

Odpravljanje šuma v MR slikah z avtoenkoderji

Irinej Papuga, mentor: Žiga Špiclin

Univerza v Ljubljani, Fakulteta za elektrotehniko, Tržaška 25, 1000 Ljubljana, Slovenija
E-pošta: irinej.papuga@gmail.com

Povzetek. V seminarskem delu je predstavljen način zmanjševanja šuma v magnetno resonančnih slikah z uporabo avtoenkoderjev. V vsaki zajeti magnetno resonančni sliki je prisoten šum, ki omejuje možnosti avtomatske razgradnje slik. Pregledali smo klasične postopke odprave šuma iz magnetno resonančnih slik in razvili globoko nevronske mreže z namenom odpravljanja šuma v slikah. Iz slik magnetno resonančnih naprav treh različnih proizvajalcev smo ustvarili učno in testno množico ter generirali model avtoenkoderja. Autoenkoder smo učili za dva različna testna primera. Pri 1. Primeru smo učili autoenkoder le na slikah ene naprave, v 2. Pa na slikah dveh naprav. Slike tretje naprave so bile kontrolna skupina. Na podlagi metrik smo pokazali, da je pri učenju avtoenkoderja bistvena velikost in raznolikost učne množice. Za klinično uporabo avtoenkoderjev je bistveno, da smo gotovi v anatomsko pravilne rekonstrukcije magnetno resonančnih slik.

Ključne besede: magnetna resonanca, šum, globoke nevronske mreže, avtoenkoder, ocenjevanje kakovosti

Denoising MR images by autoencoders

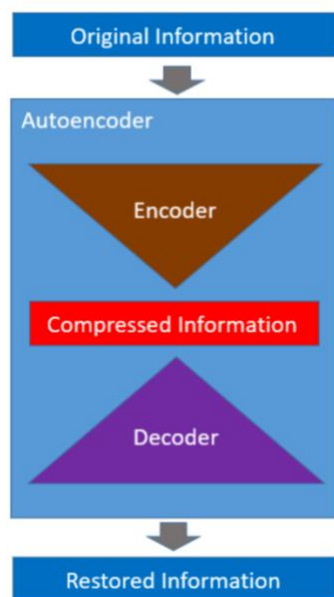
In this seminar paper we presented a method of magnetic resonance image denoising by the use of deep neural networks called autoencoders. In each magnetic resonance image we can't avoid some noise. We presented current methods for image denoising and developed deep neural network topology to allow us to denoise magnetic resonance images. From magnetic resonance images from devices of three different developers we developed training and test sets for autoencoder. Autoencoder training took place in two different ways. In 1st we used only images from one device to train autoencoder, while in 2nd we use images from two different devices. Images from 3rd device were control group. By the use of different quality assessment metrics we showed importance of training set size and diversity in training autoencoders. For clinical use of such denoising it is critical, that we are certain in out reconstruction of denoised magnetic resonance images.

Keywords: magnetic resonance, noise, deep neural networks, autoencoder, quality assesment

1 UVOD

Magnetno resonančno slikanje je metoda zajema medicinskih slik, ki ima široko področje uporabe v diagnostiki in raziskavah. Z razvojem tehnologije se je tudi kvaliteta magnetno-resonančnih (MR) slik izboljševala. Vendar pa se kljub napredku ne moremo izogniti določeni stopnji šuma, ki nastane pri zajemu. Stopnja šuma, ki se pojavi v MR slikah je odvisna od proizvajalca naprave, nastavitve protokola zajema, modela same naprave ipd.. Šum v zajetih MR slikah pa lahko otežuje avtomatsko razgradnjo slik in s tem podaljšuje diagnosticiranje bolnikovega stanja. S tem so

se razvili različni pristopi odpravljanja šuma v MR slikah. S grobem jih delimo na filtrirne, transformacijske in statistične [1]. Pri prvi kategoriji se lotimo odpravljanja šuma z uprabo različnih filtrov, kot so filtri nelokalnega povprečenja [2], filtriranje z gradientno anizotropno difuzijo in uporaba filtriranja v različnih časovnih in prostorskih domenah. Drugi pristop vsebuje uporabo transformacij kot je valčna transformacija, kjer preslikamo celotno sliko v drugo matematično domeno v kateri lažje ločimo šum od signala v sliki. Kljub svoji široki uporabnosti pa ima valčna preslikava tudi svoje slabosti. Največkrat se težave pojavijo, če so na sliki ostri robovi, ali pa če robove lahko opišemo s krivuljami. Takrat z matematičnimi orodji dopolnimo valčno preslikavo za bolj učinkovito odpravo šuma. Tretja kategorija odprave šuma v MR slikah pa je statističen pristop, kjer se osredotočamo na oceno vrste šuma v sliki. To je pomembno tako za ovrednotenje MR naprave, kot tudi za odpravo šuma in kasnejšo razgradnjo slike. Za ovrednotenje šuma uporabljamo oceno največje verjetnosti, srednjo kvadratično napako, ocena fazne napake in neparametrično vrednotenje. Z razvojem konvolucijskih nevronske mreže in njihovo vse bolj pogosto uporabo pa se ponuja še dodaten pristop k zmanjšanju šuma v MR slikah. Šum lahko zmanjšamo tako, da sliko rekonstruiramo z nevronske mreže imenovano avtoenkoder [3]. Mreža opravlja funkcijo kodiranja oz. zgoščevanja podatkov in rekonstrukcijo teh podatkov oz. dekodiranje kot je prikazano na sliki 1.



Slika 1: Shema avtoenkoderja

Kljub preprosti shemi nevronske mreže pa moramo avtoenkoder prilagoditi specifičnemu problemu in primerno določiti topologijo omrežja.

Namen seminarja je implementacija nevronske mreže za odpravo šuma v MR slikah in vrednotenje izboljšanja kvalitete slik. Na podlagi prejšnjih del {dva seminarja od lani} smo se tudi pozanimali kakšne so obstoječe metode ocene kakovosti MR slik in izbrali metrike za vrednotenje MR slik. Zanimala nas je tudi primerjava med uspešnostjo odprave šuma na slikah, ki so bile zajete na različnih napravah, in kako je potrebno trenirati nevronske mreže za čimboljše rezultate.

2 MATERIALI IN METODE

Raziskovali smo kakšna topologija avtoenkoderja je najbolj primerna za odpravo šuma v MR slikah. Postopek uporabe globokih nevronske mreže za zmanjšanje šuma v MR slikah je opisan v članku [4]. Glede na vira [5] in [6] pa smo definirali avtoenkoder v programskem jeziku Python na način, da se je lahko mreža v čim krajšem času naučila rekonstruirati MR slike z manjšo vsebnostjo šuma kot vhodne slike. Zaradi hitrejšega učenja mreže smo se odločili delati v okolju Google Colaboratory, ki je bilo povezano z osebnih Google Drive-om.

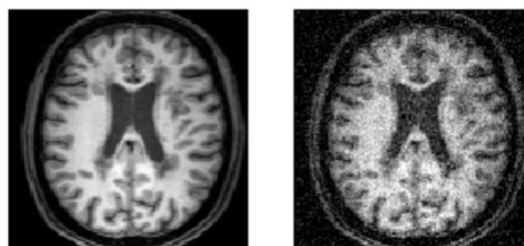
2.1 Podatki

Podatki, ki smo jih uporabili so bil 3D MR slike zajete na treh različnih napravah. Naprave so bile od proizvajalcev General Electric, Philips in Siemens. Vse slike smo naložili na Google Drive, do njih pa smo dostopali iz Jupyter skripte. Vse slike je bilo potrebno najprej pregledati in standardizirati 2D rezino na kateri bomo učili nevronske mreže. Zaradi biološke

variabilnosti in različnih dimenzij slik iz posameznih naprav so bili vsi indeksi 2D rezine za posameznega pacienta izbrani ročno. Kljub temu, da smo imeli na voljo več slik kot smo jih potrebovali smo od vsakega proizvajalca uvozili slike 27 pacientov. Vse slike so zajete v T1 modaliteti in je že na prvi pogled vidno, da prihajajo iz različnih naprav. Po tem, ko smo iz posamezne 3D slike izvozili 2D rezino smo določili učno in validacijsko množico. Odločili smo se za t.i. križno testiranje zaradi majhnega števila vzorcev. Pri križnem testiranju najprej učimo mrežo na 2/3 celotne množice, validiramo pa na preostali tretjini. Postopek še dvakrat ponovimo, pri čemer prerazporedimo slike, tako da so v učni množici vedno druge slike. S tem postopkom testiramo nevronske mreže na celotni množici slik. Poleg križnega testiranja smo opazovali tudi kako nevronska mreža naučena na slikah iz ene naprave obdela slike iz druge naprave. S tem namenom smo zastavili 2 testa. Pri prvem smo učili avtoenkoder samo na slikah iz naprave General Electric, testirali pa smo na slikah iz vseh treh naprav. Pri drugem testu smo avtoenkoder učili hkrati na slikah iz naprav General Electric in Philips. Mrežo smo prav tako testirali na slikah iz vseh treh naprav. V našem primeru 27 slik za vsako napravo je učna množica pri 1. testu sestavljena iz 18 slik, testna pa iz 9 slik. Pri drugem testu kjer za učenje uporabljamo slike dveh naprav pa je učna množica velika 36 slik, testna pa še vedno 9. Učenje avtoenkoderja še na slikah iz Siemens naprave se ni zdelo smiselno, saj nam ta množica predstavlja proizvajalca naprav za katerega nevronska mreža ni bila naučena. Tako lahko bolj točno preverimo kako se nevronska mreža odziva na slike za katere ni bila naučena.

2.2 Predprocesiranje

Slike smo pred vnašanjem v nevronske mreže morali obrezati in prevzorčiti na enako velikost. Nato pa smo morali še normalizirati sivinske vrednosti, na interval med 0 in 1. Vse slike smo obrezali na velikost 128x128 pri čemur smo si pomagali s knjižnico SimpleITK. Normalizirane in prevzorčene MR slike smo razdelili v učne in testne množice, nato pa smo slikam, ki predstavljajo vhod v nevronske mreže dodali šum normalne porazdelitve. Šum smo ravno tako standardizirali, amplitudo pa smo določili na 10% dinamičnega območja slik. Primerjava med zašumljeno in originalno MR sliko je prikazana na sliki 2.



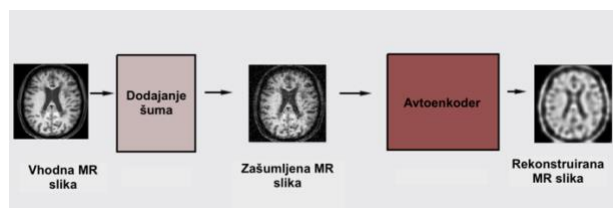
Slika 2: Normalizirana slika (levo) in zašumljena slika (desno)

2.3 Nevronska mreža

Sprva smo poskušali naučiti mrežo s preprostim avtoenkoderjem vendar nismo bili uspešni pri rekonstrukciji MR slik. Zato smo se odločili za dodajanje konvolucijskih slojev v mrežo preprostega avtoenkoderja. Mreža je tako sestavljena iz 3 konvolucijskih slojev, sledijo pa jim 2 dekonvolucijska sloja. Prvi konvolucijski nivo ima 64 filtrov, druga dva pa 32. Prvi dekonvolucijski nivo je sestavljen iz 64 filtrov drugi pa iz 1 filtra. Zadnji dekonvolucijski nivo je hkrati tudi izhod iz nevronske mreže. Po vsaki konvoluciji razpolovimo velikost vhodne slike v konvolucijski sloj. Pri tem uporabljamo metodo dodajanja ničel, tako da lažje predvidevamo dimenzije podatkov v posameznem sloju. Velikost konvolucijskega jedra je dimenzij 3x3. Med dekonvolucijskimi sloji podatkom večamo dimenzijo, tako da je izhod avtoenkoderja enakih dimenzij kot vhod. Celotno število učnih parametrov v nevronske mreži je enako 47425.

Konvolucijskim slojem smo poskušali dodati še preoblikovanje slike v vektor, izvesti kodiranje in dekodiranje na vektorju, ter ponovno rekonstruirati sliko iz vektorja. Ta pristop se tudi ni izkazal za ustreznega saj je bilo število učnih parametrov preveliko, učinek pri odpravljanju šuma pa je bil zanemarljiv, zato smo ostali pri zgoraj opisanem konvolucijskem modelu avtoenkoderja.

Za učenje smo izbrali 'Adam' optimizator, izgube pa smo pračunali po principu binarne križne entropije. Za vsak test smo naučili 3 nevronske mreže za namene križnega testiranja avtoenkoderja. Učenje posamezne nevronske mreže je trajalo 200 iteracij, izgube pa so pri tem bile reda ~0,3. Na sliki 3 je prikazana shema eksperimenta.



Slika 3: Shema poteka eksperimenta

2.4 Ocena kakovosti slik

Za oceno kakovosti MR slik je znanih veliko metrik, ki so podrobneje opisane v člankih [7] in [8]. Kakovost rekonstruiranih slik smo ocenjevali z dvema metrikama, ki sta vgrajeni v TensorFlow knjižnico. In sicer smo uporabljali t.i. Peak Signal to Noise Ratio ali PSNR in Structural Similarity Index ali SSIM. Prva metrika PSNR nam pove kakšno je največje razmerje med signalom in šumom. Pri obdelavi slik je razmerje signal-šum definirano z enačbo

$$SNR = \frac{\mu}{\sigma},$$

kjer je μ srednja vrednost, σ pa standardni odklon slikovnih vrednosti. V našem primeru opazujemo kakšno je največje razmerje med signalom in šumom med vhodno sliko in izhodom iz avtoenkoderja.

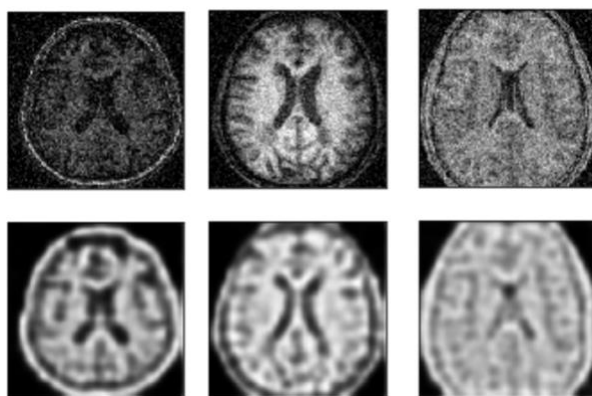
Druga metrika SSIM definira mero podobnosti med dvema slikama po enačbi

$$SSIM = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)},$$

pri čemer sta μ_x in μ_y srednji vrednosti slik x in y, σ_x in σ_y standardna odklona slik x in y, σ_{xy} pa je kovarianca med dvema slikama. Prosta člena c_1 in c_2 sta odvisna od dinamičnega območja slik in sta namenjena stabilizaciji enačbe pri manjhnih imenovalcih.

3 REZULTATI

S poskušanjem različnih modelov avtoenkoderja smo uspešno rekonstruirali slike vseh treh različnih naprav. Vendar pa so očitne razlike med testom 1 in 2. Najboljši rezultati po indeksu SSIM in PSNR se kažejo na testni množici slik iz naprave General Electric kar je smiselno, saj smo slike iz te naprave uporabili kot učno množico pri obeh testih. Nasprotno so slike iz naprave Siemens pri testu 1 najslabše rekonstruirane kar je jasen indikator, da je za uspešno odpravo šuma potrebno učiti avtoenkoder na slikah iz iste naprave. Rekonstrukcija slik iz Philips naprave je pri prvem testu primerljiva z slikami iz naprave Siemens. Pri drugem testu kažejo slike iz naprave Philips komaj vidno izboljšanje SSIM in PSNR, kjer smo kot učno množico uporabili tudi slike te naprave. Na sliki 4 so prikazane vhodne zašumljene slike in njihova rekonstrukcija spodaj za naprave General Electric, Philips in Siemens(od leve proti desni).



Slika 4: Zašumljene in rekonstruirane slike iz različnih naprav

Iz zgornjih slik vidimo, da je stopnja šuma vidno manjša na izhodnih rekonstruiranih slikah. Vendar pa bi bilo za vse tri vrste naprav potrebno izboljšati topologijo nevronske mreže, da bi dobili relevantno rekonstrukcijo anatomije. Tudi napaka treniranja nevronske mreže, ki

jo izračunamo po principu binarne križne entropije, je reda ~0,3 kar jasno govori o tem, da je potrebno izboljšati geometrijo avtoenkoderja. S testiranjem sem ugotovil, da se napaka ne znižuje dovolj hitro in tako ne moremo rešiti problema samo s povečevanjem števila iteracij učenja. Poleg tega lahko s predolgim učenjem dobimo preveč naučeno nevronske mreže, kar pomeni, da smo modelirali tudi šum učne množice in naša mreža ni zares uporabna na slikah izven učne množice.

Za jasno opredelitev uspešnosti odpravljanja šuma iz danih MR slik smo določili povprečne vrednosti PSNR in SSIM za posamezno napravo pri testu 1 in 2. Rezultati so prikazani v tabeli 1.

	GE	Philips	Siemens
Test 1-PSNR	19,36	11,29	14,85
Test 1-SSIM	0,19	0,17	0,14
Test 2-PSNR	25,65	24,45	24,12
Test 2-SSIM	0,34	0,43	0,38

Vidimo, da se je indeks strukturne podobnosti pri 2. testu povišal za vse naprave kar lahko nakazuje na dejstvo, da moramo povečati učno množico. Konvolucijsko nevronske mreže bi lahko izboljšali z večjim številom učnih parametrov kar pa bi pomenilo daljši čas računanja in sama konstrukcija topologije omrežja presega vsebino tega seminarja.

4 DISKUSIJA

Za nevronske mreže je znano, da se bolje naučijo z velikim številom podatkov, pri čemer moramo paziti, da ne prenaučimo mreže oz. modeliramo šuma in biološke variabilnosti v mreži. Možna rešitev za izboljšanje nevronske mreže bi tako lahko bila, da povečamo učno množico in učimo z istimi slikami za različne stopnje šuma. Na podlagi metrik SSIM in PSNR smo pokazali, da ni vseeno kako definiramo učno množico in kako velika je naša množica. Potrebno bi bilo še bolj podrobno raziskati vpliv različnih vrst šuma v učni in testni množici. Namen nadaljnjih raziskav iz tega področja bi bila konstrukcija karseda splošnega avtoenkoderja za široko in hitro uporabo pri odpravi šuma v MR slikah. Pri tem je potrebno vedeti, da je točna rekonstrukcija bistvenega pomena pri klinični uporabi in je zato toliko bolj pomembno, da smo gotovi v pravilno delovanje našega avtoenkoderja.

LITERATURA

- [1] J. Mohan, V. Krishnaveni, in Y. Guo, „A survey on the magnetic resonance image denoising methods“, *Biomedical Signal Processing and Control*, let. 9, str. 56–69, jan. 2014, doi: 10.1016/j.bspc.2013.10.007.
- [2] J. Manjon, J. Carbonellcaballero, J. Lull, G. Garciamarti, L. Martimonmati, in M. Robles, „MRI denoising using Non-Local Means“, *Medical*

Image Analysis, let. 12, št. 4, str. 514–523, avg. 2008, doi: 10.1016/j.media.2008.02.004.

- [3] N. Shibuya, „How to Reduce Image Noises by Autoencoder“, *Medium*, 30-nov-2017. [Na spletu]. Dostopno na: <https://medium.com/activating-robotic-minds/how-to-reduce-image-noises-by-autoencoder-65d5e6de543>. [Pridobljeno: 13-jan-2020].
- [4] J. V. Manjón in P. Coupe, „MRI Denoising Using Deep Learning“, v *Patch-Based Techniques in Medical Imaging*, let. 11075, W. Bai, G. Sanroma, G. Wu, B. C. Munsell, Y. Zhan, in P. Coupé, Ur. Cham: Springer International Publishing, 2018, str. 12–19.
- [5] „Building Autoencoders in Keras“. [Na spletu]. Dostopno na: <https://blog.keras.io/building-autoencoders-in-keras.html>. [Pridobljeno: 13-jan-2020].
- [6] R. Vasudev, „Understanding and Calculating the number of Parameters in Convolution Neural Networks (CNNs)“, *Medium*, 11-feb-2019. [Na spletu]. Dostopno na: <https://towardsdatascience.com/understanding-and-calculating-the-number-of-parameters-in-convolution-neural-networks-cnns-fc88790d530d>. [Pridobljeno: 13-jan-2020].
- [7] L. Brodnik, „Ocenjevanje kakovosti MR slik“, Seminar iz Analize Medicinskih Slik, str. 1-4., 2017
- [8] J. Podobnik, „Ocena kakovosti magnetno resonančnih slik“, Seminar iz Analize Medicinskih Slik, str. 1-4., 2016