MHD mode tracking using high-speed cameras and deep learning

OPEN ACCESS
IOP Publishing
Plasma Physics and Controlled Fusion
Plasma Phys. Control. Fusion **65** (2023) 074002 (10pp)
https://doi.org/10.1088/1361-6587/acd581

MHD mode tracking using high-speed cameras and deep learning

Y Wei[1,*], J P Levesque[1], C Hansen[1,2], M E Mauel[1] and G A Navratil[1]

[1] Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY 10027, United States of America
[2] Department of Aeronautics & Astronautics, University of Washington, Seattle, WA 98195, United States of America

E-mail: yw2714@columbia.edu

Received 23 January 2023, revised 25 April 2023
Accepted for publication 15 May 2023
Published 30 May 2023

Abstract

We present a new algorithm to track the amplitude and phase of rotating magnetohydrodynamic (MHD) modes in tokamak plasmas using high speed imaging cameras and deep learning. This algorithm uses a convolutional neural network (CNN) to predict the amplitudes of the $n = 1$ sine and cosine mode components using solely optical measurements from one or more cameras. The model was trained and tested on an experimental dataset consisting of camera frame images and magnetic-based mode measurements from the High Beta Tokamak - Extended Pulse (HBT-EP) device, and it outperformed other, more conventional, algorithms using identical image inputs. The effect of different input data streams on the accuracy of the model's predictions is also explored, including using a temporal frame stack or images from two cameras viewing different toroidal regions.

Keywords: MHD mode tracking, fast camera, diagnostics, control, machine learning, HBT-EP

(Some figures may appear in colour only in the online journal)

1. Introduction

The ability to accurately and robustly identify and track magnetohydrodynamic (MHD) instabilities and other modes is an important diagnostic capability for future tokamak-based fusion reactors. Such an observer is a necessary feature of proposed techniques to forecast and avoid disruptive stability boundaries [1–4] and active feedback systems for the control of internal [5–7] and external instabilities [8–11]. In present machines this is usually done with arrays of nearby magnetic Mirnov coils. However, such diagnostics may not be well suited for the long-pulse and harsh environment present in reactors [12]. As a result, observers capable of utilizing new and distant (e.g. optical) diagnostics are desirable to support future tokamaks.

Using the High Beta Tokamak - Extended Pulse (HBT-EP) device, extensive studies have been done on MHD mode diagnostics and real-time mode tracking. Beyond established approaches using in-vessel magnetic probes [8, 13], methods using non-magnetic sensors have also been explored. These include using arrays of extreme ultraviolet (EUV) sensors [14, 15], electrodes placed in the plasma scrape-off-layer [16], as well as using a visible-range high-speed imaging camera [17]. Among these methods the high-speed camera has the advantage of being an external optical diagnostic. It is relatively easy to implement using mostly off-the-shelf components, and it can be adjusted and maintained independent of the machine's maintenance cycles. The challenge, however, is that measurements on the camera have a non-trivial relation to the plasma's 3D emission structure. This also couples to the complex emission mechanism which depends on plasma and wall parameters, many of which may not be obtainable with sufficient accuracy in real-time. These make first-principle

* Author to whom any correspondence should be addressed.

implementations of real-time mode tracking using camera data particularly challenging and motivates the study of algorithms using data-driven methods such as deep learning.

Deep learning [18] as well as other machine learning techniques have been applied to a variety of areas in fusion and plasma science. While initial studies have mainly focused on forecasting transient events, in particular disruptions in tokamaks, through either supervised [19–23] or unsupervised [1, 3] approaches, recently additional areas have also been explored. These include but are not limit to tomographic reconstruction [24, 25], equilibrium profile modeling [26–29], and the identification of reduced-order models [30–32]. Novel techniques such as reinforcement learning [33–35], reservoir computing [36], and object detection algorithms [37, 38] have also been demonstrated in fusion-relevant applications.

In this work, we present the first deep learning based MHD mode tracking algorithm using high-speed camera diagnostics. This algorithm utilizes the convolutional neural network (CNN) architecture: it takes in a single or multiple images from one or two cameras to predict the amplitudes of the sine and cosine components of long-wavelength $n = 1$ kink and tearing modes, similar to those described in [10, 11]. We developed this model using a dataset of camera images and magnetic mode signals measured during plasma discharges on HBT-EP, and we found this algorithm to perform better than the other two tested algorithms using linear regression and singular value decomposition (SVD) dominant mode-pair reconstruction with the same inputs.

The remaining sections are organized as follows: section 2 introduces the HBT-EP device, the high-speed camera diagnostic system, and the experimental database. Section 3 describes the implemented CNN model and two other comparison models. The results of these three models using individual frames from one camera as inputs are compared in section 4, and section 5 explores methods on further improving the CNN model's predictions through including additional temporal or spatial information. We conclude with a summary and proposed future work.

2. Experiment and database

The data used in this study were collected from a dedicated run campaign conducted on HBT-EP. HBT-EP is a circular, ohmically heated, large aspect ratio tokamak ($R/a \sim 6$). It has an on-axis toroidal magnetic field of 0.33 T, typical plasma currents of 15 kA, a major radius of 0.92 m, and a plasma minor radius of 0.15 m. Figure 1 shows the hardware of the high-speed camera diagnostics on HBT-EP. The two cameras (Phantom S710) are placed away from the toroidal field (TF) coils and collect light through object lenses that are mounted close to an optical viewing port. These cameras measure visible light emission from the plasma edge due to plasma-neutral and plasma-wall interactions, predominantly at the D_α line [17]. The object lenses are connected to the cameras using coherent optical fiber bundles. Between each fiber bundle and camera, a custom 2:1 demagnification relay setup was installed using a 25 mm f/0.95 and a 50 mm f/1.4 prime lens. This relay setup increases the light intensity received by the camera sensor at the expense of reduced image resolution.

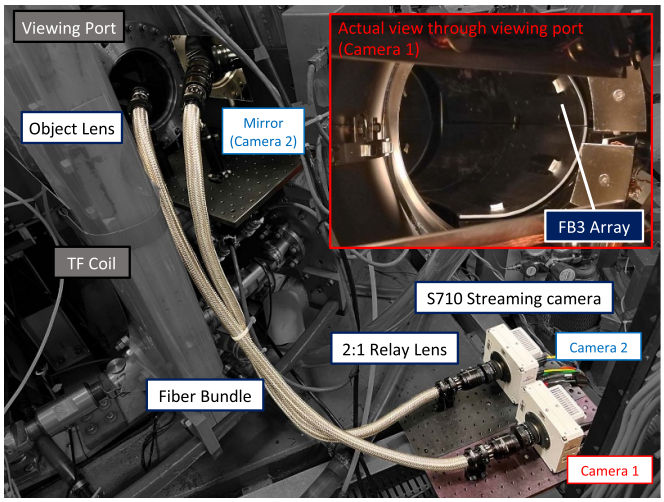

Figure 1. Picture of the high-speed camera diagnostics on HBT-EP. Each camera diagnostic consists of an object lens, an optical fiber bundle, a custom relay setup, and a streaming camera placed on an optical breadboard away from the TF coil. An actual view through the glass viewing port similar to the view of Camera 1 is shown in the top right inset. The location of one sensor of the in-vessel, low-field-side toroidal magnetic sensor array (FB3) used to determined the ground truth $n = 1$ sine and cosine components is indicated in the inset figure.

Figure 2 shows the camera setup used in the run campaign for this work. The two cameras observed the plasma with views in opposite toroidal directions, and their central line of sights were about 90° apart toroidally. Both cameras were set to run at 250 kfps, 128 × 64 pixels (width × height) resolution, 12-bit grayscale bit depth, and 3 μs exposure time. Figure 3 shows samples of camera images taken during a discharge, along with an example SVD-based reconstruction of the dominant rotating mode. Note that the SVD reconstruction method used for mode tracking over multiple shots in the following sections (as described in section 3) is different than for this example; the example shown here is purely to aid in visualization. As a mode distorts the plasma, the associated density and temperature perturbations cause variations in local interactions between the plasma, neutral gas, and the walls, producing variations in local light emission. This temporal and spatial variation in local emission is how the cameras observe the mode, through direct observation of emitting volumes and also via reflections from the walls. We used the unprocessed camera images as the inputs to the algorithm in order to avoid the time complexity of applying additional spatial or temporal filters in a real-time system.

We used the sine and cosine components of the $n = 1$ mode determined by a least squares fit to the signals measured from an in-vessel, low-field-side magnetic sensor array (FB3) as the ground truth (or labels), which the model is trained to reproduce using the camera frames as the inputs. To remove high frequency noises in the ground truth signals we applied a

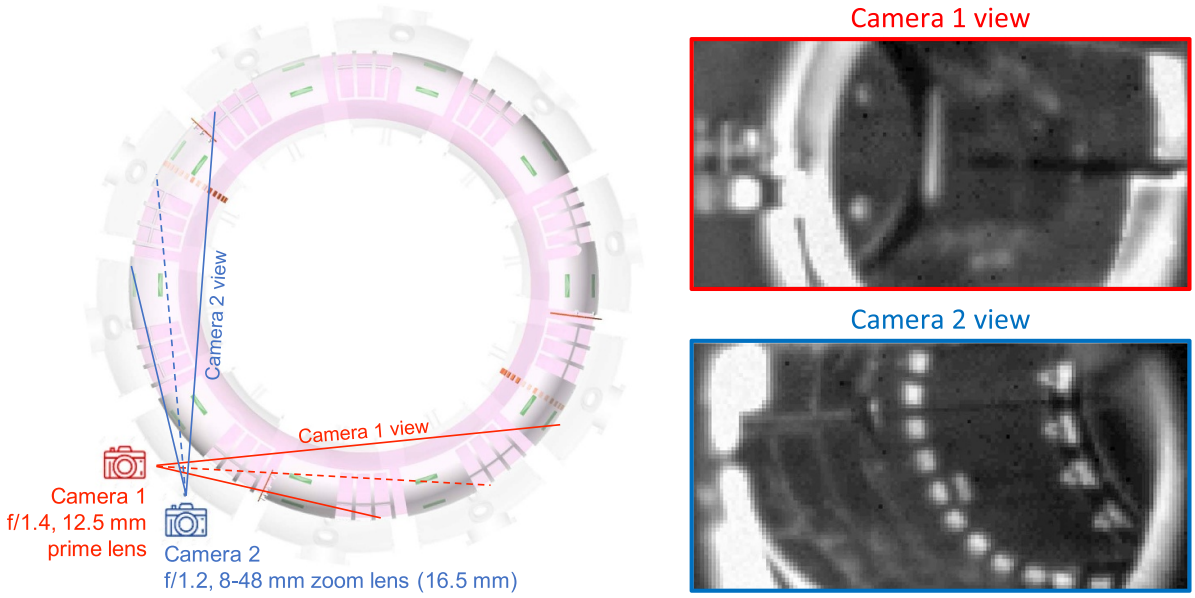

Figure 2. Left: illustration of the camera views in the HBT-EP device (top–down). The two cameras look at the cross-section of the plasma (purple) toward opposite toroidal directions. Parameters of the object lenses are shown under each camera. Right: camera images (128 × 64 pixels resolution, exposure & contrast adjusted) taken during alignment using the experimental setup.

5 kHz low-pass filter to the mode amplitude determined from the mode components, before converting the filtered amplitude and the unfiltered phase back to sine and cosine components. This data processing routine has been used extensively in previous mode control experiments on HBT-EP and used for evaluating other non-magnetic mode tracking methods. By accurately reproducing the mode signal measured by these sensors, the outputs of the camera-based tracking algorithm could be fed directly to the existing downstream active feedback controllers and should be able to achieve similar mode suppression results. It should be noted that mode measurements from the magnetic sensor array contain intrinsic errors due to resolution limitations, sensor noise, and/or the utilized decomposition algorithm compared to the properties of the physical mode present in the plasma; however, such an analysis is beyond the scope of this work and in the following sections we define the errors of the camera-based mode tracking algorithm as the difference between its predictions and the magnetic mode signals.

In these experiments, we targeted a specific discharge evolution from previous mode control experiments. Figure 4 shows selected plasma parameters of a typical discharge of this shot style. We used the entire duration of each discharge from about 0.6 ms after breakdown up to the thermal quench as indicated by the current spike event. Because of evolution of the edge safety factor (q_{edge}), both $m/n = 4/1$ and 3/1 modes (where m and n are the poloidal and toroidal mode number respectively) were present in these shots, with timing and characteristics dependent on the details of each shot. Within this database, the rotation frequencies of the modes were typically in the range around 8–12 kHz. We did not target a specific mode shape and included both the 4/1 and 3/1-dominant time periods.

The final compiled database consists of 45 shots which corresponds to 86 275 time slices, also known as *samples*. Each sample consists of a frame image from each camera and the amplitudes of the $n = 1$ sine and cosine components interpolated at the corresponding time point. We used the first 40 shots as the training set (76 975 samples) and the last five shots as the testing set (9300 samples) to mimic the data acquisition process during a control experiment. To optimize the neural network model's hyperparameters, we further split the training data by samples (instead of by shots) randomly using a 9:1 ratio (training:validation). The final models were trained using all data from the 40 training shots.

3. Algorithms

We implemented the deep learning model in Tensorflow [39] using a CNN. The CNN model consists of two parts: a feature extractor using three Conv2D and MaxPooling2D layers, and a regressor using two Dense (Fully Connected) layers. Table 1 summarizes the hyperparameters of the optimized CNN model. The model takes in a single image from a camera and returns the predicted sine and cosine components of the $n = 1$ mode. During training the model's predictions are compared to the ground truth values from the database, and their differences (mean squared error) are used to update the network's coefficients through back-propagation. We used the Adam optimizer and a step decay learning rate scheduler for training the network.

In addition to the deep learning model, two conventional algorithms were implemented for comparison. These are:

1. A baseline regression model using linear regression. This model was trained on the same set of inputs and labels as

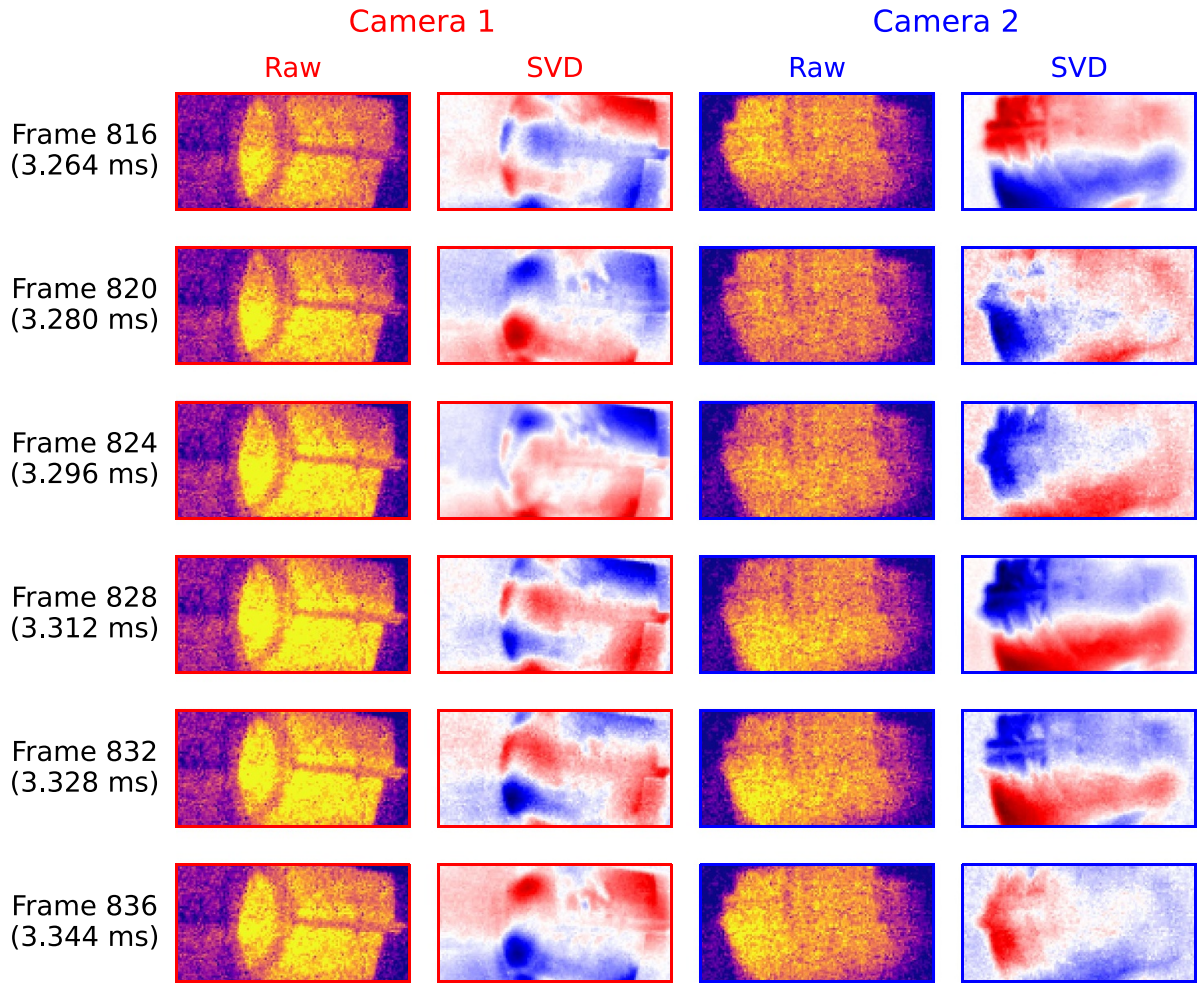

Figure 3. Raw and SVD-reconstructed frame images of shot 114461, using the camera views shown in figure 2. The SVD-reconstructed images are included to better illustrate the structure and phase of the observed MHD mode. Left sub-columns: raw camera measurements (adjusted for dynamic range). Right sub-columns: reconstructed images using the dominant SVD mode-pair (symmetric color scale). SVD modes here are calculated using the raw frame images of each of the cameras over the duration of 3–4 ms of this discharge. The time range shown corresponds to roughly a half-period of mode rotation, as revealed by the shifting of red and blue stripes in the SVD-reconstructed images over time.

the CNN model, using data from the 40 training shots. The model takes in a single frame image and returns the predicted mode components.

2. An algorithm using the SVD dominant mode-pair reconstruction method as proposed in [17] and implemented in [14] for different diagnostics. The SVD modes were calculated from the 4/1-dominant period (2–4 ms) of shot 114461. This shot, denoted as the "training" shot, was taken during the same run day right before the testing shots so that the plasma conditions would be as similar as possible. The spatial modes of the first SVD mode pair were taken as the sinusoidal basis of the dominant MHD mode. Performing a dot-product of this basis with a given image produces the predicted component amplitudes of the optical fluctuation. From our tests, we found this method only gave valid results if the decomposition window of the "training" shot contained a clear and distinct m/n mode shape. For this reason we specifically chose the 4/1-dominant period instead of using the entire shot duration. An additional gain and phase shift correction matrix was introduced to match the optical fluctuations with actual magnetic mode signals and was determined from the 4/1-dominant period of the following shot (114462).

4. Results

The following section compares the predictions of the 3 candidate models discussed in the previous section. All 3 models take in the same input data (a single frame from Camera 1) and return the predicted $n = 1$ sine and cosine components. We then converted the predictions back to amplitude ($A_t = \sqrt{{s_t}^2 + {c_t}^2}$) and phase ($\phi_t = \arctan(s_t/c_t)$, s_t and c_t being the predicted sine and cosine components at time t from each of the mode tracking algorithms) to give a more intuitive comparison.

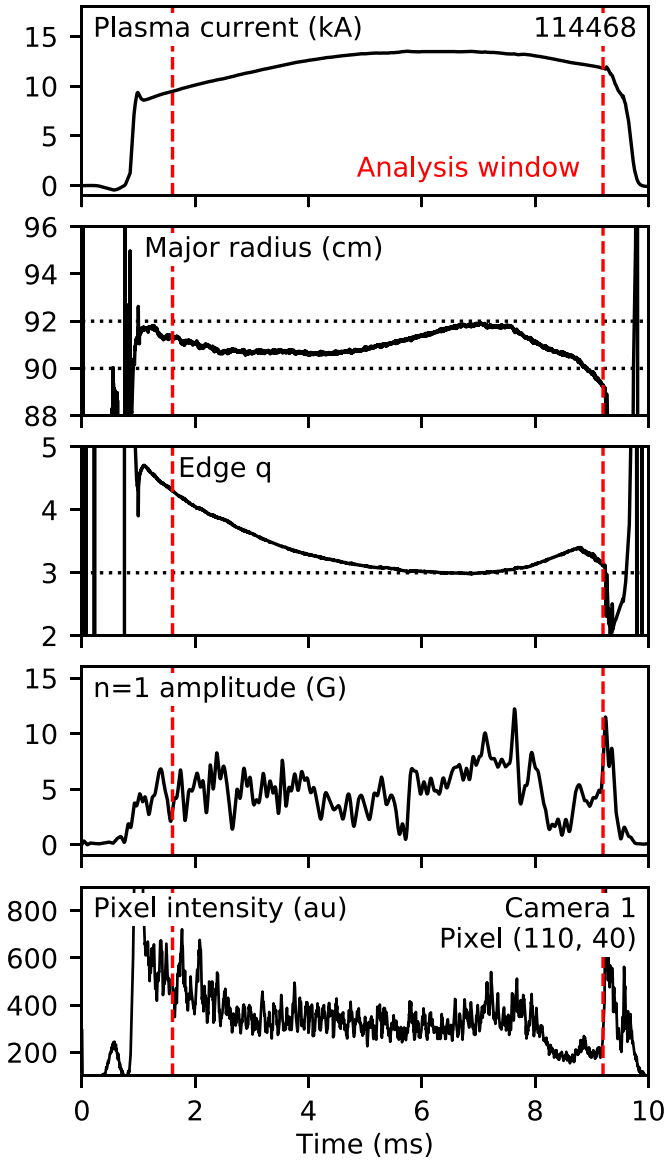

Figure 4. Time evolution of the plasma current (I_p), major radius (R), edge safety factor (q_{edge}), $n = 1$ mode amplitude, and intensity of Camera 1 pixel (110, 40) of sample shot 114468 from the run campaign. The analysis window is indicated by the two dashed lines, the first one after the plasma breakdown, and the second one at the thermal quench. The plasma parameters (I_p, R, q_{edge}, etc) were only used during analysis and are not accessible to any of the mode tracking algorithms.

4.1. Single shot tracking

Figure 5 shows the amplitude and phase tracking results of testing shot 114467 given by the 3 candidate models. The 4/1 and 3/1-dominant periods (cyan and yellow respectively) are also highlighted in the figures. These were identified from an in-vessel poloidal magnetic sensor array, and the information was not given to any of the models during training or when performing inference.

We found the CNN model (blue, second row in figure 5) gave the most accurate amplitude and phase predictions with the smallest amount of error. Its amplitude prediction closely matches the ground truth across both the 4/1 and 3/1-dominant periods as well as in the transition period during which the mode amplitude was lower. The phase prediction shows a uniform error bound independent of the variation in mode amplitude, except in the short time periods when the amplitude was very small, such as at 6 and 8.4 ms. This may be caused by the lower signal-to-noise ratio of the magnetic measurement during similar low-amplitude periods from the training shots which reduced the quality of the training labels and resulted in worse prediction given by the trained model. For the purpose of mode control we are only interested in the time periods when the mode amplitude is significant, therefore we do not expect these issues would affect the outcome during a control experiment.

The predictions of the other two methods are less accurate than those given by the CNN model. In the case of the linear regression model (orange, third row in figure 5), its predictions show similar behaviors as the CNN model, although the noise levels are substantially higher in both amplitude and phase. For the SVD mode-pair reconstruction method (red, fourth row in figure 5), both its amplitude and phase predictions are only relatively accurate, although with significant amount of noise, during the 4/1-dominant period (aqua shading in figure 5). In the rest of the shot its amplitude prediction misses the two periods of rapid mode growth at 5.5 and 7 ms, and the phase prediction is very inaccurate. As has been discussed in section 3, the SVD mode basis was obtained from the same 4/1-dominant period of the "training" shot, and because of this patterns in the new input images need to be sufficiently similar to the 4/1 structure represented in the SVD basis in order for the algorithm to give valid predictions. We have also tested performing SVD on the 3/1-dominant period and have observed the reciprocal result; the SVD model then performs poorly during the 4/1-dominant period while being better in the 3/1-dominant period which the SVD basis is computed on.

Table 1. List of hyperparameters of the CNN model. The neural network consists of three Conv2D-MaxPooling2D layers followed by two hidden Dense (Fully Connected) layers and the output layer. The model takes in a single camera image of dimension 128×64 and returns predicted $n = 1$ sine and cosine component amplitudes.

Input layer	Grayscale image stack, shape = (64, 128, n_frames)
Convolution layers	Conv2D(8) - MaxPooling2D Conv2D(8) - MaxPooling2D Conv2D(16) - MaxPooling2D kernel = (3,3), activation = 'relu', padding = 'valid', pool = (2,2)
Dense layers	Dense(256) - Dense(64), activation = 'relu'
Output layer	Dense(2, 'linear')
Loss function	mean squared error
Optimizer	Adam with step decay schedule (initial = 1×10^{-3}, reduce by 0.5 every 15 epochs)
Epochs	50

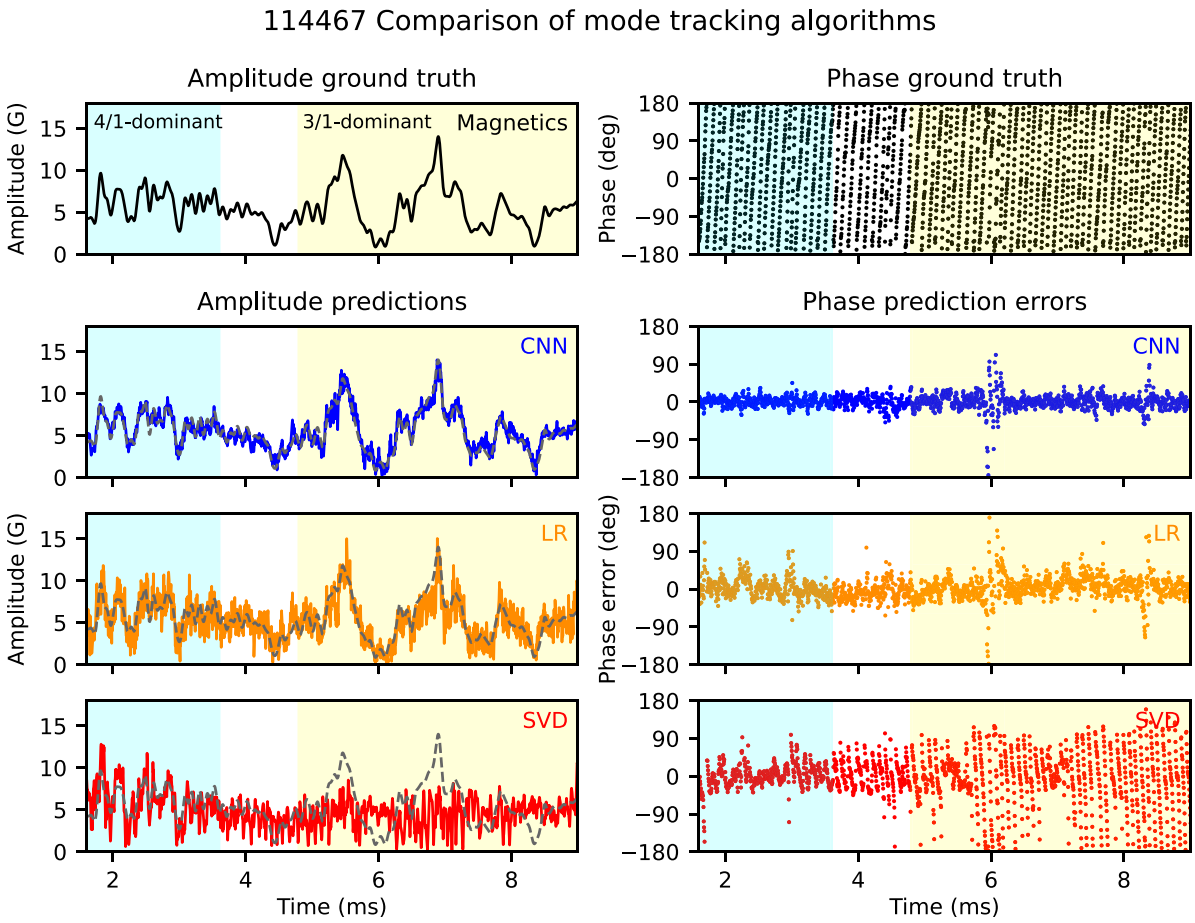

Figure 5. Single shot tracking results of testing shot 114467 using CNN (blue), linear regression (orange), and SVD dominant mode-pair reconstruction (red). The 4/1 and 3/1-dominant periods are highlighted in cyan and yellow respectively. Top left: amplitude ground truth from magnetic sensors. Lower left: amplitude predictions given by the candidate models (solid) overlaid onto the ground truth (dotted). Top right: phase ground truth. Lower right: errors of phase predictions given by the candidate models.

4.2. Testing set distribution

We tested the three candidate models using all data from the testing set. Figure 6 shows the distributions of amplitude and phase errors of the three models. These distributions exclude data from time periods where the mode amplitude was insignificant (defined as true amplitude <3 G) so that the results are relevant for control experiments. For each model the solid lines show the distributions of data from the entirety of each shot, and the dashed lines show data from only the 4/1-dominant periods.

In accordance with the results from single shot tracking, we found the CNN mode gives the narrowest error distributions in both amplitude and phase predictions and significantly outperforms the other two methods. From these distributions the errors of the CNN model can be estimated to be around ± 1 G in amplitude and $-15°$ to $+20°$ in phase (with respect to ground truth magnetic measurements), both are estimated at the half maximum of the distributions. The linear regression model, while having wider error distributions than the CNN model, shows a slight advantage over the SVD-based model even if we only consider the 4/1-dominant period of the shots.

5. Investigating the effect of multiple input frames

For the previously discussed results, we have considered a CNN model using only individual frames from one camera as input. To further explore possible inputs, we tested supplying the model with additional temporal information by using multiple sequential frames from a single camera, or spatial information by using images from both cameras simultaneously. figure 7 illustrates the modified network architectures of these two approaches. The convolution and dense blocks have the same number of layers and sizes as described in table 1, although the number of parameters in each layer may be different from the original model due to the increased input size. In addition, while it is possible to also add temporal information to the two-camera model, we chose not to investigate that feature in this work as it would further increase the network's complexity. We trained each modified model using the same 40 training shots and tested them on the five testing shots.

First, to add temporal information we modified the model inputs by stacking the current frame at time t_i with a number of historic frames at t_{i-1}, t_{i-2}, etc (figure 7(A)). This way the

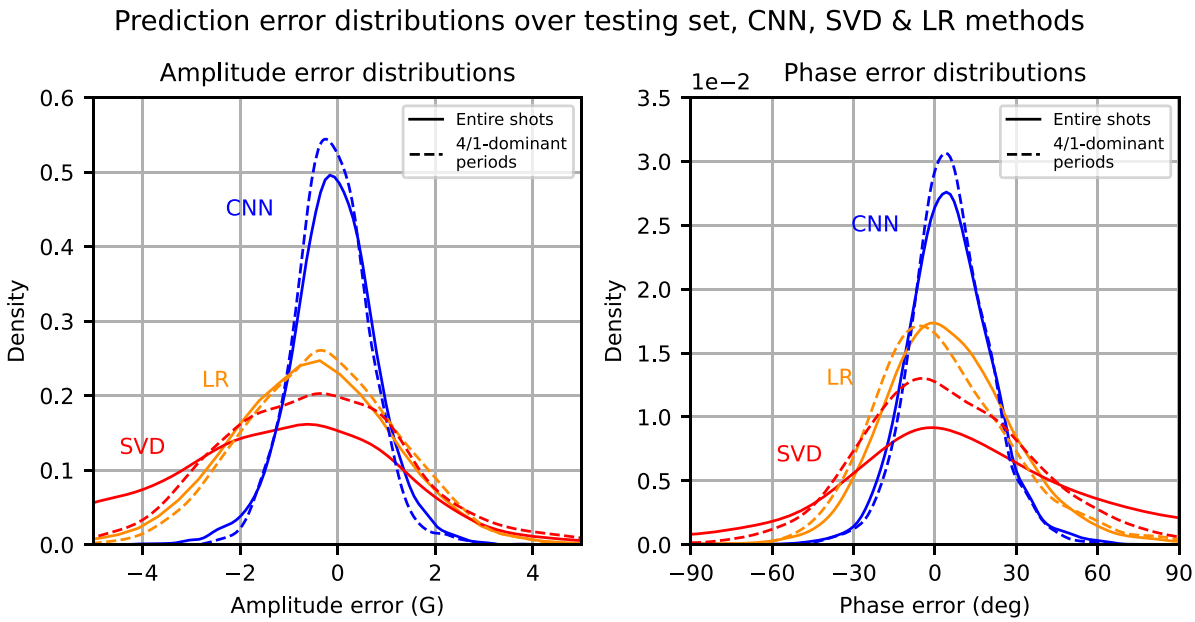

Figure 6. Distributions of amplitude (left) and phase (right) errors for the predictions given by CNN (blue), SVD mode-pair reconstruction (red), and linear regression (orange) over the five testing set shots where ground truth amplitude >3 G. For each method, the solid lines include data over the entire shots, while the dashed lines use a subset of data from the 4/1-dominant period.

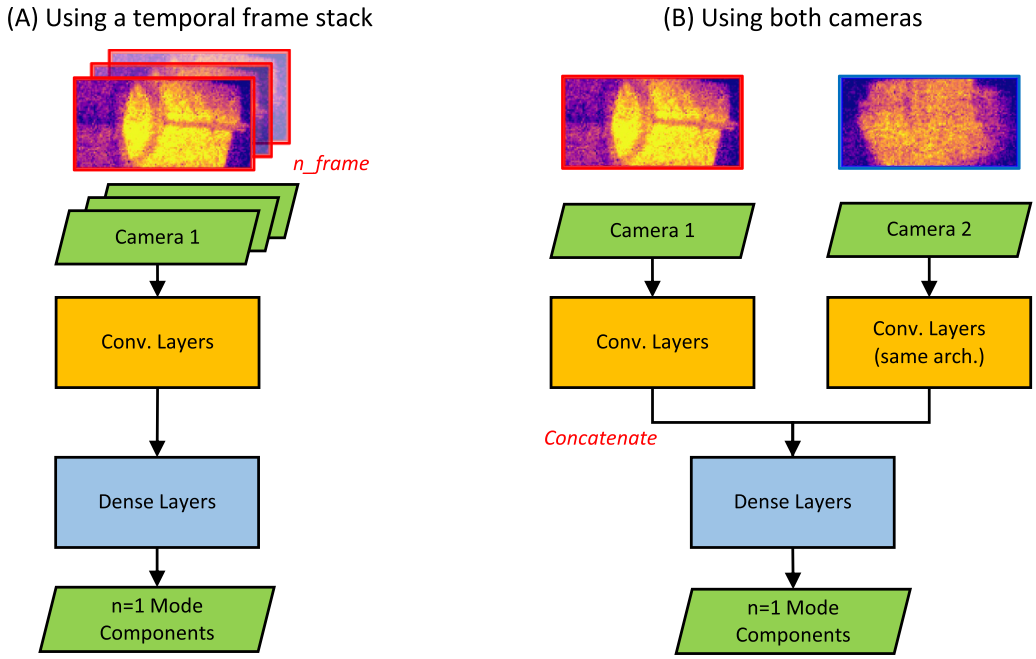

Figure 7. Illustrations of the modified neural network architectures with additional (A) temporal and (B) spatial information added to the CNN model. The convolutional and dense layers for each network have the same hyperparameters given in table 1.

shape of the input array becomes (128, 64, `n_frame`) where `n_frame` indicates the depth of the frame stack. The modified model then takes in this three-dimensional frame stack and predicts the same mode components at the current time point t_i. We used a rolling window so that the number of samples in each shot remained identical. The results are shown in figure 8, in which we compared the error distributions of the original one-camera one-frame model shown in section 4 with three modified models with `n_frame` = 2, 4, and 6. Figure 8 shows that using a historic frame stack slightly improves the prediction in both amplitude and phase; however the effect appears to saturate beyond using a stack of four frames.

Second, to include additional spatial information we added the frame image from the second camera at the same time t_i into the model inputs (figure 7(B)). The line of sight of this camera ("Camera 2" in figure 2) was roughly 90° toroidally apart from Camera 1, which we assumed would give a quadrature effect and help improve the model's predictions. The

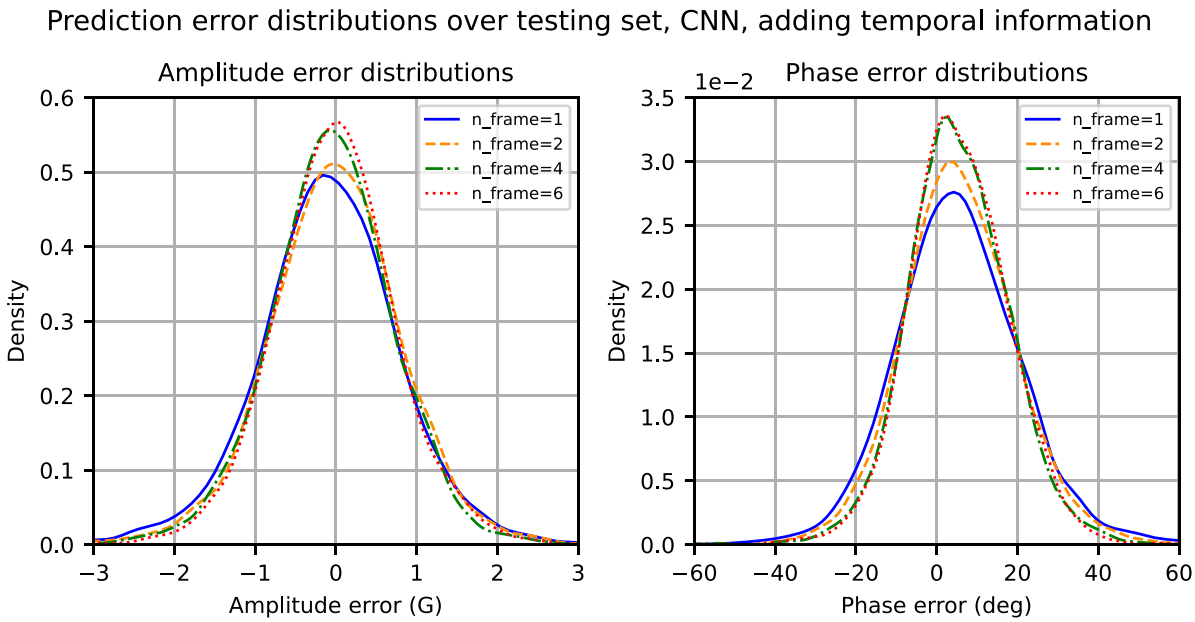

Figure 8. Distributions of the amplitude (left) and phase prediction (right) errors over the five testing set shots (ground truth amplitude >3 G), using the modified CNN models with temporal frame stack inputs (figure 7(A)). The four distributions show models using `n_frame` = 1 (blue, i.e. the original CNN model), 2 (orange), 4 (green), and 6 (red).

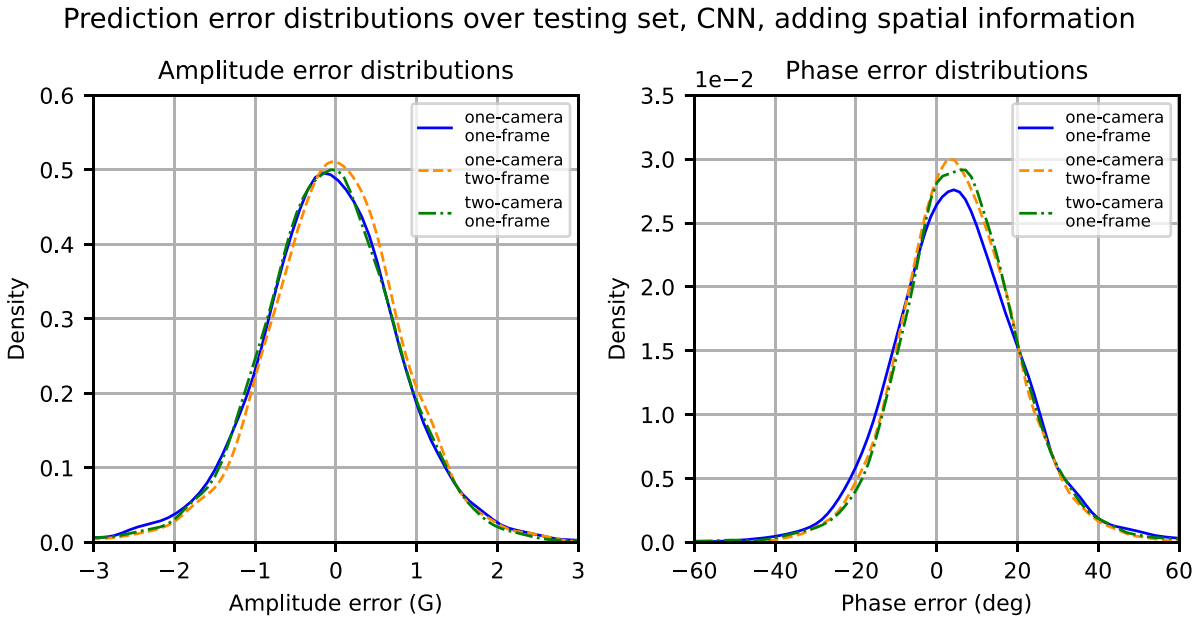

Figure 9. Distributions of the amplitude (left) and phase prediction errors (right) over the five testing set shots (amplitude ground truth >3 G), using the one-camera one-frame model (blue, i.e. the original CNN model), the one-camera two-frame model (orange), and the new two-camera one-frame model (green). The two-camera one-frame model uses the network shown in figure 7(B).

viewing cones differ by more than a direct toroidal rotation due to the use of different hardware (in particular the object lenses) as well as their positions on the machine (figure 1), which forbid us from combining them into a single frame stack as the plasma is located at a different position on each camera. Therefore we added a second stack of convolution layers with the same hyperparameters for processing data from Camera 2, and we concatenated the outputs into a single array before sending it to the downstream dense layers. Figure 9 compares the prediction results of this new model with the one-camera one-frame model and the one-camera two-frame model shown previously. We found that while adding the second camera does give marginal improvements especially in phase prediction, the effects are similar to adding one historic frame from the existing camera. This could be because the images from a single camera already contain sufficient mode information, thus making information from the second camera effectively redundant. Nevertheless, this may be unique to our specific choice of viewing angles, and different camera positioning should be tested before drawing final conclusions. For

example, determination of the mode phase might be improved by setting the second camera to instead have a zoomed-in view of a specific boundary feature whose local emission strongly depends on the mode phase, such as a limiting surface.

6. Discussion

In this section we briefly discuss aspects of the present work and future work that we are currently pursuing which are relevant to real-time control applications.

First, while the model presented in this study was designed with low-latency deployment as a primary consideration, additional work will be necessary for actual hardware deployment. Recent works from the high energy physics community have demonstrated a workflow [40–42] for deploying neural network models on field-programmable gate array (FPGA) hardware and achieved an estimated inference latency between 5 and 24 μs for convolutional networks [43, 44] similar to the one presented in this study. This workflow requires further optimizing the model through network compression and quantization [45–47] as well as potentially reducing the input dimensions in order to fit the network within the resource constraints of an FPGA device. Assuming we can achieve a similar latency while maintaining an acceptable level of accuracy, this implementation could satisfy the requirement for mode feedback control on HBT-EP and other devices.

In addition to optimizing the model for FPGA implementation, additional components are also needed for integrating the controller into a feedback control system. These includes but are not limited to: reading the data stream through frame grabber hardware, interfacing the frame grabber with the FPGA controller, and sending the control request to downstream actuators. A custom-built framework will be necessary, as currently there is not an established solution that satisfies the latency requirements for mode control on HBT-EP [13]. These considerations are currently being investigated and will be reported in future publications.

Beyond the camera-to-magnetics setup previously discussed for HBT-EP, the method presented in this study could also be applied to mode tracking using other combinations of either 1D or 2D optical diagnostics as the inputs, and other sources of mode information as the ground truth. For example, on HBT-EP a straightforward study would be to compare the prediction accuracy of a similar deep-learning-based algorithm with existing methods using identical input data from arrays of EUV sensors [14]. For other and future fusion devices, alternative data sources and model training schemes should also be explored as actual diagnostic measurements may be scarce. Some potential approaches could be through using synthetic diagnostic data and transfer learning, through using temporary diagnostics to gather batches of training data, or through using online learning as more data become available during reactor operation. In addition, alternative deep learning architectures, such as recurrent neural networks [48, 49] or transformers [50] for processing time series data, should also be explored. However, for a real-time control system a successful implementation will require balancing the model's accuracy with its inference latency and initiation intervals when running on the targeted computing hardware with consideration of the MHD timescale of the specific device.

7. Conclusion

We implemented an algorithm to track the amplitude and phase of rotating MHD modes in tokamak plasmas using high speed imaging cameras and deep learning. This algorithm uses a CNN architecture. It takes in camera images and predicts the amplitudes of $n = 1$ sine and cosine mode components. We developed this algorithm using a dataset from the HBT-EP tokamak, and we found it to perform better than the other conventional methods that were tested. Analysis of including additional temporal or spatial information to increase the size of the data stream showed modest levels of improvement in the models' predictions, but did indicate a clear difference between additional temporal and spatial information in terms of model performance.

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgments

This work was supported by U.S. Department of Energy, Office of Fusion Energy Sciences Grant Numbers DE-FG02-86ER53222 and DE-SC0021325.

ORCID iDs

Y Wei https://orcid.org/0000-0002-8868-0017
J P Levesque https://orcid.org/0000-0003-1193-475X
C Hansen https://orcid.org/0000-0001-6928-5815
M E Mauel https://orcid.org/0000-0003-2490-7478
G A Navratil https://orcid.org/0000-0002-2930-5196

References

[1] Pau A, Fanni A, Carcangiu S, Cannas B, Sias G, Murari A and Rimini F (The JET Contributors) 2019 *Nucl. Fusion* **59** 106017
[2] Sabbagh S, Berkery J, Park Y, Ahn J, Jiang Y, Riquezes J, Butt J and Bialek J 2020 Tokamak disruption event characterization and forecasting research and expansion to real-time application *2020 IAEA Fusion Energy Conf.*
[3] Wei Y, Levesque J P, Hansen C J, Mauel M E and Navratil G A 2021 *Nucl. Fusion* **61** 126063
[4] Boyer M D, Rea C and Clement M 2022 *Nucl. Fusion* **62** 026005
[5] Sen A K, Singh R, Sen A and Kaw P K 1997 *Phys. Plasmas* **4** 3217–21
[6] Kolemen E, Welander A S, La Haye R J, Eidietis N W, Humphreys D A, Lohr J, Noraky V, Penaflor B G, Prater R and Turco F 2014 *Nucl. Fusion* **54** 073020

[7] Kong M, Felici F, Sauter O, Galperti C, Vu T, Ham C J, Hender T C, Maraschek M and Reich M (The EUROfusion MST1 Team) 2022 *Plasma Phys. Control. Fusion* **64** 044008
[8] Hanson J M, De Bono B, Levesque J P, Mauel M E, Maurer D A, Navratil G A, Pedersen T S, Shiraki D and James R W 2009 *Phys. Plasmas* **16** 056112
[9] Chu M S and Okabayashi M 2010 *Plasma Phys. Control. Fusion* **52** 123001
[10] Rath N *et al* 2013 *Plasma Phys. Control. Fusion* **55** 084003
[11] Peng Q, Levesque J P, Stoafer C C, Bialek J, Byrne P, Hughes P E, Mauel M E, Navratil G A and Rhodes D J 2016 *Plasma Phys. Control. Fusion* **58** 045001
[12] Orsitto F *et al* 2016 *Nucl. Fusion* **56** 026009
[13] Rath N, Kato S, Levesque J P, Mauel M E, Navratil G A and Peng Q 2014 *Rev. Sci. Instrum.* **85** 045114
[14] Levesque J, Brooks J, DeSanto S, Mauel M and Navratil G 2021 Active control of kink modes using a non-magnetic, extreme ultraviolet sensor array *47th EPS Conf. Plasma Physics* p P4.1053 (available at: https://ocs.ciemat.es/EPS2021PAP/pdf/P4.1053.pdf)
[15] Chandra R, Levesque J, Wei Y, Li B, Saperstein A, Stewart I, Hansen C, Mauel M and Navratil G 2023 *Fusion Eng. Des.* **191** 113565
[16] Levesque J P, Brooks J W, Abler M C, Bialek J, Byrne P J, Hansen C J, Hughes P E, Mauel M E, Navratil G A and Rhodes D J 2017 *Nucl. Fusion* **57** 086035
[17] Angelini S M, Levesque J P, Mauel M E and Navratil G A 2015 *Plasma Phys. Control. Fusion* **57** 045008
[18] LeCun Y, Bengio Y and Hinton G 2015 *Nature* **521** 436–44
[19] Rea C, Montes K J, Erickson K G, Granetz R S and Tinguely R A 2019 *Nucl. Fusion* **59** 096016
[20] Montes K J *et al* 2019 *Nucl. Fusion* **59** 096015
[21] Kates-Harbeck J, Svyatkovskiy A and Tang W 2019 *Nature* **568** 526–31
[22] Zhu J *et al* 2021 *Nucl. Fusion* **61** 114005
[23] Dong G, Felker K G, Svyatkovskiy A, Tang W and Kates-Harbeck J 2021 *J. Mach. Learn. Model. Comput.* **2** 49–64
[24] Matos F, Svensson J, Pavone A, Odstrčil T and Jenko F 2020 *Rev. Sci. Instrum.* **91** 103501
[25] Ferreira D R, Carvalho P J, Sozzi C and Lomas P J (JET Contributors) 2020 *Fusion Sci. Technol.* **76** 901–11
[26] Abbate J, Conlin R and Kolemen E 2021 *Nucl. Fusion* **61** 046027
[27] Wan C, Yu Z, Wang F, Liu X and Li J 2021 *Nucl. Fusion* **61** 066015
[28] Boyer M D and Chadwick J 2021 *Nucl. Fusion* **61** 046024
[29] Wei X, Dong G, Sun S, Tang W, Lin Z and Du H 2022 arXiv:2208.08730
[30] Kaptanoglu A A, Morgan K D, Hansen C J and Brunton S L 2020 *Phys. Plasmas* **27** 032108
[31] Miller M A, Churchill R M, Dener A, Chang C S, Munson T and Hager R 2021 *J. Plasma Phys.* **87** 905870211
[32] Alves E P and Fiuza F 2022 *Phys. Rev. Res.* **4** 033192
[33] Seo J, Na Y S, Kim B, Lee C Y, Park M S, Park S J and Lee Y H 2021 *Nucl. Fusion* **61** 106010
[34] Seo J, Na Y S, Kim B, Lee C Y, Park M S, Park S J and Lee Y H 2022 *Nucl. Fusion* **62** 086049
[35] Degrave J *et al* 2022 *Nature* **602** 414–9
[36] Jalalvand A, Abbate J, Conlin R, Verdoolaege G and Kolemen E 2022 *IEEE Trans. Neural Netw. Learn. Syst.* **33** 2630–41
[37] Han W, Pietersen R A, Villamor-Lora R, Beveridge M, Offeddu N, Golfinopoulos T, Theiler C, Terry J L, Marmar E S and Drori I 2022 *Sci. Rep.* **12** 18142
[38] Jacobus C, Choi M J and Kube R 2022 arXiv:2201.07941
[39] Abadi M *et al* 2015 TensorFlow: large-scale machine learning on heterogeneous systems (available at: www.tensorflow.org/)
[40] Duarte J *et al* 2018 *J. Instrum.* **13** 07027
[41] Ngadiuba J *et al* 2020 *Mach. Learn.: Sci. Technol.* **2** 015001
[42] Fahim F *et al* 2021 arXiv:2103.05579
[43] Aarrestad T *et al* 2021 *Mach. Learn.: Sci. Technol.* **2** 045015
[44] Jwa Y J, Di Guglielmo G, Arnold L, Carloni L and Karagiorgi G 2022 *Front. Artif. Intell.* **5** 855184
[45] Courbariaux M, Bengio Y and David J P 2015 arXiv:1511.00363
[46] Jacob B, Kligys S, Chen B, Zhu M, Tang M, Howard A, Adam H and Kalenichenko D 2017 arXiv:1712.05877
[47] Coelho Jr C N, Kuusela A, Li S, Zhuang H, Ngadiuba J, Aarrestad T K, Loncar V, Pierini M, Pol A A and Summers S 2021 *Nat. Mach. Intell.* **3** 675–86
[48] Hochreiter S and Schmidhuber J 1997 *Neural Comput.* **9** 1735–80
[49] Sherstinsky A 2020 *Physica* D **404** 132306
[50] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A N, Kaiser L and Polosukhin I 2017 Attention is all you need (arXiv:1706.03762)