APPLICATION OF SIAMESE NETWORKS TO ANOMALY DETECTION IN MULTIBEAM SONOGRAMS

Jonathan Farr[1], Abdullah Olcay[1], Paul White[1], Jonathan Bull[2]

[1] *Institute of Sound and Vibration Research, Southampton University, United Kingdom*
[2] *School of Ocean and Earth Science, University of Southampton, Southampton, UK*

Abstract

Effective anomaly detection in active sonar data is critical for the long-term monitoring of offshore Carbon Capture and Storage (CCS) sites, where early identification of leaks is essential for environmental safety and regulatory compliance. Multibeam sonar systems offer high-resolution imaging of the water column but produce large amounts of data, rendering conventional reconstruction-based anomaly detection methods computationally demanding and unsuitable for bandwidth-limited deployments. This study investigates the application of Siamese convolutional neural networks for anomaly detection in multibeam sonar imagery. Instead of reconstructing input images, the proposed method learns a similarity measure from feature embeddings derived from pairs of sonar observations, enabling anomaly detection based on Euclidian distance comparisons. A dataset comprising labelled anomalous and non-anomalous multibeam sonar images was collected at an open-water test facility, with anomalies generated using controlled scattering and wake-producing events. This study explores how model configuration and training data composition influence Siamese-network-based anomaly detection in multibeam sonar imagery. Experiments were carried out using different network sizes and varying fractions of the available pairwise training combinations to examine their impact on model behaviour and training efficiency. By considering how changes in architectural complexity and combination count affect learning stability and generalisation, the work seeks to identify configurations that are well suited to resource-constrained, edge-computing environments typical of long-term CCS monitoring deployments.

Corresponding author: *Jonathan.Farr@soton.ac.uk.*

Keywords: *Anomaly Detection, Deep Learning, Siamese Networks, Multibeam Sonar*

1. Introduction

One component of climate change mitigation strategies is carbon capture and storage where CO2 is stored in depleted hydrocarbon reservoirs or in an aquifer [1]. It is important for public confidence that in the unlikely event of CO2 escape from the sub-surface that this can be detected, quantified and verified. Offshore, a number of studies have used controlled releases of CO2 through the seabed to determine the best monitoring methodologies [2]-[3]. Within the water column it has been shown that acoustic methods are best at detecting the gaseous phase, while chemical sensing techniques are optimal for the dissolved phase [3]. Passive acoustic methods can be used to detect and quantify gas bubbles emerging from the seabed provided the acoustic sensors are close enough to the source of bubbles (within c. 10 m for a single hydrophone detector; Li et al., 2020). Active acoustic methods are ideal for detecting bubble plumes within the water column [4], and can be used to quantify fluxes if the plume is insonified with a broad range of acoustic frequencies [5]. The high acoustic impedance contrast between a gaseous bubble and the water column makes them efficient scatterers of incident sound. Plumes in the water column are bright in active sonar data due to their highly increased levels of backscatter [4],[5]. The data used in this paper comes from a single frequency, multibeam active sonar which is commonly applied to bathymetry as well as water column imaging. These sonars record backscatter over substantial ranges, but create large datasets.

Near real time communication of data for the identification of leaks for investigation and remediation is important for public confidence. Edge computing of recorded active sonar data to identify potential leaks (anomalies) would limit the data required to be communicated, for example from an active sonar mounted on a seabed lander to a surface buoy. Therefore, robust automated anomaly detection methods are essential for identifying potential leak related anomalies within sonar imagery in near real time.

Autoencoder-based approaches are commonly used for anomaly detection in sonar images [6]-[7]. In these methods, a convolutional autoencoder is typically trained only on normal (non-anomalous) data [6]. The process involves reconstructing the input image so that the network learns the underlying structure of typical observations [6],[7]. During inference, anomalous inputs are expected to produce larger reconstruction errors because the network has not learned to represent such patterns during training [6], [7]. This principle has also been widely applied in various imaging tasks where anomalies are rare or poorly represented in the available datasets [8], [9] which is often the case in the standard data collection processes. It is important to note that reconstruction-based anomaly detection methods present several limitations. Anomaly detection requires a full forward pass through both the encoder and decoder followed by a pixel-wise comparison between the input and reconstructed images. In addition, these methods commonly rely on mean squared error reconstruction loss, which tends to emphasize large background regions while being less sensitive to small foreground anomalies [7]. This limitation becomes particularly important for sonar imagery, where anomalous objects typically occupy only a small portion of the image. Similar limitations of reconstruction-based anomaly detection approaches have also been discussed in broader anomaly detection literature [7]. An alternative strategy is to learn a similarity metric between observations rather than reconstructing the input image. Siamese neural networks provide a natural framework for this purpose. Bromley et al. [10] first introduced the method to signature verification, where two subnetworks were used to compare pairs of inputs and determine whether they belonged to the same class. Modern Siamese architectures typically consist of two parallel subnetworks that extract feature embeddings from paired inputs and compare them within a learned feature space [11]. In many implementations, the subnetworks share weights to enforce symmetric feature extraction, but partial or non-weight-sharing configurations are also possible as dependent on the application [12]. Siamese networks have since been successfully applied to various visual comparison tasks [13]-[14].

In this study, we use a Siamese convolutional architecture to compare normal and anomalous sonar images through their learned embeddings. Instead of reconstructing the input image, the network learns a representation in which the distance between feature vectors reflects the similarity between observations. Anomaly detection can therefore be performed through embedding comparisons rather than reconstruction errors. In practice, embeddings of normal observations can also be precomputed and stored, which means that anomaly detection during inference requires only a single encoder pass for the test image followed by a simple distance calculation. These properties make Siamese networks particularly suitable for near real time sonar anomaly detection scenarios.

This study explores the viability of Siamese networks for use in CCS monitoring by using a dataset collected at the National Physical Laboratory's open water test facility on Wraysbury Reservoir, Surrey, England.

2. Methods

2.1. Data

The data set was extracted from data collected using a Kongsberg M3 500 KHz multi-beam sonar deployed off the National Physical Laboratory's open water test facility on Wraysbury Reservoir, Surrey, England. Data was acquired at a 1 Hz ping rate with a 10,000 µs pulse duration, covering ranges between 10 and 250 m.

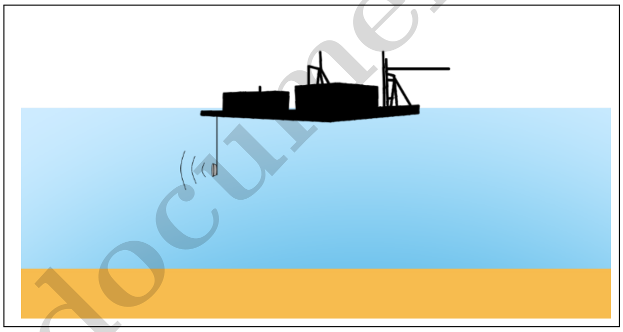

Figure 1: Diagram of dataset collection configuration at Wraysbury Reservoir, U.K.

The sonar was deployed at a depth of 5 m with the sonar swath directed towards the centre of the reservoir. Normal image samples were extracted across a 12 hour recording period. Using the same sonar position, anomalous samples, were created by positioning a boat as a scattering object on the reservoir[1]; or producing a cavitating wake in the water column[2]; or deploying small scattering objects[15]. These events were created over the total swath of the multibeam. Normal and anomalous image events were manually labelled.

2.2. Dataset arrangement

The dataset was comprised of 560 anomalous samples and 560 normal samples. These were divided into training and validation subsets using a 78:22 split, giving 436 anomalous and 436 normal samples for training, and 124 anomalous and 124 normal samples for validation.

Siamese-network inputs were generated as sample pairs. Pairs containing two normal samples were labelled as negative pairs, while pairs containing one anomalous and one normal sample were labelled as positive pairs. Pairs containing two anomalous samples were not included, as not relevant to this comparison task.

For the training set, the 436 normal samples produced 94,830 possible normal–normal pairs. To maintain

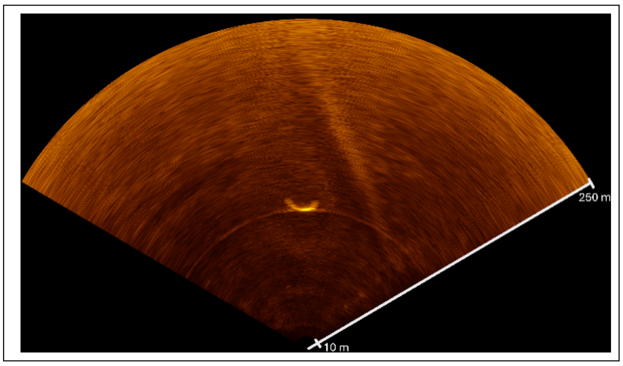

Figure 2: Example image of an anomalous sample with range labelled and the anomaly highlighted by red circle from small scattering object.

class balance, an equal number of anomalous–normal pairs were selected from the full set of possible positive pairs. This resulted in a balanced master training-pair list containing 189,660 pairs in total. Smaller training-pair lists were then generated by sampling fixed fractions of this master list while preserving the positive-to-negative class balance. These fractions and their corresponding numbers of pairs are shown in Table 1. For the validation an equal number of normal-normal and anomalous–normal pairs were selected to produce a balanced validation set of 1,952 pairs in total.

2.3. Model

The model used in this study is a Siamese convolutional neural network designed for binary anomaly classification from paired sonar images. The network takes two grayscale input images of size 150×300×1, which are normalised to the range [0, 1] prior to training. Each input image is processed by its own convolutional pathway within the subnetwork, with both pathways having identical architecture but without weight sharing. Each pathway consists of two convolutional layers followed by max-pooling operations. The first convolutional layer uses 32 filters of size 3×3 with Rectified Linear Unit (ReLU) activation function [12] with zero padding, followed by a 2×2 max-pooling layer. The second convolutional layer uses 64 filters of size 3×3, again with ReLU activation and zero padding, followed by another 2×2 max-pooling layer. The resulting feature maps are then flattened to produce a one-dimensional feature vector. The feature vectors obtained from the two pathways are compared by using Euclidean distance layer, which produces a scalar value representing the dissimilarity between the two inputs. This value is then passed to a fully connected layer with two output units and a softmax activation function to perform binary classification (Figure 3). In terms of model complexity, each convolutional pathway contains 18,816 trainable parameters, 37,632 parameters across both pathways and 37,636 trainable parameters with the final dense classification layer. It should be noted that the number of trainable parameters vary in model size experiment, the values reported here correspond to the baseline configuration used in this study.

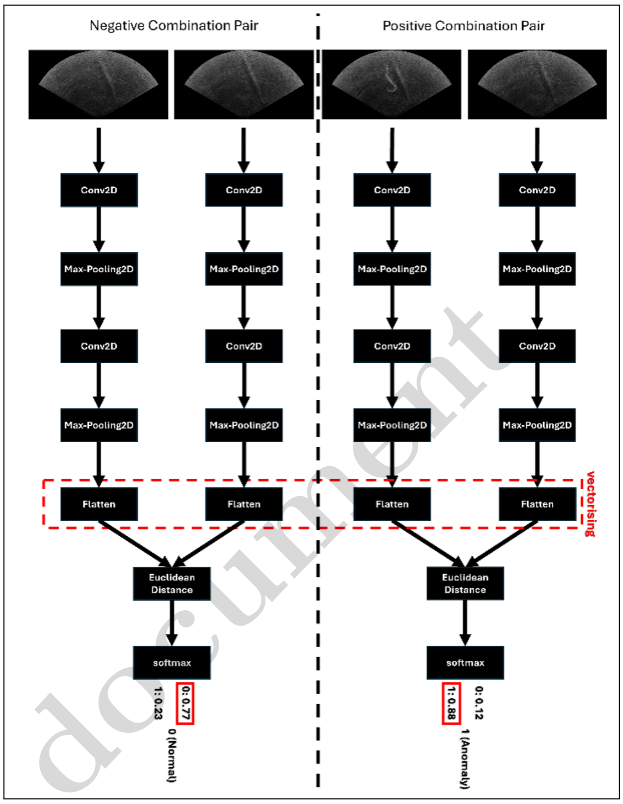

Figure 3: Siamese convolutional neural network architecture used in this study for anomaly detection. Two input images are processed by independent convolutional pathways, and their feature embeddings are compared by using Euclidean distance followed by a softmax classifier. The anomaly pair on the left (negative) comprises two normal images, while the anomaly pair on the right (positive) comprises one anomalous image (second from right) and one normal image. After the softmax classifier the image pairs are correctly identified as normal (left) or anomalous (right) based on their binary classifier percentage.

2.4. Training setup

The model is trained by using sparse categorical cross-entropy loss function with the Adaptive Moment Estimation (Adam) optimiser [16]. Training was performed with a batch size of 16 for up to 100 epochs. Early stopping with a patience of 10 epochs was applied based on validation loss, and the best-performing model was saved. The reported validation accuracy values in the study are the maximum validation accuracies achieved during the training. This approach was used as validation performance can fluctuate across epochs due to the randomness in model parameter initialisation and mini-batch sampling during the training. Also, early stopping criteria has been applied, which can result in slightly different stopping points across the training runs due to fluctuations in the validation error curve.

3. Results

3.1. Training dataset size effect

The Siamese model described in Section 2.3 was trained using different numbers of combinations. The peak validation accuracy achieved across 6 training attempts was recorded and displayed in Table 1. Across the training dataset size experiments, validation accuracy values remain above 99 %, with only small variations (0.48 % standard deviation). This shows the effectiveness and robustness of the proposed model. The small performance differences are mainly due to changes in the diversity of the combined samples included in each training subset. In some cases, the added combinations may reduce the representativeness of the training data with respect to the validation dataset, which slightly affects the model's generalisability.

Table 1: Validation accuracy for different fractions of the available balanced combinations.

Fraction of total possible balanced combinations	Number of combinations	Validation accuracy (%)
1/2	47 414	99.6
1/4	23 706	99.4
1/8	11 852	99.6
1/10	9 482	98.8

3.2. Model Size

Another important factor for understanding the efficacy of the Siamese network is the model size, as it directly affects the computational cost required to achieve efficient anomaly detection performance. The Siamese model described in Section 2.3 was adapted by changing the number of filters in the 2D convolutional layers. Models with each number of parameters were trained 6 times with the dataset containing ¼ of the total possible data sample combinations. The total number of parameters, number of filters used in each convolutional layer of the model and the maximum validation accuracy achieved are shown in Table 2. The validation accuracy consistently increases with the model size from 5k to 40k reaching a peak value of 99.6 %. As model size continues to increase, detection performance slightly degrades up to 400k followed by a steep decline in accuracy at 800k. Considering that larger models are generally more prone to overfitting [17], our results suggest that overfitting starts to emerge for the Siamese network after approximately 40k. However, even the 10k model achieves above 99 % validation accuracy. This indicates that the Siamese network can show effective anomaly detection performance with a relatively small model size, which makes the method suitable for real time implementations. The increase for the 400k observed is likely due to the randomness in training or an optimisation artifact, where 160k model may have converged to a poorer local minimum. Nevertheless, the overall results indicate that increasing model size beyond a relatively small network does not consistently improve performance.

Table 2: Number of model parameters and the maximum validation accuracy achieved.

Model Parameters	Number of Filters in Layers		Validation Accuracy (%)
5k	12	24	93.5
10k	16	32	99.2
40k	32	64	99.6
80k	47	94	97.1
160k	66	132	93.4
400k	105	210	96.8
800k	150	300	69.3

4. Conclusion

This study demonstrates that Siamese convolutional neural networks may provide an effective approach for anomaly detection in multibeam sonar imagery for CCS monitoring. By learning a similarity-based embeddings rather than relying on reconstruction, the method overcomes key limitations of autoencoder approaches in multibeam sonar data. The results show consistently high performance when trained on multibeam sonar data of simulated gaseous-like anomalies, with validation accuracies exceeding 99 % across varying dataset and model sizes. High performance was also shown with compact models and small datasets. This reinforces the stability of the approach for computationally constrained, near real-time edge computing applications. Future work should focus on validating the method using data from different near-seabed datasets, including above CCS storage complexes, to improve generalisation and operational reliability in real-world monitoring scenarios.

5. Acknowledgments

This work was supported by the National Physical Laboratory, UK Government Department of Energy Security and Net Zero OESEA fund, and the Danish Agency for Science, Technology and Innovation.

References

[1] B. Metz, O. Davidson, H. de Coninck, M. Loos, and L. Meyer, "Carbon dioxide capture and storage," *The Intergovermental Panel on Climate Change*, Jan. 2005.

[2] J. Blackford et al., "Detection and impacts of leakage from sub-seafloor deep geological carbon dioxide storage," *Nature Climate Change*, vol. 4, no. 11, pp. 1011–1016, 2014, ISSN: 1758-6798. DOI: `10.1038/nclimate2381`. [Online]. Available: `https://doi.org/10.1038/nclimate2381`.

[3] A. Flohr et al., "Towards improved monitoring of offshore carbon storage: A real-world field experiment detecting a controlled sub-seafloor co2 release," *International Journal of Greenhouse*

Gas Control, vol. 106, p. 103 237, 2021, ISSN: 1750-5836. DOI: https://doi.org/10.1016/j.ijggc.2020.103237. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1750583620306629.

[4] M. Dean, J. Blackford, D. Connelly, and R. Hines, "Insights and guidance for offshore co2 storage monitoring based on the qics, eti mmv, and stemm-ccs projects," *International Journal of Greenhouse Gas Control*, vol. 100, p. 103 120, 2020, ISSN: 1750-5836. DOI: https://doi.org/10.1016/j.ijggc.2020.103120. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1750583620305454.

[5] J. Li et al., "Broadband acoustic inversion for gas flux quantification—application to a methane plume at scanner pockmark, central north sea," *Journal of Geophysical Research: Oceans*, vol. 125, no. 9, e2020JC016360, 2020. DOI: https://doi.org/10.1029/2020JC016360.

[6] J. Coffelt and J. Christensen, "Anomaly detection in side-scan sonar," Sep. 2021, pp. 1–6. DOI: 10.23919/OCEANS44145.2021.9705947.

[7] R. Bouman and T. Heskes, *Autoencoders for anomaly detection are unreliable*, 2025. arXiv: 2501.13864 [cs.LG]. [Online]. Available: https://arxiv.org/abs/2501.13864.

[8] A. Nawaz, S. S. Khan, and A. Ahmad, "Ensemble of autoencoders for anomaly detection in biomedical data: A narrative review," *IEEE Access*, vol. 12, pp. 17 273–17 289, 2024. DOI: 10.1109/ACCESS.2024.3360691.

[9] S. S. Khan et al., "Unsupervised deep learning to detect agitation from videos in people with dementia," *IEEE Access*, vol. 10, pp. 10 349–10 358, 2022. DOI: 10.1109/ACCESS.2022.3143990.

[10] J. Bromley et al., "Signature verification using a "siamese" time delay neural network," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 7, p. 25, Aug. 1993. DOI: 10.1142/S0218001493000339.

[11] Y. Li, C. L. P. Chen, and T. Zhang, "A survey on siamese network: Methodologies, applications, and opportunities," *IEEE Transactions on Artificial Intelligence*, vol. 3, no. 6, pp. 994–1014, 2022. DOI: 10.1109/TAI.2022.3207112.

[12] M. Mohazzebi, M. Eftekhari, and M. Shariatzadeh, "How weight-sharing mechanisms affect the performance of deep siamese networks," in *2022 9th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS)*, 2022, pp. 1–6. DOI: 10.1109/CFIS54774.2022.9756423.

[13] R. Caye Daudt, B. Le Saux, and A. Boulch, "Fully convolutional siamese networks for change detection," in *2018 25th IEEE International Conference on Image Processing (ICIP)*, 2018, pp. 4063–4067. DOI: 10.1109/ICIP.2018.8451652.

[14] X. Chen and K. He, *Exploring simple siamese representation learning*, 2020. arXiv: 2011.10566 [cs.CV]. [Online]. Available: https://arxiv.org/abs/2011.10566.

[15] D. Connelly et al., "Assuring the integrity of offshore carbon dioxide storage," *Renewable and Sustainable Energy Reviews*, vol. 166, p. 112 670, 2022, ISSN: 1364-0321. DOI: https://doi.org/10.1016/j.rser.2022.112670. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364032122005627.

[16] D. P. Kingma and J. Ba, *Adam: A method for stochastic optimization*, 2017. arXiv: 1412.6980 [cs.LG]. [Online]. Available: https://arxiv.org/abs/1412.6980.

[17] M. Belkin, D. Hsu, S. Ma, and S. Mandal, "Reconciling modern machine-learning practice and the classical bias–variance trade-off," *Proceedings of the National Academy of Sciences*, vol. 116, no. 32, pp. 15 849–15 854, Jul. 2019, ISSN: 1091-6490. DOI: 10.1073/pnas.1903070116. [Online]. Available: http://dx.doi.org/10.1073/pnas.1903070116.