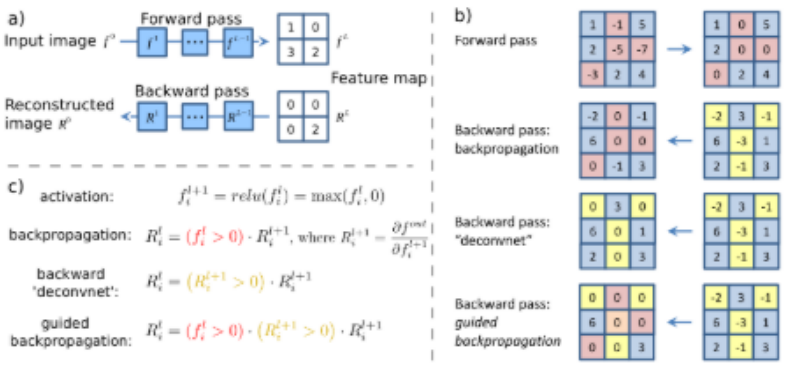
Striving for Simplicity: The All Convolutional Net | Paper | Notes | Implementation

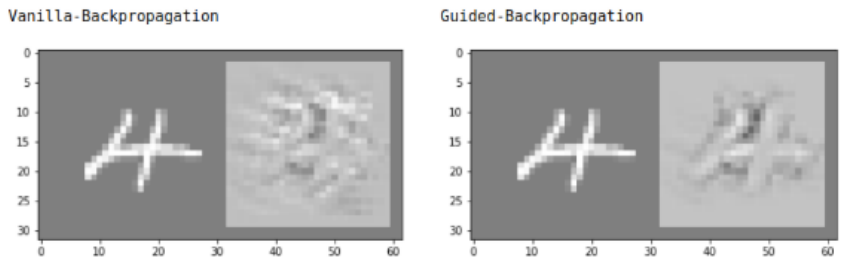
---

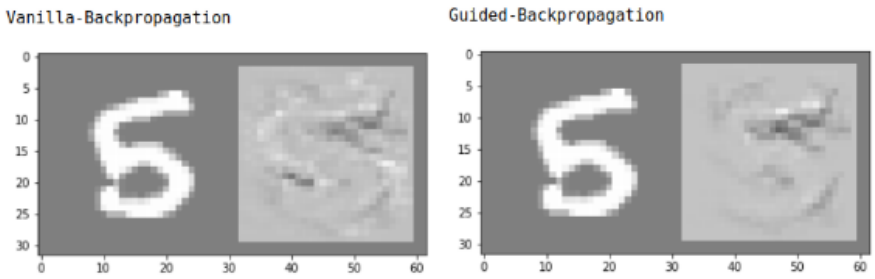
Guided Backpropagation

Deconvnet doesn't work well without max-pooling layers. To visualize the concepts learned by a network without pooling layers, the author proposes the method guided backpropagation. During backward pass, guided backpropagation handles **gradient from ReLU** in a combination of deconvnet and vanilla backpropagation. Gradients will be masked out when either corresponding entries of (1) **the top gradient** (**deconvnet**) or (2) **bottom data** (**vanilla backpropagation**) is **negative**. This prevents backward flow of negative gradients, corresponding to the neurons which decrease the activation of the higher layer unit we aim to visualize.



Different ways about how vanilla backpropagation, deconvent, and guided backpropagation handle ReLU nonlinearity in backward pass.





Comparison on vanilla backpropagation and guided backpropagation

Discussion on Deconvnet

The pooling layers in the deconvnet help the network to compute switches (positions of maxima within each pooling region). Using these switches in the deconvent backward pa ss makes it conditioned on an input image and the network doesn't directly visualize lear ned features. On the other hand, if the network doesn't contain pooling layers, the decon vnet isn't conditioned on an input image. This way we get insights into learned features i n lower layers. However, the network fails to visualize recognizable

feature maps for higher layers, because higher layers learn more invariant representation s, and there is no single image maximally activating those neurons. Hence to get reasona ble reconstructions for higher layers it is necessary to condition on an input image. Gradi ent based method is an alternative way that is conditioned on an image through activatio n functions (ReLU).

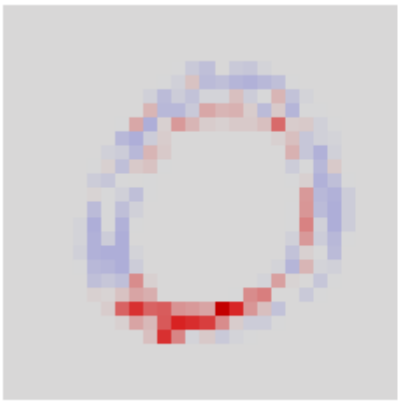
Q's

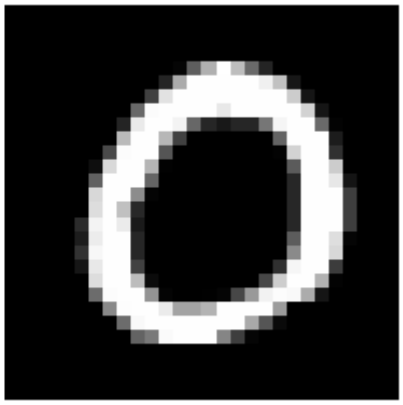
- What does the term "invariant representation" mean in page 8?

LRP Transpose

Implementation:

Exemplary Results:

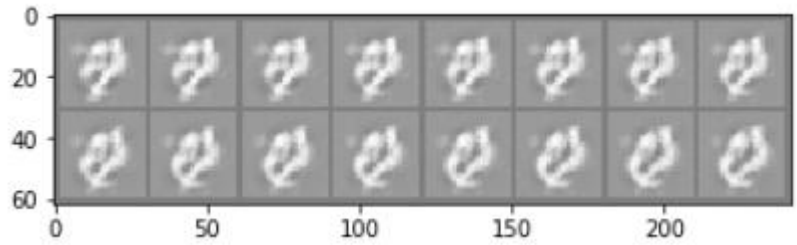




Activation Maximization in Code Space

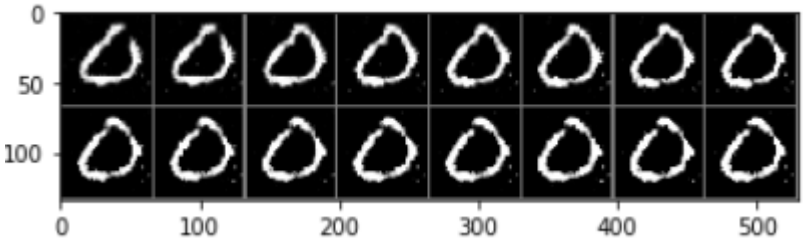
GAN Implementation

Exemplary Results for AM in CodeSpace for 0 (Last 16 runs of AM):



DCGAN Implementation

Exemplary Results for AM in CodeSpace for 0 (Last 16 runs of AM):


0
50
100
0
100
200
300
400
500