Recurrent Residual U-Net: Short Critical Review

Muhammad Zubair Khan
School of Computing and EngineeringUniversity of Missouri-Kansas City
Kansas City, USA

Abstract:- [Summary] Alom *et al.* in their article with title Recurrent Residual U-Net for Medical Image Segmentation, published in March 2019 **in Journal of Medical Imaging, proposed two deep network architectures for medical image semantic segmentation [1]. These models are evaluated using existing benchmark medical image datasets. This work aims to penetrate the deep learning concept in medicine to minimize human intervention in medical diagnosis. To achieve this goal, the author utilized the power of existing state-of-the-art deep network architectures designed for medical image segmentation, including U-NET, residual network, and recurrent convolutional neural network. It is found that deep learning generally anddefined deep neural architectures particularly has an enormous impact to accurately perform medical image analysis.**

Keywords:- *Deep Learning, Critical Review, Image Segmenta- Tion, Convolution Neural Network Sequence.*

I. INTRODUCTION

Unlike shallow architectures, which operate on handcrafted features designed to analyze and detect the object of interest using a comparatively small set of data samples, deep archi- tecture, inspired by the human brain, autonomously operates layer-wise to generate feature maps from provided input data [2], [3]. Several challenges are associated with deep networks; more prominent is its demand to data-in-bulk for generating fruitful results. Fortunately, various sensors made this possible to get a large amount of data with less effort. Hence, Deep learning has found great appreciation in computer vision, particularly in medicine, to diagnose life-threatening diseases with no or reduced human interference. Also, it helps practitioners in accurate and fast diagnosis. Deep neural architectures developed for natural images are ample due to ease in data availability—recent research analyzes its impact in medicine. Unlike images captured from RGB cameras, medical images are more not that simple to deal with. IBM researchers [4], [5] found that 90 percent of data sources available in the health industry comprise medical images, but a significant challenge resides in analyzing this data mainly due to its privacy concern. Therefore, health centers are reluctant to share this valuable information conveniently among researchers even. Moreover, expert practitioners and efficient tools are required for its annotation to make it cost- effective [6]–[8]. In the underline article, the author proposed two deep semantic segmentation networks called Recurrent U-Net (RU-Net) and Recurrent Residual U-Net (R2U-Net) for medical image segmentation. The proposed model utilized the foundational structure of U-Net [9] along with residualunits [10], which helps in training deep architectures along with recurrent convolutions, which assure improved feature representation for segmentation [11], [12]. Moreover, proposed architectures are executed in a multi-modal segmentation environment where retinal blood vessels, skin cancer, and lungs segmentation corpora are used as benchmark datasets to analyze models. For retinal vessel segmentation, the authors considered the patch-based method, while for skin lesion and lung segmentation end-to-end image-based approach is utilized to evaluate underline architectures. Moreover, the comparison has been made by keeping several parameters constant, and it has been observed that the proposed models outperform all state-of-art approaches with the same parameter count. Furthermore, the robustness of R2U-Net is also examined empirically concerning SegNet [13] and U-Net architectures.

II. EVALUATION

Researchers empirically observed that the deeper neural networks, the better the results achieved to a certain level, after which performance degrades. The reason behind this degradation is the vanishing and exploding gradient. The same would be the case with simple U-Net architecture if it would have been used [14]. As, if we keep increasing encoder decoders layers without taking specific preventive measures, at a certain point, training accuracy will degrade. Moreover, it is more effective to keep a record of previous input and current pixel information to utilize in predicting future output; in this way, it helps the model integrate context information that is significant in semantic segmentation. Since the proposed technique is a variant of U-Net and contains residual blocks and recurrent units, it could easily be extended to several multiple convolution de-convolution layers by in- corporating long/short skip connections. It helps in avoiding overall performance degradation in training accuracy that is caused by vanishing/exploding gradients [15]–[22]. More- over, the proposed approach can withhold reasonable long dependencies among pixel values by considering contextual data, which helps the network predict upcoming pixels and ultimately improve representation in the segmentation map. It is interesting to notice that the authors applied both patch and end-to-end image-based methods for claiming the proposed approach universal. Long-skip connections across the encoder- decoder path and short-skip connection within residual block not only resolve the vanishing gradient problem but also help the system avoid lossy compression with identity mapping.

III. RESPONSE

Multiple directions are identified after thoroughly analyzing the strategy pointed in a reviewed article. It is observed that the article sustains basic versions of recurrent units (vanilla version). With enough mathematical evidence and prior research conducted by researchers, other variants of the simple recurrent unit with memory cells and multiple gates like Gated Recurrent Unit (GRU) [23], Long Short-Term Memory (LSTM) [24] have shown enormous response when applied to Fully Convolutional Neural Network. Similarly, Bi-directional Recurrent Neural Network for capturing complete after-before context knowledge can be other alternatives along with its variant in terms of GRU and LSTM. Moreover, in residual convolutional networks, only identify mapping is considered between feature maps. Empirical study has shown that mixed identity and projection mapping boost the network perfor- mance with a negligible increase in parameters. Therefore, when feature maps are identical across convolutions, identity mapping works. However, nodes where dimension change with an increased in the number of filters hence feature maps, using projection mapping would boost the overall performance of underline architecture.

IV. CONCLUSION

Medical image analysis using a deep learning approach has opened new doors to the health care domain and is a significant step to replace human intervention with machine diagnosis. Deep learning has found great appreciation in computer vision, more specifically in medicine for diagnosing life-threatening diseases. Also, it helps practitioners in the accurate and fast patient examination. In the underline article, the author proposed two models, named RU-Net and R2U- Net, for medical image semantic segmentation. Models are evaluated using existing benchmark medical image datasets and evaluated empirically using state-of-the-art approaches with the same parameter count. It is observed that wrapping residual block and recurrent unit to U-Net boosts the overall performance by removing degradation caused by vanishing gradient without increasing additional parameters. Also, it is observed that the article sustain basic versions of the recurrent unit and residual blocks, alteration to which mathematically may increase training accuracy. Overall, it is observed that the proposed approach is considered to be one of the valuable contributions for medical image semantic segmentation.

REFERENCES

[1]. M. Z. Alom, C. Yakopcic, T. M. Taha, and V. K. Asari, "Nuclei segmentation with recurrent residual convolutional neural networks based u-net (r2u-net)," in *NAECON 2018-IEEE National Aerospace and Electronics Conference*. IEEE, 2018, pp. 228–233.

[2]. S. Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz, and D. Terzopoulos, "Image segmentation using deep learning: A survey," *arXiv preprint arXiv:2001.05566*, 2020.

[3]. ˙I. Atli and O. S. Gedik, "Sine-net: A fully convolutional deep learning architecture for retinal blood vessel segmentation," *Engineering Science and Technology, an International Journal*, vol. 24, no. 2, pp. 271–283, 2021.

[4]. D. Learning, "Deep learning," *High-Dimensional Fuzzy Clustering*, 2020.

[5]. M. Z. Khan and U. Qamar, "Towards service evaluation and ranking model for cloud infrastructure selection," in *2015 IEEE 12th Intl Confon Ubiquitous Intelligence and Computing and 2015 IEEE 12th Intl Conf on Autonomic and Trusted Computing and 2015 IEEE 15th Intl Conf on Scalable Computing and Communications and Its Associated Workshops (UIC-ATC-ScalCom)*. IEEE, 2015, pp. 1282–1287.

[6]. G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. Van Der Laak, B. Van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Medical image analysis*, vol. 42, pp. 60–88, 2017.

[7]. S. A. Kamran, K. F. Hossain, A. Tavakkoli, S. L. Zuckerbrod, K. M. Sanders, and S. A. Baker, "Rv-gan: retinal vessel segmentation from fundus images using multi-scale generative adversarial networks," *arXiv preprint arXiv:2101.00535*, 2021.

[8]. Z. S. Shaukat, R. Naseem, and M. Zubair, "A dataset for software requirements risk prediction," in *2018 IEEE International Conference on Computational Science and Engineering (CSE)*, 2018, pp. 112–118.

[9]. O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[10]. D. Li, D. A. Dharmawan, B. P. Ng, and S. Rahardja, "Residual u-net for retinal vessel segmentation," in *2019 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2019, pp. 1425–1429.

[11]. M. Liang and X. Hu, "Recurrent convolutional neural network for object recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3367–3375.

[12]. P. M. Samuel and T. Veeramalai, "Vssc net: Vessel specific skip chain convolutional network for blood vessel segmentation," *Computer Methods and Programs in Biomedicine*, vol. 198, p. 105769, 2021.

[13]. P. Kumar, P. Nagar, C. Arora, and A. Gupta, "U-segnet: fully convolu- tional neural network based automated brain tissue segmentation tool," in *2018 25th IEEE International Conference on Image Processing (ICIP)*. IEEE, 2018, pp. 3503–3507.

[14]. M. Islam, T. N. Poly, B. A. Walther, H. C. Yang, Y.-C. J. Li *et al.*, "Ar- tificial intelligence in ophthalmology: a meta-analysis of deep learning models for retinal vessels segmentation," *Journal of clinical medicine*, vol. 9, no. 4, p. 1018, 2020.

[15]. W. L. Alyoubi, W. M. Shalash, and M. F. Abulkhair, "Diabetic retinopa- thy detection through deep learning techniques: A review," *Informatics in Medicine Unlocked*, p. 100377, 2020.

[16]. Y. Wu, Y. Xia, Y. Song, Y. Zhang, and W. Cai, “Nfn+: A novel network followed network for retinal vessel segmentation,” *Neural Networks*, vol. 126, pp. 153–162, 2020.
[17]. D. Wang, A. Haytham, J. Pottenburgh, O. Saeedi, and Y. Tao, “Hard attention net for automatic retinal vessel segmentation,” *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 12, pp. 3384–3396, 2020.
[18]. T. Yang, T. Wu, L. Li, and C. Zhu, “Sud-gan: Deep convolution generative adversarial network combined with short connection anddense block for retinal vessel segmentation,” *Journal of digital imaging*, vol. 33, no. 4, pp. 946–957, 2020.
[19]. D. Chen, Y. Ao, and S. Liu, “Semi-supervised learning method of u-net deep learning network for blood vessel segmentation in retinal images,” *Symmetry*, vol. 12, no. 7, p. 1067, 2020.
[20]. D. Hao, S. Ding, L. Qiu, Y. Lv, B. Fei, Y. Zhu, and B. Qin, “Sequen- tial vessel segmentation via deep channel attention network,” *Neural Networks*, vol. 128, pp. 172–187, 2020.
[21]. S. A. Taghanaki, K. Abhishek, J. P. Cohen, J. Cohen-Adad, and
[22]. G. Hamarneh, “Deep semantic segmentation of natural and medical images: a review,” *Artificial Intelligence Review*, pp. 1–42, 2020.
[23]. M. Siam, S. Elkerdawy, M. Jagersand, and S. Yogamani, “Deep se- mantic segmentation for automated driving: Taxonomy, roadmap and challenges,” in *2017 IEEE 20th international conference on intelligent transportation systems (ITSC)*. IEEE, 2017, pp. 1–8.
[24]. A. Ullah, K. Muhammad, W. Ding, V. Palade, I. U. Haq, and S. W. Baik, “Efficient activity recognition using lightweight cnn and ds-gru network for surveillance applications,” *Applied Soft Computing*, vol. 103,p. 107102, 2021.
[25]. F. Shahid, A. Zameer, and M. Muneeb, “Predictions for covid-19 with deep learning models of lstm, gru and bi-lstm,” *Chaos, Solitons & Fractals*, vol. 140, p. 110212, 2020.