Disruption Prediction with Deep Neural Networks

João Calado Vicente

Instituto Superior Técnico

MEIC-T, 78470

Abstract—Plasma disruptions remain one of the main problems in nuclear fusion reactors, such as tokamaks. The task of predicting these disruptions during an ongoing experiment presents itself as an opportunity for machine learning and deep learning approaches. Being able to mitigate these harmful events would represent a significant improvement towards nuclear fusion as a source of energy. Several methods have been implemented over the last decades, using different plasma diagnostic signals to build a predictive model for disruptions. In this work, our goal is to use a deep neural network to predict the occurrence, and estimated time of a disruption, during a plasma discharge. In contrast with existing approaches that typically use several single-channel signals from different diagnostics, here we use a multi-channel signal from a single diagnostic (the bolometer system), which is capable of determining the shape and position of the plasma inside the reactor. This can be used as the discriminant factor for disruption prediction.

Keywords—Nuclear Fusion, Plasma Diagnostics, Disruption Prediction, Deep Learning, Neural Networks

1 Introduction

Nuclear Fusion is a process that allows energy generation through fusion reactions. In order for these reactions to occur, matter needs to be sufficiently heated so that it becomes plasma. In addition, sufficient energy confinement is necessary in nuclear fusion.

The current leading designs for plasma confinement are the tokamak [1] and inertial confinement [2]. A tokamak is a device that confines plasma by using a powerful magnetic field. This design differs from inertial confinement in the sense that plasma is not compressed, instead it flows inside the reactor while only being confined to avoid contact with the walls of the reactor. The magnetic field acts as a recipient that is not affected by heat, unlike any ordinary solid container. The tokamak design represented a major improvement in plasma science, because it allowed higher temperature levels and longer plasma confinement times. The largest currently operating tokamak in the world, and the source of the data used in this work, is the Joint European Torus (JET) [3].

One of the main hazards when using a tokamak device to confine plasma during a fusion experiment is the occurrence of plasma disruptions.

1.1 Plasma Disruptions

A plasma disruption is an uncontrolled loss of plasma current and consequently of its own confinement [4]. These unstable occurrences are extremely harmful to tokamak reactors and remain one of the main threats to the general integrity of these reactors during experiments.

Disruptions can be structured into 4 different phases: initiating event, precursor phase, thermal quench and current quench [5]. In tokamak reactors, the initiating event often involves heavy metal particles being torn from the wall of the reactor, which are then carried by the plasma flow and begin concentrating at the plasma core. This event causes the plasma to become unstable and is denominated as the precursor phase, which ultimately leads to the thermal quench. The final phase is the loss of the plasma current, also called current quench, and consequently the end of the experiment.

1.2 Disruption Prediction

Disruption prediction consists of combining statistical analysis with machine learning tech-

niques in order to predict plasma disruptions during experiments conducted in fusion reactors. In this work, we use deep neural networks trained on data collected at JET, using a particular diagnostic (the bolometer system), to predict the occurrence of these disruptions during a plasma experiment.

Since these plasma disruptions can harm tokamak reactors and are one of the main threats during plasma experiments, preventing or mitigating these disruptions by deploying security measures upon their early detection would be an essential procedure to solve this problem.

2 Related Work

Disruption prediction is a research field that aims to improve the detection of plasma disruptions in fusion reactors (especially tokamaks), before they actually happen, often using machine learning techniques [6]–[11]. Typically, multiple signals from plasma diagnostics are extracted and combined in order to enable the detection of an impending disruption and, in some cases, the remaining time-to-disruption [9], [12].

Table 1
Popular diagnostic signals used in the field of disruption prediction [13]–[15]

#	Signal	Unit
1	Plasma Current	A
2	Locked Mode Amplitude	T
3	Radiated Power	W
4	Plasma Density	m^{-3}
5	Total Input Power	W
6	Plasma Internal Inductance	
7	Safety Factor	
8	Poloidal Beta	
9	Stored Diamagnetic Energy Time Derivative	W
10	Plasma Centroid Vertical Position	m

Table 1 shows most of the diagnostic signals used in the approaches further described in this section.

Many different techniques have been used throughout the years to tackle this problem [6], with some of them being more effective than others [14].

An approach based on fuzzy logic while taking into account an analysis based on Classification and Regression Trees (CART) [16] is described in [5]. This predictor was tested on a dataset with data collected from JET campaigns, which contained 292 disruptive discharges and 385 non-disruptive discharges. The adaptive predictor yielded an 8.9% ratio of missed alarms and a 7.0% ratio of false alarms.

The APODIS predictor [6], [14], [17]–[19], which is an implementation based on support vector machines, was able to successfully detect 93% of the disruptive experiments, with a 5% ratio of false alarms during its live performance on pulses in the C28 campaign, after deployment in the JET's infrastructure [17].

Some interesting aspects that are worth mentioning are related to two other approaches: The Venn predictor from scratch [20] and the SPAD (Single signal Predictor based on Anomaly Detection) [21]. The Venn predictor was able to achieve a 90.21% success rate using only 2 disruptive examples and 2 non-disruptive examples as the basis for defining the prediction model, which makes it a suitable candidate for implementation in a tokamak that does not have a high tolerance for disruptions [20]. On the other hand, the SPAD was able to outperform the APODIS when the average warning time is above 255 ms (despite the overall performance offering lower results), registering an 83.4% success rate, compared to 80.3% of the APODIS.

Finally, the most recent approach developed in the field of disruption prediction is the Fusion Recurrent Neural Network (FRNN), presented in [22], [23]. This approach managed to obtain a 93.5% rate of predicted disruptions while keeping the false alarm rate at 7.5%, after being tested in a dataset containing 1200 pulses [23].

3 The Bolometer System at JET

The approach in the present work is based on the notion of plasma tomography [24], which

consists in obtaining the plasma profile on a cross-section of the fusion device, via a reconstruction process based on the measurements collected by diagnostics systems placed around the plasma. An example of such diagnostic is the bolometer system, which measures plasma radiation along multiple lines of sight.

In particular, the goal of this work is to predict disruptions by using the radiation measurements on a cross-section of the tokamak, which offers a spatial notion of the position of the plasma inside the reactor, instead of the more existing approaches that use a single value of total radiation among other plasma diagnostic signals.

To this end, the input signals we will be using are obtained from a specific bolometer system available at JET.

3.1 The KB5 Bolometer System

The KB5 bolometer system [25]–[27] is the latest plasma radiation measurement system installed at JET.

KB5 is composed of 2 cameras (one vertical camera, KB5V, and one horizontal camera, KB5H) with 48 viewing chords and 8 backup viewing chords belonging to KB5V, making a combined total of 56 measurement channels across one section of the plasma chamber, as shown in Figure 1. Each camera (KB5V & KB5H) has 24 chords, with 8 of those directed at the divertor region of the reactor, the bottom region of the chamber that has a greater concentration of plasma during experiments [27].

This new measurement system is a significant upgrade over the first (KB1) bolometer system with its improvements including a higher number of viewing chords, a larger measurement angle, higher detectable range of radiation, higher sensitivity, lower noise and lower detectable signal [27].

All of the KB5 signals are available to the Real Time Control System implemented at JET, with a maximum sampling rate of 5 kHz [25]. This configuration allows us to obtain 5000 measurements of the plasma radiation every second, along a viewing chord in a cross-section of the reactor, for each of the 56 different channels.

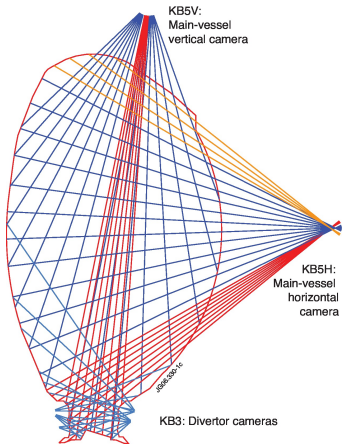

Figure 1. KB5 lines-of-sight (Source: EUROfusion figures database JG06.330-1c)

3.2 Dataset Retrieval and Preprocessing

The dataset used in this work was constructed from data colected at JET, pertaining to plasma campaigns C28 through C36 which encompass all the pulses ranging from no. 80128 to no. 92504, both disruptive and non-disruptive.

In total, the dataset contains measurements from 8115 non-disruptive pulses and 1683 disruptive pulses.

The data was retrieved with a sampling rate of 200 Hz, meaning that a value of plasma radiation was retrieved for every 5 ms of a plasma shot, for each one of the 56 channels comprised in the bolometer system. Since the original sampling rate of the KB5 bolometer system is 5 kHz, each retrieved value represents the mean value of 25 consecutive measurements across each channel. This subsampling is made especially for noise reduction, but it also contributes to reduce the amount of data, and allow for a faster training and testing of the implemented prototype.

Furthermore, a quick glance at the retrieved dataset allowed us to discard the data collected from the backup channels, as these were turned off in most plasma experiments, and did not seem to collect any data. These were the channels 25 to 32 of KB5V.

Through a more complete analysis, some more channels were discarded as they were deeply interfeering with the learning and decisions of the prototype neural networks, on account of being broken. The aforementioned

channels are the channels 15, 16 and 23 of KB5V and 19 to 24 of KB5H. These channels displayed erratic measurements, such as negative plasma radiation values and extremely high values throughout each pulse.

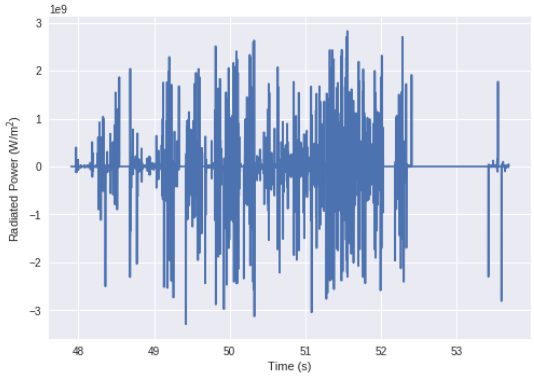

Figure 2. Broken KB5 channel readings for pulse no. 92213

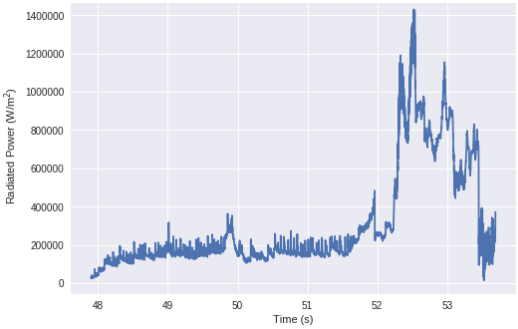

Figure 3. Standard KB5 channel readings for pulse no. 92213

The dataset was divided in three distinct parts, used for training, validating and testing the neural networks. 80% of the entire dataset was used to train the networks, while the other 20% was divided in half for testing and validation purposes. In each of these three parts, the ratio of non-disruptive to disruptive examples was approximately maintained in relation to the ratio observed in the entire dataset. No undersampling or oversampling was made in order to balance the dataset, as that would not reflect the natural ratio of disruptive pulses in relation to all the experiments made, and that factor could influence the outcome of this work. The training dataset contained 6479 non-disruptive pulses and 1289 disruptive pulses, while both the validation and test datasets contained 818 non-disruptive pulses and 197 disruptive pulses.

3.3 Plasma Tomography

The main application of the bolometer system at JET is to provide a tomographic reconstruction of the plasma radiation profile along a poloidal cross-section of the device.

At JET, the signals from the horizontal and vertical cameras of KB5 are used to recontruct the plasma profile according to the method described in [28].

Although plasma tomography is beyond the scope of this work, it is important to note that we are using the same input signals for the purpose of disruption prediction. Also, from previous work done in the field of plasma tomography, we are able to see that disruptive events are shown in the spatial behaviour of the plasma [29].

As we have seen in the previous section, disruption prediction is usually attained with several overall plasma diagnostic signals, which provide a global value of what is happening inside the device. In contrast, in this work we use the bolometer signals, which provide us with a more detailed 2D view of the plasma, including its position and shape. In principle, this should allow us to learn disruptive behavior based on the internal spatial configuration of the plasma.

Therefore, we feed the bolometer signals to a neural network that will learn to identify disruptive behaviour and predict incoming disruptions.

4 Neural Networks

Two prototype neural networks for disruption prediction were developed in order to achieve the desired goals, one that predicts the Time-to-Disruption for a given sample, and one that predicts the Probability of a Disruption happening. Both networks were developed in Python, using the Keras deep learning library with TensorFlow as the backend. GPU computing is used, through the CUDA and cuDNN libraries developed by NVIDIA, in order to achieve significantly faster training and testing times.

Given the temporal nature of this problem, a recurrent neural network architecture was considered as the basis for these prototypes.

4.1 Architecture

Both networks have a similar architecture and their input consists of several samples of the plasma radiation values measured by each of the 39 working channels belonging to the KB5 bolometer system over a pre-determined temporal window of 1 s.

The networks are composed by two LSTM layers that contain 100 units each, and two Dense layers, one with 100 neurons and another (the final layer in each of the neural networks) with a single neuron, used to provide a time-to-disruption output in one case, and a probability of disruption in the other. All of the activations used in the LSTM and Dense layers are of type ReLU, except for the final Dense layer of the Probability of Disruption network, which has a sigmoid activation. Some dropout was used in the recurrent layers of the Time-to-Disruption network in order to avoid the premature overfitting that was occurring during training. Figure 4 provides a detailed illustration of the network architecture.

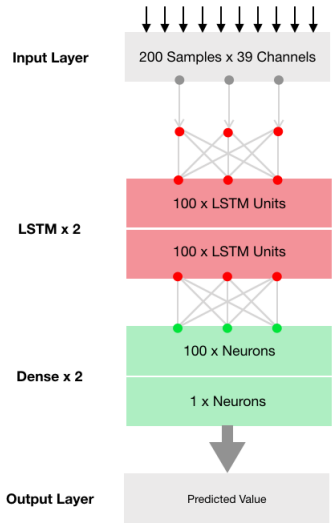

Figure 4. Base architecture for both of the disruption prediction neural networks.

Through some experiments we were able to see that overcomplicated architectures would lead to the surfacing of new problems, with little to no improvement on the performance of the network. An example of such a problem was the increase in the time it took to analyse the input samples and predict an outcome, as this would make it impossible to analyse each sample before the next sample was measured and made available for analysis. This would render it useless during a real-time application of the network, which is one of the main goals of this work.

4.2 Training

During the training phase of each network, a batch generator is used to feed the data in batches of samples with the intent to simulate a real-time input of data. This avoids constantly loading the desired samples, one by one, into GPU memory, as that represents an enormous increase in the training time due to fixed data transfer times between RAM and GPU memory. Both networks were trained on a machine that contained a single NVIDIA GeForce GTX 1070 GPU.

Each batch of samples passed to the networks consisted of 256 examples of plasma radiation measurements over a temporal window of 1 s (thus making it 200 measurements due to the sampling frequency of 200 Hz), for each of the 39 working channels of KB5. Here, the difference between both networks lies in the labels used for each sample. For the Time-to-Disruption network each example was labeled with the Time-to-Disruption, which was calculated from the last measurement in the example using the actual time of the measurement and the recorded time of disruption. For the Probability network, each example was labeled with 0 or 1, depending if the pulse it originated from was either non-disruptive or disruptive, respectively.

The training dataset used for each network was also different, as the Time-to-Disruption network was trained using only disruptive examples, in contrast with the probability network that was trained using the entire training dataset (containing examples from both disruptive and non-disruptive discharges). Each example belonging to a disruptive pulse was selected using an exponential function that

would yield an higher probability of selecting a temporal window that was closer to the time of disruption.

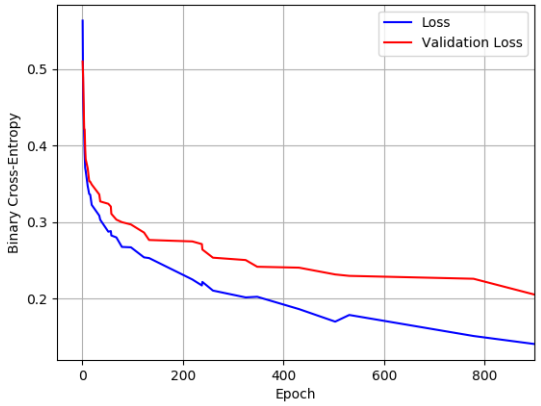

Figure 5. Loss values registered during the training phase of the Disruption Probability neural network.

During training both networks used the Adam optimizer [30], but the loss function used in the Time-to-Disruption network was the mean absolute error (MAE), given in seconds (s), and the loss function used in the Probability network was the Binary Cross-Entropy. The training stopped when no new minimum was obtained for the validation loss, during the same number of epochs as those that were necessary to reach the current minimum value. Figure 5 provides a detailed look at the evolution of the loss values during the training phase, for the Probability of Disruption network.

4.3 Joint Solution

While testing the performance of the system in the context of predicting disruptions, we decided to use the outputs of both networks to classify a given sample as disruptive or non-disruptive, and an alarm would be triggered each time a disruptive sample was detected while scanning each pulse in the test dataset. A sample is classified as disruptive by comparing the outputs from both networks to predefined threshold values. These threshold values were tweaked during the testing phase with the goal of obtaining the best possible classification accuracy. The values that reflected this were 0.85 for the Probability of Disruption output and 2.2 s for the Time-to-Disruption output.

5 Experiments and Results

The results presented in this section were obtained by running each pulse in the testing subset through the trained networks. For each of the two networks, a prediction is obtained for each consecutive 50 ms of the pulse in the form of a sliding window that contains 1 second of plasma radiation measurements. Once for every three samples analysed (every 150 ms of the pulse), the mean values of the predictions are calculated and these values are the ones that are compared to the predetermined thresholds, thus yielding the final output of the prototype, considering the pulse to be disruptive or non-disruptive for each of those instants (See Figure 6).

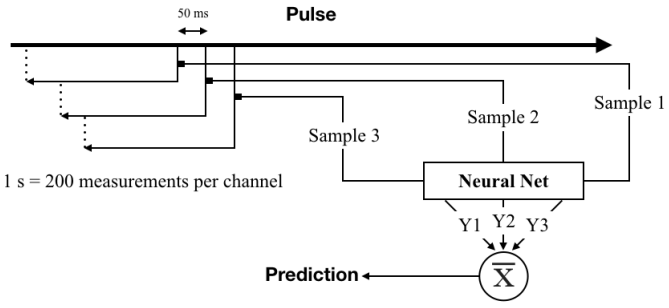

Figure 6. Analysis procedure for each pulse.

This time-frame was used in order to be able to analyse and predict the outcome of the given sample before the next sample is available for analysis, as this is a requirement for a real-time deployment of the system.

With the results from the predictor, it is possible to reach one of four outcomes: true positives, false positives, true negatives and false negatives.

5.1 True Positives (Predicted Disruptions)

In an efficient predictor, this should represent the bulk of the outcomes related to disruptive examples. A pulse is classified as disruptive if at any moment, the probability of disruption is high at the same time that the time-to-disruption registers smaller values, which

means that both predictions reach the thresholds that would trigger an alarm. In Figure 7, we can observe an example of a pulse where a correct prediction of a disruption would occur, as both networks predict the time-to-disruption and the probability of disruption, with a certain degree of accuracy, throughout the entire pulse.

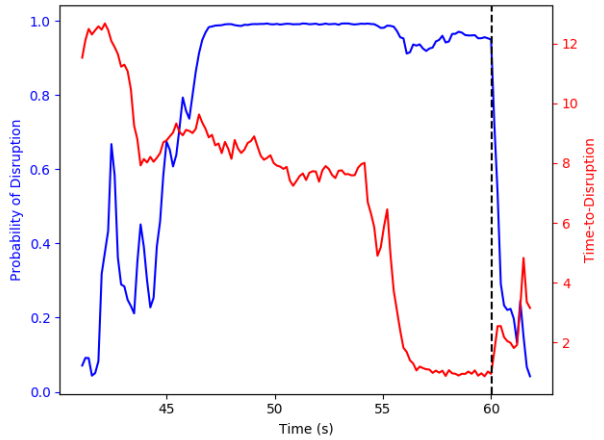

Figure 7. Predictions for pulse no. 85978. Disruption occurs at ∼60.0 s.

5.2 False Negatives (Missed Alarms)

There were some cases of missed alarms during the testing phase. A fraction of 16% of the total disruptive pulses was not identified as such, this was a result of tweaking the thresholds in order to minimize the ratio of false alarms triggered by the predictor.

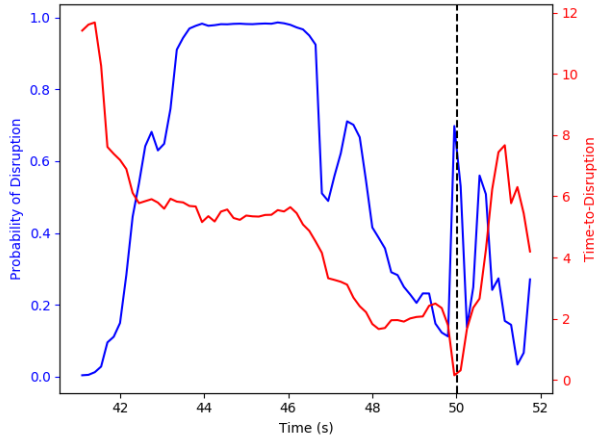

Figure 8. Predictions for pulse no. 92269. Disruption occurs at ∼50.0 s.

Figure 8 represents an example of such a case. In an earlier stage the probability of disruption is high but the time-to-disruption is also high, so an alarm would not be triggered. In a later stage, the time-to-disruption is lower but the probability of disruption plummets, which would also prevent an alarm from being triggered. Even though the probability rises in the final instants before the disruption, it does not reach high enough values for a disruption to be detected, and even if it would, it would probably be too late to take mitigating actions, thus being classified as a missed (late) alarm.

5.3 True Negatives (Correct Prediction)

This case should represent the bulk of the predictions obtained in non-disruptive pulse examples, for a reliable disruption predictor. That is indeed the case when it comes to the prototype developed in this work, as 87% of the non-disruptive pulses were tested without the predictor triggering any alarm.

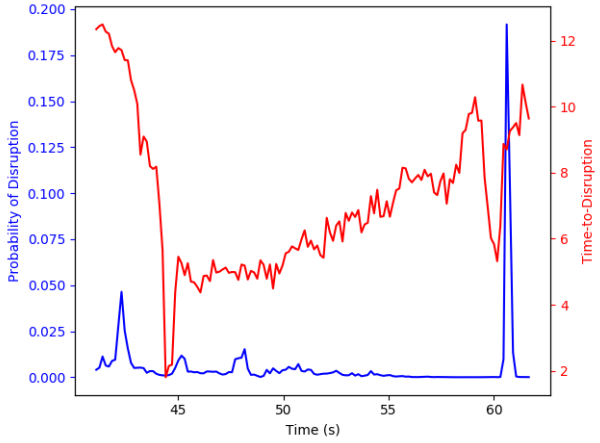

Figure 9. Predictions for pulse no. 86077

This example does not trigger any alarm due to the fact that both outputs need to be taken into account for a disruption to be detected. Even though there is a moment in which the Time-to-Disruption reaches alarm-triggering values, the probability of disruption is almost zero, and never surpasses 0.2.

5.4 False Positives (False Alarms)

One of the main issues in disruption prediction, besides missed alarms, are false alarms.

These occur when disruptive behaviour, or behaviour that could be classified as disruptive, is observed in pulses that do not result in a disruption. This behaviour could happen due to instabilities during the plasma experiment that are sufficiently relevant to be detected by the bolometer system, but that are not unstable enough to cause a full disruption in the plasma experiment. These behavioral changes in the plasma are sometimes sufficient to confound the neural networks into believing that a disruption is imminent.

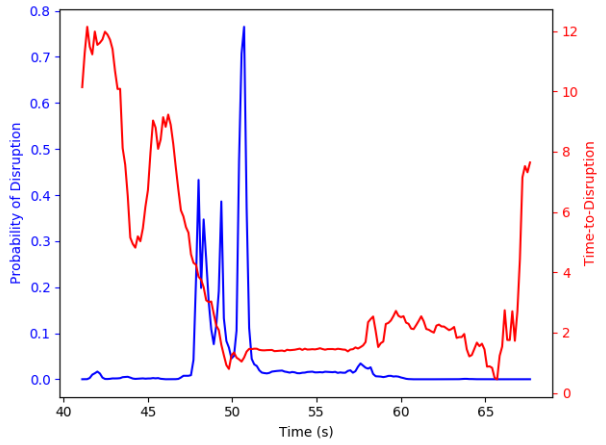

Figure 10. Predictions for pulse no. 87867

Figure 10 shows an example of a situation where a sudden spike in the probability of disruption would trigger an alarm in a non-disruptive pulse, if the threshold used for that output would be less restrictive.

During the best attempt at predicting disruptions without triggering many false alarms, it was registered that 13% of all the non-disruptive pulses in the dataset triggered an incorrect alarm.

5.5 Discussion

The prediction system was able to successfully predict 84% of the disruptions while triggering a false alarm for 13% of the non-disruptive pulses. The system had an overall accuracy of 86.4% while analysing the pulses in the test dataset.

Even though the observed results do not reflect the highest accuracy ever achieved in the field of disruption prediction, we can still consider this to be a somewhat reliable approach to disruption prediction. Figure 11 depicts the Receiver Operating Characteristic (ROC) curve for the developed system. The overall accuracy of the classifier, while considering different threshold combinations, can be measured by the area under the ROC curve (AUC) which is 0.91 for the ROC curve calculated using the results from different tests. Given the observed results, we can then conclude that the performance of the system in predicting disruptions in plasma discharges is quite satisfactory.

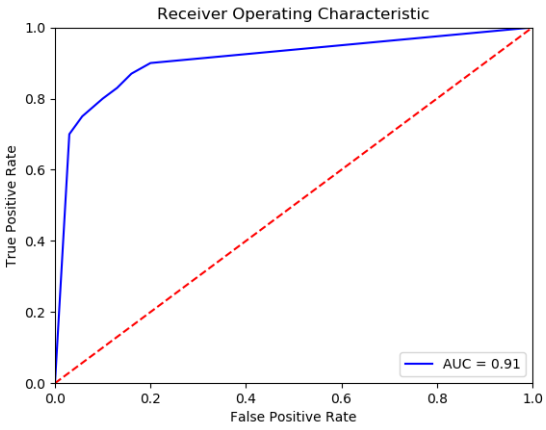

Figure 11. Receiver Operating Characteristic curve for the results obtained using several threshold values.

The results obtained by our system may not be as excellent as other approaches due to the fact that in this work only one plasma diagnostic system is used for the basis of the prediction, and the bolometer system in question (KB5) does not have an excellent signal-to-noise ratio in the context of disruption prediction. Also, by using only a single diagnostic system, we may be overlooking important information about the behaviour of the plasma that could be decisive for a correct prediction of a disruptive event. Other methods that rely on several diagnostic signals are able to complement these flaws with data that introduces several "perspectives" of the behaviour of the plasma during a live experiment, especially when a system is able to successfully decide which signal best contributes for a correct prediction in any given moment of the experiment.

The testing results registered an average prediction time of 4.33 s for the performance

Table 2
Comparison between the accuracy and average warning time of detected disruptions, for several predictors, when disruptions are detected at an earlier stage of the discharge. Data gathered from [14]

	Average Warning Time	Accuracy
APODIS	$> 400ms$	80.3%
SPAD	$\sim 288ms$	83.4%
JET Locked Mode Predictor	$255ms$	65.0%
Our Prototype Predictor	$4.33s$	84.0%

presented above, a value that is substantially higher than the ones registered by all the other methods. Table 2 offers a comparison between the performance of several predictors at higher average warning times. Considering only the earlier stages of plasma discharges, it is possible to observe that the prediction system developed in this work is able to obtain a better overall performance than some of the other methods, exhibiting a much higher average warning time. This behaviour can be justified by the diagnostic system that is used as input in order to predict the disruptions in this work, as the KB5 bolometer system is able to give the network a notion of the position of the plasma inside the reactor, and this spatial position exhibits signs of disruptive behaviour earlier than other diagnostic signals.

6 Conclusion

In this work we developed a system that consists of two neural networks in order to predict plasma disruptions, using only a single plasma diagnostic system that utimately yields information about the radiation profile of the plasma. The data used in this work was gathered at JET through the bolometer system installed at the tokamak reactor (the KB5 bolometer system) and divided into different sets used for training and testing purposes. Even though the performance registered during the testing phase of the system was not as good as other methods, being able to predict 84% of the disruptions while triggering false alarms on 13% of non-disruptive pulses, the overall results that were obtained are still positive and demonstrated that disruption prediction based on the radiation profile of the plasma is a viable method. These error rates are mostly due to the noise ratio in the diagnostic signals used as input for the neural networks which, despite the large training time, are not able to achieve a perfect generalization. However, we were able to further improve the prediction performance displayed by each single neural network by combining the outputs of the two networks and using both values in order to reach the outcome of the prediction system.

The most interesting metric displayed in the results obtained from using this method was the average warning time for the predicted disruptions. The system demonstrated a better performance than most of the other prediction methods when considering how early in the plasma discharge the disruptions were predicted, with an average warning time of 4.33s. This ability to detect disruptions on the earlier stages of the plasma experiments derives from the use of the KB5 diagnostic signals as the basis for the prediction, and more specifically because these signals yield information about the position of the plasma inside a cross-section of the reactor, which in turn displays signs of disruptive behaviour earlier than other diagnostic signals.

We can conclude that the best improvements to be made on this approach are the reduction of the noise ratio in the input signals, or the combination of the KB5 signals with other relevant diagnostic signals that are often used for predicting plasma disruptions. However, we still consider that the results obtained in this work are able to demonstrate that this is a reliable approach to disruption prediction.

References

[1] J. D. Strachan *et al.*, "High-temperature plasmas in a tokamak fusion test reactor," *Phys. Rev. Lett.*, vol. 58, pp. 1004–1007, 1987.

[2] J. Lindl, "Development of the indirect-drive approach to inertial confinement fusion and the target physics basis for ignition and gain," *Physics of Plasmas*, vol. 2, no. 11, pp. 3933–4024, 1995.

[3] P. Rebut, R. Bickerton, and B. Keen, "The joint european torus: installation, first results and prospects," *Nuclear Fusion*, vol. 25, no. 9, p. 1011, 1985.

[4] A. Pereira, J. Vega, R. Moreno, S. Dormido-Canto, G. A. Rattá, and F. Pavón, "Feature selection for disruption prediction from scratch in JET by using genetic algorithms and probabilistic predictors," *Fusion Engineering and Design*, vol. 96-97, no. Supplement C, pp. 907 – 911, 2015, proceedings of the 28th Symposium On Fusion Technology (SOFT-28).

[5] A. Murari, G. Vagliasindi, P. Arena, L. Fortuna, O. Barana, M. Johnson, and JET-EFDA Contributors, "Prototype of an adaptive disruption predictor for JET based on fuzzy logic and regression trees," *Nuclear Fusion*, vol. 48, no. 3, p. 035010, 2008.

[6] A. Murari, J. Vega, P. Boutot, B. Cannas, S. Dormido-Canto, A. Fanni, J. López, R. Moreno, A. Pau, G. Sias, J. Ramírez, and G. Verdoolaege, "Latest developments in data analysis tools for disruption prediction and for the exploration of multimachine operational spaces," in *24th IAEA Fusion Energy Conference, Proceedings*, 2012, p. 8.

[7] A. Murari, J. Vega, G. Rattá, G. Vagliasindi, M. Johnson, S. Hong, and JET-EFDA Contributors, "Unbiased and non-supervised learning methods for disruption prediction at JET," *Nuclear Fusion*, vol. 49, no. 5, p. 055028, 2009.

[8] A. Sengupta and P. Ranjan, "Forecasting disruptions in the ADITYA tokamak using neural networks," *Nuclear Fusion*, vol. 40, no. 12, p. 1993, 2000.

[9] B. Cannas, A. Fanni, G. Sias, P. Sonato, M. Zedda, and JET EFDA Contributors, "Neural approaches to disruption prediction at JET," in *31st EPS Conference on Plasma Physics*, ser. ECA, vol. 28G, 2004.

[10] B. Cannas, A. Fanni, A. Murari, A. Pau, G. Sias, and JET EFDA Contributors, "Automatic disruption classification based on manifold learning for real-time applications on JET," *Nuclear Fusion*, vol. 53, no. 9, p. 093023, 2013.

[11] G. A. Rattá, J. Vega, A. Murari, M. Johnson, and JET-EFDA Contributors, "Feature extraction for improved disruption prediction analysis at JET," *Review of Scientific Instruments*, vol. 79, no. 10, p. 10F328, 2008.

[12] G. Farias, J. Vega, S. Dormido-Canto, A. Murari, R. Moreno, H. Vargas, and J. A. Valencia, "Prediction of the time to disruption in JET with an ITER-like wall," EFDA–JET–PR(14)08, 2014.

[13] B. Cannas, A. Fanni, P. Sonato, M. Zedda, and JET-EFDA contributors, "A prediction tool for real-time application in the disruption protection system at JET," *Nuclear Fusion*, vol. 47, no. 11, p. 1559, 2007.

[14] J. Vega, R. Moreno, A. Pereira, G. A. Rattá, A. Murari, S. Dormido-Canto, S. Esquembri, E. Barrera, and M. Ruiz, "Review of disruption predictors in nuclear fusion: Classical, from scratch and anomaly detection approaches," in *IECON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society*, 2016, pp. 6375–6379.

[15] C. Rea and R. S. Granetz, "Exploratory machine learning studies for disruption prediction using large databases on diii-d," *Fusion Science and Technology*, vol. 74, no. 1-2, pp. 89–100, 2018.

[16] L. Breiman, J. H. Friedman, C. J. Stone, and R. A. Olshen, *Classification and regression trees*. Chapman & Hall/CRC, 1984.

[17] J. M. López, J. Vega, D. Alves, S. Dormido-Canto, A. Murari, J. M. Ramírez, R. Felton, M. Ruiz, and G. de Arcas, "Implementation of the disruption predictor APODIS in JET's real-time network using the MARTe framework,"

[18] G. A. Rattá, J. Vega, A. Murari, G. Vagliasindi, M. F. Johnson, P. C. de Vries, and JET EFDA Contributors, "An advanced disruption predictor for JET tested in a simulated real-time environment," *Nuclear Fusion*, vol. 50, no. 2, p. 025005, 2010.

[19] R. Moreno, J. Vega, A. Murari, S. Dormido-Canto, J. M. López, J. M. Ramírez, and JET EFDA Contributors, "Robustness and increased time resolution of JET advanced predictor of disruptions," *Plasma Physics and Controlled Fusion*, vol. 56, no. 11, p. 114003, 2014.

[20] J. Vega, A. Murari, S. Dormido-Canto, R. Moreno, A. Pereira, A. Acero, and JET-EFDA Contributors, "Adaptive high learning rate probabilistic disruption predictors from scratch for the next generation of tokamaks," *Nuclear Fusion*, vol. 54, no. 12, p. 123001, 2014.

[21] J. Vega, A. Murari, S. Dormido-Canto, R. Moreno, A. Pereira, and S. Esquembri, "Real-time anomaly detection for disruption prediction: the JET case," EUROFUSION WPJET1-PR(16) 14855, 2016.

[22] A. Svyatkovskiy, J. Kates-Harbeck, and W. Tang, "Training distributed deep recurrent neural networks with mixed precision on GPU clusters," in *Proceedings of the Machine Learning on HPC Environments*, ser. MLHPC'17, 2017, pp. 10:1–10:8.

[23] W. M. Tang, "Deep learning acceleration of progress toward delivery of fusion energy," in *Theory & Simulation of Disruptions Workshop (TSDW-2017)*. PPPL, Jul. 2017.

[24] F. A. Matos, D. R. Ferreira, and P. J. Carvalho, "Deep learning for plasma tomography using the bolometer system at JET," *Fusion Engineering and Design*, vol. 114, no. Supplement C, pp. 18–25, 2017.

[25] A. Huber, K. McCormick, P. Andrew, P. Beaumont, S. Dalley, J. Fink, J. Fuchs, K. Fullard, W. Fundamenski, L. Ingesson, F. Mast, S. Jachmich, G. Matthews, P. Mertens, V. Philipps, R. Pitts, S. Sanders, and W. Zeidner, "Upgraded bolometer system on JET for improved radiation measurements," *Fusion Engineering and Design*, vol. 82, no. 5, pp. 1327–1334, 2007.

[26] A. Huber, K. McCormick, P. Andrew, M. de Baar, P. Beaumont, S. Dalley, J. Fink, J. Fuchs, K. Fullard, W. Fundamenski, L. Ingesson, G. Kirnev, P. Lomas, F. Mast, S. Jachmich, G. Matthews, P. Mertens, A. Meigs, V. Philipps, J. Rapp, G. Saibene, S. Sanders, R. Sartori, M. Stamp, and W. Zeidner, "Improved radiation measurements on JET – first results from an upgraded bolometer system," *Journal of Nuclear Materials*, vol. 363–365, no. Supplement C, pp. 365–370, 2007.

[27] K. McCormick, A. Huber, C. Ingesson, F. Mast, J. Fink, W. Zeidner, A. Guigon, and S. Sanders, "New bolometry cameras for the JET enhanced performance phase," *Fusion Engineering and Design*, vol. 74, no. 1, pp. 679–683, 2005.

[28] L. C. Ingesson, B. Alper, B. J. Peterson, and J.-C. Vallet, "Chapter 7: Tomography diagnostics: Bolometry and soft-x-ray detection," *Fusion Science and Technology*, vol. 53, no. 2, pp. 528–576, 2008.

[29] D. R. Ferreira, P. J. Carvalho, H. Fernandes, and JET Contributors, "Full-pulse tomographic reconstruction with deep neural networks," *Fusion Science and Technology*, vol. 74, no. 1-2, pp. 47–56, 2018.

[30] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *CoRR*, vol. abs/1412.6980, 2014.