

NAPOVEDOVANJE OBRAZOV OTROK Z GAN METODO

NEJC ŠUKLJE, LUKA ŠAJN

Univerza v Ljubljani, Fakulteta za računalništvo in informatiko, Ljubljana, Slovenija
nejc.suklje@hotmail.com, luka.sajn@fri.uni-lj.si

Sinopsis Naslov "Napovedovanje slik otrok obrazov z GAN metodo" je že na prvi pogled vzbudil motivacijo, ker je potencialni rezultat zanimiv. GAN metoda je tudi dokaj novo področje, kar vse skupaj naredi še bolj zanimivo, saj ni veliko raziskav, v katerih je bila uporabljena. Na problem smo pogledali tudi s strani "uporabnika", torej nekoga, ki bi rad videl sliko svojega potencialnega otroka. Iz tega sklepamo, da bi končna rešitev na trgu doživela zanimanje, še posebej v bogatejšem sloju prebivalstva. Raziskav na to temo je bolj malo, vendar dovolj za osnovo pri izdelavi naloge. Cilj naloge je bil dobiti sliko obraza otroka na osnovi izgleda njegovih staršev.

Prispevek temelji na:

Šuklje, N. (2022).

Napovedovanje obrazov otrok z GAN metodo: diplomsko delo,

Univerza v Ljubljani, Fakulteta za računalništvo in informatiko. Ljubljana.

Ključne besede:

grebenska regresija, latentni prostor, generiranje obrazov, GAN metoda, sorodstvena podobnost

PREDICTING CHILDREN FACES WITH GAN METHOD

NEJC ŠUKLJE, LUKA ŠAJN

University of Ljubljana, Faculty of Computer Science and Informatics, Ljubljana,
Slovenia
nejc.suklje@hotmail.com, luka.sajn@fri.uni-lj.si

**The paper is
based on:**
Šuklje, N. (2022).
Napovedovanje
obrazov otrok z
GAN metodo:
bachelor's thesis,
University of
Ljubljana, Faculty
Computer and
Information
Science. Ljubljana.

Keywords:
ridge regression,
latent space,
GAN method,
generation of
faces,
kinship
resemblance

Abstract The title "Prediction of children's faces with the GAN method" aroused motivation at first glance, because the potential result is interesting. The GAN method is also a fairly new thing, which made it more interesting because there aren't many studies that have used it. We also looked at the problem from the "buyer's" side, that is, someone who would like to see a picture of their potential child. From this, we conclude that the final solution on the market would experience an increase, especially in the wealthier strata of the population. There is less research on this topic, but it still helped in the preparation of the assignment. The goal of the task was to get a picture of a child's face from its parents.

1 Uvod

GAN metoda je pritegnila veliko pozornosti med raziskovalci na področju računalniškega vida (Šuklje, 2022). Največ se uporablja v medicini za sintezo slik. Cilj je pridobiti čim večjo količino podatkov oziroma slik, ki so podobne realnim slikam. Te generirane slike so nato uporabljene skupaj z dejanskimi za potrebe učenja. Študenti medicine lahko na njih izvajajo analize in se tako učijo na novih izračunanih primerih. Zaradi raznolikosti, ki jo dodamo z namišljenimi slikami, je lahko pridobljeno znanje širše.

Delo je osnovano na že prej pripravljeni kodi Matjaža Mava (Šuklje, 2022). Osnova je zastavljena kot napovedovanje otrok na podlagi izgleda staršev z globokim učenjem (ang. deep learning). Začeli bomo s predstavitvijo GAN metode, potrebnimi splošnimi informacijami za razumevanje le-te, nato pa pridejo na vrsto postopki in metode, ki so bili uporabljeni, da smo prišli do želenega rezultata. Rezultati bodo predstavljeni in komentirani. Primerjana bo podobnost generiranih slik z resničnimi. Po analizi smo tudi podali možne ideje za nadaljnje delo.

2 Pregled področja

2.1 GAN metoda grobo

GAN metoda je relativno mlada metoda, ki se uporablja pri procesiranju slik (Šuklje, 2022). Uporablja se tako, da na začetku uporabimo učno množico slik, s katero se bo program učil obraznih značilnic. Ta množica se čez čas razširja, ker program proizvaja nove slike, ki so dodane v to množico in so približek tistim, ki so bile v množici že od začetka.

2.2 Prva omemba GAN metode

GAN metoda je bila prvič predstavljena in opisana v članku (Goodfellow in sod., 2014). Članek govori o modelu, ki je sestavljen iz dveh funkcij. Prva je zadolžena za generiranje podatkov, druga pa za oceno primernosti generiranih podatkov, in jih opredeli kot pravilne ali napačne. Cilj teh dveh funkcij je, da povečamo verjetnost prepoznave napake druge funkcije. S tem dosežemo, da je podatkov opredeljenih kot pravilnih malo, vendar so zelo podobni resničnim. Do te mere, da ne moremo ločiti med resničnimi in neresničnimi. Avtorji članka generator in diskriminator

razložita na zelo razumljiv način. Generator si lahko predstavljamo kot ponarejevalca, ki poskuša narediti čim bolj pristne ponaredke, ki bodo ostali neodkriti, diskriminator pa kot policijo, ki išče ponaredke in njihovo uporabo na trgu (Škuje, 2022). Modela nenehno izboljšujeta svoji funkciji (Goodfellow in sod., 2014).

2.3 Sorodna dela

Članek (Goodfellow in sod., 2014) je pritegnil veliko pozornosti med raziskovalci na področju računalniškega vida. Največ se uporablja v medicini za sintezo slik in posledično za pridobivanje večje količine podatkov (Yi in sod., 2019).

Članek (Rani, 2014) govori o generiranju slik otrok. Bolj specifično tistih, ki imajo avtizem. Zaradi bolezni je njihovo izražanje čustev različno od ljudi, ki nimajo avtizma. S procesiranjem njihovih izrazov na obrazu se poskuša določiti čustva, ki jih ta trenutek izražajo (Rani, 2014).

3 Kako deluje GAN metoda

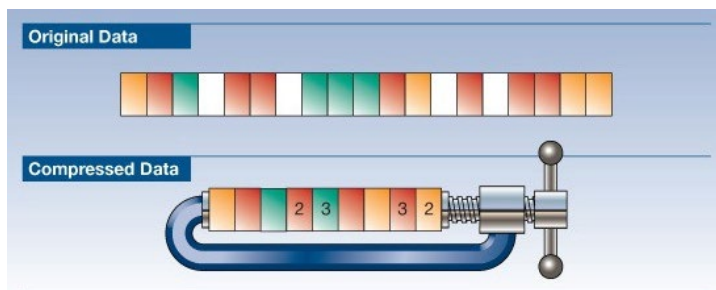
Kratice GAN pomeni "Generative Adversarial Network", v slovenščino to prevedemo kot generativna nasprotniška mreža. Metoda uporablja 2 nevronske mreže, ki delujeta ena proti drugi. Iz tega tudi izhaja beseda v imenu metode, in sicer "Adversarial". Nevronske mreže ustvarjata podatke, ki bi lahko veljali za resnične (Goodfellow in sod., 2014).

Za boljše razumevanje metode bomo navedli konkreten primer. Opisali bomo tudi generator in diskriminator, ki sta ključni funkciji v metodi (Škuje, 2022). Predpostavimo, da generativni algoritem generira nam nepoznane podatke, za katere ne poznamo njihovih značilnosti. Zdaj je naloga diskriminatorja, da nam te podatke kategorizira in posledično označi. Glede na značilnosti podatkov dobimo oznako. Vhodne podatke seveda po potrebi razdelimo na več vrst. Kot primer vzemimo e-pošto. Naloga diskriminatorja je, da nam označi posamezno sporočilo kot nezaželeno/vsiljeno ali pa običajno. Diskriminator vhodne podatke razdeli na besede v sporočilu in na naslov sporočila (zadeva). Matematično to lahko predstavimo s pomočjo verjetnostnega zapisa. Oznako označimo z z , kot značilnost oziroma z x , kot podatke. Formula, ki povezuje ti dve neznanki je $p(y | x)$. Formulo preberemo kot "verjetnost y pri x ", v našem konkretnem primeru bi se to glasilo kot "verjetnost, da je sporočilo nezaželeno glede na besede v sporočilu".

Torej diskriminator glede na značilnosti podatkov le-te označi. Zdaj se vrnemo k generatorju. Njega si lahko predstavljamo, da deluje obratno. Torej, da iz označbe dobi značilnosti. To pomeni, da bo zgeneriral podatke z določenimi značilnostmi, ki veljajo za neko podano označbo. Imamo še en način, kako lahko opišemo generator in diskriminator. Diskriminator pozna meje med razredi, generator pa pozna porazdelitev vsakega razreda posebej (Nicholson, 2020).

4 Latentni prostor

Latentni prostor je izraz, ki ga moremo razumeti, da nam bo celotna slika bolj jasna. V našem programu smo operirali s slikami in to na tak način, da smo jih predstavili vektorsko, torej s številkami. Problem, ki se tu pojavi, je količina podatkov, ki jih posamezna slika vsebuje. Zmanjšati moramo dimenzije, ki jih porabimo za predstavitev slike. Dimenzija novo nastalega vektorja mora biti občutno manjša. Vektor z zmanjšanimi dimenzijami, ki ga dobimo, se imenuje latentna predstavitev slike. Na sliki 1 je predstavljena ideja manjšanja dimenzij. Devetnajst točk stisnemo tako, da jih lahko prestavimo samo z devetimi (Manning, 2022), (Silva, 2018), (Tiu, 2020).



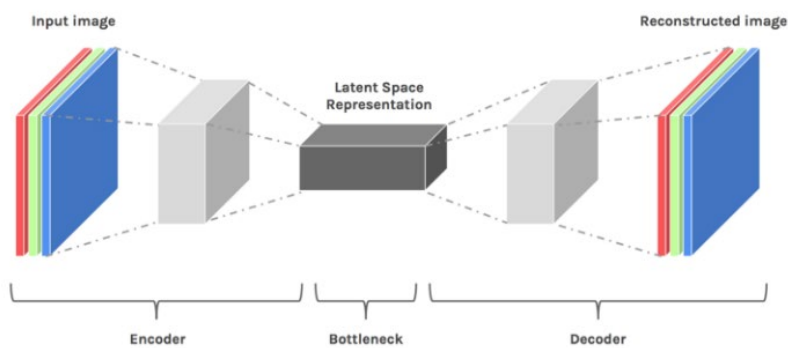
Slika 1: Primer kompresije vektorja.

Vir: (Manning, 2022)

Torej imamo za vsako sliko, ki jo bomo uporabili v našem generatorju, latentno predstavitev v obliki vektorja (Šuklje, 2022). Vsi ti vektorji sestavljajo oziroma oblikujejo prostor, ki mu rečemo latenten prostor. V tem prostoru so slike, torej njihovi vektorji, ki so si podobni bolj skupaj, slike, ki si pa niso, pa so bolj oddaljene. Dimenzija latentnega prostora je določena z velikostjo vektorja. Na primer, če imajo vektorji dolžino 3, to pomeni, da je dimenzija latentnega prostora 3. Na sliki 2 imamo prikazano vizualizacijo latentnega prostora. Slike, ki jih stisnemo, sestavljajo latenten

prostor tako, da so ohranjeni samo pomembni podatki. To je tudi osnova, da sploh lahko stiskamo. Seveda, da lahko iz latentnega prostora dekompresiramo kakšno sliko, potrebujemo dekodek, ki je to sposoben narediti (Manning, 2022), (Silva, 2018), (Tiu, 2020).

Ime »latent space« oziroma latentni prostor izhaja iz besede »hidden« oziroma skrit. To pa zato, ker si težko predstavljamo prostor, ki ima kakšno dimenzijo več kot 3, kaj šele prostore dimenzije reda 100 (Manning, 2022), (Silva, 2018), (Tiu, 2020).



Slika 2: Vizualizacija latentnega prostora.

Vir: (Manning, 2022)

5 Rezultati delovanja programa

Ko je bil program enkrat razvit smo potrebovali slike, da program uporabimo na dejanskem primeru. Za učenje smo uporabili 217 primerov. Torej 1 primer pomeni: slika očeta, slika mame in slika njunega otroka. Za testiranje pa smo imeli 20 takih primerov. Vse slike, ki smo jih uporabili, smo dobili že iz vnaprej pripravljene baze (Silva, 2018). Njihove latentne slike pa nam je zagotovil laboratorij za računalniški vid na fakulteti za računalništvo in informatiko. Napovedovalni model smo zgradili z grebensko regresijo. Testne slike smo dali v model in zadnji korak je bil le, da še te latentne slike generiramo in dobimo dejanske slike. To nam je omogočila grafična kartica NVIDIA GeForce GTX 970M. Hitrost je 924 MHz, velikost spomina pa 3GB. Generiranje 80 slik je trajalo nekaj manj kot 5 minut. Slike smo analizirali sami, za dejansko vrednotenje kakovosti smo uporabili ocene zunanjih neodvisnih ocenjevalcev, ki so podali svoje mnenje. S tem smo lažje zaključili in ocenili celotno delo.

5.1 Analiza

Na naslednji sliki, sliki 3, imamo v vsaki vrstici o leve proti desni: očeta, mamo, njunega otroka in na koncu še generiranega otroka z našim programom. Vsaka vrstica je svoj primer, torej imamo 4 primere.



Slika 3: Primeri staršev, njunega otroka in generiranega otroka

Vir: lasten

To so 4 primeri slik, ki smo jih pokazali zunanjim opazovalcem. Dobili smo mnenja ljudi od starosti 20 pa vse do 50 let. Ob prebiranju mnenj smo ugotovili, da imajo različne starostne skupine drugačno razmišljanje in, da drugače opazajo podrobnosti in celostno sliko. Mlajša starostna skupina, torej okrog 20 let, ni opazila manjših podrobnosti, so pa bila podana mnenja zelo razumljiva in smiselno sestavljena. Starejša starostna skupina, 40-50 let, pa je pokazala čisto nasprotje. Opazili so veliko manjših podrobnosti, vendar so bila njihova mnenja napisana manj enotno, večasih zmedena.

5.2 Naše mnenje

Nad podobnostjo resničnih in generiranih otrok smo bili kar malo presenečeni saj nismo pričakovali, da je možno doseči takšno podobnost. Glede na to, da so generirani otroci zgrajeni le na podlagi izgleda staršev, ne pa tudi genskega zapisa, so v večini bolj podobni samim staršem kot njihovi resnični otroci. Ne najbolj optimalna rešitev, vendar za začetek deluje dovolj dobro.

Na prvi pogled padeta v oči prva in tretja vrstica. Resničen in generiran otrok sta si tako zelo podobna, da smo prepričani, da bi ju večina opazovalcev videla kot sorojenca. Pri prvi in četrti vrstici sta si generirana otroka zelo podobna. Najverjetneje tukaj igra vlogo to, da imata podobni materi. Drugih specifičnih podobnosti ne opazimo, zato bomo zdaj podali še splošno mnenje.

Na prvi pogled zgleda, kot da je generiran otrok podoben staršu istega spola, vendar že pri prej analiziranem primeru, ko smo rekli, da imata primerka podobni materi, se zdi, da to ni čisto res. Predvidevamo, da nas k temu zavede dolžina las in zato otrok izgleda podoben staršu istega spola. Zdi se pa tudi, da so generirani otroci zelo podobni resničnim otrokom, čeprav ti niso bili uporabljeni pri generiranju. S tem dobimo potrditev, da je resničen otrok podedoval veliko videza od staršev, najverjetneje več kot od starih staršev, ki niso uporabljeni v modelu.

6 Zaključek

S končnimi rezultati smo zadovoljni, ker je podobnost smiselna. Dobre rezultate potrjujejo tudi mnenja zunanjih opazovalcev. Izprašani dajejo občutek, da so zelo presenečeni nad podobnostjo. Poleg mnenj v pisni obliki smo z njimi tudi govorili, tam pa je bila presenečenost še bolj izražena. Predvidevamo, da je presenečenost posledica tega, da nimajo predznanj o tej temi in se jim zdi vsaka podobnost izjemna.

6.1 Možne izboljšave

Za nadaljnje delo bi predlagali, da se za generacijo otrok ne uporabi enak delež očeta in matere. Predvidevamo, da bi takšni rezultati bili še bolj zanimivi, ker bi bil otrok moškega spola še bolj podoben očetu in bi lahko iskali po obrazu, kaj je pa tisto, kar je podedoval po mami. Glede na to, da je generiranje temeljilo le na izgledu staršev

bi bilo smiselno v generiranje zraven staršev vključiti tudi stare starše, vendar seveda z manjšim deležem.

Literatura

- Goodfellow I.J. in sod. (2014). *Generative Adversarial Networks*. doi: 10.48550/ARXIV.1406.2661. url: <https://arxiv.org/abs/1406.2661>
- Yi X., Walia E. in Babyn P. (2019). *Generative adversarial network in medical imaging: A review*. V: *Medical Image Analysis* 58, str. 101552. issn: 1361-8415. doi: <https://doi.org/10.1016/j.media.2019.101552>. url: <https://www.sciencedirect.com/science/article/pii/S1361841518308430>.
- Rani P. (2014). *Emotion Detection of Autistic Children Using Image Processing*. V: *2019 Fifth International Conference on Image Information Processing (ICIIP)*. Nov. 2019, str. 532–535. doi: 10.1109/ICIIP47207.2019.8985706
- Nicholson C. (2020). *A Beginner's Guide to Generative Adversarial Networks (GANs)*. url: <https://wiki.pathmind.com/generative-adversarial-network-gan> (pridobljeno 27. 10. 2022)
- Manning publications. *concept latent space in category GAN*. (2022). url: <https://livebook.manning.com/concept/gan/latent-space> (pridobljeno 29. 10. 2022)
- Silva T. (2018). *An intuitive introduction to Generative Adversarial Networks (GANs)*. url: <https://www.freecodecamp.org/news/an-intuitive-introduction-to-generative-adversarial-networks-gans-7a2264a81394> (pridobljeno 29. 10. 2022)
- Šuklje, N. (2022). *Napovedovanje obrazov otrok z GAN metodo*: diplomsko delo, Univerza v Ljubljani, Fakulteta za računalništvo in informatiko. Ljubljana.
- Tiu E. (2020). *Understanding Latent Space in Machine Learning*. url: <https://towardsdatascience.com/understanding-latent-space-in-machine-learning-de5a7c687d8d> (pridobljeno 29. 10. 2022)
- Robinson J.P. in sod. (2021.) *Families in Wild Multimedia: A Multimodal Database for Recognizing Kinship*. V: *IEEE Transactions on Multimedia* 24, str. 3582–3594
- Šuklje N. (2022). *Napovedovanje obrazov otrok z GAN metodo*. Univerza v Ljubljani, Diplomsko delo.

