

Adversarijalni model se najlakše primenjuje kada su oba modela (generator i diskriminator) višeslojni perceptroni. Da bi generator naučio distribuciju p_g iz podataka \mathbf{x} , definiše se apriorna raspodela nad promenljivom \mathbf{z} , $p_z(\mathbf{z})$ koja predstavlja random šum, a zatim predstaviti preslikavanje kao $G(\mathbf{z}; \theta_g)$, gde je G diferencijabilna funkcija predstavljena višeslojnim perceptronom sa parametrima θ_g . Takođe se definiše drugi višeslojni perceptron $D(\mathbf{x}; \theta_d)$ čiji je izlaz jedan skalar. $D(\mathbf{x})$ predstavlja

verovatnoću da je \mathbf{x} iz podataka pre nego iz raspodele p_g . reniramo D tako da maksimizujemo verovatnoću dodeljivanja ispravnih klasa i trening i uzorcima generisanih od strane G . Uporedo treniramo c da minimizuje $\log(1 - D(G(z)))$.

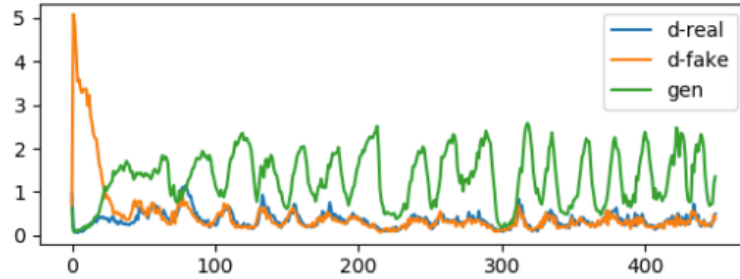
Preciznije, D i G igraju minmax igru sa ciljnom funkcijom $V(G, D)$:

$$\min_G \max_D V(D, G) = \mathbb{E}_{\mathbf{x} \sim p_{\text{data}}(\mathbf{x})} [\log D(\mathbf{x})] + \mathbb{E}_{\mathbf{z} \sim p_z(\mathbf{z})} [\log(1 - D(G(\mathbf{z})))]$$

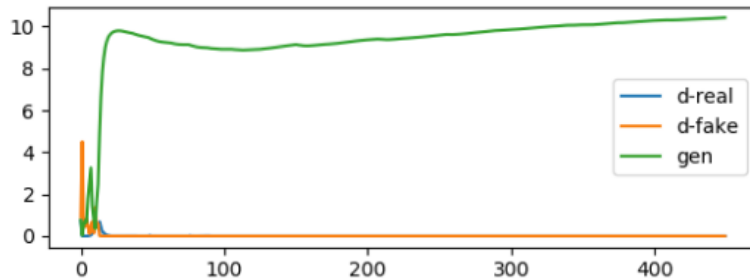
4 Prednosti i mane adversarijalnih mreža

Jedan od problema koji nastaju pri treniranju GAN mreža je *mode collapse*. To se odnosi na situaciju kada generator sposoban da generiše mali podskup instanci. Na primeru slika može da se identifikuje pregledom velikog uzorka, gde su slike pokazuju malu raznolikost, i neke od njih se iznova ponavljaju.

Na grafiku ovaj problem se ispoljava u vidu oscilacija u funkciji gubitka generatora.



Često dolazi do nekonvergencije modela. Najčešće se dešava kada generator i diskriminator ne mogu da postignu ekvilibrijum. Na samom početku procesa, generator generiše slike koje diskriminator lako prepoznaje i ima veliku vrednost funkcije gubitka, najčešći uzrok je korišćenje agresivne optimizacije ili korišćenje prevelikih odnosno premalih filtera.



Sa grafika se jasno vidi da je funkcija gubitka diskriminatora jako mala, što znači da jako lako može razlikovati prave i lažne slike, dok generator ne uspeva da napravi dovoljno uverljivu sliku i njegova funkcija greške nastavlja da raste. Prikazani grafici su dobijeni nad *MNIST* skupom.

5 Eksperimentalni rezultati

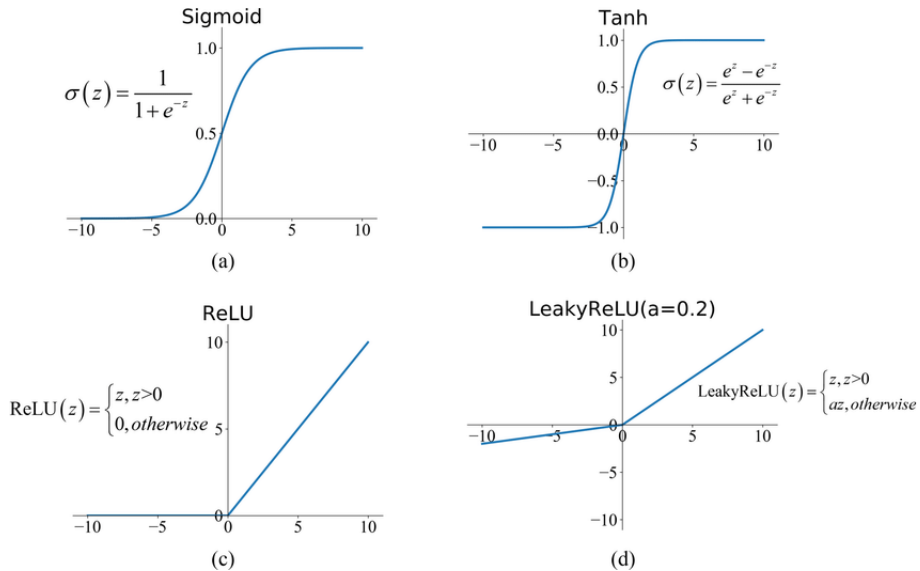
5.1 Konfiguracija modela

Korišćena je biblioteka *TensorFlow*. Generator i diskriminator su realizovane kao konvolutivne neuronske mreže. Zbog kompleksnosti i vremena trajanja treninga modela, korišćen je filter veličine 5x5 u svim konvolutivnim slojevima sa veličinom koraka 2. Takođe se koristi *batch normalization* da bi se smanjila varijansa u aktivacijama skrivenih slojeva.

Korišćena je *leaky-relu* aktivaciona funkcija zato što pri korišćenju sigmoidne funkcije postoje dva glavna problema:

1. Za ekstremno velike i ekstremno male vrednosti gradijent teži ka 0
2. U situaciji gde se radi sa slikama u RGB formatu gradijent će biti samo pozitivan, jer sigmoidna funkcija nije centrirana oko 0, što dodatno otežava nalaženje (lokalnog) minimuma ili maksimuma

Njena alternativa, tangens hiperbolički jeste centrirana oko 0 ali i dalje postoji problem 1.

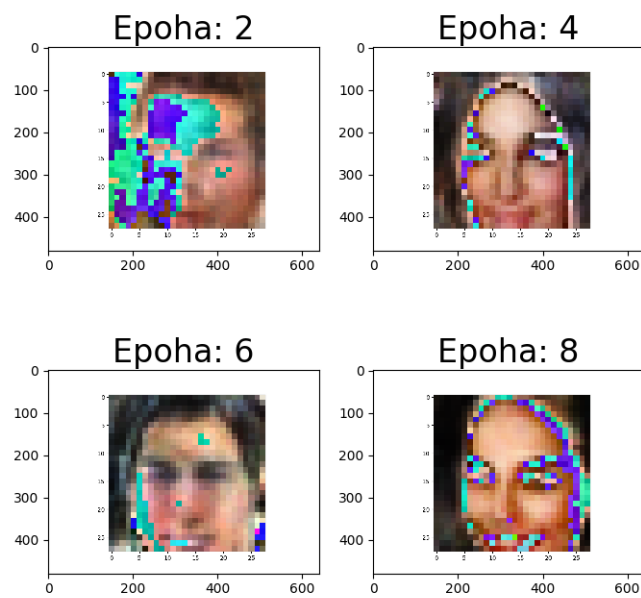
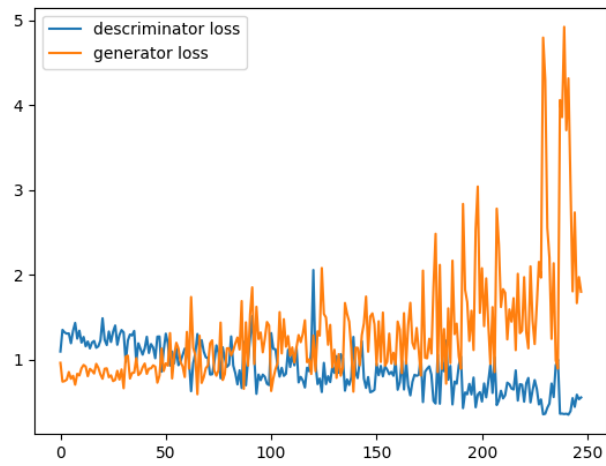


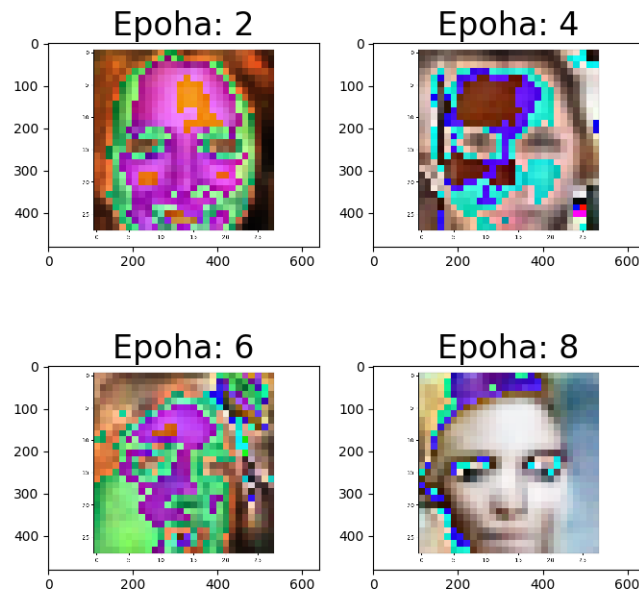
ReLU aktivaciona funkcija nema problem br. 2, a problem br. 1 ima samo na intervalu $[0, \infty]$, zbog toga se koristi l-*ReLU* koja uvodi male pozitivne otklone čime rešava oba navedena problema i u praksi konvergira i do šest puta brže.

Kao funkcija greške koristi se *cross entropy*.

5.2 Postignuti rezultati

U ovom poglavlju su prikazani neki, izabrani, postignuti rezultati uz grafik grešaka diskriminatora i generatora.





6 Zaključak

Treniranje je vršeno u 8 epoha na *Intel(R) Core(TM) i7-6700HQ CPU @ 2.60GHz* procesoru i *NVidia GeForce GTX 950* grafičkoj kartici, u trajanju od 12h, dobijene slike su veličine 12Kb. Konvergencija GAN-a se dostiže na mnogo većem broju epoha, što zahteva znatno bolje resurse. Uzevši u obzir sve probleme, sa promenama u arhitekturi i parametirima kao i vremenu treniranja potencijalno bi se dobili bolji rezultati.

7 Literatura

- [1] *Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair†, Aaron Courville, Yoshua Bengio*, Generative Adversarial Nets
- [2] *Tim Salimans, Ian Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, Xi Chen*, Improved Techniques for Training GANs
- [3] *Mladen Nikolić, Anđelka Zečević*, Mašinsko učenje