Short-circuit Neural Network for Child Pneumonia Classification

Ming Yan[1*], **Xueli Xiao**[2] and **Yu Zhang**[3]
[1]Sichuan University [2]Georgia State University
[3]Chengdu Aeronautic Polytechnic
yanmingtop@gmail.com, xxiao2@student.gsu.edu, zhangyu.cap@gmail.com

Abstract

Pneumonia is one of the top death-rate diseases in children. Classifying the complex pattern of pneumonia on X-ray images, neural network-based methods employed transfer learning or pre-training approach to promoting the performance of deep models. Different from the existing neural network-based methods, we proposed a new neural network, termed short circuit neural network (SC-Net), to classify the child pneumonia on X-ray images. Short circuit neural network conducts training by the short-circuit link, which introduced the gradient truncation of the recurrent neural network into feedforward neural networks by single back gradient propagation. This manner in short circuit neural network promotes the training efficiency by enhancing the gradient propagation of deep model from deep layers to front layers.

1 Introduction

The pneumonia is one of the top death-rate diseases in children under five-years-old. The clinical pediatrician commonly employs X-ray images associating their diagnosis. Classifying the tiny differences in X-ray images between pneumonia and the healthy always need professional knowledge. One influence factor is the X-ray original imaging mechanism with noises, low contrast in the X-ray images, which increase the difficulty of classification. Moreover, the limited training data is another important factor for data-based method.

Former computed associated diagnosis methods for pneumonia classification is a multi-step process with lung segmentation, feature selection, and classifier, which reduces the image factors and gets an accuracy performance around 80%. Neural network-based methods simplified the multi-steps into an end-to-end manner which learns discriminative features from raw data. Benefiting by the high nonlinear classification capability of neural networks, [Wang *et al.*, 2017] promotes the classification performance to 88%. However, the rare training data always limited, especially in the medical field. Transfer learning releases this constraint in the training of deep models. [Kermany *et al.*, 2018] promotes pneumonia accuracy to 92% with a transfer learning of eye dataset.

The transfer learning in deep neural network releases the rely on training data, while the data distribution of pre-train dataset and train dataset is different. In another word, transfer learning tends to reduce the accuracy of the deep neural network. Moreover, the transfer learning and pre-training initialization are not resolved the training problem of deep model. In this manuscript, we proposed a new neural network,SCNet, to promote the training efficiency and performance of deep neural networks. SCNet is a new network which introduced gradient truncation of the recurrent neural network into feed-forward neural network. Short circuit neural network promotes the training efficiency of deep neural network that got the same performance without transfer learning on the children pneumonia classification. Moreover, the short-circuit link in SCNet is a single back direction link, which is a gradient enhancement of back-propagation.

2 Short Circuit Neural Network

Short-circuit link in SCNet is the difference to other feed-forward neural networks that transfer the gradient from rear layers to front layers. The short-circuit link not only releases the gradient vanishment but also promotes the training efficiency. And this link is conducted by the gradient truncation in back-propagation of short-circuit link.

2.1 Short-circuit Link

Figure 1 shows the computation of short-circuit link in SC-Net. The motivation of this link is to reinforce the gradient of low-level layers in the deep model. And the gradient reinforcement is conducted by the gradient truncation. Besides, the only constraint is the neurons must be the same in the layers of short-circuit link.

In the feed-forward computation, SCNet is same as other feed-forward neural networks. While in the back-propagate computation, SCNet receives gradient from not only traditional back gradient of feed-forward neural network (Equation 1) but also extra rear layer from short-circuit

*Contact Author

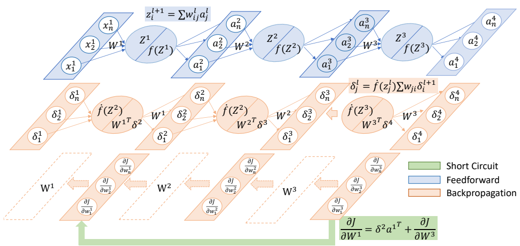


Figure 1: The computations in SCNet.

link(Equation 2).

$$\frac{\partial J}{\partial W^l} = \delta^{l+1} \cdot a^l \tag{1}$$

$$\frac{\partial J}{\partial W^l} = \delta^{l+1} {a^l}^T + \frac{\partial J}{\partial W^c} \tag{2}$$

2.2 Gradient Truncation

The extra gradient of short-circuit link is a truncated gradient other than traditional back-propagation computation. This motivation derives from long short-term memory network which faces gradient problems of vanishment and exploration with the time-delayed back-propagation computation (Equation 3). The gradient problems are solved by introducing of gradient truncation (Equation 4).

$$\begin{cases} net(t) = WS(t-1) + X(t-1) \\ S(t) = S(t-1) + f(net(t)) \\ \frac{\partial P(t)}{\partial P(t-\tau)} = \prod_{i=1}^{\tau}(1 + Wf'(net(t-i))) \end{cases} \tag{3}$$

$$\frac{\partial net(t)}{\partial S(t-1)} \overset{tr}{\approx} 0 \tag{4}$$

This gradient problem grows in the deep model with layers' increasing, which caused by the chain-rule of the back-propagation algorithm (Equation 5). The SCNet truncated the intermediate layers between the short-circuit link to be one which reduced the influence of gradient problem in Equation 6.

$$\begin{cases} Y^L = f^1(W^1, f^2(W^2, ...f^L(W^L, X))) \\ \frac{\partial J}{\partial w^l} = \frac{\partial J}{z^L}\frac{z^L}{\partial w^l} = \frac{\partial J}{\partial Z^L}\frac{\partial Z^L}{\partial Z^{L-1}}\frac{Z^{L-1}}{Z^{L-2}}\cdots\frac{Z^{l+1}}{Z^l}\frac{\partial Z^l}{\partial W^l} \end{cases} \tag{5}$$

$$\prod_{l=n}^{L-m-1} \frac{\partial Z^{l+1}}{\partial Z^l} \overset{tr}{\approx} 1 \tag{6}$$

3 Experiments

In the experiment section, SCNet is tested on the children pneumonia dataset and the CIFAR dataset.

Comparison in Pneumonia Classification Table 1 shows the accuracy comparison of SCNet with ResNet([He *et al.*, 2016]). SCNet got the best performance of 92.78%, with a sensitive of 93.33%, which outperforms the comparison method ResNet. In the evaluation, SCNet got the same performance with the state-of-art methods (accuracy 92.80% and sensitive 93.2%) which conducts transfer learning in the deep model training. Table 2 shows the performance of different depth SCNet and ResNet. SCNet has got better performance than ResNet on CIFAR10. With the layer increment, the difficulty of training makes the accuracy of ResNet decreases after 110. However, the accuracy of SCNet continuous increase, which proves SCNet promote the training of deep networks.

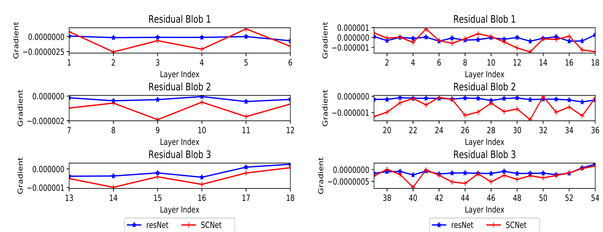
Table 1: Comparison on Children Pneumonia Classification

Lyaers	**20**	**32**	**44**	**56**	**110**
ResNet	88.46%	89.74%	91.82%	90.38%	88.14%
SCNet	89.10%	90.54%	92.31%	92.78%	89.42%

Table 2: Comparison on CIFAR10

Layers	**20**	**32**	**44**	**56**	**110**	**1202**
ResNet	91.25	92.49	92.83	93.03	93.39	92.07
SCNet	91.81	92.94	93.29	93.81	94.13	94.46

Gradient in SCNet Mean-gradient of each layer in 200-epochs is plotted in Figure 2. In the positive and negative range, the mean-gradient of SCNet and is clear enhanced than original ResNet in the rear blobs.



(a) (b)

Figure 2: Mean gradient Comparison of SCNet and ResNet

4 Reference

References

[He *et al.*, 2016] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[Kermany *et al.*, 2018] Daniel S Kermany, Michael Goldbaum, Wenjia Cai, Carolina CS Valentim, Huiying Liang, Sally L Baxter, Alex McKeown, Ge Yang, Xiaokang Wu, Fangbing Yan, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell*, 172(5):1122–1131, 2018.

[Wang *et al.*, 2017] Xiaosong Wang, Yifan Peng, Le Lu, Zhiyong Lu, Mohammadhadi Bagheri, and Ronald M Summers. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2097–2106, 2017.