BRONCHOPULMONARY SEGMENTATION OF THE LUNGS BY USING TERNARY NET WEIGHTS IN MASK-R NEURAL NETWORK

N.S. Francis, N.J. Francis, M. Saqib
Scientific Supervisor: Asst Prof. S.V. Axyonov
Tomsk Polytechnic University, Russia, Tomsk, Lenin ave., 30, 634050
E-mail: nadinesuzannefrancis@gmail.com

БРОНХОЛЕГОЧНАЯ СЕГМЕНТАЦИЯ В ЛЕГКИХ С ПОМОЩЬЮ ТЕРНАРНЫХ ВЕСОВ В НЕЙРОННОЙ СЕТИ MASK-R

Н.С. Франсис, Н.Д. Франсис, М. Сакиб
Научный руководитель: доцент, к.т.н Аксёнов С.В.
Национальный исследовательский Томский политехнический университет
Россия, г. Томск, пр. Ленина, 30, 634050
E-mail: nadinesuzannefrancis@gmail.com

Аннотация. *Цель работы заключается в разработке алгоритма для выявления бронхолегочные сегменты в легких человека при уменьшении вычислительных затрат. Алгоритм реализован без использования графического процессора.Основой алгоритма является модель Mask R-CNN с помощью троичного веса. Тройная гиперболическая касательная функция заменяет функцию активации CNN уменьшить накладные расходы. Это удобная система, созданная для помощи рентгенологам в сегментации легких с высокой точностью, а также недорого.*

Introduction. Medical image processing and modelling [1] has become a very famous sector. Many researchers are trying to identify ways and methods to help doctors with medical image analysis and simplify their process. Of all medical imaging, studies of images of the brain and lungs is given more importance by scientists. This is due to the fact that, identifying diseases [1, 2] at early stages in these organs is very vital for the patient's survival. In this research, image analysis is performed on the lungs to identify bronchopulmonary segments. Many segmentation methods [3] to identify various pathologies related to the lungs, have been performed over the years. These segmentation techniques produce output results with a percentage [4] of 43 - 75% which is very low for accurate disease analysis. Moreover, storage space can be saved by directly analyzing the dicom image, proving to be a better segmentation method for any image analysis.

In an attempt to segment the lung previously, shadowed segments of the lung were identified, and these shadowed regions were then sent into Mask R-CNN [3, 4] to identify the lung segments. During the second convolution of the CNN, U-Net was used to convolve the images and identify the segments. Though the output obtained is almost accurate in most cases, this method required an 8 GPU [4] (Graphic Processing Unit) to provide its results, due to the floating point matrix multiplications to calculate weights within the CNN. In order to help make this process with less overhead, these floating point weight functions [5] are replaced with ternary weights. To understand how this works, the entire function of Mask R-CNN is broken down and explained.

Dataset Preparation. All the medical images [6] are taken in dicom format and preprocessing is applied to them. After preprocessing, the image is obtained in .png or .jpg. These images are then analysed to identify the shadow of the bronchopulmonary segments. The shadow is analysed by Ridge based Distribution Analysis method (RAD). In this method, possible structure points are detected. These structure points define the entire border of the lung image. Ridge points or points of possible segment differentiation are then analysed. Ridge points are close to each other are finally joined together to form a histogram Ω(r). The histogram is calculated as follows.

$$A\ (r, \sigma) = T(r, \sigma s) * (\nabla\Omega(r, \sigma g) \cdot \nabla\Omega k(r, \sigma g))$$

Where σs is the neighbourhood size [7] to form the segment in the image, r is the centre ridge point within the area A, the value of $\sigma = \{\sigma s, \sigma g\}$, the number of ridge points forming the segment is T and the gaussian kernel k calculates the vector field calculus.

Method and Functioning of CNN. Each identified segment is then sent into the Mask R-CNN which consists of two stages. The first stage is called the Region Proposal Network (RPN) [7] where the input image provides a set of proposals as output with its score of accuracy. A sliding window over the convolved image provides the scores and the proposals. The output which is a bounding box consisting of each score and proposal is then stored in an ROIPool.

During the second stage, features are extracted from the ROIPool [7, 8] from each bounding box and classification and regression takes place. Parallel to identification of box offset and class, binary masks are identified for each ROI (region of interest). A small feature map is extracted from every ROI and placed in the ROI pool. Unlike the original mask R algorithm, ternary weights are quantised to the ROI, which is then divided into the spatial bins. After dividing them, max pooling is applied. The output of max pooling is then quantised and then bilinear interpolation helps produce the final segmented result. Fig1 shows the entire working model of the Mask R-CNN with ternary weights applied.

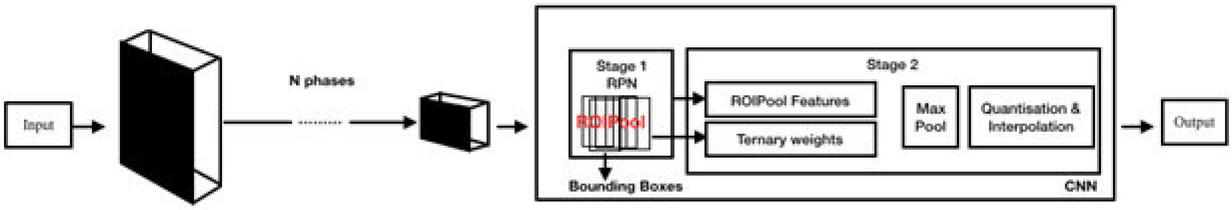

Fig. 1: Entire working model of Mask R-CNN with ternary weights

During hidden layer convolution, the ternary hyperbolic tangent [6, 8] function is used. This function also uses the ternary weights to calculate the output value for the next node in the hidden layer. These ternary weights [8] and activations are based on masked Hamming distances. The ternary weight is calculated from the following equation. The value of the ternary weight can be calculated as follows:

(i) The weight is +1 for all ternary weights greater than α

(ii) The weight is 0 for all ternary weights lesser than or equal to α

(iii) The weight is -1 otherwise

The value of α is calculated as follows for all weights from 1 to x.

$$\alpha = 0.7/x \sum |T_w|$$

Where $n_\alpha = \Sigma_i|T_w|$, and T_w is the ternary weight. The ternary activation function tanh_T is calculated with the equation.

$$\tan h_T = 0.5(2\delta y - \delta) - 0.5(2\delta y - \delta)$$

Where 2 hyperbolic tangents [8] (2δy-δ) and (-2δy-δ) are used to calculate the weights as +1 and -1, thus reducing the computational overhead and providing the required output.

Training the Dataset and Output Results. Each network is trained with around 200 iterations and with about 45 epochs each. The initial learning rate of the CNN is around 0.0025 and a cross-entropy loss for the preprocessed bitmap image is about 2%. Approximately 80 to 85 images are used for training and around 30 for testing. Out of all the images trained, the model produced an output accuracy of 86.82 percent which is 10 percent higher when compared to using the 8 GPU.

Conclusion: Most neural networks that are used for image processing require high processing GPUs and hence are expensive. If medical image analysis equipment becomes very expensive, it will be difficult for most hospitals to acquire the equipment required and research on medical image will be useless. For the purpose of reducing computational overhead, ternary weights are introduced into mask R-CNN to identify segments without the help of a GPU. The output produced using this method is in par or produces even more better bronchopulmonary segments and also makes it available to hospitals.

REFERENCES

1. Zheng, Li & Lei, Yiran. (2018). A Review of Image Segmentation Methods for Lung Nodule Detection Based on Computed Tomography Images. MATEC Web of Conferences. 232. 02001. 10.1051/matecconf/201823202001.
2. A R Amanda and R Widita (2016) Comparison of image segmentation of lungs using methods: connected threshold, neighborhood connected, and threshold level set segmentation J. Phys.: Conf. Ser. 694 012048.
3. Masood A, Yang P, Sheng B, et al. Cloud-Based Automated Clinical Decision Support System for Detection and Diagnosis of Lung Cancer in Chest CT. IEEE J Transl Eng Health Med. 2019;8:4300113. Published 2019 Dec 4. doi:10.1109/JTEHM.2019.2955458.
4. Harrison A.P, Xu Z, George K, Lu L, Summers R.M, Mollura D.J. (2017) Progressive and multi-path holistically nested neural networks for pathological lung segmentation from CT images. Medical Image Computing and Computer-Assisted Intervention MICCAI, 20th International Conference, Quebec City, QC, Canada, September 11-13, 2017, Proceedings, Part III. Springer International Publishing, 2017, pp. 621–629.
5. K. He, G. Gkioxari, P. Dollar and R. Girshick, "Mask R-CNN" in *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 42, no. 02, pp. 386-397, 2020. doi: 10.1109/TPAMI.2018.2844175 url: https://doi.ieeecomputersociety.org/10.1109/TPAMI.2018.2844175.
6. Sailesh Conjeti, Abhijit Guha Roy, Amin Katouzian, and Nassir Navab. Hashing with residual networks for image retrieval. In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 541–549. Springer, 2017.
7. vhttp://tuberculosis.by/ Last visited 19.09.2018. Under construction.
8. He, K., Gkioxari, G., Dollár, P., & Girshick, R.B. (2017). Mask R-CNN. 2017 IEEE International Conference on Computer Vision (ICCV), 2980-2988.