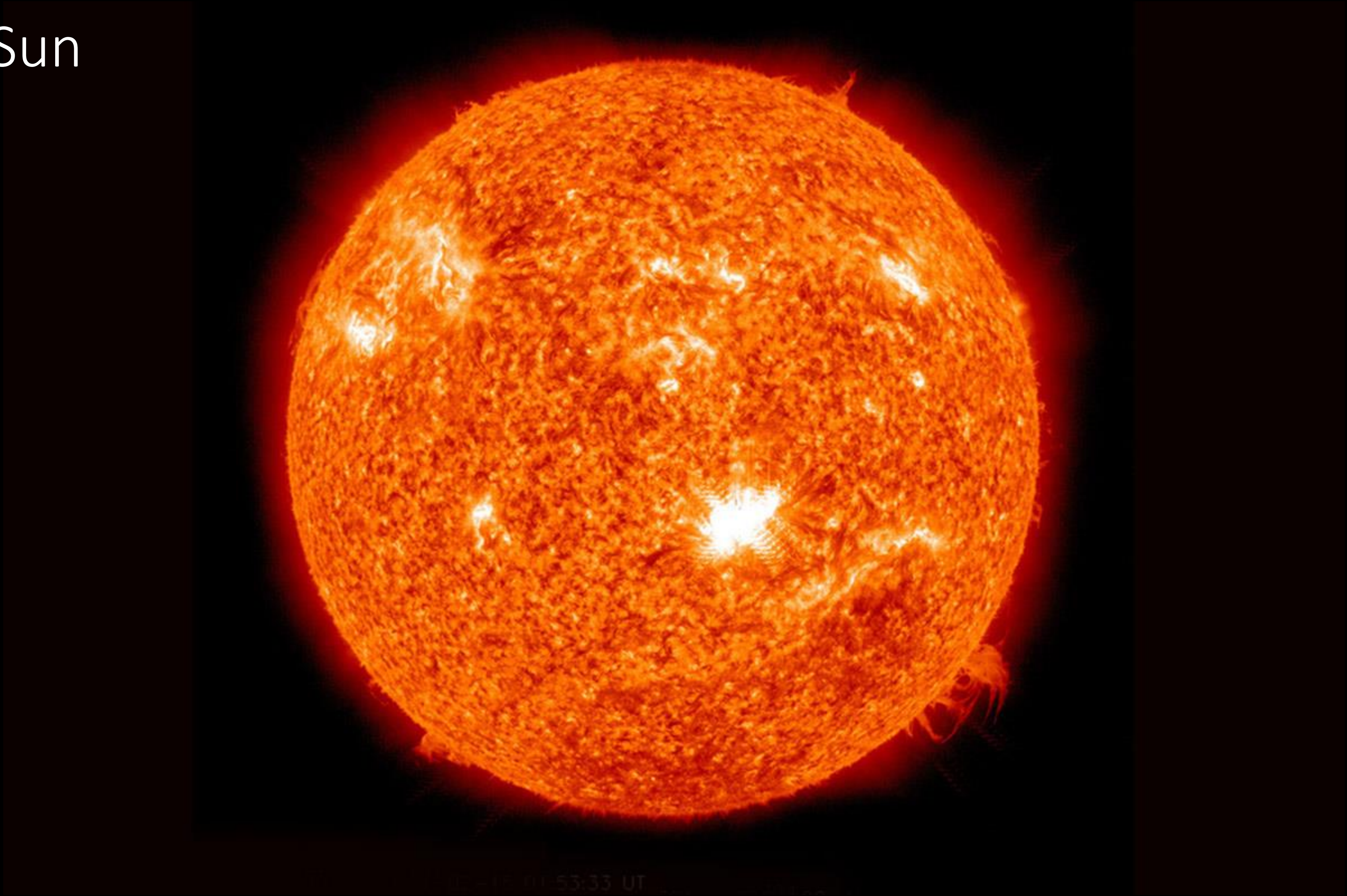
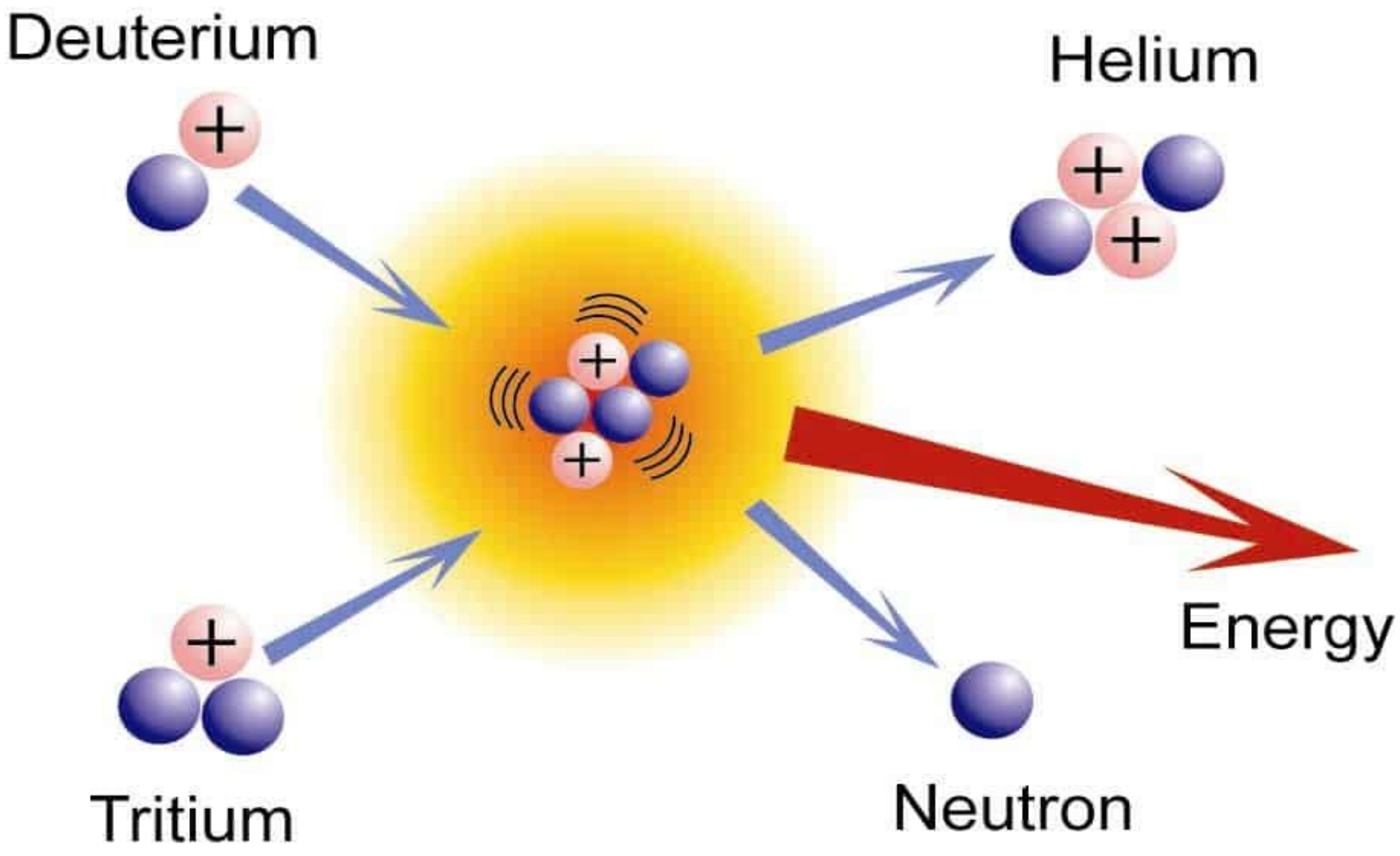
Applications of Deep Learning to Nuclear Fusion Research

Diogo R. Ferreira
IST, University of Lisbon
diogo.ferreira@tecnico.ulisboa.pt

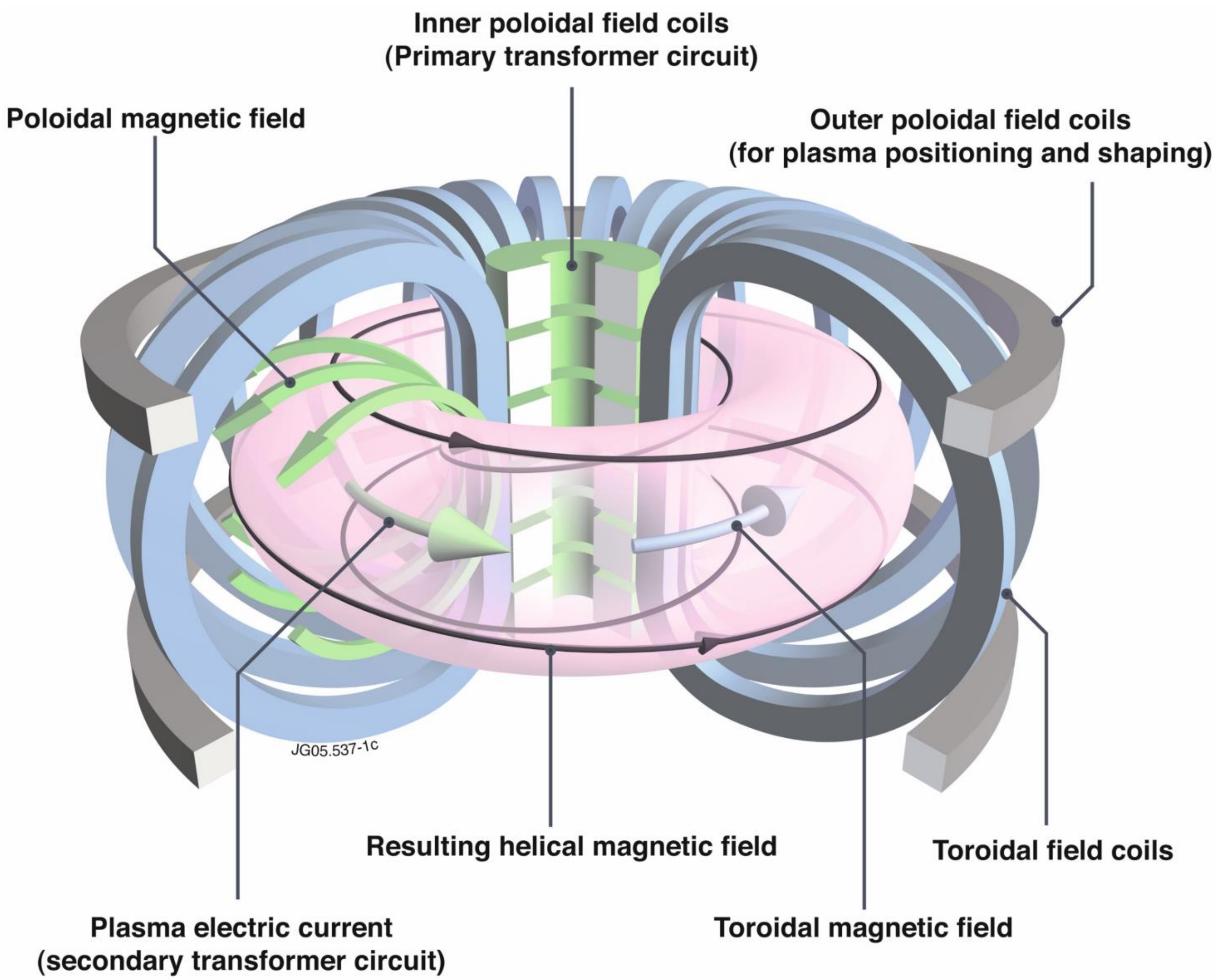
The Sun



Nuclear Fusion Reaction

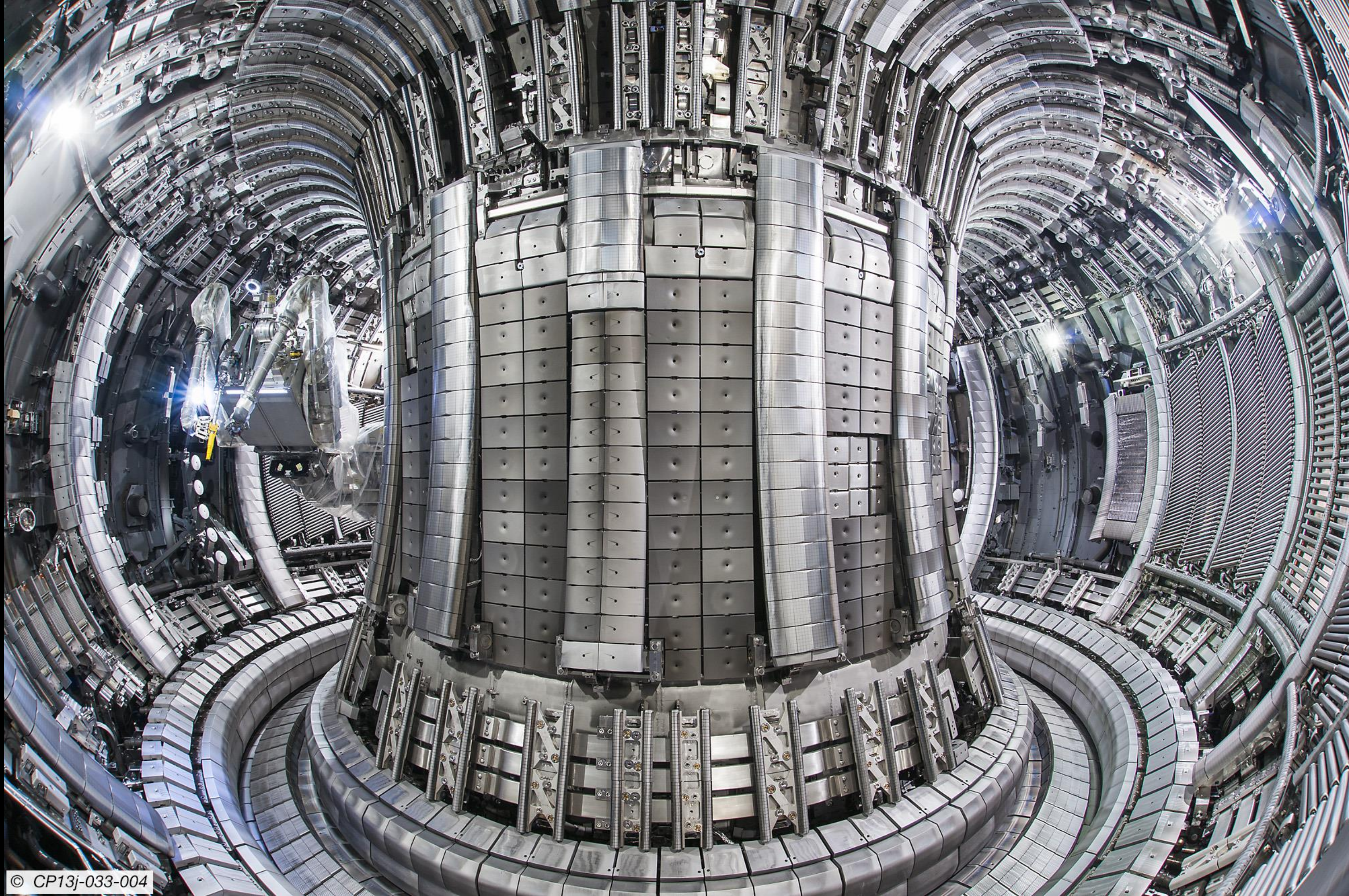


The Tokamak

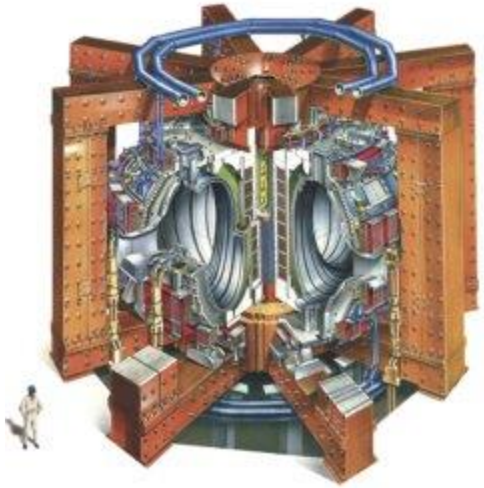


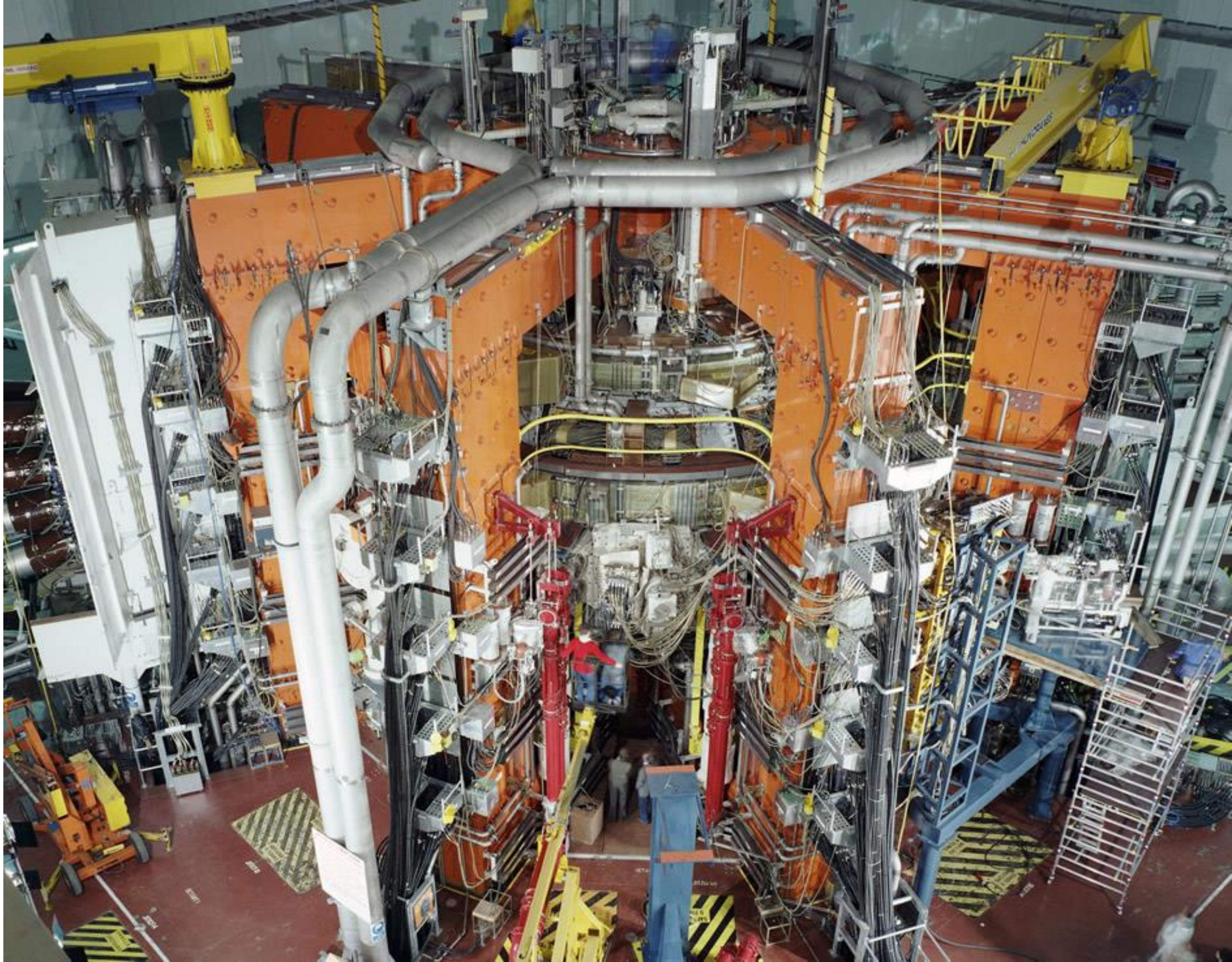
Inner poloidal field coils
(Primary transformer circuit)
Poloidal magnetic field
Outer poloidal field coils
(for plasma positioning and shaping)
JG05.537-1c
Resulting helical magnetic field
Toroidal field coils
Plasma electric current
(secondary transformer circuit)
Toroidal magnetic field

JET
(Joint
European
Torus)
near
Oxford,
UK

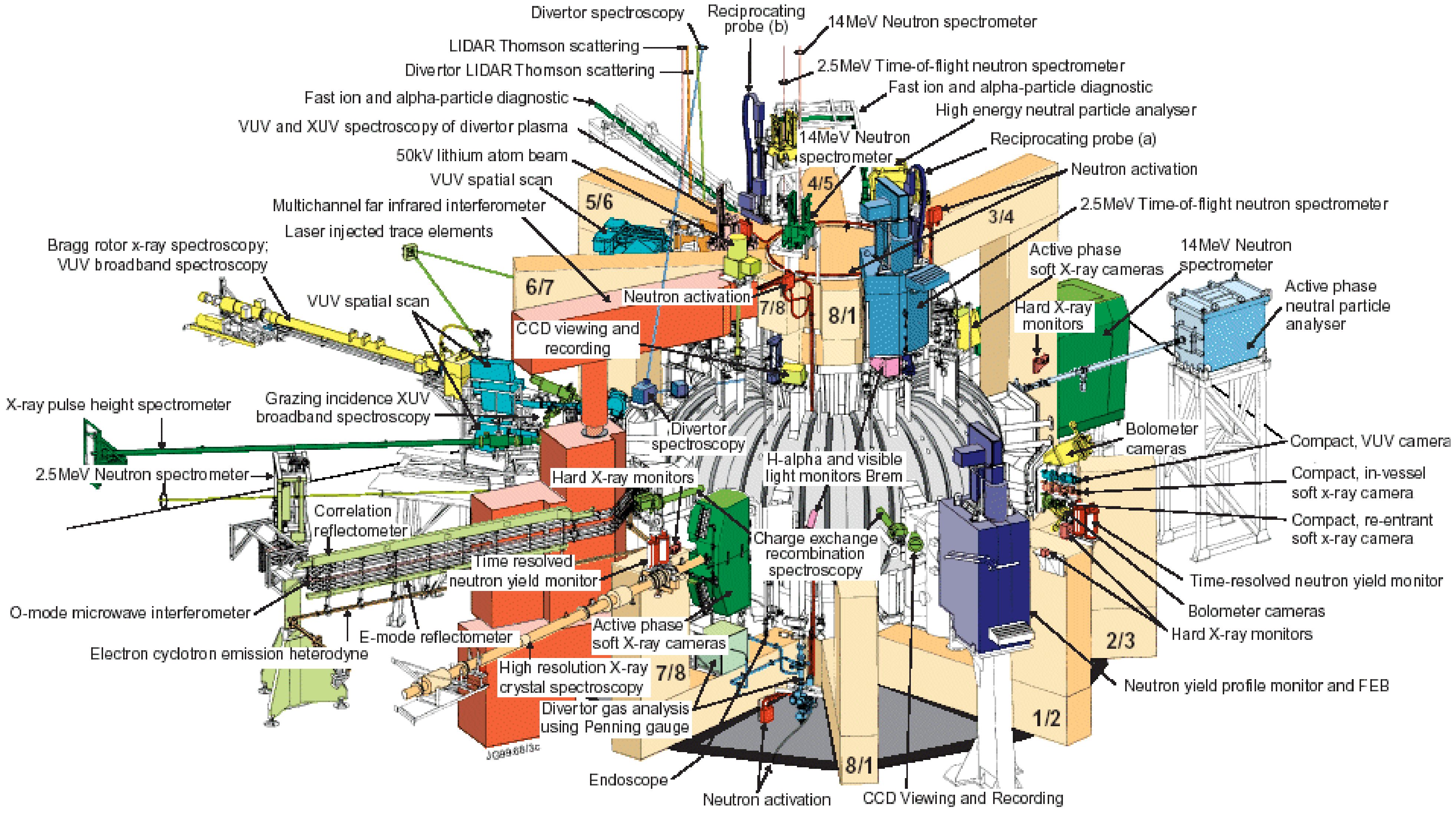


JET
(Joint European Torus) near Oxford, UK

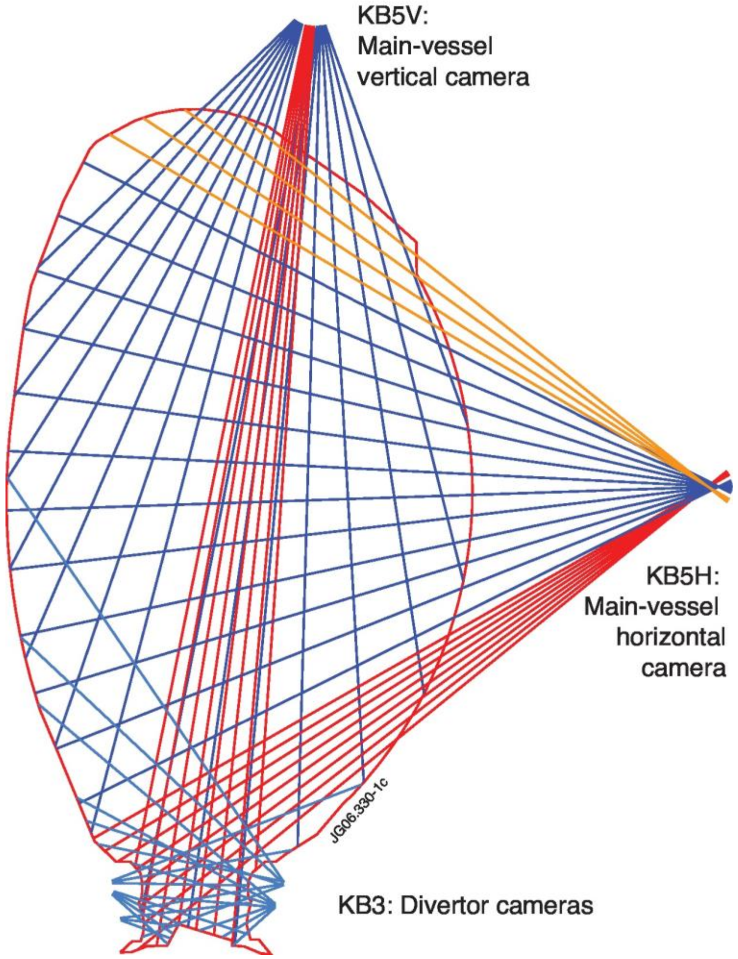


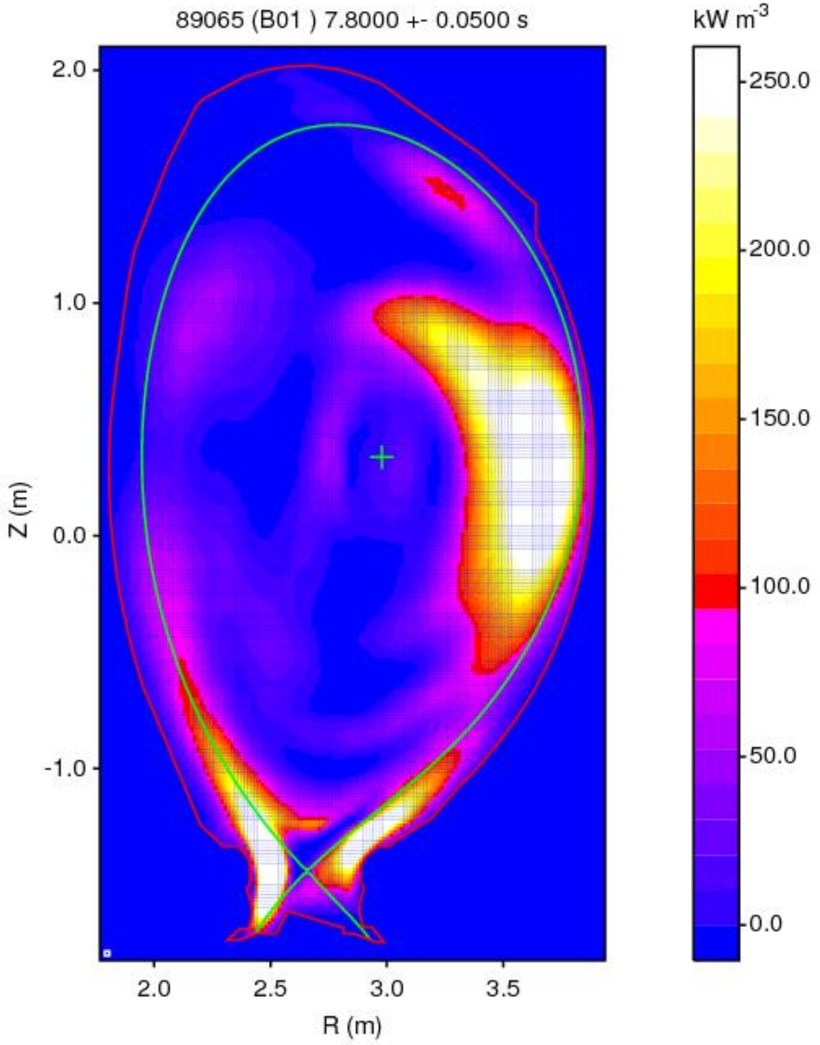


JET Diagnostics



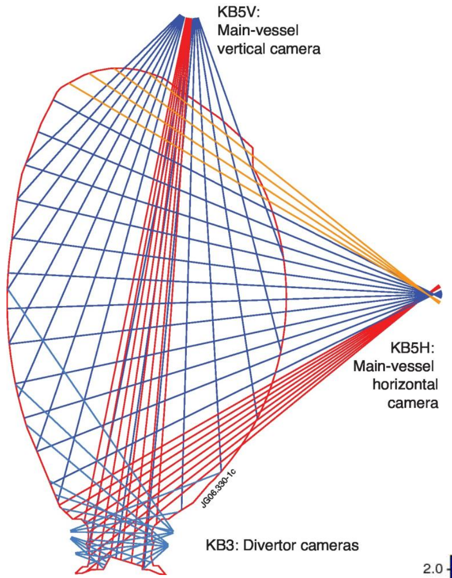
Tomography at JET

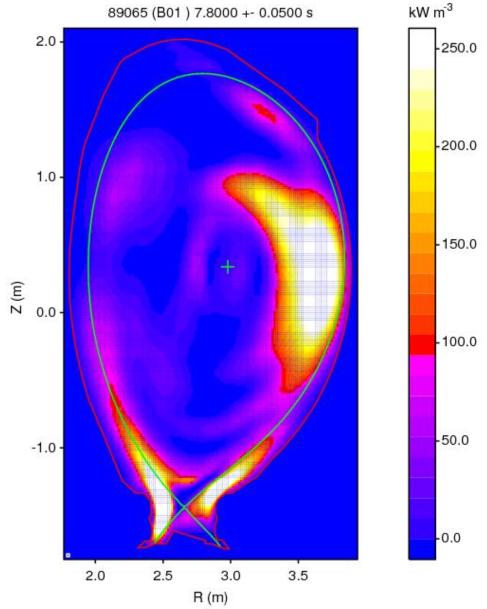




Tomography at JET

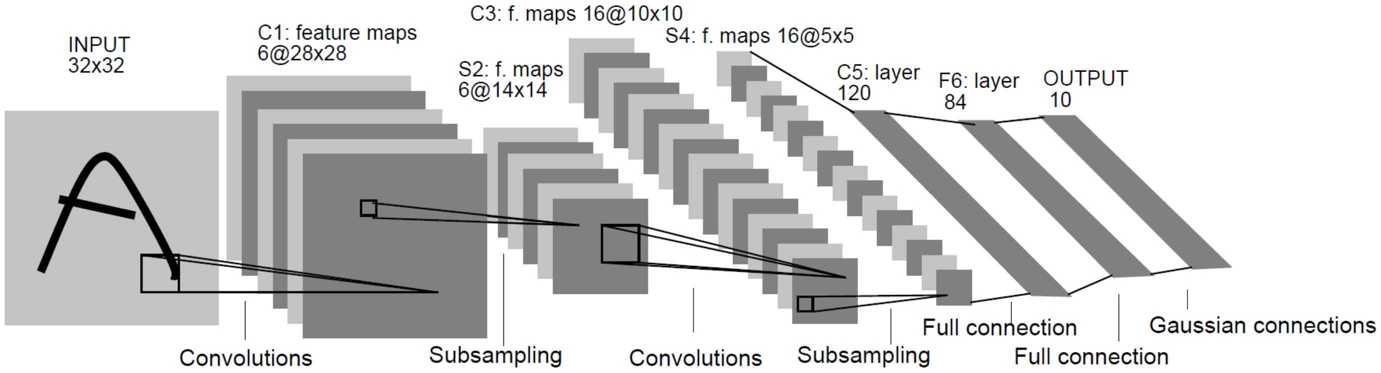
- Camera signals
  - sampling rate: 5 kHz
  - window average of 5 ms (25 samples)
  - 5 kHz / 25 = 200 Hz
- Tomographic reconstructions
  - pulse duration: ~30 sec
  - 30 sec × 200 Hz = 6000 reconstructions/pulse
  - in practice, only a few reconstructions per pulse
- Time per reconstruction
  - ~1h on average
  - 6000 × 1h = 250 days



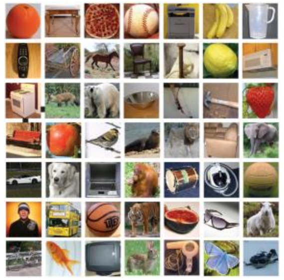


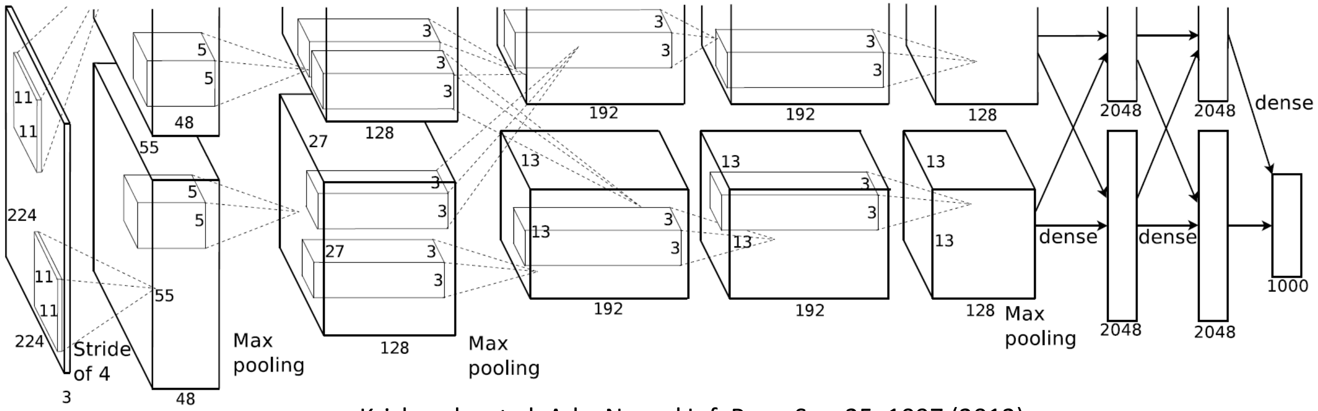
Deep Learning

- Convolutional Neural Networks (CNNs)



LeCun et al, Proc. of the IEEE 86, 2278 (1998)

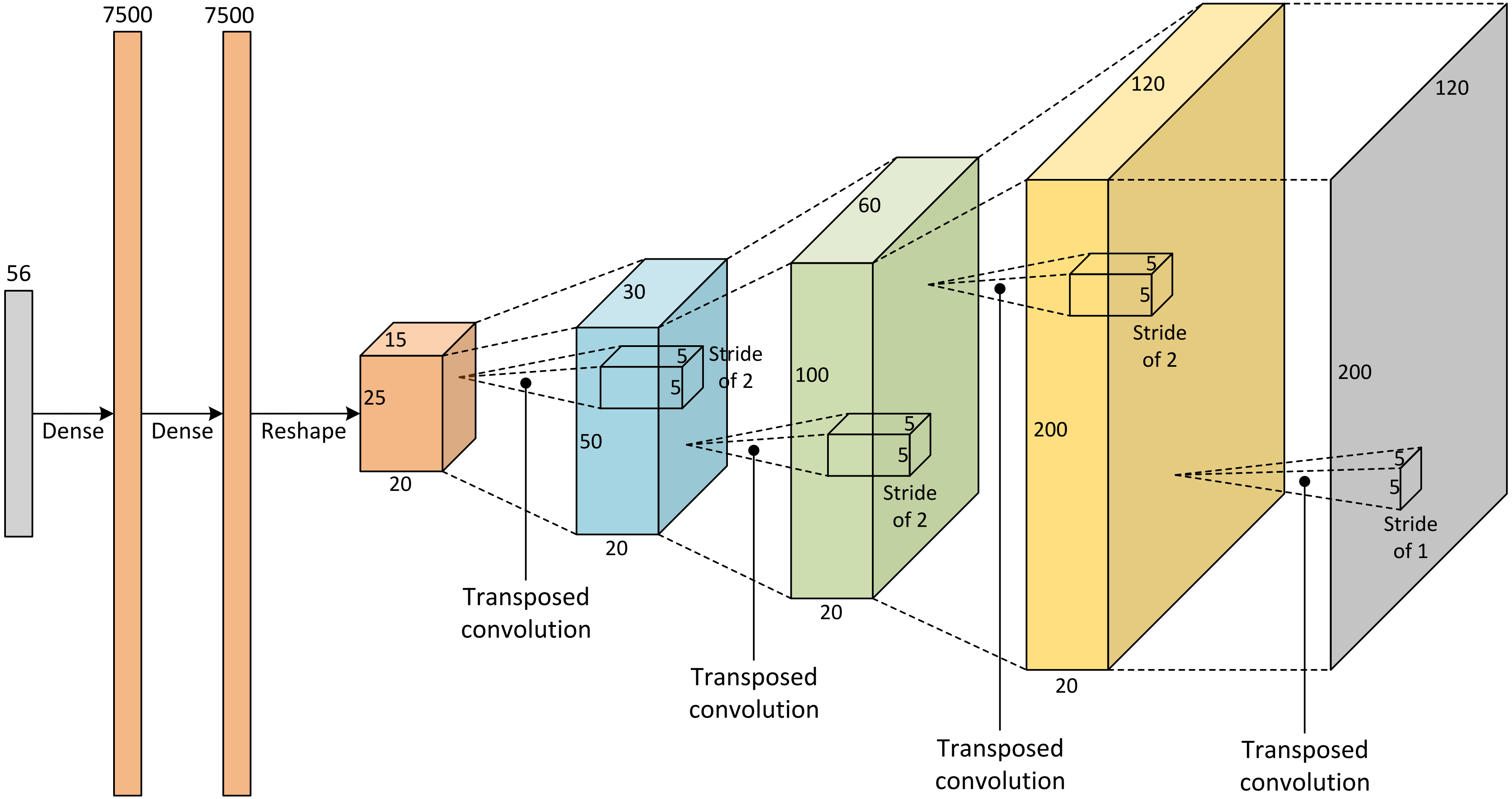




Krizhevsky et al, Adv. Neural Inf. Proc. Sys. 25, 1097 (2012)

Deep Learning

- Inverse of a CNN ~ "deconvolutional" neural network



Training

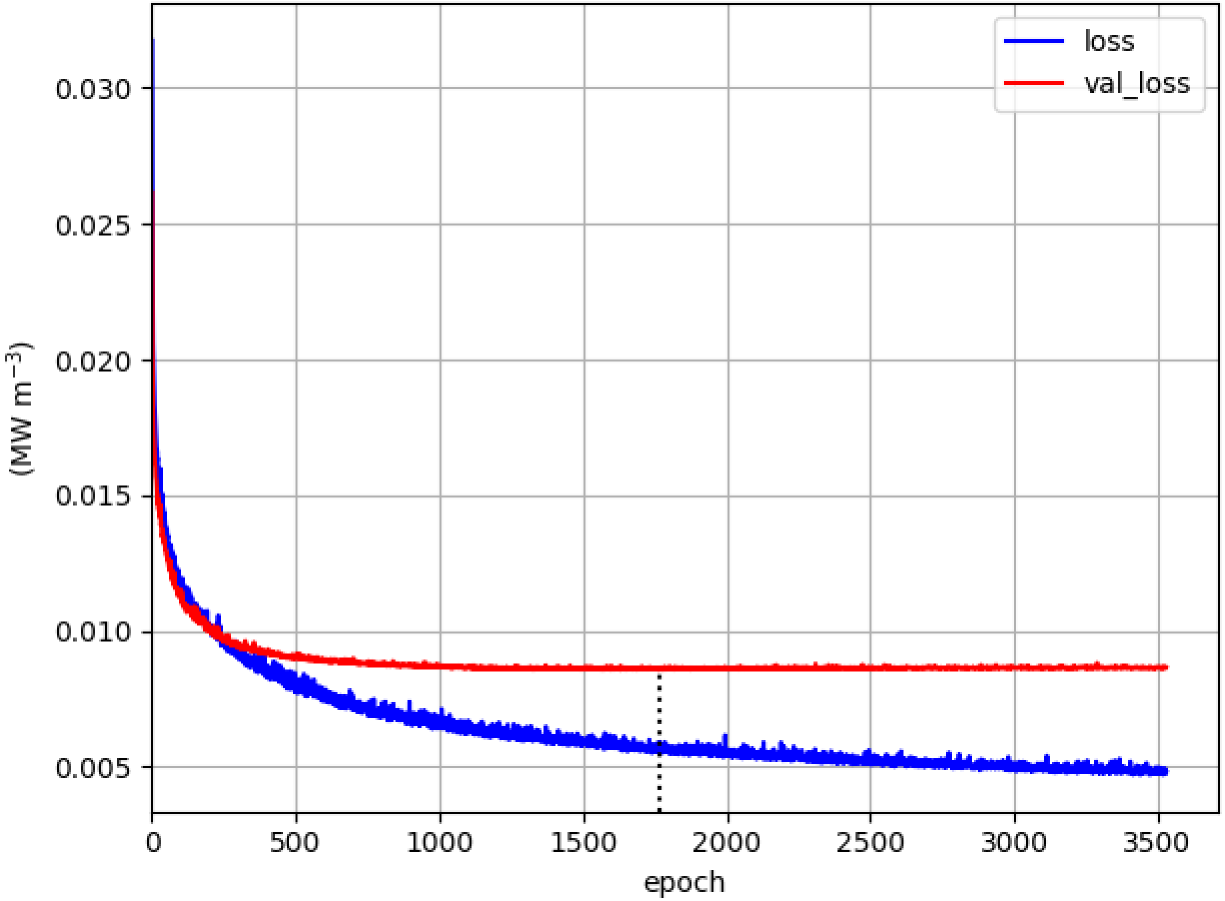
- Dataset
  - pulses 80128 to 92504 (2011-2016)
  - 25584 sample reconstructions
  - 90% for training (23025), 10% for validation (2559)
- Training
  - accelerated gradient descent (Adam)
  - learning rate: 10^{-4}
  - batch size: 307 (307 * 75 = 23025)
  - 75 batches = 75 updates/epoch

Training

Keras

TensorFlow

NVIDIA

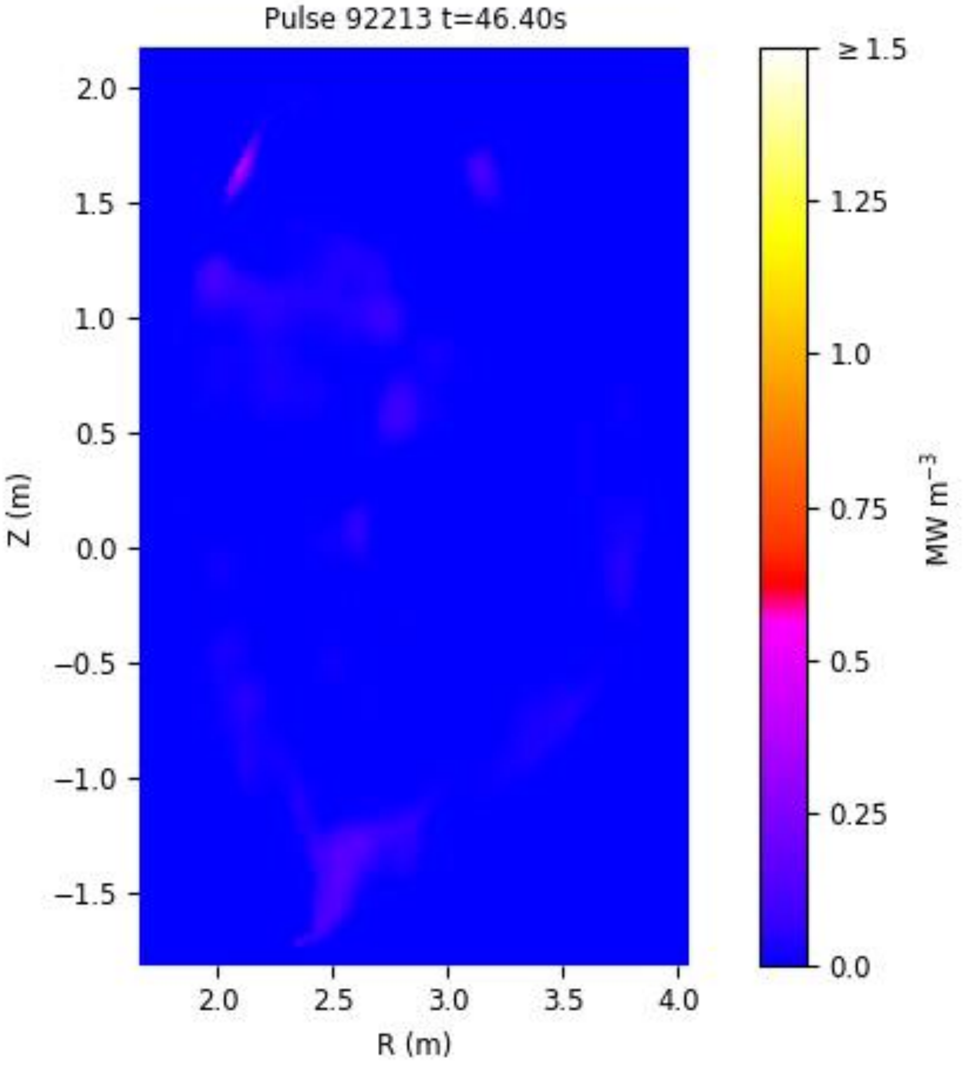


NVIDIA Tesla P100

total:
3530 epochs
20 hours

best:
epoch 1765
val_loss 0.008556 MW m^{-3}

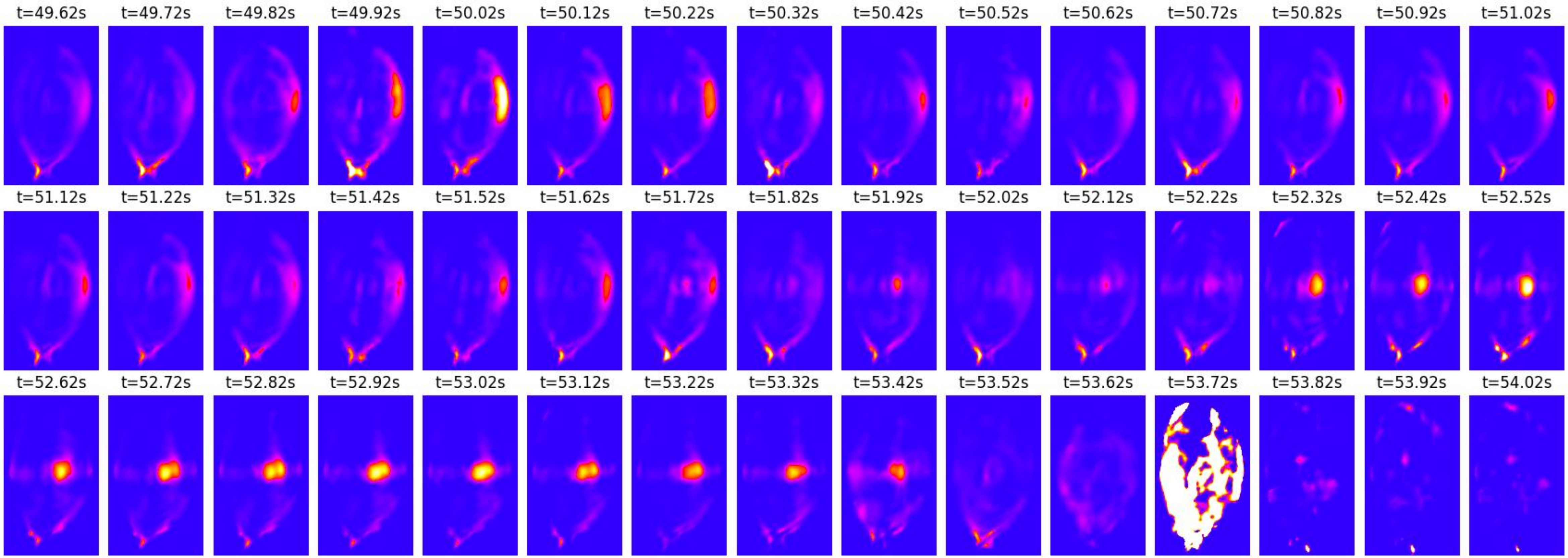
Video Demo



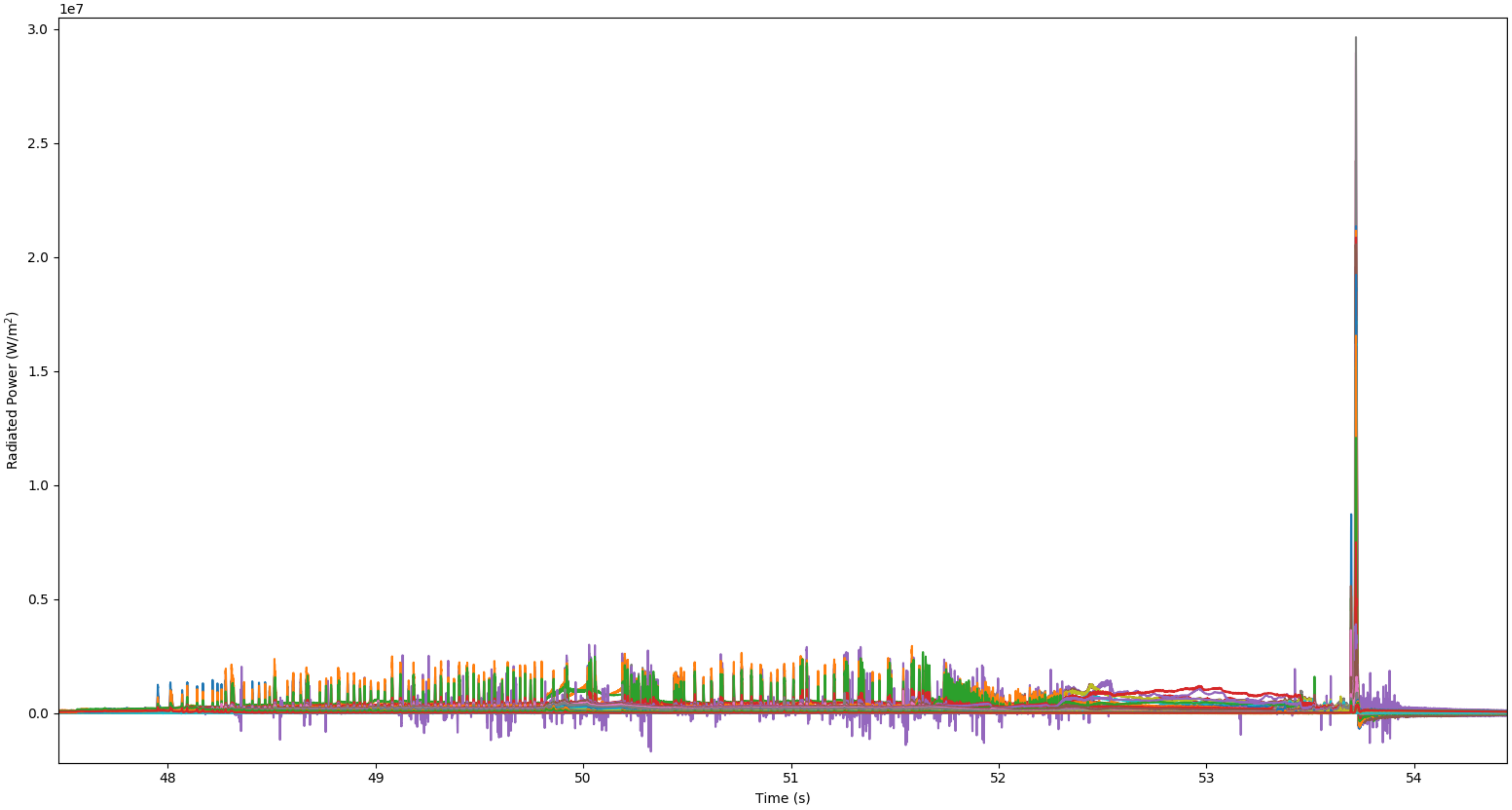
Pulse 92213 t=46.40s
Z (m)
R (m)
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
2.0
2.5
3.0
3.5
4.0
≥ 1.5
1.25
1.0
0.75
0.5
0.25
0.0
MW m^{-3}

Plasma Disruptions

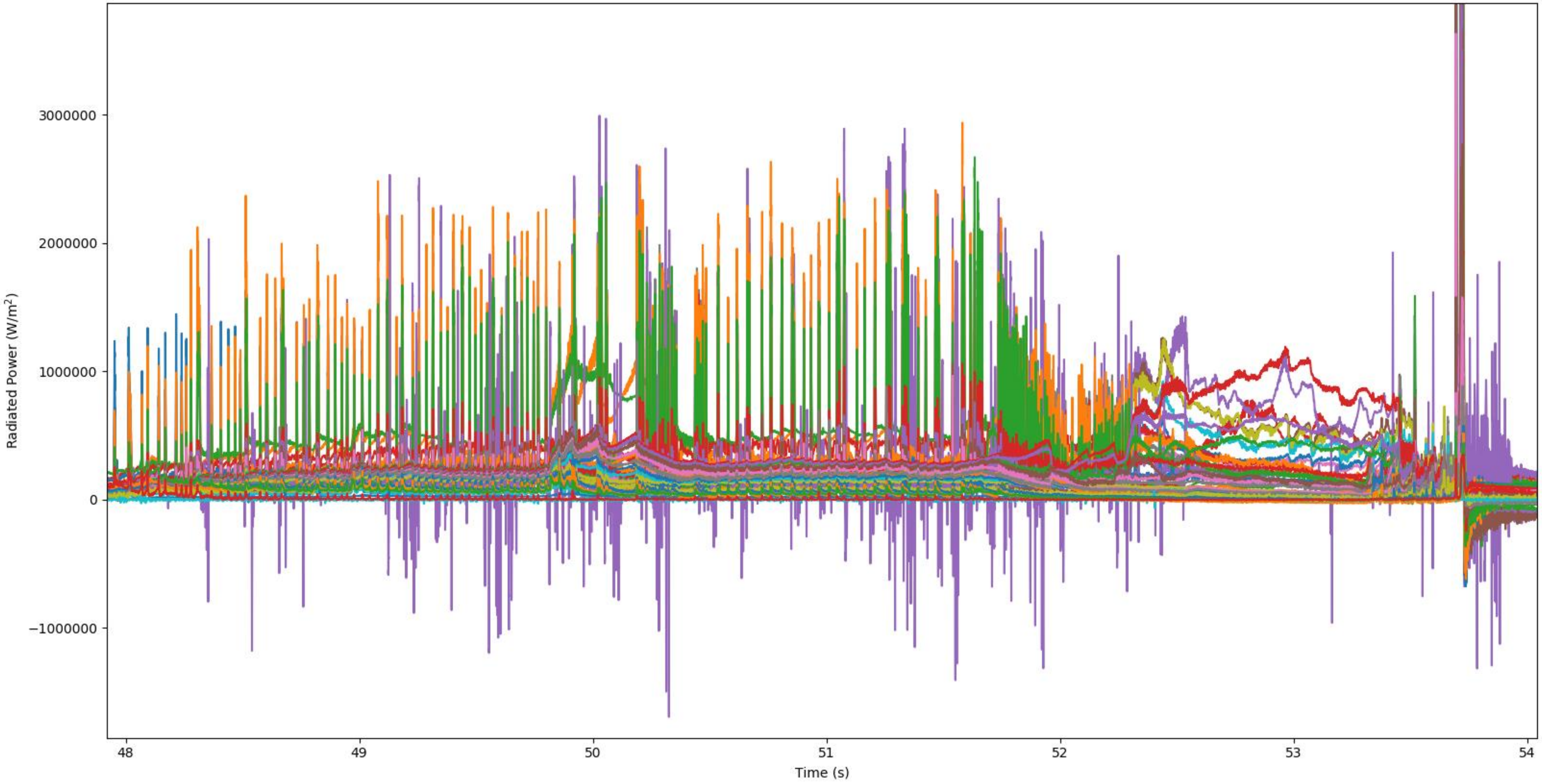
- Disruptions are a major problem in tokamaks



Raw Camera Signals



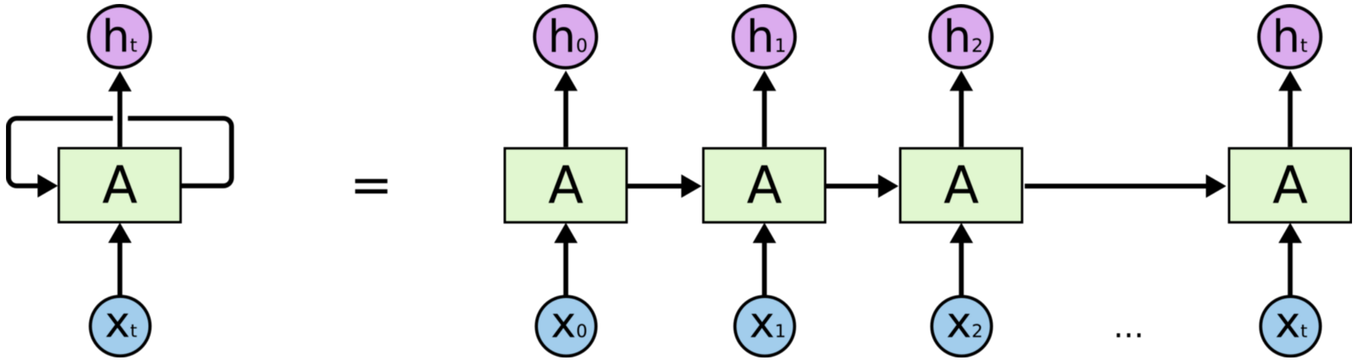
Raw Camera Signals



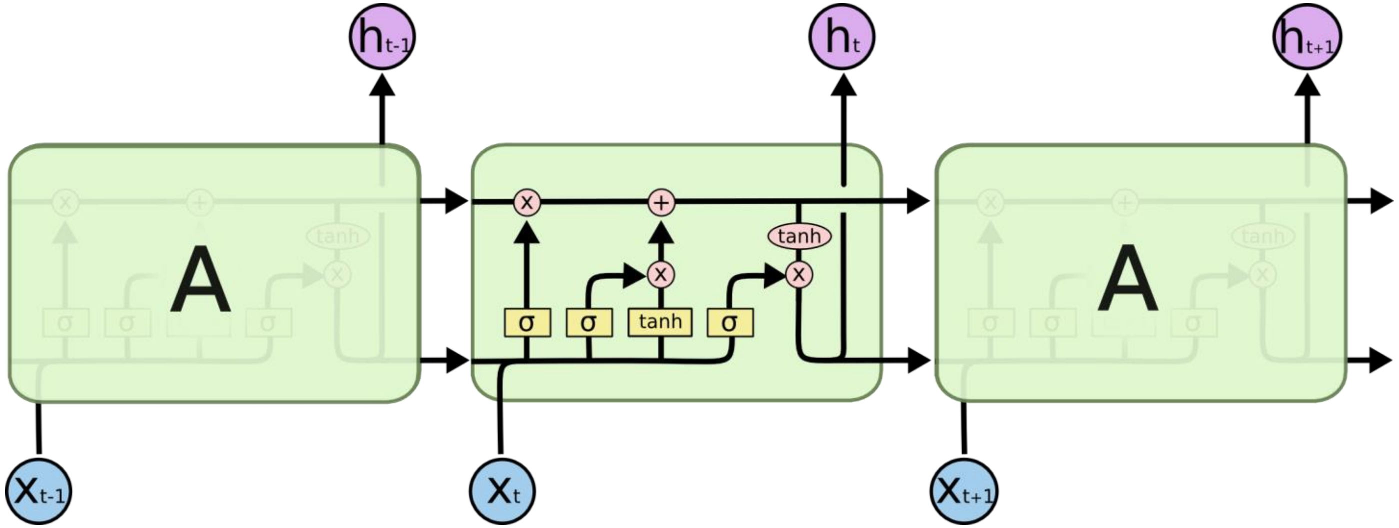
Deep Learning

- Recurrent Neural Networks (RNNs)

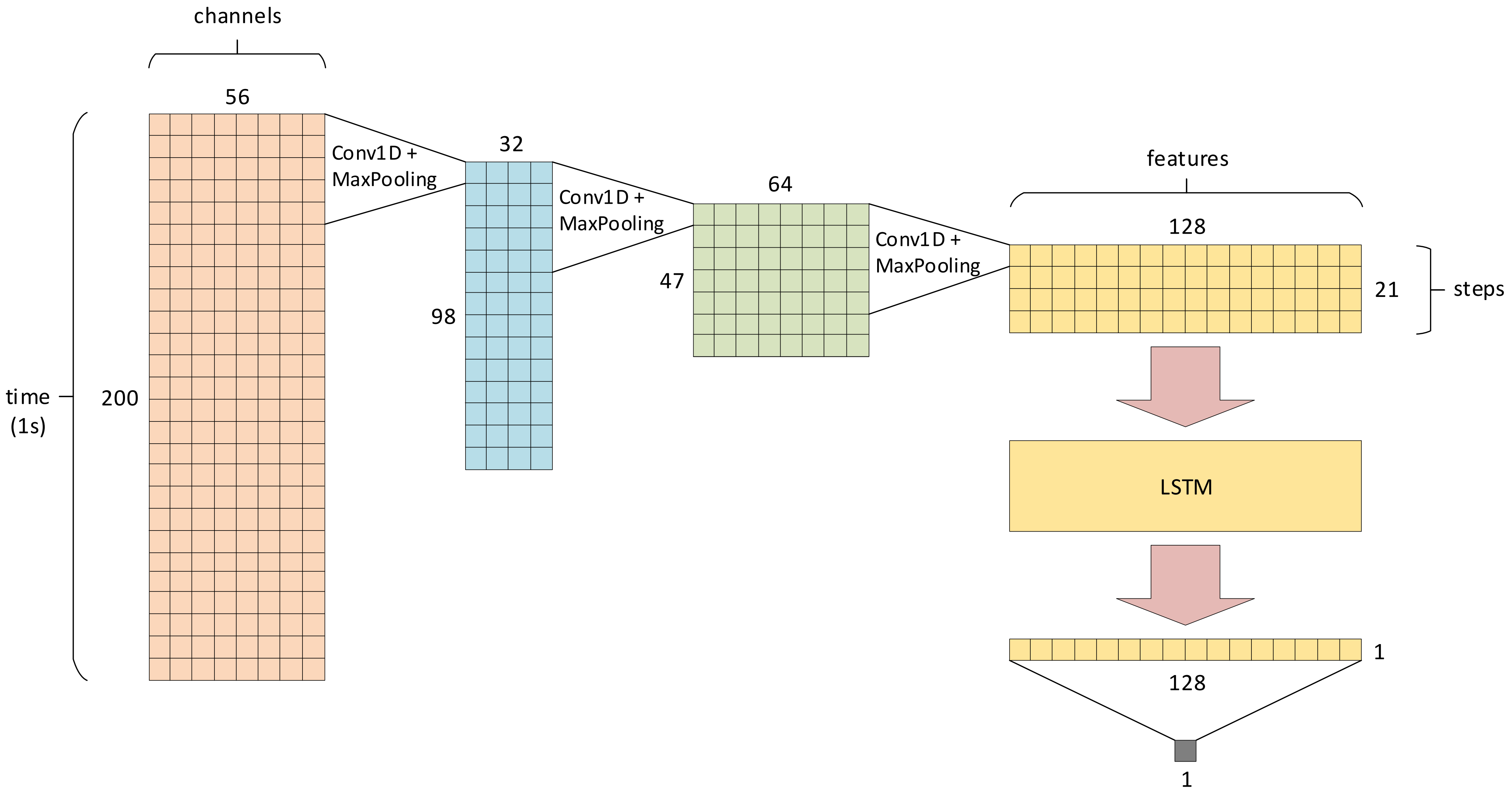
Simple RNN



LSTM (Long short-term memory)



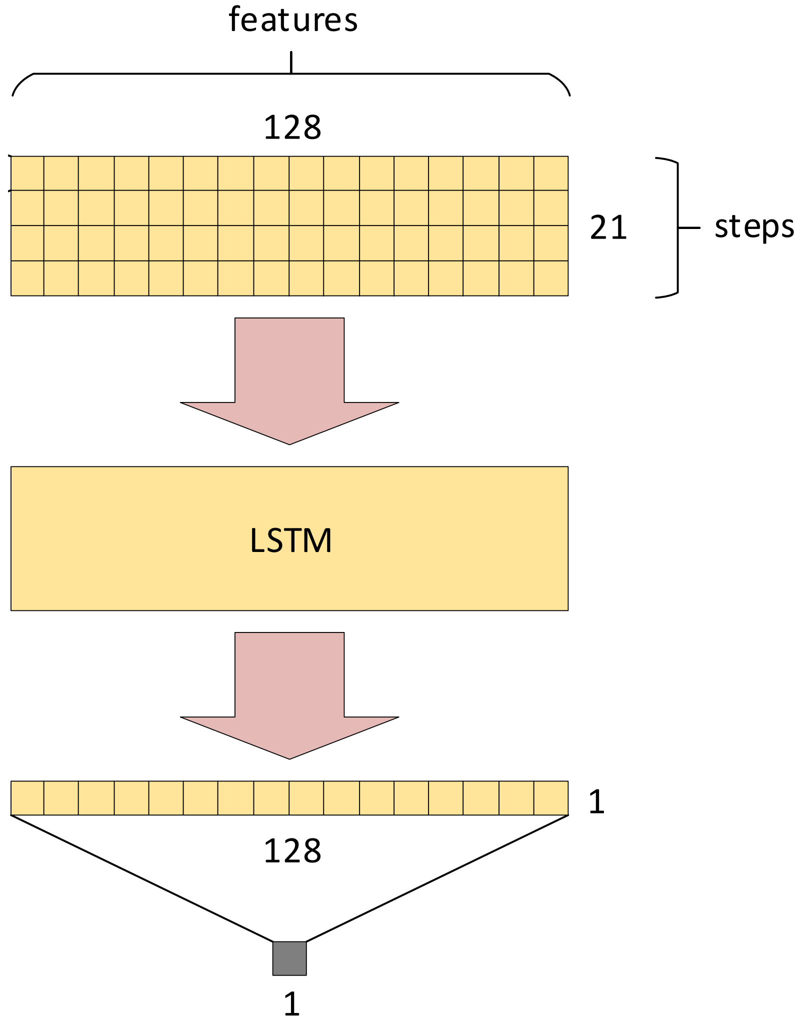
Disruption Prediction



Disruption Prediction

- Time-to-disruption (regression)
  - last layer is Dense(1) with **no activation**
  - loss function is **mean absolute error** (mae)
  - use only **disruptive** pulses

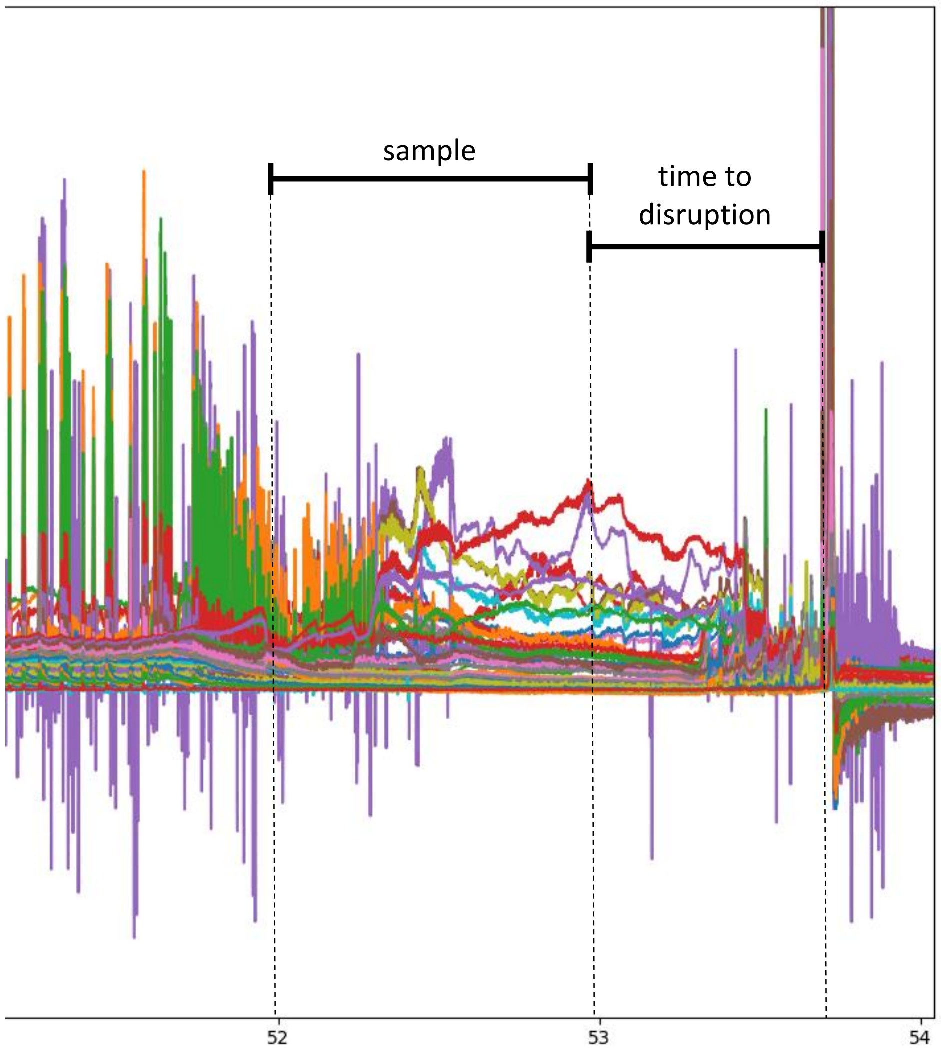
- Probability of disruption (classification)
  - last layer is Dense(1) with **sigmoid** activation
  - loss function is **binary cross-entropy**
  - use both **disruptive** and **non-disruptive** pulses



Training

- Time-to-disruption (ttd)
  - 1683 disruptive pulses
  - 90% for training, 10% for validation
  - X: draw random samples from each pulse
  - y: (disruption time) – (latest sample time)

- Probability of disruption (prd)
  - 9798 pulses (17% disruptive)
  - 90% for training, 10% for validation
  - X: draw random samples from each pulse
  - y: 1 if pulse contains disruption, 0 otherwise



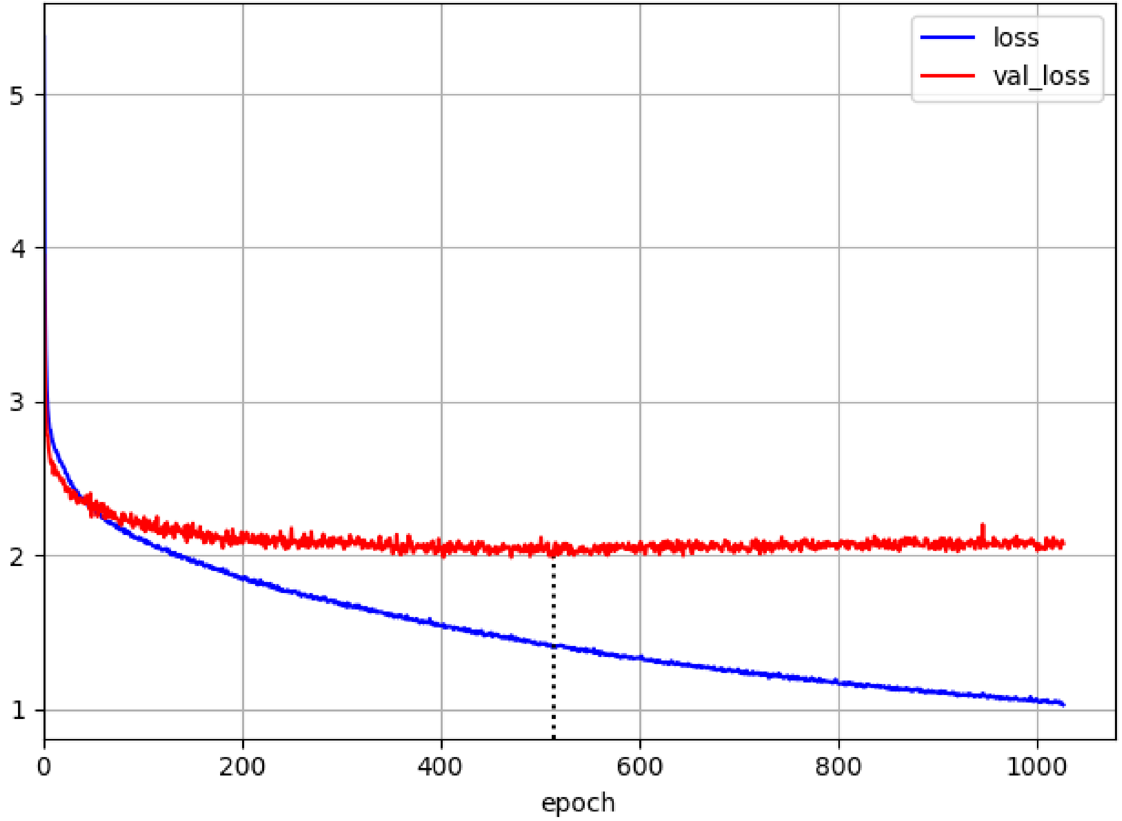
Training

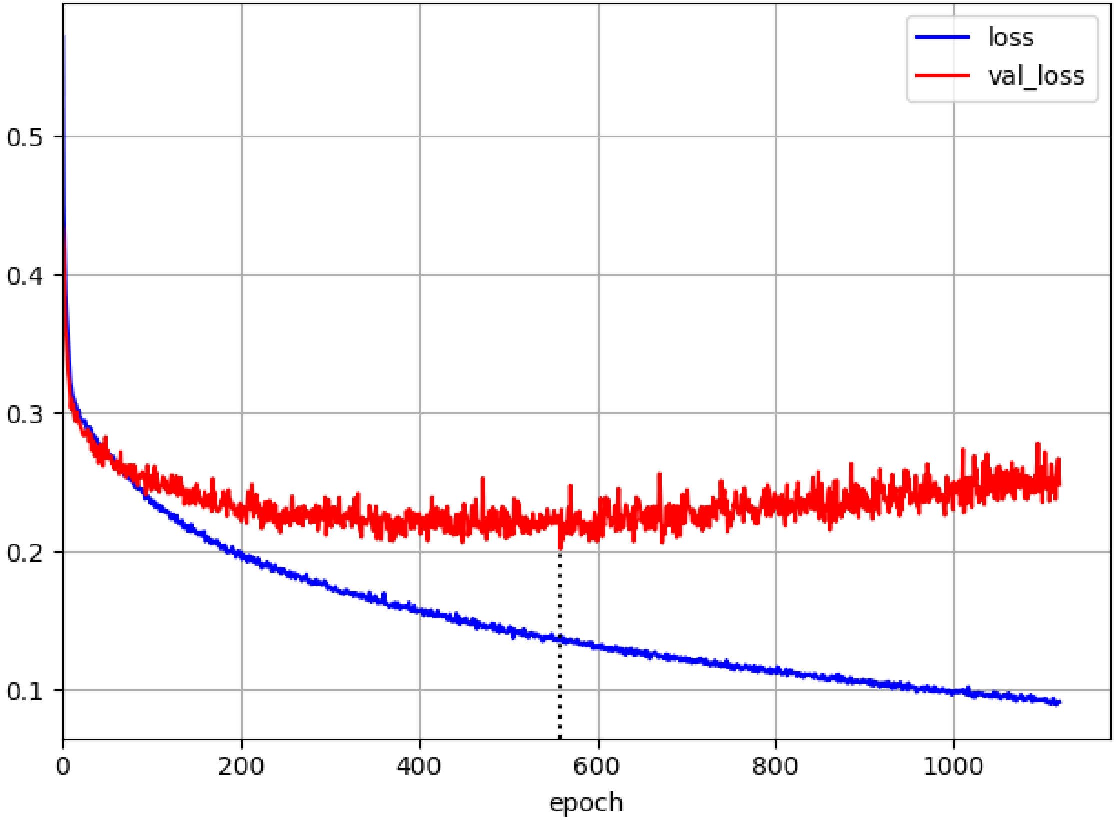
ttd (mean absolute error)
NVIDIA Tesla P100

total:	**best:**
1026 epochs	epoch 513
13 hours	val_loss 1.986 s

prd (binary cross-entropy)
NVIDIA Tesla P100

total:	**best:**
1118 epochs	epoch 559
14 hours	val_loss 0.202

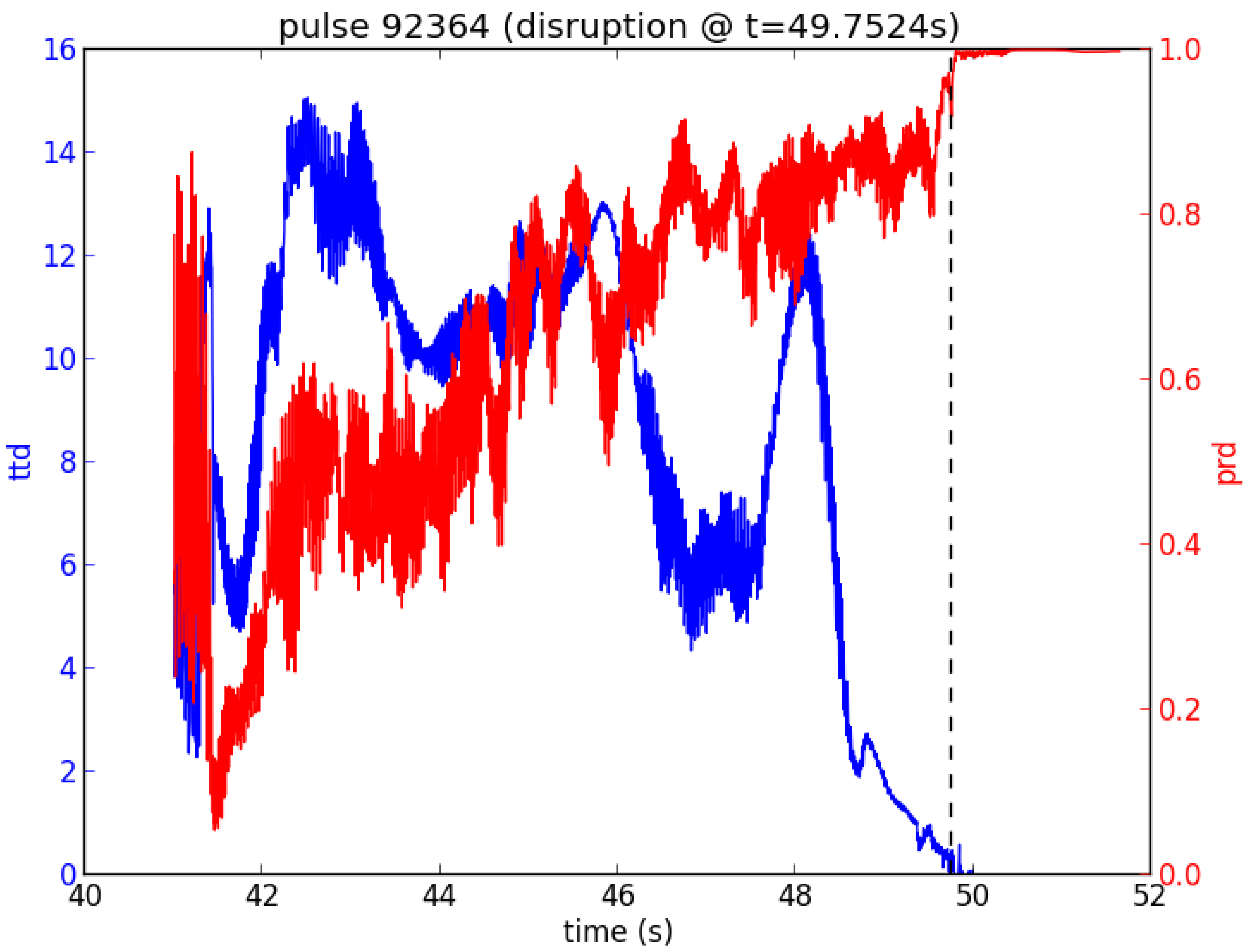




Results

low ttd
high prd

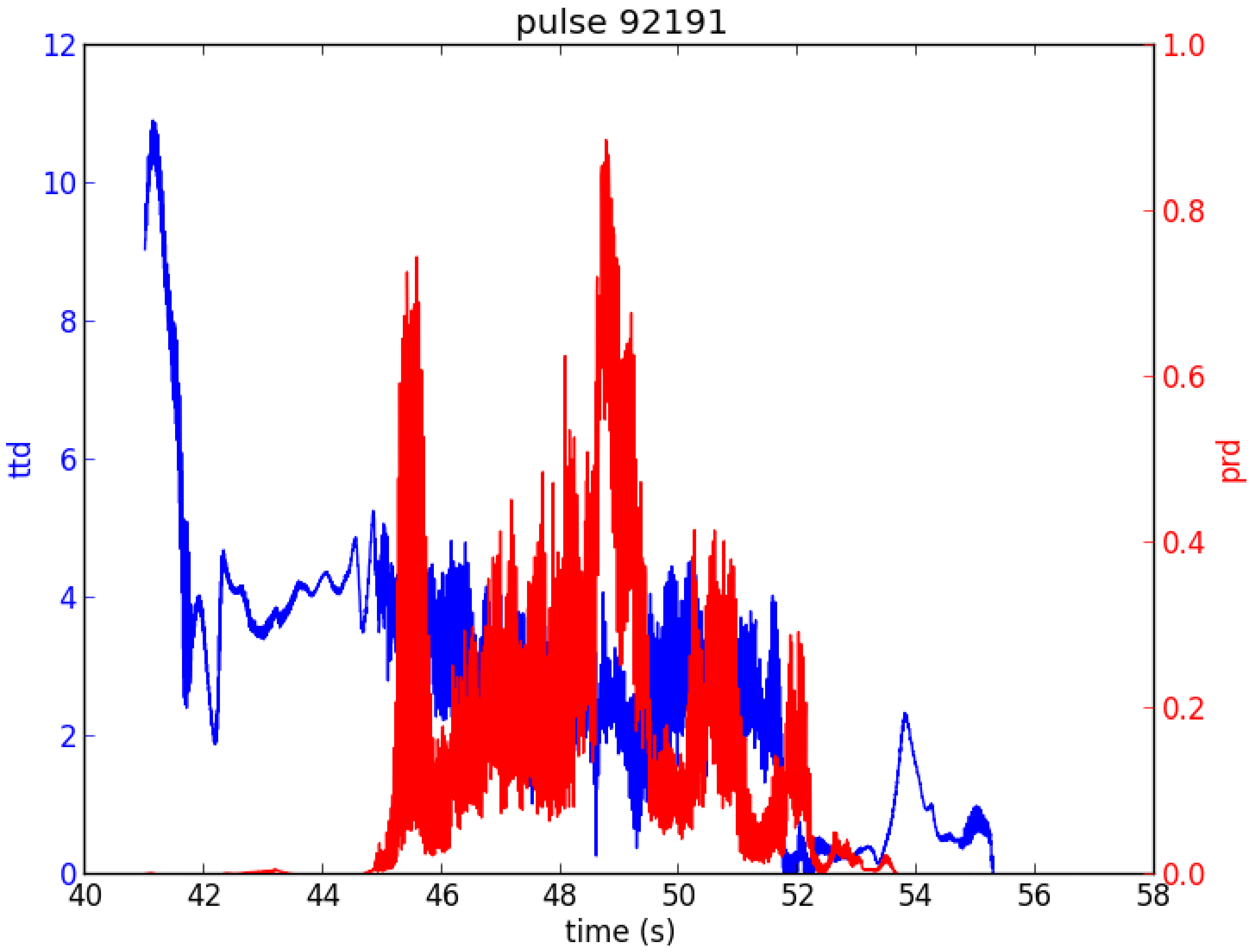
disruption



Results

low ttd
low prd

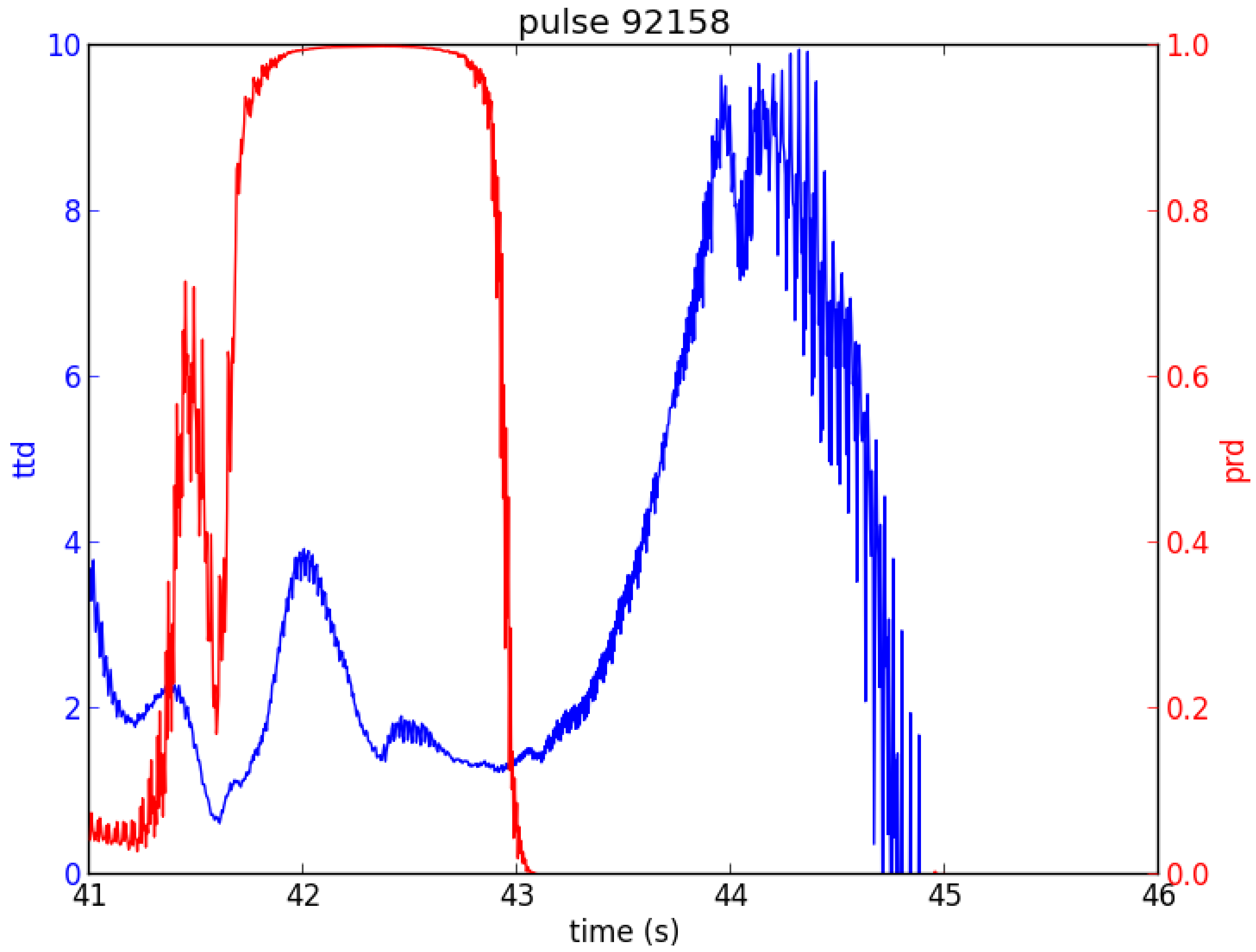
no disruption



Results

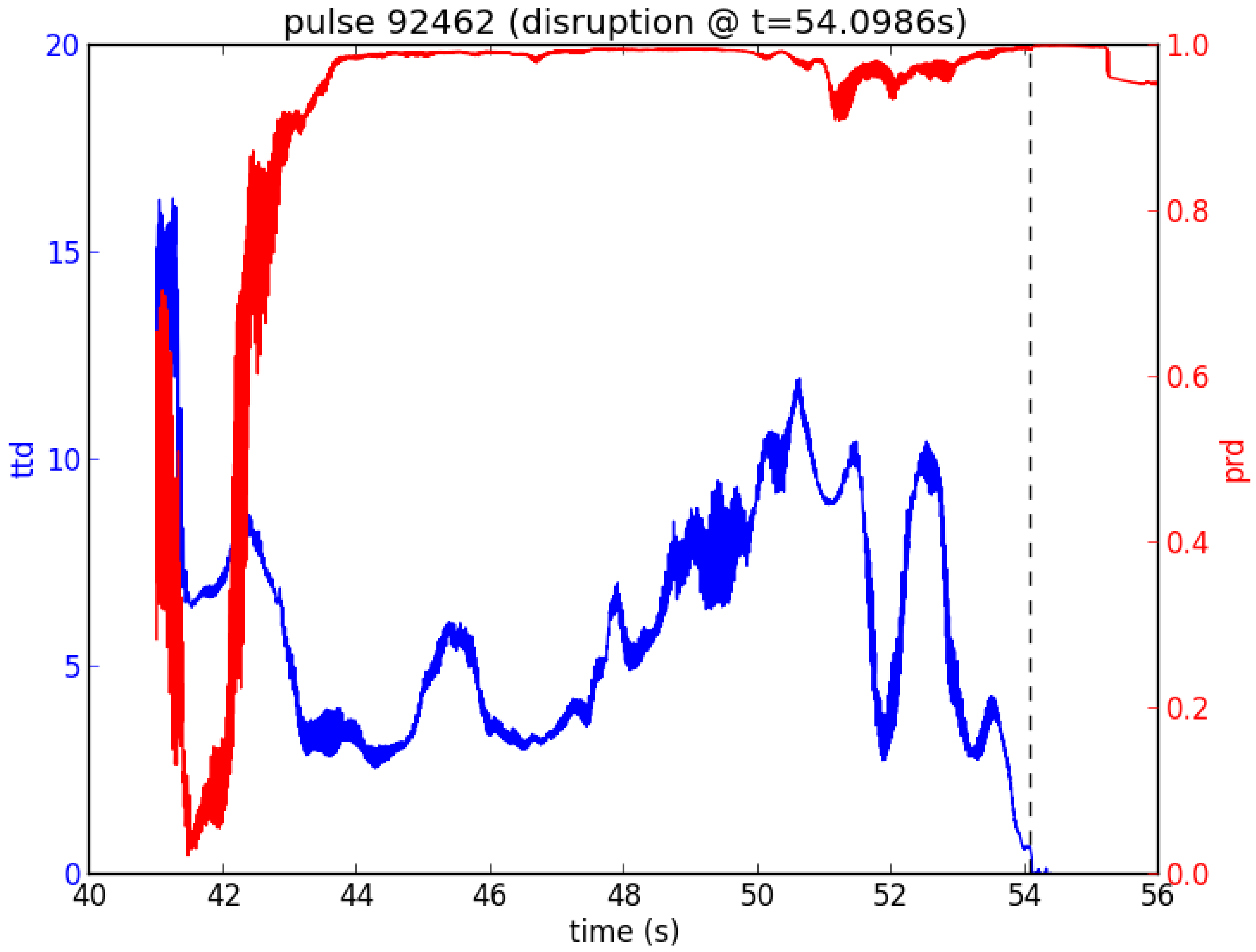
high ttd
high prd
or
high ttd
low prd

no disruption



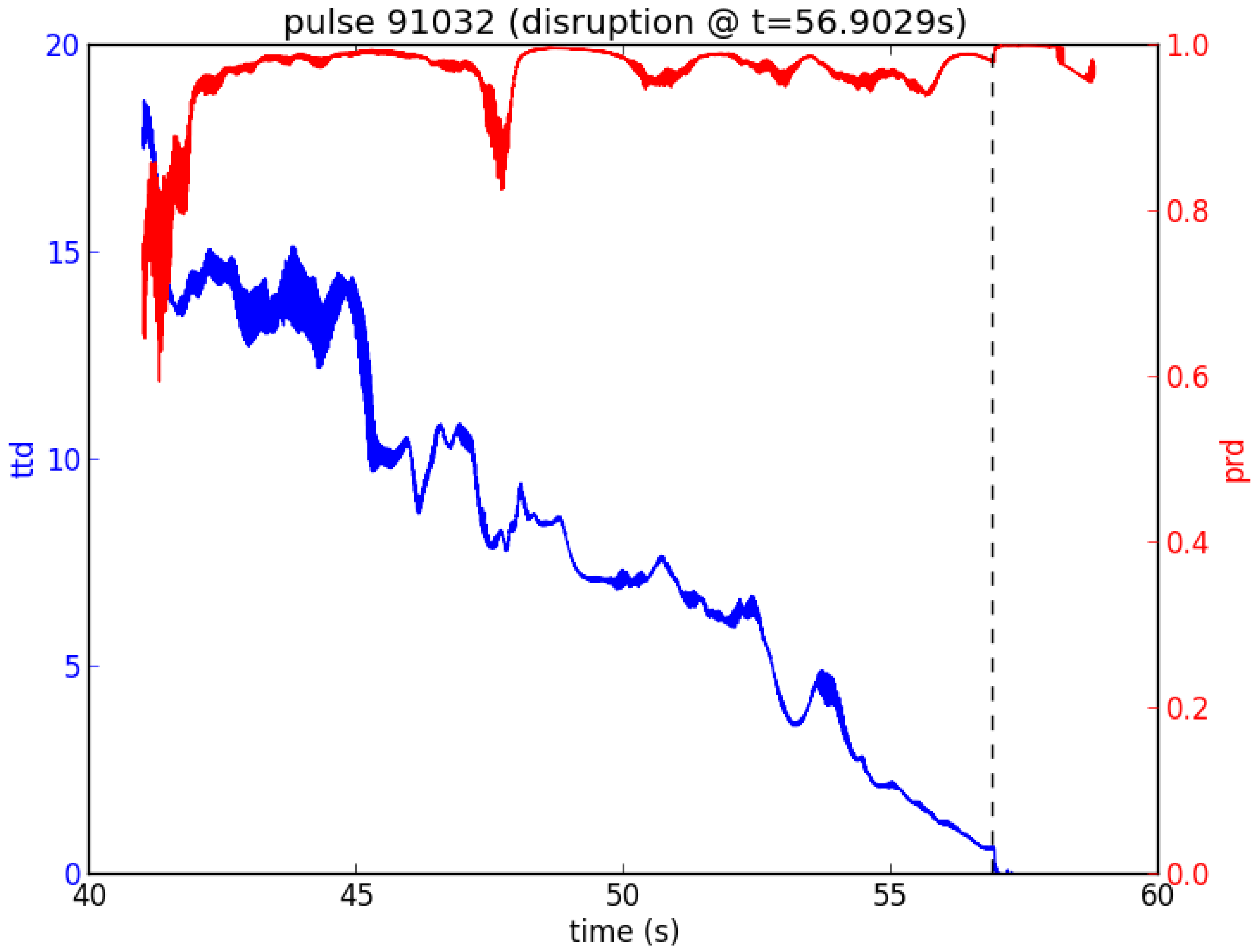
Results

early signs



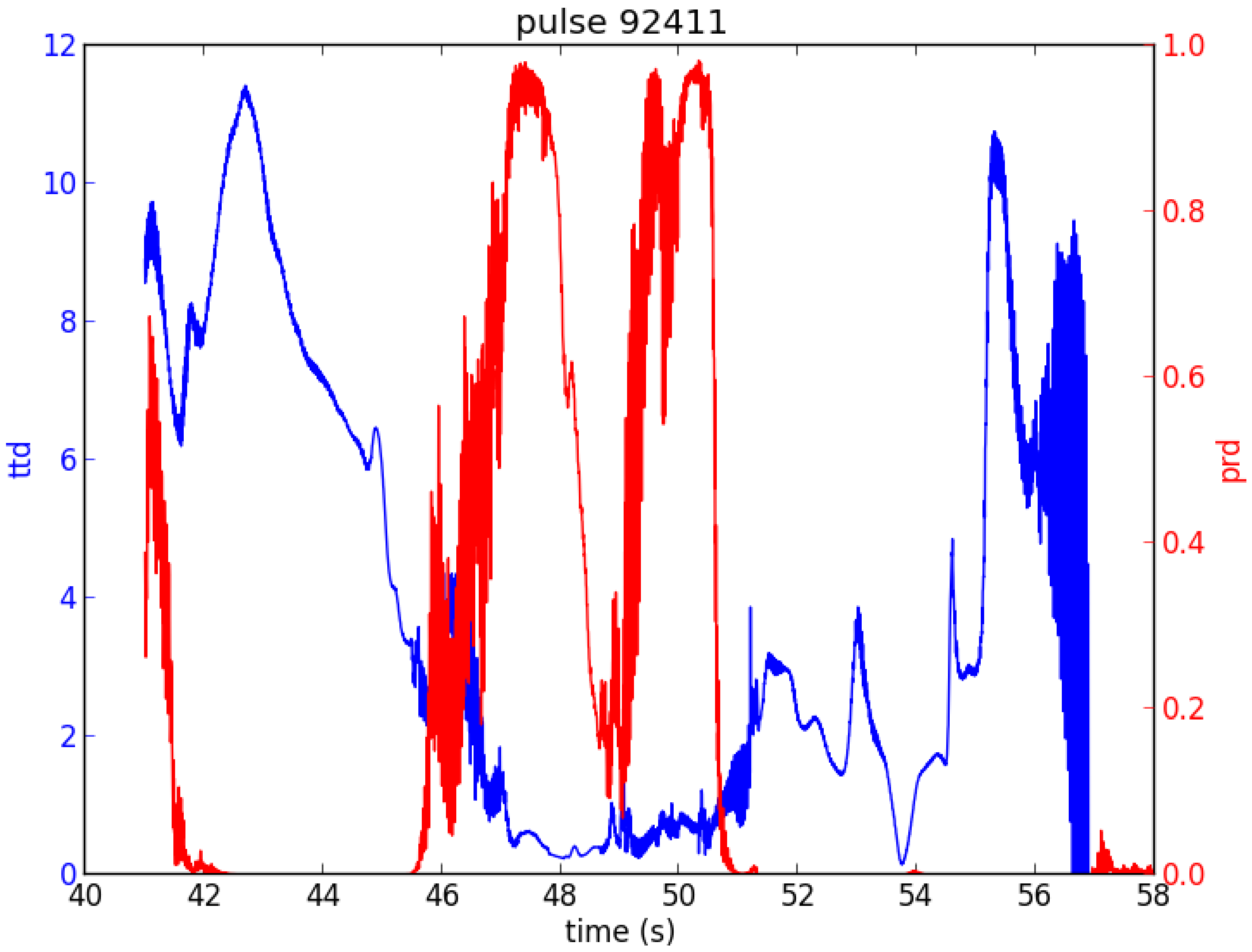
Results

early signs



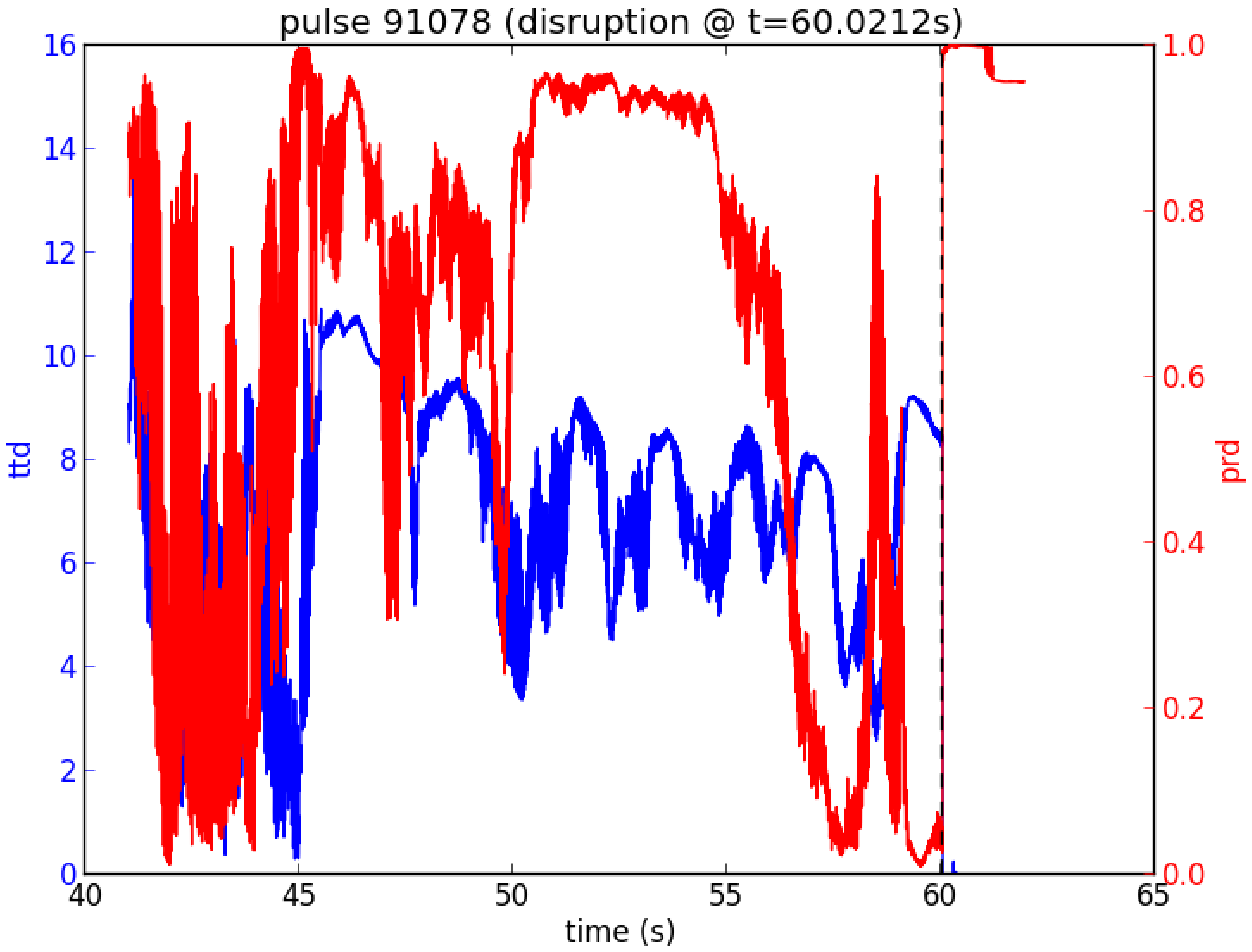
Results

false alarms
(false positives)



Results

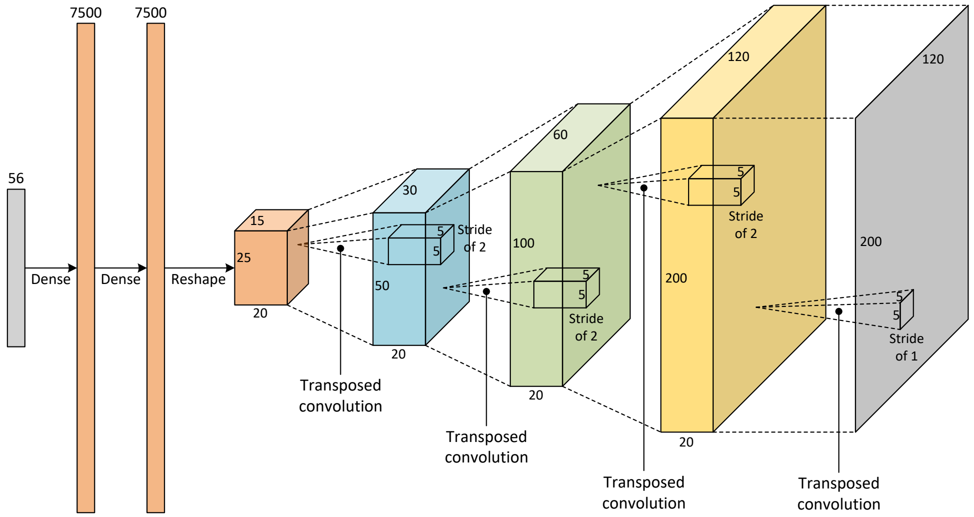
missed alarms
(false negatives)
very rare



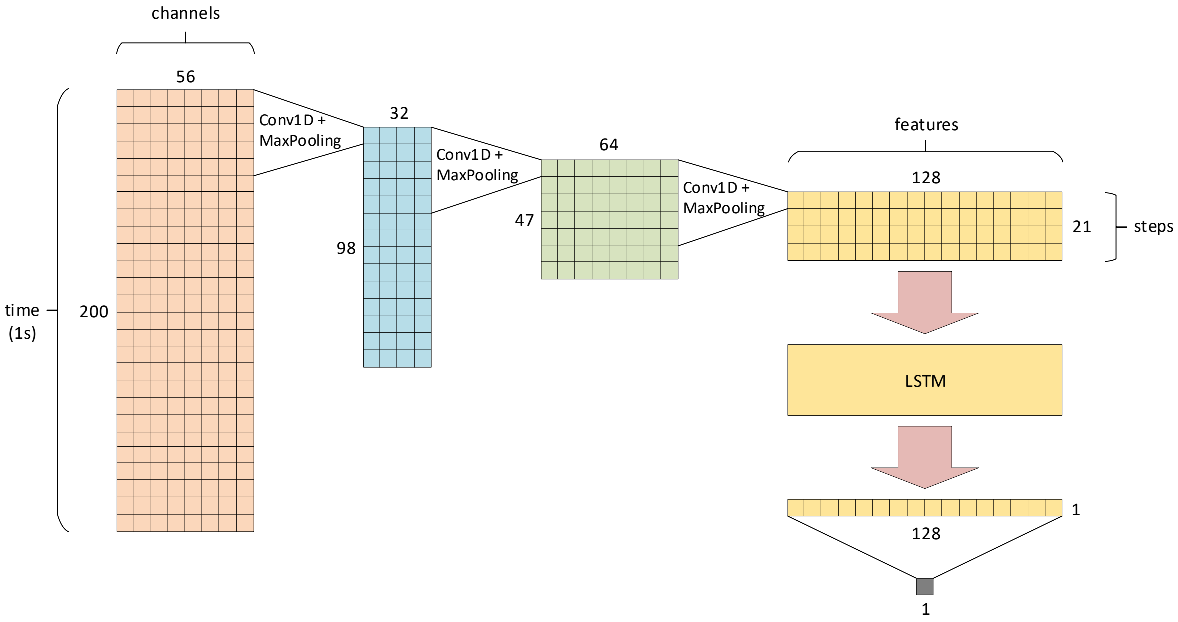
Conclusion

- Deep learning and the analysis of fusion data
  - replacing compute-intensive tasks (e.g. tomography)
  - support for tokamak operation (e.g. disruption prediction)
  - both post-processing and real-time processing

CNN for plasma tomography



RNN for disruption prediction



Pointers

- Source code
  - Plasma tomography
    - https://github.com/diogoff/plasma-tomography
  - Disruption prediction
    - https://github.com/diogoff/plasma-disruptions
- More info
  - Full-pulse Tomographic Reconstruction with Deep Neural Networks
    - https://arxiv.org/abs/1802.02242
  - Artificial intelligence helps accelerate progress toward efficient fusion reactions
    - https://www.pppl.gov/news/2017/12/artificial-intelligence-helps-accelerate-progress-toward-efficient-fusion-reactions
  - Princeton Team Using AI for Fusion Up for Global Impact Award
    - https://blogs.nvidia.com/blog/2018/03/05/ai-deep-learning-global-impact-awards-princeton/