Prediction of Radiative Collapse in the Large Helical Device Plasma Discharges using Convolutional Neural Networks

Yuya SUZUKI[1)], Mamoru SHOJI[1,2)], Naoki KENMOCHI[1,2)] and Masayuki YOKOYAMA[1,2)]

[1)]*The Graduate University for Advanced Studies, SOKENDAI, Gifu 509-5292, Japan*

[2)]*National Institute for Fusion Science, National Institutes of Natural Sciences, Gifu 509-5292, Japan*

(Received 8 May 2024 / Accepted 13 December 2024)

Predicting and preventing abrupt plasma termination incidents pose considerable challenges in nuclear fusion research. In the Large Helical Device (LHD), this occurrence is referred to as radiative collapse. During radiative collapse, impurity particles induce energy dissipation via radiation, hindering the maintenance of plasma discharges. Our approach aims to predict radiative collapse by analyzing the visible light emitted during such events. LHD uses approximately ten cameras to continuously observe plasma discharges, resulting in the accumulation of substantial video data from previous experiments. Using these images, convolutional neural network (CNN) models were trained to identify discharge states and subsequently applied to plasma discharge videos of the plasma discharges as a predictor. As a result, a determination model was developed, capable of discerning between stable and collapsed plasma discharge states with an accuracy of 91.5% ± 4% using plasma discharge images. Notably, this model demonstrated the potential to predict radiative collapse approximately three frames (66–132 ms) in advance. An examination of the model's focal points revealed consistency with findings from prior research.

Keywords: radiative collapse, deep learning, image recognition, prediction, LHD

DOI: 10.1585/pfr.20.1402021

1. Introduction

Abrupt termination of plasma discharges poses a critical challenge in nuclear fusion research, potentially leading to substantial damage to vacuum vessels and divertor tiles. One such termination event observed in the Large Helical Device (LHD) is known as "radiative collapse". This phenomenon is triggered by the influx of low-mass impurities into the plasma. These particles induce plasma cooling through intense radiation, hindering the maintenance of high-density plasmas. In LHD, the Sudo density limit represents the density threshold beyond which plasma performance is compromised. Notably, radiative collapse poses a considerable obstacle to achieving the Sudo density limit. Hence, predicting and preventing radiative collapse is crucial for sustaining plasma densities near this critical threshold. The Sudo density limit is mathematically expressed as follows [1].

$$n_e^{sudo}[10^{20}\mathrm{m}^{-3}] = 0.25 P^{0.5} B^{0.5} a^{-1} R^{-0.5}. \quad (1)$$

Here, n_e, P, B, a and R denote the electron density, the absorbed plasma heating power [MW], the magnetic field strength [T], the average minor radius [m], and the major radius [m], respectively. In practical applications, maintaining the density close to this limit can be challenging.

Previous research has elucidated the characteristics of radiative collapse [2], including its physical mechanism and the location of its precursors. Radiative collapse is induced by impurities such as carbon and oxygen, which emit radiation in the plasma. It typically manifests in the inner region of the torus and an ergodic region located near the divertor tiles.

Another study has demonstrated the efficacy of a data-driven approach in predicting and preventing radiative collapse [3]. This study identified four effective parameters—$\bar{n}_e$, CIV, OV and $T_{e,edge}$—through the application of machine-learning techniques. Here, $\bar{n}_e$ represents the line-averaged electron density near the magnetic axis, while CIV and OV denote the emission intensities of carbon and oxygen, respectively. $T_{e,edge}$ indicates the electron temperature in the edge region. By incorporating these parameters, the likelihood of radiative collapse can be predicted, thereby enabling its avoidance in actual plasma discharges in LHD.

This paper introduces a novel method for predicting radiative collapse events in plasma discharges by using images captured by visible light cameras and convolutional neural networks (CNNs), a prominent machine-learning technique. CNN models have also been used in the analysis of plasma simulations [4, 5]. Our method focuses on the strong light emission associated with radiative collapse. This study describes the process of creating datasets for model training, evaluates the model's accuracy and predictive performance, and identifies the regions of interest considered by the model when determining discharge states.

author's e-mail: suzuki.yuya@nifs.ac.jp

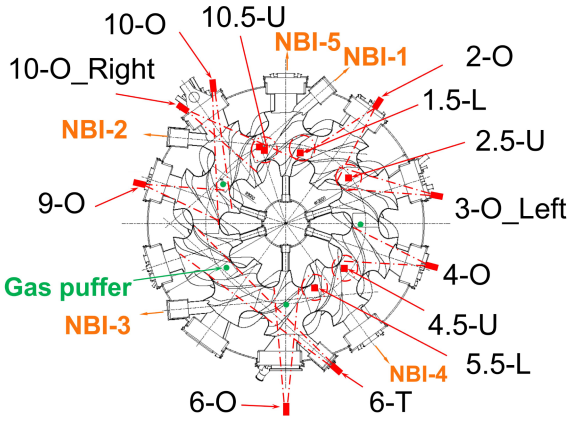

Fig. 1 Camera layout for the monitoring system in deuterium discharge experiments on LHD. The inner twisted circle represents the vacuum vessel, equipped with numerous horizontal and top-bottom viewports. Cameras for plasma monitoring are positioned to capture both central and tangential views. Additional cameras are installed to observe the divertor tiles and walls facing the plasma heating system.

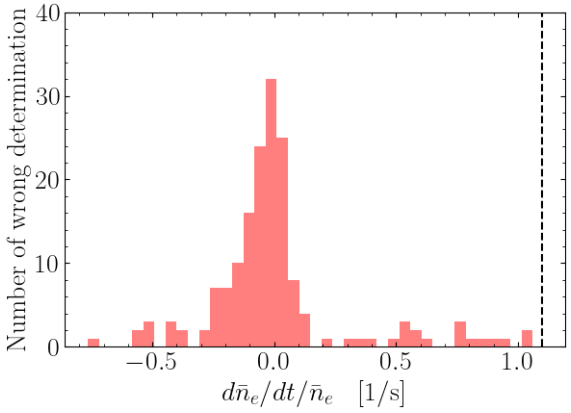

Fig. 2 Number of incorrect determinations on each $\dot{\bar{n}}_e/\bar{n}_e$. The dashed line shows $\dot{\bar{n}}_e/\bar{n}_e = 1.1$.

2. Plasma Monitoring System in LHD

LHD is equipped with a unique plasma monitoring system comprising approximately ten cameras positioned at viewports on the vacuum vessel, as illustrated in Fig. 1 [6, 7]. These cameras monitor various aspects of the plasma, including the divertor tiles, plasma heating system, and gas fueling. They use charge-coupled devices, which are highly sensitive to visible light and capable of capturing high-resolution images. However, the cameras are susceptible to damage caused by high-energy neutrons and gamma rays emitted during deuterium plasma discharges, requiring careful consideration in their placement and operation. To protect the camera and its associated equipment from harmful radioactive rays, they are housed in shielded enclosures positioned over 12 m from the center of the LHD. Images captured at the viewport are transmitted to the camera via optical fibers, minimizing the risk of radiation damage. This arrangement has proven highly effective in extending the camera's operational life, enabling it to reliably monitor the plasma, even during deuterium discharge experiments.

3. Dataset for Model Training and Evaluation

In our study, we treated the prediction of radiative collapse as a binary classification problem, distinguishing between stable and collapsed states. To train the model, we generated datasets from images of plasma discharges with radiative collapse. Notably, we used a key parameter known as the density exponent [2] to reliably detect the occurrence of radiative collapse.

$$\text{density exponent } x = (\dot{P}_{rad}/P_{rad})/(\dot{\bar{n}}_e/\bar{n}_e). \quad (2)$$

Here, P_{rad} and the dot notation denote the radiation power measured in megawatts (MW) and the time derivative of the each parameter, respectively. When the calculated value of x surpasses 3, the plasma is considered to have transitioned into a collapsed state. The efficacy of this threshold has been well-established in previous research [2, 3]. It is noteworthy that when $\bar{n}_e$ remains constant or exhibits a gradual increase, the derivatives of $\bar{n}_e$ are zero or very small values. In this case, the abovementioned equation diverges and becomes greater than 3. To avoid this incorrect decision, a threshold is required for $\bar{n}_e$. Figure 2 shows the number of incorrect determination. If the plasma did not collapse within 200 ms after the density exponent exceeded 3, this $\dot{\bar{n}}_e/\bar{n}_e$ was recorded as a false determination. This demonstrates that these errors can be eliminated by setting the threshold to $\dot{\bar{n}}_e/\bar{n}_e = 1.1$. When the density exponent exceeded the threshold, we used 0 instead.

In addition to the threshold, the following experimental conditions were chosen to automatically extract a typical plasma discharge with radiative collapse from a large volume of experimental data: B was 2.75 T, R_{ax} was 3.6 m, the fueling gas was hydrogen, both ECH and NBI were injected, and W_p was greater than 100 kJ. W_p is the amount of energy stored in the plasma. Using the threshold and the experimental conditions, plasma discharges with typical plasma discharge with radiative collapse were extracted, as shown in Fig. 3. The period from the moment the density exponent surpassed 3 until the end of the plasma discharge was considered to be in a collapsed state. The period from the time excluding start-up time to 0.15 s before the density exponent surpasses 3 was considered to be in a stable state. Determining whether a state is stable or collapsed just before a collapsed state is difficult. As a result, this transition zone was not used as a dataset for learning in this study. Images extracted from the collapsed state were classified as "collapse", while those extracted from the stable state were classified as "stable".

This approach was applied to plasma discharges with

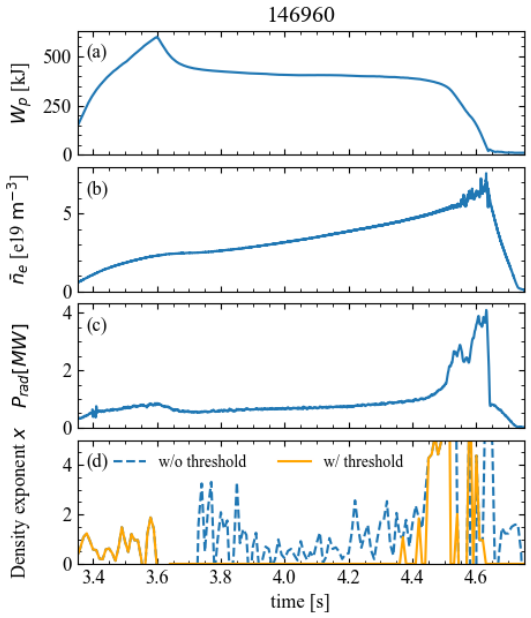

Fig. 3 Example of a plasma discharge with typical radiative collapse. In this discharge, the radiative collapse occurred at approximately 4.45 s. The blue dashed line in (d) depicts the density exponent calculated without a threshold on $\dot{\bar{n}}_e/\bar{n}_e$. The density exponent with a threshold is represented by the orange line.

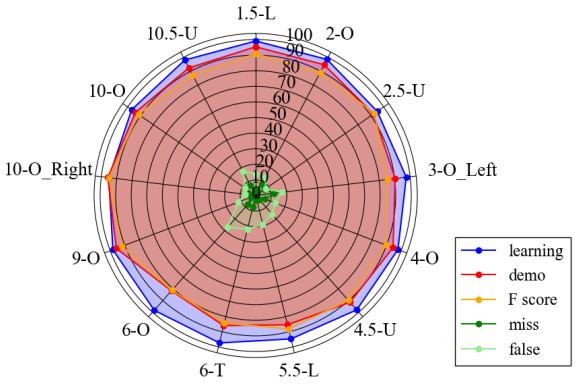

Fig. 4 Accuracy validation results for each model. The names of the models that were trained using images from the port with that name are displayed as the outer labels. The "Learning" and "Demo" scores represent the accuracy of the training and evaluation phases, respectively. The "F Score" indicates the accuracy of the evaluation data and considers the ratio of stable to collapsed state images. The "Miss" and "False" scores represent the proportion of times the model mistook a collapsed-state image for a stable-state image and vice versa.

shot numbers ranging from 157,001 to 162,000 in this work. As a result, 42 plasma discharges that ended with a typical radiative collapse were extracted. The videos of the 42 plasma discharges were divided into "learning data" and "evaluation data" in a 7:3 ratio. The evaluation data was not used in the learning process. Approximately ten stable-state images and four collapsed-state images were extracted from each video. Approximately 600 images were obtained on each camera for the entire learning dataset, consisting of 450 stable state images and 150 collapsed state images.

4. Evaluation of the Accuracy of Trained Determination Models

The models were trained using images of the plasma discharges. The model was constructed based on "EfficientNet-b4" [8]. EfficientNet is a CNN model that has been optimized for computational efficiency and learning capability. "b4" indicates the model's size. EfficientNet includes versions with different sizes, ranging from "b0" to "b7". EfficientNet-b4 has a medium level of computational cost and learning ability. We used the softmax function to normalize the model's outputs. The softmax function normalizes the sum of the outputs from 0 to 1. We developed a model for each of the 13 cameras. Each model was trained only using images from a single camera. We changed the weights of all blocks while keeping the EfficientNet-b4 structure. Batch normalization was used in this study; however, data augmentation techniques were not used. Figure 4 shows the accuracy of each model's determination. In this figure, "Learning" and "Demo" represent the accuracy when the learning and evaluation datasets were used, respectively. "F score" is the accuracy, taking into account the ratio of stable and collapsed-state images, as we will explain later. "Miss" is the proportion of collapsed-state images that the model identified as stable-state images. "False" is the proportion of stable-state images that the model identified as collapsed-state images. The F score is determined using the following formula.

$$Recall : \frac{TP}{TP + FN}, \tag{3}$$

$$Precision : \frac{TP}{TP + FP}, \tag{4}$$

$$F = \frac{2 \times Recall \times Precision}{Recall + Precision}. \tag{5}$$

Here, TP (True Positive) is the score that the model correctly identified collapsed-state images, FN (False Negative) is the score that the model incorrectly identified stable-state images, and FP (False Positive) is the score that the model incorrectly identified collapsed-state images.

Except for the model trained on images from the camera installed in the 6-O port, the models achieved near-perfect accuracy on the "Learning" dataset. These models also demonstrate an accuracy of 91.5% ± 4 % on the "Demo" dataset.

Figure 5 shows that the model trained on images from the camera installed in the 6-O port had the lowest accuracy, whereas the model trained on images from the camera

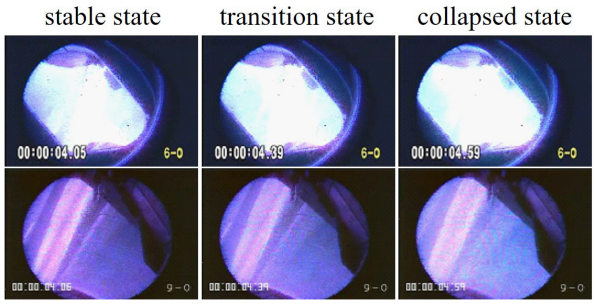

Fig. 5 The top and bottom images were captured by cameras in the 6-O and 9-O ports, respectively. The left, middle, and right images represent the stable, transition, and collapsed states, respectively. The plasma images from the 6-O port exhibit greater brightness in most regions when compared to those from the 9-O port.

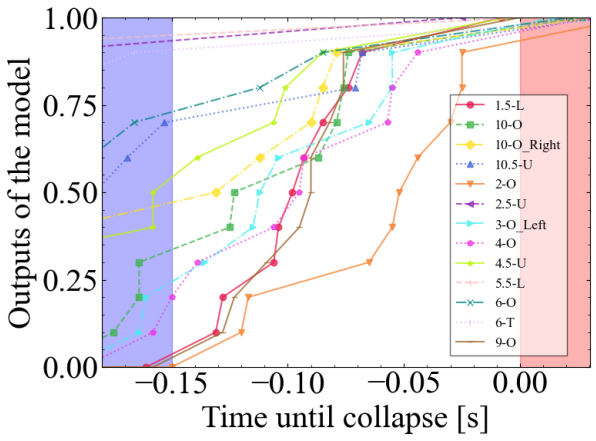

Fig. 6 Prediction performance results for the models. The model outputs were calculated by averaging the outputs across evaluation datasets. The time intervals designated as stable and collapsed states in the labeling process are represented by the blue and red hatched areas, respectively. A model that generates a high value during the time period represented by the blue-hatched area is not suitable as a predictor because its output is too premature to be used for an alarm.

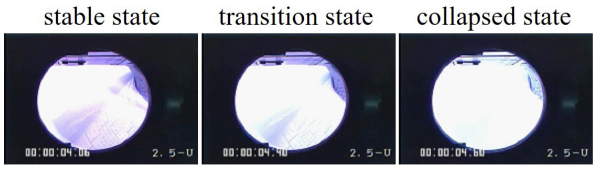

Fig. 7 Images captured by the camera at the 2.5-U port. The left, middle, and right images correspond to the stable, transition, and collapsed states, respectively. Compared to the 9-O port, the 2.5-U port images exhibit higher brightness across most areas.

installed in the 9-O port had a high score. A comparison of the input images from the cameras in the 6-O and 9-O ports suggests a possible explanation for the lower accuracy of the 6-O port. Figure 5 shows an example of input images in stable state, transition state, and collapsed state from left to right. The majority of the regions in the images from the 6-O port were extremely bright, indicating that the brightness in those regions was saturated. Reducing the sensitivity of this camera could improve accuracy. From a different perspective, hydrogen/deuterium gases were fueled from the 5.5 port, which is near the 6-O port, in the majority of the plasma discharges used in this study. This result suggests that observing plasma images from a location distant from the gas fueling point may be beneficial for predicting radiative collapse.

The higher accuracy of the model in 5.5-L compared to the one in 6-O can be attributed to the higher resolution of the camera at the 5.5-L port. Furthermore, the camera at 5.5-L can still observe the plasma-facing wall in its field of view, even when the emission intensity is high. From this vantage point, the camera at 5.5-L can observe the shape of the plasma as it begins to exhibit high emission intensity.

5. Performance as Predictor

Determination models were used as a radiative collapse predictor on 10 videos of plasma discharges in the evaluation dataset. Frames in the evaluation video were not used in learning. The video images were captured at a rate of 30 fps. Figure 6 shows the prediction performance results for the models. The time to collapse is represented on the horizontal axis. The model's output is represented on the vertical axis. The model's output indicates how close the input image is to a collapsed state. The value represents the output of the model used for classification. If this value exceeds 0.5, the input image is classified as a collapsed state; if it is below 0.5, it is classified as a stable state. In Fig. 6, the blue-hatched area represents a time interval designated as a stable state during the labeling process. The absence of any symbol for some models indicates that the outputs of these models surpassed 0.9 before −0.18 s. Models indicating a collapsed state during this time period were deemed to have made a premature determination. Certain determination models could predict 2–4 frames (66–132 ms) in advance. The plasma's energy confinement time was approximately 100 ms, determined from the ratio of stored energy to deposited energy within the experimental scope of this research. If this energy confinement time is regarded as the margin for controlling the plasma, the determination models can serve as predictors of radiative collapse by applying an appropriate threshold. The optimal threshold has not been finalized because it will be determined considering the response times of components such as cameras, heating devices, fuel injectors, computers, and the plasma itself.

Comparing the input images from the 2.5-U camera, which yielded a premature determination, to those from

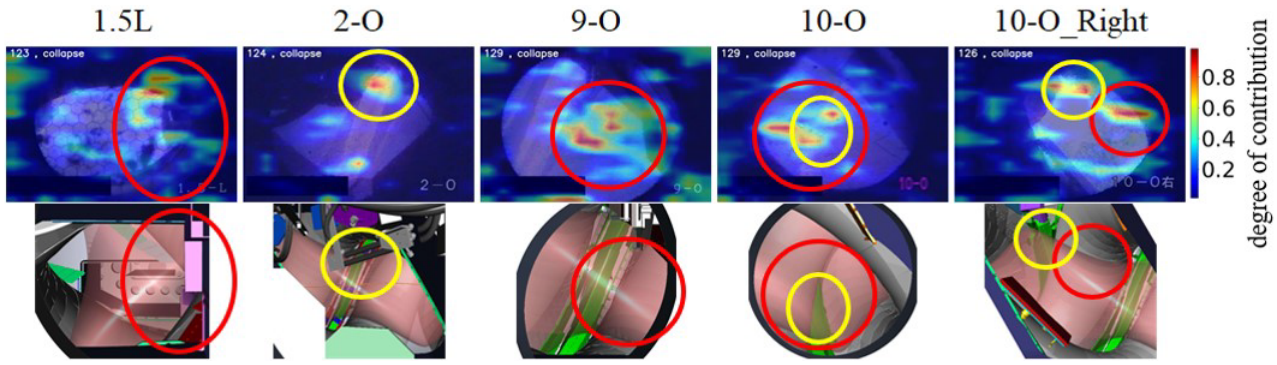

Fig. 8 Results of the focus area analysis using GradCAM. The top heatmaps represent the GradCAM results, while the bottom figures show the expected field of view as determined by CAD. High contributions to model outputs are indicated by the colored areas. The regions in the red circle indicate that the model focused on luminescence from the torus' inboard side. The locations inside the yellow circle represent luminescence from the divertor tiles.

the 9-O camera, which yielded an accurate prediction, indicates that brightness saturation was the cause of the premature determination. Figure 7 shows an example of input images from a 2.5-U port camera in the stable, transition, and collapsed states, displayed from left to right. The images from the 2.5-U camera, which resulted in a premature determination, demonstrate that the brightness in most areas was saturated, similar to the images from the 6-O camera. The accuracy decreased because the areas available for determination were reduced owing to the saturation of brightness.

The low determination accuracy in 6-O and premature determination in 2.5-U were both attributed to brightness saturation. The difference in resolution and field of view between the cameras used in 2.5-U and 6-O accounts for the variation in determination accuracy and prediction performance. Images in 2.5-U have a higher resolution than those in 6-O. In 2.5-U images, the plasma shape is clearly visible in the stable state, but it undergoes significant changes in the collapsed state, with the shape disappearing. This contrasts with 6-O images, where the emitting area's shape remains relatively unchanged during the transition from the stable to the collapsed state. The higher resolution of 2.5-U images, combined with the pronounced changes in the emission area, likely contributed to greater determination accuracy. Conversely, the minimal changes in 6-O images during the transition might be attributed to finite outputs in the transition area.

6. Analysis of the Focus Area

Analyzing the regions of interest using gradient-weighted class activation mapping (GradCAM) [9] revealed the areas that influenced the model's outputs, as shown in Fig. 8. Through the convolutional process, CNNs generate multiple feature maps, which occur before the fully connected layer, and the input image is compressed by square matrices of integers called the kernel, such as 3×3 or 5×5. The convolutional process is repeated multiple times, and the arrays generated at each step are referred to as feature maps. Finally, in the model used in this study, the input image's height and width are reduced by approximately 1/30 following the convolutional process. The depth becomes approximately 1,800, which is the result of multiplying the number of color layers by the number of kernels in each layer. The GradCAM calculates the relationship between each feature map and the output using the backpropagation method. It generates a heatmap that indicates which regions of the input image contribute considerably to the output by weighting and summing feature maps. The red-colored areas in Fig. 8 represent regions with high contributions. This technique is widely used to visualize and understand the focus areas of CNNs for particular outputs. In earlier research characterizing radiative collapse [2], radiation from plasma on the inboard side of the torus and radiation near the divertor tiles were mentioned as characteristic features. In Fig. 8, the regions highlighted by red and yellow circles represent the above-mentioned characteristic areas, respectively. This result demonstrates that the determination model uses relevant image regions to determine plasma discharge states. Furthermore, the plasma regions are consistent with previous research [2].

7. Conclusion

This study demonstrated the prediction of radiative collapse using images of plasma discharges, employing CNNs, a type of deep learning model. Several trained models were used to determine discharge states (stable or collapsed) with 91.5% ± 4 % accuracy. Some determination models predicted the collapsed states 2–4 frames (66–132 ms) in advance. Analysis of the regions of interest using GradCAM revealed that the models' focus areas aligned with previous research because they encompassed radiation from the plasma on the inboard side of the torus

and near the divertor tiles.

Future work will focus on enhancing the model and controller design to demonstrate the prediction and avoidance of radiative collapse in plasma discharge experiments. To expedite the model's determination time, we intend to optimize computational efficiency and reduce model size by examining neurons and connections that have a negligible impact on accuracy. Using images captured by multiple cameras simultaneously, images captured in sequential frames, or subtracted images obtained from sequential frames, we aim to enhance accuracy and prediction performance as well as elucidate focus areas.

Acknowledgments

We thank Dr. T. Yokoyama at QST for providing the research data that was used for previous papers. The raw data were generated at the LHD facility. The automatic Integrated Data Analysis software and the analyzed data are available from the LHD data repository located at https://www-lhd.nifs.ac.jp/pub/Repository_en.html.

[1] S. Sudo *et al.*, Nucl. Fusion **30**(1), 11 (1990).
[2] B.J. Peterson *et al.*, Plasma Fusion Res. **1**, 045 (2006).
[3] T. Yokoyama *et al*., Plasma Fusion Res. **17**, 2402042 (2022).
[4] E. Narita *et al.*, Nucl. Fusion **62**(8), 086037 (2022).
[5] J. Kates-Harbeck *et al.*, Nature **568**, 526 (2019).
[6] M. Shoji *et al*., Plasma Fusion Res. **3**, 440 (2000).
[7] M. Shoji *et al*., Plasma Fusion Res. **15**, 2402039 (2020).
[8] M. Tan and Q.V. Le, arXiv:1905.11946 (2019).
[9] R.R. Selvaraju *et al.*, arXiv:1610.02391 (2016).