

# Exploring data-driven models for spatiotemporally local classification of Alfvén eigenmodes

Alan A. Kaptanoglu<sup>1,2</sup>, Azarakhsh Jalalvand<sup>3,4,\*</sup>, Alvin V. Garcia<sup>5</sup>,  
Max E. Austin<sup>6</sup>, Geert Verdoolaege<sup>7</sup>, Jeff Schneider<sup>8</sup>,  
Christopher J. Hansen<sup>9</sup>, Steven L. Brunton<sup>2</sup>, William W. Heidbrink<sup>5</sup>  
and Egemen Kolemen<sup>4,10</sup>

<sup>1</sup> Institute for Research in Electronics and Applied Physics, University of Maryland, College Park, MD 20742, United States of America

<sup>2</sup> Department of Mechanical Engineering, University of Washington, Seattle, WA 98195, United States of America

<sup>3</sup> Dept. of Electronics and Information Systems, Ghent University, Ghent B-9052, Belgium

<sup>4</sup> Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ 08544, United States of America

<sup>5</sup> Department of Physics, University of California, Irvine, CA 92697, United States of America

<sup>6</sup> College of Natural Sciences, The University of Texas at Austin, Austin, TX 78712, United States of America

<sup>7</sup> Department of Applied Physics, Ghent University, Ghent B-9000, Belgium

<sup>8</sup> School of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213, United States of America

<sup>9</sup> Department of Aeronautics and Astronautics, University of Washington, Seattle, WA 98195, United States of America

<sup>10</sup> Princeton Plasma Physics Laboratory, Princeton, NJ 08544, United States of America

E-mail: [azarakhsh.jalalvand@ugent.be](mailto:azarakhsh.jalalvand@ugent.be)

Received 23 April 2022, revised 12 July 2022

Accepted for publication 16 August 2022

Published 31 August 2022



CrossMark

## Abstract

Alfvén eigenmodes (AEs) are an important and complex class of plasma dynamics commonly observed in tokamaks and other plasma devices. In this work, we manually labeled a small database of 26 discharges from the DIII-D tokamak in order to train simple neural-network-based models for classifying AEs. The models provide spatiotemporally local identification of four types of AEs by using an array of 40 electron cyclotron emission (ECE) signals as inputs. Despite the minimal dataset, this strategy performs well at spatiotemporally localized classification of AEs, indicating future opportunities for more sophisticated models and incorporation into real-time control strategies. The trained model is then used to generate spatiotemporally-resolved labels for each of the 40 ECE measurements on a much larger database of 1112 DIII-D discharges. This large set of precision labels can be used in future studies for advanced deep predictors and new physical insights.

Keywords: machine learning, Alfvén eigenmodes, energetic particles, tokamak

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

\* Author to whom any correspondence should be addressed.

## 1. Introduction

Future steady-state operation of nuclear fusion devices necessitates active real-time control of a number of complex plasma processes, including edge-localized modes [1–3], disruptions [4, 5], and Alfvén eigenmodes (AEs) [6]. Some of these dynamics are highly nonlinear, and can occur on fast time scales and small spatial scales where many analytic assumptions break down. The reality of fast time scales requires that many real-time plasma control schemes are limited to simple models such as those based on 1D transport [7], linearization or local-expansions [8–16], heuristics (based on prior experimental knowledge) [17, 18], the biorthogonal decomposition [19–24], and so forth. Despite the challenges posed by multi-scale dynamics and nonlinearity, many of these models have been successfully employed for real-time control in operational scenarios. However, future progress in controlled fusion hinges on models that have the potential to overcome the many barriers associated with predicting the dynamics of complex plasmas.

Fortunately, there has already been remarkable success in data-driven models for disruption identification and real-time control in tokamaks [4, 5, 25–29], including high-performance models that are not limited to a specific device [30]. Neural network models can be trained offline and then used to predict plasma activity online. There has also been recent deep learning work for automated identification of AEs in diagnostic data [31–33]. Our prior work [34], briefly reviewed in section 1.3, substantially advanced these AE studies by utilizing a large database of high-resolution spatiotemporal data for AE prediction. In addition to real-time control applications, machine learning studies open the possibility for new physical understanding of this unique class of plasma modes in terrestrial plasma experiments or even space-based probes [35].

Heuristic models for AEs have also seen some successes. For instance, Hu *et al* [18] controlled AEs by computing a measure of the strength of AEs ‘by the mean value of the coherent power spectral density from Fourier analysis of temperature measurements from ten pairs of ECE channels with a bandpass between the geodesic acoustic and TAE frequencies... averaged over the period of  $t = 400\text{--}700\text{ ms}$ ’. Despite this coarse-grained metric and lack of distinction between AE modes, it worked reasonably well for a linear feedback controller with neutral beam injection used as an actuator.

The purpose of this work is to use simple neural-network-based models to provide more sophisticated spatio-temporal predictions of AE activity. In this work, we meticulously label a database of 26 DIII-D tokamak discharges in order to train simple neural-network-based models for full spatio-temporal AE classification, improving on previous work that could not provide localized predictions in space and time. These models are then used to generate new labels for the much larger database [36] of 1112 DIII-D discharges.

In the remainder of this section, we briefly describe AEs as well as the electron cyclotron emission (ECE) diagnostics that we use to develop the data-driven models for locating and identifying AE modes. These models are introduced in section 2 and their performances are discussed in section 3.

### 1.1. Alfvén eigenmodes

AEs make up a diverse class of complex dynamics that pose many challenges for simple machine learning models, generalization to new datasets, and analytic methods. They are typically observed in tokamaks and other fusion-relevant plasma devices, although they can also occur in space plasmas [37]. Some AEs are extremely well-studied, while others are recently discovered and yet to be fully understood. Although their causes and effects vary widely, some AEs can lead to confinement loss and damage to plasma-facing components in present and future devices. Previous examples include ablating carbon and ruining the transmission of optical components [38] and lost fast ions drilling a hole in TFTR that actually vented the machine [39]. The database used in this work (described in section 2) distinguishes between several types of AE activity: low-frequency modes (LFMs have been formerly characterized as beta-induced Alfvén-acoustic modes [36], denoted BAAE), beta-induced Alfvén eigenmodes (BAE), reversed-shear Alfvén eigenmodes (RSAE), and toroidal Alfvén eigenmodes (TAE) [36, 40].

Distinguishing AE types can be done by utilizing both the analytic theory and qualitative trends in a database of discharges. Rough estimates of the typical frequency ranges in a DIII-D discharge are  $\lesssim 50\text{ kHz}$  for LFMs,  $\sim 30\text{--}150\text{ kHz}$  for BAEs,  $\sim 100\text{--}200\text{ kHz}$  for RSAEs, and  $\sim 90\text{--}200\text{ kHz}$  for TAEs. In addition to the somewhat separated frequency ranges for some of these modes, they each exhibit a different qualitative temporal character, which can be visualized in a spectrogram. RSAEs exhibit frequency-chirping, LFMs exhibit ‘Christmas-light’ patterns, TAEs tend to be longer and relatively flat, and BAEs tend to be splotchy and intermittent. Several plasma diagnostics can be jointly used for distinguishing the AEs. Some of the AEs exhibit relatively simple frequency dependence on the plasma parameters, and their appearance on a spectrogram can be predicted reasonably well [41]. TAE and RSAE are often observed concurrently (presumably with some mode-mode coupling [42]), even in discharges where each ECE channel only observes one or the other. Finally, additional analytic and heuristic criteria can be used; for instance, RSAEs are often localized near the minima in the magnetic safety factor profile. The AE modes are further described in table 1, where references to the relevant theoretical and experimental manuscripts can also be found.

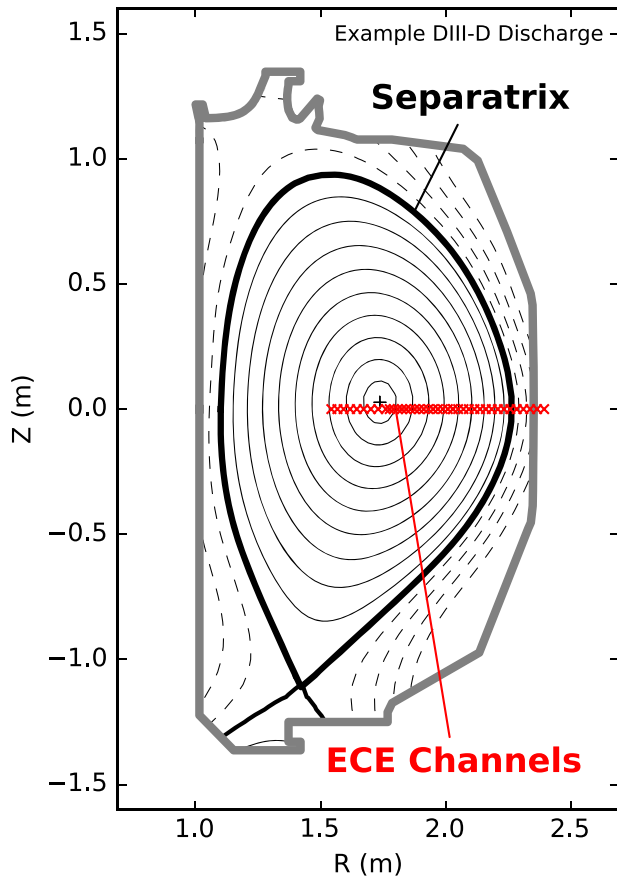
Although AEs are relatively rare events, models for AE classification can be trained because AEs appear on many high-resolution plasma diagnostics and significant troves of past data are often available on tokamaks such as DIII-D. For future real-time control applications, there are a wide range of experimental actuators that can be used to control different AEs [6, 43–45].

### 1.2. Electron cyclotron emission

ECE provides direct local measurements of the electron temperature for thermal plasmas [56], and as such, can provide spatiotemporally-localized information about AE activity. The

**Table 1.** Description of the AE activity considered in this work, adapted from Heidbrink [55]. The poloidal wave number is denoted by  $m$  and the minimum value of the safety factor is denoted  $q_{\min}$ . Other AEs are omitted because they were not common enough for adequate training using the database described in section 2.

AE name	Acronym	Cause
Beta-induced	BAE [46–48]	Compressibility
Low-frequency modes	LFM [36]	Hot electrons, $q_{\min} \sim$ rational
Reversed-shear	RSAE [49, 50]	$q_{\min}$
Toroidal	TAE [51–54]	$m$ & $m + 1$ coupling



**Figure 1.** Illustration of the 40 radial ECE measurement locations alongside the closed (solid) and open (dashed) flux surfaces for an example DIII-D discharge. The effective position of the ECE channels can vary significantly in each discharge, and measurements outside the last closed flux surface are not local or accurate. Reproduced courtesy of IAEA. Figure from [34]. Copyright (2021) IAEA.

DIII-D ECE diagnostic data is obtained at 500 kHz, in a single toroidal cross-section, at 40 different channels corresponding to varying radial locations, as shown in figure 1. Each ECE channel spans an approximately 1–2 cm radial extent, which is small compared to the radial scale of AEs. Most of the time, an AE mode can be seen across several channels.

Properly capturing the spatial correlations is difficult because the effective position of the ECE channels changes

with the magnetic equilibrium, and therefore can vary substantially during startup operation. The first few ECE channels regularly view data that is outside the last closed flux surface and this data is considered untrustworthy. Despite this spatial variability and data corruption, our prior work [34] used the full, raw, unprocessed ECE data, i.e. only the ECE channel indices corresponding to the relative major radial positions of the measurements. This has the advantage that the magnetic field evolution is not required. In the present work, we use the time-resolved magnetic field equilibrium data from EFIT [57] to map the ECE channels to the normalized plasma radius  $\rho$ , the square root of the toroidal flux normalized by the value at the last-closed flux surface. This mapping is motivated in order to make spatially-local predictions, and a normalized flux function like  $\rho$  is a natural choice. For instance, consider that ECE channels that observe different radial locations may in fact observe the same flux surface and the same AE.

### 1.3. Summary of results from prior work

Despite the simplifications that we described in the previous subsection, high classification performance was obtained in our initial work [34]. Table 2 illustrates the primary conclusions of the prior work using a reservoir computing network (RCN) [34]. RCNs are recurrent neural networks, and therefore are capable of processing temporal information while providing a simple weights initialization and training procedure [58, 59]. This work used a very large but imprecise database that contains labels for a single time slice per AE activity to indicate the approximate occurrence of AEs in the time domain. These labels are specific to the overall discharge but not specific to each ECE channel, so spatially-localized prediction of the modes was not possible. Moreover, since only a single time slice label indicates roughly where each AE occurs, significant interpolation and windowing is required to train a proper machine learning method. Otherwise, the learning architecture is attempting to learn from a dataset with very few true positives and a very large number of false negatives. Therefore, only somewhat temporally-localized predictions were possible—predictions of AE activity by the RCN were considered true positives if they were within 500 ms of the corresponding ground-truth time slice.

**Table 2.** Summary of the highest performance model from Jalalvand *et al* [34], a two-layer RCN with 8K–500 nodes per layer, on a validation set. A threshold of 0.2 has been applied to binarize the model output. There are 540 labeled AEs in this validation dataset. Definitions for the acronyms and the TPR and FPR metrics are listed below the table<sup>a</sup>.

AE	TP	FP	TN	FN	P	N	TPR	FPR
BAE	75	46 982	470 368	17	92	517 350	0.82	0.09
LFM	8	4102	587 088	3	11	591 190	0.73	0.01
RSAE	167	48 319	417 211	13	180	465 530	0.93	0.10
TAE	248	76 133	330 057	9	257	406 190	0.97	0.19

<sup>a</sup>TP  $\equiv$  True positive, FP  $\equiv$  False positive,  
 TN  $\equiv$  True negative, FN  $\equiv$  False negative,  
 P = Total # of labels with AEs (TP + FN),  
 N = Total # of labels without AEs (FP + TN),  
 TPR  $\equiv$  TP/P, FPR  $\equiv$  FP/N.

Furthermore, table 2 illustrates the highly imbalanced nature of the data. Even after interpolation and windowing of the original labels, only about 0.02% of the training or validation sets are labeled as an AE. This fact is why true-positive-ratio (TPR) and false-positive-ratio (FPR) are the primary metrics reported in that work. Reporting the *accuracy* of the model, by calculating the ratio of correctly predicted labels divided by the total number of samples in the dataset  $((TP + TN)/(P + N))$ , would be profoundly misleading; a model that *never* raises flag for any AE activity (TP = FP = 0, TN = N, FN = P) would still report accuracy above 99%. Despite the many shortcomings of the data, the strong performance with RCNs was encouraging for continued machine learning work.

In contrast to this prior work, the present work uses a much smaller but precisely-labeled database to train models that can make spatiotemporally-local predictions of AE activity. This ability facilitates the prediction of different AEs on specific flux surfaces in the plasma, and the trained model can be used to produce precise ECE labels in much larger databases.

## 2. Spatiotemporally-localized AE classification

The previous work identified that the primary avenue for further improvements was to provide improved database labels that are unique to each ECE channel, facilitating spatiotemporally-localized machine learning predictors. For instance, it was found that other machine learning classifiers (MLPs, GRUs, RNNs, CNNs, etc) that were trained with the original, large database struggle to find performance  $\gtrsim 70\%$  TPR, and very quickly overfit to the data. This poor training performance persists despite the different learning architectures, and despite large variations in the network hyperparameters. This behavior is characteristic of database labels that are of insufficient quality for training accurate machine learning models. This is an expected finding—the database labels are not unique to each ECE channel so the model is inevitably being trained on many false positives and many false negatives, providing an upper bound on the performance achievable with any machine learning model. This

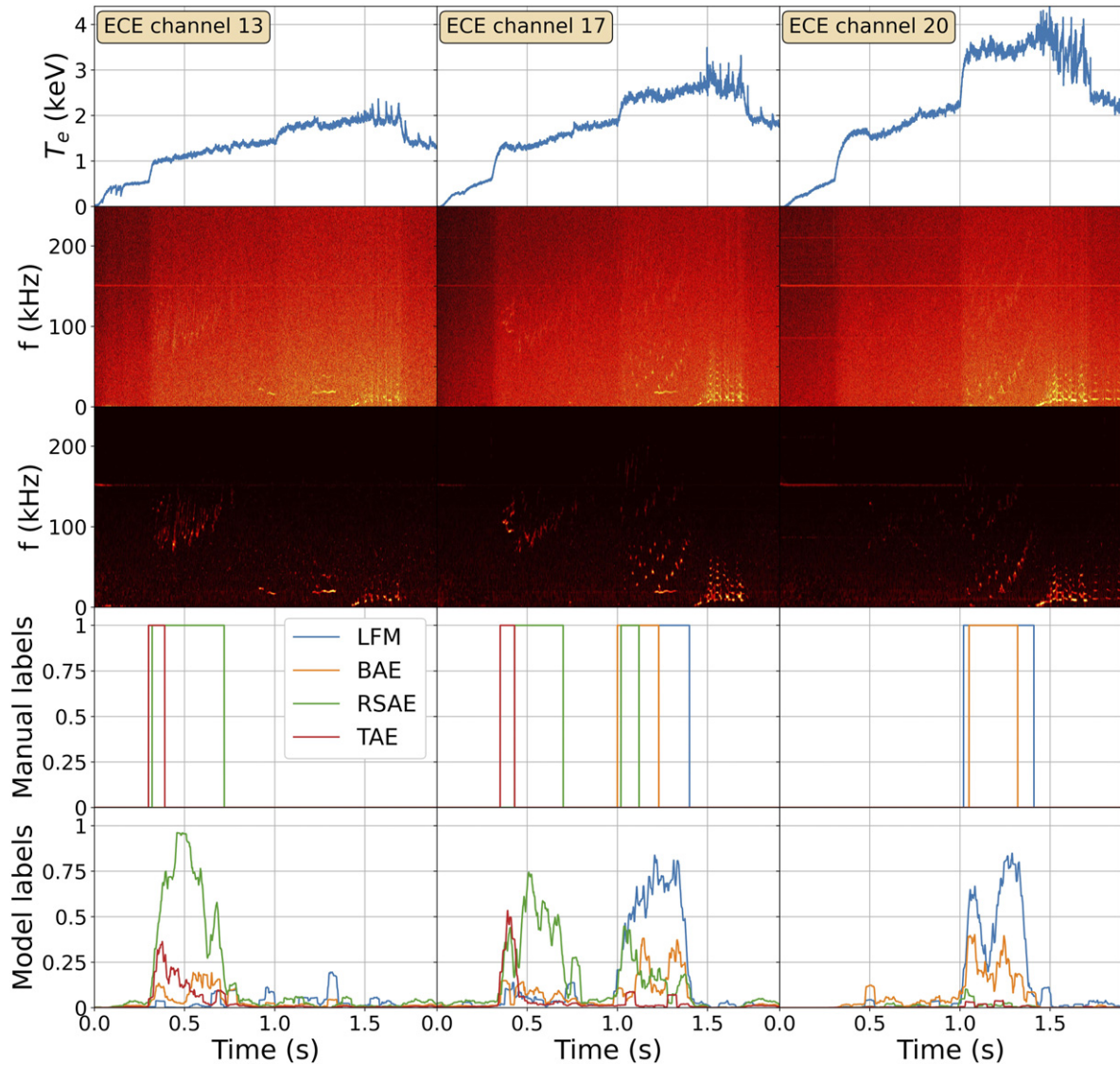
finding further motivated a second, much smaller database consisting of higher-precision, spatiotemporally-localized AE labels.

### 2.1. High precision database labels

ECE-channel-specific labels were produced by selecting 26 very active DIII-D discharges, manually examining all  $26 \times 40 = 1040$  ECE channel spectrograms, and providing specific time and frequency windows for the labels (a box is drawn on the spectrogram where the mode appears). This labeling process relied on the visual identification of the distinctive features of each AE (discussed in section 1.1) in the spectrograms. Analytic profiles for the TAE and minimum RSAE frequencies [41] were computed throughout the time evolution and superimposed on top of the spectrograms to aid in the labeling process. The determined labels were also cross-checked for reasonable consistency with the discharge-wide labels from the larger database used in prior work. These original labels were in turn often cross-checked with a few different experimental diagnostics, including ECE, CO2 interferometry, and magnetics, especially when concurrent AEs were present.

This manual examination becomes significantly easier by first using standard image processing techniques [60] to enhance the spectrograms by highlighting the AE activity [61]. Each spectrogram is shape  $3709 \times 256$  in time  $\times$  frequency space, with bin sizes of approximately  $50 \mu\text{s} \times 1 \text{ kHz}$ . The denoising process is as follows: the spectrogram values are (1) converted to the log of the values, (2) normalized to  $[0, 1]$ , (3) thresholded off if below the 90% quantile, (4) Gaussian-filtered with temporal standard deviations (31, 3), (5) mean-filtered along the frequency dimension, (6) morphologically-filtered, and finally (7) mean-filtered along the frequency dimension again. Across the vast majority of ECE spectrograms obtained from a wide range of DIII-D discharges, this image filtering sequence was empirically found to perform well at highlighting the AE activity while denoising the remainder of the image, as in figure 2.

Although this is a very small subset of the original database, each spectrogram provides 3709 time slice input vectors, each of dimension 256, meaning a total of  $\sim 4$  million input vectors



**Figure 2.** From top to bottom this illustration shows three ECE channel time series for the DIII-D test discharge 178637, along with the corresponding spectrograms, the denoised spectrograms, the manual labels corresponding to each ECE channel, and the predictions using the baseline model described in section 3.1. These predictions were facilitated by the creation of a small dataset of 26 discharges containing high-precision labels unique for each ECE channel. The model closely predicts the manual labels, albeit with weaker strength.

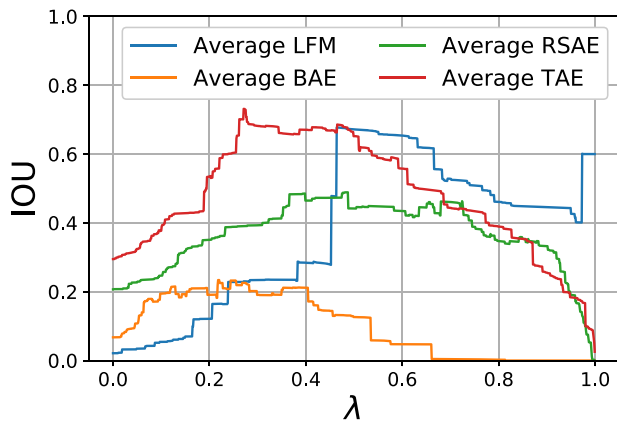
for training models. The raw signals, spectrograms, denoised spectrograms (denoised as described in the previous paragraph), labels, and predictions are illustrated for three active ECE channels during DIII-D discharge 178637 in figure 2. The model was not trained on this discharge. The denoising process strongly highlights the AE activity and improved both the manual and model-based labeling; models trained with the denoised spectrograms universally outperformed models with the raw spectrograms.

Out of the 1040 ECE channel spectrograms,  $\approx 2\%$  of the total time series contain an LFM,  $\approx 3\%$  contain BAE,  $\approx 8\%$  contain RSAE, and  $\approx 7\%$  contain TAE. However, often the AEs are concurrent, so this is an underestimate of the inactive periods. The total number of ‘events’ of each type are 143 LFM, 329 BAE, 415 RSAE, and 441 TAE (this can be reconciled with the previous percentages because BAE tends

to appear on the ECE channels as short, isolated events). Overall, approximately 95% of the samples in this sub-database has no AE activity whatsoever. Although still quite imbalanced, this database is a substantial improvement over the original one where approximately 99% of the data contained no AE activity (or at least no AE labels).

### 3. Models and results

Evaluating model performance with this database is not straightforward. Often, the TPR and FPR are used for imbalanced datasets. However, we are primarily interested if the model can provide accurate, spatiotemporally localized predictions of the AEs. In other words, metrics should be evaluated not on individual spectrograms, but on the full, 40 channel ECE data. Each discharge’s ECE data forms an ‘image’ in the space of normalized plasma radius and time ( $\rho, t$ ).



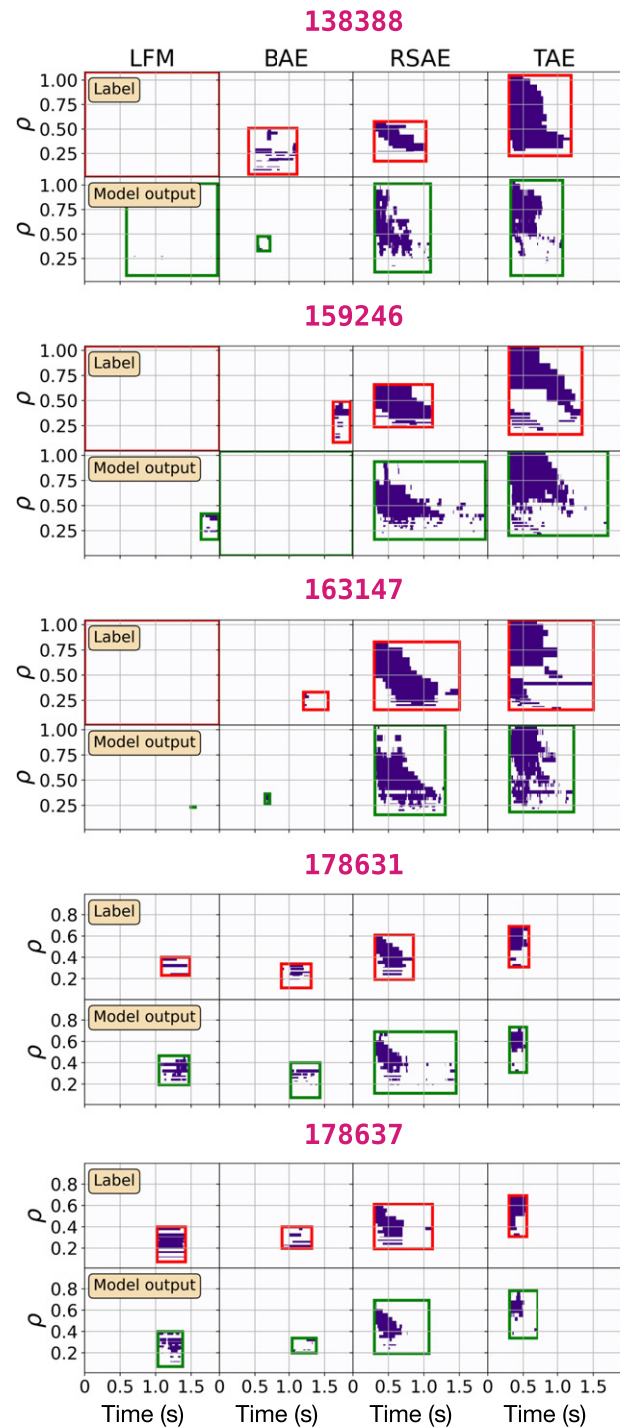
**Figure 3.** IOU values for each of the AE types as a function of the threshold  $\lambda$ . The IOU values were computed from the five test discharges.

One of the primary metrics in the field of image segmentation is the intersection-over-union (IOU) [62–64]. For each AE type, this metric places a bounding box over the entire active portion of the ground truth image, and another bounding box over all of the active image pixels that are predicted by the model to be above some threshold value  $\lambda$ . Then the IOU can be computed from these two bounding boxes, providing a similarity metric between the predictions and manual labels. IOU above 0.5 is typically considered strong performance in image identification.

### 3.1. Multi-layer perceptron models

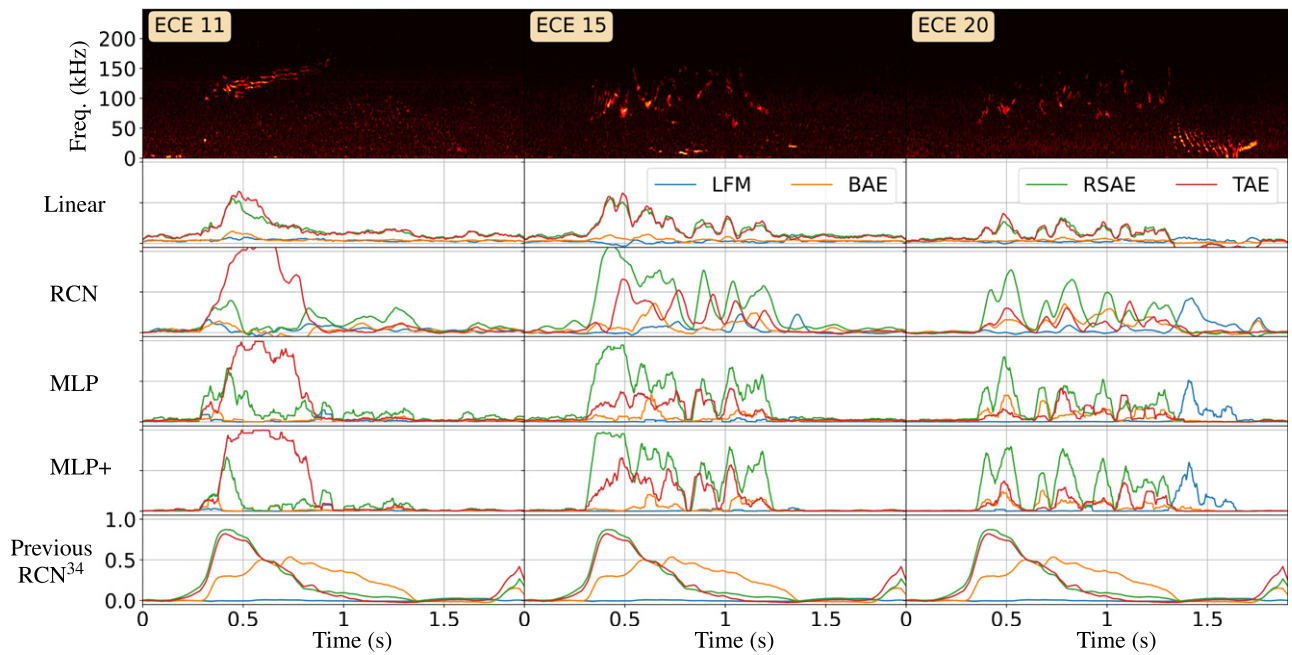
As a baseline, a multi-layer perceptron (MLP) with three layers, 512 nodes per layer, and interleaving 30% dropout layers was trained with the mean-squared error loss, directly on the denoised spectrogram data from this new database. Each input is a single column from the denoised spectrogram, i.e. a feature vector of length equal to the number of frequency points in the spectrogram. Therefore, the inputs do not contain any temporal information beyond that which is encoded in the frequency information in the spectrogram. Such a memory-less approach tends to produce rather ‘spiky’ predictions, so we smooth the outputs using a running average with a window of  $\sim 50$  ms. In order to compare some of model predictions with the manual labels, the 26 discharges are split into 16 training discharges, 5 validation discharges, and 5 testing discharges in a way that provides enough LFM and BAE instances for training. Model performance was fairly insensitive to changes in the dropout rate, additional nodes per layer, and other network hyperparameters.

The performance of the model is evaluated by computing the IOU for each of the validation discharges and for each of the AE types as a function of the prediction threshold  $\lambda$ . The full IOU validation results are shown in figure 3 and indicate that at  $\lambda \approx 0.4$ , LFM, RSAE, and TAE IOUs are  $\approx 40\%$ – $80\%$  or higher, while the BAE IOU peaks at  $\approx 20\%$ . It is possible that the lower performance for BAE is because this mode



**Figure 4.** Summary of the IOU results for the 5 test discharges, at a threshold of  $\lambda = 0.4$ . Apart from some of the BAE, it is visually apparent that the model can accurately predict the spatiotemporal dependence of the modes; the primary degradation in the LFM, RSAE, and TAE IOU tend to be isolated outlier predictions, which could be removed by additional output smoothing. It is also worth noting that the manual labels are not perfect. It is possible that in some cases, the model is actually more accurate than the provided labels. For instance, on additional inspection discharge 159246 looks like it does exhibit some weak LFM activity at the location predicted by the model, and had been labeled as such in the original database containing 1112 discharges (but not in the database of 26 discharges).

tends to be more intermittent than other AEs. We plot each validation discharge image and the corresponding bounding



**Figure 5.** Illustration of the predictions for the linear, RCN, MLP, and MLP + derivatives models on the active ECE channels in discharge 170669, along with the RCN outputs from our prior work [34]. The linear model struggles to accurately predict AEs, the new RCN and the MLPs predict AEs quite well, and the older RCN accurately predicts the ‘average’ AE content of the discharge but cannot distinguish between ECE channels. Recall that the RCN from previous work was trained on a different set of labels that is not unique to each ECE channel; despite this setback, it predicts well the timing and type of primary AE activity that occurs over the majority of ECE channels.

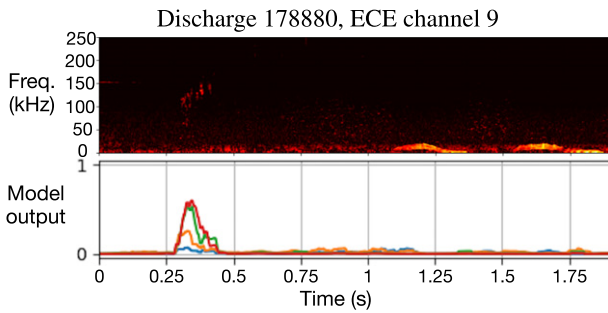
box in figure 4 for a threshold of  $\lambda = 0.4$ . The visualization illustrates that the model accurately captures the spatiotemporal location of AE activity in all the testing discharges, with most of the degradation in IOU coming from small outliers. The model IOU can frequently be further improved by using a more aggressive output smoothing than the 50 ms running average employed here.

The baseline MLP did not use spatial or temporal correlations in the data to train, and yet manages to accurately capture much of the spatiotemporal dependence of AEs in the data. The frequency content at a single time slice of the individual denoised spectrograms is apparently already very informative for AE classification. Temporal memory was also found to be fairly unimportant in our prior work. The lower performance with BAE and LFM may be a symptom of the simplicity of this MLP. The MLP is only three layers, and the input data contains no temporal memory for the MLP to extract. There is no spatial memory either, because each ECE channel is processed individually. This is not very relevant for BAE and LFM, which tend to have smaller spatial extents than TAE or RSAE. LFM in particular exhibit a complex pattern in time, meaning one might expect that identifying these modes would be more sensitive to temporal memory. A simple way to incorporate some temporal information in these models is to enrich the input by including first and second temporal derivatives, calculated from an arbitrary window of size  $\Delta t$  around each input at time  $t$ . This triples the size of the input features, but we observe some minor improvements in performance. The addition of temporal information primarily increases the confidence of the model decisions, particularly for detecting LFM.

### 3.2. Other models

The MLP models were then compared with two other relatively simple machine-learning models, one based on ridge regression and another based on a two-layer RCN. The MLP outperforms ridge regression, which struggles to identify BAE or LFM, cannot distinguish between TAE and RSAE, and struggles to fully turn off when AE activity is not present. The RCN performs similarly to the MLP model. All four of the linear, RCN, MLP, and MLP + derivatives models were tested against a set of 20 randomly chosen discharges to visually compare model performances. A representative set of spectrograms from discharge 170669 are illustrated in figure 5, along with the outputs from the original RCN used in our prior work. The RCN and MLP models accurately predict AE activity quite well. The older RCN predicts the overall AE activity in the discharge well, but cannot distinguish activity in each ECE channel since it was trained on the original database labels that are not specific to ECE channel. Nonetheless, the older RCN performs well at predicting the timing and type of primary AE activity that occurs across the ECE channels.

Encouraged with the performance of this baseline MLP, we now use the model to predict labels for the large database consisting of another 1112 DIII-D discharges. This larger database varies across a very wide range of parameter space despite the database having been compiled to focus on discharges with energetic particles. This sort of bootstrapping is critical for creating more useful datasets in the fusion community; recent work [65] provided similar labels for LH transitions in a DIII-D database. In this scenario, we denote a ‘true AE’ in a specific ECE channel if the mode is predicted to occur with



**Figure 6.** The model correctly identifies minor AE activity throughout the database, which is not usually labeled manually.

probability  $>0.2$  for more than 50 ms of the data. With this definition, the percent of spectrograms in the database with at least one occurrence of LFM, BAE, RSAE, or TAE is 3.9%, 1.0%, 22%, and 24% respectively. However, the percentage of time in the whole database where each AE is predicted to occur with probability  $>0.2$  is only 0.6%, 0.2%, 4.8%, 5.2%. In total, 8% of the data has at least one of type AE occurring. This estimate contains significantly more AE activity than was originally labeled in this database and it alone represents an interesting finding. It provides an estimate for the frequency of occurrence in AE modes across a broad parameter space of DIII-D operation. A representative example is illustrated in figure 6 where weak AE activity was originally ignored but the model correctly predicts AEs. Future work could cluster the AE activity by various plasma parameters in order to assess AE occurrence frequencies and other AE properties in different dynamical regimes.

#### 4. Conclusion

A small set of 26 DIII-D discharges were manually labeled for AE activity and used to train machine learning models to generate spatiotemporally localized classification of AEs. Initial work with this small database shows that simple machine-learned models can already produce strong spatiotemporally-localized AE classification from denoised spectrograms of individual DIII-D ECE channel data. The baseline MLP model performs favorably compared with a ridge regression model, a RCN model, and a memory-aware MLP. The baseline MLP was subsequently used to generate spatiotemporally localized labels for a much larger database of 1112 DIII-D discharges. These labels facilitate future work using more sophisticated deep learning models, as well as physical studies on the characteristics of AEs in different dynamical regimes. Additional future work includes prediction and control algorithms for AEs and combining datasets from different diagnostics such as ECE, soft x-ray (SXR) imaging, magnetics, and beam-emission spectroscopy.

#### Acknowledgments


This material is based upon work supported by the US Department of Energy, Office of Science, Office of Fusion

Energy Sciences, using the DIII-D National Fusion Facility, a DOE Office of Science user facility, under Awards DE-FC02-04ER54698, DE-SC0021275, DE-SC0020337, National Science Foundation Award Numbers 1633631 and 2108384, and Ghent University Special Research Award No. BOF19/PDO/134.

#### Disclaimer

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

#### ORCID iDs

Alan A. Kaptanoglu  <https://orcid.org/0000-0002-6337-2907>  
 Azarakhsh Jalalvand  <https://orcid.org/0000-0001-8739-1793>  
 Geert Verdoolaege  <https://orcid.org/0000-0002-2640-4527>  
 Christopher J. Hansen  <https://orcid.org/0000-0001-6928-5815>  
 William W. Heidbrink  <https://orcid.org/0000-0002-6942-8043>

#### References

- [1] Lang P.T. *et al* 2004 Frequency control of type-I ELMs by magnetic triggering in ASDEX upgrade *Plasma Phys. Control. Fusion* **46** L31
- [2] Maingi R. 2014 Enhanced confinement scenarios without large edge localized modes in tokamaks: control, performance, and extrapolability issues for ITER *Nucl. Fusion* **54** 114016
- [3] Ham C., Kirk A., Pamela S. and Wilson H. 2020 Filamentary plasma eruptions and their control on the route to fusion energy *Nat. Rev. Phys.* **2** 159–67
- [4] Rea C. and Granetz R.S. 2018 Exploratory machine learning studies for disruption prediction using large databases on DIII-D *Fusion Sci. Technol.* **74** 89–100
- [5] Fu Y. *et al* 2020 Machine learning control for disruption and tearing mode avoidance *Phys. Plasmas* **27** 022501
- [6] Garcia-Munoz M. *et al* 2019 Active control of Alfvén eigenmodes in magnetically confined toroidal plasmas *Plasma Phys. Control. Fusion* **61** 054007



- [7] Maljaars E. *et al* 2017 Profile control simulations and experiments on TCV: a controller test environment and results using a model-based predictive controller *Nucl. Fusion* **57** 126063
- [8] Kolemen E. *et al* 2015 Heat flux management via advanced magnetic divertor configurations and divertor detachment *J. Nucl. Mater.* **463** 1186–90
- [9] Albanese R., Ambrosino R., Castaldo A., De Tommasi G., Luo Z.P., Mele A., Pironti A., Xiao B.J. and Yuan Q.P. 2017 ITER-like vertical stabilization system for the EAST tokamak *Nucl. Fusion* **57** 086039
- [10] Morgan K.D., Hossack A.C., Hansen C.J., Nelson B.A. and Sutherland D.A. 2021 High-speed feedback control of an oscillating magnetic helicity injector using a graphics processing unit *Rev. Sci. Instrum.* **92** 053530
- [11] Kaptanoglu A.A., Morgan K.D., Hansen C.J. and Brunton S.L. 2020 Characterizing magnetized plasmas with dynamic mode decomposition *Phys. Plasmas* **27** 032108
- [12] Nayak I. and Teixeira F.L. 2020 Dynamic mode decomposition for prediction of kinetic plasma behavior (Online 27–31 July 2020) *2020 Int. Applied Computational Electromagnetics Society Symp. (ACES)* (IEEE) pp 1–2
- [13] Willcox K. and Peraire J. 2002 Balanced model reduction via the proper orthogonal decomposition *AIAA J.* **40** 2323–30
- [14] Ariola M., Ambrosino G., Pironti A., Lister J.B. and Vyas P. 2002 Design and experimental testing of a robust multi-variable controller on a tokamak *IEEE Trans. Control Syst. Technol.* **10** 646–53
- [15] Ariola M. and Pironti A. 2005 Plasma shape control for the JET tokamak: an optimal output regulation approach *IEEE Control Syst. Mag.* **25** 65–75
- [16] Moreau D. *et al* 2003 Real-time control of the  $q$ -profile in JET for steady state advanced tokamak operation *Nucl. Fusion* **43** 870
- [17] Goodman T.P. *et al* 2011 Sawtooth pacing by real-time auxiliary power control in a tokamak plasma *Phys. Rev. Lett.* **106** 245002
- [18] Hu W. *et al* 2018 Active real-time control of Alfvén eigenmodes by neutral beam and electron cyclotron heating in the DIII-D tokamak *Nucl. Fusion* **58** 124001
- [19] Levesque J.P. *et al* 2013 Multimode observations and 3D magnetic control of the boundary of a tokamak plasma *Nucl. Fusion* **53** 073037
- [20] Galperti C. *et al* 2014 Development of real-time MHD markers based on biorthogonal decomposition of signals from Mirnov coils *Plasma Phys. Control. Fusion* **56** 114012
- [21] Victor B.S., Akcay C., Hansen C.J., Jarboe T.R., Nelson B.A. and Morgan K.D. 2015 Development of validation metrics using biorthogonal decomposition for the comparison of magnetic field measurements *Plasma Phys. Control. Fusion* **57** 045010
- [22] Galperti C., Coda S., Duval B.P., Llobet X., Milne P., Sauter O., Moret J.M. and Testa D. 2017 Integration of a real-time node for magnetic perturbations signal analysis in the distributed digital control system of the TCV tokamak *IEEE Trans. Nucl. Sci.* **64** 1446–54
- [23] Kaptanoglu A.A., Morgan K.D., Hansen C.J. and Brunton S.L. 2021 Physics-constrained, low-dimensional models for magnetohydrodynamics: first-principles and data-driven approaches *Phys. Rev. E* **104** 015206
- [24] Kaptanoglu A.A., Callahan J.L., Aravkin A., Hansen C.J. and Brunton S.L. 2021 Promoting global stability in data-driven models of quadratic nonlinear dynamics *Phys. Rev. Fluids* **6** 094401
- [25] Cannas B., de Vries P.C., Fanni A., Murari A., Pau A. and Sias G. 2015 Automatic disruption classification in JET with the ITER-like wall *Plasma Phys. Control. Fusion* **57** 125003
- [26] Rea C., Granetz R.S., Montes K., Tinguely R.A., Eidiets N., Hanson J.M. and Sammuli B. 2018 Disruption prediction investigations using machine learning tools on DIII-D and Alcator C-Mod *Plasma Phys. Control. Fusion* **60** 084004
- [27] Murari A., Lungaroni M., Peluso E., Gaudio P., Vega J., Dormido-Canto S., Baruzzo M. and Gelfusa M. 2018 Adaptive predictors based on probabilistic SVM for real time disruption mitigation on JET *Nucl. Fusion* **58** 056002
- [28] Kates-Harbeck J., Svyatkovskiy A. and Tang W. 2019 Predicting disruptive instabilities in controlled fusion plasmas through deep learning *Nature* **568** 526–31
- [29] Aymerich E., Sias G., Pisano F., Cannas B., Carcangiu S., Sozzi C., Stuart C., Carvalho P. and Fanni A. 2022 Disruption prediction at JET through deep convolutional neural networks using spatiotemporal information from plasma profiles *Nucl. Fusion* **62** 066005
- [30] Montes K.J. *et al* 2019 Machine learning for disruption warnings on Alcator C-Mod, DIII-D, and EAST *Nucl. Fusion* **59** 096015
- [31] De Bustos A., Ascasíbar E., Cappa A. and Mayo-García R. 2021 Automatic identification of MHD modes in magnetic fluctuations spectrograms using deep learning techniques *Plasma Phys. Control. Fusion* **63** 095001
- [32] Škvára V., Šmídl V., Pevný T., Seidl J., Havránek A. and Tskhakaya D. 2020 Detection of Alfvén eigenmodes on COMPASS with generative neural networks *Fusion Sci. Technol.* **76** 962–71
- [33] Woods B.J.Q., Duarte V.N., Fredrickson E.D., Gorelenkov N.N., Podestà M. and Vann R.G.L. 2020 Machine learning characterization of Alfvénic and sub-Alfvénic chirping and correlation with fast-ion loss at NSTX *IEEE Trans. Plasma Sci.* **48** 71–81
- [34] Jalalvand A. *et al* 2021 Alfvén eigenmode classification based on ECE diagnostics at DIII-D using deep recurrent neural networks *Nucl. Fusion* **62** 026007
- [35] Zhao L.-L., Zank G.P., He J.S., Telloni D., Adhikari L., Nakanotani M., Kasper J.C. and Bale S.D. 2021 MHD and ion kinetic waves in field-aligned flows observed by Parker solar probe *Astrophys. J.* **922** 188
- [36] Heidbrink W.W. *et al* 2020 ‘BAAE’ instabilities observed without fast ion drive *Nucl. Fusion* **61** 016029
- [37] Beliën A.J.C., Poedts S. and Goedbloed J.P. 1996 Magneto-hydrodynamic continua and stratification induced Alfvén eigenmodes in coronal magnetic loops *Phys. Rev. Lett.* **76** 567
- [38] Duong H.H., Heidbrink W.W., Strait E.J., Petrie T.W., Lee R., Moyer R.A. and Watkins J.G. 1993 Loss of energetic beam ions during TAE instabilities *Nucl. Fusion* **33** 749
- [39] White R.B., Fredrickson E., Darrow D., Zarnstorff M., Wilson R., Zweben S., Hill K., Chen Y. and Fu G. 1995 Toroidal Alfvén eigenmode-induced ripple trapping *Phys. Plasmas* **2** 2871–3
- [40] Madsen B. *et al* 2020 Tomography of the positive-pitch fast-ion velocity distribution in DIII-D plasmas with Alfvén eigenmodes and neoclassical tearing modes *Nucl. Fusion* **60** 066024
- [41] Van Zeeland M.A. *et al* 2016 Electron cyclotron heating can drastically alter reversed shear Alfvén eigenmode activity in DIII-D through finite pressure effects *Nucl. Fusion* **56** 112007
- [42] Van Zeeland M.A. *et al* 2007 Coupling of global toroidal Alfvén eigenmodes and reversed shear Alfvén eigenmodes in DIII-D *Phys. Plasmas* **14** 056102
- [43] Pace D.C. *et al* 2016 Control of power, torque, and instability drive using in-shot variable neutral beam energy in tokamaks *Nucl. Fusion* **57** 014001
- [44] Li P. *et al* 2021 Dynamics between toroidal Alfvén eigenmode evolution and turbulence suppression under RMP on EAST *Nucl. Fusion* **61** 086020
- [45] Tang S.X., Carter T.A., Crocker N.A., Heidbrink W.W., Lestz J.B., Pinsker R.I., Thome K.E., Van Zeeland M.A. and

- Belova E.V. 2021 Stabilization of Alfvén eigenmodes in DIII-D via controlled energetic ion density ramp and validation of theory and simulations *Phys. Rev. Lett.* **126** 155001
- [46] Turnbull A.D., Strait E.J., Heidbrink W.W., Chu M.S., Duong H.H., Greene J.M., Lao L.L., Taylor T.S. and Thompson S.J. 1993 Global Alfvén modes: theory and experiment *Phys. Fluids B* **5** 2546–53
- [47] Heidbrink W.W., Strait E.J., Chu M.S. and Turnbull A.D. 1993 Observation of beta-induced Alfvén eigenmodes in the DIII-D tokamak *Phys. Rev. Lett.* **71** 855
- [48] Heidbrink W.W., Van Zeeland M.A., Austin M.E., Crocker N.A., Du X.D., McKee G.R. and Spong D.A. 2021 Stability of beta-induced Alfvén eigenmodes (BAE) in DIII-D *Nucl. Fusion* **61** 066031
- [49] Sharapov S.E. *et al* 2002 Alfvén wave cascades in a tokamak *Phys. Plasmas* **9** 2027–36
- [50] Kimura H. *et al* 1998 Alfvén eigenmode and energetic particle research in JT-60U *Nucl. Fusion* **38** 1303
- [51] Cheng C.Z., Chen L. and Chance M.S. 1985 High- $n$  ideal and resistive shear Alfvén waves in tokamaks *Ann. Phys., NY* **161** 21–47
- [52] Cheng C.Z. and Chance M.S. 1986 Low- $n$  shear Alfvén spectra in axisymmetric toroidal plasmas *Phys. Fluids* **29** 3695–701
- [53] Heidbrink W.W., Strait E.J., Doyle E., Sager G. and Snider R.T. 1991 An investigation of beam driven Alfvén instabilities in the DIII-D tokamak *Nucl. Fusion* **31** 1635
- [54] Wong K.L. *et al* 1991 Excitation of toroidal Alfvén eigenmodes in TFTR *Phys. Rev. Lett.* **66** 1874
- [55] Heidbrink W.W. 2008 Basic physics of Alfvén instabilities driven by energetic particles in toroidally confined plasmas *Phys. Plasmas* **15** 055501
- [56] Austin M.E. and Lohr J. 2003 Electron cyclotron emission radiometer upgrade on the DIII-D tokamak *Rev. Sci. Instrum.* **74** 1457–9
- [57] Lao L.L., St John H., Stambaugh R.D., Kellman A.G. and Pfeiffer W. 1985 Reconstruction of current profile parameters and plasma shapes in tokamaks *Nucl. Fusion* **25** 1611
- [58] Jaeger H. 2001 The ‘echo state’ approach to analysing and training recurrent neural networks *Tech. Rep. GMD Report vol 148* (German National Research Center for Information Technology)
- [59] Jalalvand A., Abbate J., Conlin R., Verdoolaege G. and Kolemen E. 2021 Real-time and adaptive reservoir computing with application to profile prediction in fusion plasma *IEEE Trans. Neural Netw. Learn. Syst.* **33** 2630–41
- [60] Chandel R. and Gupta G. 2013 Image filtering algorithms and techniques: a review *Int. J. Adv. Res. Comput. Sci. Softw. Eng.* **3** 872–8
- [61] Ahn E., Jalalvand A. and Kolemen E. 2021 Localizing Alfvén eigenmodes in plasma based on high resolution ECE spectrograms at DIII-D using autoencoders and image processing techniques *Bull. Am. Phys. Soc.* **2021** JP11.102
- [62] Rahman M.A. and Wang Y. 2016 Optimizing intersection-over-union in deep neural networks for image segmentation (Las Vegas, NV, USA December 12–14, 2016) *Int. Symp. on Visual Computing* (Springer) pp 234–44
- [63] Qiao Z., Zhao J., Zhu J., Tyree Z., Mudalige P., Schneider J. and Dolan J.M. 2020 Human driver behavior prediction based on urbanflow (Online 31 May – 31 August 2020) *2020 IEEE Int. Conf. on Robotics and Automation (ICRA)* (IEEE) pp 10570–6
- [64] Hao Z., Averbuch-Elor H., Snively N. and Belongie S. 2020 Dualsdf: semantic shape manipulation using a two-level representation *Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recognition* pp 7631–41
- [65] Montes K.J., Rea C., Tinguely R.A., Sweeney R., Zhu J. and Granetz R.S. 2021 A semi-supervised machine learning detector for physics events in tokamak discharges *Nucl. Fusion* **61** 026022