CORTEX

Core monitoring techniques and experimental validation and demonstration

Detection and Localisation of Multiple In-Core Perturbations with Neutron Noise-Based Self-Supervised Domain Adaptation

SAINT Workshop AI/ML - 13th January 2021

Aiden Durrant, University of Aberdeen / University of Lincoln
a.durrant.20@abdn.ac.uk

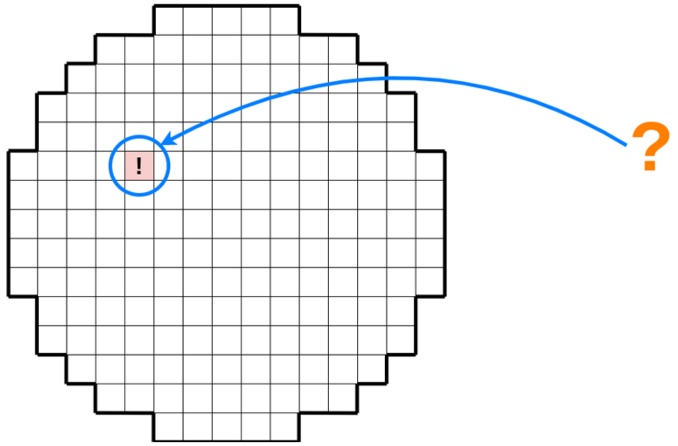
This project has received funding from the Euratom research and training programme 2014-2018 under grant agreement No 754316.

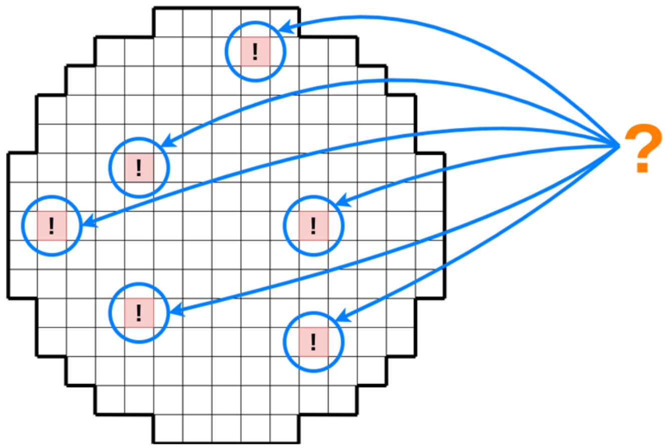
My Background

- PhD student at the University of Aberdeen
- Main research areas:
  - Machine Learning and Deep Learning theory
  - Unsupervised Representational Learning
  - Self-Supervised Learning
  - ML applied to nuclear energy
- Core Monitoring Techniques & Experimental Validation and Demonstration (CORTEX)

Outline / Questions

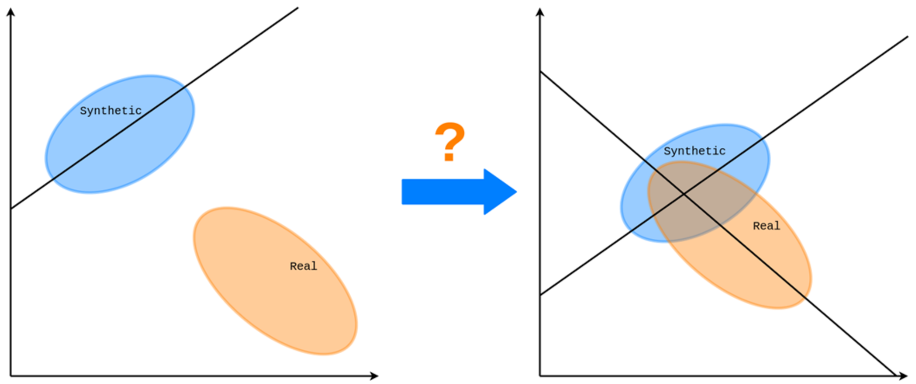
- 1) Can we unfold? Classify and Localise?





- 2) Multiple Simultaneously
  Occurring Perturbation?

- 3) How can we Leverage Simulated
  and Real Plant Measurements?



Problem Case

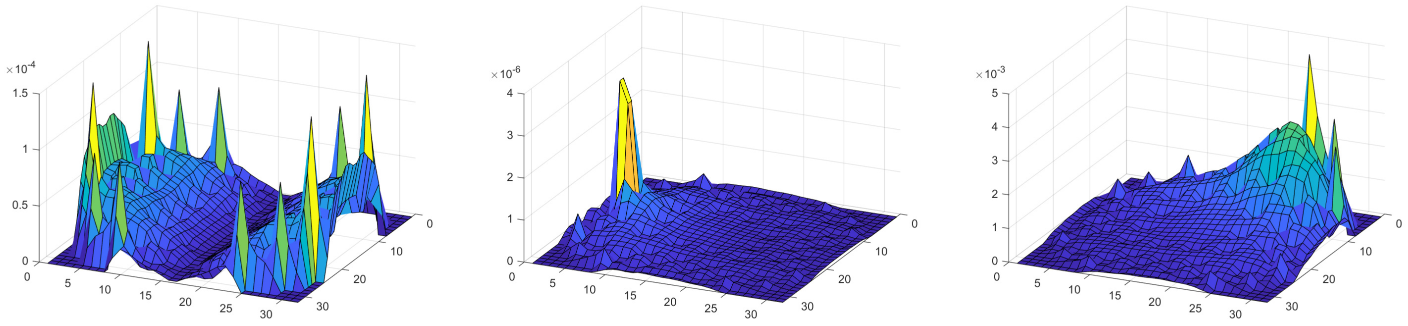
- We aim to unfold reactor transfer function to provide core diagnostics.
  - Derivation of core perturbation characteristics to **classify** and **locate** its origin.
- Yet this is challenging due to the limited number of neutron detectors in western type reactors.
- We ask, *can we use machine learning to successfully approximate the reactor transfer function?*
- However, to effectively train ML algorithms large quantities of data are required.

Data Acquisition

- Real plant measurements are difficult to obtain, unlabelled, and anomalies are thankfully rare.
- As such it is beneficial to have alternative means to train our algorithms.
    - Supervised Learning – Data where a known perturbation type and source is assumed.
- We utilise a diffusion-based core simulation tool to provide simulated training data that is both labelled and capable of producing any theoretical perturbation scenario.
    - CORE SIM + (A. G. Mylonakis, P. Vinai, and C. Demaziere. "Numerical solution of two-energy-group neutron noise diffusion problems with fine spatial meshes." Annals of Nuclear Energy,volume140, p. 107093 (2020).)

Data Prerequisites

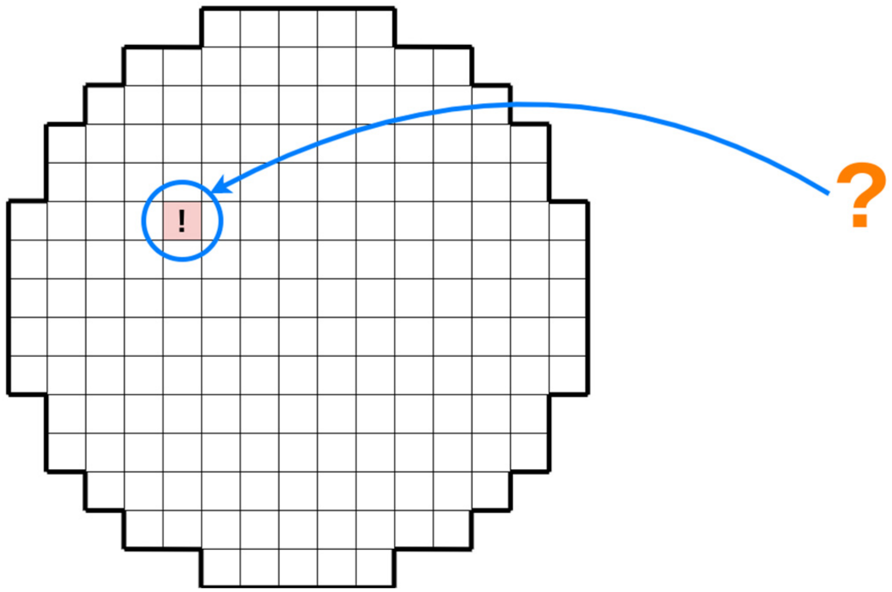
- CORESIM + produces neutron response to a induced perturbation in the frequency domain for the whole core volume.



- Yet to align with real plant measurements we only use a small number of these readings to corresponds to the neutron detectors. (48 readings from the 32x32x32 simulated volume)
- There are 9 different perturbation scenarios, each being simulated for all theoretically possible origins = Terabytes of Data !

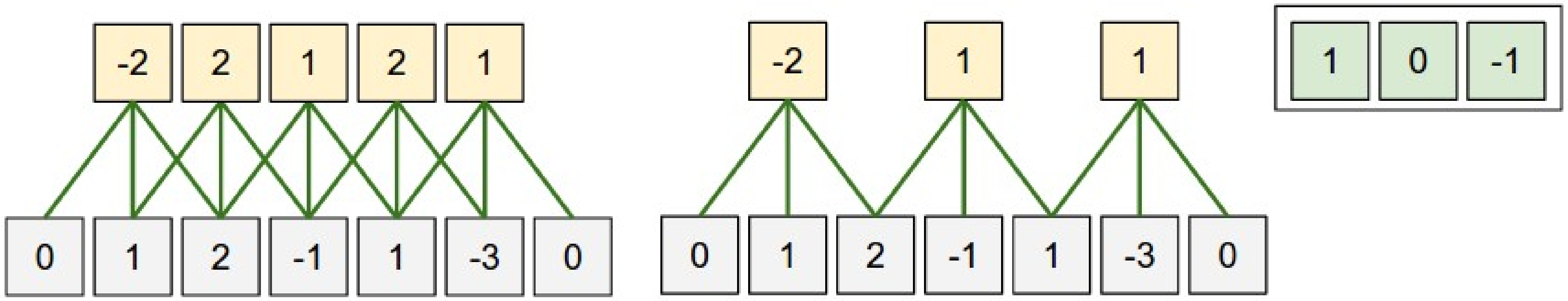
Unfolding the Reactor Transfer Function

- Single Perturbation Classification and Localisation



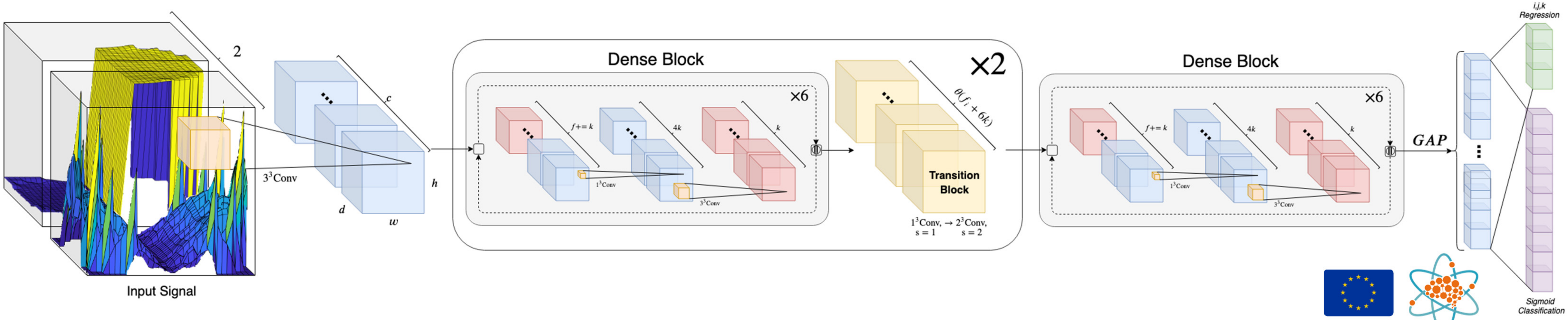
3D Convolutional Neural Network (CNN)

- Given our goal is to determine the origin within a volume, it is desirable to train our network to learn spatial features of the response.
- Inspired by medical imagine and specifically MRI / CAT scans we employed the use of 3D Convolutional Neural Networks (CNNs) to operate on the whole core volume.
- Convolutional Neural Networks (CNNs) :



3D Densely Connected CNN (1)

- The complexity of the problem and the limited detectors required deep very deep and powerful network to adequately parameterise the problem.
- A 3D extension to DenseNet was proposed.
- Dense connections allow for the greater flow of gradients / information through the network, resulting in the passing of collective knowledge.



3D Densely Connected CNN (2)

- To classify and localise, the network outputs a representational feature vector to two fully-connected (Multi-Layer Perceptron) output layers.
  - 3 Neuron - >Coordinate Regression
  - 9 Neuron Softmax Non-linear -> Classification,
- Given the two tasks we use a multi-task objective function to simultaneously train the network for both tasks.

$$\mathcal{L}(\mathcal{D};\mathbf{W},\lambda_1,\lambda_2) = -\frac{1}{M}\sum_{i=1}^{M}\Bigg[\underbrace{\frac{\lambda_1}{P}\sum_{j=1}^{P}\left[\mathbf{y}_j\log(\hat{\mathbf{y}}_j)+(1-\mathbf{y}_j)\log(1-\hat{\mathbf{y}}_j)\right]}_{\text{Classification}} \underbrace{-\frac{\lambda_2}{C}\sum_{c=1}^{C}\left\|\mathbf{y}_c-\hat{\mathbf{y}}_c\right\|^2}_{\text{Regression}}\Bigg]_i,$$

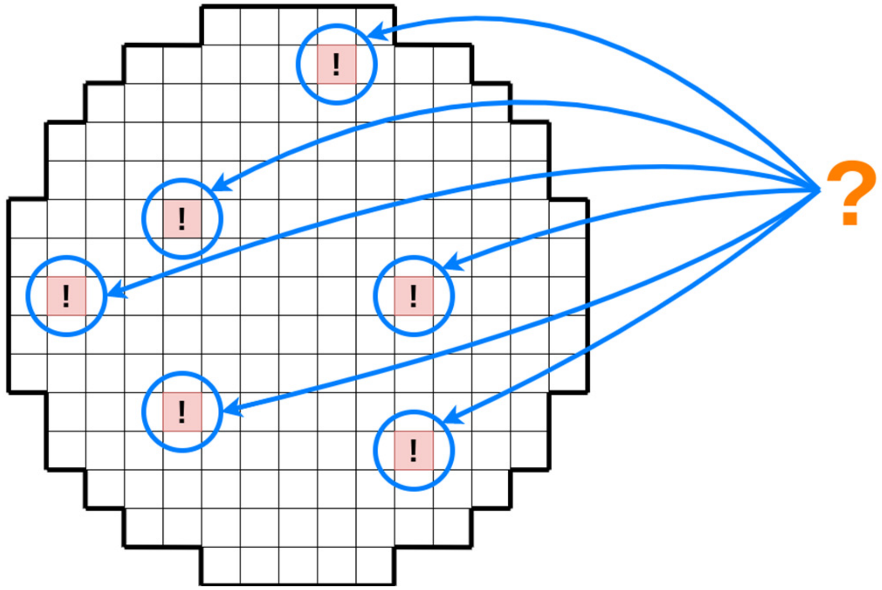
Results (Frequency Domain)

- Classification of 9 perturbation classes = **0.11% Error.**
- Localisation of the perturbation source = **Mean Absolute Error 0.2902**

  (Average error in coordinate prediction to target.
  Approx. 4cm error in (4m x 4m x 4m) volume).
- Our network performed extraordinarily well, exceeding our expectations.

Single Perturbation					
Sensors (SNR)	Train / Valid / Test (%)	Classification		Regression *(I,j,k)*	
		Error (%)	F1-Score	MAE	MSE
All	70 / 15 / 15	0.06±0.051	0.9344±0.004	0.1435±0.011	0.0342±0.008
48 In-Core	70 / 15 / 15	**0.11±0.010**	**0.9311±0.001**	**0.2902±0.011**	**0.3072±0.014**
48 In-Core	25 / 15 / 60	0.32±0.025	0.9149±0.002	0.3978±0.017	0.6407±0.052
48 In-Core	15 / 25 / 60	0.44±0.061	0.9141±0.003	0.4858±0.017	0.7727±0.006
48 In-Core (3)	70 / 15 / 15	0.15±0.006	0.9231±0.001	0.3456±0.016	0.4905±0.011
48 In-Core (1)	70 / 15 / 15	0.19±0.036	0.9225±0.002	0.3709±0.020	0.5185±0.017

Semantic Segmentation
- Multiple, Simultaneously Occurring Perturbation Classification and Localisation

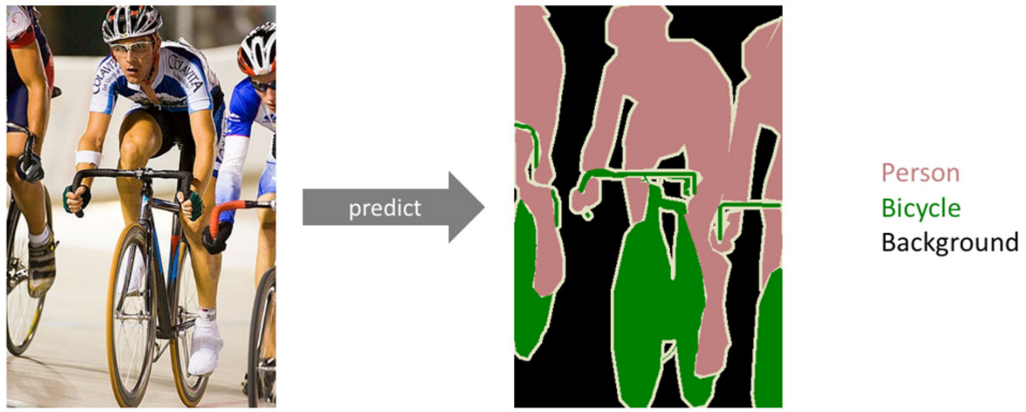


Beyond One Perturbation!

- In reality **perturbations rarely occur in isolation**, instead they are found as multiple perturbations occurring simultaneously.
- How can we develop architectures that **computationally scale** well to the large scale datasets and different reactors?
- Importantly how can we make an **arbitrary number of predictions** per sample that change between samples?

Semantic Segmentation

- Semantic segmentation is a methodology for the “linking” of each pixel in an input sample to a semantic (class) label.

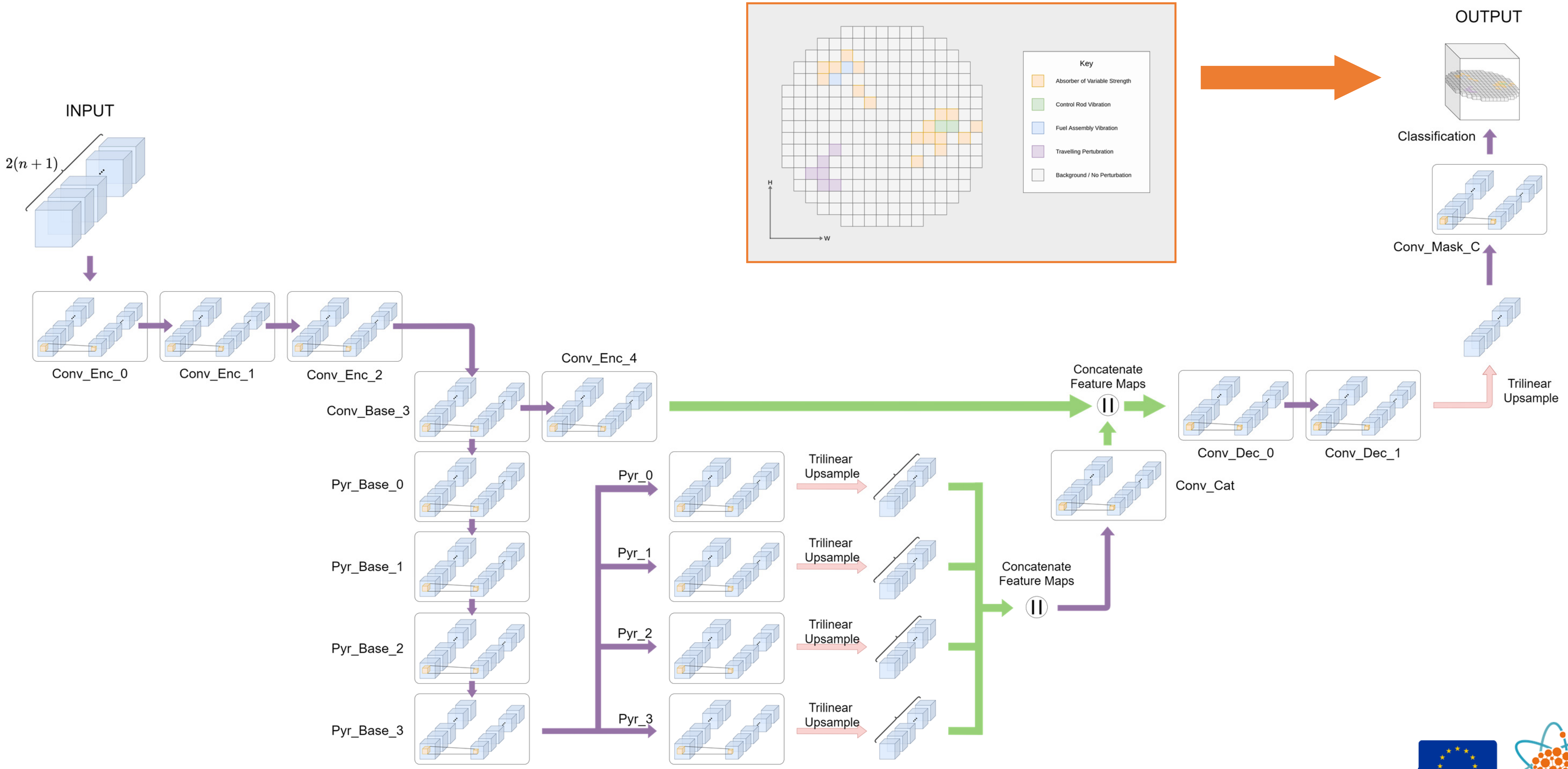


- We can use this methodology to “link” the response to multiple driving perturbations (the image) to semantic label mask (classification) where the position of that semantic label in the image represents its origin location.

Voxel-Wise Semantic Segmentation (1)

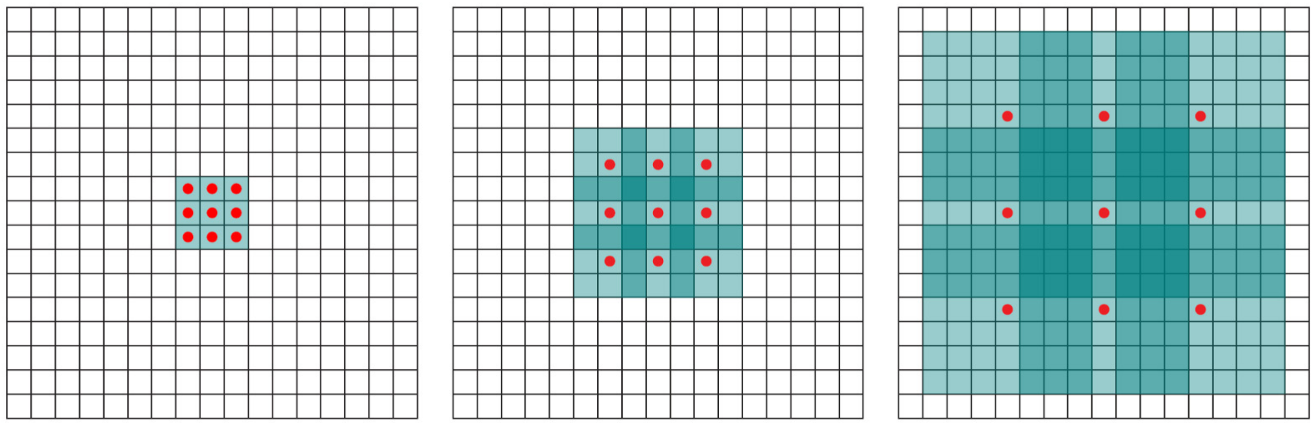
- Each voxel in the output represents an origin location of a driving perturbation, the classification of a voxel represents that a driving perturbation of the identified scenarios is present.
- Leading from our previous work, we employ a 3D Fully-Convolutional Neural Network.
- Specifically we employ an encoder-decoder architecture.
  - Encoder: Produces a feature representation matrix of the input that encodes the information describing the input.
  - Decoder: Takes this feature matrix and decodes to produce a prediction the same dimensions as the input.

Voxel-Wise Semantic Segmentation (2)



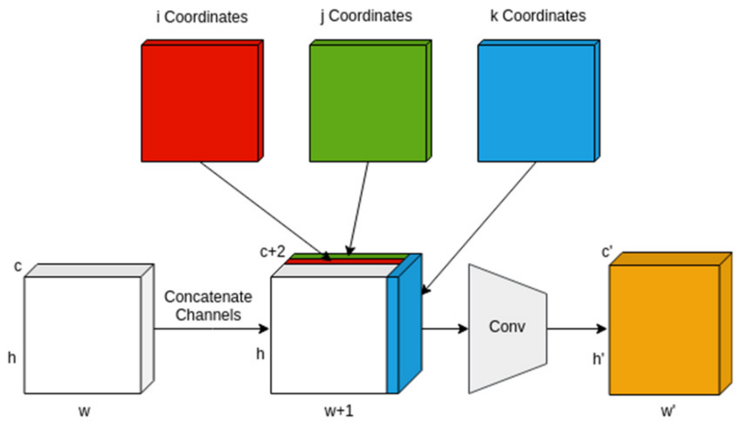
Voxel-Wise Semantic Segmentation (3)

- Our network uses some additional ‘tricks’ to assist in our task:
  - Dilated Convolutions



  - Coord Conv

-1 , -1	-1 , 0	-1 , 1
0 , -1	0 , 0	0 , 1
1 , -1	1 , 0	1 , 1



  - Strided Convolutions

Voxel-Wise Semantic Segmentation (4)

- Our major challenge lied with class imbalance, we have a large volume (34x34x34) with only a relatively small number of present perturbations.
- Focal loss helps, it gives more weight for "hard-to-classify examples".
- We train the network to minimise the average categorical focal loss of every voxel in the mask to the ground truth (the true source location of the simulated perturbation).

$$FL(y,\hat{y}) = -\frac{\lambda_1}{P}\sum_{p=1}^{P}\left[y_p(\alpha_p(1-\hat{y}_p)^{\gamma}\log(\hat{y}_p)) + (1-y_p)((1-\alpha_p)\hat{y}_p{}^{\gamma}\log(1-\hat{y}_p))\right]$$

- We also utilise a logarithmic class weighting scheme to the focal loss to reduce the impact of perturbation classification imbalance.

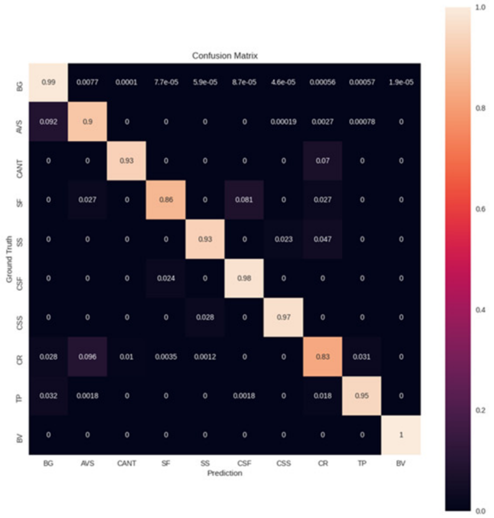
Data and Experimental Setup

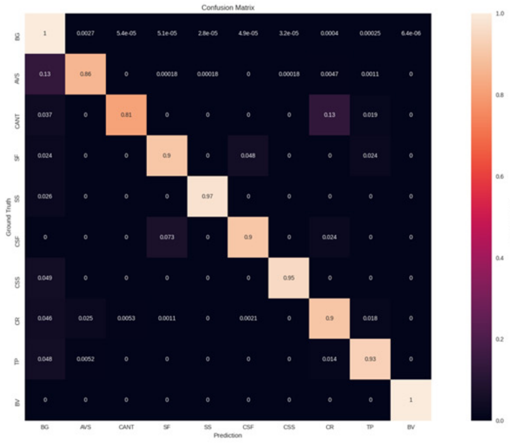
- As previously mentioned we employed CORE SIM + to produce our large simulated dataset.
- We additively combine a random number of individual driving perturbations data samples within a range [1, x] where x = {15, 30, 45}.
- We produce 500,000 combined samples per set (15, 30, 45)
- Additionally, CORE SIM+ models two different PWR reactors that correspond to conditions found in real plant measurements.
- We demonstrate the performance of our network under different values for max. combinations for different reactors.

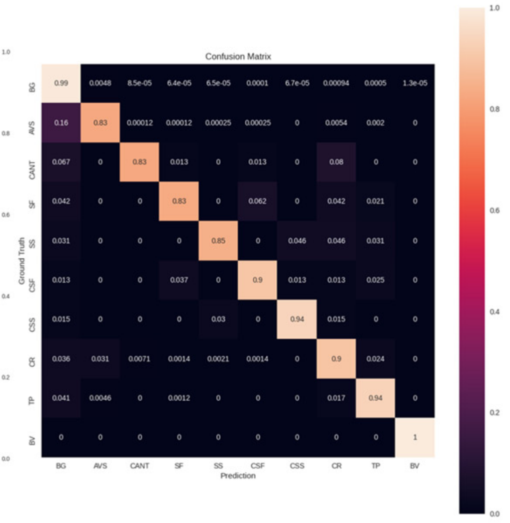
Results (German pre-KONVOI)

- Strong classification of source perturbations at their originating assemblies.
- The highest perturbation classification error is attributed to AVS, we conjecture due to the more varied source locations of our combination procedure.

Per Class Voxel Prediction Accuracies *											
No. Comb	No. Det	Accuracy (%)									
		BG	AVS	CANT	SF	SS	CSF	CSS	CR	TP	BV
15	56	99.08	90.47	92.98	86.49	93.02	97.62	97.22	83.06	94.74	100.00
30	56	99.64	85.97	81.48	90.48	97.37	90.24	95.12	90.21	93.25	100.00
45	56	99.35	82.28	88.00	87.50	89.23	90.00	92.42	88.99	93.20	100.00





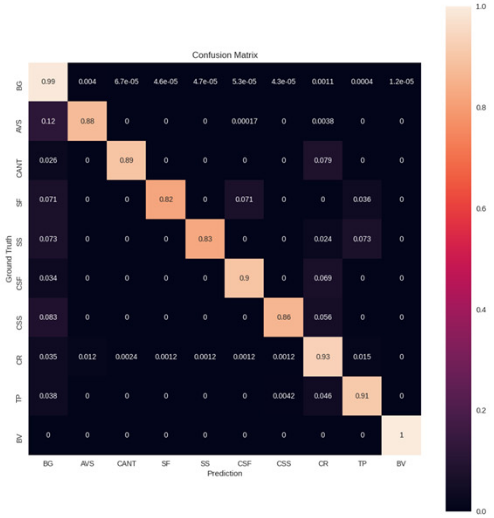


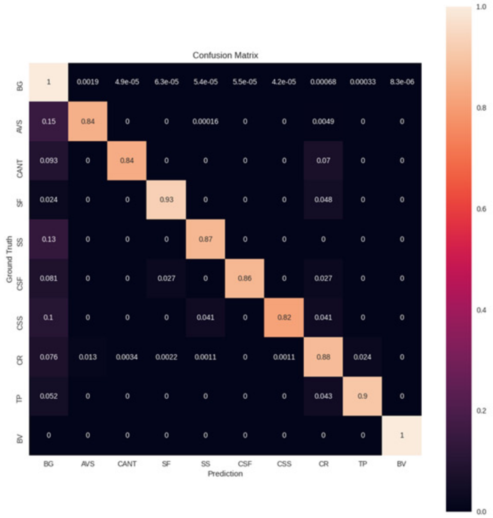
Results (Swiss pre-KONVOI)

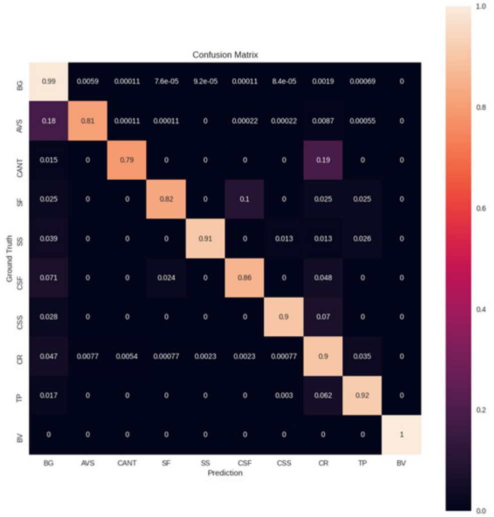
- The majority of erroneous results come from False Positive identification, around the location of the true perturbation.

- The performance drop is somewhat expected from the German pre-KONVOI, due to increased size of the reactor mesh volume.

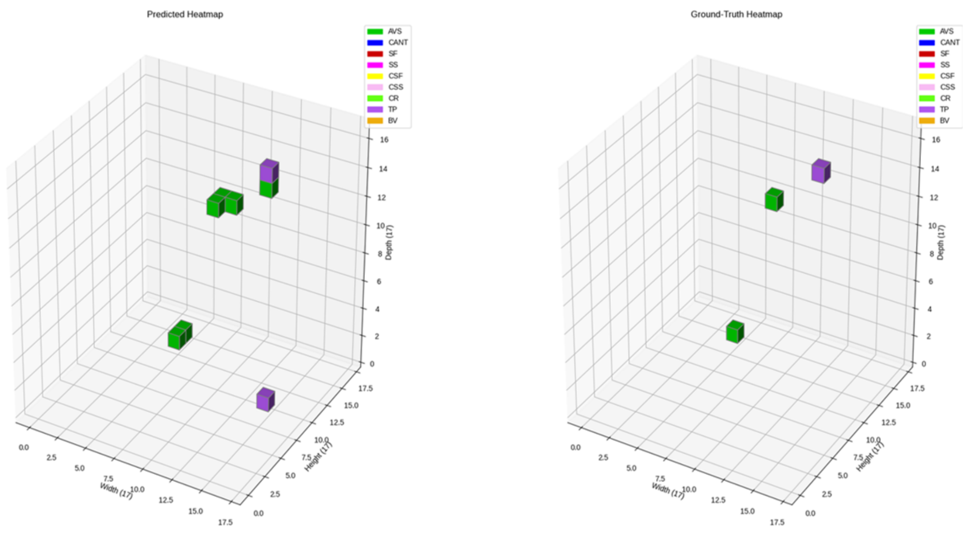
Per Class Voxel Prediction Accuracies *											
No. Comb	No. Det	Accuracy (%)									
		BG	AVS	CANT	SF	SS	CSF	CSS	CR	TP	BV
15	56	99.43	87.64	89.47	82.14	82.93	89.66	86.11	93.05	91.16	100.00
30	56	99.68	84.45	83.72	92.86	86.54	86.49	81.63	87.85	90.48	100.00
45	56	99.11	80.95	79.41	82.50	90.79	85.71	90.14	89.86	91.81	100.00

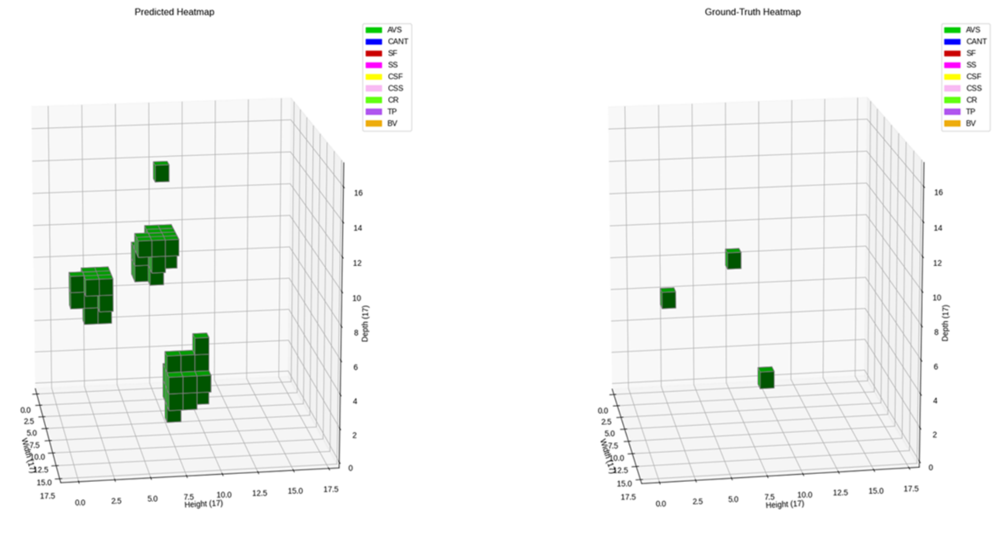




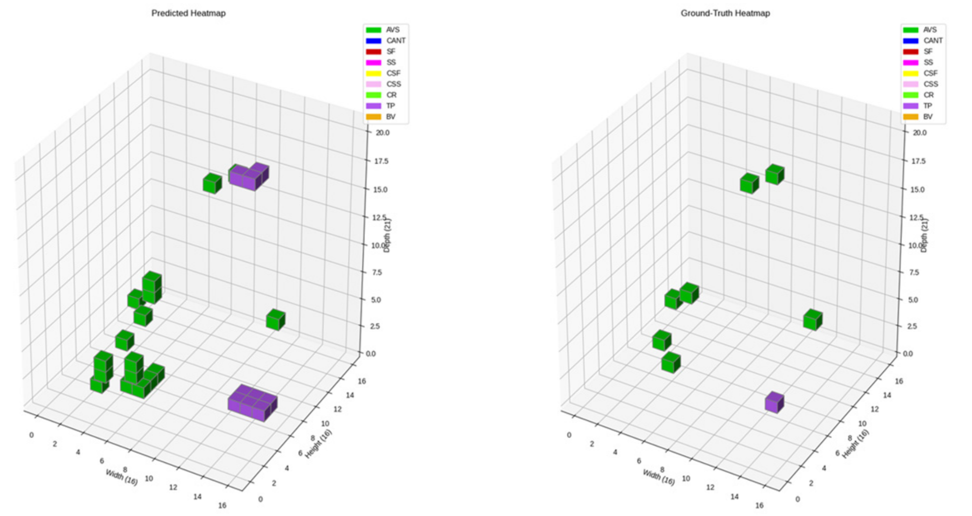


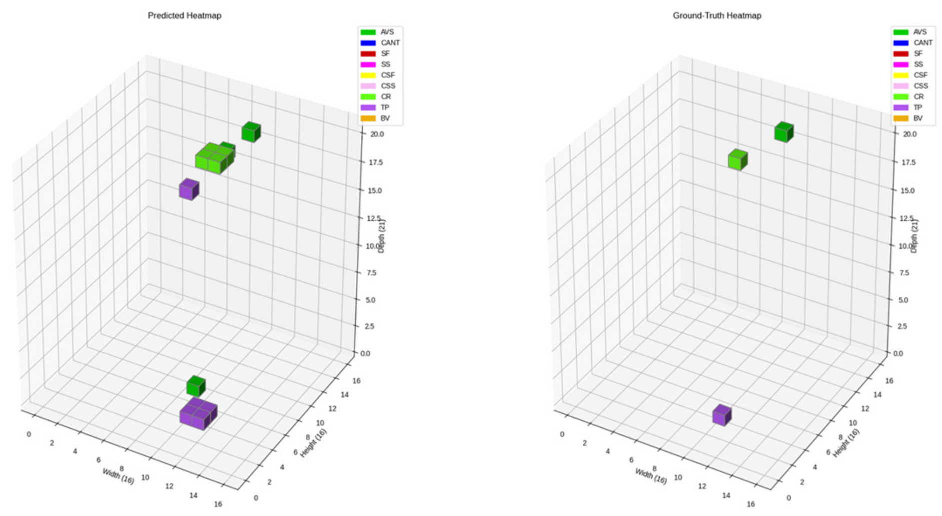
Example Prediction Masks





(German pre-KONVOI)





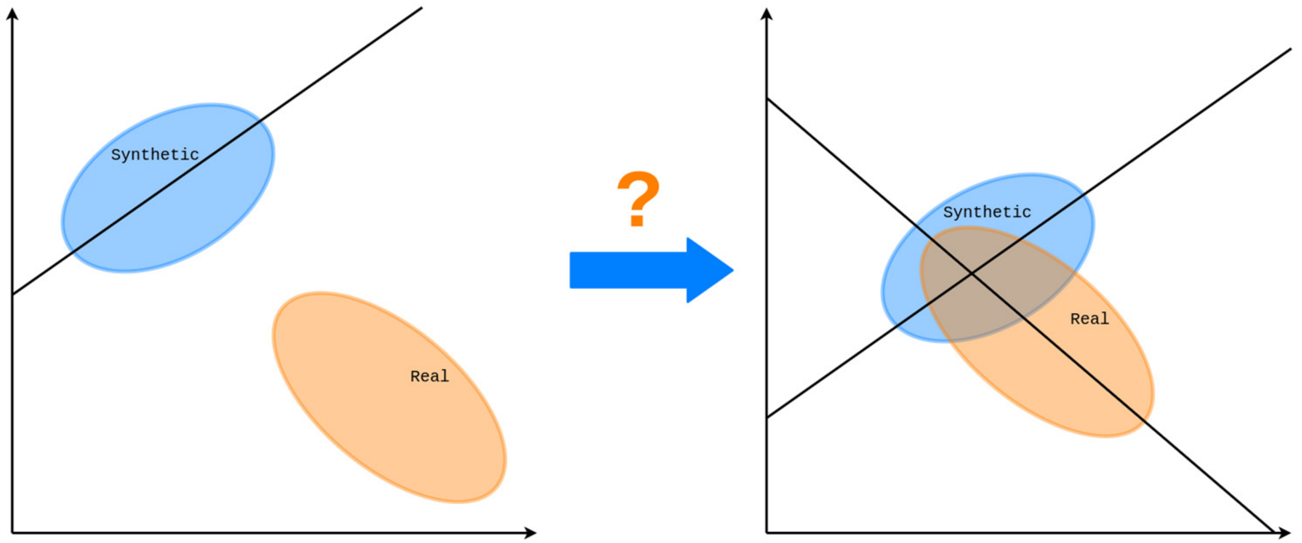
(Swiss pre-KONVOI)

Conclusions

- We provide a technique to perform the novel task to accurately (in the simulated case) classify and subsequently localize many simultaneously occurring perturbations via noise diagnostics.
- Our network requires very little additional reactor information to make these strong predictions.
- Large networks are required, and the computational effort is large.
- This approach is scalable and transferable to many reactor types as shown.
- Further work is being done to extend our approach into uncertainty prediction, and the determination of more perturbation characteristics.

How can we Leverage Synthetic and Real Plant Measurements?

- Self-Supervised Domain Adaptation
- Synthetic to Real Adaptation



Let’s Get Real!

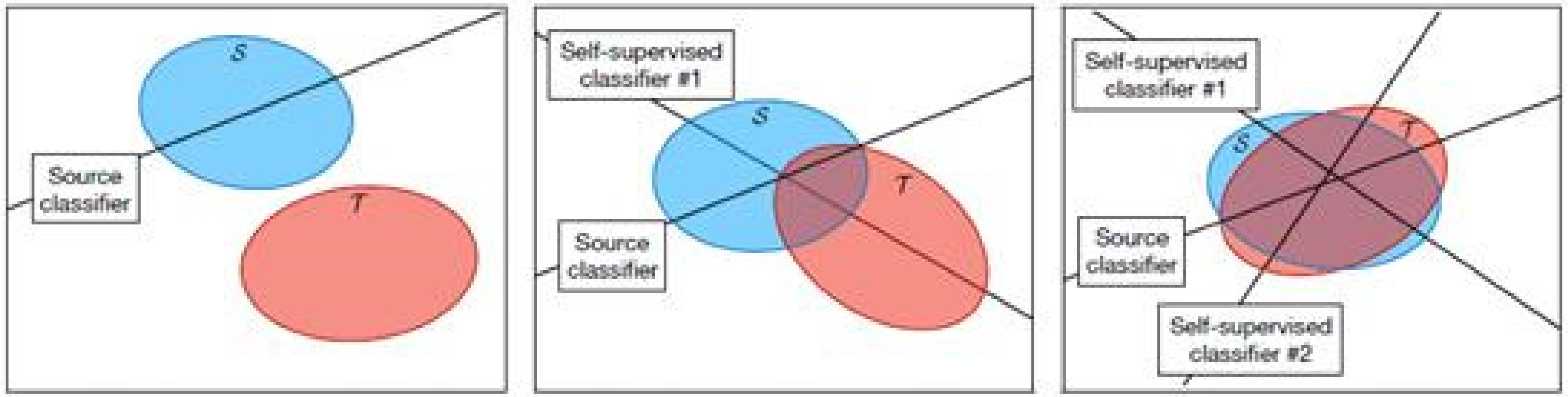
- Can we just make predictions on the real plant measurements from the network trained on simulated data?
- **Real plant data is not annotated** (unsupervised), how can we leverage the annotated simulated data that is abundant and provides clear perturbations distinctions?
- Real plant data, although modelled by the simulations, **contains some inherent differences to simulated data**, how do we minimise these differences as not to confuse our trained network?

Unsupervised Domain Adaptation

- We aim to learn a discriminative classifier (our voxel-wise semantic segmentation network) for classifying perturbations that is invariant to the presence of a domain shift from simulated to real data.
- We have no annotations in the real plant measurements so we need a method to align these different domains without semantic information rather we need to find common features across domains.
- Therefore, we opt to train our classifier to align the two domains in some shared feature space represented by the discriminative model through the process of solving a common auxiliary task that are constructed from the data itself (self-supervised learning).

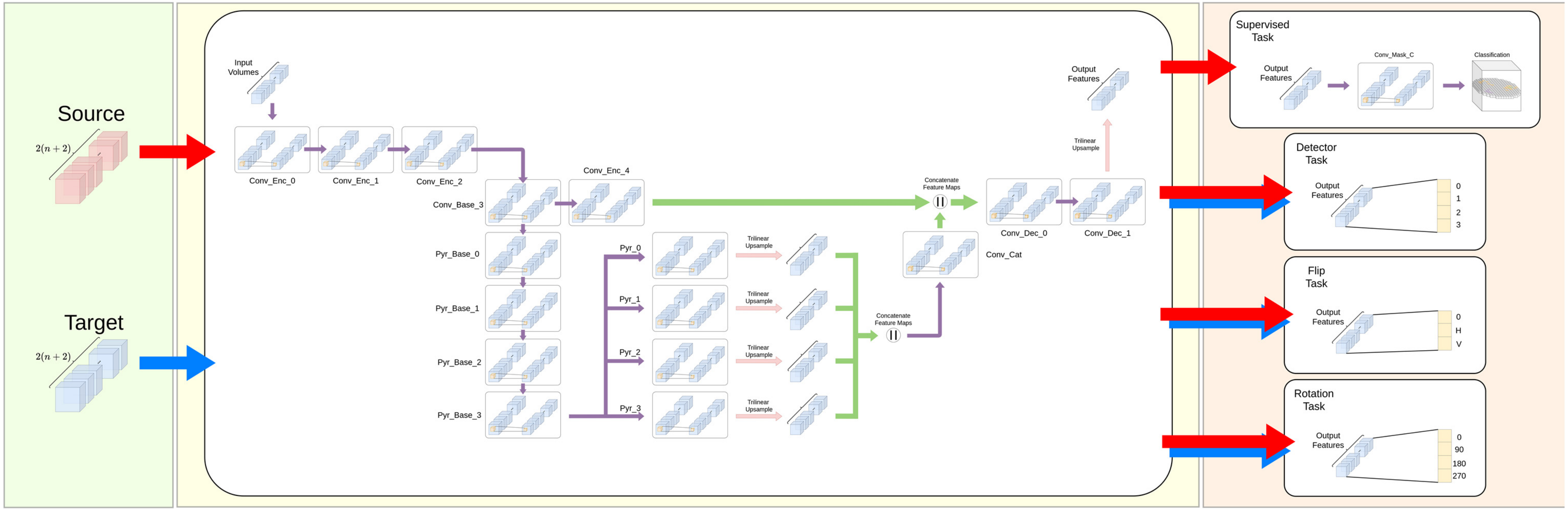
Self-Supervised Domain Adaptation (1)

- Auxiliary tasks are constructed from the input, providing feature understanding of structurally relevant info whilst not requiring annotation.
- These tasks encourage alignment between the distribution of features captured in both the simulated (S) and real measurements (T) domains.
- The feature extractor predict identical augmentations for each input, enforcing invariance to the nuances displayed between distributions.



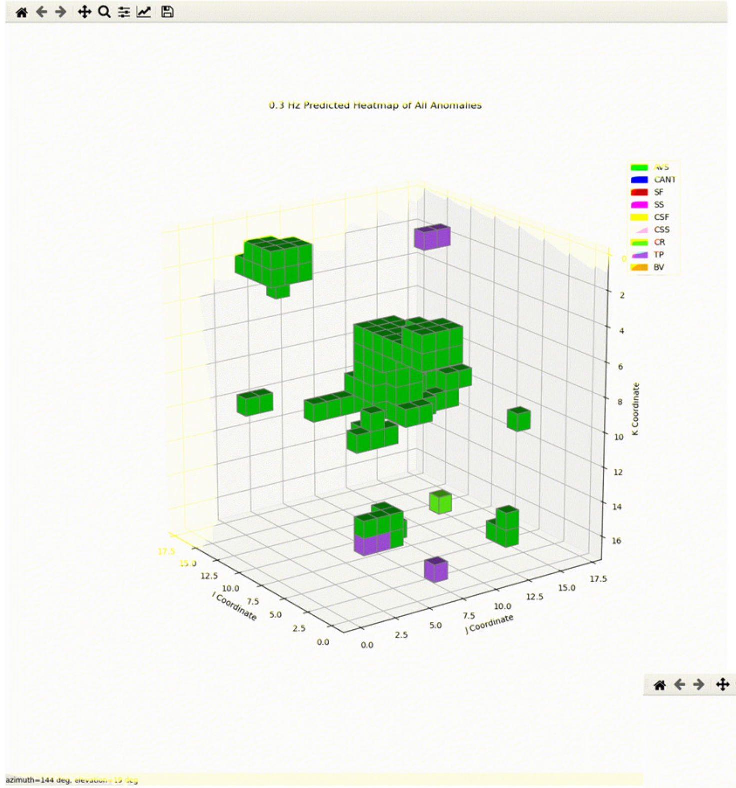
(a) Source classifier only (b) adding a self-supervised task (c) adding more tasks

Self-Supervised Domain Adaptation (2)



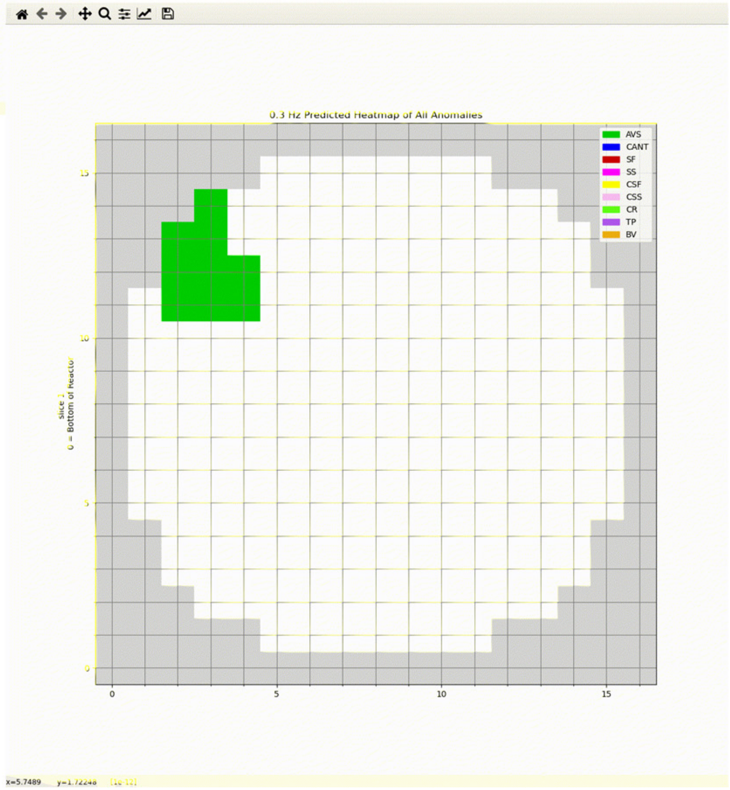
Results

- Very initial results, as we have no assured validation … Yet
- Positively, we identify vertically transporting phenomena which is also identified in signal processing analysis.
- Additionally, we observe a vertical column of AVS inline with this transport phenomena.



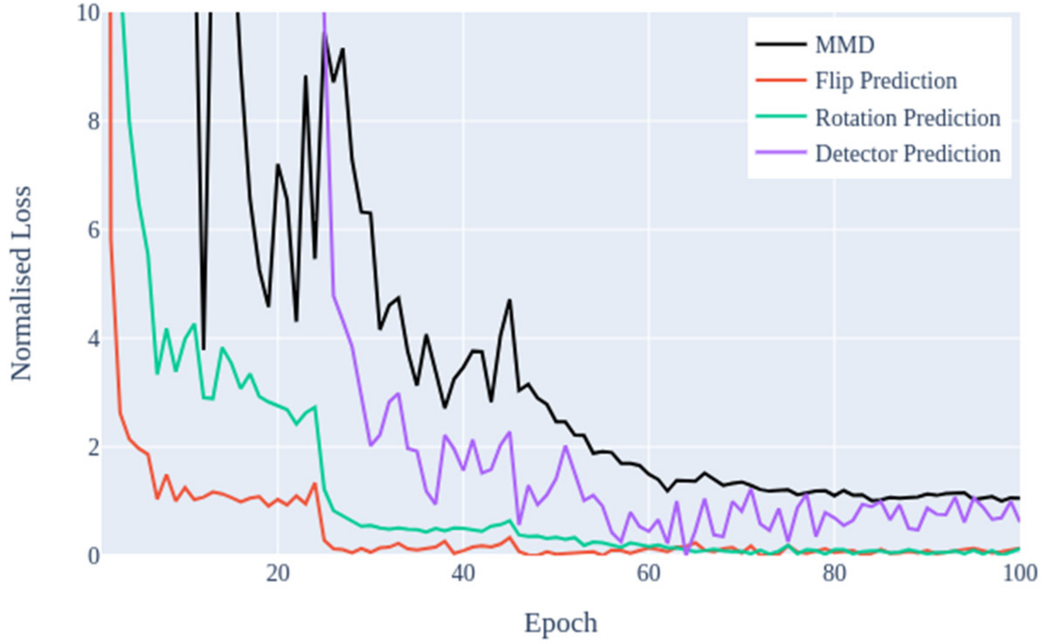
German
4-Loop pre-KONVOI

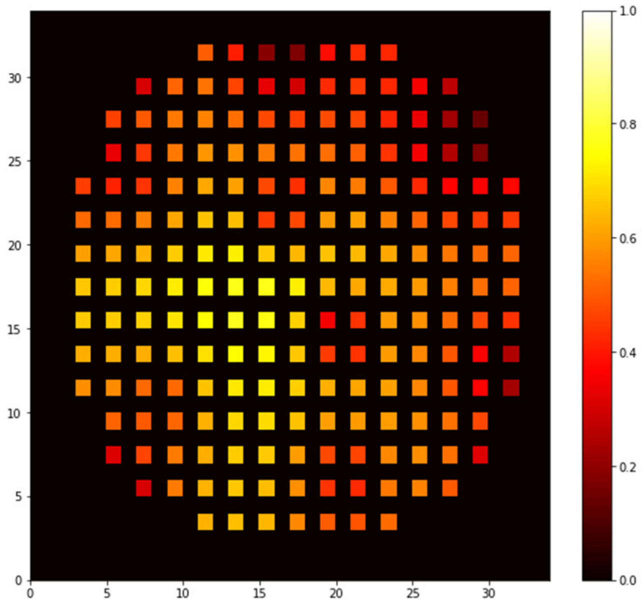
Frequency = 0.3 Hz



Results

- The MMD between the synthetic and real domains in feature space is reduced during training.
- Such convergence shows the network is reducing the distance between domains in feature space empirically showing alignment.
- Our results are further verified by other works within the project in which an unsupervised approach reports similar phenomena.





(NTUA – Later Talk)

Conclusions

- We provide a methodology to leverage both domains of data, simulated and real.
- Our model uses structurally relevant information inherent in both domains to find common features.
- This approach does not require extra-human annotation yet can use the large labelled datasets and align to the nuances of real data to get a more accurate result.
- The future is explainable results!

Thank you & Questions?

Email: a.durrant.20@abdn.ac.uk

LinkedIn: https://www.linkedin.com/in/aidendurrant/