

Globoka konvolucijska nevronska mreža za kategorizacijo delnih SAR slik

Blaž Pongrac, Peter Planinšič, Danijel Šipoš, Marko Malajner, Dušan Gleich

Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko, Inštitut za avtomatiko
E-pošta: blaz.pongrac@student.um.si

Deep Convolutional Neural Networks for SAR Patch Categorization

The categorization of Synthetic Aperture Radar (SAR) patches consist of feature extraction and classification. Recently, very good results were obtained using Convolutional Neural Networks (CNN) for categorization of image patches. This paper presents deep convolutional networks for Synthetic Aperture Radar patch categorization. Several structures of deep convolutional neural networks were introduced. We have tested convolutional networks with 10 and 20 layers and analyzed recognition rate by changing SAR patch size. We have designed a custom database of SAR patches, which were cut from several spotlight TerraSAR-X products. Database consists of 6 categories with approximately 1000 samples per category. Experimental results showed, that deep Convolutional Neural Networks can achieve 84 % accuracy using patches with a size of 200×200 pixels and it performs slightly better than algorithm for categorization, which use Dual Tree Oriented Wavelet Transformation, spectral features and classification with Support Vector Machine (SVM), which achieved accuracy of only 80 % using the same training and testing sets.

Keywords – Synthetic Aperture Radar; Patch Categorization; Deep Convolutional Network

1. Uvod

Radar s sintetično odprtino (SAR) je zračna ali vesoljska platforma za zajem slik, ki jo lahko uporabljamo v vsakem vremenu, podnevi ali ponoči. Danes lahko s pomočjo vesoljskih visoko ločljivih SAR sistemov dosežemo ločljivost do 10 cm. Zato je primeren za klasifikacijo ciljev, izvidništvo, nadzor, itd. Zaradi sipanja svetlobe in šuma v SAR slikah sta interpretacija in razumevanje le-teh drugačna od analize optičnih slik oz. fotografij. Ročno iskanje majhnih ciljev je lahko v velikih SAR slikah časovno potratno in pogosto nepraktično.

Z vpeljavo teorije globokega učenja [1] za avtomatsko prepoznavanje značilk iz podatkovnih nizov z večplastnim pristopom, so aplikacije globokega učenja postale izjemno privlačne. Klasifikacija s konvencionalno nevronska mrežo in podpornim strojnim vektorjem (SVM) potrebuje proces izločitve značilk v ustreznem prostoru, medtem, ko sistem z globokim učenjem ne potrebuje specifičnega procesa za iskanje posameznih značilk in

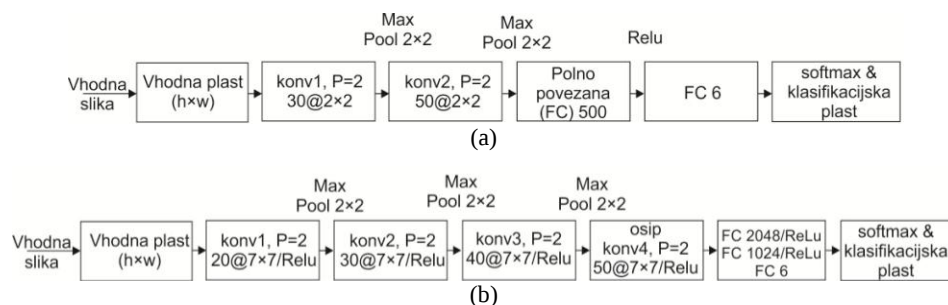
se lahko le-teh avtomatsko priuči direktno iz podatkov. Globoka nevronska mreža je bila uporabljena za daljinsko procesiranje slik [2] in klasifikacijo [3]. Konvolucijska nevronska mreža (CNN), kot ena od tipičnih metod globokega učenja, je do sedaj dosegla izjemne rezultate na več področjih [4]. Na SAR slikah le s težavo ločimo razrede, saj se na njih pojavlja kompleksno sipanje in naključni šum. V [5] je bila za namene SAR avtomatskega določevanja ciljev (ATR) uporabljena globoka konvolucijska nevronska mreža. Ta metoda je v času pisanja v primerjavi z ostalimi navedenimi metodami dosegla najvišjo natančnost (do 99 %) klasifikacije desetih razredov.

V zadnjih letih je bilo predlaganih več izvedb modelov globokih konvolucijskih nevronske mreže, npr. model AlexNet [6], model VGG-S [7], Berkeley-trained model [8], model Places-CNN [9], model polne konvolucijske semantične segmentacije [10], CNN model za ocenjevanje značilnih objektov [11], Places-CNDS modeli za prepoznavo scen [12], modeli za določitev starosti in spola [13], model GoogLeNet [14], itd. Vsak izmed omenjenih modelov vsebuje rešitev specifičnega problema, kot je npr. problem zasičenosti, problem povezav med točkami bližnjih plasti, problem doseganja zadosti velike sposobnosti (kapacitete) učenja, itd.

Pričujoči članek predstavlja uporabo globoke konvolucijske nevronske mreže za kategorizacijo delnih SAR slik. Motivacija je raziskava učinkovitosti globokih konvolucijskih nevronske mreže z 10 in 20 plastmi. Dodatno smo želeli določiti optimalno velikost SAR delne slike, saj velikost te slike vpliva na učinkovitost kategorizacije. Za namene testiranja obeh konvolucijskih nevronske mreže smo zato pripravili podatkovno bazo sestavljeno iz delnih SAR slik velikosti 200×200 , 150×150 in 100×100 slikovnih pik.

2. Konvolucijska nevronska mreža

Globoka konvolucijska nevronska mreža je večplastna nevronska mreža. Sestavljena je iz vhodne plasti, ki ji sledijo izbirne plasti, namenjene pod-vzorčenju in regularizaciji ter se končajo s popolnoma povezanimi plastmi. Vhodni podatek v globoko konvolucijsko nevronska mrežo je slika, ki jo obdelamo z več filtri. Odziv filtra pridobimo s konvolucijo. Torej konvolucija kodira vhodne podatke in določuje značilke, z uporabo pod-vzorčenja pa zmanjšujemo dimenzijo.



Slika 1. Konvolucijske nevronske mreže. (a) 10 plastna konvolucijska nevronska mreža. (b) 20 plastna konvolucijska nevronska mreža

Odzivi filtrov predstavljajo vhodne podatke aktivacijskih funkcij brez nasičenja, kakršne lahko drastično pospešijo proces učenja. To izvedemo s pomočjo usmerjenih linearnih enot (ReLU), ki so vključene v nelinearne procese nasičenja. Te funkcije so uporabljene po vsaki konvoluciji in po vsaki povezani plasti. Zadnja plast uporablja normalizirano eksponentno funkcijo “softmax” kot aktivacijsko funkcijo. S tem je multi-nominalna logistična regresija objekta maksimalna. Rezultat filtriranja je po navadi pod-vzorčen, saj s tem zagotovimo nadaljnjo reduciranje dimenzij značilk, kar te naredi invariantne translaciji. Kadar uporabimo “max-pool” diskretizacijo, ta 2×2 plast razdeli izhod v vzorce z 2×2 celici, katere se ne prekrivajo. Maksimalna vrednost aktivacijske funkcije se zapiše v vsako celico. Na ta način smo zmanjšali dimenzijo za 2, istočasno pa so značilke invariantne na translacijo objektov.

2.1. Struktura globoke konvolucijske nevronske mreže z 10 sloji

Slika (1a) prikazuje globoko konvolucijsko nevronska mrežo z 10 sloji. Gre za malenkost drugačno strukturo kot pri LaNet mreži [16], za katero je znano, da se dobro obnese pri klasifikaciji. Konvolucijska nevronska mreža ima dve konvolucijski plasti, katerima sledita dve “max-pool” plasti in dve polno povezani plasti, na izhod pa je povezana “softmax” funkcija in klasifikator. “Max-pool” plasti sta povezani za dvema konvolucijskima plastema, ki imata velikost 2×2 s korakom po 2 slikovni piki. K vsaki konvolucijski plasti je bilo dodano $P = 2$ prostorsko dodajanje ničel.

Učni in testni vzorci so v vhodni sloj potisnjeni v velikosti $h \times w$, kjer je h število stolpcev in w število vrstic vhodne slike. Po prostorskem dodajanju ničel s $P = 2$ v prvi konvolucijski plasti, podatke v velikosti $(h + 2) \times (w + 4)$ filtriramo s 30 konvolucijskimi filtri velikosti 2×2 v prvi konvolucijski plasti, kar da 30 značilk velikosti $(h + 2) \times (w + 2)$ pred “max-pool” podvzorčenjem z velikostjo 2×2 . Slednje da $(h/2 + 1) \times (w/2 + 1) \times 20$ izhodov. Izhodi iz prve “max-pool” plasti so vhodi v drugo konvolucijsko plast. Velikost filtra v drugi konvolucijski plasti je 2×2 , kar nam da 50 značilk velikosti $(h/2) \times (w/2)$, te pa postanejo

velikosti $(h/4) \times (w/4)$ po drugi “max-pool” plasti. Te izhode nato peljemo v prvo polno povezano plast s 500 enotami. Nato uporabimo ReLU nelinearnost v drugi polno povezani plasti, na izhodu katere dobimo 6 enot (število razredov). Na koncu uporabimo klasifikator s “softmax” funkcijo, s pomočjo katere določimo verjetnost vsakega razreda.

2.2. Struktura globoke konvolucijske nevronske mreže z 20 plastmi

Struktura konvolucijske nevronske mreže z 20 plastmi je prikazana na sliki (1b). Ta je sestavljena iz štirih konvolucijskih slojev in treh “max-pool” slojev z velikostjo 2×2 in korakom 2 slikovni točki, ki sledijo prvim trem konvolucijskim slojem. ReLU nelinearnost je uporabljena v vsaki skriti konvolucijski plasti. Vhodna slika je filtrirana z 20 konvolucijskimi filtri velikosti 7×7 v prvi konvolucijski plasti, česar rezultat je 20 značilk. Izhod prve max-pool plasti je vezan na vhod druge konvolucijske plasti, s konvolucijskimi filtri velikosti 7×7 , katerih rezultat je 30 značilk. Velikost konvolucijskih filtrov v tretji konvolucijski plasti je 7×7 , kar da 40 značilk, izhod četrte konvolucijske plasti, v kateri so konvolucijski filtri velikosti 7×7 , pa je 50 značilk. Pred četrtem konvolucijskim slojem je uporabljena regularizacija z izpuščanjem povezav, za vsakim konvolucijskim slojem pa je bila uporabljen še ReLU plast. Za četrtem slojem je vezana prva polno povezana plast velikosti 2048, za njo pa polno povezana plast velikosti 1024 in polno povezana plast velikosti 6. Za polno povezanimi plastema velikosti 2048 in 1024 je uporabljena ReLU plast. Na koncu te nevronske mreže sta uporabljena normalizirana eksponentna funkcija “softmax” in klasifikator. V tej nevronske mreži je uporabljena plast z izpuščenimi povezavami (osipna plast), kar vpliva na arhitekturo in preprečuje nasičenost. Ideja je, da s povezavo konvolucijskih in polno povezanih plasti med učenjem izločimo izhode skritih nevronov z verjetnostjo p . Verjetnost je bila nastavljena na $p = 0.5$. Osipna plast zmanjša dvojno prilagoditev nevrona oz. prisili nevrone, da ponudijo bolj robustni prispevek v kombinaciji z aktivnimi nevroni. Set nevronov se v vsaki epohi spreminja naključno, zasičenost pa se zmanjšuje z $1/(1-p)$, če

takšno strukturo primerjamo s strukturo brez osipane plasti.

2.3. Učenje

Nevronski mreži sta bili trenirani oz. učeni z "back-propagation algoritmom" in stohastičnim gradientnim zmanjševanjem, z naborom po 100 vzorcev. Parametri algoritma učenja so bili naslednji: momentum je bil nastavljen na 0.9, parameter upadanja uteži pa na vrednost 0.0005. Podrobnosti glede parametra upadanja uteži so predstavljeni v [6]. Spreminjanje hitrosti učenja lahko predstavimo kot: $lr (1 + \gamma k)^{-\alpha}$, kjer je lr inicializiran na 0.01, k je trenutna iteracija in kjer sta parametra izračuna stopnje učenja $\gamma = 0.00001$ in $\alpha = 0.75$. Proces inicializacije uteži je nadvse pomemben, saj lahko inicializacija vpliva na izvedbo algoritma. Inicializacija je bila avtomatsko določena glede na število vhodnih in izhodnih nevronov. Pristranskost inicializatorja je bil določena kot konstanta s privzeto vrednostjo polnjenja 0. Pri postopku nadzorovanega učenja je bila vpeljana predčasna prekinitev učenja, s čimer preprečimo zasičenost, ko prične natančnost padati, kot je to definirano v [6].

3. Eksperimentalni rezultati

3.1. Vzorci

Ustvarjena je bila podatkovna baza, sestavljena iz 10 kompleksnih podatkov z enojnim pogledom (Single Look Complex). Ti so bili pridobljeni v Spotlight načinu na TerraSAR-X satelitu. Podatki so bili pridobljeni iz različnih urbanih, ruralnih in gozdnih območij z različnimi koti incidence. Podatkovno bazo delnih SAR slik je načrtal strokovnjak, ki je zbral več delnih slik za vsak razred ali klaso, z znanim dejanskim stanjem na površini, ki je bilo potrjeno s pomočjo optičnih fotografij istega območja. Pred začetkom učenja sta obe konvolucijski nevronske mreži (10 in 20 plastna) imeli iste parametre. V vsakem eksperimentu smo uporabili 90 % vzorcev za učenje. V prejšnjih eksperimentih smo ugotovili, da velikost delnih SAR slik vpliva na natančnost prepoznav. Zaradi tega smo uporabili tri različne velikosti delnih SAR slik (200×200 , 150×150 in 100×100 slikovnih pik). Zemeljska resolucija je bila med 0.8 in 1 m.

Tabela 1. Število vzorcev v posamezni kategoriji

Kategorije	Število vzorcev
C1	1168
C2	1325
C3	1085
C4	1065
C5	1037
C6	1613

Izbrali smo 6 razredov: C01 (Urbana območja), C02 (Gozd), C03 (Travnate površine), C04 (Železniška proga), C05 (Reka) in C06 (Ceste). Število delnih slik

v vsaki kategoriji je predstavljeno v Tabeli 1. Največja kategorija je C06, ki je ima 1613 vzorcev, najmanjša pa je C05 s 1037 vzorci. V procesu učenja je bilo uporabljenih 90 % SAR delnih slik za vsako kategorijo. Preostalih 10 % je bilo uporabljenih za testiranje.

3.2. Primerjava natančnosti

Tabela 2. Povprečna natančnost kategorije v %

Kat.	200x200		150x150		100x100		200x200	
	10P	20P	10P	20P	10P	20P	SFOWT	SF
C1	82	81	74	75	69	56	82	76
C2	63	87	57	74	45	54	83	74
C3	68	86	63	72	59	57	79	82
C4	81	84	66	70	58	50	81	78
C5	73	85	68	77	46	61	79	77
C6	68	81	56	63	41	49	76	83
Skupaj	73	84	64	72	53	55	80	78

V Tabeli 2 je prikazana primerjava natančnosti nevronske mreže iz slik (1a) in (1b). Prikazane so natančnosti za različne velikosti delnih slik in število plasti v konvolucijski nevronske mreži. Obe nevronske mreži smo primerjali še s prepoznavo kategorij na osnovi dvojne drevesne valovne transformacije, kjer so za kategorizacijo bile uporabljene spektralne značilnosti in podporni strojni vektor [17]. Primerjava je bila prav tako izvedena z metodo s pomočjo statistične obdelave 2. reda, ki ji pravimo logaritmčni seštevki (komulanti) v slikovnem prostoru [18]. Iz Tabele 2 lahko sklepamo, da so razlike med uspešnostjo kategorizacije različnih metod večplastnih konvolucijskih nevronske mreže zelo majhne. Pri vseh metodah smo uporabili enako podatkovno bazo za učenje. V tabeli 2 lahko prav tako vidimo, katera velikost delnih slik ima najvišjo natančnost prepoznav razreda. Najvišja natančnost prepoznav je bila z delnimi slikami v velikosti 200×200 slikovnih pik v konvolucijski nevronske mreži z 20 plastmi. Slednja je dosegla 84, 72 in 55 % natančnost pri velikostih delnih slik 200×200 , 150×150 in 100×100 slikovnih pik. Eksperiment je pokazal, da ima 20 plastna konvolucijska nevronska mreža najboljše rezultate pri vseh treh velikostih delnih slik. Velikost selne slike 100×100 slikovnih pik je premajhna, da bi lahko z gotovostjo prepoznali zemeljske površinske značilnosti ali strukture na delni SAR sliki, saj je stopnja prepoznav 10 plastne in 20 plastne konvolucijske nevronske mreže okrog 50 %. S povečanjem velikosti delnih SAR slik povečamo tudi stopnjo prepoznav na 64 % in 72 %. Rezultati so bili ocenjeni s pomočjo SFOWT metode [17] in SF metode [18], ki sta dosegli 78 % in 80 % stopnjo prepoznav kategorij.

Tabela 3 prikazuje, koliko SAR slik je bilo klasificiranih v specifičnih razredih. Ta primerjava je zapisana s konfuzijsko matriko za 20 plastno konvolucijsko nevronske mreže in velikostjo delnih slik 200×200 slikovnih pik. Največ zamenjav je v C06 razredu (Ceste), saj se meša z vsemi ostalimi razredi, razen s C03 (včasih se cesta nahaja blizu

obale). Za C06 je algoritem neučinkovit. Zapišemo lahko vrstni red učinkovitosti: C02 (87.5 %), C03 (86 %), C05 (85,4 %), C04 (84,9 %), C01 (81,8 %) in C06 (81,6 %).

Tabela 3. Klasifikacija v specifičnih razredih

Kat	C1	C2	C3	C4	C5	C6	skupaj
C1	104	1	0	2	6	14	127
C2	1	113	3	1	2	9	129
C3	0	3	486	2	1	8	100
C4	4	1	2	90	4	5	106
C5	3	1	4	1	88	6	103
C6	9	7	7	1	6	133	163

4. Zaključek

Kategorizacija posnetkov oziroma slik radarja s sintetično odprtino in določanje značilke je pomembno za razumevanje funkcij radarja in pravilno interpretacijo vsebine. V tem članku je predstavljena raziskava večplastnih konvolucijskih nevronske mreže in vpliv velikosti delnih SAR slik na učinkovitost algoritma kategorizacije. Po navadi se ob uporabi globoke konvolucijske nevronske mreže pri procesiranju SAR slik pojavijo težave z zasičenostjo, kar je posledica premalo učnih vzorcev in vodi k slabemu rezultatu klasifikacije. Zaradi tega je bila v naš algoritem bila vstavljena osipana plast. Načrtali smo dve konvolucijski nevronske mreži, 10 in 20 plastno. Bili sta testirani z isto podatkovno bazo, ki je vsebovala delne slike velikosti 200×200 , 150×150 in 100×100 slikovnih pik. Eksperimentalni rezultati kažejo, da je algoritem učinkovitejši ob uporabi delnih SAR slik velikosti 200×200 slikovnih pik in ob uporabi 20 plastne konvolucijske nevronske mreže. Stopnja prepoznavanja pri 10 plastni mreži je bila 73 % in pri 20 plastni 84 %. Prišli smo do zaključka, da je kategorija "Ceste" vplivala na skupno učinkovitost kategorizacije globoke konvolucijske nevronske mreže, saj so testni vzorci pojavili v vseh ostalih kategorijah. Eksperimentalni rezultati kažejo na to, da je učinkovitost algoritma odvisna od SAR podatkovne baze, katere načrtovanje ni trivialno. Kategorije morajo biti skrbno izbrane. Na žalost je možno določene kategorije pomešati z drugimi, kot to kažejo tudi eksperimentalni rezultati v tem članku.

Literatura

- [1] G.E. Hinton, R.R. Salakhutdinov: Reducing the dimensionality of data with neural networks, Science, vol. 313, no. 5786, pp. 504-607, 2006
- [2] W. Diao, X. Sun, X. Zheng, F. Dou, H. Wang, K. Fu: Efficient saliency-based object detection in remote sensing images using deep belief networks, IEEE Geoscience and Remote sensing Letters, vol. 13, no. 2, pp. 137-141, 2016
- [3] F. P. S. Luus, B. P. Salmon, F. van den Bergh, B. T. J. Maharaj: Multiview deep learnign for land-use

- classification, IEEE Geoscience and Remote Sensing Letters, vol. 12, no. 12, pp. 2448-24552, 2015
- [4] X. Chen, S. Xiang, C.L. Liu, C.H. Pan: Vehicle detection in satellite image by hybrid deep convolutional neural networks, IEEE Geoscience and Remote Sensing Letters, vol. 11, no. 10, pp. 1797-1801, 2014
- [5] S. Chen, H. Wang, F. Xu, Y. Q. Jin: Target classification using the deep convolutional networks for sar images, IEEE Transactions on Geoscience and Remote Sensing, vol. 54, no. 8, pp. 4806-4817, 2016
- [6] A. Krizhevsky, I. Sutskever, G. E. Hinton: Imagenet classification with deep convolutional neural networks, Advances in neural information processing systems, pp. 1097-1105, 2012
- [7] K. Chatfield, K. Simonyan, A. Vedaldi, A. Zisserman: Return of the devil in the details: Delving deep into convolutional nets, arXiv preprint arXiv:1405.3531, 2014
- [8] P. Heit: The berkeley model, Health education, vol. 8, no. 1, pp. 2-3, 1977.
- [9] B. Zhou, A. Lapedriza, J. Xiao, A. Torralba, A. Oliva: Learning deep features for scene recognition using places database, Advances in Neural Information Processing Systems, pp. 487-495, 2014
- [10] E. Shelhamer, J. Long, T. Darrell: Fully convolutional networks for semantic segmentation, arXiv preprint arXiv:1605.06211, 2015
- [11] J. Zhang, S. Ma, M. Sameki, S. Sclaroff, M. Betke, Z. Lin, X. Shen, B. Price, R. M'ech: Salient object subitizing, in 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4045-4054, 2015
- [12] L. Wang, C.-Y. Lee, Z. Tu, S. Lazebnik: Target classification using the deep convolutional networks for sar images, arXiv preprint arXiv: 1505.02496, 2015
- [13] G. Levi, T. Hassner: Age and gender classification using convolutional neural networks, in 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 34-42, 2015
- [14] P. C. Z. Ge, C. McCooll: Content specific feature learning for finegrained plant classification, in Working notes of CLEF 2015 con, 2015
- [15] X. Glorot, A. Bordes, Y. Bengio, Deep sparse rectifier neural networks, in Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics, pp. 315-323, 2011
- [16] Y. Lecun, L. Bottou, Y. Bengio, P. Haffner: Gradient-based learning applied to document recognition, Proceedings of the IEEE, vol. 86, no. 11, pp. 2278-2324, 1998.
- [17] P. Planinsic, J. Singh, D. Gleich: Sar image categorization using parametric and nonparametric approaches within a dual tree cwt, Geoscience and Remote Sensing Letters, IEEE, vol. 11, no. 10, pp. 1757-1761, Oct 2014
- [18] J. Singh, M. Datcu: Sar image categorization with log cumulants of the fractional fourier transform coefficients, Geoscience and Remote Sensing, IEEE Transactions on, vol. PP, no. 99, pp. 1-10, 2013