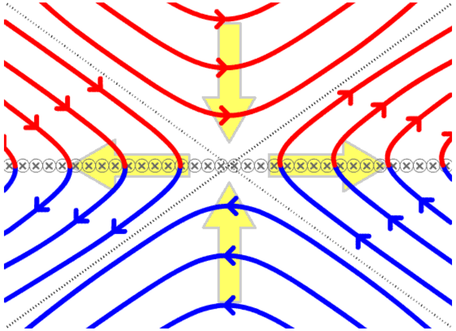
Automatic identification of magnetic reconnection events in 2D Hybrid Vlasov Maxwell simulations using Convolutional Neural Networks

A. Hu[1], M. Sisti[2,3], F. Finelli[2], F. Califano[2], J. Dargent[2], M.Faganello[3], E. Camporeale[1,4,5], J. Teunissen[1]

[1]Centrum Wiskunde & Informatica, Amsterdam, The Netherlands
[2]Dipartimento di Fisica, Universita di Pisa, Italy, EU
[3]Aix-Marseille University, CNRS, PIIM UMR 7345, Marseille, France, EU
[4]CIRES, University of Colorado, Boulder, CO, USA
[5]NOAA Space Weather Prediction Center, Boulder, CO, USA



This research is developed in the framework of the European project AIDA (Artificial Intelligence Data Analysis)

The AIDA project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 776262.

Significance

- Magnetic reconnection is a fundamental process in space and laboratory plasmas in which magnetic energy is converted into kinetic energy, released in the form of accelerated particles, flows and heating. Although the process itself is highly localized, it eventually leads to a global change of the magnetic field topology.

The AIDA project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 776262.

Goals

Now

Recognize reconnection in 2D simulations

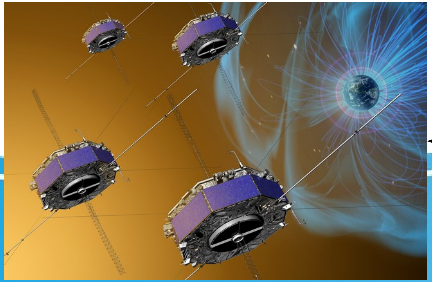
Future

'Virtual satellites'

3D simulations

1D time series

MMS data



The AIDA project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 776262.

Simulations performed at CINECA on Marconi

- 2D Hybrid Vlasov-Maxwell model
  - Ions: Vlasov (distribution function not yet used)
  - Electrons: fluid

Sim 1

- Resolution: 3072x3072 x 51^3
- $\Delta L/d_i : 0.1$
- $N_{samples} = 2024$
- Cost: ~5Mh core hours
- Memory: 10Tb

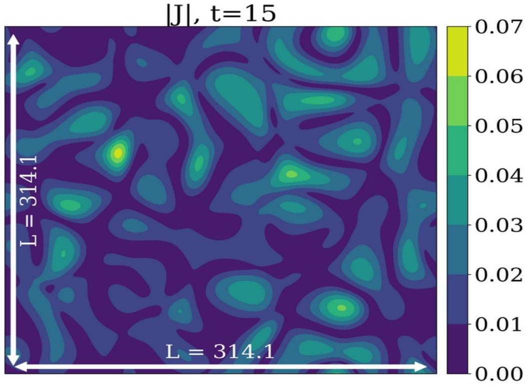
Sim 2

- Resolution: 2048x2048x51^3
- $\Delta L/d_i : 0.15$
- $N_{samples} = 124$
- Cost: ~1Mh core hours
- Memory: 4Tb

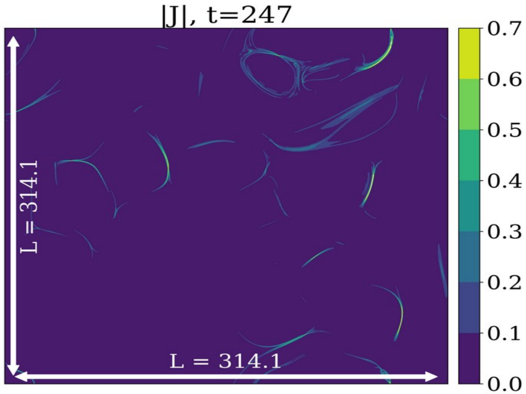
The AIDA project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 776262.

Current sheet temporal changes

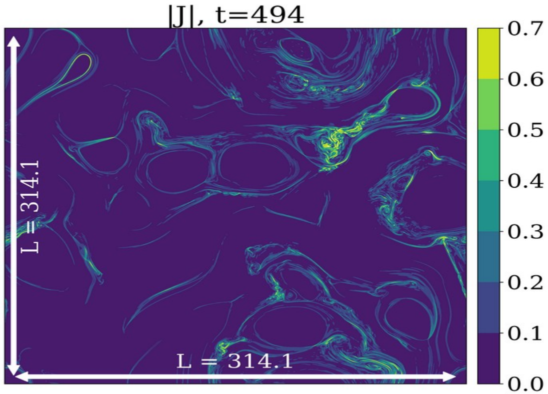
Initial set up



CS formation phase



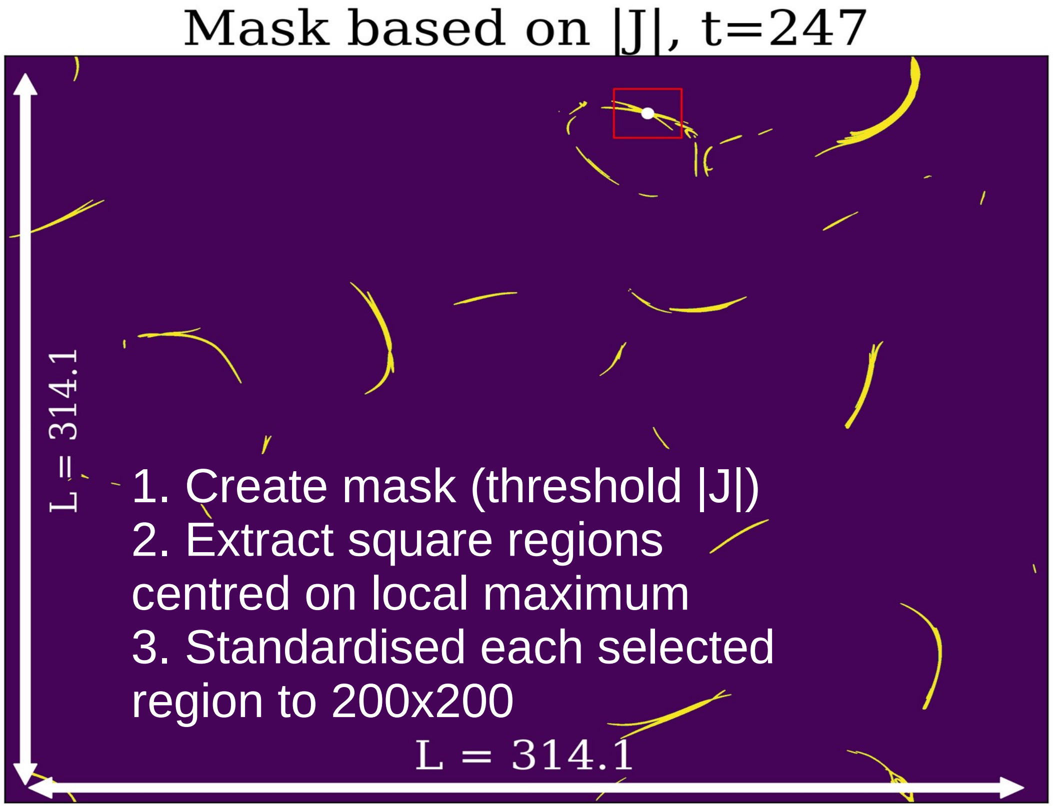
Fully turbulent phase



How the current sheet changes with the time passing by in Sim 1.

The AIDA project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 776262.

Current Sheet and Selected Variables

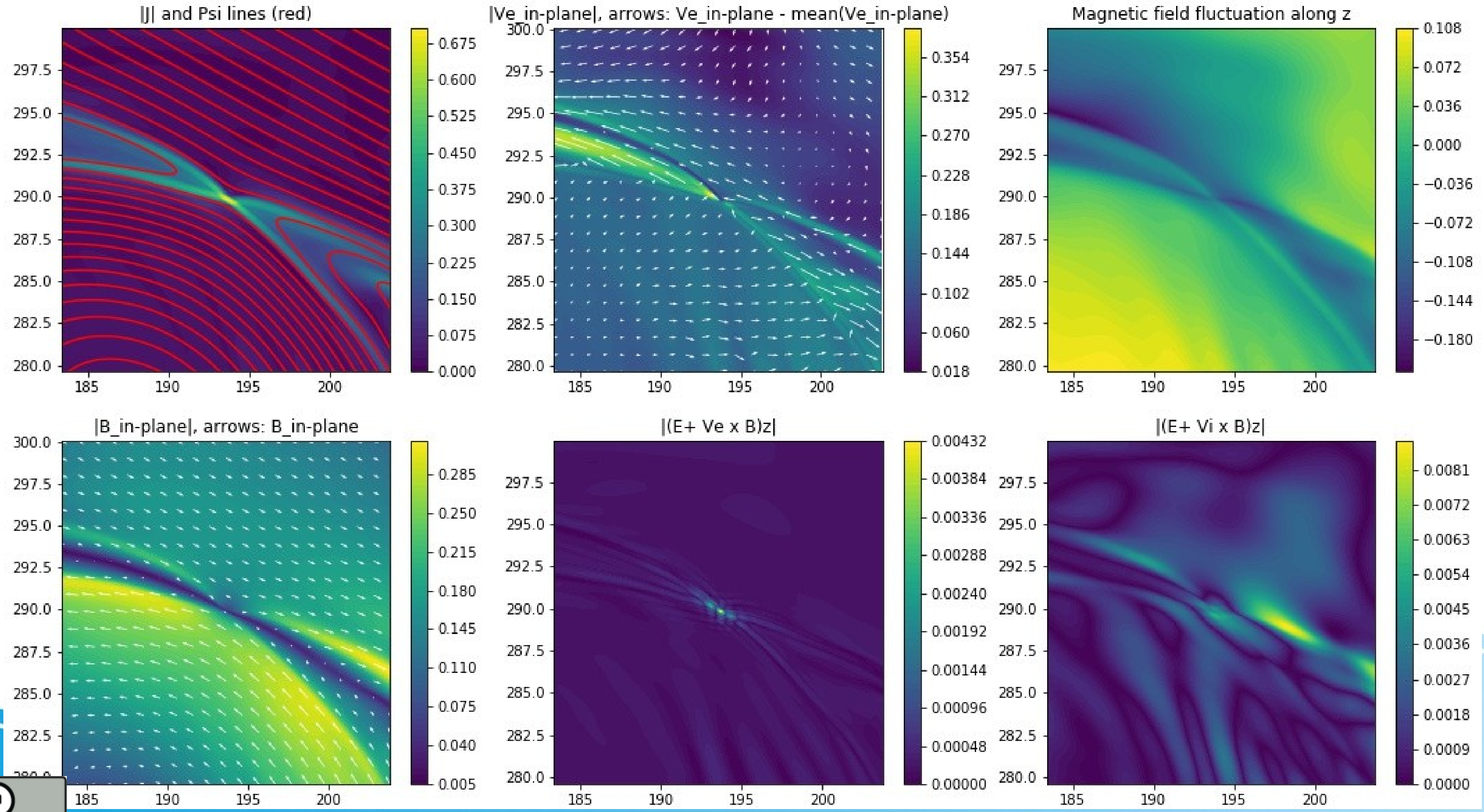


1. Create mask (threshold |J|)
2. Extract square regions centred on local maximum
3. Standardised each selected region to 200x200

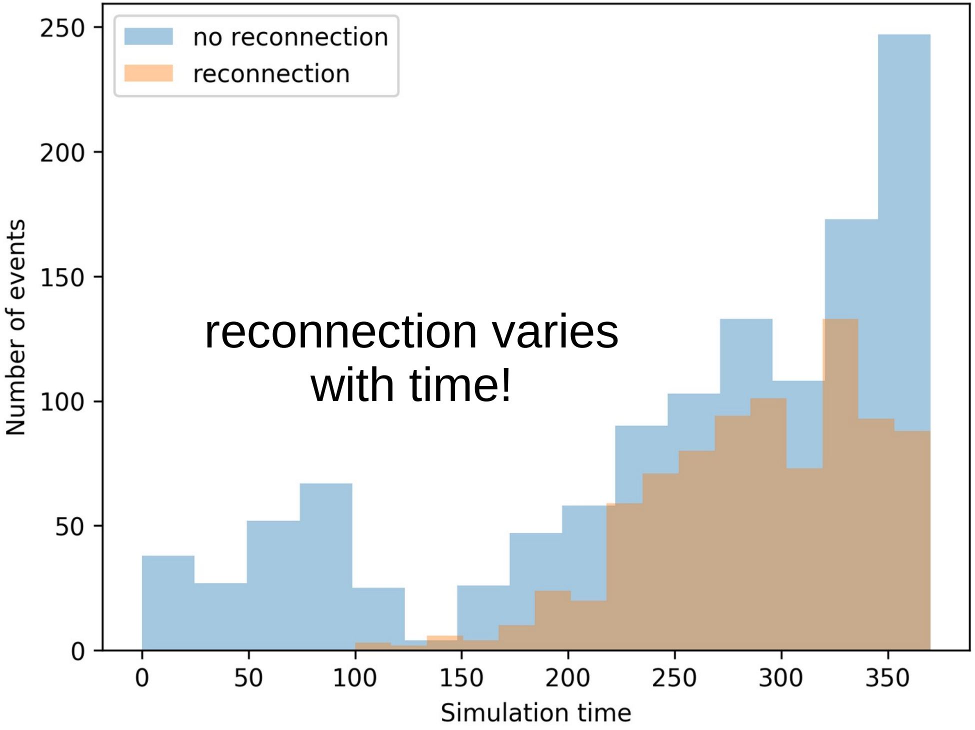
Variable name	Description
$\lvert\vec{J}\rvert$	L_2-norm of total current density $\vec{J}$
Ψ	flux function, $\vec{B} = \nabla\Psi \wedge \hat{z}$, $\vec{J} = -\nabla^2\Psi$
$V_{e,x}$	electron x-velocity
$V_{e,y}$	electron y-velocity
$V_{e,z}$	electron z-velocity
$V_{e,\text{plane}}$	$\sqrt{V_{e,x}^2 + V_{e,y}^2}$
B_z	z-component of magnetic field
B_{plane}	$\sqrt{B_x^2 + B_y^2}$
$E_{\text{dec},e}$	$(\vec{E} + V_e \times \vec{B})_z$ (decoupling electrons)
$E_{\text{dec},i}$	$(\vec{E} + V_i \times \vec{B})_z$ (decoupling ions)

The AIDA project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 776262.

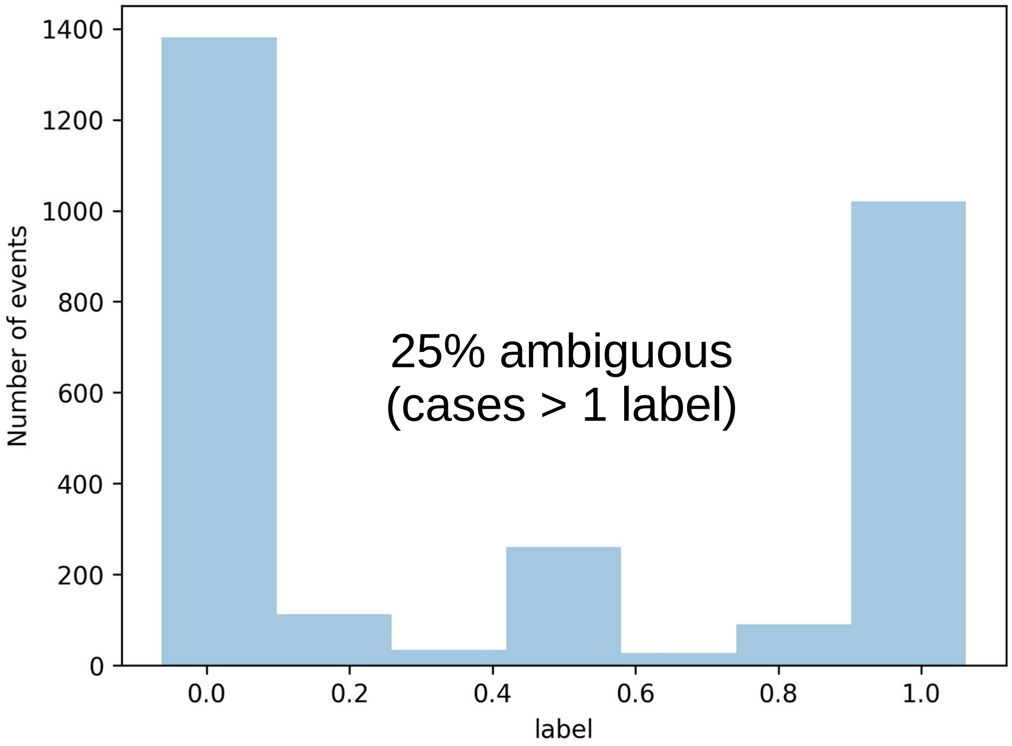
Variables in a selected region for human labelling



Statistics Sim 1



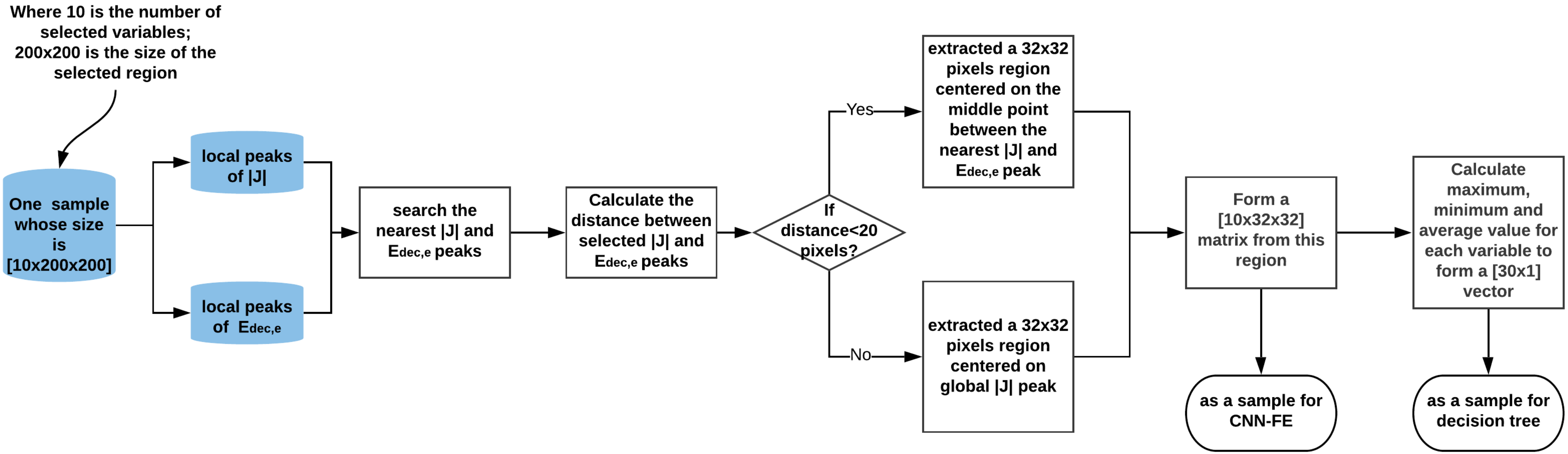
reconnection varies with time!



25% ambiguous (cases > 1 label)

train/test split in time

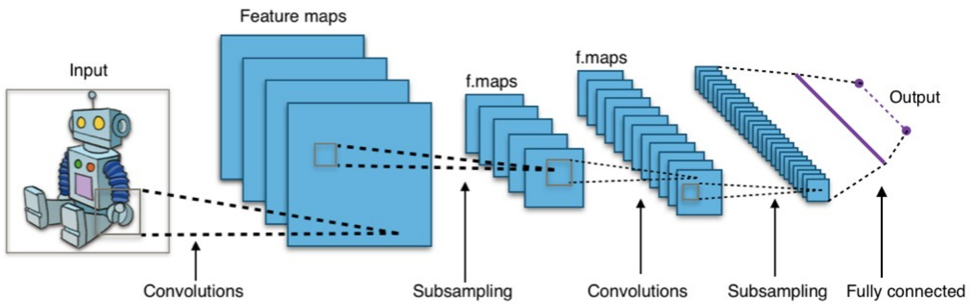
Feature engineering



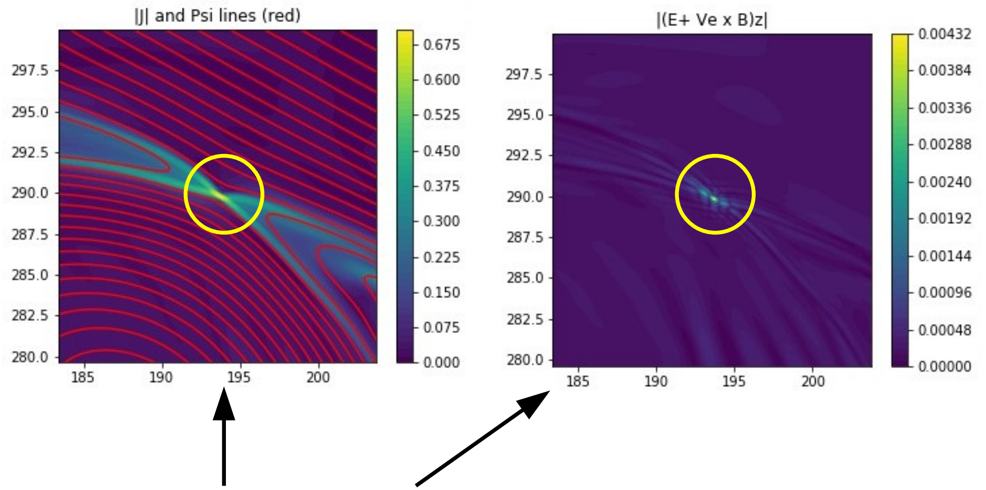
The AIDA project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 776262.

Machine learning models

Convolutional Neural Network (CNN)



Heuristic model



1. Detect peaks
2. Extract min/max/mean around peaks if they are close together
3. Use decision tree

Machine learning models

Used model in this study

name	feature engineering	input (per variable)
CNN-origin	×	200^2 area
CNN-FE	✓	32^2 area
decision tree	✓	min, max, mean of 32^2 area

The AIDA project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 776262.

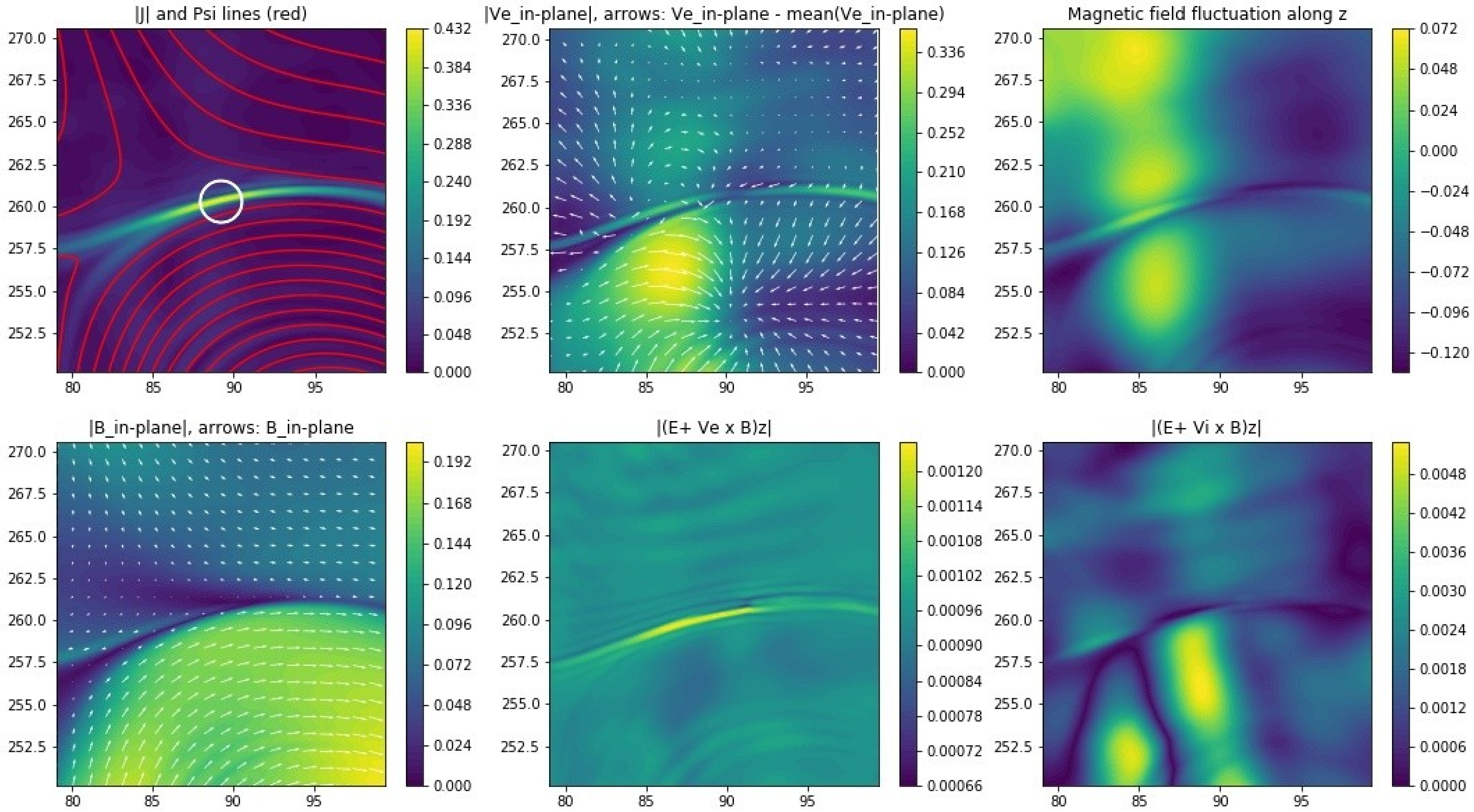
Results

No.	Training set	Test set	Model	TP	FP	FN	TN	F1 score	accuracy	MCC
1	Sim 1	Sim 2	CNN-FE	58	12	22	32	0.73	0.77	0.44
2	(33%-100%)	(0-33%)	CNN-FE	-	-	-	-	0.76	0.85	0.56
3	(0-33%), (67%-100%)	(33%-67%)	CNN-FE	-	-	-	-	0.74	0.84	0.55
4	(0-67%)	(67%-100%)	CNN-FE	239	52	134	196	0.72	0.70	0.42
5	(0-67%)	(67%-100%)	CNN	223	58	150	190	0.68	0.67	0.36
6	(0-67%)	(67%-100%)	Decision tree	266	88	107	160	0.62	0.68	0.35

Except Scheme No. 1, all the percentage of both training and test sets are referring to Sim 1

The AIDA project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 776262.

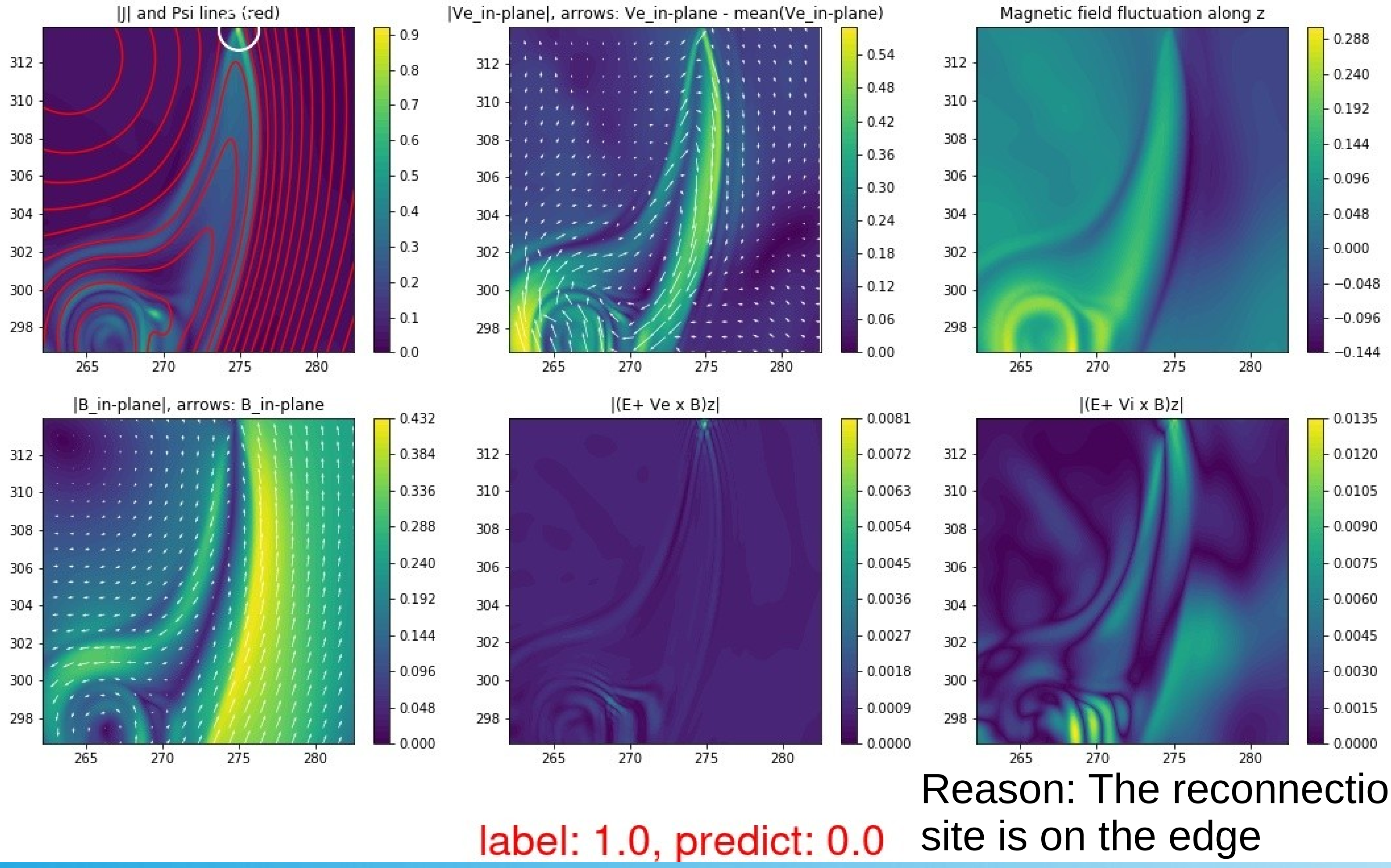
False-Negative (FN) case



label: 1.0, predict: 0.0

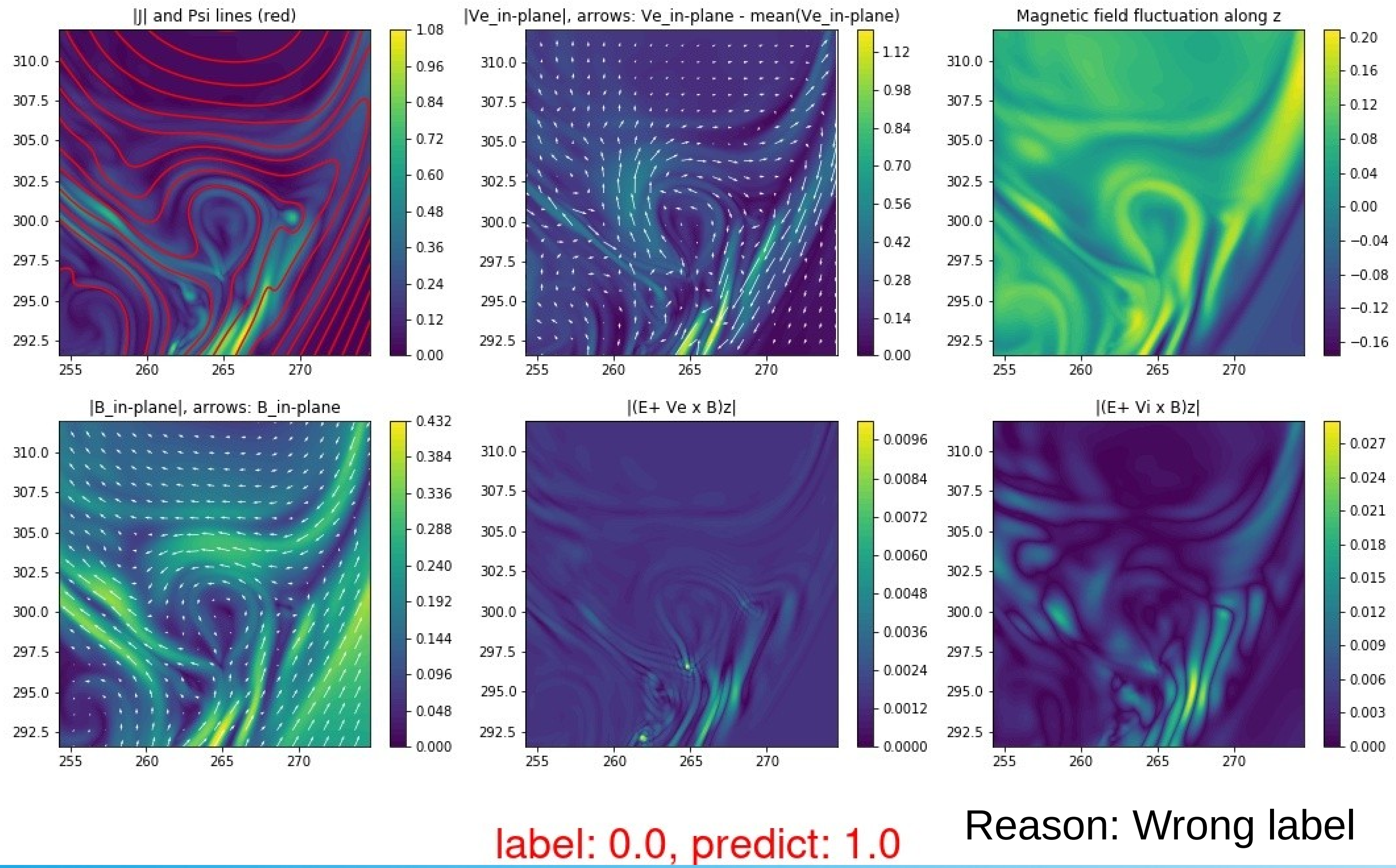
Reason: The figure is ambiguous

False-Negative (FN) case



Reason: The reconnection site is on the edge

False-Positive (FP) case



Reason: Wrong label

Conclusion & Future

1. The CNN-FE model is generic to other simulations.

2. The influence of each variable to reconnection label has also been investigated in this study. The results show that |J| and $V_{e,z}$ contribute most to the reconnection classification. Ψ , B_{plane} and $V_{e,plane}$ are slightly less significant in comparison with the first two variables. The other variables, $V_{e,x}$, $V_{e,y}$, B_z and $E_{dec, i}$ are not very significant to the classification.

3. The developed CNN-FE model also manages to find good reconnection events even which an human expert might make a mistake with.

The AIDA project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 776262.