Encoding and Verifying Guidelines for Automated Contouring in Radiotherapy using `VoxLogicA`

Gina Belmonte[*], Giovanna Broccia[**], Laura Bussi[**,***],
Vincenzo Ciancia[**], Diego Latella[**], and Mieke Massink[**]

[*]S.C. Fisica Sanitaria, Azienda USL Toscana Nord Ovest, Lucca.
[**]Istituto di Scienza e Tecnologie dell'Informazione, Consiglio Nazionale delle Ricerche, Pisa.
[***]Dipartimento di Informatica, Università di Pisa.

Abstract

Contouring of lesions and organs at risk is a crucial step in the Radiotherapy workflow. Protocols and guidelines for such workflows are published, and regularly updated. Several promising research results cater for fully-automated contouring, often based on machine-learning approaches. How to check whether such approaches obey to protocols and guidelines is an open research question. The tool `VoxLogicA` can be used to encode protocols and automatically identify the cases in which auto-contouring significantly diverts from the relevant guidelines. Future work will be directed to this application of `VoxLogicA`, which is of uttermost importance for the widespread adoption of automated contouring in Radiotherapy. In this paper, we summarise the current case study devoted to brain tumour segmentation, and the challenges and benefits in following this approach.

1 Introduction

Contouring of lesions and organs at risk is a crucial step in Radiotherapy. The current best practices encompass manual segmentation by domain experts, which is quite time consuming and may create queues, and unwanted delays in treatment. For this reason, fully automated contouring ("auto-contouring") procedures are a hot topic for research.

This paper takes into account concerns related to *quality-assurance* of fully-automated contouring procedures in the radiotherapy workflow. The research question we tackle is that of monitoring an autocontouring procedure (often based on so-called "black-box" machine-learning) to assess and report whether it respects recommendations, protocols and guidelines that are regularly published and updated.

As a running example, we consider automated segmentation of brain tumours (see [Baid *et al.*, 2021]). In the area, guidelines are described in a quite formal way (see [Niyazi *et al.*, 2016]) as a checklist for the Radiologist. Such a checklist may include anatomical constraints such as the organs that should *not* be included in the contour (e.g. the cerebellum), topological considerations (e.g. an enlargement radius that should be adopted), indications on preferred imaging modalities for the identification of specific features, and so on.

The results we present have their origin in the EU FP7 QUANTICOL project[1]. Currently, the project is partially funded by the Italian PRIN IT-MATTERS[2].

2 Declarative Imaging via Spatial Model Checking

The Free and Open Source spatial model checker `VoxLogicA`[3] (see [Ciancia *et al.*, 2021] for an introduction) proved to be suitable to express domain knowledge in medical imaging in a concise, executable form, and delivers top-notch performance due to the usage of state-of-the-art imaging libraries, automatic parallel scheduling, *memoization* and more recently, a GPU-based engine [Bussi *et al.*, 2021]. In [Belmonte *et al.*, 2019; Banci Buonamici *et al.*, 2020] a declarative specification of about 10 lines of text has been used for fully automated segmentation of brain tumours against the BraTS dataset [Baid *et al.*, 2021] with accuracy in par with the state of the art, and very fast run-time (about 5 seconds per case, which may be reduced by one-two orders of magnitude using the GPU variant). In [Belmonte *et al.*, 2021], the tool has been used, with accurate results, for segmentation of *nevi* against a well-known dataset.

Our proposed approach embraces a classic tenet of AI: that of *spatial logics*. More precisely, our work stems from the so-called *topological spatial logics* [Aiello *et al.*, 2007], where the object of reasoning are points, not regions[4]. However, differently from more traditional applications in *Symbolic Artificial Intelligence* we do not take *deduction* into account, but rather propose *model checking*. Model checking is a technique traditionally used to statically verify system behaviour against functional requirements, formalised using modal logics. We adapted and augmented such paradigm, obtaining a computational imaging application that can *automatically* execute logical specifications that describe

[1]EU FP7 Grant 600708, see https://cordis.europa.eu/project/id/600708/it.

[2]MIUR PRIN 2017FTXR7S, see http://itmatters.imtlucca.it/.

[3]See https://github.com/vincenzoml/VoxLogicA

[4]See also [Ciancia *et al.*, 2019] encoding of the *region calculus* of [Randell *et al.*, 1992] into an extended topological spatial logic.

sets of points (therefore, ultimately regions) satisfying user-specified properties. We defined the *Spatial Logic of Closure Spaces* (SLCS) [Ciancia *et al.*, 2016], first used in applications related to smart cities (see e.g. [Ciancia *et al.*, 2018]). It is also worth noting that the logic SLCS has been adopted, and extended, also in other contexts and by other groups of authors; for instance, for cyber-physical systems ([Tsigkanos *et al.*, 2017]), or for run-time monitoring [Nenzi *et al.*, 2018; Bartocci *et al.*, 2020].

3 Encoding Radiotherapy Guidelines: Challenges and Benefits

The accuracy of the procedure first described in [Belmonte *et al.*, 2019][5] is mostly due to the encoding of set constraint which is not a traditional kind of imaging feature (e.g. texture, statistical features) but rather described as a fact coming from domain knowledge: "the tumour mostly consists of the *very intense* areas of an MRI-FLAIR image of the brain that have an *hyper-intense* core, whereas those very intense areas that do not have an hyper-intense core are usually not part of the lesion". Such constraint is not only based on thresholds, but additionally on a global connectivity property, that can be easily expressed in SLCS, as the language includes reachability predicates.

Challenges Radiotherapy guidelines such as [Niyazi *et al.*, 2016] contain several indications similar to the one above. Spatial logics is, generally speaking, a good instrument to formalise such requirements. However, for some constraints (e.g. those requiring the identification of specific organs such as the cerebellum) we envisage the combination of declarative methods (mostly seen as *coordination languages*) with black-box methods (performing simple, specific tasks), resulting in a hybrid-artificial-intelligence approach that is explainable at higher levels of abstraction, and becomes black-box in the handling of lower-level details. Similarly, hybrid artificial intelligence can be used to calibrate the numeric parameters of declarative procedures on each image, obtaining more precise fully automated methods that are still deterministic, unambiguous, and completely explainable. Designing and validating such hybrid methods is part of our planned future work and constitutes the main challenge in attaining our stated goal of encoding contouring guidelines.

Benefits Once radiotherapy guidelines can be satisfactorily encoded as mostly-declarative (possibly hybrid) procedures, which are explainable by construction, such *executable specifications* can be used in pair with black-box methods that software or hardware vendors will provide in the near future, to monitor their behaviour and signal divergence from guidelines, either based upon agreement using numerical scores (e.g. the Dice/Sorensen similarity index) or on topological constraints (e.g. containment). One of the advantages of this approach, besides explainability, is that short, unambiguous, fully-automated declarative specifications can be easily distributed, discussed/published and collectively improved by the community, therefore empowering healthcare providers and researchers with their own instruments to evaluate and assess dependability of vendor-specific autocontouring procedures.

[5]Average Dice/Sorensen score on the High-Grade glioblastomas from the dataset of [Baid *et al.*, 2021] (259 cases): 0.83. Best possible accuracy with per-image calibration of just one numerical parameter: 0.87.

4 Conclusions

In our work on using spatial logics and model checking to identify features of images, we learned that declarative specifications can be extremely appealing for their simplicity, for their fully formalised execution semantics, and for their accuracy. Using model checking, execution speed is generally speaking much faster than other forms of analysis and the workflow is quite simple, as the interpreter we designed is portable and easy to set up, especially in comparison to the state-of-the-art machine learning methods.

Given the widespread concerns about quality assurance of machine learning in Radiotherapy, and the availability of semi-formal protocols and guidelines, it seems natural at this point to direct our attention to such problems. However, some challenges, such as the combination of black-box and declarative methods, arise. In the near future we plan to tackle such challenges and develop a hybrid-artificial-intelligence approach for monitoring and quality assurance of black-box autocontouring procedures in Radiotherapy.

5 Acknowledgments & Contacts

Research partially supported by the MIUR Project PRIN 2017FTXR7S "IT- MaTTerS".

The authors are listed in alphabetical order.

E-mail addresses: Gina Belmonte: gina.belmonte@uslnordovest.toscana.it – Giovanna Broccia, Vincenzo Ciancia, Diego Latella, Mieke Massink, Laura Bussi: name.surname@isti.cnr.it.

References

[Aiello *et al.*, 2007] M. Aiello, I. Pratt-Hartmann, e J. Benthem, van. *Handbook of Spatial Logics*. Springer, 2007.

[Baid *et al.*, 2021] Ujjwal Baid *et al.* The RSNA-ASNR-MICCAI brats 2021 benchmark on brain tumor segmentation and radiogenomic classification. *CoRR*, abs/2107.02314, 2021.

[Banci Buonamici *et al.*, 2020] Fabrizio Banci Buonamici, Gina Belmonte, Vincenzo Ciancia, Diego Latella, e Mieke Massink. Spatial logics and model checking for medical imaging. *Int. J. Softw. Tools Technol. Transf.*, 22(2):195–217, 2020.

[Bartocci *et al.*, 2020] Ezio Bartocci, Luca Bortolussi, Michele Loreti, Laura Nenzi, e Simone Silvetti. Moonlight: A lightweight tool for monitoring spatio-temporal properties. In *Runtime Verification - 20th International Conference*, volume 12399 of *Lecture Notes in Computer Science*, pages 417–428. Springer, 2020.

[Belmonte *et al.*, 2019] Gina Belmonte, Vincenzo Ciancia, Diego Latella, e Mieke Massink. Voxlogica: A spatial model checker for declarative image analysis. In *Tools and Algorithms for the Construction and Analysis of Systems - 25th International Conference, TACAS 2019*, volume 11427 of *Lecture Notes in Computer Science*, pages 281–298. Springer, 2019.

[Belmonte *et al.*, 2021] Gina Belmonte, Giovanna Broccia, Ciancia Vincenzo, Diego Latella, e Mieke Massink. Feasibility of spatial model checking for nevus segmentation. In *Proceedings of the 9th International Conference on Formal Methods in Software Engineering (FormalieSE'21)*, pages 1–12. IEEE, 2021.

[Bussi *et al.*, 2021] Laura Bussi, Vincenzo Ciancia, e Fabio Gadducci. Towards a spatial model checker on GPU. In *FORTE 2021 - 41st International Conference on Formal Techniques for Distributed Objects, Components, and Systems*, volume to appear, 2021.

[Ciancia *et al.*, 2016] V. Ciancia, D. Latella, M. Loreti, e M. Massink. Model Checking Spatial Logics for Closure Spaces. *Logical Methods in Computer Science*, Volume 12, Issue 4, October 2016.

[Ciancia *et al.*, 2018] V. Ciancia, S. Gilmore, G. Grilletti, D. Latella, M. Loreti, e M. Massink. Spatio-temporal model checking of vehicular movement in public transport systems. *International Journal on Software Tools for Technology Transfer*, Jan 2018.

[Ciancia *et al.*, 2019] Vincenzo Ciancia, Diego Latella, e Mieke Massink. Embedding RCC8D in the collective spatial logic CSLCS. In *Models, Languages, and Tools for Concurrent and Distributed Programming - Essays Dedicated to Rocco De Nicola on the Occasion of His 65th Birthday*, volume 11665 of *Lecture Notes in Computer Science*, pages 260–277. Springer, 2019.

[Ciancia *et al.*, 2021] Vincenzo Ciancia, Gina Belmonte, Diego Latella, e Mieke Massink. A hands-on introduction to spatial model checking using voxlogica. In Alfons Laarman e Ana Sokolova, editors, *Model Checking Software*, pages 22–41, Cham, 2021. Springer International Publishing.

[Nenzi *et al.*, 2018] L. Nenzi, L. Bortolussi, V. Ciancia, M. Loreti, e M. Massink. Qualitative and Quantitative Monitoring of Spatio-Temporal Properties with SSTL. *Logical Methods in Computer Science*, 14(4):1–38, 2018.

[Niyazi *et al.*, 2016] Maximilian Niyazi, Michael Brada, Anthony J Chalmers, Stephanie E Combs, Sara C Erridge, Alba Fiorentino, Anca L Grosu, Frank J Lagerwaard, Giuseppe Minniti, René-Olivier Mirimanoff, Umberto Ricardi, Susan C Short, Damien C Weber, e Claus Belka. ESTRO-ACROP guideline "target delineation of glioblastomas". *Radiother. Oncol.*, 118(1):35–42, January 2016.

[Randell *et al.*, 1992] David A. Randell, Zhan Cui, e Anthony G. Cohn. A spatial logic based on regions and connection. In *Proceedings of the 3rd International Conference on Principles of Knowledge Representation and Reasoning (KR'92).*, pages 165–176. Morgan Kaufmann, 1992.

[Tsigkanos *et al.*, 2017] Christos Tsigkanos, Timo Kehrer, e Carlo Ghezzi. Modeling and verification of evolving cyber-physical spaces. In *Proceedings of the 2017 11th Joint Meeting on Foundations of Software Engineering*, ESEC/FSE 2017, pages 38–48. ACM, 2017.