Artificial Neural Network-Based Tomography Reconstruction of Plasma Radiation Distribution at GOLEM Tokamak

S. Abbasi[1] · J. Mlynar[1] · J. Chlum[1] · O. Ficker[1,2] · V. Svoboda[1] · J. Brotankova[1]

Abstract

The paper presents an artificial neural network-based model for tomography reconstruction of visible plasma radiation distribution at the GOLEM tokamak. The model was trained using a dataset from emissivity phantoms and associated synthetic measurements from a poloidal cross-section of the GOLEM tokamak. The model validation was performed on the prediction of various unseen phantom samples with shapes similar to those in the training dataset. The backfit of line-integrated measurements indicates the considerable potential of the proposed model for reconstructing the position, size, shape and intensity of the radiation function of one cross section. Additionally, the neural network-based model offers a significantly shorter prediction time compared to traditional tomography methods, providing a substantial advantage.

Keywords Artificial neural networks · Tomography reconstruction · Plasma diagnostics

Introduction

The interpretation of radiation as a diagnostic tool to characterize plasma properties is a critical aspect of fusion plasma confinement. In tokamak plasma, the tomographic inversion methods arise in the field of the reconstruction of the local radiation emissivity using the plasma projections measurements [1]. The reconstruction accuracy of plasma parameters mainly contributes to precise monitoring and control, which are essential for achieving an efficient plasma confinement [2]. However, the limited field of view causes the measured data sparse, leading the implementation of regularization method introducing computationally expensive inversion process [3]. On the other hand, the rapid changes occurring during the plasma discharge underline the importance of achieving high temporal resolution [4].

Beyond the inversion result accuracy, the important role of the real-time monitoring [5] has prompted researchers to leverage Machine Learning (ML) techniques [6], such as Artificial Neural Network (ANN)-based models [7]. In tomography, the neural networks are implemented to train a model to reconstruct the value associated with each pixel with a high accuracy, effectively modeling the entire grid pattern [8]. The ability to yield a large number of reconstructions per second at high resolution is a considerable advantage of the trained model, enabling the detection of the plasma profile throughout an entire discharge. This capability provides a promising potential for real-time control and tokamak disruption prediction [9]. Several methods, such as Feedforward Neural Networks (FNNs) with fully-connected layers [10] and deconvolutional neural networks, which are the inverse of Convolutional Neural Networks (CNNs) [11], have demonstrated high accuracy in plasma tomography. Besides the optimal neural network structure and the proper training method, the performance of the model critically depends on the coverage and completeness of the training data [12]. The model can be trained using a real experimental dataset, which captures a variety of the most common conditions [13], or a representative synthetic dataset, which offers advantages in fast data creation and less risk of overfitting.

In previous work [14], the images captured by two visible cameras were used for tomography reconstruction of the radiation function for a single cross-section by implementing the Tomotok package [15]. Then, in order to reach a model with shorter reconstruction time, the tomography

✉ S. Abbasi
    saraabbasi2@gmail.com

[1] Faculty of Nuclear Sciences and Physical Engineering, CTU in Prague, Brehova 7, Prague 11519, Czech Republic

[2] Institute of Plasma Physics of the Czech Academy of Sciences, U Slovanky 2525/1a, Prague 182 00, Czech Republic

reconstruction result of different plasma discharges was used as a dataset to train a ML-based model [16]. However, due to the role of the training dataset in model learning, the inversion errors commonly encountered in traditional tomography method can negatively impact the model's accuracy. In order to eliminate the mentioned inversion error, and achieve relatively higher accuracy, this study attempts to apply a representative synthetic dataset to train an ANN-based tomography model. The synthetic dataset is constructed by samples consisting of emissivity phantoms and associated synthetic measurements corresponding to one poloidal cross-section of the GOLEM tokamak. The model is trained to predict the radiation function corresponding to the images captured by two Photron Mini UX high-speed cameras with crossed fields of view, placed in the same poloidal cross-section at GOLEM [17].

This paper is organized as follows: Section Overview of Neural Network gives a brief overview of artificial neural network. Section ANN-based Tomography at GOLEM Tokamakdescribes an ANN-based tomography model at the system under investigation, the GOLEM Tokamak with installed fast visible cameras, Training Dataset and Training Process. Section Results and Discussion details and discusses the reconstruction result provided by the trained ANN-based model. Finally, section Conclusion presents a summary of the conclusions.

Artificial Neuron and Neural Network

ANN offers robust methods for solving problems by extracting and interpreting the patterns within a dataset. They simulate the neural structure of human visual processing by means of high-speed processing artificial neurons that produce a series of real-valued activations to learn solutions to a given problem [18]. Figure 1 (left side) illustrates an artificial neuron structure, comprising input data $(x_1, x_2, ..., x_j)$ with corresponding weights $(w_{i1}, w_{i2}, ..., w_{ij})$, the activation function (f), and the resulting output (a_i). The activation function computes the output of the neuron using the formula $a_i = f(\sum_j w_{ij} x_j + b_i)$, where b_i is the neuron bias. Figure 1 (right side) visualizes an ANN composed of interconnected layers of neurons, consisting of an input layer, two hidden layers, and an output layer, highlighting the distinct layers and the flow of information between neurons. In this architecture, the output of each neuron in one layer represents the input for neurons in the subsequent layer, enabling the propagation of information through the network.

Neural networks learn by iteratively adjusting their parameters based on the error in predictions of the input data using a method called backpropagation. In this method, during the training process, the errors between the predicted and actual outputs are calculated and propagated backward through the network to update the weights, improving the network's accuracy. A cost function is employed to measure the disparity between the actual output and the calculated output.

ANN-based Tomography at GOLEM Tokamak

GOLEM Tokamak with Installed Fast Visible Cameras

The GOLEM tokamak is located at the faculty of Nuclear Physics an Physical Engineering (Czech Technical University in Prague). The diagnostic system to detect the visible plasma radiation consists of two crossed visible color cameras installed on the same poloidal cross-section. On the left side of Fig. 2, the schematic of one GOLEM's circular cross-section illustrates the Line of Sight (LoS) layout of the Radial (R) and Vertical (V) cameras, represented in pink
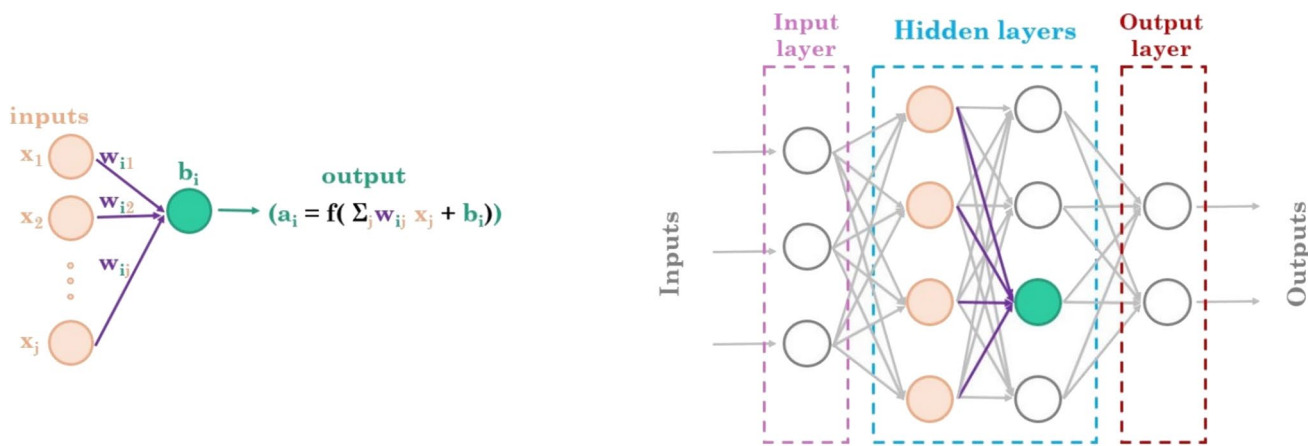

Fig. 1 An artificial neuron structure (left side) and a neural network structure with the different layers (right side)
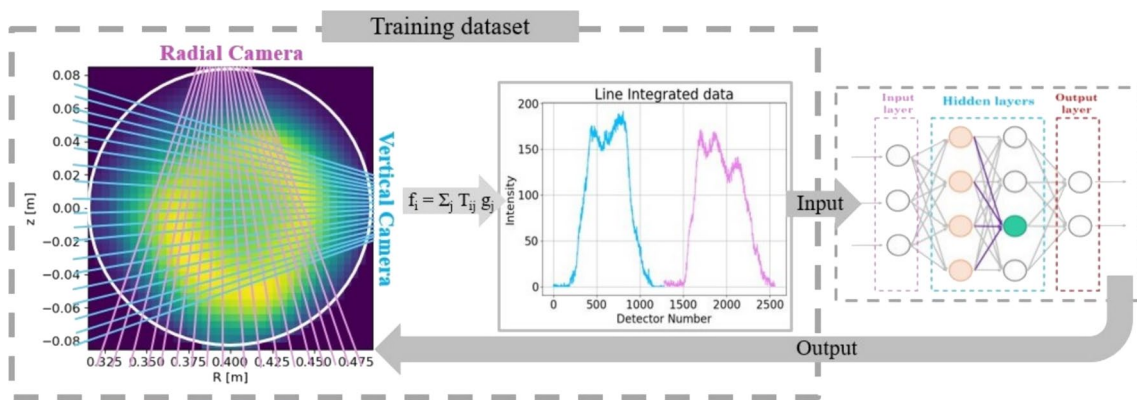
Fig. 2 Schematic diagram of the ANN model performed on the GOLEM tokamak. Left side: Schematic figure of one poloidal cross-section of the GOLEM tokamak and the simulated phantom distribution with the LoS layout of R and V cameras, represented in pink and blue, respectively. Middle side: The line integrated measurements associated to the phantom data. Right side: The ANN architecture consisting of four layers of neurons that are fully connected. Neurons in the input layer and the output layer correspond to the line integrated measurements and the grid pixels of the phantom, respectively

and blue, respectively. These cameras can achieve speeds of up to 204,800 frames per second (fps) with a resolution of 1280×8 pixels in 12-bit ADC dynamic range [17]. Each pixel has a size of $10\mu m \times 10\mu m$, and the cameras operate in the visible spectral range. In the current work, an ANN-based tomography model is trained to predict radiation function of one cross-section using the images captured by these cameras.

Training Dataset

To construct a synthetic training dataset, 4000 emissivity phantoms of one GOLEM poloidal cross-section with associated line integrated data was used. The left side of Fig. 2 shows the phantom simulated on a square rectilinear grid with the appropriate pixel size. The line integrated measurements represent the data measured by the LoS of the cameras detectors, considered as the input of the neural network as illustrated on the middle side of Fig. 2. The intensity of the incident light radiation on the i-ith detector of each camera is given by

$$\mathbf{f_i} = \sum_j \mathbf{T_{ij}} \, \mathbf{g_j},\tag{1}$$

where T_{ij} is the element of the geometric matrix describing how the radiation emitted from the plasma located in j-th pixel of phantom contributes in the data measured by i-th detector [1]. In such training dataset, f_i, calculated using the known function g_j of phantom, and the phantom itself are respectively considered as the input and output. Then the trained ANN model will be able to reconstruct the unknown function g_j for an unseen sample from the measured data f_i.

Specifically, two images captured by R and V cameras and the corresponding radiation distribution of one cross section will be the input and output data of the trained model, respectively. In this representation, the input and the output are described by arrays $X = [R_{1o}, ..., R_{io}, ..., R_{Io}, V_{1o}, ..., V_{ko}, ..., V_{Ko}]$ and $Y = [Z_{11}, ..., Z_{mn}, ..., Z_{MN}]$, respectively. The array elements R_{io}, V_{ko} and Z_{mn} are, respectively, the data corresponding to the middle line (o) of the discretized R image, V image and the plasma region grid with an M × N resolution, where the middle line of each image is considered for tomography reconstruction. By selecting a spatial resolution of 1280×56 pixels for the cameras and a phantom with a square rectilinear grid size of 40×40 pixels, the number of input and output features are 2560 and 1600, respectively.

While ANN's are generally less reliable at predicting outside the range of the training data, the database was diversified to include various shapes (such as Gaussian, Hollow, and Banana shapes) in a wide range of intensity and position. However, ANN's, especially deeper ones, can combine the detected features of dataset in non-linear ways [19]. Furthermore, each feature in the training dataset is normalized to ensure equal contribution to the model. The test data are then normalized using the parameters obtained during training.

Training Process

To train the predictive model, a neural network architecture specifically designed for the training dataset was developed. The neural network was modeled with an input layer of 2560 neurons (number of input features), two hidden layers of 640 and 320 neurons each, and an output layer of 1600 neurons (number of output features). A schematic representation of

the ANN framework is depicted on the right side of Fig. 2 illustrating the inputs and outputs used to train the ANN model.

The key considerations in setting the training parameters include the choice of optimization algorithm, epochs, batch size, learning rate, regularization techniques, and selection of an appropriate loss function. The number of epochs specifies how many times the entire training dataset is passed through the network which should be addressed for effective learning from the training data. A fixed learning rate promotes a stable training process, facilitating better generalization through consistent and gradual updates to the model's parameters. Additionally, regularization techniques like dropout [20] help prevent overfitting by randomly deactivating some neurons during training. Early stopping is another critical regularization technique that halts training when validation loss diverges from the training loss, preventing model overfitting.

To optimize the training process, we employed Adam, a popular variant of stochastic gradient descent (SGD). The training setup was empirically selected, featuring a mini-batch size of ten samples, a learning rate of 0.0001, and 1500 epochs. The model was trained by using eighty percent of the dataset (training dataset) and the remaining twenty percent (validation dataset) was used to validation of the trained model. The cost function chosen to evaluate the deviation between the actual output value and the predicted value obtained by the network was the Mean Square Error (MSE). Figure 3 illustrates the trend of the loss function value for the training and validation datasets during the training process of the neural network. It shows that the losses in both the training and validation processes decrease gradually and are close to each other. Such variation indicates that the model is learning patterns without memorizing the training data (overfitting) that generalize to new, unseen data.
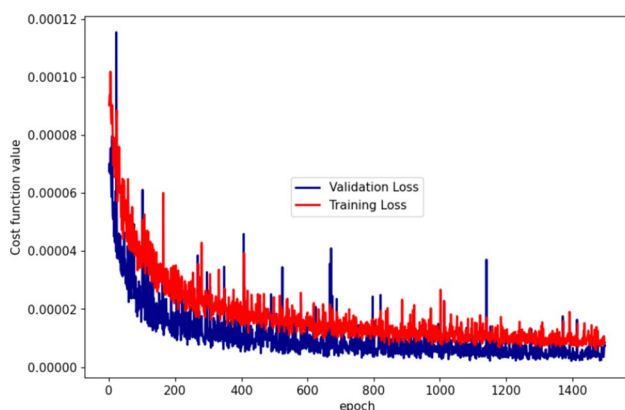

Fig. 3 The trend of the loss function value for the training and validation dataset during the training process of the neural network

Results and Discussion

The trained ANN model was applied on three unseen phantom samples to predict the radiation function corresponding to their line integrated measurements. The samples have various shapes similar to those in the training dataset. Subsequently, the backfit was evaluated to compare the line-integrated measurements of the ANN's predictions with those of the phantom samples. Figure 4 shows (from left to right) the phantom sample, the ANN prediction of radiation function and the corresponding backfit of line-integrated measurements for three unseen samples (from top to bottom). The result shows that the trained ANN model predicts the radiation function of samples very near to corresponding phantom. The backfit analysis confirms the reliability of the proposed ANN model in reconstructing the radiation function. However, noticeable fluctuations in backfit are observed in certain spatial coordinates of the grid pixels for ring shape sample.

To evaluate the performance of the ANN model predictions on the samples that are dissimilar to the training dataset but represent a mix of those samples, two ANN models was trained by two different training dataset. The first model, Model_1 was trained by using the training dataset consisting of various shapes such as Gaussian, Hollow, Banana shapes and mix of them. The second model, Model_2 was trained by using the training dataset consisting of various shapes such as Gaussian, Hollow, Banana shapes but without mix of them. The two trained models was performed to predict one sample which exist in the first training dataset (Model_1) but not in the second one but it is a mix of samples existing in the second one. As it is shown in Fig. 5, the phantom is a mix of Gaussian and banana shapes. Figure 5 shows (from left to right) the phantom sample, the ANN prediction of Model_1, the ANN prediction of Model_2 and the corresponding backfit of line integrated measurements, respectively. The result shows that the model trained with a dataset dissimilar to the unseen samples (Model_2) can recognize the mixed shapes, but it does not provide accurate predictions in certain spatial coordinates. Further training with a more diverse training dataset like Model_1 may be necessary to improve the model's accuracy in these specific areas.

The ANN model requires time in order of 10 ms for prediction, which is significantly faster compared to the traditional tomography reconstruction time of around 3 s. The computations were done on the same device-regular laptop. This demonstrates a remarkable improvement in terms of speed, highlighting the efficiency of the ANN model in this application.

The optimization of the ANN model's performance is influenced by various factors, including the quantity and quality of training data, data pre-processing, and feature
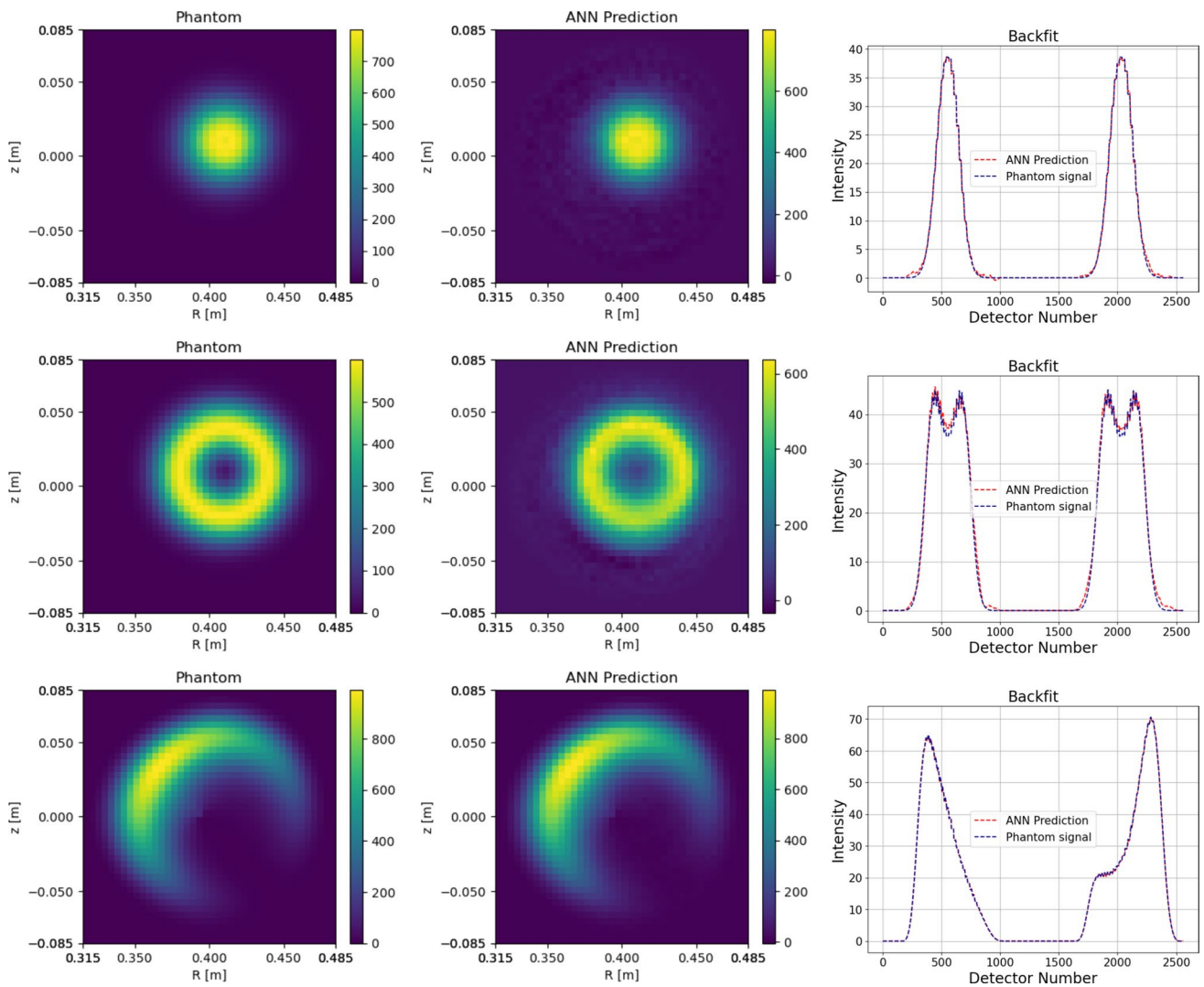
Fig. 4 From left to right: the phantom sample, the ANN prediction of radiation function and the corresponding backfit of line integrated measurements for three unseen data samples (from top to bottom)
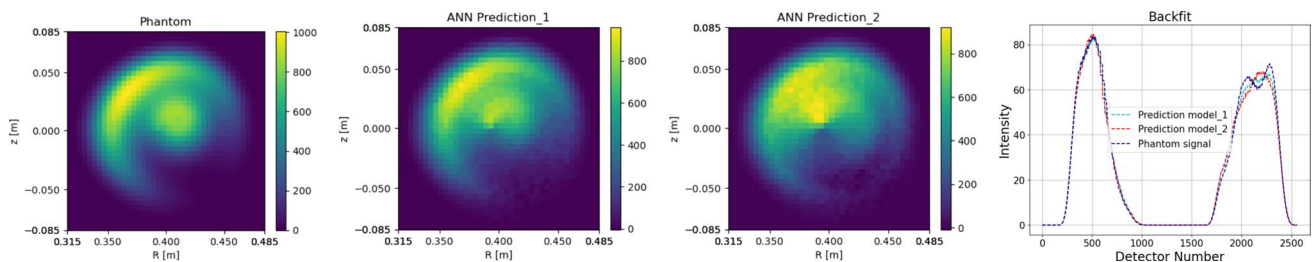

Fig. 5 From left to right: Phantom sample, the prediction of Model_1, the prediction of Model_2 and the corresponding backfit of line integrated measurements, respectively

reduction and selection. The work plan for the future involves enhancing these aspects. Additionally, pre-processing methods to handle missing values to replace these missing values can improve the results in the context of being sparse the data. This can be achieved by incorporating additional diagnostic data. Furthermore, validating prediction data with n diagnostic inputs can further enhance model robustness and accuracy.

Conclusion

The paper presents an artificial neural network model applied to predict the visible plasma radiation distribution at the GOLEM tokamak. The training dataset was constructed using samples consisting of emissivity phantoms and associated line integrated measurements corresponding to one poloidal cross-section of GOLEM tokamak. The dataset was defined with different parameterization in distribution shape (Gaussian, Hollow, and Banana shapes) and with a different range of intensity value and size.

The backfit analysis of line-integrated measurements confirms the reliability of the trained ANN model in reconstructing the radiation function. However, significant variations in backfit are observed at certain spatial coordinates of the grid pixels and also in unseen samples dissimilar to the training dataset. To address this, future work will focus on optimizing the ANN model's performance, considering factors such as the quantity and quality of training dataset.

One of the key advantages of the ANN prediction model is its significantly shorter prediction time (approximately 10 ms) compared to traditional tomography reconstruction methods (approximately 3 s).

References

1. J. Mlynar, G. Bonheure, V. Weinzettl, A. Murari, JET-EFDA CONTRIBUTORS, Inversion techniques in the soft-x-ray tomography of fusion plasmas: toward real-time applications. Fusion Sci. Technol. **58**(3), 733–741 (2010)

2. P.J. Carvalho, H. Thomsen, R. Coelho, P. Duarte, C. Silva, H. Fernandes, Isttok plasma control with the tomography diagnostic. Fusion Eng. Des. **85**(2), 266–271 (2010)

3. J. Mlynar, T. Craciunescu, D.R. Ferreira, P. Carvalho, O. Ficker, O. Grover, M. Imrisek, J. Svoboda, JET contributors, Current research into applications of tomography for fusion diagnostics. J. Fusion Energ. **38**, 458–466 (2019)

4. P. Clemente Angioni, T. Pütterich, M. Mantica, M. Valisa, E.A. Baruzzo, P. Belli, F.J. Belo, C. Casson, P. Drewelow. Challis et al., Tungsten transport in jet h-mode plasmas in hybrid scenario, experimental observations and modelling. Nucl. Fusion **54**(8), 083028 (2014)

5. Diogo R. Ferreira, Pedro J. Carvalho, Ivo S. Carvalho, Chris Stuart, Peter J. Lomas, J.E.T. Contributors, Monitoring the plasma radiation profile with real-time bolometer tomography at jet. Fusion Eng. Des. **164**, 112179 (2021)

6. W. Zheng, X.U. Fengming, S.H. Chengshuo, Y. Zhong, A.I. Xinkun, C.H. Zhongyong, D.I. Yonghua, M. Zhang, Y.A. Zhoujun et al., Overview of machine learning applications in fusion plasma experiments on j-text tokamak. Plasma Sci. Technol. **24**(12), 124003 (2022)

7. D. Wroblewski, G.L. Jahns, J.A. Leuer, Tokamak disruption alarm based on a neural network model of the high-beta limit. Nucl. Fusion **37**(6), 725 (1997)

8. K.H. Jin, M.T. McCann, E. Froustey, M. Unser, Deep convolutional neural network for inverse problems in imaging. IEEE Trans. Image Process. **26**(9), 4509–4522 (2017)

9. D.R. Ferreira, P.J. Carvalho, H. Fernandes, J.E.T. Contributors, Full-pulse tomographic reconstruction with deep neural networks. Fusion Sci. Technol. **74**(1–2), 47–56 (2018)

10. A. Jardin, J. Bielecki, D. Mazon, J. Dankowski, K. Król, Y. Peysson, M. Scholz, Neural networks: from image recognition to tokamak plasma tomography. Laser Part. Beams **37**(2), 171–175 (2019)

11. D.R. Ferreira, P.J. Carvalho, H. Fernandes, Deep learning for plasma tomography and disruption prediction from bolometer data. IEEE Trans. Plasma Sci. **48**(1), 36–45 (2019)

12. X. Liang, Z. Liu, H. Chang, L. Zhang, Wireless channel data augmentation for artificial intelligence of things in industrial environment using generative adversarial networks. in *2020 IEEE 18th International Conference on Industrial Informatics (INDIN)*, IEEE, vol. 1, pp. 502–507 (2020)

13. L.S. Van Leeuwen, Machine learning accelerated tomographic reconstruction: for multispectral imaging on TCV. Master's thesis, Eindhoven University of Technology (2022)

14. S. Abbasi, J. Chlum, J. Mlynar, V. Svoboda, J. Svoboda, J. Brotankova, Plasma diagnostics using fast cameras at the golem tokamak. Fusion Eng. Des. **193**, 113647 (2023)

15. J. Svoboda, J. Cavalier, O. Ficker, M. Imríšek, M. Hron, Tomotok: python package for tomography of tokamak plasma radiation. J. Instrum. **16**(12), C12015 (2021)

16. S. Abbasi, J. Mlynar, J. Chlum, V. Svoboda, J. Svoboda, O. Ficker, J. Brotankova, Machine-learning-based reconstruction of spatial distribution of plasma radiation using color visible cameras at golem tokamak. in *21st Conference of Czech and Slovak Physicists*, Proceedings. Slovak Physical Society, ISBN 978-808985521-6, pp. 59–60 (2023)

17. Photron Europe Limited. Product datasheet Mini UX Fastcam series by photron, (2021)

18. C.M. Bishop. *Neural networks for pattern recognition.* Oxford university press, (1995)

19. Y. LeCun, Y. Bengio, G. Hinton, Deep learning. Nature **521**(7553), 436–444 (2015)

20. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. **15**(1), 1929–1958 (2014)