DeepImageTranslator V2: analysis of multimodal medical images using semantic segmentation maps generated through deep learning

En Zhou Ye[1*], En Hui Ye[2*], Maxime Bouthillier[3*], and Run Zhou Ye[4*]

From the [1] Saint-Laurent Academy, Ottawa, Ontario, Canada. [2] Saint-Josephine High School, Ottawa, Ontario, Canada. [3] Département de radiologie, radio-oncologie et médecine nucléaire, Faculté de médecine, Université de Montréal, Montréal, Québec, Canada. [4] Division of Endocrinology, Department of Medicine, Centre de recherche du Centre hospitalier universitaire de Sherbrooke, Université de Sherbrooke, Sherbrooke, Québec, Canada

* These authors contributed equally to this work.

Corresponding author:

Run Zhou Ye

Division of Endocrinology, Department of Medicine, *Centre de recherche du Centre hospitalier universitaire de Sherbrooke*, *Université de Sherbrooke*, Sherbrooke, Quebec, Canada

3001, 12th Ave North, Sherbrooke, Quebec, Canada, J1H 5N4

Tel: 873-888-7979

E-mail: Run.Zhou.Ye@usherbrooke.ca

Running title: Tool for segmentation analysis of multimodal medical images

Keywords: multimodal medical imaging, semantic segmentation, regions of interest, deep learning, software, PET/CT imaging

ABSTRACT

Introduction: Analysis of multimodal medical images often requires the selection of one or many anatomical regions of interest (ROIs) for extraction of useful statistics. This task can prove laborious when a manual approach is used. We have previously developed a user-friendly software tool for image-to-image translation using deep learning. Therefore, we present herein an update to the DeepImageTranslator software with the addition of a tool for multimodal medical image segmentation analysis (hereby referred to as the MMMISA).

Methods: The MMMISA was implemented using the Tkinter library; backend computations were implemented using the Pydicom, Numpy, and OpenCV libraries. We tested our software using 4188 slices from whole-body axial 2-deoxy-2-[18F]-fluoroglucose-position emission tomography/computed tomography scans ([^{18}F]-FDG-PET/CT) of 10 patients from the American College of Radiology Imaging Network-Head and Neck Squamous Cell Carcinoma (ACRIN-HNSCC) database. Using the deep learning software DeepImageTranslator, a model was trained with 36 randomly selected CT slices and manually labelled semantic segmentation maps. Utilizing the trained model, all the CT scans of the 10 HNSCC patients were segmented with high accuracy. Segmentation maps generated using the deep convolutional network were then used to measure organ specific [^{18}F]-FDG uptake. We also compared measurements performed using the MMMISA and those made with manually selected ROIs.

Results: The MMMISA is a tool that allows user to select ROIs based on deep learning-generated segmentation maps and to compute accurate statistics for these ROIs based on coregistered multimodal images. We found that organ-specific [^{18}F]-FDG uptake measured using multiple manually selected ROIs is concordant with whole-tissue measurements made with segmentation maps using the MMMISA tool.

INTRODUCTION

Analysis of multimodal medical images (*e.g.*, position emission tomography/magnetic resonance imaging [PET/MRI] and PET/computed tomography [PET/CT]) often requires the selection of one or many anatomical regions of interest (ROIs) for extraction of useful statistics [1-6]. The use of spherical or ellipsoid ROIs may be sufficient for large organs such as the liver and large muscle groups. However, for organs/tissues with complex shapes (*e.g.*, the intestines and adipose tissues), manual ROI segmentation is not a scalable approach. One possible method is the use of deep learning for automated segmentation. Nevertheless, most deep learning pipelines for semantic image segmentation generate color-coded segmentation maps stored as image files, while most free software programs for medical image analysis (*e.g.*, 3D-Slicer, OsiriX Lite, and AMIDE) cannot use these files to generate ROI statistics of multimodal images stored as DICOM files.

We have previously developed a user-friendly software tool for image-to-image translation using deep learning (DeepImageTranslator, described in [7], released at: https://sourceforge.net/projects/deepimagetranslator/). Therefore, we present herein an update to the DeepImageTranslator software with the addition of a tool for multimodal medical image segmentation analysis (hereby referred to as the MMMISA). We then demonstrate the use of the program for the measurement of 2-deoxy-2-[18F]fluoroglucose ([^{18}F]-FDG) uptake by the lungs and subcutaneous adipose tissue using whole-body [^{18}F]-FDG-PET/CT scans from the ACRIN-HNSCC-FDG-PET/CT database [8-10]. Furthermore, we also compare measurements performed using the MMMISA and those made with manually selected ROIs.

METHODS

Development of the MMMISA program

The MMMISA program presented herein was written in Python 3.8 and distributed under the GNU General Public License (version 3.0). The graphical user interface was developed using the Tkinter library, which is the most commonly used Python package for graphical user interface creation. Image analysis algorithms were implemented using the Pydicom, Numpy, and OpenCV libraries. The program is included as part of version 2 of the DeepImageTranslator software (https://sourceforge.net/projects/deepimagetranslator/) and is also available as a standalone program (https://sourceforge.net/projects/mmmisa/) for Windows. The source codes are available at: (https://github.com/runzhouye/MMMISA).

PET/CT image dataset

Whole-body CT and FDG-PET images from 10 patients (numbers 001, 002, 003, 007, 008, 010, 012, 018, 019, and 027—which were the first 10 whole-body scans with the same windowing and devoid of significant radiographic artifacts) were downloaded from the ACRIN-HNSCC-FDG-PET/CT (ACRIN 6685) database [8, 9] via the Cancer Institute Archive [10]. We arbitrarily chose the ACRIN-HNSCC-FDG-PET/CT database since it was one of the few databases that contain coregistered whole-body PET and CT scans.

Manual extraction of multimodal image data

CT and FDG-PET images were loaded into the AMIDE software [11]. For each patient, 11 spherical ROIs (10 mm diameter) in the subcutaneous adipose tissue and 3 spherical ROIs (50 mm

diameter) in the lungs were drawn at different axial positions based on whole-body CT images. ROI statistics were subsequently generated for the coregistered PET images.

Semantic image segmentation

Thirty-six axial slices were randomly chosen from the 4188 axial CT images from the 10 patients for manual semantic segmentation with the GIMP (GNU Image Manipulation Program) software of the background, lungs, bones, brain, subcutaneous and visceral adipose tissue, and other soft tissues by labelling these regions in black (RGB=[0,0,0]), yellow (RGB=[255,255,0]), white (RGB=[255,255,255]), cyan (RGB=[0,255,255]), red (RGB=[255,0,0]), green (RGB=[0,255,0]), and blue (RGB=[0,0,255]). Thirty-six training samples were considered more than sufficient since we have previously shown that models can be trained to achieve high accuracy with as little as 17 images [12]. CT image-segmentation map pairs were then loaded into the DeepImageTranslator software to train a deep convolutional neural network as previously described in Ye et al. [12] with 1000 training epochs. The final model was used to perform automatic semantic segmentation of the 4188 axial CT images from the 10 patients in less than 10 minutes. The generalizability of such an approach for automated segmentation has previously been shown to be excellent [12].

Automated extraction of multimodal image data

For each patient, the original PET/CT scans were loaded into the MMMISA program along with the semantic segmentation maps produced by the convolutional neural network. In this study, we chose to extract FDG uptake from the lungs and subcutaneous adipose tissue by extracting regions of the model-generated segmentation maps containing yellow and red pixels, respectively using the MMMISA program. Lower and upper color threshold were set at (R,G,B) = (150,150,0) and

(R,G,B) = (255,255,150), respectively, for the lungs, and (R,G,B) = (150,0,0) and (R,G,B) = (255,150,150), respectively, for the subcutaneous adipose tissue. ROI statistics were then generated for the FDG-PET scans using the MMMISA software.

Statistical analyses

Statistical analyses were carried out using GraphPad Prism version 9. Pearson's R values were computed for the correlation between organ-specific FDG uptake measured using multiple manually selected ROIs and FDG uptake determined using deep learning-generated segmentation maps.

Data availability

The source code for the DeepImageTranslator is publicly available at:

https://github.com/runzhouye/MMMISA

The compiled standalone software is available for Window10 at:

https://sourceforge.net/projects/deepimagetranslator/ and https://sourceforge.net/projects/mmmisa/

The datasets generated during and/or analyzed during the current study are available at:

https://doi.org/10.6084/m9.figshare.16800925

RESULTS

The MMMISA plugin for the DeepImageTranslator

The MMMISA program is included in version 2 of the DeepImageTranslator and is also available as a standalone software. The main window (**Fig.1**) allows for the user to visualize single- and dual-modality images written in the standard DICOM (Digital Imaging and Communications in Medicine) file format, the most commonly used file format in medical imaging. When images from a second modality are loaded into the program, they are automatically matched, along with the corresponding segmentation map, to the image of the first modality that is being currently displayed. When necessary, the program also performs image registration of modality 2 images based on modality 1 images through translation and/or stretching such that objects in both image sets overlap. This allows for simultaneous visualization of both image sets and segmentation maps.

A second, ROI selection, window (**Fig.2**) displays user settings for the extraction of ROIs based on pixel colors of the segmentation maps. Specific regions of the color-coded semantic segmentation maps can be extracted by setting lower and upper thresholds for the red, green, and blue color component values using the ROI selection window. The user can also choose to only include the left or right side of the patient for analysis, which can be useful in order to exclude the strong signals from of certain radiotracers injected into the left or right arm. When the "Apply" button is pressed, ROIs are generated based on the color thresholds using the segmentation maps and applied to corresponding slices of modality 1 and 2 images. The cropped images are then displayed in the main window for visualization.

When "Save analysis" is selected, data will be extracted from modality 1 and 2 images, including the name of the scan, time at which each slice was produced, position of image slices, total area of the ROIs on each slice, total pixel values in the ROIs, average and standard deviation of values of pixels inside the ROIs, and pixel size. Results are then written in an excel file and stored under the user-designated directory.

Semantic segmentation of PET/CT images

Segmentation results for images outside of the training set obtained with the convolutional neural network trained using the DeepImageTranslator were illustrated in **Fig.3**. Our final model was able to accurately segment the lungs, brain, bone, subcutaneous and visceral adipose tissue, and other soft tissues.

Increase in number of manually selected ROIs increases accuracy of organ-specific FDG uptake approximations compared to true organ-specific FDG uptake measured using deep learning-generated segmentation maps

Next, we tested the concordance of organ-specific FDG uptake measured using multiple manually selected ROIs *versus* FDG uptake determined using deep learning-generated segmentation maps. In general, regardless the number of ROIs used, manually measured FDG uptake in the lungs and subcutaneous adipose tissue was well correlated with that calculated with segmentation maps using the MMMISA program (**Fig.4**). For subcutaneous adipose tissue FDG uptake, the correlation coefficient and the -log of the P-value increased sharply once values from more than 4 ROIs were combined (**Fig.4A**). Increase in measurement accuracy (determined by the correlation coefficient) through increasing numbers of manually selected ROIs plateaued after more than 8 ROIs were used. Nevertheless, the P-value of the correlation between manual measurement and that using segmentation maps continued to decrease when more ROIs were used (**Fig.4B**). Similar results were obtained for the measurement of FDG uptake in the lungs (**Fig.4C-D**).

DISCUSSION

In recent years, numerous open-source software tools have been reported in the field of medical image processing [13-17]. One growing area of development is the popularization of deep learning methods through the creation of user-friendly tools with a graphical interface. Nevertheless, most deep learning pipelines for semantic image segmentation generate color-coded segmentation maps stored as image files, while most free software programs for medical image analysis cannot use these files to generate ROI statistics of multimodal images stored as DICOM files.

Nonetheless, selection of ROIs is an important aspect of *in vivo* metabolic studies involving PET/CT imaging [18-21]. In particular, measurements of volume and radiotracer uptake of adipose tissues of different regions may prove to be important for future studies on the metabolic syndrome, as hypertrophic obesity is related to changes in adipose tissue distribution and alterations in metabolic endpoints [22, 23].

Therefore, we have presented herein an update to the DeepImageTranslator software [7] by including a tool for multimodal medical image segmentation analysis based on semantic segmentation maps generated using a deep convolutional neural network. Our program can be accessed through a graphical interface and allows users to extract ROI statistics of multimodal images (*e.g.*, PET/CT and PET/MRI) based on color-coded semantic segmentation maps. We showed that organ-specific FDG uptake measured using multiple manually selected small, spherical ROIs is concordant with whole-tissue measurements made with segmentation maps using the MMMISA program. Furthermore, we found that increase in number of manually selected ROIs increases the accuracy of organ specific FDG uptake approximations. Therefore, our pipeline constitutes a simple, automated, and scalable approach to obtain ROI statistics using multimodal scans. Although the accuracy of the neural network would never surpass the accuracy of manually labelled segmentation maps used for model training, our approach would eventually greatly

simplify the task of researchers and radiologists performing whole-body semantic segmentation of multimodal tomography data.

DISCLOSURES

The authors declare no competing interests.

AUTHOR CONTRIBUTIONS

Software development: EZY, EHY, and RZY. Statistical analyses: EHY. Interpretation and manuscript drafting: EZY, MB, and RZY.

REFERENCES

1. Lensen, K.D.F., et al., *Variability in quantitative analysis of atherosclerotic plaque inflammation using 18F-FDG PET/CT.* PLoS One, 2017. **12**(8): p. e0181847.
2. Cao, Y., et al., *Subcutaneous adipose tissue free fatty acid uptake measured using positron emission tomography and adipose biopsies in humans.* Am J Physiol Endocrinol Metab, 2019. **317**(2): p. E194-E199.
3. Noll, C., et al., *Seven-day overfeeding enhances adipose tissue dietary fatty acid storage and decreases myocardial and skeletal muscle dietary fatty acid partitioning in healthy subjects.* Am J Physiol Endocrinol Metab, 2020. **318**(2): p. E286-E296.
4. Hojjati, M., et al., *Role of FDG-PET/MRI, FDG-PET/CT, and Dynamic Susceptibility Contrast Perfusion MRI in Differentiating Radiation Necrosis from Tumor Recurrence in Glioblastomas.* J Neuroimaging, 2018. **28**(1): p. 118-125.
5. Montastier, E., et al., *Increased postprandial nonesterified fatty acid efflux from adipose tissue in prediabetes is offset by enhanced dietary fatty acid adipose trapping.* Am J Physiol Endocrinol Metab, 2021. **320**(6): p. E1093-E1106.
6. Labbe, S.M., et al., *Increased myocardial uptake of dietary fatty acids linked to cardiac dysfunction in glucose-intolerant humans.* Diabetes, 2012. **61**(11): p. 2701-10.
7. Ye, R.Z., et al., *DeepImageTranslator: a free, user-friendly graphical interface for image translation using deep-learning and its applications in 3D CT image analysis*. bioRxiv, 2021.
8. Kinahan, P., Muzi, M., Bialecki, B., & Coombs, L. (2019). Data from the ACRIN 6685 Trial HNSCC-FDG-PET/CT [Data set]. TCIA. https://doi.org/10.7937/K9/TCIA.2016.JQEJZZNG.
9. Lowe, V.J., Duan, F., Subramaniam, R. M., Sicks, J. D., Romanoff, J., Bartel, T., Yu, J. Q. (Michael), Nussenbaum, B., Richmon, J., Arnold, C. D., Cognetti, D., & Stack, B. C., Jr. (2019). Multicenter Trial of [18F]fluorodeoxyglucose Positron Emission Tomography/Computed Tomography Staging of Head and Neck Cancer and Negative Predictive Value and Surgical Impact in the N0 Neck: Results From ACRIN 6685. Journal of Clinical Oncology, 37(20), 1704–1712. https://doi.org/10.1200/jco.18.01182.
10. Clark, K., Vendt, B., Smith, K., Freymann, J., Kirby, J., Koppel, P., Moore, S., Phillips, S., Maffitt, D., Pringle, M., Tarbox, L., & Prior, F. (2013). The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository. Journal of Digital Imaging, 26(6), 1045–1057. https://doi.org/10.1007/s10278-013-9622-7.
11. Loening, A.M. and S.S. Gambhir, *AMIDE: a free software tool for multimodality medical image analysis.* Mol Imaging, 2003. **2**(3): p. 131-7.
12. Ye, R.Z., et al., *DeepImageTranslator: A free, user-friendly graphical interface for image translation using deep-learning and its applications in 3D CT image analysis.* SLAS Technol, 2022. **27**(1): p. 76-84.
13. Lessard, R., et al., *An open-source software for monitoring intrafraction motion during external beam radiation therapy based on superimposition of contours of projected ROIs on cine-MV images.* J Appl Clin Med Phys, 2020. **21**(8): p. 173-182.
14. Yushkevich, P.A., et al., *User-Guided Segmentation of Multi-modality Medical Imaging Datasets with ITK-SNAP.* Neuroinformatics, 2019. **17**(1): p. 83-102.
15. Bucking, T.M., et al., *From medical imaging data to 3D printed anatomical models.* PLoS One, 2017. **12**(5): p. e0178540.
16. Debus, C., et al., *MITK-ModelFit: A generic open-source framework for model fits and their exploration in medical imaging - design, implementation and application on the example of DCE-MRI.* BMC Bioinformatics, 2019. **20**(1): p. 31.

17. Razeghi, O., et al., *CemrgApp: An interactive medical imaging application with image processing, computer vision, and machine learning toolkits for cardiovascular research.* SoftwareX, 2020. **12**: p. 100570.
18. Croteau, E., et al., *Image-derived input function in dynamic human PET/CT: methodology and validation with 11C-acetate and 18F-fluorothioheptadecanoic acid in muscle and 18F-fluorodeoxyglucose in brain.* Eur J Nucl Med Mol Imaging, 2010. **37**(8): p. 1539-50.
19. Lee, S.Y., et al., *Comparison of 3T diffusion-weighted MRI and (18)F-FDG PET/CT in musculoskeletal tumours: quantitative analysis of apparent diffusion coefficients and standardized uptake values.* Br J Radiol, 2019. **92**(1102): p. 20181051.
20. Richard, G., et al., *Contribution of perfusion to the (11) C-acetate signal in brown adipose tissue assessed by DCE-MRI and (68) Ga-DOTA PET in a rat model.* Magn Reson Med, 2021. **85**(3): p. 1625-1642.
21. Carreau, A.M., et al., *Bariatric Surgery Rapidly Decreases Cardiac Dietary Fatty Acid Partitioning and Hepatic Insulin Resistance Through Increased Intra-abdominal Adipose Tissue Storage and Reduced Spillover in Type 2 Diabetes.* Diabetes, 2020. **69**(4): p. 567-577.
22. Ye, R.Z., et al., *Fat Cell Size: Measurement Methods, Pathophysiological Origins, and Relationships with Metabolic Dysregulations.* Endocr Rev, 2021.
23. Carpentier, A.C., *100(th) anniversary of the discovery of insulin perspective: insulin and adipose tissue fatty acid metabolism.* Am J Physiol Endocrinol Metab, 2021. **320**(4): p. E653-E670.

FIGURES

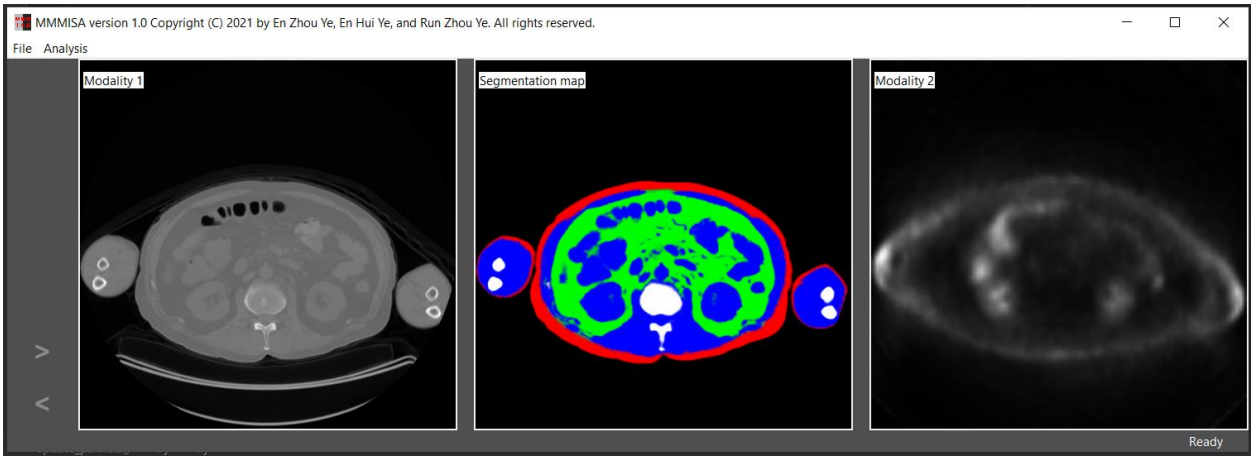

Fig.1: Main window of MMMISA, showing (from left to right), modality 1 (CT) images, segmentation maps generated with convolutional neural network, and modality 2 (PET) images.

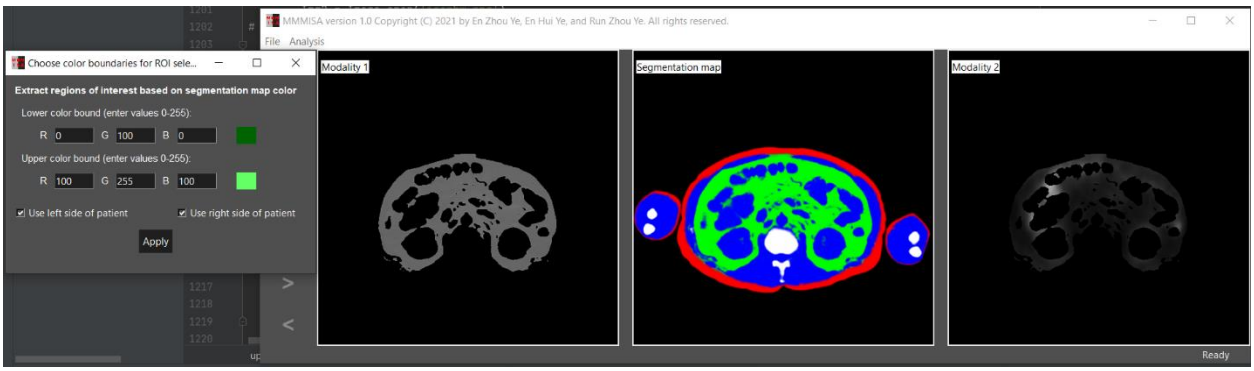

Fig.2: ROI selection window and main window with updated modality 1 and 2 images.

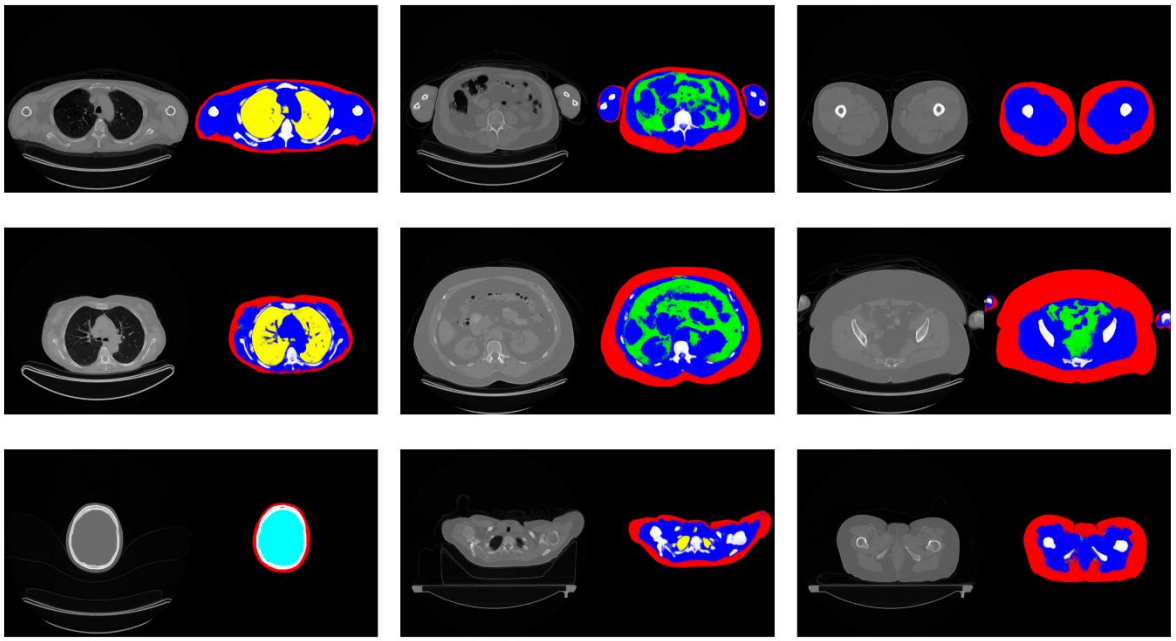

Fig.3: Pairs of CT images outside of the training set and sematic segmentation maps generated with deep convolutional neural network. The background, lungs, bones, brain, subcutaneous and visceral adipose tissue, and other soft tissues were labelled in black, yellow, white), cyan, red, green, and blue, respectively.

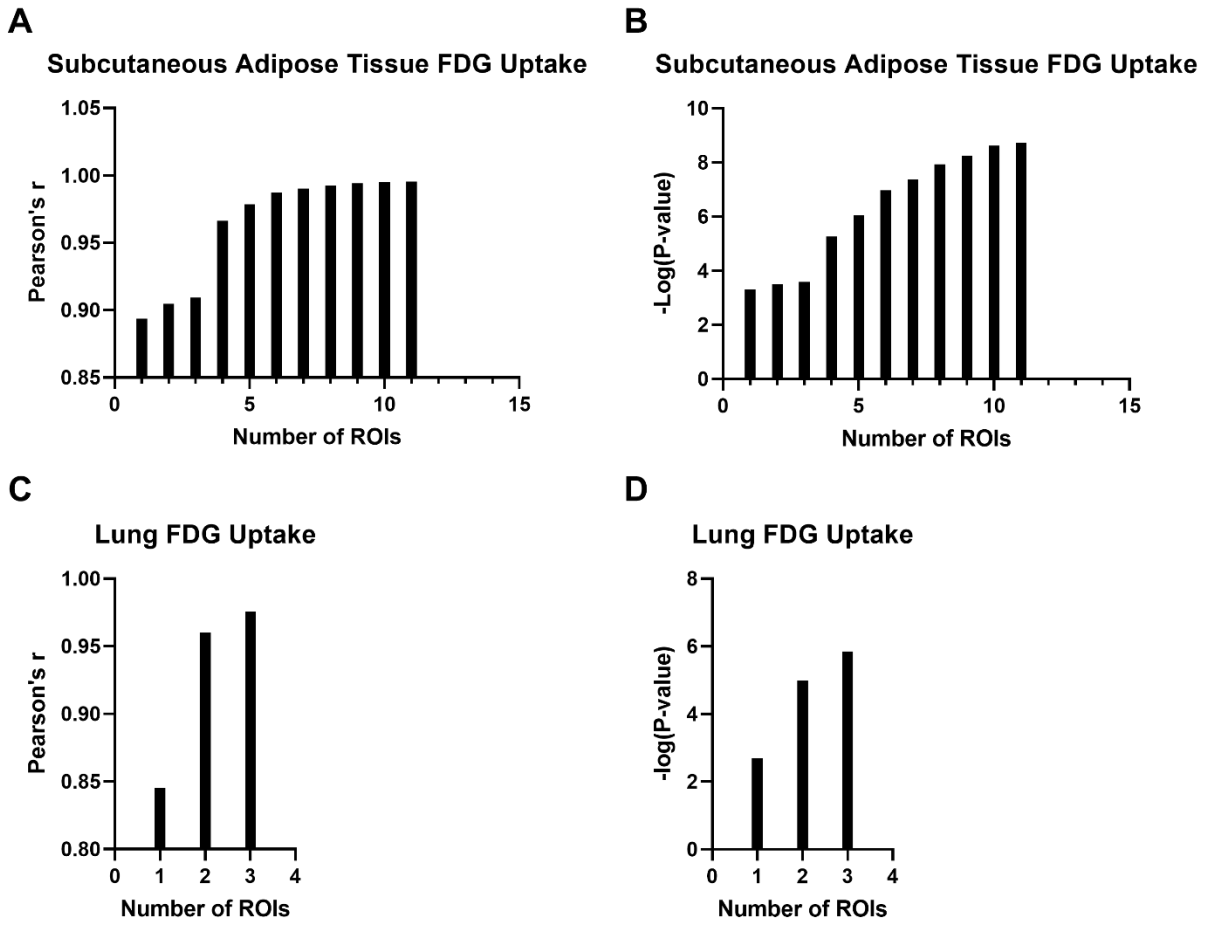

Fig.4: Correlation coefficient (**A** and **C**) and P-value (**B** and **D**) for the association between organ-specific FDG uptake measured using multiple manually selected ROIs and FDG uptake determined using deep learning-generated segmentation maps, as a function of number of manually selected ROIs, for the subcutaneous adipose tissue (**A** and **B**) and lungs (**C** and **D**). ROI: region of interest.