Reading Report

2019.4.22

1 Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift (BN)

1.1 BN

Algorithm 1 BN

Input: Values of x over a mini-batch: $\beta = x_{1...m}$; Parameters to be learned:γ, β

Output: $\{y_i = BN_{\gamma,\beta}(x_i)\}$

1: $\mu_\beta = \frac{1}{m}\Sigma_{i=1}^{m} x_i$; //mini-batch mean

2: $\sigma_\beta^2 = \frac{1}{m}\Sigma_{i=1}^{m}(x_i - \mu_\beta)^2$ // mini-batch variance;

3: $\widetilde{x_i} \leftarrow \frac{x_i - \mu_\beta}{\sqrt{\sigma_\beta^2 + \varepsilon}}$ // normalize

4: $y_i \leftarrow \gamma\widetilde{x_i} + \beta \equiv BN_{\gamma,\beta}(x_i)$ // scale and shift

advantages

1.regularize the model instead of L_1, L_2, or dropout

2.enable higher learning rate

3.Improve the precision of model training

1.2 Group Normalization(GN)

Nowadays this group presented a new paper, about Group Normalization

2 Deconvolutional Networks

2.1 DeConv

DeConv can be seemed as convolution matrix translation.
1.we reshape the kernel $(n * n)$ to sparse matrix $m * (m^2)$
2.reshape the image $(m * m)$ to $(m * m) * 1$
3.times it and reshape the answer

2.2 directions and development of DeConv

（1） Unsupervised learning is actually covolutional sparse coding[1][2]： the deconv here is only the reverse of the traditional conv in concept, and the traditional conv generates feature map from images, while the deconv finds a group of kernel and feature map by unsupervised method and lets them reconstruct images
（2）CNN visualization: using deconv to restore the feature map obtained by conv in CNN into pixel space to observe which pattern images are sensitive to the specific feature map. In fact, the deconv here is not an invertible operation of conv, but an transpose of conv, which is called *transpose_conv* in tensorflow.
（3）In the pixel - wise prediction such as image segmentation [4] and image generation [5], because you need to do to the original image size prediction of the space, and convolution due to stride tend to reduce the image size, so often need to restore to the original image size by using the method of upsampling, deconv ACTS as an upsampling role.(FCN and DC GAN)

3 Deep Residual Learning for Image Recognition(ResNet)

For the original network, simply increasing the depth will result in gradient dispersion or gradient explosion. We always use Batch Normalization to solve the problem but it may cause degradation, The number of network layers increased, but the accuracy of training set was saturated or even decreased

3.1 ResNet

图 1: ResNet

The local response normalization layer performs a kind of “lateral inhibition”by normalizing over local input regions

$$y = \mathcal{F}(x, \{W_i\}) + x$$

$$\mathcal{F} = W_2 * \sigma(W_1 x)$$

$$y = \mathcal{F}(x, \{W_i\}) + W_s x$$

and we always use shortcut to mapping.

3.2 advatanges

This residual learning structure can be realized by a shortcut connection add a forward neural network, as shown in the structure diagram. And shortcut connections are simply equivalent to performing an equivalent mapping, with no additional arguments or additional computational complexity. Moreover, the entire network can still be trained through end-to-end back propagation.

3.3 Research directions

DenseNet....
Also ResNet can diverse using Euler discretization(Jens Behrmann、Will Grathwohl, Ricky T.Q.Chen etc.)