Detecting and Classifying Neurofibromas using Deep Learning

Tuong Hieu Huynh
School of Information Technology and Electrical Engineering
The University of Queensland, Qld., 4072, Australia

Abstract

Cutaneous Neurofibromas (cNFs), a subtype of Neurofibromatosis type 1 (NF1), are benign tumors arising from nerve tissues. These prevalent genetic disorders present a significant challenge in medical imaging due to their variable appearance, severity. Since they cause a huge impact on the life quality of many patients with cosmetic disfigurement, the accurate detection and assessment of cNFs or NF1 severity is crucial for guiding appropriate treatment interventions.

The U-net architecture, a convolutional neural network (CNN) structure introduced in 2015 for biomedical image segmentation, has presented an achievement in precise segmentation of small targets, making it well-suited for neurofibroma detection and classification. Through presentations and discussions on U-net-based deep learning techniques used for semantic segmentation problem, this conference aims to highlight the application of U-net in automating the identification and severity measurement of neurofibromas in medical imaging data through a graphical user interface (GUI).

1 Introduction

NF1 is a genetic disorder that affects individuals globally, with an incidence ranging from about 1 in 2600 to 1 in 4500. One of the most common subtypes of NF1 is cNFs, which affect over 99% of adult patients [1]. These tumors vary widely in size, ranging from 2 mm to 3 cm, and in quantity. They are well-defined tumors closely associated with nerves in the skin. Despite being benign, these tumors often cause discomfort, itching, and aesthetic concerns, leading to psychosocial challenges for NF1 patients and significantly impacting life quality [1].

Predicting the severity of NF1 presents a challenge because of its diverse presentations across various body regions and ages, highlighting the crucial need for accurate detection and measurement to personalize treatment strategies effectively. Traditional approaches to detect and classify neurofibromas in medical imaging typically involve manual interpretation, resulting in subjective evaluations and diagnostic variability. Recent endeavors have focused on using edge detection filters and mathematical algorithms to address pixel-wise semantic segmentation tasks in the field of medical imaging [2].

Artificial intelligence (AI) is increasingly applied in medical imaging analysis, aiding doctors in diagnostic tasks and enhancing diagnosing efficiency. Thanks to advancements in hardware, deep learning methodologies proved their significant capabilities in tasks related to image processing. Specifically, recent advancements in deep CNN models have revolutionized this field. CNNs outperformed the state of the art in this task due to their capacity to extract intricate hierarchical features and their end-to-end trainable framework, marking a significant breakthrough in the field [3]. With CNN algorithms, deep learning (DL), a new subset of machine learning (ML), can efficiently process high-dimensional data in image classification, object detection, and semantic segmentation [4].

Deep learning operates through a convolutional multi-layer neural network characterized by numerous hidden layers and free parameters. Within this architecture, each input image undergoes processing through convolution layers, pooling layers, filters/kernels, and fully-connected layers, culminating in a final decision-making process facilitated by a functional operation. Thanks to their ability for approximation, feature extraction, optimization, and classification, deep learning-based segmentation methods are widely selected in medical field analysis.

Various deep learning network structures, including CNN, deep residual network (DRN), deep feed-forward network, deep convolutional neural network (DCNN), and U-net, have been introduced [5]. Among these, U-net stands out as a prominent method for medical image segmentation. Its versatility has been demonstrated in various medical fields such as brain tumor segmentation, cardiac analysis, liver examination, breast imaging, and retinal vessel delineation [6]. U-net consistently outperforms other solutions, achieving a much higher average Intersection over Union (IOU) of 92% in cell segmentation tasks [7]. Therefore, the U-net structure is selected in this study.

This study aims to utilize a U-net-based deep learning approach to detect NF1 tumors and evaluate their severity across various body regions. This will be accomplished by computing the ratio of detected NF1 to the cross-sectional area of unaffected skin. The paper further identifies the most efficient model by analyzing loss and Intersection over Union (IOU) evaluation metrics through experimental analysis.

Additionally, a graphical user interface (GUI) has been developed to demonstrate the potential application of the proposed system.

2 Segmentation techniques background and proposed model

Semantic segmentation involves assigning a categorical label to every organ or lesion pixel in a medical image from its background, which is essential for determining patient treatment. It efficiently extracts critical features regarding organ or lesion shapes and volumes, drawing significant attention from researchers [2].

In biomedical image processing, the need to label each pixel individually with precise location in the image poses a highly complex challenge, particularly due to the training data limit [7]. Recent approaches have tackled these limitations by integrating features from multiple layers, enabling effective localization and comprehensive context utilization simultaneously. One such approach is U-net, introduced in 2015, specifically designed to address these challenges [8]. U-net fulfills the medical image segmentation task through its encoder-bottleneck-decoder paths and skip connections. This makes it well-suited for tasks of neurofibromas detection and classification in this study.

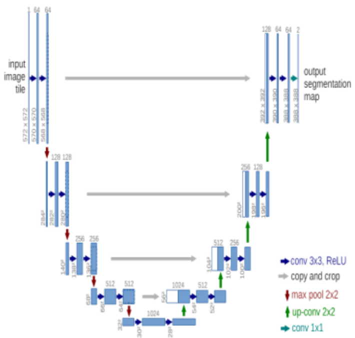

Fig 1. U-net architecture (from page 2 of [7])

The left side of the U-Net depicts encoder path, the conventional structure of a CNN. This involves iteratively applying two unpadded convolutions of size 3x3, then processed by a rectified linear unit (ReLU), and then downsampling using a 2x2 max-pooling layer with a stride of 2. The feature channels number is doubled after downsampling process.

On the right side of the U-Net architecture is the decoder path. Here, the high-resolution features extracted from the encoder path are then concatenated with the upsampled output to incorporate spatial information. This results in an increase in the number of feature channels. These features then undergo two 3x3 upsampling convolution procedures, effectively halving the feature channels number but increasing the resolution of the output. Following this step, the output undergoes processing by a Rectified Linear Unit (ReLU), facilitating the U-net model's acquisition of contextual information from higher resolution layer. This U-shaped architecture symmetrically reconstructs detailed localizations.

The output is a one-channel convolutional layer, generating probabilities for the foreground class through using sigmoid function. The network operates as a fully convolutional network, relying solely on valid segments from each convolution while excluding fully connected layers. Consequently, the segmentation map only encompasses pixels where complete context is accessible in the input image.

U-net-based approaches for neurofibromas detection typically involve training the U-net network on annotated medical imaging data. During training, the network learns to segment regions of interest corresponding to neurofibromas from background. Transfer learning techniques are used to adapt pre-trained U-net models to the specific characteristics of neurofibromas imaging data, enhancing performance with limited labeled data.

Following the forward pass of U-net structure, the segmentation map is generated, assigning a probability score to each pixel indicating its likelihood of belonging to the foreground class. During the backward pass, the disparity between the predicted output and the actual ground truth labels is calculated using a loss function, typically using Binary Cross-Entropy for binary classification tasks.

The gradients of the loss function with respect to the model parameters, computed through the chain rule of calculus, are utilized to update the model parameters, including weights and biases, in the direction that minimizes the loss function. This update process is typically performed using an optimization algorithm. Specifically, Adaptive Moment Estimation (Adam) algorithm is used in this study. Through iterating over multiple batches of training data, the process involves computing gradients and updating model parameters gradually, improving the model's fit to the training data.

In the context of binary classification distinguishing between NF1 and healthy skin, where NF1 is labeled as 1 and healthy skin as 0, the energy function involves applying a pixel-wise sigmoid operation over the final feature map. This sigmoid operation squashes the output of each pixel to a value between 0 and 1, representing the probability of belonging to the positive class (NF1). The result of this operation is then used with the Binary Cross-Entropy Loss function to calculate the discrepancy between the predicted probability and the true binary label (0 for healthy skin, 1 for NF1) for each pixel.

Through combining sigmoid function with the Binary Cross-Entropy Loss, the loss is computed, resulting in more stable and preventing potential numerical instabilities that can arise when applying the Sigmoid function directly to the logits.

In binary classification tasks in medical images, the Intersection over Union index (IOU) is used to quantify the intersection between the predicted and ground truth mask. During model training, IOU indices are computed for batches and take their average over the entire dataset to display the metric of the model's performance. This index needs to be maximized to increase the model's performance.

3 Databases and Pre-Processing

Detecting and Classifying NF1 for measuring its density on the patient’s skin requires access to relevant and comprehensive datasets of other patients. One confidential dataset is the images data collected from 13 different patients with different NF1 conditions and provided by the University of Newcastle and the University of Queensland for this project study. This comprehensive dataset aids researchers in understanding different NF1 patterns, using in manually labelling NF1 by Labelme software for creating ground truth masks. The training dataset comprises images and gound truth masks from 9 different patients, while the testing dataset comprises the ones from 4 different patients not included in the training dataset to train and evaluate U-net model.

Image patching serves to expand the training dataset, followed by augmentation to enhance diversity, and align with the model architecture. This approach is particularly beneficial for scenarios with limited data or where invariance to transformations is required, helping to prevent overfitting by training the model to a broader data variations range. After preprocessing, the data is passed through a DataLoader, a utility commonly used in machine learning frameworks like PyTorch. The DataLoader streamlines dataset management during training and validation by loading data in batches, shuffling for randomness, and parallelizing data loading across epochs. This simplifies data handling, allowing developers to focus on model building and training.

4 Results and Discussion

Studies employing U-net for neurofibromas detection and classification have demonstrated promising results through analysing IOU index and loss during training epoch. Specifically, the loss of the training model has been gradually decreased to 0.003 whereas the IOU index has been increased and then reached nearly 95% over the training epochs as shown in figure 2.

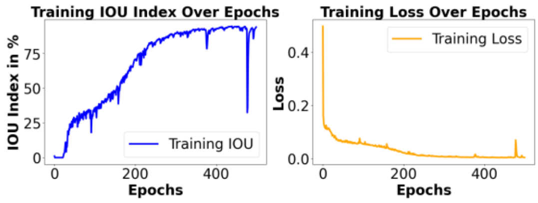

Fig 2. IOU and Loss of model during training

The study tested the model's performance with 63 image patches from the data of 4 different patients not included in the training dataset through using validation mode. Remarkably, in many cases, the IOU index exceeded 80%, as shown in figures 3.

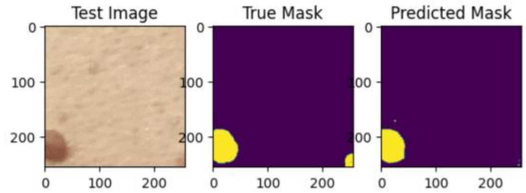

Fig 3. IOU of 84% with a test image with big size

In some cases, the IOU index is lower because the model successfully detects many NF1 instances with small sizes and blurry NF1 tumors as in figure 4 (*IOU of 69%)*, while the ground truth mask may not accurately represent these instances due to NF1 boundary issues or inaccurate labelling encountered during the manual labeling process. Therefore, the involvement of doctors or specialists in data labeling is very essential for improving result accuracy.

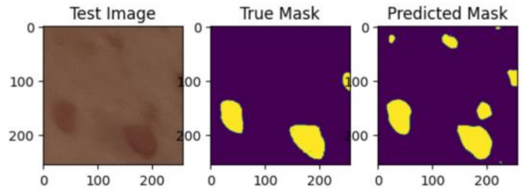

Fig 4. Lower IOU due to incorrect ground truth mask

The trained model achieved an average IOU of 52%, successfully detecting NF1 of various sizes and conditions. Expanding the diversity of datasets in both the training and testing sets is needed to enhance IOU, reducing the risk of model overfitting. Severity measurement in skin images is computed based on the number of NF1 detected for further treatment.

The findings demonstrated that the U-net architecture's capability to capture intricate details while maintaining spatial context effectively enables robust segmentation of neurofibromas, even in challenging cases with limited data and purposes aimed at detecting pixel-level targets. Furthermore, the modular nature of U-net allows for easy integration with other deep learning techniques, enabling further improvements in performance through ensemble methods or multi-scale analysis. For further research, ResNet or Swin Transformer could be explored as potential replacements for the encoder path.

5 Applications of the proposed model

Thanks to the patching procedure of the test image, the created patches (256x256 per patch) will be passed to the trained model to generate the predicted mask patches, indicating the detection of NF1 (or cNFs) for the image patches. Subsequently, through the unpatching process, these predicted patches will be reconstructed to match the original size of the test image. This approach ensures that the GUI can work with different test image sizes (multiples of 256x256 size) without a resizing step, thus maintaining the resolution of the original image for model processing. As can be seen in figure 5, this research proposed a GUI for NF1 detection and severity measurement in skin images, offering detailed information to doctors and specialists for better understanding and treatment selection, potentially reducing costs and improving diagnostic efficiency in healthcare.

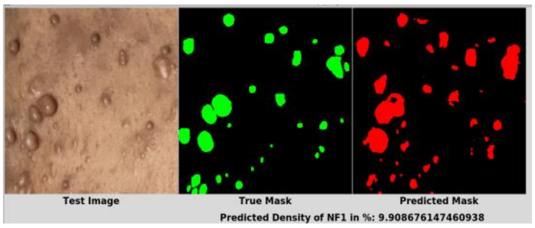

Fig 5. Graphical user interface for the NF1 detection and density measurement

Additionally, the model's application in research and clinical trials facilitates analysis of large medical image datasets, extracting insights on tumor biology and treatment responses, thereby aiding NF1 prediction efforts through early detection and personalized interventions.

6 Conclusion

In summary, the application of U-net-based deep learning techniques holds significant promise for enhancing the detection and classification of neurofibromas in medical imaging through biomedical segmentation tasks. Utilizing data patching and augmentation methods, U-net efficiently maximizes the utility of limited image datasets. With its unique architectural features and adaptability, U-net enables more accurate and efficient automated analysis of neurofibromas imaging data, potentially improving patient severity measurement and healthcare delivery.

References

[1] Robert J. Allaway, Sara J. C. Gosline, Salvatore La Rosa, Pamela Knight, Annette Bakker, Justin Guinney and Lu Q. Le, *"Cutaneous neurofibromas in the genomics era: current understanding and open questions"*, British Journal of Cancer. Available: https://doi.org/10.1038/s41416-018-0073-2.

[2] Mohammad Hesam Hesamian, Wenjing Jia, Xiangjian He, Paul Kennedy, *"Deep Learning Techniques for Medical Image Segmentation: Achievements and Challenges"*, Journal of Digital Imaging (2019) 32:582–596. Available: https://doi.org/10.1007/s10278-019-00227-x.

[3] Panqu Wang, Pengfei Chen, Ye Yuan, Ding Liu, Zehua Huang, Xiaodi Hou, Garrison Cottrell, TuSimple, Carnegie Mellon University, University of Illinois Urbana-Champaign, UC San Diego, *"Understanding Convolution for Semantic Segmentation", 2018 IEEE Winter Conference on Applications of Computer Vision*. Available: DOI 10.1109/WACV.2018.00163.

[4] Ruixin Yang and Yingyan Yu, "*Artificial Convolutional Neural Network in Object Detection and Semantic Segmentation for Medical Imaging Analysis*", REVIEW article, Sec. Cancer Imaging and Image-directed Interventions, Volume 11 - 2021 | https://doi.org/10.3389/fonc.2021.638182.

[5] Gokay Karayegen, Mehmet Feyzi Aksahin, "*Brain tumor prediction on MR images with semantic segmentation by using deep learning network and 3D imaging of tumor region*", Biomedical Signal Processing and Control, journal homepage: www.elsevier.com/locate/bspc.

[6] Di Yuan, Zhenghua Xu, Biao Tian, Hening Wang, Yuefu Zhan, Thomas Lukasiewicz, *"μ-Net: Medical image segmentation using efficient and effective deep supervision"*, Computers in Biology and Medicine, journal homepage: www.elsevier.com/locate/compbiomed.

[7] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, *"U-Net: Convolutional Networks for Biomedical Image Segmentation"*, Computer Science Department and BIOSS Centre for Biological Signalling Studies, University of Freiburg, Germany, WWW home page: http://lmb.informatik.uni-freiburg.de/.

[8] Xiao-Xia Yin, Le Sun, Yuhan Fu, Ruiliang Lu, and Yanchun Zhang, "*U-Net-Based Medical Image Segmentation*", Journal of Healthcare Engineering, Volume 2022, Article ID 4189781, 16 pages. Available: https://doi.org/10.1155/2022/4189781.