

Problem Set 4: Neural Networks

Posted: Thursday, September 28, 2017

Due: Thursday 23:59, October 5, 2017

Please submit a report named `<your_kerberos>.pdf`.

Late Submission Policy: We do not accept late submissions. The submission deadline has a 50-minute soft cut-off; after midnight Thursday, submissions are penalized 2% per minute late.

Collaborators: You are free to discuss problems with other students but all writing must be done individually. Please list all collaborators at the top of your report.

Readings: Lecture slides 6 & 7 and Pg. 200-220 (Section 6.5) from the Deep Learning Book, freely available at <http://www.deeplearningbook.org/contents/mlp.html>.

Problem 1 *Convolution Layers*

Convolution layer is the most popular module in computer vision tasks. In this question, you will derive the equations for its forward and backward propagations.

- Consider your input x_{in} and output x_{out} are both 1-D signals with the same dimension N , and your kernel W has size k . Find the equation for forward propagation.
- Consider the back propagation process, with learning rate η , and the gradients from the last layer is $\frac{\partial C}{\partial x_{out}}$. Find the gradients of the input $\frac{\partial C}{\partial x_{in}}$, and the update rule for the kernel weights W^{i+1} .
- Discuss how you handle the boundaries and explain your choice.

Problem 2 *Pooling Layers (6.869 required; optional for 6.819)*

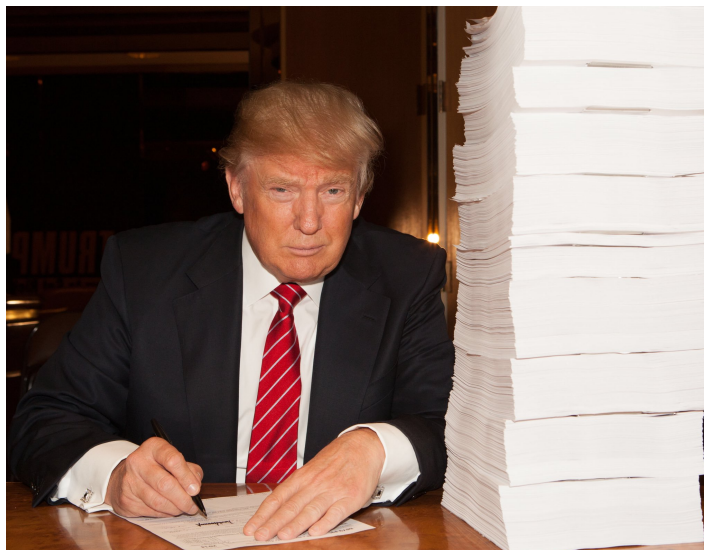
Pooling layer is a popular layer without trainable parameters. In this question, the pooling is a max pooling operator with stride 1.

- Consider your input x_{in} and output x_{out} are 1-D signals with the different size. Please find the equations or pseudo code for its forward and backward propagations.
- Discuss how you handle the boundaries and explain your choice.

Problem 3 *Research Problem (Open question, optional for everyone)*

Here is an interesting practical problem. We do not have standard answers for it, but we encourage you to think and try.

Can you train a system to recover the text in the document? Propose your methods.



Hint 1: This problem is similar to image deblurring/superresolution; you could build a neural network that learns to recover the high-resolution document from this blurry image. For training, you can use the parts of the document for which the content is known (it is a standardized form, see image below. Original images are attached in the Pset folder).

Form 8879-PE	IRS e-file Signature Authorization for Form 1065	OMB No. 1545-0123
Department of the Treasury Internal Revenue Service	<p>► Do not send to the IRS. Keep for your records.</p> <p>► Information about Form 8879-PE and its instructions is at www.irs.gov/form8879pe.</p>	2014
Name of partnership	For calendar year 2014, or tax year beginning ., 2014, and ending ., 20 .	Employer identification number
Part I Return Information (Whole dollars only)		
1	Gross receipts or sales less returns and allowances (Form 1065, line 1c)	1
2	Gross profit (Form 1065, line 3)	2
3	Ordinary business income (loss) (Form 1065, line 22)	3
4	Net rental real estate income (loss) (Form 1065, Schedule K, line 2)	4
5	Other net rental income (loss) (Form 1065, Schedule K, line 3c)	5
Part II Declaration and Signature Authorization of General Partner or Limited Liability Company Member Manager (Be sure to get a copy of the partnership's return)		
<p>Under penalties of perjury, I declare that I am a general partner or limited liability company member manager of the above partnership and that I have examined a copy of the partnership's 2014 electronic return of partnership income and accompanying schedules and statements and to the best of my knowledge and belief, it is true, correct, and complete. I further declare that the amounts in Part I above are the amounts shown on the copy of the partnership's electronic return of partnership income. I consent to allow my electronic return originator (ERO), transmitter, or intermediate service provider to send the partnership's return to the IRS and to receive from the IRS (a) an acknowledgement of receipt or reason for rejection of the transmission and (b) the reason for any delay in processing the return. I have selected a personal identification number (PIN) as my signature for the partnership's electronic return of partnership income.</p>		
General Partner or Limited Liability Company Member Manager's PIN: check one box only		
<input type="checkbox"/> I authorize _____ ERO firm name to enter my PIN _____ as my signature <small>do not enter all zeros</small>		

Hint 2: You might want to warp the photograph of the document first so that it appears rectangular and parallel to the camera plane, with the same aspect ratio as the original document.

- (a) Please provide your proposed processing pipeline for this challenging task, and give explanations for every step.
- (b) Try your best to recover the text in the image, and show your result. It does not matter if it does not look good.