CNN basic

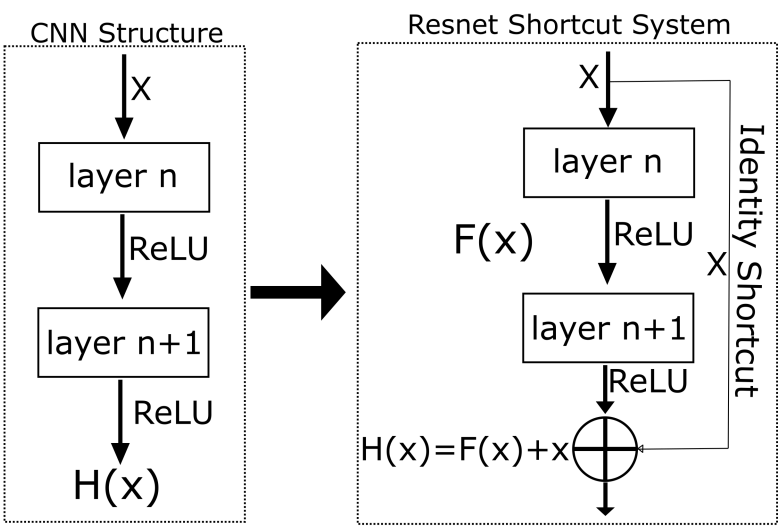
After the first effective usage of CNN models, the deeper structure of models starts to be problematic when back-propagated gradients are being quite small. The main solutions in this area come with the introduction of Residual Network Models(ResNet). By identifying a shortcut connection between multiple layers, it is possible to use deeper models without sacrificing the performance of a model. During their study, He et al. [1] proposed an identity function to link between several layers to be able to optimize deeper networks in much easier ways. After their first findings, they showed how to obtain better performance with 1001 layer deep neural networks [2].

Here, as a preliminary study to understand how a basic transfer learning-based Resnet model can differentiate cells, 954 images are cropped from each cell with a cropping size of 200x200. These crops are selected randomly through the channels. Train:Val:Test splits ratio set to become 754:100:100. In this training; train, validation, and test datasets are put into different folders. This selection is performed randomly instead of in the cross-validation step. Because the respective test dataset is compared with the final quantitative results. Starting from the Transfer Learning models, different Resnet architectures (Resnet18, Resnet 34, Resnet 50, Resnet 152) are used to train created cells datasets. As a final model, the Resnet-34 structure is selected when the time and performance are considered. A batch size of 32 is used in this training. The initial learning rate set to be 0.001, momentum is taken as 0.9. As an optimizer, SGD model is used. During training, Google Colabs' TESLA P100-PCI-E GPU is used and it takes 286 seconds to process 50 epochs. The traditional structure of CNN and replaced Resnet shortcut model is given in the Figure 2.8.a. The CNN structure based on the transfer learning model with the corresponding layers and filters is given in the Figure 2.8.b.

REFERENCES

1. He, K., Zhang, X., Ren, S. & Sun, J. *Deep Residual Learning for Image Recognition* in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, June 2016). https://doi.org/10.1109/cvpr.2016.90.
2. He, K., Zhang, X., Ren, S. & Sun, J. in *Computer Vision – ECCV 2016* 630–645 (Springer International Publishing, 2016). https://doi.org/10.1007/978-3-319-46493-0_38.

a.



b.

Layer Name	Resnet-34 Model Structure	Output Size
conv1	7x7, 64, Stride 2	112x112
pool1	3x3, max poll, stride 2	112x112
conv2_x	$\left[\frac{3x3,64}{3x3,64}\right]x3$	56x56
conv3_x	$\left[\frac{3x3,128}{3x3,128}\right]x4$	28x28
conv4_x	$\left[\frac{3x3,256}{3x3,256}\right]x6$	14x14
conv5_x	$\left[\frac{3x3,512}{3x3,512}\right]x3$	7x7
fully connected	average pool, 2d fc, softmax	1x1

Figure 1: Classification process. a. CNN and Resnet Shortcut System Comparison. b. Resnet 34 model