

# Načrtovanje in analiza globokih nevronske mreže za določanje modalitete MR slik

**Gašper Podobnik, mentor: doc. dr. Žiga Špiclin**

Univerza v Ljubljani, Fakulteta za elektrotehniko, Tržaška 25, 1000 Ljubljana, Slovenija

E-pošta: gp3354@student.uni-lj.si

**Povzetek.** Avtomatski postopki analize MR slik glave za svoje delovanje navadno potrebujejo informacijo o modaliteti vhodne MR slike. Ob dejstvu, da od razvoja MR skenerja število zajetih slik vsako leto vztrajno narašča in da se ne moremo zanašati na to, da je modaliteto mogoče razbrati iz imena datoteke MR slike ali DICOM značke, je razvoj modela za avtomatsko detekcijo modalitete MR slike smiselna rešitev tega problema. Ker gre za temeljni podatek o sliki in od njega zavisi tudi uspešnost vseh nadaljnjih analiz, sta ključnega pomena kvaliteta ter robustnost takega algoritma. V pričujočem poročilu predstavim robustni model na osnovi konvolucijskih nevronske mreže in prenesenega učenja, ki s povprečno natančnostjo 97,7 % uvrsti sliko v enega izmed razredov:  $T_1$ -w,  $T_2$ -w, *FLAIR* in *OTHER* ter analiziram področja, ki največ prispevajo h končni odločitvi modela.

**Ključne besede:** konvolucijske nevronske mreže, magnetna resonanca, razvrščanje, področja največjega pomena za odločitev nevronske mreže, robustnost

## Design and analysis of deep neural networks for classifying MR modalities

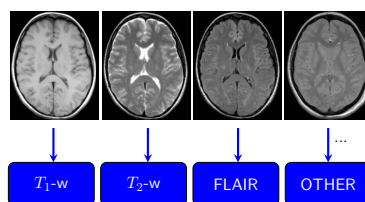
Systems for automatic analysis of MR images require information about modality of the input image. Due to the fact that there is no robust way to include information about image modality in MRI metadata, an automatic recognition of image modality is needed. I present a robust model based on convolutional neural network and transfer learning approach that achieves mean accuracy of 97,7 % on test data, classifying modalities in four categories:  $T_1$ -w,  $T_2$ -w, *FLAIR* and *OTHER*.

## 1 UVOD

Z magnetno resonanco lahko zajamemo slike, ki so si med seboj precej različne v intenzitetah posameznih tkiv ter v kontrastih med njimi. To s spreminjanjem različnih parametrov MR skenerja počno radiologi v vsakodnevni praksi z namenom izluščiti čim več informacije iz MR slike. Z najbolj pogosto uporabljenimi kombinacijami parametrov dobimo MR slike, ki jih tipično imenujemo sekvence ali modalitete MRI, na primer:  $T_1$ -w,  $T_2$ -w, *FLAIR* in druge... Ker pa nastavitve teh parametrov niso standardizirane in enake za vse skenerje (to velja celo za skenerje istega proizvajalca) pa ob zajemu MR slike ni mehanizma, ki bi v metapodatke k sliki zapisal modaliteto slike. Ker je poznavanje modalitete MR slike bazična informacija, ki je ključnega pomena za vse nadaljnje (avtomatske) postopke obdelave magnetnoresonančnih slik je nastala potreba po rešitvi tega problema.

Na prvi pogled se zdi problem enostavno rešljiv na način, da radiolog ob zajemu MR slike v ime datoteke zapiše ime modalitete. V praksi se izkaže, da se v dnevni naglici pri poimenovanju datotek pogosto pripe-

tijo napake, ki pa zaradi kasnejših postopkov avtomatske obdelave niso dopustne. Kot smiselna alternativa se zdi računalniški model, ki bi na podlagi parametrov zajema MR slike sklepal o modaliteti. Taka rešitev bi bila sicer dobra, vendar ne ponuja rešitve za razlikovanje med MR slikami pri katerih pacientu v kri vbrizgajo kontrastno sredstvo (KS) in slik brez KS, saj pri tem uporabijo povsem iste parametre.



Slika 1: Slike različnih MR slik ter pripadajoča imena modalitet. Oznako OTHER sem uporabil za modalitete, ki jih ni mogoče uvrstiti v ostale tri razrede

Zaradi naštetih dejavnikov je poleg/namesto metapodatkov (parametrov) slike smiselno analizirati dejansko slikovno informacijo MR slike. V zadnjem času so se za analizo slik uveljavile konvolucijske nevronske mreže, ki za razliko od bolj klasičnih pristopov ne zahtevajo razvoja značilnic, vendar jih v fazi učenja s pomočjo optimizacijskega algoritma določijo kar same.

Z nevronske mreže so se problema določanja modalitet MR slike lotile tudi druge raziskovalne skupine. V članku [1] so s t.i. PhiNet 3D konvolucijsko nevronske mrežo (CNN) podprobleme razlikovanja med a)  $T_1$ -w,  $T_2$ -w, *FLAIR*, b)  $T_1$ -w brez kontrasta in  $T_1$ -w

s kontrastom ter c) *FLAIR* brez kontrasta in *FLAIR* s kontrastom rešili zelo uspešno (z natančnostjo tudi nad 99 %). Nekoliko drugačen pristop so uporabili v [2], kjer so načrtali model sestavljen iz treh konvolucijskih nevronske mreže, ki vzporedno analizirajo tri različne slike: dve iste modalitete in eno sliko druge modalitete ter ob tem optimirajo parametre mreže tako, da je funkcija podobnosti med slikama istih modalitet čim večja in čim manjša pri slikah različnih modalitet. Poročajo, da je bistvena prednost tovrstne arhitekture modela v tem, da za učenje potrebujejo manj slik kot pri uporabi standardnih CNN, kar rešuje težavo pomanjkanja učnih podatkov za v praksi redko uporabljene modalitete.

Kljub relativno dobrim rezultatom iz literature, motivacija za gradnjo lastnega modela ostaja:

- naučeni modeli niso javno dostopni,
- postopki predprocesiranja so zahtevni/lahko odpo-vejo\*,
- predvsem pa ostaja vprašanje robustnosti predlaganih modelov.

Robustnost ali s tujko *splosnost generalizacije* (ang. generalization) je ena izmed ključnih preprek v razvoju modelov za analizo (biomedicinskih) slik. V primeru klasifikacije modalitete MR slike je robusten model tak, ki bo  $T_1$ -w sliko uvrstil v razred  $T_1$ -w ne glede na to a) ali je bila slika zajeta na MR skenerju v Ljubljani ali na drugem koncu Slovenije/Evropa/sveta†, b) ali je imel pacient atipično obliko glave, c) ali so na sliki vidni znaki nevrodegenerativne bolezni, so na sliki artefakti, ipd.

Bistveni prispevek mojega seminarja je a) razvoj robustnega modela, dobro validiranega na skrbno pripravljenih množicah podatkov, b) zanesljiv postopek predobdelave MR slik in c) kvalitativna analiza področij aktivacije (v nadaljevanju uporabljam kratico CAM, ang. class activation maps), ki pokažejo kateri del vhodne slike je najbolj prispeval h končni odločitvi nevronske mreže.

Nadaljevanje tega poročila je strukturirano na sledeči način: v poglavju 2 predstavim vhodne podatke, postopek predprocesiranja MR slik, strukture in parametre nevronske mreže, ki sem jih uporabil v eksperimentih in na kratko povzamem idejo o CAM. V poglavju 3 predstavim najpomembnejše rezultate in jih v poglavju 4 komentiram ter postavim v kontekst ostale literature.

## 2 MATERIALI IN METODE

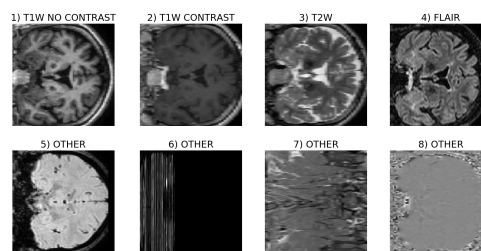
### Vhodni podatki in predobdelava

Za eksperiment sem imel na voljo 1135 slik, ki so imele vsaj 20 rezin in največjo dimenzijo voksla manjšo od 3,3 mm. Referenčne podatke za modaliteto vsake slike je deloma pripravil dr. Tim Jerman, del pa sem

\*Primer: postopek avtomatskega obrezovanja področja vratu iz MR slike glave, ki so ga uporabili v [1], je lahko problematičen

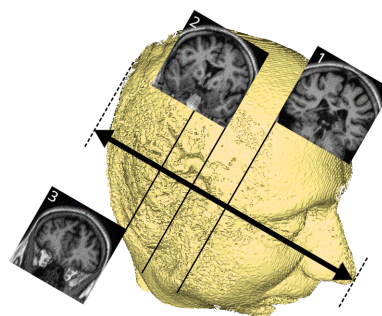
†Različni skenerji imajo namreč različno porazdelitev šuma, različne artefakte, različne jakosti magnetnega polja...

jih ročno anotiral sam. Od 1135 slik je bilo 330  $T_1$ -w brez kontrasta, 206  $T_1$ -w s kontrastom, 144  $T_2$ -w, 380 *FLAIR* in 75 *OTHER* slik. V razred *OTHER* sem uvrstil vse slike, ki ne spadajo v enega izmed prvih štirih razredov ali pa se je na njih pojavila kakšna anomalija. Primer: v skupino *OTHER* sem uvrstil slike hrbtenice (primer 7 na sliki 2) ali pa slike, ki pri katerih je bila glava razločno vidna le v aksialnem prerezu (primer 6: v koronalnem prerezu je slika povsem nerazpoznavna).



Slika 2: Primeri predobdelanih vhodnih slik: rezina nekje na sredini glave v koronalnem prerezu

Da bi iz vsake od 1135 slik, ki sem jih predhodno prevzorčil tako, da bila velikost voksla v vseh slikah (1, 1, 1) mm, dobil čim bolj podobne rezine, sem uporabil sledeči postopek: z uporabo Otsu upravljanja sem dobil masko, ki je glavo ločila od temnega ozadja kot je z rumeno barvo prikazano na sliki 3. Potem sem na vsaki sliki poiskal skrajno zadnjo točko, ki še leži na maski in skrajno sprednjo točko (v sagitalnem prerezu to pomeni skrajno levo in skrajno desno točko na maski). Iz teh dveh točk sem izračunal kje je sredinska koronalna rezina. Potem sem poleg sredinske rezine vzel še rezini za 15 % širine glave anteriorno in posteriorno od sredinske rezine kot je prikazano na sliki 3.

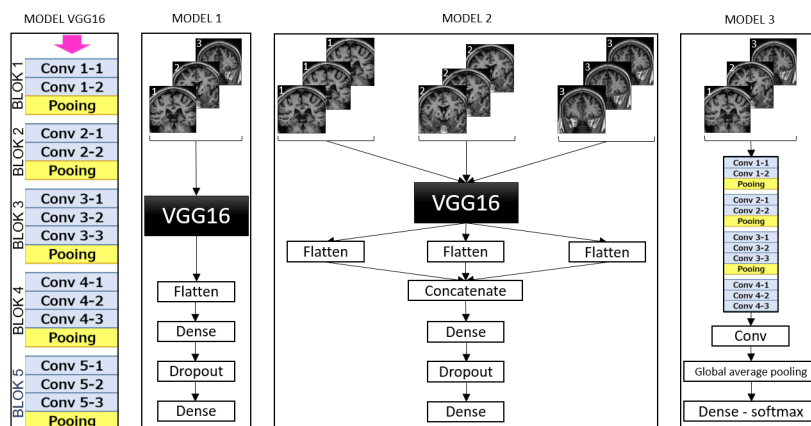


Slika 3: Shema izbire treh koronalnih rezin iz MR slike glave

Iz vsake rezine sem nato izluščil okno velikosti 128x128 in sicer na način, da sem sredino okna postavil simetrično glede na frontalni prerez in jo poravnal tako, da je bilo v oknu vidnega čim manj črnega ozadja. Če je bila slika manjše velikosti od 128x128 sem ostale pikse postavil na 0 (primer 8 na sliki 2). Na koncu sem normiral intenzitete na interval [0, 1].

### Struktura konv. nevronske mreže

Vsem trem strukturam nevronske mreže s katerimi sem



Slika 4: Strukture vseh treh modelov

eksperimentalno je skupno, da temeljijo na konv. nev. mreži imenovani VGG16 [3]. VGG16 je sestavljena iz večih konv. plasti ter t.i. max-pooling (združevalnih) nivojev in na koncu goste povezanih plasti (dense). Razlog za uporabo VGG16 je v tem, da je bila mreža že naučena na bazi slik imenovani ImageNet. Uteži v VGG16 so torej prilagojene za prepoznavanje živali, stavb, predmetov.

Uporaba prednaučenih modelov je v literaturi znana pod pojmom preneseno učenje (ang. transfer learning). Modeli nevronske mreže se v prvih plasteh tipično naučijo prepoznavati nizkonivojske vzorce (robove, kontraste,...), ki so uporabni ne glede na domeno iz katere prihajajo vhodne slike, v kasnejših plasteh pa se naučijo bolj specifičnih značilnic. Zaradi tega sem v svojih modelih za prve štiri bloke modela VGG16 (glej sliko 2) uporabil že prednaučene uteži, uteži v petem bloku in v sledečih t.i. dense plasteh pa sem v postopku učenja na novo naučil (začetne vrednosti uteži so bile enake prednaučenim).

#### Področja aktivacije nevronske mreže - CAM

Ena od bistvenih slabosti nevronske mreže je njihova interpretativnost. Eden od načinov, da dobimo vsaj vpogled v področja slike, ki največ prispevajo h končni odločitvi modela je uporaba CAM - teorija je predstavljena v [4]. To je bila tudi motivacija za uporabo modela 3, katerega struktura omogoča izris področij aktivacije CNN.

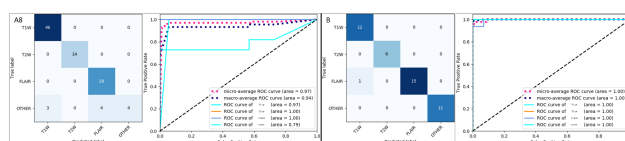
### 3 REZULTATI

V procesu razvoja modelov sem eksperimentalno s tremi strukturami nevronske mreže predstavljenimi na sliki 2, njihovimi parametri in pa z izhodnimi razredi na sledeči način: razvrščal sem a) v razrede [ $T_1$ -w,  $T_2$ -w,  $FLAIR$ ,  $OTHER$ ] ( $T_1$ -w slike brez in s KS sem združil pod oznako  $T_1$ -w) in b) v razrede [ $T_1$ -w brez KS,  $T_1$ -w KS,  $T_2$ -w,  $FLAIR$ ,  $OTHER$ ]. S poskušanjem sem ugotovil, da 20 epochov, optimizacijski algoritem Adam, učna konstanta  $10^{-4}$  in kriterijska funkcija categorical\_crossentropy delujejo dobro za

dani problem.

Ključni korak za dobro validacijo modela leži v zasnovi množic vhodnih podatkov (glej tabelo 1): v vsako od množic A1-9 sem uvrstil slike, ki so bile zajete na istem (samo enem) MR skenerju. To sem storil za skenerje, za katere sem imel vsaj 40 slik. V množico B sem uvrstil slike iz vseh ostalih 12-ih skenerjev za katere sem imel po manj kot 40 slik. Množico C sem pripravil tako, da sem iz vsake od množic A1-9 vzel po pet slik. Iz vsake slike sem nato s postopkom predobdelave izluščil po 3 rezine.

Princip učenja in testiranja je bil sledeči: vsakokrat sem model učil na slikah iz osem izmed devetih A1-9 množic in preostalo množico uporabil za testiranje modela. Za eksperimenta B in C sem model naučil na vseh slikah iz množic A1-9 ter ga nato testiral na množici B in C. Na ta način sem z eksperimenti A1-9 in B preverjal kako robusten je model, množica C pa je bila validacija ali model dobro deluje na slikah, ki jih sicer ni bilo v učni množici, so pa bile v učni množici slike drugih bolnikov zajete z istimi skenerji (s podobnimi artefakti, šumom,...).



Slika 5: Matriki razvrščanja in ROC krivulji za A8 in B

Tabela 1 in slika 5 prikazujeta rezultate za model 2, ki se je izkazal za najboljšega. Natančnost modela na učni množici (glej stolpec Train ACC) je za vse množice skoraj 100 % (vsi rezultati so zaokroženi na dve decimalni mesti). Najbolj objektivna informacija o kakovosti modela je število napačno razvrščenih slik na testni množici, ki je skoraj v vseh primerih nad 95 %. Pregled napačno klasificiranih slik za množico A4 in A8, pri katerih je natančnost najslabša (93 %), je pokazal, da večina teh slik v resnici spada v razred  $OTHER$ , kar ne preseneča, saj so s to oznako labelirane zelo raznolike

Tabela 1: Rezultati najboljšega modela (modela 2) za razvrščanje v razrede [ $T_1$ -w,  $T_2$ -w, *FLAIR*, *OTHER*]. Stolpci 2-7 prikazujejo porazdelitev slik po razredih za testne množice. ACC = natančnost, AUC = ploščina pod ROC krivuljo

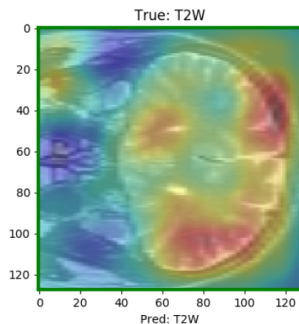
Eksp. A	Skupaj slik	$T_1$ -w	$T_1$ -w KS	$T_2$ -w	<i>FLAIR</i>	<i>OTHER</i>	Train ACC	Test ACC	AUC
A1	85	18	25	19	22	1	1,00	0,99	1,00
A2	34	14	8	3	6	3	1,00	1,00	1,00
A3	195	63	57	6	69	0	1,00	1,00	1,00
A4	220	57	21	44	84	14	1,00	0,93	0,99
A5	35	6	11	0	18	0	1,00	1,00	1,00
A6	223	97	24	26	76	0	1,00	0,98	1,00
A7	87	20	13	6	19	29	1,00	0,99	1,00
A8	99	24	22	14	29	10	1,00	0,93	0,97
A9	67	12	12	13	26	4	1,00	0,96	0,99
Povp. A							1,00	0,90	0,98
B	45	7	5	6	16	11	1,00	0,98	1,00
C	45	12	8	7	15	3	1,00	1,00	1,00
Povp. vsi							1,00	0,98	0,98

slike (od slik hrbtenice do DWI in drugih modalitet, glej sliko 2).

Podatki v stolpcu AUC pomenijo ploščino pod ROC krivuljo, za katero želimo, da je čim bližje 1. Številске vrednosti v tem stolpcu so višje od Test ACC, saj se pri izračunu AUC upošteva število slik iz posameznega razreda (ta metrika ostreje kaznuje napake v razredu, kjer je več slik), kar pa v tem primeru, ko so problematične predvsem *OTHER* slike, ki jih je malo, nima bistvenega vpliva.

Tudi modela 1 in 3 sta razvrščanje v razrede [ $T_1$ -w,  $T_2$ -w, *FLAIR*, *OTHER*] opravila precej dobro. Dejstvo, da je pri modelu 3, pri katerem je število parametrov nevronske mreže (4,723,716) bistveno manjše od ostalih dveh (model 1: 15,473,156; model 2: 13,372,164) in povprečna vrednost natančnosti na vseh testnih množicah enaka 96 %, pomeni, da bi verjetno lahko pri modelih 1 in 2 brez bistvenega poslabšanja kvalitete zmanjšal število dimenzij nekaterih plasti.

Slika 6 prikazuje primer pravilno klasificirane  $T_2$ -w slike z modelom 3 na katero je superponirana slika, ki z rdečo barvo prikazuje področja, ki so najbolj vplivala h končni odločitvi CNN.



Slika 6: Primer pravilno klasificirane  $T_2$ -w slike ter CAM

V tem primeru se lepo vidi, da so z rdečo obarvana področja bele in sive možganovine (WM in GM), ki mejijo na maščobno tkivo in likvor. To so področja na podlagi katerih sem se tudi sam pri pripravi referenčnih podatkov odločal o modaliteti. Za  $T_2$ -w slike je namreč značilno, da imata tekočina in maščoba izrazito visoke intenzitete, WM in GM pa sta temnejše barve. Ta primer

deloma služi kot potrditev, da se mreža osredotoča na smiselne regije. Pri pregledu CAM slik pri modelu, ki razvršča tudi v  $T_1$ -w brez KS in  $T_1$ -w KS sem ugotovil, da se model osredotoča na zunanji del parietalnega režnja, ki je temen na  $T_1$ -w brez KS slikah, ampak svetel zaradi žil na  $T_1$ -w KS slikah. Ker pa so na tem predelu zaradi maščobnega tkiva svetle tudi  $T_2$ -w slike, je bilo znatno število  $T_2$ -w slik napačno klasificiranih kot  $T_1$ -w KS.

Ta ugotovitev nakazuje, da je smiselno zgraditi model iz dveh delov: 1. del loči med [ $T_1$ -w,  $T_2$ -w, *FLAIR*, *OTHER*], 2. del pa med [ $T_1$ -w brez KS in  $T_1$ -w KS], kar je stvar eksperimentov, ki jih nameravam izvesti v prihodnosti.

## 4 DISKUSIJA

Predlagani model za klasifikacijo modalitet [ $T_1$ -w,  $T_2$ -w, *FLAIR*, *OTHER*] je po natančnosti na testnih slikah primerljiv z rešitvama predstavljenima v člankih [1] in [2] ter ima pomembno prednost: dobro preverjeno robustnost delovanja. Rešitev za izboljšanje natančnosti razlikovanja med [ $T_1$ -w brez KS in  $T_1$ -w KS] vidim v novem modelu, katerega vhod bi bile slike, ki bi jih obstoječi model klasificiral kot  $T_1$ -w.

CAM analiza je deloma validirala model št. 3 in omogočila zanimivo možnost interpretacije modela. V prihodnjem delu bi bilo smiselno na sistematičen način ovrednotiti CAM področja in na ta način dobiti bolj celovit vpogled v delovanje mreže.

## LITERATURA

- [1] S. Remedios, S. Roy, D. L. Pham, and J. A. Butman, "Classifying magnetic resonance image modalities with convolutional neural networks," p. 89, SPIE-Intl Soc Optical Eng, feb 2018.
- [2] S. Puch, I. Sánchez, and M. Rowe, "Few-shot Learning with Deep Triplet Networks for Brain Imaging Modality Recognition," aug 2019.
- [3] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, International Conference on Learning Representations, ICLR, 2015.
- [4] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, "Learning Deep Features for Discriminative Localization," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2016-Decem, pp. 2921–2929, 2016.