Deep Learning for Plasma Tomography in Nuclear Fusion

Anonymous Author(s)
Affiliation
Address
email

Abstract

Tomography is arguably one of the most representative examples of an inverse problem, where the shape of an object must be reconstructed from its projections over a limited number of lines of sight. The regularization that must be imposed to solve such an ill-posed problem often results in iterative algorithms that are computationally expensive and do not meet the requirements of real-time applications. Deep learning offers a promising approach to perform such reconstruction with sufficient accuracy, while being several orders of magnitude faster, to the point that it becomes possible to use tomography in real-time. In this paper, we give an example of how real-time tomography based on deep learning is being used to reconstruct the plasma radiation profile inside a nuclear fusion device. The availability of such profile in real-time allows setting up new alarms in the real-time control system, with a view towards anticipating plasma disruptions.

1 Introduction

Tomography is best known in the context of medical applications, where it is usually referred to as *Computed Tomography* [1]. In this context, tomography provides the means to observe the interior of an object by the successive application of X-rays at different angles. While in medical applications the object is traversed by externally produced X-rays, in other applications the object itself emits X-rays and possibly other types of electromagnetic radiation as well. In any case, the principle of tomography is the same: by collecting projections of an object from multiple viewpoints, it should be possible to reconstruct the internal structure of that object.

Figure 1(a) illustrates the simplest possible example of a tomographic reconstruction problem. Let $\mathbf{x}$ be an image of the object, and let $\mathbf{y}$ be its projections along certain lines of sight. The relationship between $\mathbf{x}$ and $\mathbf{y}$ can be represented in matrix form as $\mathbf{y} = \mathbf{A}\mathbf{x}$, where $\mathbf{A}$ is the transformation matrix between an image and its projections along those lines of sight. In practice, we are given the projections $\mathbf{y}$ but not the original image $\mathbf{x}$, so the purpose is to reconstruct $\mathbf{x}$ from $\mathbf{y}$. In the simple case of Figure 1(a), it appears that it is possible to do that via $\mathbf{x} = \mathbf{A}^{-1}\mathbf{y}$.

Figure 1(b) shows a more realistic scenario, where the number of projections is insufficient to have a fully-determined system that can be directly inverted. In this case, there is no information about certain details of the image. However, it is often possible to assume that the object is smooth or has some properties that allow the missing elements to be derived from the ones that can be observed. This eventually results in some form of regularization that provides a unique solution by minimizing some criteria. For example, minimizing the differences between adjacent elements of $\mathbf{x}$ leads to the well-known scheme of *total variation* [2].

Submitted to 34th Conference on Neural Information Processing Systems (NeurIPS 2020). Do not distribute.

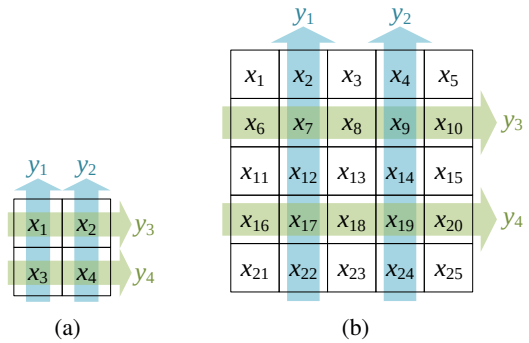

Figure 1: (a) Fully-determined vs. (b) under-determined system.

A common way to solve an under-determined system is by Tikhonov regularization [3], where the solution to $\mathbf{y} = \mathbf{A}\mathbf{x}$ is found by minimizing the functional:

$$\phi(\mathbf{x}) = \|\mathbf{y} - \mathbf{A}\mathbf{x}\|^2 + \lambda \|\mathbf{\Gamma}\mathbf{x}\|^2 \tag{1}$$

In Eq. (1), λ is a scalar regularization parameter (i.e. a regularization weight) and $\mathbf{\Gamma}$ is a regularization matrix applied over the elements of $\mathbf{x}$. By setting $\mathbf{\Gamma}$, it is possible to implement different forms of regularization. For example, $\mathbf{\Gamma} = I$ (identity matrix) leads to the traditional L_2 regularization. By careful choice of $\mathbf{\Gamma}$ it is also possible to implement variational regularization schemes [4], such as those that take into account the differences between neighboring elements of $\mathbf{x}$.

An analytical solution to the minimization of Eq. (1) can be found by equating $\partial\phi/\partial\mathbf{x} = 0$ and solving for $\mathbf{x}$, yielding:

$$\mathbf{x} = \underbrace{(\mathbf{A}^\top\mathbf{A} + \lambda\mathbf{\Gamma}^\top\mathbf{\Gamma})^{-1}\mathbf{A}^\top}_{\mathbf{M}}\,\mathbf{y} \tag{2}$$

If matrix $\mathbf{M}$ were available, then reconstructing $\mathbf{x}$ would be a simple matter of multiplying $\mathbf{M}$ by the projections $\mathbf{y}$. In practice, however, a simple linear model $\mathbf{x} = \mathbf{M}\mathbf{y}$ is unable to capture the non-linearities involved in the measurement process, such as sensor saturation, calibration, malfunctions, etc. A more accurate model is $\mathbf{x} = \mathbf{M}(\mathbf{y})$, where $\mathbf{M}$ is a non-linear function, such as a neural network that is trained to transform the projections $\mathbf{y}$ into a reconstruction $\mathbf{x}$.

In the following sections, we describe how we devised and trained a deep neural network to perform the tomographic reconstruction in a specific scenario, where the original reconstruction process was too time-consuming and computationally intensive to be employed in real-time. By using deep learning, it became possible to produce similar results with sufficient accuracy, while being several orders of magnitude faster. This opens up new possibilities in terms of utilizing those results for the real-time monitoring and control of a large-scale scientific experiment.

2 Plasma tomography at JET

The Joint European Torus (JET), located near Oxford, in the UK, is currently the largest operating tokamak, to be soon superseded by ITER [5], which is being built in the south of France. In essence, a tokamak is a torus-shaped fusion device where a plasma is heated to very high temperatures while being confined by a magnetic field. Despite its relatively simple configuration, the plasma inside a tokamak has a rich, non-linear behavior, and is prone to magnetohydrodynamic instabilities which often lead to plasma disruptions [6].

As in many other tokamaks, the plasma behavior at JET is monitored through a wide range of diagnostic systems. Some of these diagnostics measure the power that the plasma emits by electromagnetic radiation in a broad spectrum that also includes X-rays. One of such diagnostics is the *bolometer system* [7], which comprises a horizontal camera and a vertical camera to measure the radiated power on a poloidal cross-section of the fusion device. Its geometry is illustrated in Figure 2.

Each camera has 24 sensors, and each sensor measures the line-integrated radiation along a particular line of sight. The geometrical arrangement of these lines of sight is such that, for both the horizontal

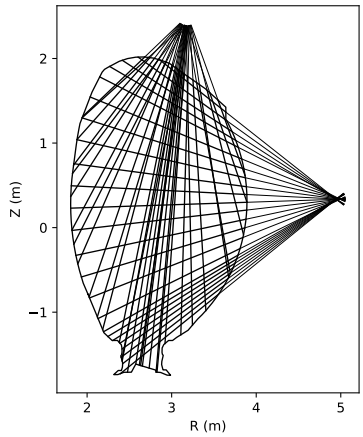

Figure 2: Lines of sight for the bolometer system at JET

and the vertical camera, 16 channels cover the whole plasma region, and 8 channels provide a more fine-grained resolution of the divertor region, at the bottom.

From the sensor measurements, there are several tomography techniques that can be applied to reconstruct the plasma radiation profile [8]. The method that is used at JET uses an iterative constrained optimization algorithm that minimizes the error with respect to the measurements, while requiring the solution to be non-negative [9]. This iterative procedure takes a significant amount of computation time, typically on the order of minutes to produce a single reconstruction.

Despite this run-time performance, thousands of reconstructions have been computed at JET over the years, and these constitute a valuable resource not only for the analysis of past experiments, but also for comparing different approaches and training new models to produce similar reconstructions.

3 Deep learning for plasma tomography

Figure 3 shows the deep neural network that we developed for plasma tomography. It receives the bolometer measurements as input, and produces a plasma radiation profile as output, with the same resolution as the tomographic reconstructions that are routinely produced at JET (196×115 pixels).

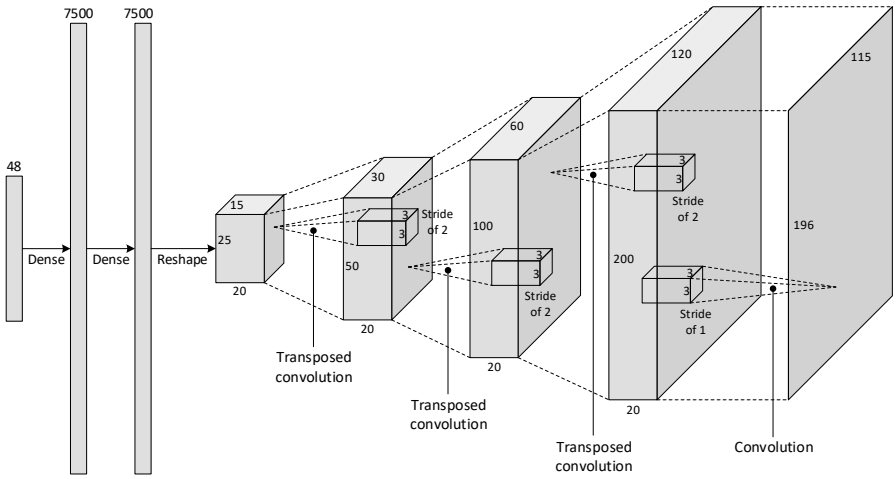

Figure 3: Deconvolutional neural network for plasma tomography

At the beginning of the network, there are two dense layers with 7500 nodes, which are reshaped into a 3D tensor of size $25\times15\times20$. This 3D shape can be interpreted as comprising 20 features maps of size 25×15. By applying a series of transposed convolutions, the feature maps are brought up to a size of 200×120, from which the output image is generated by one last convolution.

The upsampling of feature maps from 25×15 up to 200×120 is achieved by having each transposed convolution operate with a stride of two pixels, except for the very last convolution which uses a stride of one to merge the features maps into a single output image. The output shape is trimmed to 196×115 by a simple array-slicing operation, which is not shown in the figure.

The network was trained on about 10 000 sample reconstructions, where we exercised extra care to select high-quality samples. The dataset was split into 90% for training and 10% for validation. With a proper choice of batch size and learning rate, the network can be trained in about 1 hour on a machine with 8 Nvidia P100 GPUs. It converges to a mean absolute error of about 0.014 MW/m^3 on the validation set (and half of that value on the training set), which is quite low when considering that the dynamic range of the plasma radiation profile is typically of the order of 1 MW/m^3.

4 Plasma monitoring in real-time

Figure 4 shows a sequence of frames illustrating the reconstructions produced by the model. In this example, it is possible to observe the development of a radiation blob at the outboard edge, followed by the development of a radiation blob at the plasma core, which eventually leads to a disruption. The two radiation blobs seem to be related by impurity transport from edge to core, since at around $t=52.0$ s the outboard radiation decreases, followed by an increase in core radiation.

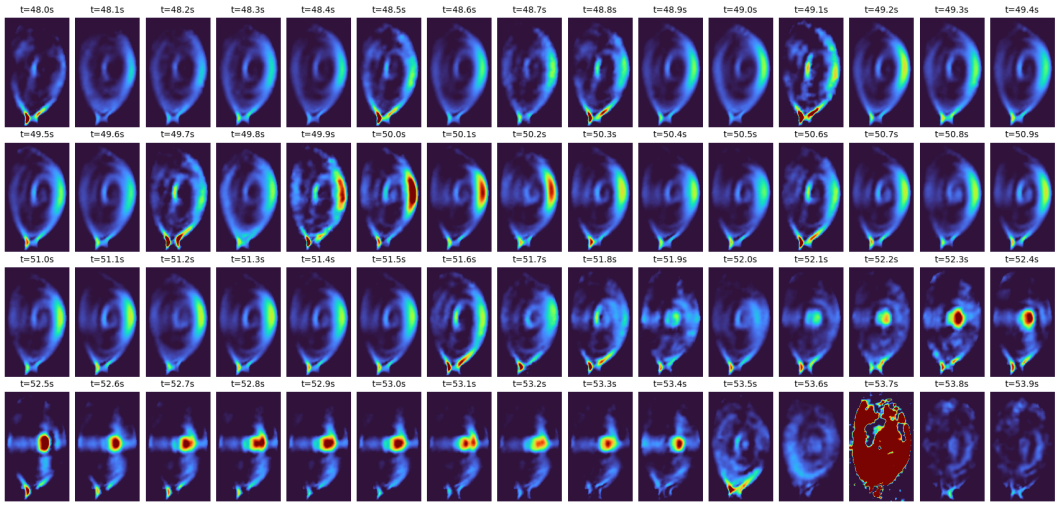

Figure 4: Plasma profiles for pulse 92213 from $t=48.0$ s to $t=53.9$ s with a time step of 0.1s

Some of the reconstructions in Figure 4 may appear to be problematic, such as the profile that is obtained at the disruption time ($t=53.7$ s) or the profiles that are obtained when there are divertor events (e.g. $t=49.1$ s). These phenomena are probably rare or even non-existent in the training dataset, so it is likely that the neural network finds them particularly difficult to reconstruct. In any case, this does not prevent using the model in real-time, since the purpose is to monitor the radiated power in certain regions, rather than knowing the exact values at precise locations.

The main advantage of using plasma tomography in real-time is that it allows monitoring the radiated power at the outboard edge, at the plasma core, and at the divertor region. In particular, it turns out that a significant fraction of disruptions at JET are due to high core radiation [10], so one of the main features to be monitored is the radiated power in the core region. For this purpose, we created an alarm in the real-time control system at JET, such that the alarm is triggered when the amount of core radiation exceeds a threshold of 3 MW. Our analysis of past experiments suggests that such alarm is able to anticipate about 50% of disruptions with only a 5% false alarm rate.

Acknowledgements

This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014-2018 and 2019-2020 under grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

References

[1] T. M. Buzug. *Computed Tomography: From Photon Statistics to Modern Cone-Beam CT*. Springer, 2008.

[2] D. Strong and T. Chan. Edge-preserving and scale-dependent properties of total variation regularization. *Inverse Problems*, 19(6):S165–S187, November 2003.

[3] G. H. Golub, P. C. Hansen, and D. P. O'Leary. Tikhonov regularization and total least squares. *SIAM Journal on Matrix Analysis and Applications*, 21(1):185–194, 1999.

[4] C. R. Vogel. *Computational Methods for Inverse Problems*. Frontiers in Applied Mathematics. SIAM, 2002.

[5] M. Claessens. *ITER: The Giant Fusion Reactor*. Springer, 2020.

[6] H. Zohm. *Magnetohydrodynamic Stability of Tokamaks*. Wiley-VCH, 2015.

[7] A. Huber, K. McCormick, P. Andrew, P. Beaumont, S. Dalley, J. Fink, J.C. Fuchs, K. Fullard, W. Fundamenski, L.C. Ingesson, F. Mast, S. Jachmich, G.F. Matthews, Ph. Mertens, V. Philipps, R.A. Pitts, S. Sanders, and W. Zeidner. Upgraded bolometer system on JET for improved radiation measurements. *Fusion Engineering and Design*, 82(5):1327–1334, 2007.

[8] J. Mlynar, T. Craciunescu, D. R. Ferreira, P. Carvalho, O. Ficker, O. Grover, M. Imrisek, and J. Svoboda. Current research into applications of tomography for fusion diagnostics. *Journal of Fusion Energy*, 38(3):458–466, 2019.

[9] L. C. Ingesson, B. Alper, H. Chen, A. W. Edwards, G. C. Fehmers, J. C. Fuchs, R. Giannella, R. D. Gill, L. Lauro-Taroni, and M. Romanelli. Soft X ray tomography during ELMs and impurity injection in JET. *Nuclear Fusion*, 38(11):1675–1694, 1998.

[10] P. C. de Vries, M. Baruzzo, G. M. D. Hogeweij, S. Jachmich, E. Joffrin, P. J. Lomas, G. F. Matthews, A. Murari, I. Nunes, T. Pütterich, C. Reux, and J. Vega. The influence of an ITER-like wall on disruptions at JET. *Physics of Plasmas*, 21(5):056101, 2014.