

Konvolučné neurónové siete v rozpoznávaní 3D objektov

Šimon Javorský, doc. RNDr. Gabriela Andrejková, CSc.

Prírodovedecká fakulta, Univerzita Pavla Jozefa Šafárika v Košiciach

ABSTRAKT

Rozpoznávanie a klasifikácia 3D objektov je problém, pre ktorý nájdeme v literatúre viacero algoritmov pracujúcich s rôznou presnosťou a spoľahlivosťou. V práci navrhujeme a implementujeme algoritmus, ktorý používa konvolučnú neurónovú sieť. Jej úlohou je rozoznávať 3D objekty a kategorizovať ich do príslušných, vopred naučených tried. Algoritmus je testovaný a vyhodnotený na množne 3D objektov z benchmarkovej databázy objektov.

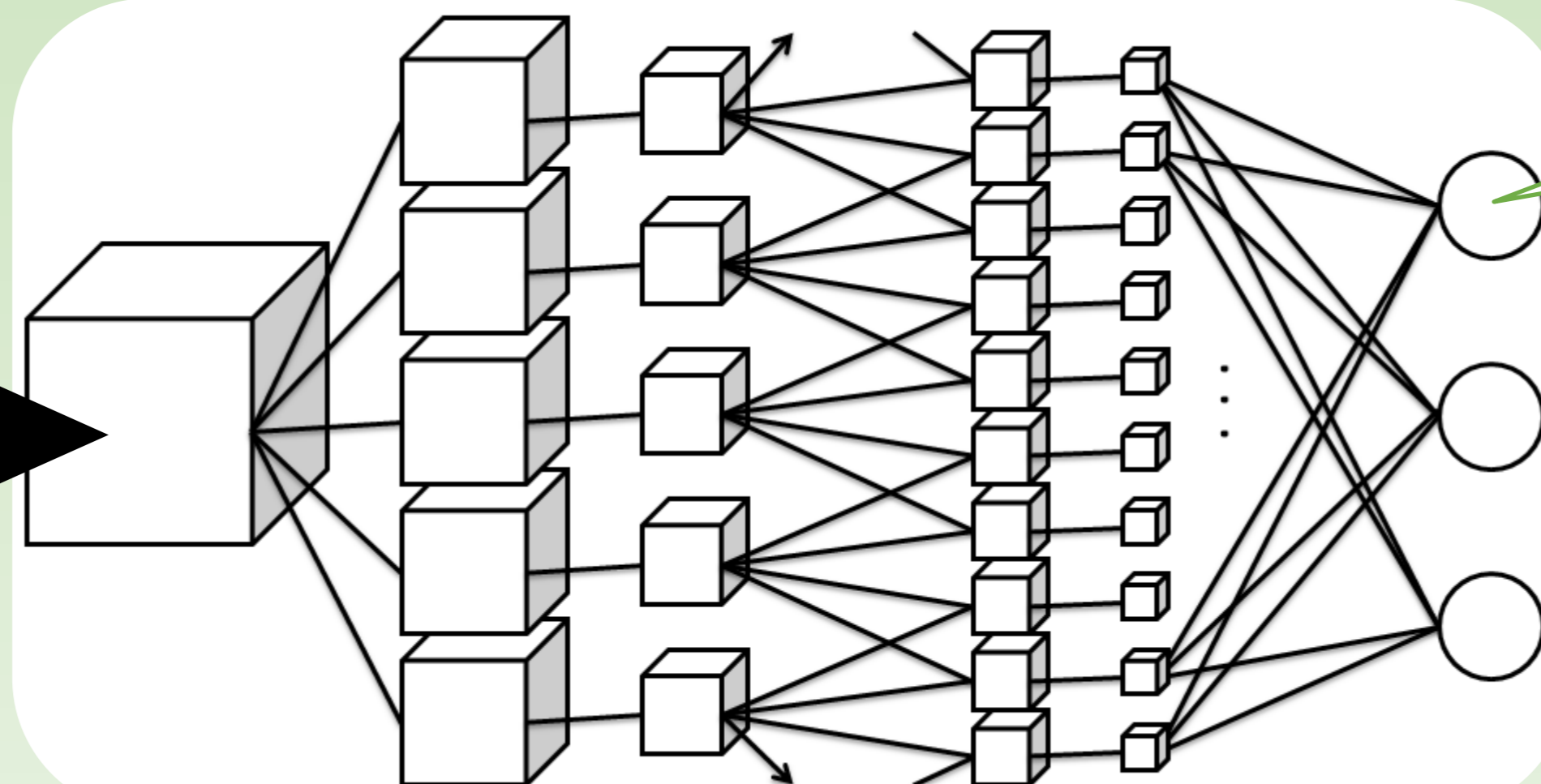
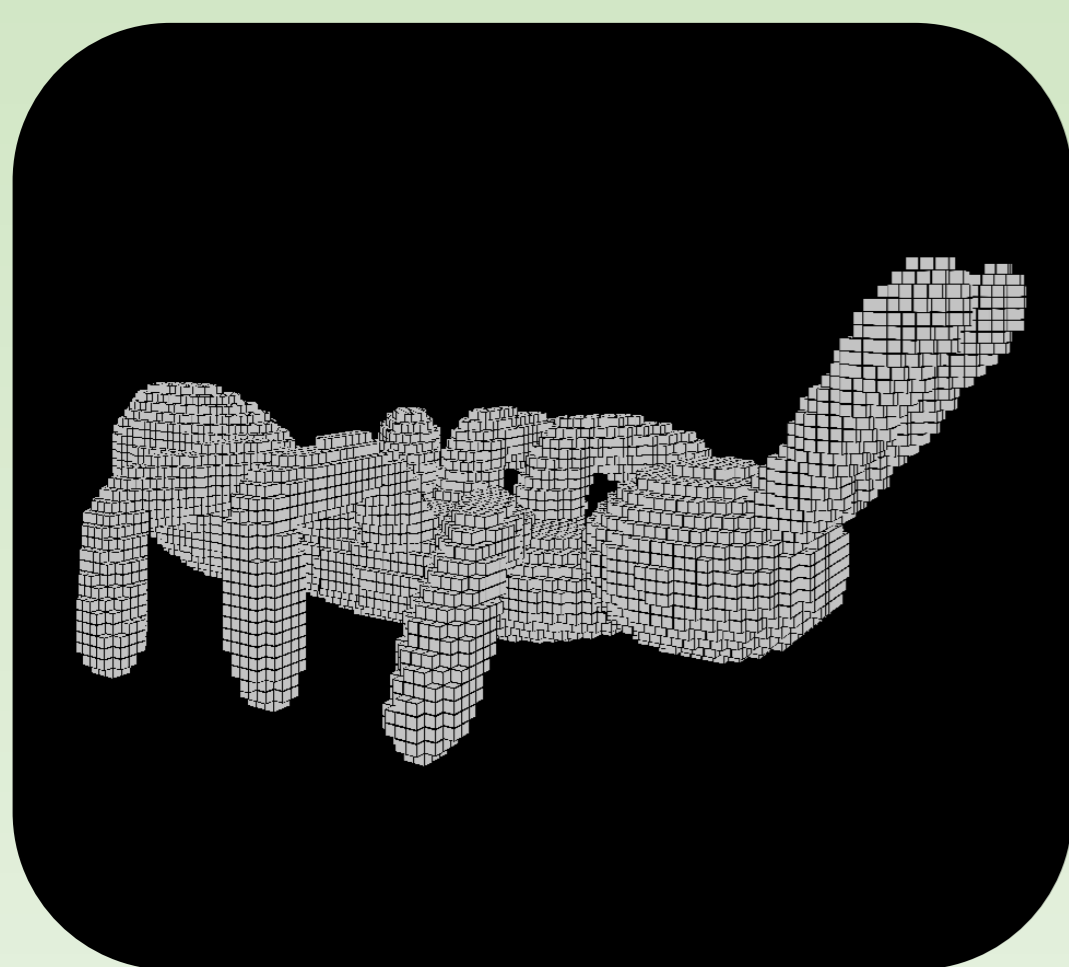
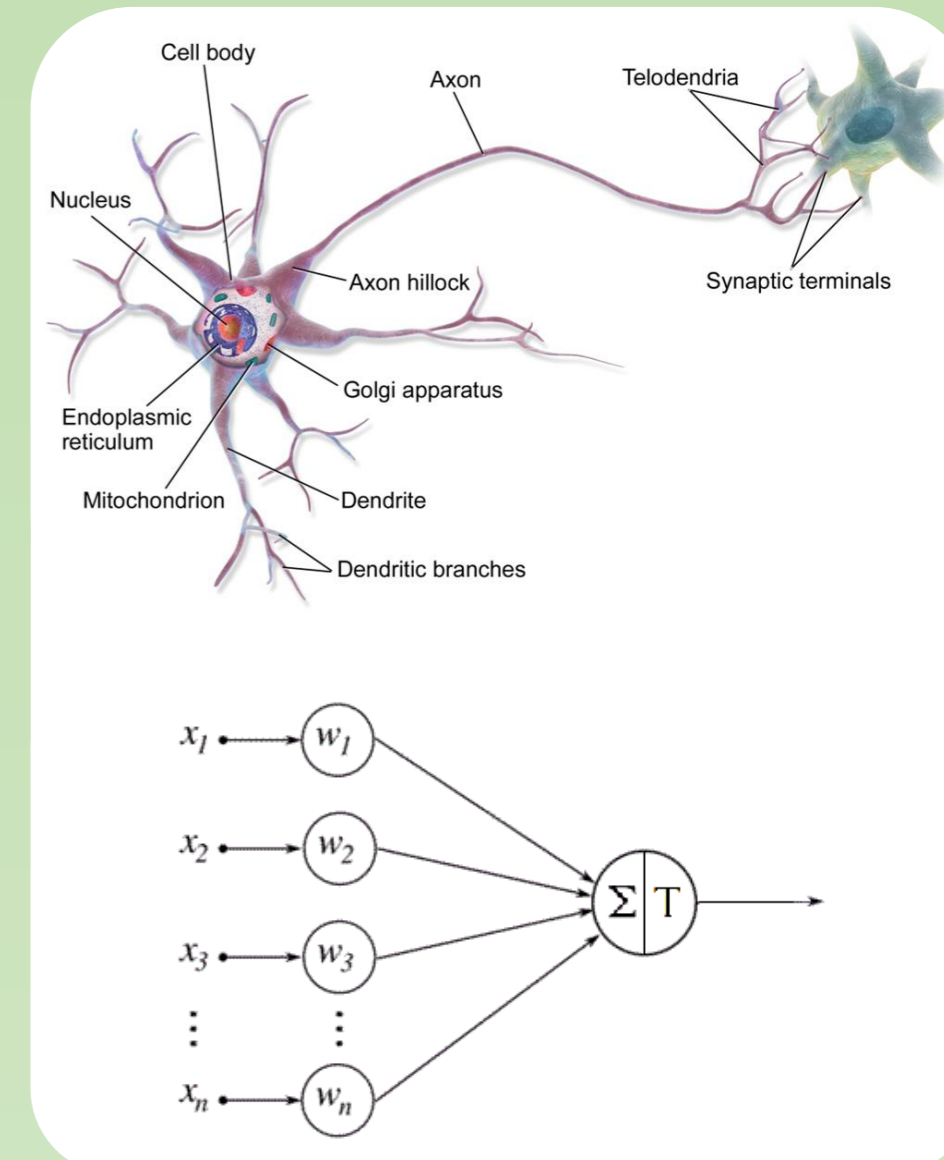
Ľudský mozog sa skladá z približne 86 miliárd neurónov. Každý jeden je prepojený s tisíckou ďalších. Stimuly z vonkajšieho prostredia, prípadne z vnútorných orgánov vytvárajú elektrické impulzy, ktoré dokážu rýchlo prejsť celou sieťou neurónov. Mozog takto dokáže spracovať vonkajšie podnety, stav v akom sa nachádza telo a vie na to adekvátne reagovať. Okrem toho, má úžasnú schopnosť naučiť sa ovládať aj vysoko komplexné procesy, akými sú napríklad udržanie tela v rovnováhe na lane, plynulá komunikácia vo viacerých cudzích jazykoch, a podobne. Niet preto divu, že sa mozog stal vzorom pri návrhu umelo inteligentných systémov.

Jedným z nich je aj neurónová sieť. Je to výpočtový model inšpirovaný biologickými neurónovými sieťami. Tak ako klasický neurón prijíma impulzy z okolia, spracováva ich a následne ich posiela ďalej, tak pracuje aj matematický model neurónu. Keď sa viacero takýchto neurónov pospája dokopy, vytvorí umelú neurónovú sieť.

Tá sa vie naučiť pravidlá medzi vstupom a výstupom, ak jej ich poskytneme, vďaka čomu je možné ju ďalej používať napríklad na klasifikáciu, aproximáciu, predikciu a podobne.

Znalosti siete sú uložené vo váhach. Tie určujú s akou silou sa pre jednotlivé neuróny berú do úvahy impulzy z okolia (viď obrázok vpravo). Z toho ale vyplýva, že z neurónovej siete nevieme priamo vyextrahovať algoritmus akým počíta, ale funguje skôr ako čierna skrinka, ktorej dáme na vstupe dáta, a na výstupe dostaneme správny výsledok s určitou presnosťou.

NEURÓNOVÁ SIEŤ



uh... pavúk?

VÝSTUP

Očakávaný výstup siete je jednotka na neuróne reprezentujúceho model posunutý siete a nuly na všetkých ostatných výstupných neurónoch.

Výhodou konvolučnej siete je schopnosť správne rozoznávať aj posunuté, prípadne inak zašumené dáta.

VSTUPNÉ DÁTA

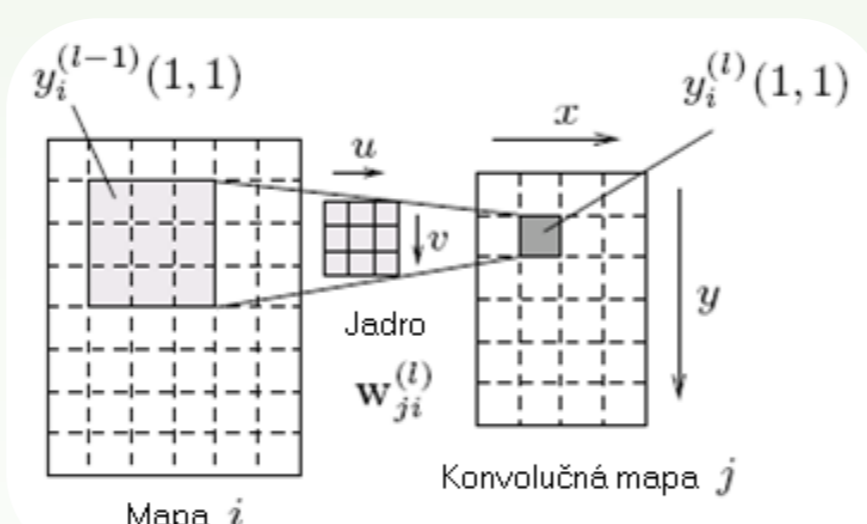
Vstupné dáta máme k dispozícii zo stránky McGillovej univerzity. Modely sa delia do dvoch skupín. Prvá skupina obsahuje modely s pohyblivými časťami, napríklad ruka ukazujúca rôzne gestá. Druhá skupina obsahuje modely bez pohyblivých častí, napríklad rôzne druhy stoličiek, šálok a podobne.

Dáta sú vo voxelovom formáte, čo znamená, že po ich načítaní zo súboru máme trojrozmernú maticu plnú núl s výnimkou miest, kde sa nachádza model, tam sú jednotky. Takéto jednotvárne dáta sa pred posunutím siete môžu rôznym spôsobom deformovať, či už ide o zašumenie celého modelu alebo systematickú zmenu jednotiek na hodnoty z intervalu (0,1).

KONVOLÚCIA

V našej sieti sa odohráva na druhej a štvrtej vrstve. Služi na vytiahnutie príznakov z objektu a to tým, že sa postupne prechádza konvolučným jadrom po celom objekte. Každá bunka jadra má vlastnú váhu, ale počas posunov sa váhy nemenia, uplatňuje sa teda princíp zdieľania váh.

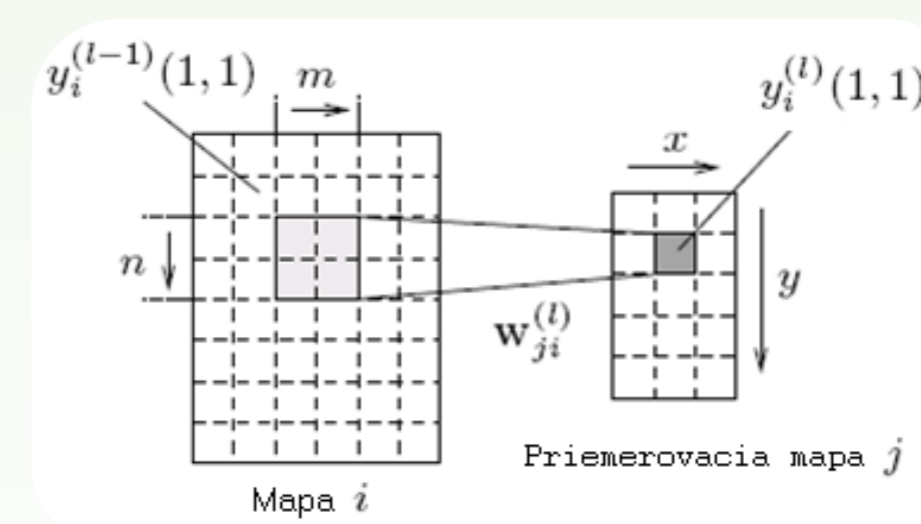
Na obrázku nižšie je príklad konvolúcie pomocou 3x3 jadra. Kvôli jednoduchšej vizualizácii je ukážka 2D. Rovnaký princíp platí aj pri pridaní tretieho rozmeru, s ktorým už pracuje aj naša sieť.



PRIEMEROVANIE

V našom modeli sa odohráva na tretej a piatej vrstve, teda hneď po konvolúcii. Tieto dve vrstvy sa navzájom striedajú a posledná priemerovacia vrstva je úplnou neurónovou sieťou napojená na výstupnú vrstvu.

Priemerovanie slúži na redukovanie dát, s ktorými pracujeme. Rovnako ako pri konvolúcii, aj tu sa využíva princíp zdieľania váh, no v tomto prípade je ešte očividnejší, pretože pre celú znakovú mapu sa používa iba jedna váha. Na obrázku nižšie je príklad spriemerovania oblasti 2x2 do jedného pixelu.



BACK-PROPAGATION

Back-propagation je algoritmus slúžiaci na učenie neurónových sietí. Je to iteratívny proces, počas ktorého sa pomaly menia váhy. Proces začína na konci siete, kde sa určí chyba, ktorú spravila sieť pri počítaní. Následne sa táto chyba distribuuje do predchádzajúcich vrstiev. Na základe týchto chýb je potom možné upravovať váhy, a teda pri ďalšom počítaní bude chyba siete menšia.

Pre konvolučné siete je tento algoritmus prispôbený, keďže každá vrstva pracuje trochu inak. Výpočet chyby výstupnej vrstvy ostáva rovnaký. Odlišne sa však počítajú chyby na konvolučnej a priemerovacej vrstve. Pri priemerovacej navyše závisí, či za ňou nasleduje konvolučná alebo výstupná vrstva. Následne po prepočítaní váh sa prepočítavajú odchýlky na konvolučných a priemerovacích vrstvách.

ZDROJE

1. V. Mařík, O. Štěpánková, J. Lažanský akol.: Umělá inteligence (4), Akademie věd ČR, Praha, 2003.
2. M. H. Hassoun: Fundamentals of artificial neural networks. MIT Press, Cambridge, 1995.
3. Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. Howard, W. Hubbard, I. Jackel: Handwritten digit recognition with a back-propagation networks. In: D. Touretzky, Advances in Neural Information Processing Systems, 1990.
4. McGill 3D Shape Benchmark, www.cim.mcgill.ca/~shape/benchMark/
5. J. Šimal: Rozpoznávanie biometrických údajov pomocou neurónových sietí. 2013
6. P. Kažimír: Rozpoznávanie biometrických údajov pomocou neurónových sietí – identifikačné znaky ucha, 2009
7. D. Medera: Klasifikácia chromozómov pomocou konvolučných neurónových sietí a inkrementálne prístupy k učeniu týchto neurónových sietí. 2007