Putting Together Wavelet-based Scaleograms and Convolutional Neural Networks for Anomaly Detection in Nuclear Reactors

Thanos Tagaris
Institute of Communication and Computer Systems
National Technical University of Athens
15780, Zografou Campus, Athens, Greece
thanos@islab.ntua.gr

George Ioannou
Institute of Communication and Computer Systems
National Technical University of Athens
15780, Zografou Campus, Athens, Greece
geoioannou@islab.ntua.gr

Maria Sdraka
Institute of Communication and Computer Systems
National Technical University of Athens
15780, Zografou Campus, Athens, Greece
masdra@islab.ntua.gr

Georgios Alexandridis
Institute of Communication and Computer Systems
National Technical University of Athens
15780, Zografou Campus, Athens, Greece
gealexandri@islab.ntua.gr

Andreas Stafylopatis
Institute of Communication and Computer Systems
National Technical University of Athens
15780, Zografou Campus, Athens, Greece
andreas@cs.ntua.gr

ABSTRACT

A critical issue for the safe operation of nuclear power plants is to quickly and accurately detect possible anomalies and perturbations in the reactor. Defects in operation are principally identified through changes in the neutron flux, as captured by detectors placed at various points inside and outside of the core. While wavelet-based analysis of the measured signals has been thoroughly used for anomaly detection, it has yet to be coupled with deep learning approaches. To this end, this work presents a novel technique for anomaly detection on nuclear reactor signals through the combined use of wavelet-based analysis and convolutional neural networks. In essence, the wavelet transform is applied to the signals and the corresponding scaleograms are produced, which are subsequently used to train a convolutional neural network that detects possible perturbations in the reactor core. The overall methodology is experimentally validated on a set of simulated nuclear reactor signals generated by a well-established relevant tool. The obtained results indicate that the trained network achieves high levels of accuracy in failure detection, while at the same time being robust to noise.

ICAAI 2019 The 3rd International Conference on Advances in Artificial Intelligence, October 26–28, 2019, Istanbul, Turkey

ISBN 978-1-4503-7253-4/19/10…15.00

DOI: 10.1145/3369114.3369121

CCS Concepts

•**Computing methodologies** → **Anomaly detection;** Supervised learning by classification; •**Applied computing** → **Physics;** *Engineering;*

Keywords

Anomaly Detection; Nuclear Power Plants; Signal Processing; Deep Learning; Convolutional Neural Networks; Wavelet Transformation; Scaleograms

1. INTRODUCTION

Anomaly detection is a very active research area with applications in a large variety of domains, such as fault diagnosis in factory systems, health monitoring, ecosystem disturbances, intrusions in network data and event detection in sensor networks [10]. In all of the aforementioned cases, the objective is to identify samples or events exhibiting behavior that deviates from what it could be considered as "normal" or "ordinary" within each domain. Recently, an increase in anomaly detection methodologies based on machine learning techniques has also been witnessed [6].

This work focuses on anomaly detection on *Nuclear Power Plant* (NPP) data. Timely detection of incipient faults and condition monitoring are crucial for operational safety in NPPs, as well as for improving their performance. For example, a drift in steam generator feedwater flow sensors can result in reactor power output reduction by as much as 3% [9]. These faults can be caused by either single or multiple component defects. The task of detecting the faulty components can be quite difficult, because the anomaly must be identified and located in a timely manner in order to be handled [13]. However, changes in reactors usually occur at slow paces and by the time the anomaly is clearly seen, it may be too late to tackle the problem. Therefore, it is

more than necessary to develop systems that automatically monitor the state and condition of nuclear reactors.

One of the principal ways of assessing a nuclear reactor's state is inspecting the level of the *neutron noise*, which can be defined as the neutron flux fluctuations around a mean value [29]. Changes in neutron noise level inside the core are related to processes within the reactor, such as the thermo-hydraulic oscillations or mechanical and fuel assembly vibrations. The induced neutron noise can be measured by two types of sensors; those located inside the core (*in-core*) and those outside (*ex-core*).

More specifically, the type of nuclear reactors examined in this work are the *KWU pre-Konvoi Pressurized Water Reactors* (PWRs) [36], designed by Siemens AG. KWU PWRs have been exhibiting high values of neutron noise since their beginning [22]. However, this is not alarming because it has not resulted in problems related to the high flux fluctuations. It is important, though, to monitor these values frequently for safety reasons. Safety authorities set maximum thresholds of neutron noise just to ensure that no accidents will happen. Some research studies in the past opted to use special filters to prevent limitations of the system, due to the high values of neutron noise exceeding the defined thresholds [16].

Therefore, the application of intelligent techniques to monitor the reactor's state are not only expected to be beneficial to the safety of the NPP but are also important for the performance of the whole system [31]. In Section 2 below, related work on anomaly detection in NPP data and machine learning techniques are presented. The proposed methodology is discussed in Section 3 and the experimental framework is described in Section 4. Finally, the results are outlined in Section 5 and the paper concludes in Section 6.

2. RELATED WORK

Various studies have been performed on anomaly detection in NPP data. Specifically, many *fault detection and diagnosis* (FDD) applications on NPPs are presented in [26]. The FDD methods are divided in *model-based* and *model-free* ones and also according to how they are applied to different NPP problems and intricacies. A real-time data-driven tool that uses *symbolic dynamic filtering* (SDF) to extract meaningful features from time series is introduced in [18] and compared to *principal component analysis* (PCA). SDF constructs a probabilistic finite state automaton whose output (a lower-dimensional vector) is used to train a classifier. In the aforementioned work, it was demonstrated that SDF-based classifiers exhibited a lower test error rate than PCA-based ones.

Support vector data description (SVDD) [35] has also been employed in NPP data in an effort to identify outliers on a dataset consisting of signals of neutron flow in a nuclear reactor channel [27]. Another approach based on *deep learning* (DL) techniques [23] that localize the specific point in the reactor core that a perturbation originates from, has been presented in [8]. In this case, the signals are transformed to the frequency domain and subsequently provided in slices to a *convolutional neural network* (CNN) [23], yielding very good results. Additionally, a clustering technique and a denoising method (using a *denoising autoencoder*) have been proposed to tackle this problem [8].

In the current work, anomalies in the reactor core (mostly neutron flux perturbations) are identified through the combined use of wavelet-based scaleograms [33] (a form of spectrograms) and DL methodologies. These two techniques have also been used together in other domains as well. For example, complex spectrograms have been used to train a speech-enhancement CNN model, which denoises the input signals [14]. In the field of acoustic scene classification, a DL approach based on scaleograms and spectrograms paired with a pretrained CNN and a *generalized regression neural network* has been proposed in [30], providing excellent results in the DCASE 2017 challenge [28]. Besides that, DL methods have been applied to speech emotion recognition problems. In [2], a CNN model is trained with spectrograms from speech signals to recognize emotions, such as anger, boredom, etc, extracted from the Berlin dataset [7]. Another methodology of emotion recognition has been presented in [32], where the speech signal has been split in parts, with each of them producing a spectrogram. This sequence of spectrograms has been provided to a convolutional-recurrent model, constructed by CNNs and *bi-directional Long-Short Term Memory Networks* (bi-LSTMs), exhibiting promising results in clean, as well as noisy conditions. Another DL technique with convolutional recurrent neural networks for emotion recognition in *multi-channel electroencephalography* data has been implemented in [25], where the signals have been transformed into spectrograms prior to model training.

3. THE PROPOSED METHODOLOGY

The main objective of the proposed methodology is to be able to identify the driving perturbation from neutron flux signals captured by the in-core and ex-core sensors of a nuclear reactor. This is achieved through the application of the *discrete wavelet transform* (DWT) [12] to the signals and the generation of the corresponding wavelet scaleograms. Subsequently, those scaleograms are provided to a CNN that learns to classify them to the different perturbation types.

3.1 Discrete Wavelet Transform

The DWT is a technique that decomposes a given signal to a set of wavelet basis functions. These, in contrast to the sinusoidal functions of the *Discrete Fourier transform* (DFT) [5] are *spatially localized* (i.e.being non-zero for only over a part of their length). The wavelet functions are dilated, scaled and translated versions of an archetype function called the *mother wavelet* [12].

One key advantage of DWT over DFT is the property of temporal resolution, i.e. its ability to capture both frequency and location information. However, in contrast to DFT, DWT refers to a set of different transformations with respect to the chosen *mother wavelet*. In other words, different mother wavelets, decompose the signal in different ways, so before performing a wavelet analysis, the *optimal mother wavelet* should be identified.

3.1.1 Mother Wavelet Selection

In principle, the optimal mother wavelet should be one whose shape "matches" that of the signal. If this is the case, then large wavelet coefficients will appear at just a few levels (associated with the desired signal), while noise will be evenly spread among all levels [24].

The optimal mother wavelet may be determined according to several criteria; however, the most popular ones are *cross-correlation* (a statistical metric) [24] and *energy-to-entropy* (an information theoretical metric) [3, 19]. The former aims

at measuring the similarity between the mother wavelet and the signal. The cross-correlation coefficient γ between signals X and Y, is calculated according to Equation 1 below:

$$\gamma(X,Y) = \frac{\sum(X-\bar{X})(Y-\bar{Y})}{\sqrt{\sum(X-\bar{X})^2 \sum(Y-\bar{Y})^2}} \quad (1)$$

where $\bar{X}, \bar{Y}$ are the mean values of X, Y respectively.

Energy-to-entropy maximizes the ratio of the decomposed signal's energy to its entropy. The *energy* can be through of as the amount of information that the signal's coefficients carry, while (Shannon's) *entropy* is a measure of uncertainty that produces high values for "noisy" decompositions [3]. This criterion aims at maximizing the amount of energy the wavelet can extract from the signal, while minimizing the entropy of the decomposed signal [19]. The energy (Equation 2) and entropy (Equation 3) of the decomposed signal s are calculated as follows

$$Energy\,(s) = \sqrt{\frac{1}{N}\sum_{i=1}^{N} s_i^2} \quad (2)$$

$$Entropy\,(s) = \sum_i \left(s_i^2 \cdot log\left(s_i^2\right)\right) \quad (3)$$

where s_i are the decomposed signal's coefficients. The *energy-to-entropy* ratio is given in Equation 4

$$\zeta\,(s) = \frac{Energy\,(s)}{Entropy\,(s)} \quad (4)$$

An example energy-to-entropy analysis of a signal is depicted in Figure 1, showing the value of the ratio for each mother wavelet function. The wavelet with the highest energy-to-entropy ratio is the optimal one (in this case, *Biorthogonal 5.5*).

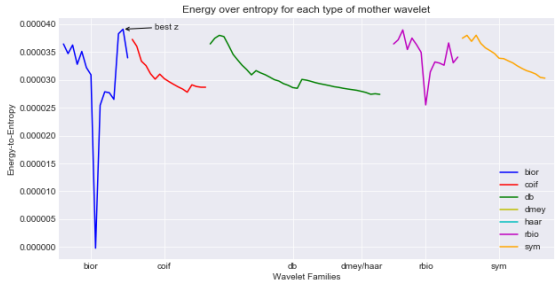

Figure 1: Energy-to-entropy criterion for various mother wavelet functions for a given signal.

3.1.2 Scaleograms

A *scaleogram* [33] is a type of a two-dimensional heatmap that depicts the spectrum of DWT frequencies as they vary through time. It is very useful for visualizing the strength of a signal at each frequency and for each time step. In contrast to *spectrograms* produced by DFT, which depict time and frequency on the x and y axis respectively, the frequency bands of a scaleogram (Figure 3) have uniform height, which makes much more sense in the context of DWT, because the y-axis now represents the *levels* of the decomposition.

A sample signal with four distinct states of equal length in time is presented in Figure 2. More specifically, the second and fourth states are the result of superimposing three distinct single-frequency signals, the first of two and the third of only one.

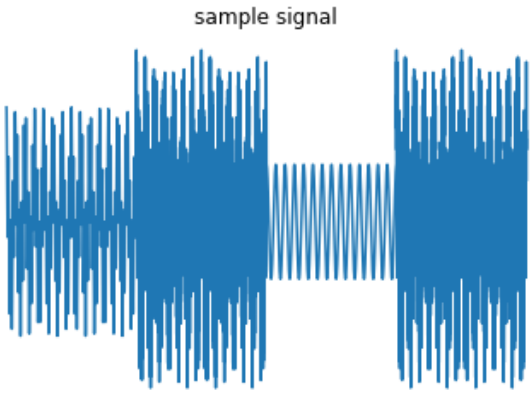

Figure 2: Sample signal with 4 distinct states.

The signal of Figure 2 has been decomposed using DWT with a *Daubechies 4* mother wavelet and its scaleogram is displayed on Figure 3.

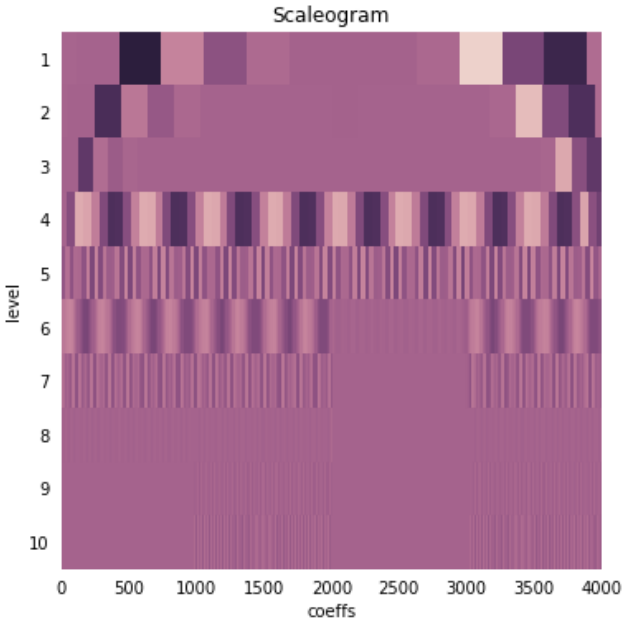

Figure 3: Scaleogram of the signal of Figure 2, with 10 levels of decomposition.

The different frequencies of the sample signal of Figure 2 become apparent in Figure 3. The highest frequency, which is present only in the second and fourth parts of the signal, can be seen at level 10 (i.e. the bottom level) of the scaleogram. Likewise, a frequency present only in the first two and the fourth parts of the signal can be viewed at level 6. Finally, the underlying frequency which is present in the whole duration of the signal can be seen at level 4.

3.2 Convolutional Neural Networks

CNNs [33] are a category of neural networks, consisting of convolutional, pooling and fully connected layers, that are most commonly used in image processing. Their main strengths are *weight sharing* and *sparse connectivity*, which significantly reduce the number of parameters per layer and allow for the creation of very deep networks. CNNs are considered to be the state-of-the-art for most image-related tasks and have been applied to many domains [17, 21].

Due to the simple nature of the scaleograms used in the subsequent experiments, a very deep CNN was not deemed necessary. Instead, a simpler CNN architecture of a depth of 6 layers has been selected. The network is described in more detail in Section 4.3.

4. EXPERIMENTAL FRAMEWORK

The purpose of the conducted experiments has been two-fold; (i) to investigate whether scaleograms could be used for anomaly detection in NPP data and (ii) to identify the optimal time window that should be used when performing this type of analysis. Intuitively, smaller windows would segment the input signal into more parts and consequently produce a larger training set for the CNN. However, segmenting the signal too much could lead to the perturbation not being captured, thereby affecting performance.

4.1 Data simulation and preprocessing

The proposed methodology has been evaluated on data generated by the SIMULATE-3K tool [11, 15], modelling some basic types of perturbations occurring in nuclear reactors. This tool has been built to simulate the steady state operation of a PWR under specific scenarios. These included phenomena such as *fuel assembly vibrations* (both single and in groups), as well as *coolant flow* and *temperature oscillations*. Each of these conditions has been simulated for different frequencies and amplitudes.

Specifically, the scenarios have been grouped in four main categories:

1. **Fuel Assembly Vibrations**, describing a vibration of a single fuel assembly in one direction. Four distinct vibrating fuel assemblies have been simulated, for different conditions (i.e. type of vibration and amplitude). A total of 32 scenarios have been considered in this process.

2. **Cluster of Fuel Assemblies synchronously vibrating**, in the center of the reactor. 12 scenarios with different combinations of vibration frequencies and amplitudes have been examined in this category.

3. **Coolant Flow Oscillations**, simulating random oscillations in the flow of the coolant by up to $\pm 1\%$.

4. **Coolant Temperature Oscillations**, simulating random oscillations in the temperature of the coolant by up to $\pm 1^{\circ}C$.

Each simulated scenario has been "captured" by 48 in-core and 8 ex-core detectors. Depending on the scenario, the signals have a length of either $35secs$ or $100secs$ and a sampling rate of $100Hz$. The in-core sensors are located at 8 different axial locations, each taking measurements at 6 different heights [4]. The ex-core sensors, on the other hand, are placed in 4 locations at 2 different levels. It has been reported that due to their spatial symmetry (the 4 ex-core sensors are placed at an angle of 90° from one another) some vibrations may not be detected at all [1]. It is for this reason that in-core measurements should also be taken into account.

Despite the fact that the examined signals are simulated, an initial preprocessing step is still necessary. Firstly, possible trends in the signals have to be removed, using *detrending techniques* [29]. Secondly, in an effort to construct a unified evaluation procedure, the signals have been resampled so as to have an equal length. Finally, the mother wavelet that best describes the data has been computed for each scenario (Section 3.1.1), along with its respective hyper-parameters.

4.2 Evaluation

The employed evaluation protocol has been k-fold cross-validation on the time dimension of each signal (Figure 4) and for $k \in \{2, 4, 8\}$. That is, for each iteration of the protocol, the total length of the signal is split into k parts, $k-1$ of which are used for training and one for testing. It should also be noted that the selection of the testing part has been consistent for all signals in a given run (e.g. the n^{th} part). Subsequently, a scaleogram has been extracted from each of the k parts (Section 3.1.2). Due to the imbalanced number of samples between the classes, both the training and evaluation of the model have been weighted. The final accuracy reported has been the weighted average of the k-folds.

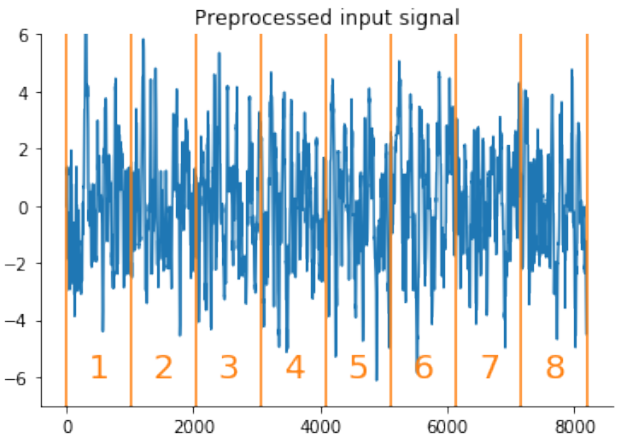

Figure 4: A preprocessed (i.e. detrended and resampled) signal, partitioned into $k = 8$ parts.

4.3 Model Selection

Since the produced scaleograms are rather simple (i.e. not exhibiting complex patterns), a very deep CNN architecture has not been considered necessary. After extensive experimentation, a network of 6 layers and around 2.5 million tunable parameters has been selected (Figure 5).

More specifically, the first convolutional layer has 32 filters, a kernel of 11×11 and strides of 4×4, while the second layer has 64 filters and a kernel size of 5×5. After both convolutions, max pooling operations are applied with kernels of 3×3 and strides of 2×2. Finally, dropout [34] is applied with a probability of $p = 0.35$. The architecture is finalized

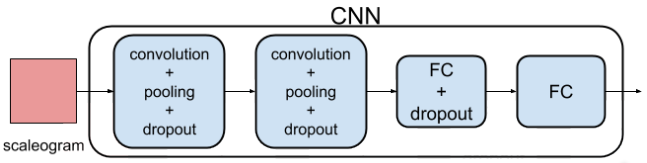

Figure 5: The CNN architecture for classifying the scaleograms.

with two *fully-connected* (FC) layers, having 256 and 4 neurons, respectively (dropout with $p = 0.3$ is applied after the first FC layer).

The *Adam* optimizer [20] has been used for network training (initial learning rate of $\eta = 10^{-4}$), with *cross-entropy* being the selected loss function. As mentioned previously, the samples have been weighted during training to alleviate the class imbalance problem.

4.4 Noise Experiments

After assessing the ability of the model to discriminate between the four classes of Section 4.1, further experiments have been performed in order to examine the robustness of the proposed technique to external noise. Specifically, each input signal has been imputed with white noise of different strengths, with the derived, "noisy" signals being processed according the previously outlined pipeline (i.e. detrending, resampling, splitting into k-folds, decomposing through DWT, converting into scaleograms and then being used for CNN training).

A total of 10 different signal-to-noise ratio (SNR) cases have been examined. The SNR of a signal is defined as the ratio of the power of the clean signal to that of the noise (i.e. the ratio of the useful information to the noise), according to Equation 5 below

$$SNR = \frac{P_{signal}}{P_{noise}} \tag{5}$$

Signal-to-noise ratios of 10^6, 10^5, 10^4, 10^3, 10^2, 10, 1, 10^{-1}, 10^{-2} and 10^{-3} have been considered for this experiment; the higher the SNR value, the "cleaner" the signal. A sample, noise-imputed input signal is depicted in Figure 6. SNRs higher than 10^3 have little effect to the signal, while for SNR lower than 10, the noise's presence in the signal is very significant.

5. RESULTS

Out of the two types of signals discussed previously (in-core and ex-core), the former are preferable (Section 4.1). However, the respective detectors are not constantly available throughout the reactor operation; they are inserted to the core for a limited amount of time. This is the reason that both types of sensors have been examined independently.

5.1 Clean signal analysis

The first experiment involves the "clean" signals of the reactor. For each type of sensor, 3 different values of k are considered, as discussed in Section 4.2. Their mean accuracy is summarized on Table 5.1.

As expected, for larger values of k, more input signals are produced and in turn more scaleograms. Furthermore, even though in-core sensors clearly outperform the ex-core ones,

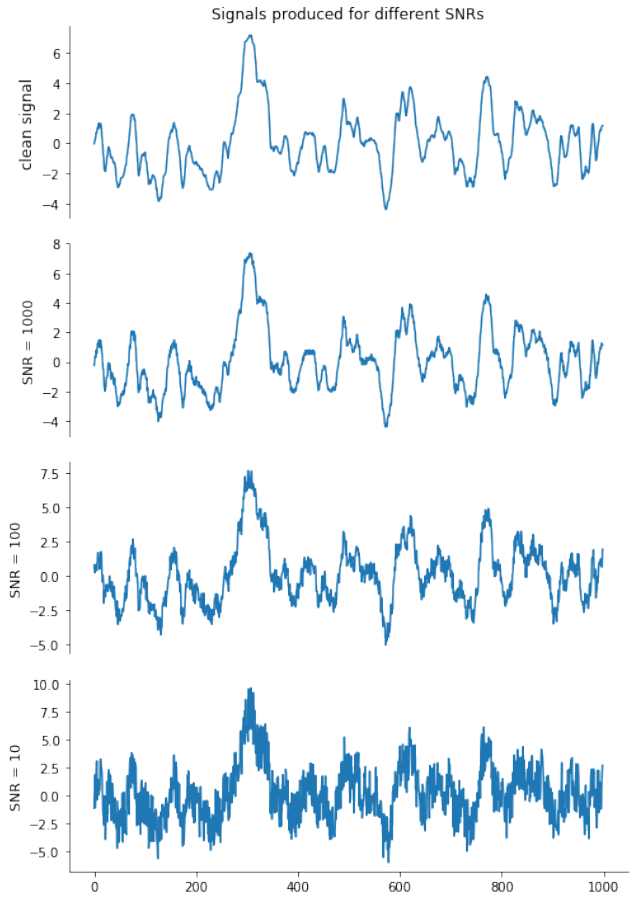

Figure 6: Input signal for different signal-to-noise ratios.

k	In-core	Ex-core
2	98.04%	83.33%
4	99.63%	83.44%
8	99.85%	93.88%

Table 1: Mean accuracy of each detector type for different values of k.

the latter remain reliable and can detect the anomalies in most cases.

5.2 Noise analysis

Table 5.2 displays the results of the noise experiments. The framework's overall performance remains unchanged for noisy signals with high values of SNR (i.e. low noise). For signals with SNR < 10, the performance starts deteriorating. From this point onwards, the noise has a strong presence in the overall signal, making it hard for the network to discriminate between samples. For very low SNR values (i.e. below 0.1), the signal is dominated by the noise, making it virtually impossible to classify it without applying denoising techniques. The aforementioned deterioration is illustrated in Figure 7. Finally, in-core signals exhibit better performance than ex-core ones in this experiment, as well.

6. CONCLUSION

	k					
	2		4		8	
SNR	In	Ex	In	Ex	In	Ex
clean	92.80	70.52	99.10	80.47	99.83	82.40
10^6	92.48	71.15	98.99	81.46	99.80	83.26
10^5	92.19	68.02	99.13	80.73	99.84	82.94
10^4	91.46	73.33	98.95	81.15	99.84	82.97
10^3	92.88	66.56	98.90	80.78	99.81	82.94
10^2	93.18	67.71	98.74	80.99	99.70	85.05
10	90.68	73.02	98.06	80.78	99.41	82.84
1	85.40	68.44	89.75	77.24	95.96	78.52
10^{-1}	72.36	23.12	73.13	53.85	70.39	57.86
10^{-2}	13.59	4.06	18.94	4.27	16.24	4.51
10^{-3}	4.08	3.33	3.64	4.06	3.64	3.98

Table 2: Weighted accuracy for both in-core and ex-core sensors, for various signal-to-noise ratios.

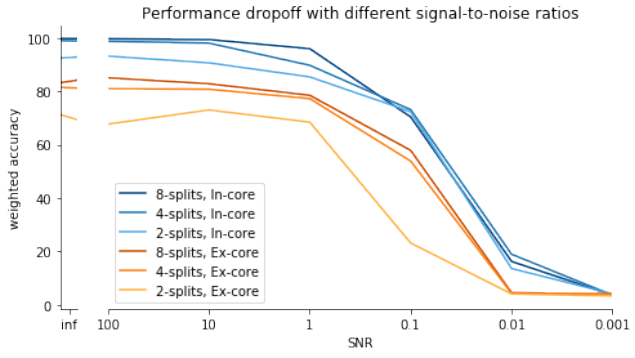

Figure 7: Performance dropoff for different signal-to-noise ratios.

In this work, a framework for detecting anomalies in nuclear reactors has been presented. Input signals are preprocessed and decomposed through DWT, in order to produce the respective scaleograms. Then, those scaleograms are treated as images that are used to train a CNN which ultimately identifies possible anomalies in the reactor. This procedure has been tested on simulated data for a PWR that contained four different types of perturbations. Both in-core and ex-core signals have been examined and the drawbacks of each have been discussed. The experimental results prove that the proposed methodology is capable of distinguishing between different anomalies in both types of signals and that it is also robust to noise.

Future research directions include the evaluation of the presented framework on actual measurements from PWRs, incorporating more types of perturbations in the model and assessing its performance on combinations of different perturbations occurring simultaneously.

Acknowledgments

The research conducted has been 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 express our gratitude to the Paul Scherrer Institute (PSI), Switzerland, for supplying the data used in our experiments, within the framework of the aforementioned CORTEX project. The data have been generated by PSI using their SIMULATE-3K tool [15, 11].

7. REFERENCES

[1] V. Arzhanov and I. Pázsit. Diagnostics of core barrel vibrations by in-core and ex-core neutron noise. *Progress in Nuclear Energy*, 43(1-4):151–158, 2003.

[2] A. M. Badshah, J. Ahmad, N. Rahim, and S. W. Baik. Speech emotion recognition from spectrograms with deep convolutional neural network. In *2017 international conference on platform technology and service (PlatCon)*, pages 1–5. IEEE, 2017.

[3] H. H. Bafroui and A. Ohadi. Application of wavelet energy and shannon entropy for feature extraction in gearbox fault detection under varying speed conditions. *Neurocomputing*, 133:437–445, 2014.

[4] P. Bernitt. *In-Core Neutron Noise Analysis for Diagnosis of Fuel Assembly Vibrations*. Chalmers University of Technology, 2008.

[5] E. O. Brigham and E. O. Brigham. *The fast Fourier transform and its applications*, volume 448. prentice Hall Englewood Cliffs, NJ, 1988.

[6] A. L. Buczak and E. Guven. A survey of data mining and machine learning methods for cyber security intrusion detection. *IEEE Communications Surveys Tutorials*, 18(2):1153–1176, Secondquarter 2016.

[7] F. Burkhardt, A. Paeschke, M. Rolfes, W. F. Sendlmeier, and B. Weiss. A database of german emotional speech. In *Ninth European Conference on Speech Communication and Technology*, 2005.

[8] F. Caliva, F. S. De Ribeiro, A. Mylonakis, C. Demaziere, P. Vinai, G. Leontidis, and S. Kollias. A deep learning approach to anomaly detection in nuclear reactors. In *2018 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 2018.

[9] A. Chan and A. Ahluwalia. Feedwater flow measurement in us nuclear power generation stations. 11 1992.

[10] V. Chandola, A. Banerjee, and V. Kumar. Anomaly detection: A survey. *ACM Comput. Surv.*, 41(3):15:1–15:58, July 2009.

[11] D. Chionis, A. Dokhane, L. Belblidia, M. Pecchia, G. Girardin, H. Ferroukhi, and A. Pautz. Simulate-3k analyses of neutron noise response to fuel assembly vibrations and thermal-hydraulics parameters fluctuations. In *M&C 2017-International Conference on Mathematics & Computational Methods Applied to Nuclear Science & Engineering, at Jeju, Korea*, 2017.

[12] I. Daubechies. The wavelet transform, time-frequency localization and signal analysis. *IEEE Trans. Inf. Theor.*, 36(5):961–1005, Sept. 2006.

[13] C. Demaziere and G. Andhill. Identification and localization of absorbers of variable strength in nuclear reactors. *Annals of Nuclear Energy*, 32(8):812 – 842, 2005.

[14] S.-W. Fu, T.-y. Hu, Y. Tsao, and X. Lu. Complex spectrogram enhancement by convolutional neural network with multi-metrics learning. In *2017 IEEE 27th International Workshop on Machine Learning for Signal Processing (MLSP)*, pages 1–6. IEEE, 2017.

[15] G. Grandi. Simulate-3k input specification (ssp-98/12, rev. 17), 2015.

[16] G. Grondey, R. Harms, G. Winderl, and H. Kumpf. Low frequency noise in a pwr and its influence on the normal operational characteristics of the plant. In *In-core instrumentation and reactor core assessment*. 1992.

[17] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[18] X. Jin, Y. Guo, S. Sarkar, A. Ray, and R. M. Edwards. Anomaly detection in nuclear power plants via symbolic dynamic filtering. *IEEE Transactions on Nuclear Science*, 58(1):277–288, 2010.

[19] P. K. Kankar, S. C. Sharma, and S. P. Harsha. Fault diagnosis of ball bearings using continuous wavelet transform. *Applied Soft Computing*, 11(2):2300–2312, 2011.

[20] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[21] D. Kollias, A. Tagaris, A. Stafylopatis, S. Kollias, and G. Tagaris. Deep neural architectures for prediction in healthcare. *Complex & Intelligent Systems*, 4(2):119–131, 2018.

[22] E. Laggiard and J. Runkel. Detection of subcooled boiling in a pwr using noise analysis and calculation of the steam void fraction. *Annals of Nuclear Energy*, 24(1):49–54, 1997.

[23] Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. *nature*, 521(7553):436, 2015.

[24] W. Li. Research on extraction of partial discharge signals based on wavelet analysis. In *2009 International Conference on Electronic Computer Technology*, pages 545–548. IEEE, 2009.

[25] X. Li, D. Song, P. Zhang, G. Yu, Y. Hou, and B. Hu. Emotion recognition from multi-channel eeg data through convolutional recurrent neural network. In *2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 352–359. IEEE, 2016.

[26] J. Ma and J. Jiang. Applications of fault detection and diagnosis methods in nuclear power plants: A review. *Progress in nuclear energy*, 53(3):255–266, 2011.

[27] C. K. Maurya and D. Toshniwal. Anomaly detection in nuclear power plant data using support vector data description. In *Proceedings of the 2014 IEEE Students' Technology Symposium*, pages 82–86. IEEE, 2014.

[28] A. Mesaros, A. Diment, B. Elizalde, T. Heittola, E. Vincent, B. Raj, and T. Virtanen. Sound event detection in the DCASE 2017 challenge. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2019. In press.

[29] I. Pázsit and C. Demazière. *Noise Techniques in Nuclear Systems*, pages 1629–1737. Springer US, Boston, MA, 2010.

[30] Z. Ren, K. Qian, Y. Wang, Z. Zhang, V. Pandit, A. Baird, and B. Schuller. Deep scalogram representations for acoustic scene classification. *IEEE/CAA Journal of Automatica Sinica*, 5(3):662–669, 2018.

[31] F. D. S. Ribeiro, F. Caliva, D. Chionis, A. Dokhane, A. Mylonakis, C. Demaziere, G. Leontidis, and S. Kollias. Towards a deep unified framework for nuclear reactor perturbation analysis. In *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 120–127, Nov 2018.

[32] A. Satt, S. Rozenberg, and R. Hoory. Efficient emotion recognition from speech using deep learning on spectrograms. In *INTERSPEECH*, pages 1089–1093, 2017.

[33] E. Sejdic, I. Djurovic, and L. Stankovic. Quantitative performance analysis of scalogram as instantaneous frequency estimator. *IEEE Transactions on Signal Processing*, 56(8):3837–3845, Aug 2008.

[34] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 15(1):1929–1958, 2014.

[35] D. M. Tax and R. P. Duin. Support vector data description. *Machine learning*, 54(1):45–66, 2004.

[36] L. Torres, D. Chionis, C. Montalvo, A. Dokhane, and A. García-Berrocal. Neutron noise analysis of simulated mechanical and thermal-hydraulic perturbations in a pwr core. *Annals of Nuclear Energy*, 126:242–252, 2019.