Demystifying the Black Box of Deep Neural Networks: Showcase on Medical Image Translation

<u> 邱維辰</u> 1

¹National Chiao Tung University

Abstract

Recent advances in deep learning have brought the magic leap to various research areas. Especially, with the growth of medical needs in the aging society, the integration of deep learning techniques into medical applications for assisting in the medical data analysis or imaging is one of the most important topics. Medical imaging, as the most prominent approach to the acquisition of spatially-resolved information of organs and tissues in vivo, has attracted great attempts to develop various deep-learning-based toolkits on processing the medical images. Particularly, the image-to-image translation across different medical image domains could provide additional diagnostic scans and benefit the clinical decision thus being an active research problem. However, the deep-learning-based translation models in themselves are inscrutable black boxes. We therefore propose to demystify the image translation process from a medical point of view, where we particularly focus on the translation from T1-MR images to PET ones, via adopting the representational similarity analysis. We discover that the process of T1-MR to PET image translation includes stages of recognition on various brain parts. Based on our findings, we build up an explainable model to demonstrate the capability of deep learning models for extracting medically meaningful information of brain imaging data, which untangles the biological plausibility hidden in deep learning models.

Keyword: Deep Learning, Image-to-Image Translation, Medical Images