Journal of Innovative Image Processing (ISSN: 2582-4252)
www.irojournals.com/iroiip/

Root CT Segmentation Using Incremental Learning Methodology on Improved Multiple Resolution Images

K. Geetha

Dean, Academics and Research, JCT College of Engineering and Technology, Coimbatore, India

E-mail: geetha.arulmani@gmail.com

Abstract

The real-time issue of reliability segmenting root structure while using X-Ray Computed Tomography (CT) images is addressed in this work. A deep learning approach is proposed using a novel framework, involving decoders and encoders. The encoders-decoders framework is useful to improve multiple resolution by means of upsampling and downsampling images. The methodology of the work is enhanced by incorporating network branches with individual tasks using low-resolution context information and high-resolution segmentation. In large volumetric images, it is possible to resolve small root details by implementing a memory efficient system, resulting in the formation of a complete network. The proposed work, recent image analysis tool developed for root CT segmented is compared with several other previously existing methodology and it is found that this methodology is more efficient. Quantitatively and qualitatively, it is found that a multiresolution approach provides high accuracy in a shallower network with a large receptive field or deep network in a small receptive field. An incremental learning approach is also embedded to enhance the performance of the system. Moreover, it is also capable of detecting fine and large root materials in the entire volume. The proposed work is fully automated and doesn't require user interaction.

Keywords: Plant phenotyping, root system analysis, deep learning, image segmentation, X-ray computed tomography

1. Introduction

The process of characterizing, quantitatively and objectively, the plants' root system is known as ROOT phenotyping [1]. It indicates the way the root system develops, changes with environmental effect, external influence, and soil maintenance. In general, previously existing methodologies involved cleaning the roots by washing them and further examining the

DOI: https://doi.org/10.36548/jiip.2021.4.006
Received: 04.11.2021, received in revised form: 02.12.2021, accepted: 13.12.2021, published: 23.12.2021

material. Though methodology provides a high-throughput output, it will result in loss of some finer roots and alterations during the washing process [2]. Moreover, it is also popular for these roots to be flattened and then used as images resulting in the loss of architectural information. This is where multi-viewing of images goes a long way in recovering the 3D structure of the root system. The development and growth of the root are different, based on the interaction between the soil and the root [3] which cannot be perceived in the absence of soil. In this proposed work, volumetric segmentation of the root system is performed using segmentation method taking into consideration the soil in CT. Changes are incorporated in the architecture of the encoder-decoder to tackle the shortcomings of deep learning [4].

Instead of compromising between a high-resolution input and large receptive field, two branch networks are proposed in this methodology where two types of examination take place:

1. Increased input size in lower resolution volume
2. Small receptive field in high resolution volume

The two branches are further combined to form an enhanced output with wider context and higher resolution. Multiple auxiliary loss functions are used to train the network, incorporating local and semantic features in every branch. With the help of a micro CT scanner [5], large CT images are captured and used as a dataset for training. An expert segments these images manually and the segmented images act as the ground truth for evaluation and training purposes. Analysis is conducted extensively on segmentation accuracy [6] which indicates that the proposed work is more accurate than the already existing methodologies that are used for semantic segmentation. In this paper, hard negative examples are used to illustrate incremental learning which further enhances the performance of the system. On account of large volumes, this methodology does not require parameter selection or user interaction, resulting in progress towards full automation of root CT [7]. The major contributions of this paper are as follows:

1. A new network architecture made up of encoder-decoders situated in a multi-resolution parallel pipeline is implemented for segmentation of the root system. A low resolution large receptive field and a higher resolution small receptive field are incorporated for training the network. This results in a significant improvement of the resolution of the network as far as large volume is concerned.

2. In this work, a multi loss training approach is introduced wherein every individual pipeline is made to branch and learn several tasks which contribute towards improving the segmentation accuracy.
3. The proposed architecture is adapted to suit volumetric segmentation and the simulated outputs indicate a positive improvement in all metrics.

2. Literature Review

In recent years, root-soil imaging has paved the way to analyze and closely study the interaction between plant and soil [8]. However, this has resulted in several challenges to accurately follow the data generated. To retrieve a non-destructive root image [9], Computed Tomography (CT) [10] is the apt technology that doesn't cause harm or destroy the structure of soil or root. The bottleneck faced is due to the large volumetric data generated using CT system and analysis of the same proves to be time-consuming [11-13]. The complex root structure of heterogeneity is difficult to analyze with the help of automated systems which is also the reason why semi-manual and manual approaches are preferred. This not only involves large investment of time but will also be difficult to execute as the volume of material increases. To determine the root system accurately [14], the root material should be segmented in a reliable manner. However, segmentation of the root soil CT images will be difficult due to the complex nature of the material [15]. A valuable proportion is provided by the CT scanners such that the root material with similar density are deemed to be misclassified with respect to risk. The presence of organic material in the soil will lead to confusion in computational approaches and human approaches [16].

However, the presence of the root branch [17-19] will be different in length based on the distribution of material, scanning hardware and reflection of age in root material. For successful segmentation [20], a degree of local contextual information is necessary. Based on the given image value [21], the distribution of noise will vary. Similarly, based on the scan resolution, the appearance of the root may be sparse or in contiguous structures such that only limited pixels root material are adjacent or overlapped in the adjacent slices [22]. The most challenging aspect is the high resolution of CT volumes wherein only a small proportion involves root material. This will result in computationally efficient image processing.

Irrespective of the layers used [23], most of the places have incorporated separation of the downsampling and consolation layer from the deconvolutional and upsampling layer. This way a symmetric structure is enabled wherein, unpooling and deconvolutional layers hold similar resolution as that of the corresponding pooling and convolutional layers. Thus the features are recognized as 'normal' and are not included in the decoder-encoder network [24], thereby resulting in low-level image data at the system level where the output is obtained. The data from the downsampling path is copied by the skip layers in correlation with the upsampling resolution layer, which are then combined with the help of concatenation or sum operation. In some methodologies, end-to-end segmentation [25] is performed with the help of stacks of encoders-decoders.

3. Methodology

Semantic segmentation is often performed using the encoder-decoder networks. Stacked hourglass, U-Net, Segnet, and FCN are some of the common techniques used for this purpose. A stacked hourglass like architecture is used as the base of the proposed network with modification to accommodate volumetric data handling capability. Figure 1 shows the proposed network architecture. When compared to the conventional CNNs, an additional dimension is present in some of the memory intensive volumetric networks. The feature size or depth must be reduced as only small input volumes of the range 16^3 are accepted by the network before the prohibition of memory consumption. Discerning the soil matter and root material, is challenging with limited field view when the input size is larger than the individual roots. Before encoder-decoder, down sampling is performed in the secondary network. With 1283 pixels, a larger input field of vision is allowed with this architecture. The resolution for several networks are constrained with this architecture as shown in Fig.1.

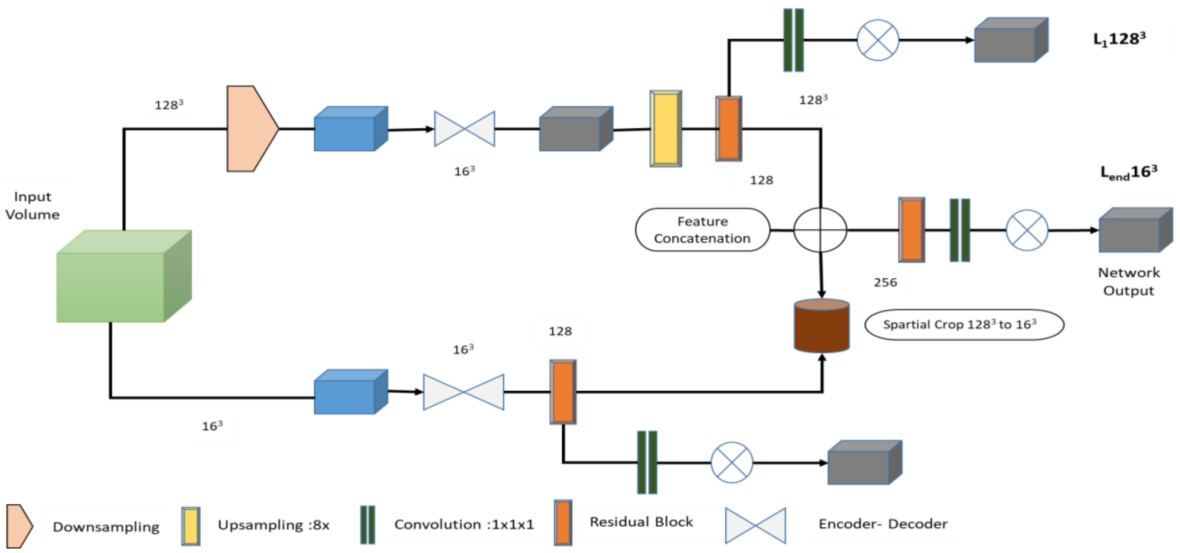

Figure 1. Architecture of proposed Deep Learning Methodology with encoders-decoders

A deep learning approach is proposed using a novel framework, involving decoders and encoders. The encoders-decoders framework is useful to improve multiple resolution by means of upsampling and downsampling images. This methodology is enhanced by incorporating network branches with individual tasks using low-resolution context information and high-resolution segmentation. In large volumetric images, it is possible to resolve small root details by implementing a memory efficient system, resulting in the formation of a complete network.

A parallel architecture that is functional at multiple resolutions is proposed to utilize the wider contextual data and local pixel data of the surrounding field of vision. Different input volume sizes are considered for the two parallel pipelines used in this network. The spatially aligned outputs of these paths are combined and the features from both are used to perform the final segmentation. Alternating Direction Method of Multipliers (ADMM) technique is used for training the network. Binary Cross Entropy (BCE) loss is used for training a sigmoid function before passing the network output through it. Wider and coarse field of view is considered with the down sampling path while high resolution images are handled in small patches with native path. The native path is spatially aligned to the wider field of view features that are appropriately cropped and then concatenated to be used by the final network layers for segmentation purposes.

4. Results

Volumetric CT images that are captured from soil grown crop roots are used for the evaluation in this work. Volume Graphics software package is used with region growing segmentation approach manually by an expert to obtain the ground truth annotation of each volume. The proposed multi-resolution network performance is investigated by comparing certain networks with the existing bottom-up tool. In the medical domain, the proposed volumetric segmentation model is compared with three state-of-the-art models. The production of data along with ground truth, setting up of the network, configuration for training and comparison of experimentation results are essential segments of this work. 50 volumetric images are used as a dataset that have captured crop roots from the soil. A micro CT scanner is used for the image capturing purpose. In each axis, the resolution of the CT scanner is set at 54µm per voxel while capturing images of a 15 day old crop grown in soil consisting of sand and cocopeat. 12GB RAM, Nvidia Titan X trained Torch7 is used for implementation of the network. For the internal encoder-decoder, at each layer, the total features are set at 128. In a random manner, smaller batches of 3 x 500 crops are obtained during training. The root material is observed in a small portion of voxel in large volumetric images.

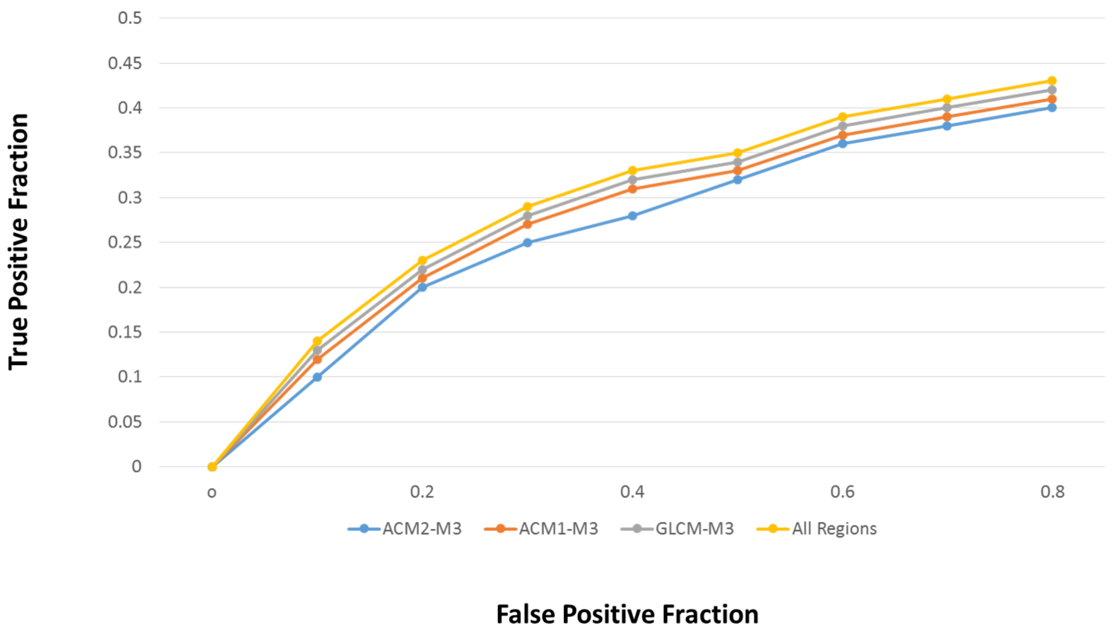

Figure 2(a). Performance of Training and Validation with False and True Positive Fraction

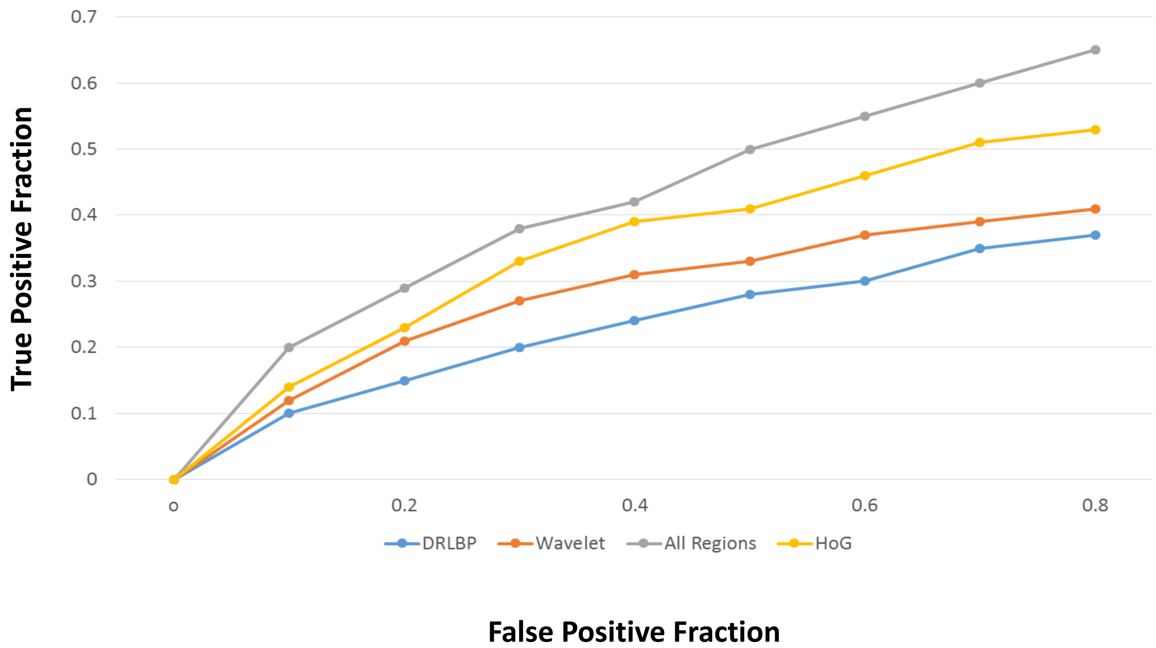

Figure 2(b). Performance of Training and Validation with False and True Positive Fraction for DRLBP, Wavelet, all regions and HoG

With a learning rate of $3x10^{-4}$, rmsprop is used for training each network. In order to ensure a fair test, the utilization of the GPU memory available is maximized by adjusting each network appropriately with respect to the parameters and/or batch size. Training of each network is performed for up to 150 epochs while retaining epoch based on the highest validation performance. During testing, every location is segmented and the final layer's output undergoes thresholding to obtain the final predictions. For various network architectures, the performance of training and validation are presented in Figure 2(a) and 2(b). For every 5 epochs, calculation of the validation loss is performed while estimating the training set loss value at each epoch. Stable validation performance is observed at the multi-loss multi-resolution network.

When compared to the conventional architectures, faster convergence is observed at the multi-resolution networks. The overall training process time is maintained while improving the network precision through hard negative images that undergo incremental learning. The conventional medical segmentation networks, volumetric U-Net, stacked hourglass and other volumetric segmentation networks that are commonly used and popular root phenotyping tools are compared with the proposed model. The proposed deep learning model outperforms the conventional imaging schemes. It is observed that based on the configuration, either precision or recall parameters are often compromised with the single resolution deep architectures. The noise factor is increased while better segmentation is achieved with fine roots and root boundaries with native resolution. Despite achieving noise reduction, the root materials are under segmented with wider fields.

5. Conclusion

CT images of crop roots from soil are segmented volumetrically using the proposed multiple resolution encoder-decoder model presented in this paper. Expert users have annotated the volumetric images used in this work in advance. Images with multiple resolutions are processed using two encoding and decoding paths. The overall network performance is enhanced with the contribution from the features learnt by each path. Thereby the system is trained with the multiple loss functions. When compared to the conventional techniques, IoY and Dice based overall scores are high along with high precision and recall, with successful segmentation by the proposed model. The soil CT images can be used for automatic segmentation of plant roots. In CT image segmentation, the human involvement can be largely reduced with the adoption of the proposed model. Furthermore, the throughput is improved in plant selection and root-soil interaction-based knowledge enhancement required for breeding programs. Different species can be analyzed with multiple scanner settings and imaging techniques in the future work for exploring the networks using transfer learning and provide finer segmentation details while improving the network architecture. The deep learning techniques are enhanced by improving the robustness with a diverse training dataset considering a larger variety of soil types.

References

[1] Verbitskiy, S., Berikov, V., & Vyshegorodtsev, V. (2021). Eranns: Efficient residual audio neural networks for audio pattern recognition. arXiv preprint arXiv:2106.01621.

[2] Adam, E. E. B. (2020). Deep Learning based NLP Techniques In Text to Speech Synthesis for Communication Recognition. Journal of Soft Computing Paradigm (JSCP), 2(04), 209-215.

[3] Xu, K., Zhu, B., Kong, Q., Mi, H., Ding, B., Wang, D., & Wang, H. (2019). General audio tagging with ensembling convolutional neural networks and statistical features. The Journal of the Acoustical Society of America, 145(6), EL521-EL527.

[4] Rodrigo, W. U. D., H. U. W. Ratnayake, and I. A. Premaratne. "Identification of Music Instruments from a Music Audio File." In Proceedings of International Conference on Sustainable Expert Systems: ICSES 2020, vol. 176, p. 335. Springer Nature, 2021.

[5] Vijayakumar, T. (2019). Comparative study of capsule neural network in various applications. Journal of Artificial Intelligence, 1(01), 19-27.

[6] de Benito-Gorron, D., Lozano-Diez, A., Toledano, D. T., & Gonzalez-Rodriguez, J. (2019). Exploring convolutional, recurrent, and hybrid deep neural networks for speech and music detection in a large audio dataset. EURASIP Journal on Audio, Speech, and Music Processing, 2019(1), 1-18.

[7] Sungheetha, Akey, and Rajesh Sharma. "Design an Early Detection and Classification for Diabetic Retinopathy by Deep Feature Extraction based Convolution Neural Network." Journal of Trends in Computer Science and Smart technology (TCSST) 3, no. 02 (2021): 81-94.

[8] Sankar, MS Arun, Tharak Sai Bobba, and PS Sathi Devi. "Stage Audio Classifier Using Artificial Neural Network." In International Conference on Communication, Computing and Electronics Systems, pp. 139-147. Springer, Singapore, 2020.

[9] Nanni, L., Maguolo, G., Brahnam, S., & Paci, M. (2021). An ensemble of convolutional neural networks for audio classification. Applied Sciences, 11(13), 5796.

[10] Balasubramaniam, Vivekanadam. "Artificial Intelligence Algorithm with SVM Classification using Dermascopic Images for Melanoma Diagnosis." Journal of Artificial Intelligence and Capsule Networks 3, no. 1 (2021): 34-42.

[11] Ghubaish, A., Salman, T., Zolanvari, M., Unal, D., Al-Ali, A. K., & Jain, R. (2020). Recent advances in the internet of medical things (iomt) systems security. *IEEE Internet of Things Journal*.

[12] Vachhani, Hrishikesh, Mohammad S. Obiadat, Arkesh Thakkar, Vyom Shah, Raj Sojitra, Jitendra Bhatia, and Sudeep Tanwar. "Machine learning based stock market analysis: A short survey." In International Conference on Innovative Data Communication Technologies and Application, pp. 12-26. Springer, Cham, 2019.

[13] Hameed, S. S., Hassan, W. H., Latiff, L. A., & Ghabban, F. (2021). A systematic review of security and privacy issues in the internet of medical things; the role of machine learning approaches. *PeerJ Computer Science*, *7*, e414.

[14] Sathye, Rohit, Sumedh Deshprabhu, Mandar Surve, and Deepak C. Karia. "Smart Medicine Distributing Tray." In International Conference on Innovative Data Communication Technologies and Application, pp. 57-66. Springer, Cham, 2019.

[15] Razdan, S., & Sharma, S. (2021). Internet of Medical Things (IoMT): Overview, Emerging Technologies, and Case Studies. *IETE Technical Review*, 1-14.

[16] Nene, Rajas, Pranay Narain, M. Mani Roja, and Medha Somalwar. "Fourier Descriptors Based Hand Gesture Recognition Using Neural Networks." In International Conference on Innovative Data Communication Technologies and Application, pp. 140-147. Springer, Cham, 2019.

[17] Smys, S., and Wang Haoxiang. "Data Elimination on Repetition using a Blockchain based Cyber Threat Intelligence." IRO Journal on Sustainable Wireless Systems 2, no. 4 (2021): 149-154.

[18] Raj, Jennifer S. "Security Enhanced Blockchain based Unmanned Aerial Vehicle Health Monitoring System." Journal of ISMAC 3, no. 02 (2021): 121-131.

[19] Huang, X., & Nazir, S. (2020). Evaluating security of internet of medical things using the analytic network process method. *Security and Communication Networks*, *2020*.

[20] Hiremath, Shivarajkumar, and R. Sanjeev Kunte. "Public Auditing Scheme for Cloud Data Integrity Verification." In International Conference on Innovative Data Communication Technologies and Application, pp. 237-246. Springer, Cham, 2019.

[21] Chakrabarty, Navoneel. "A Regression Approach to Distribution and Trend Analysis of Quarterly Foreign Tourist Arrivals in India." Journal of Soft Computing Paradigm (JSCP) 2, no. 01 (2020): 57-82.

[22] Shirley, D. R. A. (2014, July). Systematic diagnosis of power switches. In *2014 International Conference on Embedded Systems (ICES)* (pp. 32-34). IEEE.

[23] Bhalaji, N. "Reliable Data Transmission with Heightened Confidentiality and Integrity in IOT Empowered Mobile Networks." *Journal of ISMAC* 2, no. 02 (2020): 106-117.

[24] Valanarasu, Mr R. "Comparative Analysis for Personality Prediction by Digital Footprints in Social Media." Journal of Information Technology 3, no. 02 (2021): 77-91.

[25] Senthilkumar, M., Kavitha, V. R., Kumar, M. S., Raj, P. A. C., & Shirley, D. R. A. (2021, March). Routing in a Wireless Sensor Network using a Hybrid Algorithm to Improve the Lifetime of the Nodes. In *IOP Conference Series: Materials Science and Engineering* (Vol. 1084, No. 1, p. 012051). IOP Publishing.

Author's biography

K. Geetha is presently working as a Dean-Academics and Research at JCT College of Engineering and Technology, Coimbatore, India. Her area of research includes image processing, computer vision, artificial intelligence, optimization techniques, power systems, process control and automations.