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

F. ALMUHISEN[1], D. MIDOU[1], N. FEDORCZAK[1], L.FEVRE [2]
1 CEA, IRFM, F-13108 Saint Paul-lez-Durance, France.
2 Université de Lorraine, Institut Jean Lamour, CNRS, Nancy, 54000, France

Introduction

WEST tokamak experimental fusion plasma discharges rely on a variety of diagnostic tools to study plasma behavior. Among these, **two high-definition cameras** are deployed to capture and monitor the **plasma's activity** within the vacuum vessel in real-time. [1]

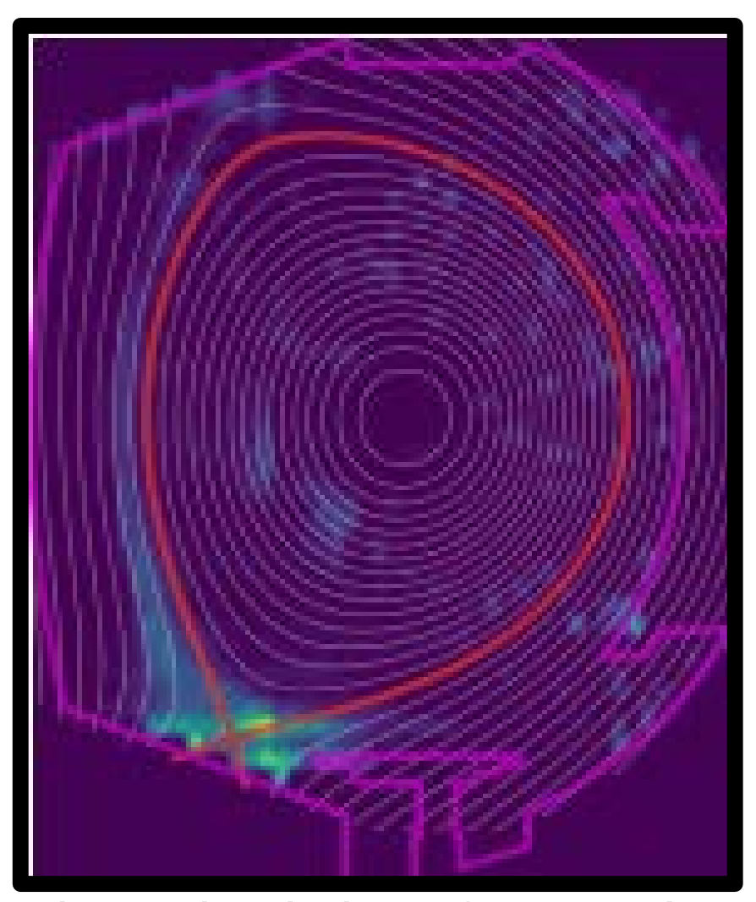


Fig 1: simulation of magnetic configuration of the plasma

Computational tools primarily based on magnetic configurations fig1, have been developed to **detect certain plasma states**, but they don't capture the full states desired by researchers in a rapid and real-time manner. Being able to do this quickly and accurately is very important for both the success and safety of the fusion operations.

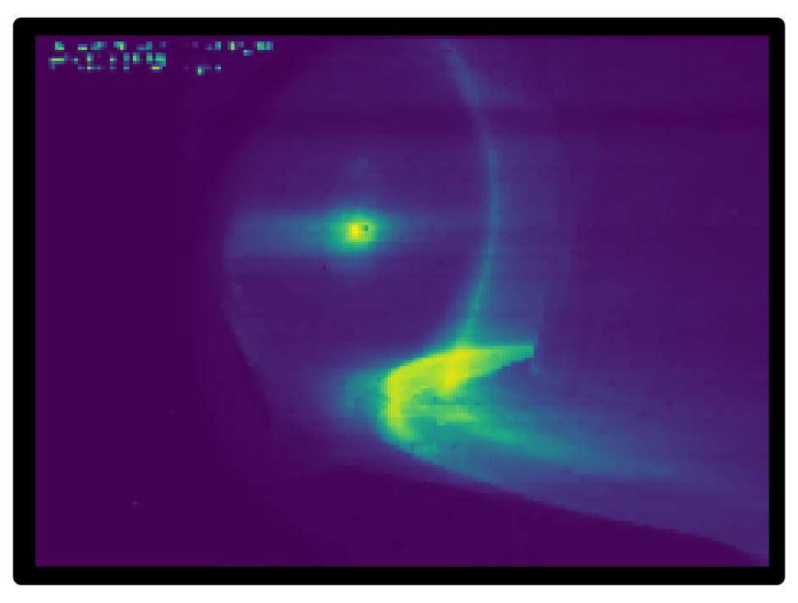


Fig 2: Plasma from visible camera

AI as Potential Solutions

Traditional methods of plasma state detection are limited in their ability to cope with the complexity and speed required for real-time analysis. This limitation opens up opportunities for applying **advanced AI techniques**, particularly **deep learning algorithms**, to analyze the rich visual data captured by cameras.

- We aim to develop an automated tool based on machine learning and AI techniques to identify the different plasma states from visible camera video, focusing on the position of the plasma's contact point against the vessel walls.
- Deploy out tools to our platform IA as service [2]

Plasma states

In the context of the WEST tokamak, **5 distinct plasma states** have been identified for this study: **current ramp-up** in limited configuration, **diverted lower single null**, **upper single null**, **double single null**, and **no plasma**. Each state presents unique characteristics and challenges for detection and analysis.

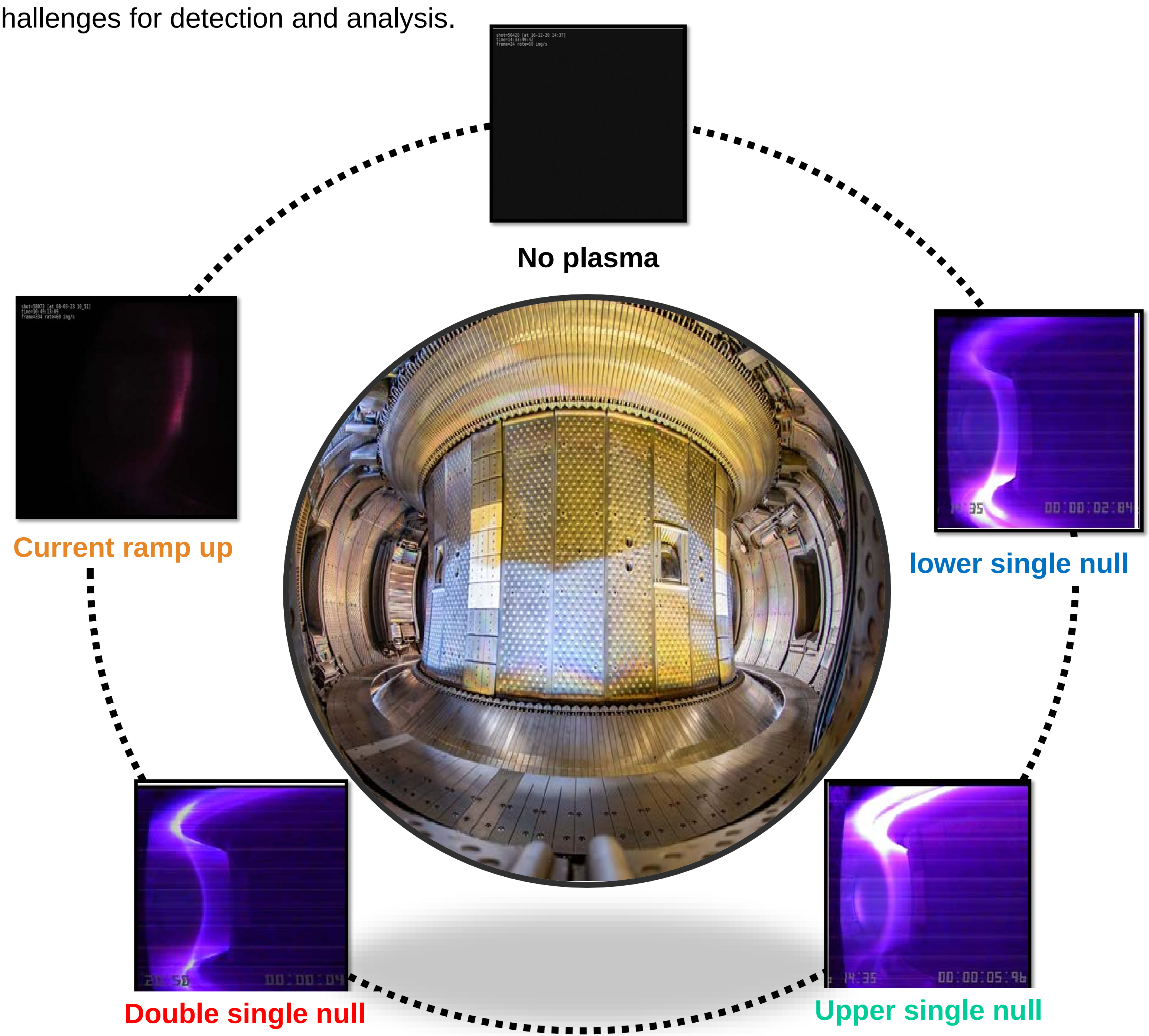


Fig 3:Example of different plasma states

The STARE Project: (STAte plasma REcognition)

Phase 1
Building labeling Tool using Unsupervised learning method

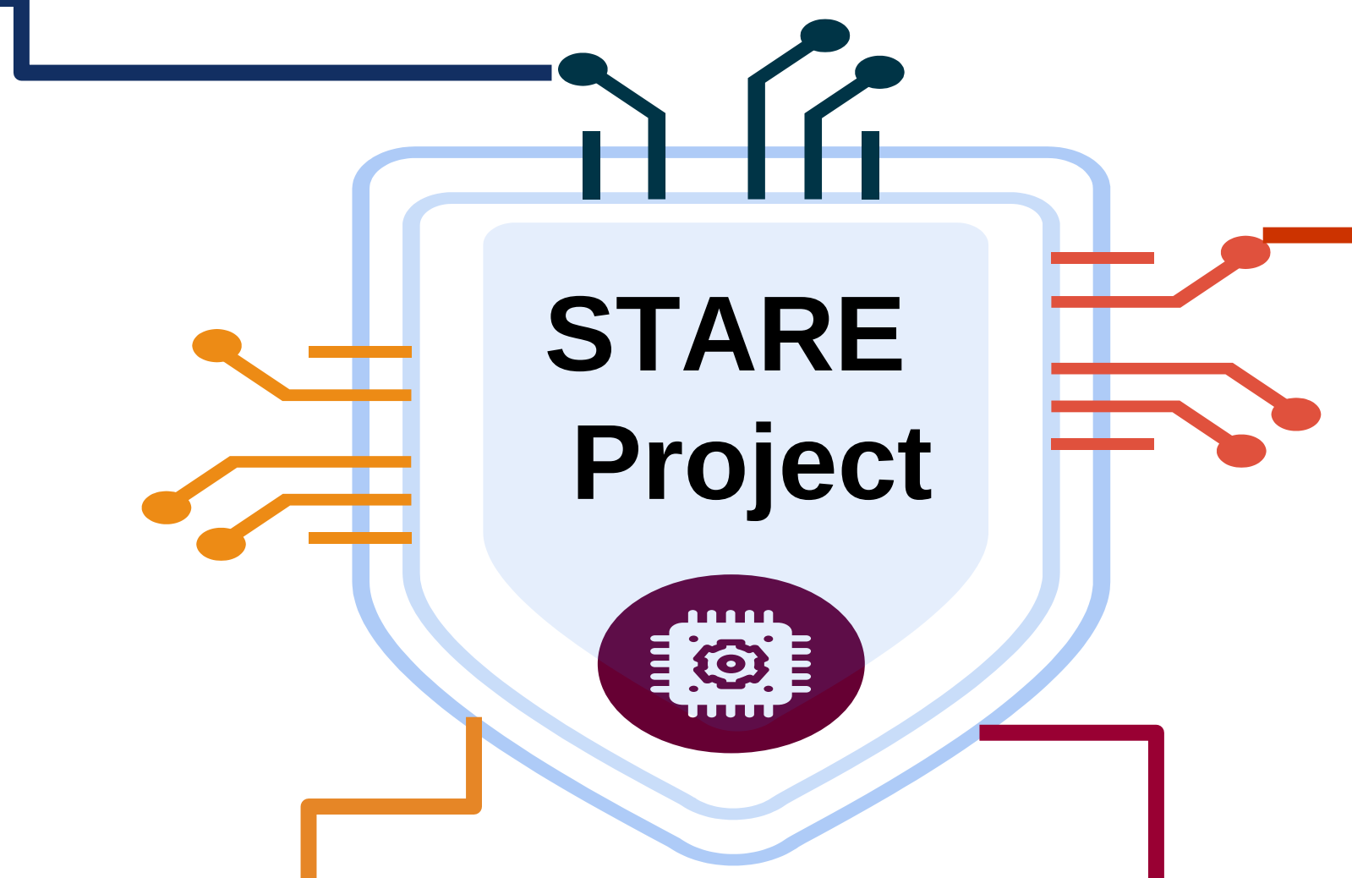
Image Grouping Based on Visual Similarity:
unsupervised Machine learning algorithm analyzes plasma images, grouping them based on visual characteristic.
We used K-means algorithm : a partitioning algorithm its goal is to minimize the within-cluster sum of squares (WCSS), which is the sum of squared distances between each data point and its corresponding centroid. Mathematically, this can be expressed as minimizing the objective function:

$$J=\sum_{i=1}^{k}\sum_{x\in Si}\|x-\mu\_i\|^2$$

where Si is the set of data points in cluster i, and μi is the centroid of cluster i.
Challenges and Solutions : Main challenge involves determining the optimal number of clusters. **Silhouette score** and **elbow method** are employed to determine the right number of clusters for the images.

[================ 80%]

Project Road map



Phase 2
Building dataset

Creating Preliminary Dataset Represent Plasma States:
Result is a set of grouped images each image is annotated by a potential plasma state and its magnetic diagnostic.
c. The dataset serves as the foundation for in-depth analysis, aiding in understanding various visual patterns of plasma behavior in the tokamak.

[==== 20%]

Phase 3
Building Classifier

Supervised learning
(deep learning Methods)

Phase 4
Deployment

Real time classification

Clustering results

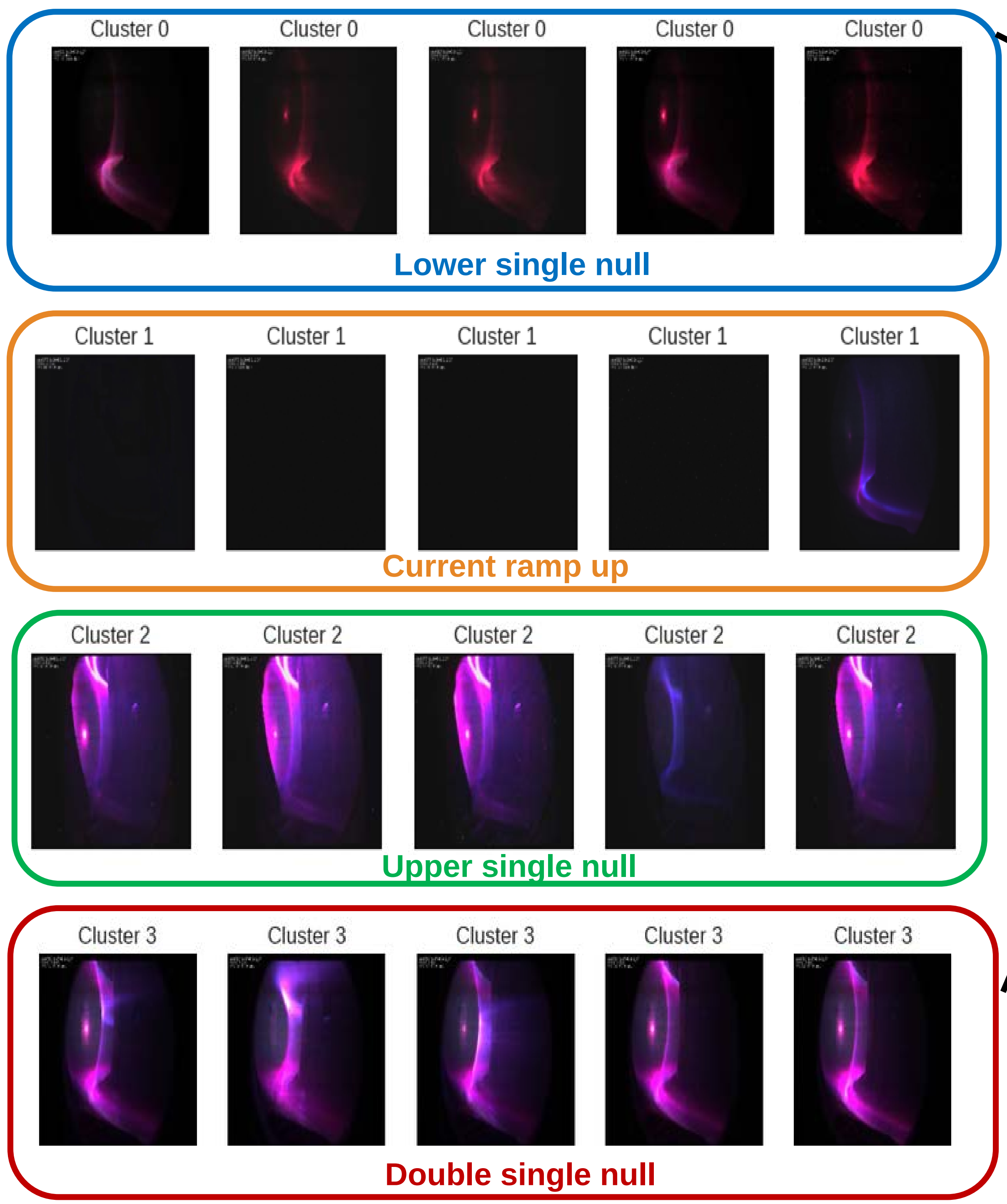


Fig4 : Clustering results for k =4

Preliminary Findings

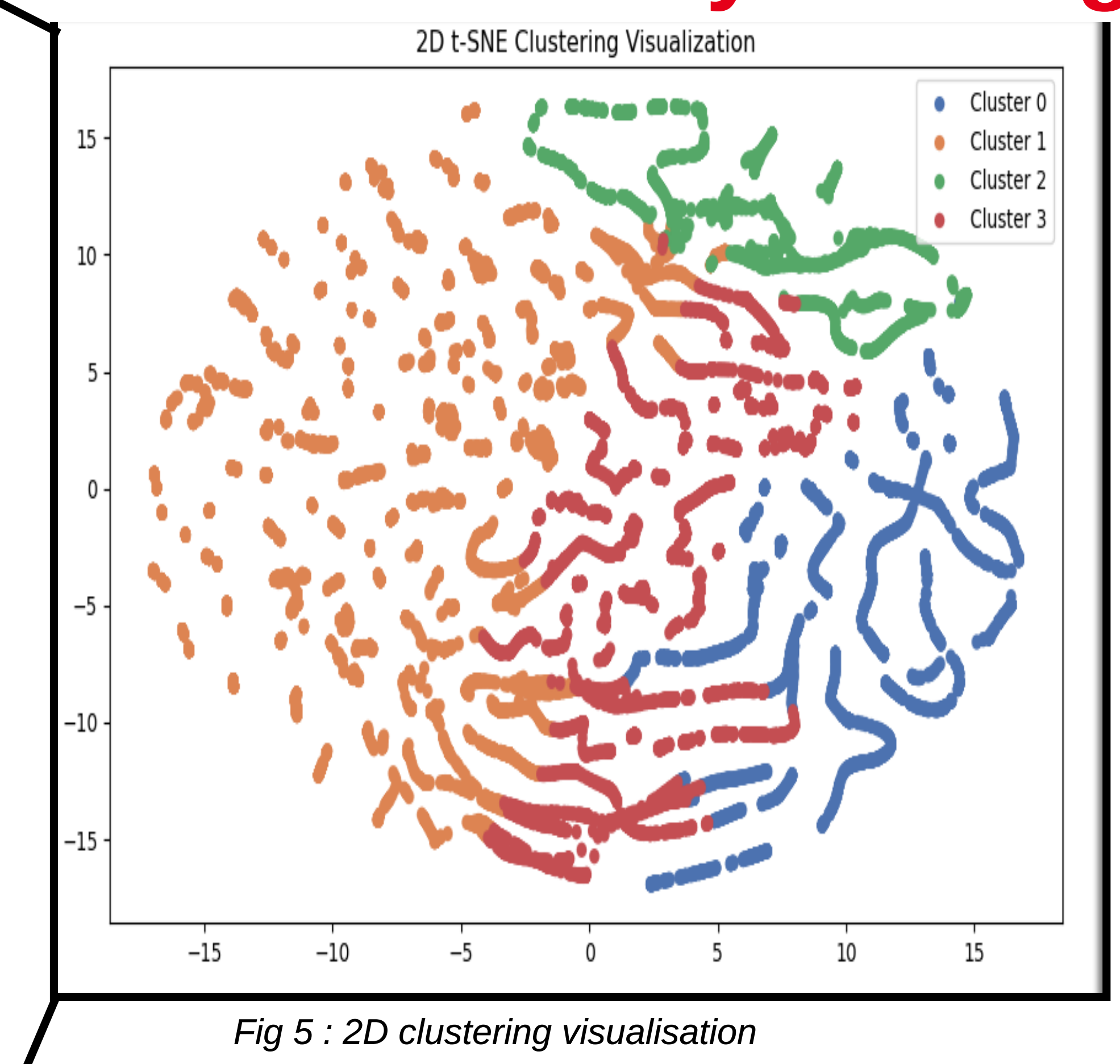


Fig 5 : 2D clustering visualisation

- Fig 5 shows a t-distributed Stochastic Neighbor Embedding (2D t-SNE) ML dimensionality reduction algorithm visualization for clusters
- T-SNE found structures within the data and distinct clusters, labeled from Cluster 0 to Cluster 3
- Some cluster are denser than other (see cluster1)

Validation

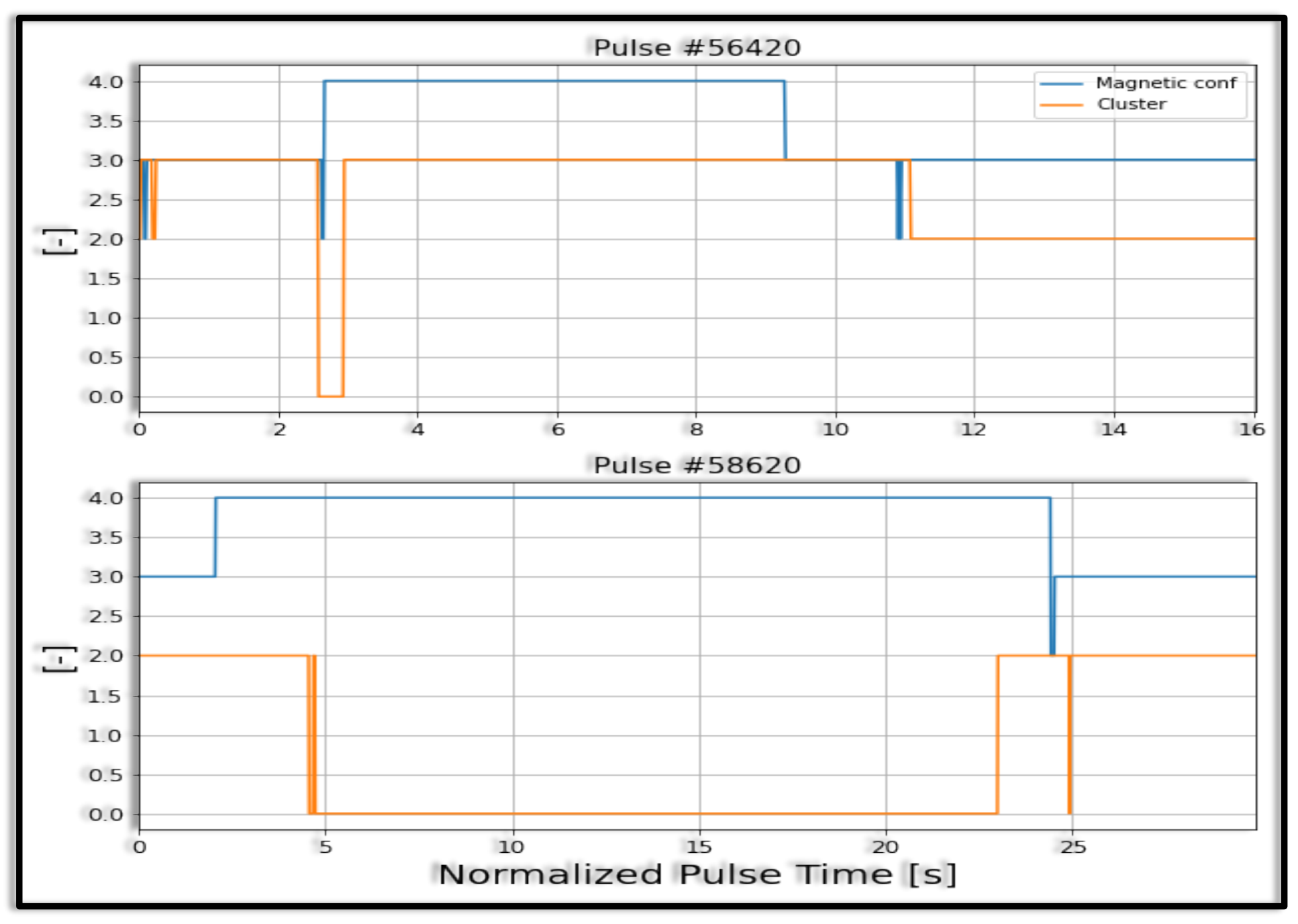


Fig 6: Method evaluation regarding diagnostic

Fig 6 shows the evaluation in between magnetic configuration and the clustering results, For Pulse #56420 there's a change in both the magnetic configuration and the cluster at approximately the same time (around the 2-second mark)Then, between 4 and 12 seconds, the magnetic configuration changes back and forth, which is not reflected in the cluster state. This could indicate that the magnetic configuration is more sensitive to certain conditions,In Pulse #58620, there are fewer changes. Both the magnetic configuration and the cluster change states at around the 5-second and remain in that new state for the duration of the observed time. The clustering algorithm is accurately reflecting changes in the magnetic configuration

Conclusion and Future Work

The K-means clustering, has set an encouraging results for identifying different plasma states in the WEST Tokamak. This approach has grouped images from camera videos into distinct clusters, each corresponding to a unique plasma state. While the current results are promising, we should highlight the need for expert evaluation to ensure the accuracy and relevance of the results.

Future Work:

- Improving the Labeling Tool: implement self-supervised learning techniques.
- Develop a classifier leverage the labeled dataset generated in the second phase to recognize and classify plasma states as they occur during tokamak operations.[3]

References

1. Bucalossi *et al* 2022 *Nucl. Fusion* **62** 042007
2. Almuhisen et al 2023, Optimizing tokamak operations using Machine learning methods as a service, 5th Technical Meeting on Fusion Data Processing, Validation and Analysis, IAEA, Jun 2023,Gand,Belgium.
3. M. Falato et al 2022 Journal of Plasma Physics, 88(6), 895880603