

# Machine Learning Approaches in Plasma State Recognition: An Overview of the STARE

Wednesday, 29 November 2023 13:30 (20 minutes)

During the WEST experimental fusion plasma discharges, several diagnostics are employed to collect diverse data. Among these diagnostics, two high-definition cameras, operating within the visible spectrum, broadcast in real time the plasma inside the vacuum vessel.

The intent of our study is to investigate the possibility of determining the plasma state from this live video for each discharge. According to the characteristics of the WEST tokamak, five potential plasma states have been identified for this first application: current ramp up in limited configuration, diverted lower single null, upper single null, double single null, and no plasma.

However, real time detection of these states is a complex task. Artificial intelligence (AI) vision techniques, including deep learning algorithms, present a potential solution by enabling the analysis of images and identification of relevant patterns. In this context, the STARE (STate plasma REcognition) project aims to develop an automated tool that identifies the different plasma states from visible camera video, where the critical attribute for identification is the position of the plasma's contact point against the vessel walls. In this project, we plan to apply a data-driven approach, where machine learning models will be trained on Tokamak discharge images extracted from camera videos. The main objective is then to train these models to recognize the relationships between visual patterns and contact point locations, i.e. plasma state, and use them to classify the current state of the plasma from the live video feed.

The challenge in training classification models lies in providing proper annotated data. In particular within this project: for each plasma discharge, several hundreds of images are generated and need to be annotated. As a first step, we used an unsupervised learning technique to develop an automatic labeling tool to overcome this challenge and facilitate the generation and validation of the potential labels.

The data used in this study were collected from a variety of plasma discharges: videos from twelve experimental discharges were collected and go into several preprocessing steps for features extraction. In order to understand and classify the complex states of plasma during a discharge experiment in the WEST tokamak, we turned to using the K-means based approach to categorize our video frames dataset into distinct clusters based on their visual characteristics and similarities.

The K-means approach carries an inherent limitation: the necessity to predefine the number of clusters(K). To validate and guide the choice of K for our specific application, we employed several metrics such as silhouette score, Calinski-Harabasz index, and the elbow method.

The clustering results provided an initial exploration of the data that allows building a preliminary labeled dataset for training. The results were visualized using a homemade tool that allows the researchers to automatically examine the clustering groups. However, these initial clusters are not optimal, and the next steps of the project aim to refine the results by incorporating domain expertise and WEST diagnostics information for a better feature selection. Also, we aim to apply active learning and exploring supervised classification methods.

## Speaker's Affiliation

CEA, IRFM, F-13108 Saint-Paul-lez-Durance, France

## Member State or IGO/NGO

France

**Primary authors:** Mr MIDOU, Dorian (CEA, IRFM, F-13108 Saint-Paul-lez-Durance, France); Mrs ALMUHISEN, Feda (cea cadarache)

**Co-author:** Mr FEDORCZAK, Nicolas (CEA, IRFM, F-13108 Saint-Paul-lez-Durance, France)

**Presenters:** Mr MIDOU, Dorian (CEA, IRFM, F-13108 Saint-Paul-lez-Durance, France); Mrs ALMUHISEN, Feda (cea cadarache)

**Session Classification:** Posters Session

**Track Classification:** AI