



PHYSICS

Editors' Pick

Incorporating human and learned domain knowledge into training deep neural networks: A differentiable dose-volume histogram and adversarial inspired framework for generating Pareto optimal dose distributions in radiation therapy

Dan Nguyen, Rafe McBeth, Azar Sadeghnejad Barkousaraie, Gyanendra Bohara, Chenyang Shen, Xun Jia, Steve Jiang

Med Phys 2019 Dec 10. doi: 10.1002/mp.13955. [Epub ahead of print]

<https://aapm.onlinelibrary.wiley.com/doi/full/10.1002/mp.13955>

What was your motivation for initiating this study?

When we first began our venture into deep learning for radiotherapy applications back in July 2017, we founded the Medical Artificial Intelligence and Automation (MAIA) Laboratory, which is a multiple investigator lab designed to facilitate collaboration and ideas-sharing among lab members. A number of us began to investigate treatment planning, and how artificial intelligence (AI) and automation technologies could be applied to improve the quality and time of treatment planning, and treatment deliverability. We found that deep learning was extremely suitable to predict the clinically delivered dose distribution and we published a paper that showed its feasibility (arXiv preprint) (publication). Soon we investigated its capability in dose prediction for several sites, such as lung, head and neck, etc.

As we progressed our knowledge of deep learning, we ran into an interesting conundrum: was it better to solve an approximate problem exactly, or to solve an exact problem approximately? In the current era of data-driven deep learning, the answer seems to lean towards the latter. Deep-learning models are capable of modelling data in extreme detail and accuracy. So, if we're already leaning towards trying to solve exact problems approximately, and we are already throwing out global optimum guarantees by having a highly non-linear AI model, why don't we take it a step further and redesign loss functions that directly reflect what we care about as well? In other words, we can create use and domain specific losses that we avoided in the past due to their non-convexities and the ill-behaved nature of their gradients. We hypothesised that deep learning would still be able to utilise such losses, as long as they showed defined gradients. Used this way, the model would be focused on minimising the errors specific to our domain, rather than just the overall error that a domain agnostic loss, such as mean squared error (MSE), would provide.

Thus we dived straight into a study that focused on investigating the effects of domain-specific losses on the AI's performance. We wanted to see the effects on the model of adding one of the most commonly used forms of evaluation data in radiation oncology, the dose-volume histogram (DVH), as a loss function. In addition we wanted to see how the performance would be affected by the addition of an adversarial loss, which would effectively allow for the AI framework to learn its own domain specific features to minimise.

What were the main challenges during the work?

There were three main challenges during the work:

1. Loss-function modelling – inherently a histogram is not differentiable. We focused efforts on redesigning a smooth differentiable approximation of the DVH. A key emphasis in the design of this function was that a loss function that used this approximate DVH would have exactly the same minima as the true DVH. By having the same minima, we could be sure that this approximate DVH could still be used to solve the exact problem.

2. Data generation – we tested the effects of the additional domain losses for a model trained to navigate the trade-off space of plans. Because this requires the network to learn how to map many dose distributions from a single anatomy, the neural network must learn to differentiate the potentially small nuances between different doses that may have substantial clinical consequences. This can be challenging for MSE loss alone, since MSE is only focused on minimising the mean error. This particular type of problem would be a good arena for testing our hypothesis.
3. To create data that would teach a model to navigate the trade-off space. We had to simulate many plans for a single patient. In our study, we generated 1200 pseudo-randomised plans, leading to a total of 84000 plans for our 70 patients. The ability to design a quick optimisation algorithm was the key to the creation of these plans in a reasonable time. We eventually turned to a proximal algorithm that came out in 2010, referred to as Chambolle-Pock or the Primal Dual Hybrid Gradient (PDHG) method.

What is the most important finding of your study?

We showed that the addition of the DVH metric (human domain knowledge) and adversarial loss (learned domain knowledge) could greatly improve the performance of the model in the areas relevant to our domain.

What are the implications of this research?

This study paved the way for us to rethink how we design and train AI frameworks. In particular, instead of using generic losses, and hoping for accurate AI models, we should focus on rewriting the loss function from the ground up, concentrating on the exact problem we wish to solve. Expert human-domain-specific knowledge can be the largest driver in the performance improvement, and adversarial learning can be used to capture further these particular subtle attributes as part of the loss. We also now have a trained AI model that can enable a physician quickly to navigate the trade-off space for a patient in real time, and produce a dose distribution as a tangible endpoint for the dosimetrist to use for planning. This is expected to reduce considerably the treatment-planning time, while improving the treatment-planning quality, enabling clinicians to focus their efforts on the difficult and demanding cases. In addition, the resulting deep-learning model from this study can be incorporated with fast automatic treatment-planning frameworks. This model would efficiently find the appropriate tradeoffs in the dose distribution, and a deliverable plan of it could be generated automatically.

Dan Nguyen & Steve Jiang

Medical Artificial Intelligence &
Automation (MAIA) Laboratory
Department of Radiation Oncology
UT Southwestern Medical Center
Dallas, USA



Dan Nguyen



Steve Jiang

