

Super resolucija MRI slik s pomočjo generativnih modelov

Osnove strojnega učenja in podatkovnega rudarjenja

2024/25

Jan Panjan

Contents

1. Uvod	1
1.1. Problemi MRI slik	1
1.2. O projektu	1
2. Predstavitev GAN strukture	1
2.1. Funkcija izgube	2
3. SRResCycGAN	3
3.1. CycleGAN	4
3.2. HR Generator	4
3.3. Residual GAN	5
4. Predstavitev podatkov	5
5. Podrobnosti implementacije	5
6. Rezultati	5
7. Zaključek	5
Bibliography	6

1. Uvod

Uporaba magnetne resonance za slikanje pacientov je izjemnega pomena za diagnosticiranje bolezni in sledenju boleznim, ki že motijo zdravje pacientov. Podatki o organih, mehkih tkivih in kosteh, katere pridobijo z MRI slikanjem, omogočajo zdravnikom, da bolj učinkovito ocenijo stopnjo bolezni in posledično primerno prilagodijo način zdravljenja. Pacienti so tako deležni bolj kvalitetnega zdravljenja, kar je še predvsem pomembno pri raznih kompleksnih boleznih. MRI slikanje pa kljub temu, da je izjemno orodje, ni enostavno za uporabljati oziroma ni enostavno pridobiti kvalitetnih podatkov. K temu pripomore več različnih faktorjev, tako človeških kot strojnih.

1.1. Problemi MRI slik

Čas priprave na slikanje je dolg, saj mora biti naprava kalibrirana na pacienta. Čas slikanja je dolg, saj potrebuje naprava dovolj časa, da zajame toliko informacij, da jih lahko zdravniki natančno ocenijo. Naprava narediti mnogo slik (t.i. *slices*) iz različnih smeri. Vse te slike se na koncu združijo v smiselno celoto (t.i. *volume*). Med slikanjem mora biti pacient na miru, saj premikanje vstavi v slike nezaželen šum in razne artefakte. Čas pridobivanja slike je tudi sorazmeren s končno kvaliteto slik (manjši čas pridobivanja, manjša ločljivost). Med drugim so MRI slikanja tudi zelo draga za zdravstvene klinike kot posledica oskrbe naprave. Posledično morajo zato tudi pacienti plačati več. Tu pripomorejo t.i. *low-field MRI scanners*, ki so cenejši. To omogoča, da je MRI slikanje dostopno vsem, vendar so slike pridobljene s temi napravami relativno nižje resolucije.

Pri tem problemu lahko pomagajo SR (super resolution) metode, ki so zmožne rekonstruirati slike nizkih ločljivosti v slike visokih ločljivosti. Tu so aktualni tako imenovani GAN (generative adversarial networks) modeli. Z globokim učenjem so zmožni rekonstruirati slike v večjih resolucij z visoko natančnostjo. Razvitih je bilo več GAN modelov katerih namen je super resolucija slik, npr. SRGAN, ESRGAN, Real-ESRGAN, itd. Med njimi je tudi SRResCycGAN oziroma *Super Resolution Residual Cycle-consistent GAN*. [1]

1.2. 0 projektu

Za ta projekt sem si zadal implementirati SRResCycGAN model v jeziku Python z uporabo knjižnice TensorFlow. Za vso programsko kodo je bila uporabljena različica 2.15. Programska koda je dostopna na [github](#). Model je bil treniran na [FastMRI](#) podatkih, specifično na t.i. *singlecoil* slikah kolen.

2. Predstavitev GAN strukture

GAN modeli so v osnovi sestavljeni iz dveh nevronske mreže - generator in diskriminator, ki med sabo tekmujeta. Cilj **generatorja** je, da se nauči porazdelitev podatkov, tako da bo sposoben generirati resnične slike. Ker sam po sebi ne more prepoznati kdaj so njegove slike resnične je tu potreben **diskriminator**. Njegova naloga je, da se nauči razlikovati med generiranimi in resničnimi slikami.

Z drugimi besedami, cilj generatorja je, da za nek vzorec LR (low resolution) slik ustvari neresnične HR slike, ki bodo prepričale (preslepile) diskriminatorja v to, da jih oceni kot resnične. Po drugi strani pa je diskriminatorjev cilj, da pravilno oceni resničnost slike. Njegove ocene so potrebne za učenje obeh modelov skozi *backpropagation*, kjer bo diskriminatorja funkcija izgube kaznovala ob napačnih ocenah, generatorja pa ko bodo njegove slike ocenjene (pravilno) kot neresnične.

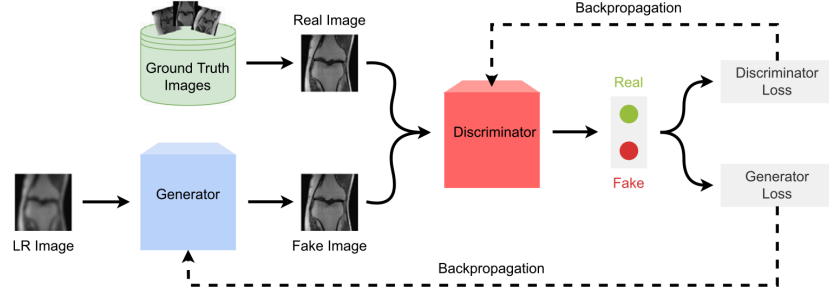


Fig. 2. Main concept behind GANs.

2.1. Funkcija izgube

Standardna funkcija izgube, ki jo uporabljajo GAN modeli se imenuje **adversarial loss**. Deluje na *min-max* principu, saj poiskuje generator vrednosti funkcije čimbolj zmanjšati, diskriminator pa zvečati.

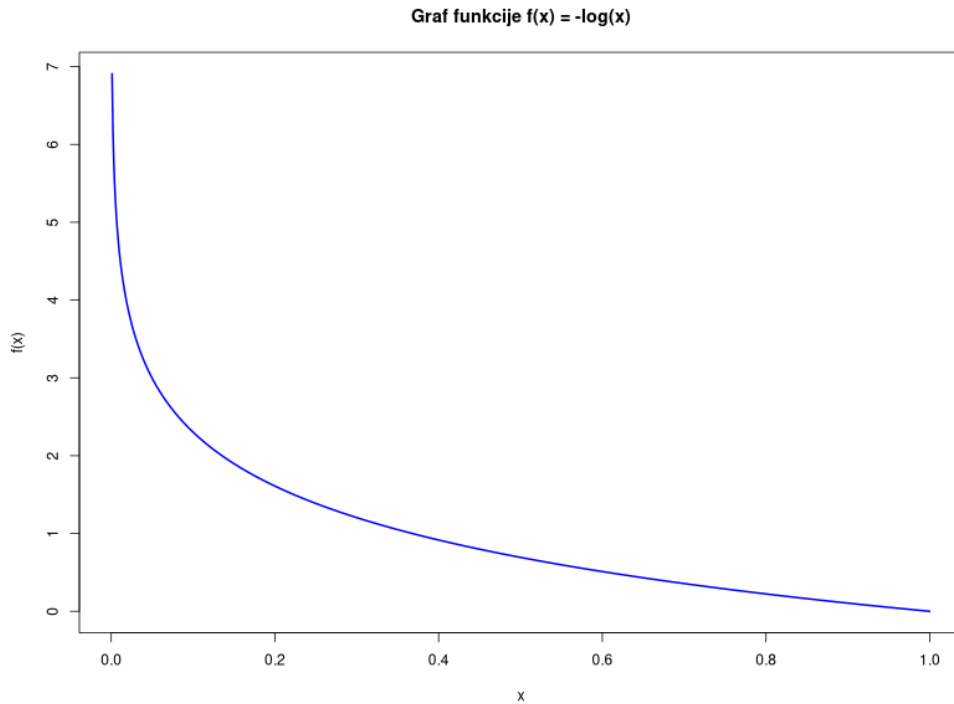
$$\min_{\theta_G} \max_{\theta_D} \mathbb{E}_{x_r} [\log(D_{\theta_D}(x_r))] + \mathbb{E}_y [\log(1 - D_{\theta_D}(G_{\theta_G}(y)))] \quad (1)$$

- x_r zaznamuje resnično sliko in $G_{\theta_G}(y)$ generirano neresnično HR sliko z vhodno LR sliko y . \mathbb{E}_{x_r} je pričakovana vrednost od vseh resničnih slik in $D_{\theta_D}(x_r)$ diskriminatorjeva ocena verjetnosti, da je vhod x_r resničen.
- \mathbb{E}_y je pričakovana vrednost vseh vhodnih LR slik y in posledično pričakovana vrednost od vseh generiranih slik $G_{\theta_G}(y)$.
- $D_{\theta_D}(G_{\theta_G}(y))$ je diskriminatorjeva ocena verjetnosti, da je generirana slika resnična.
- θ_G in θ_D predstavljata uteži in biase generatorja G in diskriminatorja D . Oba modela sta skupaj optimizirana z zgornjo funkcijo.

Taka kot je deluje kot funkcija izgube za diskriminatorja, medtem ko generatorja zanima samo ocena diskriminatorja na njegovih neresničnih podatkih, zato je levi člen med učenjem generatorja opuščen. Njegova funkcija izgube je torej

$$L_G = \frac{1}{N} \sum_{i=1}^N -\log(D_{\theta_D}(G_{\theta_G}(y_i))) \quad (2)$$

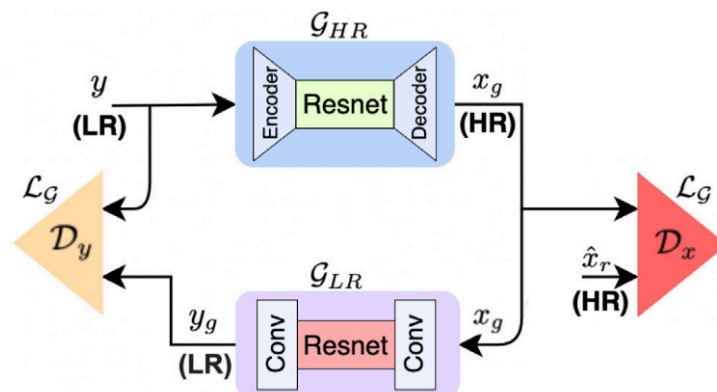
kjer je N število LR vzorcev za učenje (en batch) in y_i vhodna LR slika. Bolj kot je ocena diskriminatorja blizu 1, bolj je vrednost logaritma (ocena verjetnosti) blizu 0.



3. SRResCycGAN

Večino SR GAN metod uporablja parne podatke LR in HR slik, kjer so LR slike pridobljene z bikubično interpolacijo HR slike. To prisili modele v to, da se naučijo izničiti rezultat tega procesa, kar pa ne odraža realnega sveta - **v realnem svetu je degradacija odvisna od veliko različnih faktorjev**, ki na veliko načinov degradirajo sliko (šum, artefakti zaradi kompresije, zamegljenost zaradi premikanja, napake v lečah, itd.). Modeli zato postanejo zelo dobri v obračanju npr. bikubične interpolacije, a se ne obnese dobro na pravih *umazanih* slikah.

SRResCycGAN je poiskovala rešiti ta problem tako, da se ne uči samo super resolucije (LR \rightarrow HR), ampak tudi realistično degradacijo (HR \rightarrow LR). Zaradi tega je njegova struktura bolj kompleksna - namesto dveh nevronske mreže ima štiri - dva generatorja in dva diskriminatorja.



(a) Global structure, adapted from Umer and Micheloni (2020).

Figure 1: Struktura SRResCycGAN

Par G_{HR} , D_{HR} skrbi za učenje super resolucije, medtem ko par G_{LR} , D_{LR} skrbi za učenje degradacije. [2]

3.1. CycleGAN

Struktura je osnovana, med drugimi, na CycleGAN modelu, ki se poiskuje naučiti preslikave med vhodno in izhodno sliko v primerih, ko pravi pari slik za treniranje niso na voljo, kar je primer tudi pri npr. *style transfer* metodah. Cilj modela je se naučiti preslikave G med domeno umazanih LR slik X in domeno čistih slik visoke resolucije Y , kar zapišemo kot $G: X \rightarrow Y$, tako da je porazdelitev generiranih slik $G(X)$ nerazločljiva od porazdelitve Y z uporabo adversarialne izgube. Ker je preslikava sama po sebi slabo omejena in obstaja neskončno možnih preslikav, je združena z inverzno preslikavo $F: Y \rightarrow X$.

3.1.1. Cycle consistency loss

Uvedena je tudi nova funkcija izgube, ki se imenuje **cycle consistency loss**, ki zagotavlja, da je $F(G(X)) \approx X$ in obratno. Intuitivno si lahko predstavljamo to kot “prevod stavka iz slovenščine v angleščino mora biti identičen prevodu stavka nazaj v slovenščino”. Za $\forall x \in X$ in $\forall y \in Y$ so cikel preslikav $x \rightarrow G(x) \rightarrow F(G(x)) \approx x$ poimenovali **forward cycle consistency** ter $y \rightarrow F(y) \rightarrow G(F(y)) \approx y$ **backward cycle consistency**. Skupaj tvorita omenjeno funkcijo izgube:

$$L_{\text{cyc}}(G, F) = \mathbb{E}_x[\|F(G(x)) - x\|_1] + \mathbb{E}_y[\|G(F(y)) - y\|_1] \quad (3)$$

V sklopu strukture SRResCycGAN (glej Figure 1) vzame model G_{HR} za vhod LR sliko in vrne generirano HR sliko, medtem ko G_{LR} vzame HR sliko in vrne generirano LR sliko. G_{HR} nadzoruje D_{HR} , ki ocenjuje kako resnična je vhodna HR slika (bodisi prava ali generirana), G_{LR} pa D_{LR} , ki ocenjuje kako resnična je vhodna LR slika. [3]

3.2. HR Generator

Struktura generatorja HR slik je prevzeta iz modela SRResCGAN. Njegova struktura je prikazana spodaj (glej Figure 2). Deluje na principu “povečaj \rightarrow izboljšaj”, kar pomeni, da v začetnem delu poveča resolucijo vhodne LR slike na željeno ločljivost, skozi vmesni **residual network** (ali ResNet) se izpostavijo napake v sliki (šum, artefakti), v zadnjem delu pa se te napake od povečane slike odstranijo.

3.2.1. Residual Network

ResNet je arhitektura, ki poiskuje rešiti težave učenja globokih nevronske mreže, npr. problem izginjajočih gradientov (vanishing gradients) in manjšanje natančnosti pri večjem številu slojev. To doseže preko t.i. **globokega residualnega učenja** - namesto, da bi vsak sloj aproksimiral preslikavo izhoda prejšnjega sloja v naslednjega, ResNet uči te sloje, da se prilagodijo drugi, **residualni preslikavi**.

Formalno to pomeni, če je $H(x)$ zelena preslikava, potem ResNet uči svoje sloje, da aproksimirajo preslikavo $F(x) := H(x) - x$, $H(x) = F(x) + x$. Avtorji so postavili hipotezo, da je **lažje optimizirati residualno preslikavo** kot originalno. To si lahko predstavljamo, da če bi bila v skrajnem primeru optimalna rešitev identična preslikava (t.j. $H(x) = x$) naj bi bilo lažje

potisniti $F(x)$ proti θ , kot pa se naučiti identične preslikave preko zaporednih slojev. V praksi je formulacija $H(x) - x$ realizirana preko t.i. **preskočnih povezav (skip connections)**. [4]

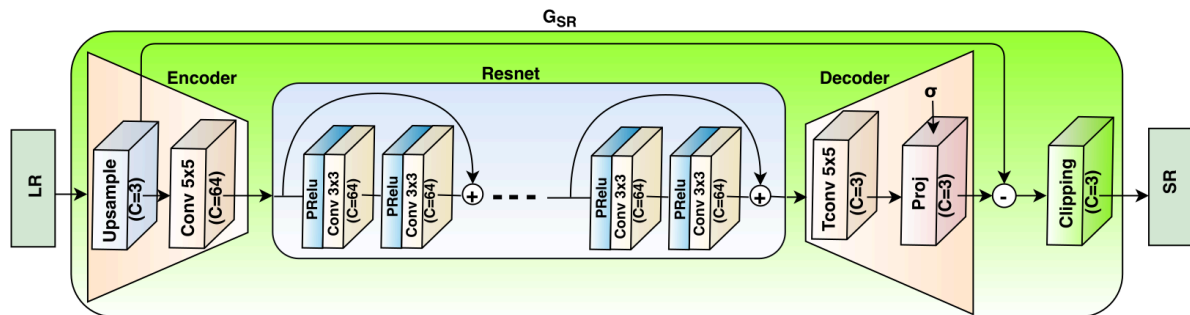


Figure 2: HR generator model

3.3. Residual GAN

4. Predstavitev podatkov

5. Podrobnosti implementacije

6. Rezultati

7. Zaključek

Bibliography

- [1] J. Guerreiro, P. Tomás, N. Garcia, and H. Aidos, “Super-resolution of magnetic resonance images using Generative Adversarial Networks,” March 15, 2023. doi: <https://doi.org/10.1016/j.compmedimag.2023.102280>.
- [2] R. Muhhamad Umer and C. Micheloni, “Deep Cyclic Generative Adversarial Residual Convolutional Networks for Real Image Super-Resolution,” September 7, 2020. doi: <https://doi.org/10.48550/arXiv.2009.03693>.
- [3] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, “Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks,” August 24, 2020. doi: <https://doi.org/10.48550/arXiv.1703.10593>.
- [4] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” December 10, 2015. doi: <https://doi.org/10.48550/arXiv.1512.03385>.