Tensor Radiomics: Paradigm for Systematic Incorporation of Multi-Flavoured Radiomics Features

Arman Rahmim, Senior Member, IEEE, Amirhossein Toosi, Mohammad R. Salmanpour, Natalia Dubljevic, Ian Janzen, Isaac Shiri, Mohamad A. Ramezani, Ren Yuan, Cheryl Ho, Habib Zaidi, Fellow, IEEE, Calum MacAulay, Carlos Uribe, Fereshteh Yousefirizi

Abstract—Radiomics features extract quantitative information from medical images, towards the derivation of biomarkers for clinical tasks such as diagnosis, prognosis, or treatment response assessment. Different image discretization parameters (e.g. bin number or size), convolutional filters, segmentation perturbation, or multi-modality fusion levels can be used to generate radiomics features and ultimately signatures. Commonly, only one set of parameters is used, resulting in only one value or ‘flavour’ for a given feature. We propose ‘tensor radiomics’ (TR) where tensors of features calculated with multiple combinations of parameters (i.e. flavours) are utilized to optimize the construction of radiomics signatures. We have applied TR to PET/CT, MRI, and CT imaging, invoking machine learning or deep learning solutions, and reproducibility analyses. Given space limitations, here we present example results on PET/CT imaging: (1) TR via varying bin sizes on PET-CT images of head & neck cancer (HNC) for overall survival prediction. A hybrid deep neural network, referred to as ‘TR-Net’, along with two ML-based flavour fusion methods showed improved accuracy compared to regular radiomics features; and (2) TR via multiple PET/CT fusions in HNC: flavours were built from different fusions using methods such as Laplacian pyramids and wavelet transforms. We also tried TR based on segmentation perturbation flavours and preprocessing-filter flavours in CT and MR images. TR showed improved reproducibility as well as overall survival prediction compared to single-flavoured radiomics. Our results suggest that the proposed TR paradigm has the potential to improve performance capabilities in different medical imaging tasks.

I. INTRODUCTION

Radiomics features capture information about tissues and lesions. Radiomics features can be assessed individually or combined to create radiomics signatures. A radiomics feature can be generated using different parameters (e.g. discretized bin number or size; different fusion techniques). There have been significant efforts to establish the best values of parameters suitable for different tasks.

Finding appropriate predictive variables, however, might be challenging. Depending on the scanner used to acquire the imaging data, specific features or versions of features (with particular parameters) might not be noise-resistant or might change. This could lead to a lack of reproducibility between different scanners and institutions. Features of radiomics [1]–[5] and may not be resistant to noise, inter-center protocol, and scanner variability [6], [7]. The wide range of feature-selec tion techniques employed in radiomics investigations [8][9], [10] confirms to the necessity of careful feature selection beforehand and the challenge of this task.

Here we aimed to tackle radiomics analyses using a paradigm that we call ‘tensor radiomics’ (TR) in which multiple versions (or flavours) of the same radiomics features are generated by varying parameters; e.g. different bin sizes, convolutional filters, segmentation perturbations, and fusion techniques. We hypothesize that TR has the potential to overcome some shortcomings of current radiomics models, and to enable improved clinical task performances.

II. MATERIAL AND METHODS

In TR, we build feature tensors using many flavours (Fig. 1) as a method towards the optimized construction of radiomics signatures. We have applied this to different PET/CT, MRI, and CT imaging tasks, but due to space limitations, here we only elaborate two TR techniques in PET/CT imaging.

This work was in part supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant RGPIN-2019-06467, the Canadian Institutes of Health Research (CIHR) Project Grants PJT-162216 and PJT-173231, the Swiss National Science Foundation Grant SNRF 320030_176052, and the BC Cancer Foundation.

Arman Rahmim is with the Department of Radiology, and Department of Physics & Astronomy, University of British Columbia, Vancouver, Canada. He is also with the Department of Integrative Oncology, BC Cancer Research Institute, Vancouver, Canada (e-mail: arman.rahmim@ubc.ca).

Amirhossein Toosi is with the Department of Integrative Oncology, BC Cancer Research Institute, Vancouver, Canada (e-mail: gtoosi@bccrc.ca).

Mohammad R. Salmanpour is with the Department of Integrative Oncology, BC Cancer Research Institute and Technological Virtual Collaboration (TECVICO Corp), Vancouver, Canada (m.salmanpoo6@gmail.com).

Natalia Dubljevic is with Department of Physics & Astronomy, University of British Columbia, Vancouver, Canada (natalia.dubljevic@gmail.com).

Ian Janzen is with BC Cancer Research Institute, Vancouver, Canada (ijanzen@bccrc.ca).

Isaac Shiri is with the Division of Nuclear Medicine and Molecular Imaging, Geneva University Hospital, Geneva, Switzerland (issa.sh92@gmail.com).

Mohamad A. Ramezani is with the Technological Virtual Collaboration (TECVICO Corp), Vancouver, Canada (amirramezanil376@gmail.com).

Ren Yuan and Cheryl Ho are with BC Cancer, Vancouver, Canada (emails: ren.yuan@bccancer.bc.ca and CHo@bccancer.bc.ca).

Habib Zaidi is with the Division of Nuclear Medicine and Molecular Imaging, Geneva University Hospital, Geneva, Switzerland (email: Habib.Zaidi@hcuge.ch).

Calum MacAulay is with BC Cancer Research Institute, Vancouver, Canada (email: cmacaulay@bccrc.ca).

Carlos Uribe is with BC Cancer, Vancouver, Canada (e-mail: curibe@bccrc.ca).

Fereshteh Yousefirizi is with the Department of Integrative Oncology, BC Cancer Research Institute, Vancouver, Canada (e-mail: firez@bccrc.ca).

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
We considered whether TR may enhance HNC outcome prediction. 224 baseline HNC PET/CT scans were obtained from The Cancer Imaging Archive (TCIA), accompanied with tumor expert-manual segmentations. We formulated outcome prediction as a binary classification problem (2-year progression).

### Discretization flavours

TR features were extracted from PET images using 10 different bin sizes (0.1-1.0 SUV) using the PyRadiomics package [1] (106 features used).

1) **ML-based flavour selection framework**: We trained two ML models, SVM with RBF kernel and logistic regression, on concatenated features for all possible combinations of 10 flavours (1013 different combinations with 2 or more flavours). The goal of this comparison is to determine whether combining "all" features from different flavors improves classification performance without performing feature selection. It also seeks to determine whether combinations that enhance classification performance are meaningful; for instance, determining whether combining the smallest and largest bin sizes can improve performance. In order to achieve this, we looked at every possible combination and listed some of the ones that performed best in terms of single bin size features.

2) **DL-based flavour fusion framework**: We applied TR-Net to perform end-to-end flavour fusion on all features extracted using all flavours. TR-Net consists of arms and body: each arm is a stack of multiple dense layers separately fed with the features of each flavour in its input layer. All arms are then concatenated and connected to a few more dense layers to complete the network architecture. A sigmoid function in the last layer performs the binary outcome classification (Fig. 2). Random selection search using a nested k-fold cross-validation was utilized for hyperparameter tuning. To tackle class imbalance during training the network, SMOTE was applied to the training set only, to up-sample the minority (positive) class. Average balanced accuracy and F1 score metrics were computed using stratified 5-fold cross-validation (CV) for both ML & DL methods. Scaled Exponential Linear Unit (SELU) activation function is used in all the dense layers of the arms and the body along with dropout.

3) **ML-based feature selection framework**: We studied the effect of applying feature selection (for any given flavour across different features) prior to performing flavour fusion, on performance of the classification task at hand. A range of 5 to 25 features were selected from each of the 10 flavours using sequential forward feature selection method [2] based on the mean f1 score over a nested 5-fold cross validation setup. These selected features were then combined for all possible combinations of 10 flavours to form our TR features, and then ranked based on mean balanced accuracy and f1 score of nested 5 fold cross validation using an RBF kernel SVM and a Logistic Regression model (Fig. 3).

### Fusion flavours

Fusion radiomics [11], a rapidly growing field, has so far involved investigating many fusions and selecting the most suitable one. In TR, we modify the paradigm and incorporate different fusions into the radiomics tensor before building the model.

First, we separately entered radiomics features extracted (by SERA radiomics package [4]) from each PET-only, CT-only, and 15 fused PET/CT images to three classifiers, including Logistic Regression (LR), Random Forest (RF), and multilayer perceptron (MLP) classifiers. We used 9 same classifiers with different optimized parameters in each fold to vote. For a TR approach, in the proposed TR paradigm, we first incorporated 211 radiomics features with 17 flavours, removing highly correlated features, use Polynomial Feature Transforms to fuse the selected relevant flavours, employ Analysis of variance (ANOVA) to select most relevant fused tensor radiomics (FTR) features. Finally, we applied these selected FTR features to the 3 above mentioned classifiers to predict survival outcome. In this work, we performed 5-fold testing validation (4-fold training and 1-fold testing). In the training process, we further divided training dataset into 2 sections with 80% of datapoints for training the model and remaining 20% for model selection.

### Segmentation and pre-processing filter flavours

We developed TR from different segmentation perturbations and different bin sizes for classification of late-stage lung cancer response to first-line immunotherapy using CT images. In MR images, we developed filter-flavour TR built from different hyper-parameters of pre-processing filters, such as Laplacian pyramids and wavelet transforms.

### III. RESULTS

#### A. Results of using discretization flavours

We first evaluated the effectiveness of combining features from various combinations of the flavours (bin size) using two classification score metrics: balanced accuracy and f1 score (we also studied area under ROC curve, with similar trends; not shown). Five different flavor combinations, chosen from the top 20 possible combinations, are used in Fig. 4 to show the enhanced metric values.
Results show the improved performances from (red) conventional radiomics to (blue) TR in ML pipelines. Further improvements were obtained when utilizing (yellow) our DL TR-Net pipeline.

We also studied the effect of feature selection prior to ML methods applied to conventional vs. TR models, with similar trends (see Fig. 5). Our findings showed that TR features perform better in ML pipelines than ordinary radiomics features when switching from (red) conventional radiomics to (blue) TR. Utilizing (yellow) our end-to-end R-Net pipeline resulted in further gains. Similar findings were observed in our study of the impact of feature selection before ML techniques were applied to conventional vs. TR models (see Fig. 5). A corrected t-test was performed on the results, but no significant improvement over the baseline was found.

**B. Results of using fusion flavours**

For our fusion-based TR framework, we included 211 TR features. After feature fusion, we selected the relevant FTRs and applied those to the mentioned ensemble classifiers. As shown in Fig. 6, the highest nested testing performance of 71.8 ± 4.8% achieved from FTR followed ensemble MLP.

**C. Results on other flavours**

Our results (not shown here due to space limitations) on segmentation-perturbation flavours in CT images showed that TR improved predicted patient responses. Furthermore, results on filter-flavour TR in MR imaging showed improved reproducibility when compared to single-flavour features.
The areas that we can consider improving TR include: determining how well TR methods perform on other, larger datasets, developing new methods (such as using techniques for longitudinal data analysis), and discovering new flavors (e.g. deep features). TR paradigm makes it possible to look back on conventional radiomics approaches and reevaluate them. For instance the sub-regional intra-tumor radiomics that are constructed based on individual- and population-level clustering [12]; by combining feature flavours extracted from various tumor partitions, a new type of TR can be created. Additionally, from various sizes of peritumoral regions, shell features can be retrieved that reflect the tumor microenvironment [13]; their combinations can then be investigated in the suggested framework of TR.

ACKNOWLEDGMENT

This work was in part supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant RGPIN-2019-06467, the Canadian Institutes of Health Research (CHIR) Project Grants PJT-162216 and PJT-173231, the Swiss National Science Foundation Grant SNRF 320030_176052, and the BC Cancer Foundation.

REFERENCES

[1] R. Da-Ano, D. Visvikis, and M. Hatt, “Harmonization strategies for multicenter radiomics investigations,” Phys. Med. Biol., vol. 65, no. 24, p. 24TR02, Dec. 2020.

[2] A. Zwanenburg, “Radiomics in nuclear medicine: robustness, reproducibility, standardization, and how to avoid data analysis traps and replication crisis,” Eur. J. Nucl. Med. Mol. Imaging, vol. 46, no. 13, pp. 2638–2655, Dec. 2019.

[3] S. Ha, H. Choi, J. C. Paeng, and G. J. Cheon, “Radiomics in Oncological PET/CT: a Methodological Overview,” Nucl. Med. Mol. Imaging, vol. 53, no. 1, pp. 14–29, Feb. 2019.

[4] R. Li, L. Xing, S. Napel, and D. L. Rubin, Radionomics and Radiogenomics: Technical Basis and Clinical Applications. CRC Press, 2019.

[5] G. F. R. Cook, G. Azad, K. Osyczcarzyk, M. Siddique, and V. Goh, “Challenges and Promises of PET Radiomics,” Int. J. Radiat. Oncol. Biol. Phys., vol. 102, no. 4, pp. 1083–1089, Nov. 2018.

[6] P. E. Galavis, C. Hollensen, N. Jallow, B. Paliwal, and R. Jeraj, “Variability of textural features in FDG PET images due to different acquisition modes and reconstruction parameters,” Acta Oncol., vol. 49, no. 7, pp. 1012–1016, Oct. 2010.

[7] I. Shiri, A. Rahimim, P. Ghaffarian, P. Geramifir, H. Abdollahi, and A. Bitarafan-Rajabi, “The impact of image reconstruction settings on 18F-FDG PET radiomic features: multi-scanner phantom and patient studies,” Eur. Radiol., vol. 27, no. 11, pp. 4498–4509, Nov. 2017.

[8] M. E. Mayerhoefer et al., “Introduction to Radiomics,” J. Nucl. Med., vol. 61, no. 9, pp. 488–495, Apr. 2020.

[9] I. S. Klyuzhin, Y. Xu, A. Ortiz, J. M. L. Ferres, G. Hamarneh, and A. Rahimim, “Testing the ability of convolutional neural networks to learn radiomic features,” bioRxiv, medRxiv, 23-Sep-2020.

[10] M. Hatt, M. Vallieres, D. Visvikis, and A. Zwanenburg, “IBSI: an international community radiomics standardization initiative,” J. Nucl. Med., vol. 59, no. supplement 1, pp. 287–287, May 2018.

[11] W. Lv, S. Ashrafinia, J. Ma, L. Lu, and A. Rahimim, “Multi-Level Multi-Modality Fusion Radiomics: Application to PET and CT Imaging for Prognostication of Head and Neck Cancer,” IEEE J Biomed Health Inform, vol. 24, no. 8, pp. 2268–2277, Aug. 2020.

[12] J. J. Lewis, R. J. O’Callaghan, S. G. Nikolov, D. R. Bull, and N. Canagarajah, “Pixel-level and region-based image fusion with complex wavelets,” Inf. Fusion, vol. 8, no. 2, pp. 119–130, Apr. 2007.

[13] F. Nencini, A. Gazzelli, S. Baronti, and L. Alparone, “Remote sensing image fusion using the curvelet transform,” Inf. Fusion, vol. 8, no. 2, pp. 143–156, Apr. 2007.