3D CONVOLUTIONAL AND RECURRENT NEURAL NETWORKS FOR REACTOR PERTURBATION UNFOLDING AND ANOMALY DETECTION

A. DURRANT, G. LEONTIDIS, S. KOLLIAS
University of Lincoln, School of Computer Science, Machine Learning Group
Brayford Pool, Lincoln LN6 7TS – United Kingdom

ABSTRACT

With Europe's ageing fleet of nuclear reactors operating closer to their safety limits, the monitoring of such reactors through complex models has become of great interest to maintain a high level of availability and safety. Therefore, we propose an extended Deep Learning framework as part of the CORTEX Horizon 2020 EU project for the unfolding of reactor transfer functions from induced neutron noise sources. The unfolding allows for the identification and localisation of reactor core perturbation sources from neutron detector readings in Pressurised Water Reactors. A 3D Convolutional Neural Network (3D-CNN) and Long Short-Term Memory (LSTM) Recurrent Neural Network (RNN) have been presented, each to study the signals presented in frequency and time domain respectively. The proposed approach achieves state-of-the-art results with the classification of perturbation type in the frequency domain reaching 99.89% accuracy and localisation of the classified perturbation source being regressed to 0.2902 Mean Absolute Error (MAE).

1. Introduction

The early detection, classification, and localisation of anomalies within the reactors' core is vital to ensure the safe and efficient operation of the increasingly aging fleet of Europe's reactors. Monitoring of these reactors at nominal conditions provides vital and valuable insights into the functional dynamics of the core, consequently allowing for early identification of anomalies. Analysis of the core operation is achieved through non-intrusive measuring of neutron flux around their mean values from in-core and ex-core detectors. These fluctuations more commonly referred to as noise are induced primarily from turbulent characteristics in the coolant flow in the core, coolant boiling, or mechanical vibrations of reactor's internal components.

Given detailed descriptions of the reactor core geometry, properties of physical perturbations, and probabilities of neutron interactions, by using a Green's function as the reactor transfer function, simulations can be constructed to show the effect of the neutron noise. Green's function holds the relationship between a locally induced perturbation and the response of the neutron flux within the core, therefore, the inversion of this function from noise readings can localise and classify such induced perturbations. This inversion known as the *backwards problem* or *unfolding* is trivial given measurements at every position within the core, however, the limited number of in-core and ex-core detectors makes it a complex challenge [1].

Machine learning (ML) is a data analytical process for the approximation of functions mapping a set of inputs to outputs. Therefore, the use of ML to approximate such reactor functions given limited detector readings is advantageous, learning high and low-level patterns given substantial training examples. This work presents an extended 3D-Convolutional and Recurrent neural network approach to unfold the reactor transfer function and classify induced perturbation types and their source locations in both time and frequency domains.

2. Related Work

The application of ML approaches in the field of nuclear safety has been of recent scientific interest, with nuclear energy essential to meeting fast changing climate goals. The ML community has been keen on predicting climate change [2] utilising a variety of approaches across all energy sectors. Nuclear energy relies on safety and availability to achieve such goals, and many recent works have been proposed to ensure this.

In [3] the authors utilised deep convolutional neural networks and Naïve-Bayes approaches for vision-based crack detection for reactor component surfaces from video sequences. A diagnosis system monitoring the condition of sensors using auto-associative kernel regression and sequential probability was proposed in [4]. Deep rectifier neural networks were implemented in [5] for the accident or transient scenario identification of pressurised water reactors (PWR), whereas others solved similar problem employing artificial neural networks improving condition-based maintenance [6]. Further ML approaches were implemented by [7] in the form of Adaptive Neuro-Fuzzy Inference System (ANFIS) for the prediction of critical heat flux. For unfolding, ANFIS approaches have also been utilised for the localisation of simulated induced neutron noise sources in VVER-100 rectors, given neutron pulse height distributions as training input [8-9].

Work proposed in [10] unfolds reactor transfer functions by the means of CNNs from simulated neutron noise readings in the frequency domain at differing perturbation types and frequencies. Classification and localisation of the perturbations had been achieved with low error by the means of a 2D-CNN. The localisation of the perturbation source was achieved through the spatial splitting of the core volume into 12 and 48 subsections for classification of source perturbation belonging to a particular subsection. Furthermore, an increased unfolding resolution for localisation was implemented, utilising the extracted latent variables from the CNN and clustering. [11] proposed a 3D-CNN approach to combat the limitations of the 2D implementation in [10] from the loss of spatial information through the conversion of the 3D volume into a 2D input. Moreover, [11] included the classification of time domain signals processed to extract temporal information via RNNs. This work extends the approaches previously developed in [10] and [11] to larger, more complex simulation scenarios, including the localisation of perturbations in the time domain.

3. Simulated Scenarios and Data Pre-Processing

The process of training ML models requires large amounts of training data, representing instances for which known perturbations are assumed and the corresponding induced neutron noise readings are estimated. The known data allows the system to learn the function mapping detector readings to their classification and origin, i.e. transfer function inversion, or unfolding. To obtain this amount of training data it is necessary to simulate scenarios to practically provide enough examples of differing anomaly types and source locations for effective unfolding. To achieve this, simulations determining the reactor transfer function or Green's function, providing detector readings of the induced neutron noise of a given perturbation scenarios for pressurised water reactors (PWR) have been employed in both the time and frequency domain.

3.1 Frequency Domain

Modelling of fluctuations in neutron flux given known perturbations in the frequency domain was achieved through the CORE SIM [12] reactor physics codes, generating neutron detector readings of the induced neutron noise in a PWR for five perturbation scenarios. CORE SIM models the effects of a noise source for a three-dimensional reactor core, of cylindrical shape in Cartesian geometry for a reactor transfer function – considered to be the Green's function of the system – capturing the response of the fluctuations of the induced neutron flux from known perturbation distributions. The Green's function provides a one-to-one relationship between any location of perturbation and the response of the neutron flux at any position within the core. CORE SIM models a PWR with a radial core of size 15x15 fuel assemblies, utilising

a fine volumetric mesh of 32x32x34 voxels modelling sub-assembly response, including boundary sources. For further details, consult the CORE SIM user manual [13],[12].

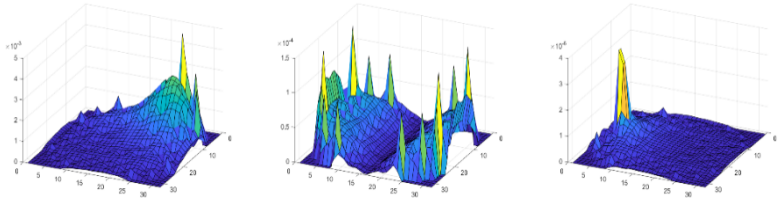

Fig 1. Examples of the amplitude induced neutron flux in the frequency domain for a single azimuthal slice on the 10th axial plane. *Left*: Absorber of Variable Strength. *Middle*: Core Barrel Vibration - *Right*: Vibrating Fuel Assembly, cantilevered.

CORE SIM provides five perturbations scenarios in 34 frequencies (0.1-1.0Hz with a step of 0.1Hz and 1.0-25.0Hz with a step of 1.0Hz) each with two energy groups, i.e. high and low energy spectrum, referred to as *Fast* and *Thermal* groups respectively. The five scenarios include: *Absorber of Variable Strength*, the perturbation of the thermal macroscopic absorption cross-section; *Axial-Travelling Perturbations*, perturbation of the coolant at the velocity of the coolant flow; *Fuel Assembly Vibrations*, vibration of a fuel assembly in the x- and/or y-direction for differing modes cantilevered beam, simply supported for the first mode (0.8-4.0Hz), simply supported in the second mode (5.0-10.0Hz), and cantilevered beam and simply supported for both modes; *Control Rod Vibrations*, vibration of a one-dimensional structure along the z-direction vibrating perpendicularly to the two-dimensional (x,y) plane; *Core Barrel Vibrations*, perpendicular or beam mode of vibration in both the in-phase and out-of-phase modes. Examples of these perturbations can be seen in Fig 1 for an axial cross section of the core volume.

3.1.1 Data Pre-Processing

The signals produced are complex 3D volumes of the size of the fine volumetric mesh (*32x32x34* voxels), representing the induced neutron noise at every point within the core volume for a given perturbation originating from a specific positional coordinate within the core (i, j, k). The signal volumes are provided as the response in both fast and thermal groups, however, for our experimentation only the thermal group is utilised due to neutron detectors being more sensitive to thermal neutrons. The dataset is comprised of 34 frequencies each containing a minimum of 106176 data examples across all scenarios, and have been split into training, validation and testing sets via frequency and source location per scenario.

To mimic the signals from real plant detectors, a pre-determined number of voxel locations have been selected from the whole *32x32x34* volume to emulate the number of detectors within the simulated core. In our case 48 in-core and 8 ex-core detectors have been used from their volumetric positions for the modelled core layout. Furthermore, to emulate reality, the *Auto-Power Spectral Densities* (APSD) and *Cross-Power Spectral Densities* (CPSD) for the simulated signals have been calculated to coincide with real plant readings. Additionally, to demonstrate the robustness of the proposed network white Gaussian noise has been added to the signals in two *signal-to-noise ratios* (SNR), SNR=3 and SNR=1. Finally, as *Deep Neural Networks* (DNNs) currently cannot easily implement complex signals, each of the complex 3D volumes is decomposed to its amplitude and phase. The now two volumes are concatenated together channel-wise to form a *2x32x32x34* volume.

3.2 Time Domain

The determination of the reactor transfer function in the time domain was employed by the Simulate-3K (S3K) algorithm [14], modelling 48 in-core and 8 ex-core neutron detectors for the four-loop, Westinghouse, PWR mixed core of the OECD/NEA transient benchmark. S3K has been utilised to perform 27 different scenarios comprised of 6 perturbation settings and their combinations: *Fuel Assembly Vibration* of the central 5x5 cluster, vibrating synchronously in the x- or y-direction at a frequency of 1.5Hz (sine wave) or random (white noise); *Fluctuations*

of the Coolant Flow, at ±1% from the relative mean amplitude; *Fluctuations of the Coolant Temperature*, at ±1ºC from the mean value of 286.7 ºC. These perturbations distributions have been performed with core operating conditions similar to the aforementioned frequency domain model.

S3K simulates each of the scenarios with a duration of 100 seconds sampled at 0.01 time steps for each of the 48 in-core and 8 ex-core detectors. The detectors are positioned at 8 azimuthal locations at 6 axial levels for in-core and distributed at 4 azimuthal locations at 2 different axial locations for the ex-core. In addition to the above classification scenarios, individual fuel assembly vibrations for all 193 azimuthal locations within the core have been modelled for 5 different scenarios of 4 perturbation settings including combinations of the 4: *Fuel Assembly Vibration* in the x-direction at a frequency of 1.5Hz (sine wave) or random (white noise); *Fluctuations of the Coolant Flow*, at ±1% from the mean value; Fluctuations of the *Coolant Temperature*, at ±1ºC from the mean value of 286.7 ºC. These scenarios have been experimented for the classification and localisation of the perturbing fuel assembly. For further technical details on S3K refer to the user manual [14].

3.2.1 Data Pre-Processing

The signals produced by S3K are presented as 10001-dimensional vectors per each of the 56 detectors for each scenario, representing the neutron readings of the induced neutron flux. Due to the limited number of data samples available, data augmentation was performed to increase the number of samples per detector per scenario, and to reduce the large input size into the DNN. To achieve this a sliding window of width 100 time-steps and stride 25 was used to represent a 1 second input to the network, this produced the vector $x \in \mathbb{R}^{396\times100}$ per detector. Furthermore, splitting the data into training, validation, and testing sets has been accomplished via the position of the detector, this means specific detector locations have been split into differing sets to the description in Fig 2 per axial position of the detectors. Finally, to further test the robustness of the proposed networks, white Gaussian noise has been added to the signals at two SNRs, SNR=5 and 10.

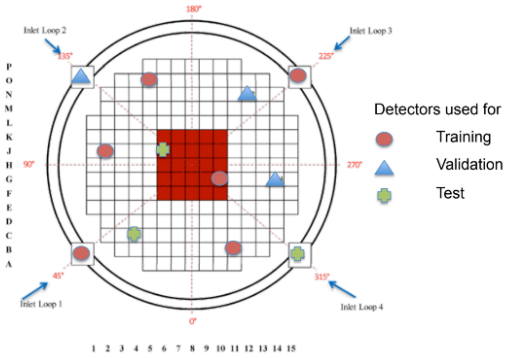

Fig 2. Modelled core layout with 8 in-core and 4 ex-core detector locations shown for one axial plane. Corresponding train, test and validation detector splits shown, with central *5x5* FA cluster shown in red.

Additionally, for the localisation of fuel assembly vibrations, the same sub-sampling process has been undertaken; however, all 56 detectors for a 1 second sample are considered to be one input into the network. Therefore, the split of data has been achieved through the source location of the vibrating fuel assembly, ensuring the same assembly is not present between sets. The same process of applying white Gaussian noise have also been applied to study the effect on the network at SNR=3 and SNR=1, at higher levels of noise, due to the added robustness of utilising all possible 56 detectors as input.

4. Approach

ML and more specifically *Deep Learning* (DL) are a set of powerful algorithmic approaches for data analytics and pattern recognition, applying iteratively learnt knowledge to unseen data for decision making tasks without being explicitly programmed. DL is a subset of ML, utilising multiple stacked layers of *Artificial Neural Networks* (ANN) – inspired by biological neurons – to extract varying levels of information, hence the term *deep*. The proposed approaches utilise modern deep learning techniques and architectures extracting valuable pattern information from the input signals to iteratively learn the inverse of the reactor transfer functions.

4.1 3D Convolutional Neural Network

Convolutional Neural Networks (CNNs) [15] are specialised ANNs designed for spatial feature extraction from data with known grid-like topologies, i.e. images. CNNs replace the traditional matrix multiplication of ANNs with the convolution operation extracting spatial features. Moreover, improving efficiency with the capability of learning coarse to fine features through the addition of more CNN layers, extracting complex hierarchical concepts from such features. Convolutional layers utilise a set of kernels, learning a corresponding number of filters that to capture these spatial patterns pertaining to the given input. Formally, computing the activation of a convolutional layer ℓ and feature-map f at positions i, j, k is given by

$$a_{i,j,k}^{[\ell,f]} = \phi(n_{i,j,k}^{[\ell,f]} + b^{[\ell,f]}) \quad (1)$$

where ϕ is a non-linear activation function such as Rectified Linear Units (ReLU: $f(x) = max(0, x)$) and b is a learnt bias $n_{i,j,k}^{[\ell]}$ is given by

$$n_{i,j,k}^{[\ell]} = \sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} \sum_{z=0}^{Z-1} W_{x,y,z}^{[\ell]} \cdot A_{i+x,j+y,k+z}^{[\ell-1]} \quad (2)$$

where $W^{[\ell]}$ is a kernel of learnt weights in layer ℓ with dimensions $X \times Y \times Z$, convolved with the activations from the previous layer $W^{[\ell]} * A^{[\ell-1]}$. This produces a weighted sum per location of all points within a kernels receptive field of the previous layers' activations. Visual examples of the features learnt via the convolution operation can be seen in Fig 4.

Given the volumetric nature of the signals in the frequency domain and the task of localisation, it is necessary to obtain spatial relationships and patterns within the data volume. Therefore, this work proposes a modified, densely-connected, 3D-CNN for the volumetric feature extraction of simulated neutron detector readings seen depicted in Fig 3.

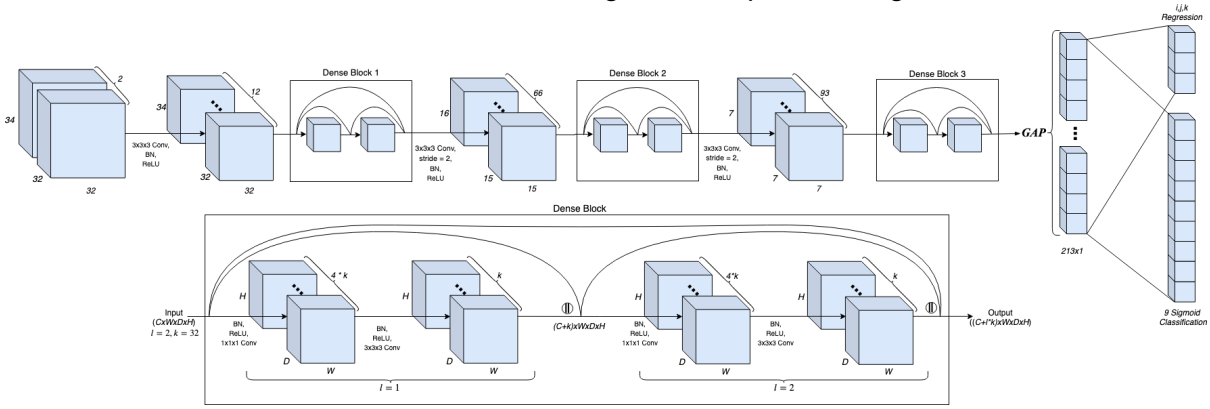

Fig 3. The proposed Densely-connected 3D CNN architecture, depicting an example dense block of 2 layers and growth rate of 32. The Fully-connected and output layers can be seen right of the GAP, each unit represents a classification perturbation type or the source (*i,j,k)* location to be regressed.

The network depicted in Fig 3. shows the architectural construction of the 3D CNN, comprised of three dense blocks modified from the 2D variant to allow for the 3D volumetric input. Dense blocks [16] are an DNN architectural design, utilising several CNN developments, with its main advantage being the use of dense connections. These connections allow for a greater flow of information between layers during the forward and backward pass of the backpropagation procedure, resulting in the reduction of vanishing gradients and achieving better performance. These connections are simply concatenations, where the ℓ^{th} hidden layer H_ℓ receives as input the feature-maps all preceding layers within that block.

$$X_\ell = H_\ell([X_0, X_1, \ldots, X_{\ell-1}]) \quad (3)$$

In addition to the dense connections, the network employs *1x1x1* kernel convolutions with stride 1 for the reduction in feature dimensionality following dense connections, furthermore, *1x1x1* kernels reduce network parameters whilst increasing network complexity, further

assisting the parameter large 3D convolution operation [17]. The dense blocks each contain $l = 20$ layers with growth rate of $k = 6$, for further details please refer to [16]. All convolutional layers are followed by the commonplace procedure: *convolutional layer → Batch Normalization (BN) → and ReLU activation*. BN normalises the activations output by the convolutional layer improving network stability, ReLU is a non-linear activation function with sparse activation, further assisting in the reduction of vanishing gradients. Furthermore, the proposed network replaces the pooling operation with strided convolutions for dimensionality reduction, retaining spatial structural information from the input vital for the localisation of perturbation sources.

The last convolutional layer of the network outputs a representational feature vector of the input of size 256 via *Global Average Pooling* (GAP) layer [17], fully connected to two output layers for perturbation classification and localisation. GAP directly outputs the spatial average over the feature maps, resulting in a vector $V \in \mathbb{R}^m$ where m is the number of feature maps. The output layer for classification is comprised of 9 non-linear, sigmoid units each for the occurrence of the individual perturbation types (nine types as modes of fuel assembly vibration are considered as classes of perturbation). For localisation three linear units have been employed each representing the (*i, j, k*) coordinates of the perturbation source to be regressed.

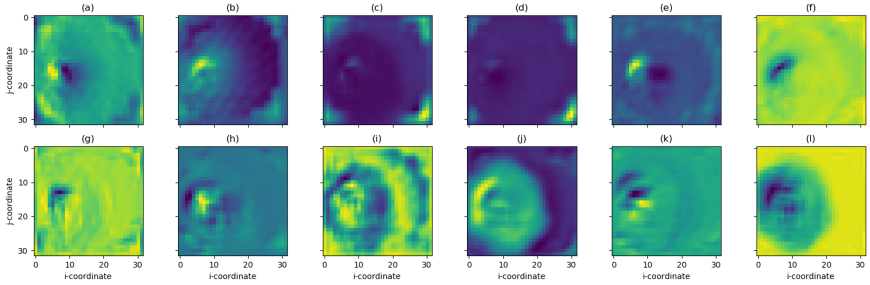

Fig 4. Sample of 12 learnt feature-maps from the output of first dense block for the input of vibrating fuel assembly at (*8,16*) given all possible detectors. Visually depicting how the differing layers highlight different features of the image. (a) shows a peak at the source of vibration, (d) the response on the core barrel, (j) the noise dissipating throughout the core.

Training the network has been achieved via implementing the multi-task loss approach from [11], minimising the weighted sum of losses per task (classification and localisation) with a weight coefficient identifying the impact each tasks loss in the training procedure. For classification the network aims to minimise the *negative log-likelihood* (NLL)

$$NLL = -\frac{1}{N}\sum_{i=1}^{N} y_i \cdot \log(\hat{y}_i) + (1 + y_i) \cdot \log(1 - \hat{y}_i) \quad (4)$$

and for localisation regression, minimises the L2 loss, or mean squared error (MSE)

$$MSE = \frac{1}{N}\sum_{i=1}^{N} \|y_i - \hat{y}_i\|^2 \quad (5)$$

where y_i and $\hat{y}_i$ are the true and predicted values of the network for N number of examples. As previously alluded the 3D CNN network is trained minimising a weighted sum of losses

$$\mathcal{L}(X; \boldsymbol{W}, \lambda_1, \lambda_2) = -\frac{1}{N}\sum_{i=1}^{N}\left[\frac{\lambda_1}{P}\sum_{p=1}^{P}\left[y_1^p \cdot \log\left(\hat{y}_1^p\right) + \left(1 - y_1^p\right) \cdot \log\left(1 - \hat{y}_1^p\right)\right] - \frac{\lambda_2}{C}\sum_{c=1}^{C}\|y_2^c - \hat{y}_2^c\|^2\right]_i \quad (6)$$

where P and C are the number of perturbation classes and source location coordinates respectively, λ_1 and λ_2 are the manually tuned hyper-parameter weight coefficients for each task loss, classification and localisation regressing respectively. This objective is minimised given X as input data with respect to $\boldsymbol{W}$ parameters (weights and biases).

4.2 Long Short-Term Memory, Recurrent Neural Network

Time domain signals hold temporal information within their sequential structure, therefore, a differing approach to previously described is necessary to capture these time-dependent

features. To more appropriately capture the relationships within the detector signals, *Recurrent Neural Networks* (RNN) have been employed. RNNs utilise recurrence to allow information about previous time-steps to persist within the network informing current and future time-step cells across the sequence. RNNs in principle formulate a non-linear output A_t from both the input data x_t at that given time-step and the activation of the previous timesteps cell A_{t-1}, where ϕ is a non-linear activation function such as hyperbolic tangent (tanh):

$$A_t = \phi(x_t, A_{t-1}) \tag{7}$$

Long Short-Term Memory (LSTM) [18], a variation of RNNs have been incorporated in this work for their ability to learn long term dependencies across long sequences, ideal for the 100 time-step sequences in question. It achieves this ability with the use of memory gates, regulating and learning how much to 'remember' from previous cell states and how much to contribute from the current data input. Initially, the forget gate determines what to remember from the previous cell state C_{t-1} given activation A_{t-1}. To decide what new information will be added to the current cell state, an input gate i_t and candidate values $\tilde{C}_t$ are generated.

$$\begin{aligned} \mathrm{f_t} &= \sigma(\mathbf{W_f} \cdot \left[\mathrm{A_{t-1,} x_t}\right] + \mathrm{b_f}) \\ \mathrm{i_t} &= \sigma(\mathbf{W_i} \cdot \left[\mathrm{A_{t-1,} x_t}\right] + \mathrm{b_i}) \\ \tilde{\mathrm{C}}_\mathrm{t} &= \tanh(\mathbf{W_C} \cdot \left[\mathrm{A_{t-1,} x_t}\right] + \mathrm{b_C}) \\ \mathrm{C_t} &= \mathrm{f_t} \odot \mathrm{C_{t-1}} + \mathrm{i_t} \odot \tilde{\mathrm{C}}_\mathrm{t} \end{aligned} \tag{8}$$

The outputs of these gates are combined to create an update the previous cell state to the cell state $\mathrm{C_t}$ via the forgetting and updating previously computed through learnt weights. The output gate is employed to control what should be output from the newly computed cell states, outputting a non-linear activation $\mathrm{A_t}$ to the subsequent cells.

$$\begin{aligned} \mathrm{o_t} &= \sigma(\mathbf{W_o} \cdot \left[\mathrm{A_{t-1,} x_t}\right] + \mathrm{b_o}) \\ \mathrm{A_t} &= \mathrm{o_t} \odot \tanh(\mathrm{C_t}) \end{aligned} \tag{9}$$

Further details of the intuition of LSTMs can be found in [18], with the above process visually depicted in Fig 5 within each of the LSTM cells.

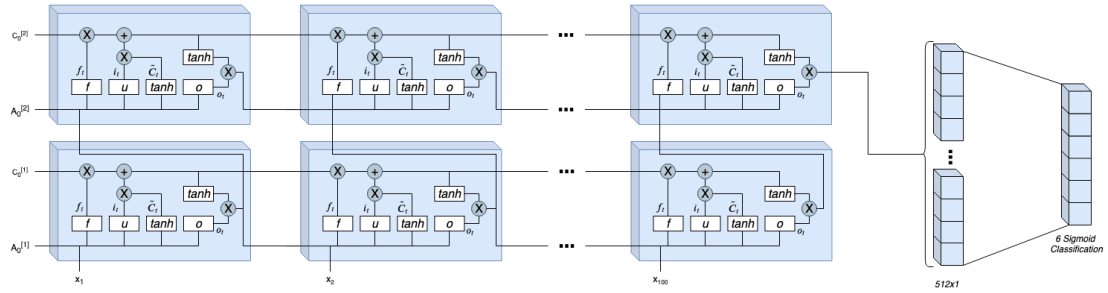

Fig 5. LSTM RNN architecture proposed for the classification task, outputting a 512-dimensional representational vector of the input to a 6-unit classification layer. The LSTM units take in input from the bottom, $\mathrm{x_t}$, with all gates depicted in each LSTM cell.

The network proposed solely for the classification task incorporates a LSTM network comprised of 2 stacked layers. Each cell within those layers contains 512 units, outputting a 512-dimensional feature representation vector of the single sensor input for 1 second, depicted in Fig 5. This network outputs to 6 non-linear sigmoid units for the classification of the presence of individual perturbations from one detector reading. Dropout [19] of 25% drop probability, has been employed in the LSTM network regularising the effects of overfitting, setting a percentage of the unit activations to zero, limiting the networks learning capacity. The LSTM network has been trained to minimise the negative log-likelihood with respect to the parameters $\mathbf{W}$ and input x as noted in (6).

Localising vibrating fuel assemblies has been achieved employing the same core LSTM architecture as aforementioned, with the addition of a linear output layer, fully connected to the 512-dimensional representation vector for the regression of azimuthal coordinates. The training of this network has been achieved by minimizing the weighted sum of each loss per task, as to the definition in (6).

5. Experimental Results

5.1 Frequency Domain

The subsequent experiments show the results of reactor transfer function unfolding for the classification and localisation of induced perturbations given the neutron flux from simulated neutron detectors in the frequency domain from the proposed densely connected 3D CNN. The experiments have been implemented utilising the Pytorch numerical computation library trained via backpropagation, minimising the multi-task loss criterion in 4.1 with the Adam optimizer with its proposed parameters as in [20]. A batch size of 32 has been used, trained on an 8-core, 16-thread Intel CPU system, with 4 Nvidia 1080ti GPUs and 94GB of RAM, each model being trained 3 times and the mean and standard deviation being taken as the result.

Two experiments were conducted on the volumetric signals, the first using different sized splits of training, validation, and testing data to more appropriately represent the limited amount of data available from real plant readings, the subsequent results can be seen in Tab 1. Furthermore, the results from the utilisation of detector readings from all possible voxel positions within the reactor core and only 48 in-core detectors are also shown, where the 48 in-core detectors are located corresponding to the layout of the core modelled in 3.1. For the latter experiment, the volumetric signals were corrupted with white Gaussian noise, as described in 3.1.1 to test the robustness of the proposed system in adverse conditions.

3D-CNN Results of Classification and Localisation

Sensors	Train / Valid / Test (%)	Classification		Localisation	
		Accuracy (%)	F1-Score	MAE	MSE
All	70 / 15 / 15	99.94 ± 0.051	0.9344 ± 0.004	0.1435 ± 0.011	0.0342 ± 0.008
48 In-Core	70 / 15 / 15	99.89 ± 0.010	0.9311 ± 0.001	0.2902 ± 0.011	0.3072 ± 0.014
48 In-Core	25 / 15 / 60	99.68 ± 0.025	0.9149 ± 0.002	0.3978 ± 0.017	0.6407 ± 0.052
48 In-Core	15 / 25 / 60	99.56 ± 0.061	0.9141 ± 0.003	0.4858 ± 0.017	0.7727 ± 0.006

Tab 1. Results of the proposed 3D-CNN for the classification and localisation of perturbation type and source location (*i,j,k*). Mean and standard deviation of 3 runs.

3D-CNN Results of Classification and Localisation with the addition of noise.

Noise	Train / Valid / Test (%)	Classification		Localisation	
		Accuracy (%)	F1-Score	MAE	MSE
No Noise	70 / 15 / 15	99.89 ± 0.010	0.9311 ± 0.001	0.2902 ± 0.011	0.3072 ± 0.014
SNR = 3	70 / 15 / 15	99.85 ± 0.006	0.9231 ± 0.001	0.3456 ± 0.016	0.4905 ± 0.011
SNR = 1	70 / 15 / 15	99.81 ± 0.036	0.9225 ± 0.002	0.3709 ± 0.020	0.5185 ± 0.017

Tab 2. Results of the proposed 3D-CNN for the classification and localisation of perturbation type and source location (*i,j,k*) with the corruption of input signals at SNR=3 and SNR=1.

The results in Tab 1 show that the proposed 3D CNN models perform highly in the classification task across all testing splits, with 99.89 ± 0.010% accuracy in the best case and 99.56 ± 0.061% in the worst, respectively achieving an F1-score of 0.9311 ± 0.001 and 0.9141 ± 0.003. F1-score is an alternative measure of accuracy of prediction and target, as a function of precision and recall

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{10}$$

where

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \quad \text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \tag{11}$$

computed from the confusion matrix of predicted values of the network and true values of the data. F1-score lies within the range [0.0,1.0] where 1 is perfect precision and recall. The regression results of the perturbation source coordinates observed in Tab 1 show low error

was achieved, with a best case of 0.2902 ± 0.011 and 0.3072 ± 0.014 for the mean absolute error (MAE) and mean squared error (MSE) respectively. In relation to the core volume, this is approximately 4cm localisation error in an 4m x 4m x 4m reactor core utilising only 48 detectors. Tab 2 shows the results with the addition of signal corruption of the volumetric signals, with a worst case of 99.81 ± 0.036% accuracy, 0.9225 ± 0.002 F1-score and 0.3709 ± 0.020 MAE for classification and localisation respectively, demonstrating the robustness of the proposed approach with minimal deviation from the best performance of no corruption.

5.2 Time Domain

Experimentation in the time domain for the unfolding of the reactor transfer function for the classification of perturbation type has been achieved via individual neutron detector measurements as described in 3.2.1. Tab 3 displays the results of the one second samples for the 27 scenarios of 6 perturbation settings under different SNRs of signal noise corruption. The finalised results are the mean and standard deviations of 3 training runs, trained via backpropagation with the RMSprop optimizer [20] with default settings and learning rate of 0.0001, and utilising a batch size of 64. The results show that given just 1 second readings from one neutron detector our approach can accurately classify the perturbation type with a best case of 96.41 ± 0.021% accuracy, the addition of noise has shown that although performance degrades, the system is robust given such minimal data input.

LSTM Classification Results

Noise	Accuracy (%)	F1-Score
Clean Signal	96.84 ± 0.491	0.9342 ± 0.003
SNR = 10	91.88 ± 0.254	0.8107 ± 0.007
SNR = 5	88.87± 0.279	0.7469 ± 0.006

Tab 3. Classification of perturbation type in the time domain under differing levels of input signal noise corruption from induvial detector inputs.

Localisation of vibrating fuel assembly source takes a similar approach utilising the same training procedure except for the minimisation criterion, replacing with the multi-task loss in (6). Additionally, all 56 detectors have been utilised – compared to the previous experiment of individual detectors – to obtain spatial information between the detectors to infer the perturbing fuel assembly location. Corrupting the signals with white Gaussian noise has also been applied to test the robustness of the proposed approach, the resulting error of localisation can be seen in Tab 4. Localisation in the time domain has been achieved with low localisation error with a worst case of 1.2304 ± 0.102 and 3.2340 ± 0.612 under SNR=1, and a best of 1.0737 ± 0.006 and 2.3682 ± 0.065 for MAE and MSE respectively.

LSTM Classification and Localisation Results

Noise	Classification		Localisation	
	Accuracy (%)	F1-Score	MAE	MSE
Clean Signal	99.89 ± 0.396	0.9976 ± 0.003	1.0737 ± 0.006	2.3682 ± 0.065
SNR = 3	99.87 ± 0.032	0.9980 ± 0.001	1.1191 ± 0.008	2.7316 ± 0.006
SNR = 1	99.46 ± 0.318	0.9962 ± 0.004	1.2304 ± 0.102	3.2340 ± 0.612

Tab 4. Localisation of the coordinates of a vibrating fuel assembly (*i,j*), in the time-domain utilising the proposed LSTM architecture, under input signal corruption. Mean and standard deviation of 3 runs.

6. Conclusions and Future Work

This work proposed an extended approach to the unfolding of reactor transfer functions for the classification and localisation of reactor core perturbations from neutron detector readings produced by simulated core models. The proposed models accurately classify perturbation types and source locations in the time and frequency domain, with extended and more complex simulated perturbation scenarios than previous work [11,12]. Our approach outperforms previous approaches for the same task localising such perturbations to a finer voxel mesh and with fewer detectors available, i.e. 48 in-core detectors for a *32x32x34* core volume.

Our experiments further solidify the applicability and capability of deep learning approaches in the domain of nuclear reactor anomaly detection, specifically for the non-trivial task of reactor

transfer function unfolding given very spare neutron flux detector readings. We will continue to extend our approaches to localising and classifying large combinations of perturbations simultaneously. Furthermore, investigations will be made to apply our model to real plant data providing further validation of the capability of our approach for on-line anomaly detection.

Acknowledgements

The research conducted was made possible through funding from the Euratom research and training programme 2014-2018 under grant agreement No 754316 for the 'CORe Monitoring Techniques And EXperimental Validation And Demonstration (CORTEX)' Horizon 2020 project, 2017-2021. We would like to thank the Chalmers University of Technology, particularly Dr C. Demaziere, Dr P. Vinai, Dr A. Milonakis and the Paul Scherrer Institute, particularly Dr A. Dokhane and Dr V. Verma for providing the frequency and domain data respectively, for assisting us with their understanding and for collaborating with us in the analysis process.

Author contribution statement

All authors have contributed equally to the conceptualisation and the technical developments pertaining to the machine learning components and the formulation of the problem.
The data were provided by some of the consortium members of the EU-H2020 project Cortex, mentioned in the Acknowledgements section. Aiden Durrant led the programming aspects of the proposed deep neural network technique and the first draft of the manuscript. Georgios Leontidis and Stefanos Kollias supervised the implementations presented. All authors contributed equally to the evaluation, validation, presentation, review and approval of the final manuscript.

References

[1] C. Demazière et al., Overview of the CORTEX project, Proc. Int. Conf. Physics of Reactors – Reactor Physics paving the way towards more efficient systems (PHYSOR2018), Cancun, Mexico, April 22-26, 2018 (2018).
[2] D. Rolnick, et al. Tackling Climate Change with Machine Learning. arXiv preprint arXiv:1906.05433, (2019).
[3] F.C. Chen, and M.R. Jahanshahi, N NB-CNN: deep learning-based crack detection using convolutional neural network and Naïve Bayes data fusion. IEEE Transactions on Industrial Electronics, 65(5), pp. 4392-4400 (2017).
[4] W. Li, et al., Design of comprehensive diagnosis system in nuclear power plant. Ann. Nucl. Energy, 109, pp. 92-102, (2017).
[5] M.C. dos Santos, et al., Deep rectifier neural network applied to the accident identification problem in a PWR nuclear power plant Ann. Nucl. Energy, 133, pp.400-408, (2019).
[6] R.M. Ayo-Imoru, and A.C. Cilliers, Continuous machine learning for abnormality identification to aid condition-based maintenance in nuclear power plant. Ann. Nucl. Energy, 118, pp.61-70, (2018).
[7] S. Zaferanlouei, et al., Prediction of critical heat flux using anfis. Ann. Nucl. Energy, 37(6), pp. 813–821, (2010).
[8] S.A. Hosseini, and I.E.P. Afrakoti, Neutron noise source reconstruction using the adaptive neuro-fuzzy inference system (anfis) in the vver-1000 reactor core. Ann. Nucl. Energy, 105, pp. 36–44, (2017).
[9] S.A. Hosseini, and I.E.P. Afrakoti, Evaluation of a new neutron energy spectrum unfolding code based on an Adaptive Neuro-Fuzzy Inference System (ANFIS). Journal of radiation research, 59(4), pp.436-441, (2018).
[10] F. Calivà F et al., A deep learning approach to anomaly detection in nuclear reactors. Proc. 2018 Int. Joint Conf. Neural Networks (IJCNN2018), Rio de Janeiro, Brazil, July 8-13, 2018 (2018).
[11] F. De Sousa Ribeiro et al., Towards a deep unified framework for nuclear reactor perturbation analysis. Proc. IEEE Symposium Series on Computational Intelligence (SSCI 2018), Bengaluru, India, November 18-21, (2018).
[12] C. Demazière, Core sim: a multi-purpose neutronic tool for research and education. Ann. Nucl. Energy, 38(12):2698–2718, (2011).
[13] C. Demazière, User's manual of the core sim neutronic tool. Technical report, Chalmers University of Technology, (2011).
[14] G. Grandi, et al. Simulate-3k models and methodology. SSP-98013, 6, (2006).

[15] Y. LeCun, et al., Generalization and network design strategies. Connectionism in perspective, pp. 143–155, (1989).
[16] G. Huang, et al., Densely connected convolutional networks. Proc. IEEE Conf. on computer vision & pattern recognition, Honolulu, Hawaii, USA, July 22-26, 2017 (2017).
[17] M. Lin, et al., Network in Network. arXiv preprint arXiv:1312.4400, (2013).
[18] S. Hochreiter, and J. Schmidhuber, Long short-term memory. Neural computation, 9(8), pp. 1735–1780, (1997).
[19] N. Srivastava, et al., Dropout: a simple way to prevent neural networks from overfitting. The journal of machine learning research, 15(1), pp. 1929-1958, (2014).
[20] D.P. Kingma, and J. Ba, Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980, (2014).s