Physics in Medicine & Biology

EDITORIAL

Focus on machine learning models in medical imaging

IOP Publishing Phys. Med. Biol. 68 (2023) 010301 https://doi.org/10.1088/1361-6560/aca069

Physics in Medicine & Biology

CrossMark

PUBLISHED
16 December 2022

EDITORIAL

Focus on machine learning models in medical imaging

Giorgos Papanastasiou[1], **Alba García Seco de Herrera**[1], **Chengjia Wang**[2], **Heye Zhang**[3], **Guang Yang**[4] and **Ge Wang**[5]

[1] University of Essex, United Kingdom
[2] University of Edinburgh, United Kingdom
[3] Sun Yat-sen University, People's Republic of China
[4] Imperial College London, United Kingdom
[5] Rensselaer Polytechnic Institute, United States of America

Medical imaging is a key component in clinical diagnosis and treatment optimization, as well as clinical trial design, accounting for almost 90% of all healthcare data (Beutel *et al* 2000, Kevin Zhou *et al* 2020, Shukla-Dave A *et al* 2019). Medical image analysis techniques aim to enhance the diagnostic performance of visual assessments in medical imaging, improving the early diagnosis of various diseases and helping to obtain a deeper understanding of physiology and pathology (Beutel *et al* 2000, Kevin Zhou *et al* 2020).

Medical image computing is an important aspect of computational medical imaging, as it aims to directly measure (through end-to-end approaches) or to indirectly guide (via segmentation, synthesis, denoising, reconstruction, etc) the extraction of clinically useful imaging biomarkers (Kevin Zhou *et al* 2020, Shukla-Dave A *et al* 2019). Recently, convolutional neural networks (CNN) achieved performance gains across numerous medical image computing tasks (Cheplygina *et al* 2019, Duncan *et al* 2019, Yi *et al* 2019, Chen *et al* 2021, Kevin Zhou *et al* 2020). CNNs can efficiently model local pixel interactions and be trained on either big- or small-scale data (the latter was made possible through data augmentation and other pre-processing techniques) (Yi *et al* 2019, Kevin Zhou *et al* 2020). CNN models revolutionised multiple tasks in medical image computing, such as image segmentation, registration, reconstruction, denoising, synthesis and pathology detection/evaluation, through extensive analysis of medical imaging data (Cheplygina *et al* 2019, Duncan *et al* 2019, Yi *et al* 2019, Haskins *et al* 2020, Kevin Zhou *et al* 2020).

Depending on the medical image analysis application, the CNN architectures may vary substantially (Kevin Zhou *et al* 2020). In this special issue named 'Focus on machine learning models in medical imaging', we aimed to select diverse medical imaging modalities and downstream clinical applications and encourage the collection of different CNN architectures and methodologies. This special issue collection can be a potentially important and useful material for future surveys and reviews. It can also be useful towards developing checklists for artificial intelligence criteria in medical imaging (CLAIM) and a logic (rule-set) for artificial intelligence (AI) algorithm design in medical imaging (Mongan *et al* 2020), according to the modality and downstream clinical application. The special issue encompasses fifteen peer-reviewed articles covering different imaging modalities and model architectures.

The first paper of the special issue by Cui *et al* is a study on medical image registration which is the task that aims to learn the deformation field which aligns a moving medical image to a reference image (from the same or another imaging modality), to support extracting complementary information and improve medical image diagnosis (Kunpeng *et al* 2022). The authors engineered a singular value decomposition (SVD) denoising layer on a U-net architecture. The SVD component helped the registration model to learn low-level image representations and optimally removed noise thus, enhancing registration accuracy. The authors also demonstrated that an exponential linear unit (ELU) activation function was robust to noise against other activation function options. By maintaining comparable registration times, the authors showed that their registration model performed higher registration accuracy, compared to state-of-the-art (SOTA) methods.

The second paper by Zou *et al* proposed an unsupervised motion-compensated model using smoothness regularization on motion deformation maps derived from different time points (images) for the reconstruction of high-resolution free-breathing dynamic lung magnetic resonance imaging (MRI) (Qing *et al* 2022). MRI reconstruction through machine learning techniques aims to accelerate image acquisition without losing spatiotemporal information, which is critical for clinical imaging. The authors assume the deformation maps to be the output of a CNN-based generator, whose inputs are time-dependent low-dimensional latent vectors which capture the motion information. Multichannel nonuniform Fourier transformation is recruited to generate the k-space measurements of the images. Unlike previous techniques, the authors formulate the joint recovery of the latent vectors, deformation maps, and the template image directly from the measured k-t space data, as a single nonlinear optimization scheme. The authors demonstrate that their technique is more accurate than previous SOTA methods in addressing bulk motion events during the scan, which translates to less blurred reconstructions in MRI datasets with extensive motion. The proposed approach may be particularly useful for paediatric and neonatal patients for which MRI acquisition is considerably challenged by bulk motion.

The third paper by Karkalousos *et al* also addresses the concept of MRI data reconstruction on T1-weighted and FLAIR brain and T2-weighted knee data (both static MRI) (Karkalousos *et al* 2022). The authors proposed a technique that combines recurrent inference machines (RIM) blocks sequentially connected through cascades, with the Independently Recurrent Neural Network (IndRNN) model being the backbone recurrent unit. The cascades allowed to train a deep but balanced RNN for improved denoising and de-aliasing, while maintaining stable gradient calculations. The proposed technique was compared against several CNN SOTA methods and conventional Compressed Sensing reconstruction. The CIRIM performed best on all training datasets, undersampling schemes tested and acceleration factors. Also, it demonstrated robust performance on reconstructing accelerated FLAIR data containing multiple sclerosis lesions, reaching good lesion contrast and efficient denoising compared to previous SOTA methods.

In the fourth paper, Rutherford *et al* developed a CNN-based model on the Inception-ResNet and InceptionTime architectures for *in vivo* 1D dose estimation using in-beam positron emission tomography (PET) in carbon ion therapy (Harley *et al* 2022). Their method was initially evaluated using Geant4 Monte Carlo simulations, and then demonstrated with experimental data acquired at the Heavy Ion Medical Accelerator in Chiba (HIMAC), Japan, using a small, preclinical in-beam PET scanner prototype. The CNN performance was compared with a recent iterative optimisation-based technique for dose estimation of ^{12}C ion spread-out Bragg peak profiles developed by the same group. The Inception CNN network outperformed the iterative approach for all simulated beam deliveries, in addition to estimating beam configurations which fell far outside of the training dataset (out-of-distribution ability). The author stated that they currently extend their Inception CNN framework to perform 3D dose estimations from 4D PET radionuclide distributions.

Rezaeifar *et al* developed a dual-CNN method involving a pre-training and a subsequent transfer learning (fine-tuning) step, to accurately predict the centre of mass (CoM) of glioblastoma tumours with different geometries, from Monte Carlo-simulated bioluminescence tomography imaging data (Behzad *et al* 2022). Their technique can open new horizons for improving pre-clinical image-guided radiotherapy and even help biologists investigate efficiently the tumour response to treatment with bioluminescence markers. The proposed model can predict the tumor's CoM with submillimeter accuracy, for tumours $>10\ \mathrm{mm}^3$. CNN-based BLI targeting may also reduce the planning time compared to physics model-based counterparts. Since this paper mainly focused on developing the Monte Carlo simulations to generate the necessary training database, it sets the framework for further studies in this field.

In the sixth paper, Jing *et al* proposed a self-supervised methodology for low-dose computed tomography (LDCT) denoising, composed of a RED-CNN backbone model equipped with self-attention mechanisms, and by recruiting perceptual and self-supervised losses (Jie *et al* 2022). Unlike previous SOTA methods, the proposed architecture did not require the use of labelled and normal-dose paired (registered) computed tomography (CT) images for LDCT denoising. The proposed approach was compared against supervised and unsupervised SOTA models and showed strong denoising ability without any high-quality reference data. Ablation studies proved the additive value of the perceptual loss and the self-attention block. The authors suggest that the perceptual loss and attention block guided the model to pay more attention to structural details, rather than just focusing on high-level pixel similarities in abdominal and cardiac LDCT data denoising.

Liu *et al* developed a two-step deep learning method which first employed a U-net structure to learn the cone beam CT (CBCT) to CT mapping using phantom images (Yuxiang *et al* 2022). In the second step, a CycleGAN model was recruited leveraging information from the well-trained U-net model to generate accurate CT images (named as 'intermediate CT images') from patient CBCT images. In the CycleGAN model, the style of the patient CBCT was transformed into the intermediate CT image generated using the phantom model to finally produce a synthetic CT image, that is close to real patient data. CBCT images are commonly used to perform radiotherapy planning; however, radiotherapy planning is commonly compromised by low CBCT image quality. The proposed two-step method effectively improved the CBCT image quality to the level of CT scans. It outperformed the SOTA baseline methods (U-net and CycleGAN alone) for region-of-interest contouring and HU calibration, which are critical for radiotherapy planning applications. The authors therefore propose an efficient method for CBCT denoising that is promising for future radiotherapy planning.

In the eighth paper, Jones *et al* developed a computer-aided diagnosis approach that combined handcrafted feature extraction and automated feature extraction using a VGG-16 model-based transfer learning method (Meredith *et al* 2022). This approach aimed to subsequently guide SVM-based models to classify lesions as benign or malignant from full-field digital mammography images. The authors suggest that this first-of-its-kind approach combining handcrafted and automated feature extractions has merit to enhance the clinical imaging diagnosis and decision making. The authors discuss several methodology improvements to follow up.

Liang *et al* proposed first data-driven CNN-based inverse Monte Carlo (MC) model, to estimate the functional form of the phase function in the MC simulations of eleven typical biological tissues with known biomedical optic properties (Yuxuan *et al* 2022). The backbone structure of the CNN model was composed of a ResNet-18 model adaptation. The authors demonstrated that larger field-of-view and higher spatial resolution provide more information about the phase function. The authors set the foundation for extended research into the sub-diffusive domain and for future *in vivo* measurements using biomedical optics.

Orlando *et al* developed a SOTA U-net++ model involving a ResNet-50 backbone, that provided fast (<1.5 s) and accurate 3D prostate segmentations across clinically diverse 3D transrectal ultrasound (TRUS), used in prostate cancer diagnosis and treatment (Nathan *et al* 2022). A 3D TRUS image quality grading scale with three factors (acquisition quality, artifact severity, and boundary visibility) was also developed to assess the impact on segmentation performance. The authors showed that the proposed technique demonstrated generalizability and efficiency when employed on smaller datasets, supporting the potential for widespread use even in cases where data are scarce. The authors also suggested that the development of an image quality grading scale provided a quantitative tool for assessing segmentation performance, which can overall support medical image segmentation tasks.

In the eleventh paper, Lappas *et al* developed a 3D U-net model for delineation of relevant normal tissues, for image-guided precision radiotherapy workflows in small animals (Georgios *et al* 2022). Standardized 3D micro CBCT (μCBCT) volumes were used as inputs. The authors aimed to develop a fully automatic, generalizable method for normal tissue contouring in preclinical studies. The 3D U-net was trained to contour organs in the head (whole brain, left/right brain hemisphere, left/right eye) and thorax (complete lungs, left/right lung, heart, spinal cord, thorax bone) regions. In the pre-processing, Hounsfield units (HUs) were converted to mass density (MD) values, to remove the energy dependency of the μCBCT scanner. The authors demonstrated accurate segmentation results with low variability, suggesting an accurate and generalizable method for multi-organ segmentation from μCBCT data.

Iyer et al developed a sequential DeepLabV3+ -based ensemble for localization and segmentation of organs at risk, to be considered and carefully protected during radiotherapy planning (Aditi *et al* 2022). The proposed method outperformed a series of previous SOTA methods utilized as baseline models. The sequential localization strategy aided in segmenting structures with complex morphology and low soft-tissue contrast. Multi-view ensemble models were found to improve the worst-case segmentation errors and could potentially be applied to improve segmentation quality in other sites. The authors therefore demonstrated a fully-automatic, accurate, and time-efficient method to segment swallowing and chewing structures in CT images and demonstrated its suitability for clinical use.

In the thirteenth paper, Guo *et al* proposed an approach that is based on ordinal regression for stage-sensitive classification of COVID-19 CT images, aiming to help clinical doctors assess the status of COVID-19 patients and find an optimal treatment plan (Xiaodong *et al* 2021). The model was composed of a modified ResNet backbone used to extract imaging features, followed by three different ordinal regression methods (multi-binary, neuro-stick breaking method, soft-labels method) to reach independent predictions. The authors showed that their ensemble learning method performed better than the traditional classification algorithm, therefore suggesting that their approach has the potential to suggest guidelines for staging COVID-19 chest CT images. The authors stated that they work towards incorporating other clinical information next to CT, to reach more accurate diagnosis in the future.

Chen *et al* developed an improved CycleGAN involving two adversarial discriminators to generate the synthetic kilovoltage CT (skV-CT) from megavoltage CT (MV-CT). The authors then proposed a two-branch network that extracts features from both skV-CT and MV-CT images for segmentation (Xinyuan *et al* 2021). Their study was the first to combine skV-CT and MV-CT for tumour segmentation in helical tomotherapy. MV-CT is used for adaptive radiotherapy in tomotherapy. However, this process is compromised by low contrast and high noise that leads to poor image quality. This study aimed to develop a deep-learning-based method to generate synthetic kilovoltage CT (skV-CT) and then evaluate its ability to improve image quality and tumour segmentation. The results demonstrated that the proposed method significantly improved image quality and tumour segmentation and can potentially be implemented for adaptive radiotherapy.

In the last paper, Olberg *et al* proposed a GAN model consisted of a generator and a discriminator, to learn synthetic CT from MRI, across multiple abdominal organs, in the setting of MR-guided radiotherapy planning (Sven *et al* 2021). The authors demonstrate that when well-matched MRI and CT images exist for each patient in the training dataset in terms of abdominal geometry (demonstrating little involvement of intestinal gas), the model could establish dosimetric accuracy. However, when intestinal gas is involved, this poses considerable complications for MR-guided dosimetry calculations, as the MR to CT mapping is substantially compromised.

We would like to thank all the authors who submitted their work to our special issue and contributed to advance the field of machine learning in medical imaging through their studies, particularly during the challenging times of COVID-19. We are grateful for the reviewers and editors who maintained high publication standards and ensured the quality of the published work. Last not least, we would like to thank the IOP journal of Physics in Medicine and Biology for their support in the publication of this special issue.

ORCID iDs

Chengjia Wang https://orcid.org/0000-0003-2345-7364
Guang Yang https://orcid.org/0000-0001-7344-7733
Ge Wang https://orcid.org/0000-0002-2656-7705

References

Aditi I *et al* 2022 *Phys. Med. Biol.* **67** 024001
Behzad R *et al* 2022 *Phys. Med. Biol.* **67** 144003
Beutel J, Kundel H L and Van Metter R L 2000 *Handbook of Medical Imaging* (Bellingham, WA, USA: SPEI Press) 1
Chen X, Diaz-Pinto A, Ravikumar N and Frangi A F 2021 Deep learning in medical image registration: a survey *Prog. Biomed. Eng.* **3** 012003
Jie J *et al* 2022 *Phys. Med. Biol.* **67** 084002
Cheplygina V, de Bruijne M and Pluim J P W 2019 Not-so-supervised: a survey of semi-supervised, multi-instance, and transfer learning in medical image analysis *Med. Image Anal.* **54** 280–96
Duncan J S, Insana M F and Ayache N 2019 Biomedical imaging and analysis in the age of big data and deep learning *Proc. IEEE* **108** 3–10
Georgios L *et al* 2022 *Phys. Med. Biol.* **67** 044001
Harley R *et al* 2022 *Phys. Med. Biol.* **67** 194001
Karkalousos D *et al* 2022 *Phys. Med. Biol.* **67** 124001
Kevin Zhou S *et al* 2020 A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises *Proc. IEEE* **109** 820–38
Kunpeng C *et al* 2022 *Phys. Med. Biol.* **67** 144002
Meredith A J *et al* 2022 *Phys. Med. Biol.* **67** 054001
Mongan J, Moy L and Kahn C E 2020 Checklist for artificial intelligence in medical imaging (CLAIM): a guide for authors and reviewers *Radiol. Artif. Intell.* **2** e200029
Nathan O *et al* 2022 *Phys. Med. Biol.* **67** 074002
Qing Z *et al* 2022 *Phys. Med. Biol.* **67** 144001
Shukla-Dave A *et al* 2019 Quantitative imaging biomarkers alliance (QIBA) recommendations for improved precision of DWI and DCE-MRI derived biomarkers in multicenter oncology trials *JMRI* **49** e101–21
Sven O *et al* 2021 *Phys. Med. Biol.* **66** 204001
Xiaodong G *et al* 2021 *Phys. Med. Biol.* **66** 244001
Xinyuan C *et al* 2021 *Phys. Med. Biol.* **66** 224001
Yi X, Walia W and Babyn P 2019 Generative adversarial network in medical imaging: a review *Med. Image Anal.* **58** 101552
Yuxiang L *et al* 2022 *Phys. Med. Biol.* **67** 084001
Yuxuan L *et al* 2022 *Phys. Med. Biol.* **67** 074001