The 9th European Congress on Computational Methods in Applied Sciences and Engineering
ECCOMAS Congress 2024
3 – 7 June 2024, Lisboa, Portugal

ON THE TRAINING OF AI ALGORITHMS USING FINITE-ELEMENT COMPUTATION DATA FOR DAMAGE IDENTIFICATION IN SENSORISED COMPOSITE STRUCTURES

ROHAN CHABUKSWAR*, CHLOE MULLEN†, KOSTAS KOURAMAS‡

Collins Aerospace, Cork T23 XN53, Ireland
*email: rohan.chabukswar@collins.com
†email: chloe.mullen@collins.com
‡ email: konstantinos.kouramas@collins.com

Key words: Structural health monitoring, embedded microwires, machine learning, intelligent structure.

Summary. Embedding microwires into composite materials is a novel technology that can provide in-situ and remote data on the structural health of the system. Damages such as delamination, disbond, and cracks in the composite structure induce a change in the stresses and strains on the microwires, modifying their electromagnetic response. Analysing this response with artificial intelligence can potentially allow us to not only detect these damages, but also characterise, localise, and quantify the damage, and possibly predict the remaining useful life of the structure. As part of the Horizon Europe project INFINITE, this work aims to develop methodologies for in-service structural health monitoring of composites equipped with microwire sensors.

This work was funded by the European Union under the Horizon Europe grant 101056884.

1. INTRODUCTION

In 2015, the top 13 US airlines spent USD 2 billion in delays and cancellations, and USD 3 billion in unscheduled maintenance, schedule interruptions thus garnering a total cost of USD 5 billion. 70% of the delays and cancellations were driven by failures of Line Replaceable Units (LRUs), 25% of the drivers of which, such as environmental factors (severe and nominal weather, pollution), malfunction, and wear-and-tear can be predicted [1]. Airlines spend an average of USD 870 in direct maintenance costs for every flight hour. 40% of this cost is for engine maintenance, and 30% each for airframe and component maintenance. Of these costs, 22% are labour cost, and 60% are material costs [2]. In 2016, the aircraft maintenance supply chain held an inventory of about USD 44 billion for an active global fleet of nearly 17,000 aircraft - approximately USD 2.5 million inventory per active aircraft [3]. In 2016, the warranty reserves held worldwide by US-based manufacturers was USD 2.1 billion for aerospace Original Equipment Manufacturers (OEMs) and USD 2.3 billion for aerospace suppliers [4]. In case of failure, analysis to find the fault and awaiting replacement parts can incur long lead-times thereby causing schedule interruptions for airlines.

Structural Health Monitoring (SHM) involves the observation and analysis of a structural system over time using periodically sampled response measurements to monitor changes of the

material and geometric properties. Electrical [5] or optical [6] strain gauges need to be permanently connected to a data acquisition system (DAQ) to monitor the development of strains in structural members to record and evaluate the gradual development of the strains during the loading process. The size of these DAQs or the presence of extra electronic circuity along with sensors such as batteries or wireless connection precludes the embedding of such sensors inside materials. Moreover, this further limits the inclusion of such sensors in aerospace applications, where weight and volume are severely restricted [7].

Giant Magneto Impedance (GMI) effect measurement using magnetostriction can be used for mechanical stress detection. Using amorphous bistable glass-coated microwires, it is feasible to perform contactless measurements of strain, even when embedded inside composite structures [8]. Any damage to the composite structures that causes a change in the internal strains of the microwires can thus be identified, characterised, quantified, and localised – but the propagation of these damage properties into the values of the strain is a complex physical process, the data-driven inverting of which can be handled by machine learning (ML) approaches such as neural networks (NNs).

The project proposes to employ hybrid physics-based, and ML based methodologies for predictive condition-based maintenance to enable continuous monitoring of components and systems, specifically using embedded microwires to automatically estimate asset health, detect degradation, and predict remaining useful life (RUL). A proactive approach to maintenance will help to reduce loss for airlines.

Motivating examples of aircraft composite components that would benefit from in situ health monitoring are fan cowls, thrust reversers, inner inlet barrel, etc. Fan cowls are easily accessible and can be monitored with non-destructive testing, however, as they receive the widest range of damages the most frequently, data collected from this component will be extremely useful in identifying damages in the other two motivating components which aren't easily accessible. Common composite damages include nicks, scratches, gouges, dents, delamination, disbond, and leading-edge erosion.

2. CHALLENGES & INNOVATIONS

Neural networks can identify complex patterns in data that humans may not be able to perceive. However, the black-box nature of artificial intelligence (AI) systems regarding the path to conclusions or decisions presents a significant impediment to the acceptance of AI-based systems [9], especially in the aerospace domain. Abstracting key patterns identified in a deep learning model as actual features by using domain knowledge would make it possible to breakdown the model into more explainable pieces [10], paving a path towards validation and certification.

Convolution leverages three important ideas that can help improve a ML system: sparse interactions, parameter sharing, and equivariant representations [11]. Convolutional Neural Networks (CNNs) use convolutional operations such as general pooling, subsampling, averaging, and many geometric operators, to extract features from images [12]. Since large depth-2 neural networks approximate a continuous function on a bounded domain [13, 14, 15], these seemingly abstract convolutional operators can be reinterpreted as mathematical functions being carried out on the data. Conversely, any physical restrictions on relationships in the data can be represented as convolutional filters. If these filters are being trained, constraints on these

filters in terms of dimensions, sizes, and weight values can be implemented to improve the training and performance. These restrictions have a secondary effect of reducing the learnable parameters of the network, thereby reducing overfitting to sparse data. Moreover, since each filter learned in this way can be interpreted as a physical function, it is possible (albeit unwieldy) to compare and adjust the weights using domain knowledge.

Another challenge encountered while carrying out this work was the lack of experimental data. Due to the concurrent nature of the project activities, and an unexpected amount of effort required to achieve an acceptable level of signal-to-noise ratio, the manufacture of portable readers capable of acquiring the required data has been delayed. This necessitated the use of available simulation data to down select the structure of the CNN, characterise, and coarsely tune the hyperparameters, while leaving the fine-tuning to such a time as experimental data would be available. Moreover, since simulation data is deterministic and noise-free, and NNs don't train or perform well under low data and noise-free conditions, the simulation data needed to be augmented by addition of random noise. This further exacerbated the requirement of restricting the NNs appropriately to prevent overfitting to variables such as simulation parameters and noise characteristics.

3. EXPERIMENTAL AND SIMULATION SETUP

The setup to acquire electromagnetic data is shown in **Figure 1** [16]. Further in the project, this setup will be done in test conditions where the composite sample will be under a loading stress. This stress will be replicated in the Finite Element Method (FEM) simulation, so that the data will be comparable.

Since there are no strain sensors (apart from the microwires), the experiments will not give any independent data about the strains within the composite coupon. Conversely, due to the nature of FEM simulations, it is not possible to model the variation in EM behaviour of the microwires due to strain. As a result, the electromagnetic data in the experiments will be assumed to correspond directly to the strains as simulated for the same conditions.

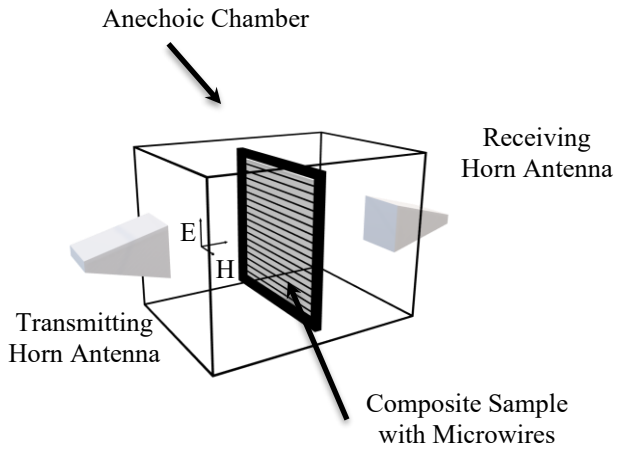

Figure 1: Experimental setup *[16]*. A free space measurement system is used consisting of two broadband horn antennas fixed to a mini anechoic chamber covered inside with a microwave absorber. To provide external stimuli the composite samples were externally loaded or embedded into a planar magnetic coil.

The simulation data available for initial efforts was generated using FEM. The simulated structure consists of a composite coupon of 8 plies of 30 mm × 30 mm × 0.25 mm with one microwire between each layer at different angles to the fibre direction: -45° at the top and bottom of the structure, 45° between plies 1 & 2 and 7 & 8, 90° between plies 2 & 3 and 6 & 7, and

0° between plies 3 & 4 and 5 & 6. There was no microwire simulated between plies 4 & 5. The finite elements were sized at 1 mm × 1 mm × 0.0625 mm for the composite (for 30 × 30 × 32 = 28800 elements) and ⌀ 0.02 mm × 0.625 mm for the microwire (for 60 elements). The structure was constrained at one end, and a bending moment of 2 N·m was applied at the other to simulate loading stress. The simulation was repeated by adding delamination centred at one of the two constrained corners between plies 1 & 2, once with radius 10 mm and once with 20 mm (resemble a quarter circle of delamination).

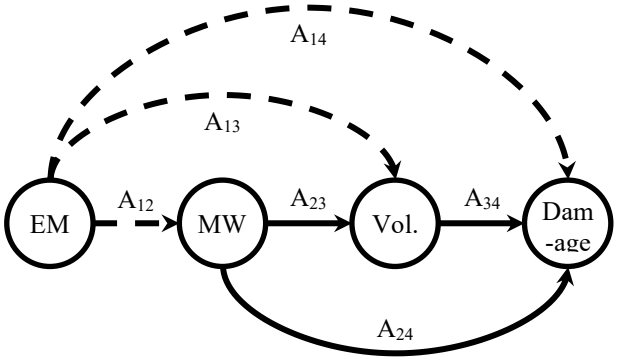

Figure 2: Several pathways can be taken from measurement of electromagnetic response to damage identification, through none, one, or both of microwire and volumetric strains or stresses. Due to the unavailability of electromagnetic data, the paths A_{12}, A_{13}, and A_{14} (denoted by the dashed lines) cannot currently be trained, leaving only A_{23}, A_{34}, and A_{24}.

The aim of this work is to construct a minimal explainable CNN to differentiate between the three conditions (henceforth referred to as undamaged, 10 mm, and 20 mm), and characterise the hyperparameters in anticipation of availability of electromagnetic data. The 3D distribution of von Mises strain and stress in the composite (referred to as volumetric data) and the axial strain and stress along each microwire were used separately to this end.

4. EXPLAINABILITY

The physics, in this case, proceeds in the following manner: the specific damage (or fault) causes a change in the volumetric strain, which causes a change in the microwire strain. This change in strain causes a change in the microwire electromagnetic response. The analysis, therefore, must start with the electromagnetic response, proceed to estimate the microwire and thence the volumetric strains, and thus identify the damage.

There are several pathways that could be taken: $A_{12} \circ A_{23} \circ A_{34}$, $A_{13} \circ A_{34}$, $A_{12} \circ A_{24}$, or A_{14}.The intermediate strain estimations serve to increase the modularity of the algorithms – the same networks could theoretically be used with a different method of measuring strain. They also enable a sanity-check of the network outputs, increasing their explainability; however, they require more training data.

Due to the unavailability of electromagnetic data, the paths A_{12}, A_{13}, and A_{14} (denoted by the dashed lines) cannot currently be trained, leaving only A_{23}, A_{34}, and A_{24}. This work focuses on training A_{34}, and A_{24}, leaving the choice of the final path and the fine-tuning to a future time when electromagnetic data is available.

5. PHYSICS-DRIVEN SIMPLIFICATIONS

The volumetric strain data has a 3-dimensional structure and would ideally be analysed in all three dimensions. The base starting CNN structure is shown in Figure 3. Given the dimensions of the simulation data, this structure consists of ~550K learnable parameters and tries to learn physical functions in all 3 dimensions.

However, the layers of the composite are loosely coupled and have similar internal relationships. Thus, rather than looking at the intra-layer coupling, a simplification of the

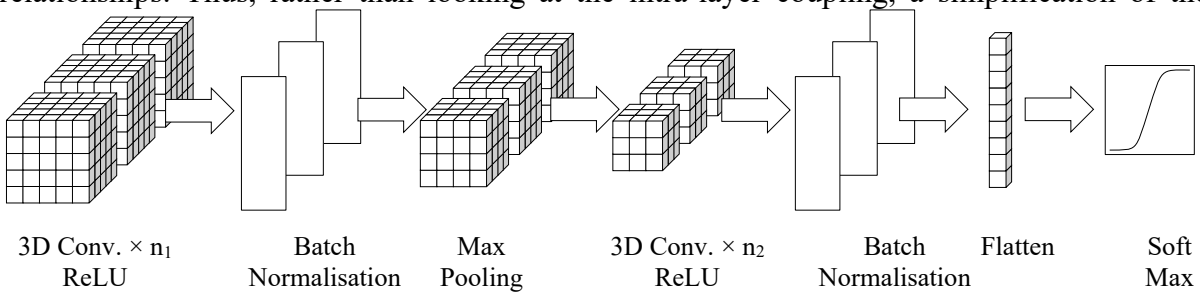

Figure 3: Original 3D CNN Structure for Volumetric Data showing two sets of convolution and normalisation layers separated by a maximum pooling layer and followed by a flattening layer and a softmax layer for output.

network is possible, where the CNN is structured to use 3D convolution with a 2D kernel (denoted in shorthand as 2.5D). This would also reduce the need for the second convolution layer, thereby yielding just ~85K learnable parameters, and thus reducing the chances of overfitting.

The microwire strain data also has a 3-dimensional span, however, each microwire has a single dimension with the same number of elements (despite a variation in length). The microwires themselves can therefore be compacted as a 2-dimensional structure. Since this is already a smaller data size, the second convolution layer can be dropped, using ~18K learnable parameters.

Once again, each microwire is loosely coupled to the others. Moreover, in the final stages of the project, the composite coupon will only have one microwire across the thickness, to prevent cross-contamination of measurements. Thus, the network can be simplified further, using a 2D convolution with a 1D kernel (denoted in shorthand as 1.5D). This further reduces the number of learnable parameters to ~10K.

6. HYPERPARAMETERS AND THEIR CHARACTERISATION

This section lists the hyperparameters in the CNN structure and their characterisation, to enable fine-tuning of the networks with experimental data. The hyperparameters were optimised in three steps based on their effects:

1. Standard deviation of noise for data augmentation,
2. Optimisation algorithm & activation function, and finally,
3. Number of filters and convolution filter size.

The tuning for each step was finalised before proceeding to the next.

6.1. Data Augmentation

Simulated models inherently result in noiseless data. To make the data as realistic as possible as well as to enable training of neural networks, varying degrees of white Gaussian noise was added to each simulation (undamaged, 10 mm, and 20 mm) to create 1000 augmented samples. The different standard deviations for noise parameters can be seen below in **Table 1** for microwire data and **Table 2** for volumetric data.

Table 1: Standard Deviation of Noise added to Microwire Data

		Noise in Strain Data (σ) (με)				Noise in Stress Data (σ) (Pa)			
		25	**50**	**100**	**180**	**2.5**	**3**	**5**	**10**
1.5D	**Loss**	0.049	0.072	0.37	1.54	0.067	0.049	0.08	0.39
1.5D	**Accuracy**	0.998	0.987	0.83	0.551	1	1	0.998	0.82

Table 2: Standard Deviation of Noise added to Volumetric Data

		Noise in Strain Data (σ) (με)			Noise in Stress Data (σ) (Pa)		
		20	**40**	**80**	**5**	**10**	**20**
2.5D	**Loss**	0	0.00015	0.0003	0.0006	0.00016	0.0012
2.5D	**Accuracy**	1	1	1	1	1	1
3D	**Loss**	0.000003	0.000125	0.00015	4.89	0.066	0.025
3D	**Accuracy**	0.327	0.98	0.993	0.33	0.98	0.99

6.2. Optimisation Algorithm

While using Deep Learning, there are several options of optimisation algorithms to use. The ones considered are:

- Stochastic Gradient Descent (SGD): Uses single records to update the weights. SGD is usually slow to converge and the path to reach global minima becomes very noisy.
- Adaptive Gradient Algorithm (AdaGrad): Uses different learning rates for each weight base on iteration. While using AdaGrad, due to monotonically decreasing learning rates, the model stops learning after some time.
- Adaptive Learning Rate with Momentum (AdaM): AdaM is extension of Adagrad that attempts to solve its radically diminishing learning rates, as well as using exponentially weighted averages to compute gradients, to minimise the effects of oscillations.

Each of the algorithms have their pros and cons that depend on the data available [17].

6.3. Activation Functions

An activation function is one that maps the output of a node to values in a fixed range. A neural network may learn to simulate nonlinear relationships between the input and output variables by introducing nonlinearity through an activation function, thereby recognising more intricate patterns. The activation functions considered are given in **Table 3**.

Table 3: Activation Functions & Their Properties (summarised from *[18]*)

Function	Mathematical Form	Range	Monotonic	Potential Issues
ReLU	$R(x) = \max(0, x)$	$[0, \infty)$	✗	Dead neurons
softmax	$S(x_i) = \frac{e^{x_i}}{\sum_{j=1}^{k} e^{x_j}}$	(0, 1)	✓	Numerical instability
tanh	$\tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}$	(-1, 1)	✓	Output saturation

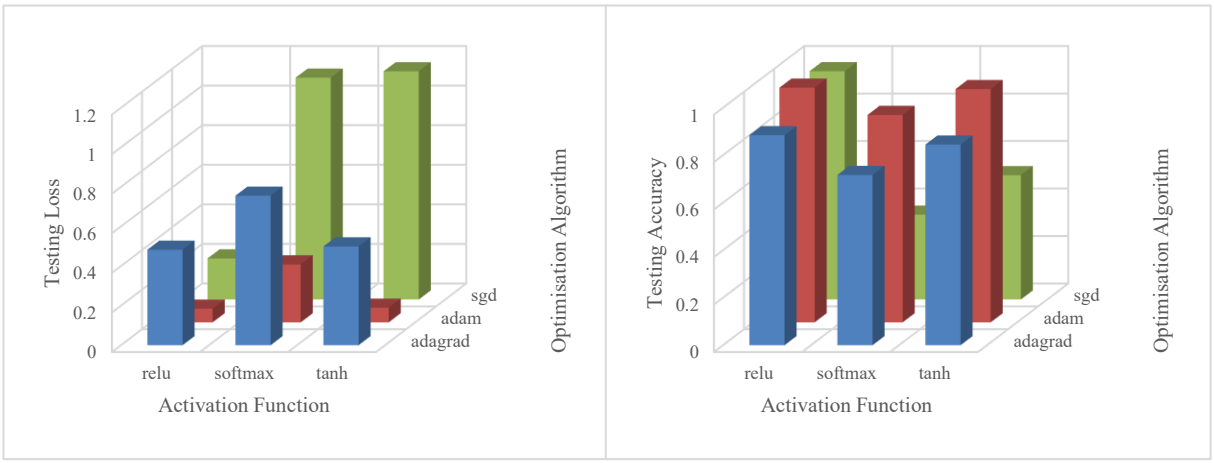

Figure 4: Average testing loss & accuracy for microwire data across optimisation algorithms and activation functions over 100 runs. The combination of AdaM optimisation algorithm and ReLU activation function performed significantly better.

In general, the effects of changing the optimisation algorithm and the activation function go together – it is difficult to separate out the effects of one from the other. Thus, for each of the combinations of an optimisation algorithm and an activation function, the network was trained and validated 100 times for both microwire (**Figure 4**) and volumetric data (**Figure 5**), and the testing loss and accuracy were compared. It can be seen that in both cases, the combination of the AdaM optimisation algorithm and the ReLU activation function performed significantly better. These hyperparameters were chosen for the last tuning step.

6.4. Number of Filters

The CNN training process starts with a number (denoted by n_1 in **Figure 3**) of randomised convolutional filters, whose weights are optimised as the training progresses, to mimic physical functions. Increasing the number of filters increases the number of physical functions that can be utilised in the final expression, as well as increasing the probability that the right physical

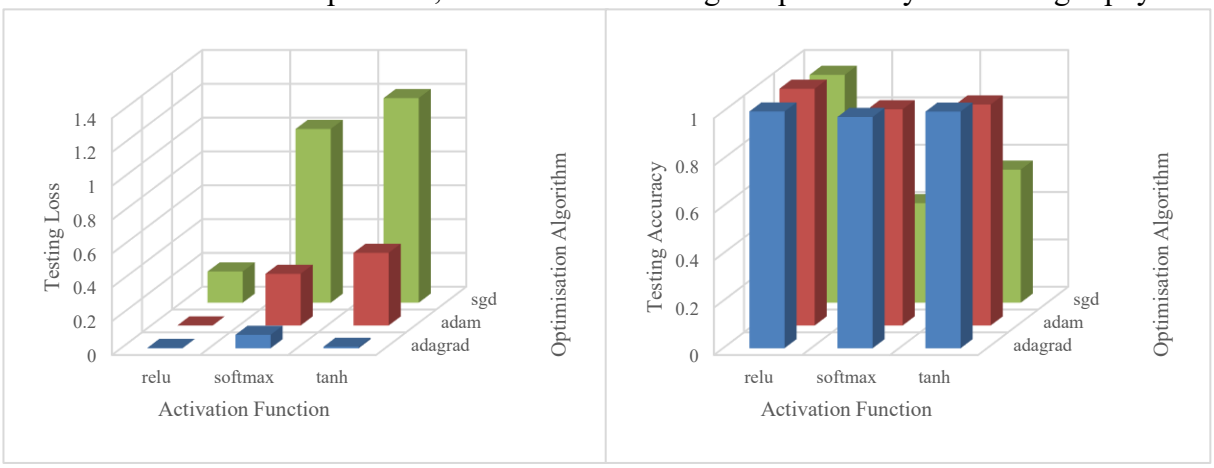

Figure 5: Average testing loss & accuracy for volumetric data across optimisation algorithms and activation functions over 100 runs. The combination of AdaM optimisation algorithm and ReLU activation function performed significantly better.

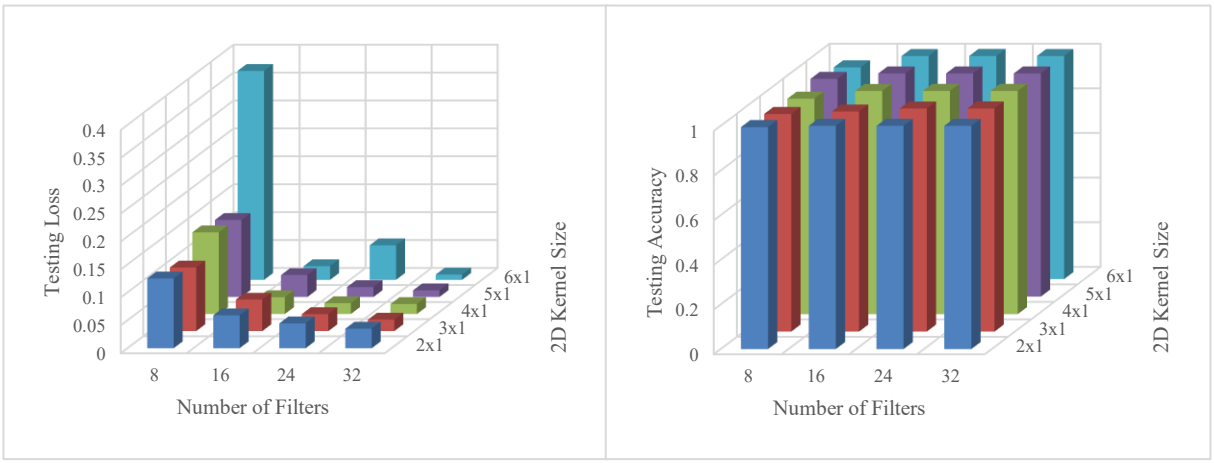

Figure 6: Average testing loss & accuracy for microwire data across number of filters and kernel size over 100 runs. The accuracy for all cases was almost 100%, while the loss decreased almost monotonically with increasing number of filters and increasing kernel size (the variations can be attributed to stochastic error).

functions are learnt. However, it also increases the training time and computational requirements, as well as increasing the chances of overfitting to the training data and reducing the explainability of the final network. Thus, initialising the network with a judicious number of filters is essential. For the microwire data, the number of filters was varied from 8 through 32 in steps of 8, whereas for the volumetric data, the number of filters was varied from 16 through 128 in powers of 2.

6.5. Kernel Size

The kernel size denotes the length and height of the matrix used as the convolutional filter. A larger kernel increases the complexity of the filters, allowing for more complex physical functions to be learnt during training. However, a larger kernel increases training time and computational requirements, as well as increasing the chances of overfitting to the training data and reducing the explainability of the final network. Thus, choosing the right size of the kernel

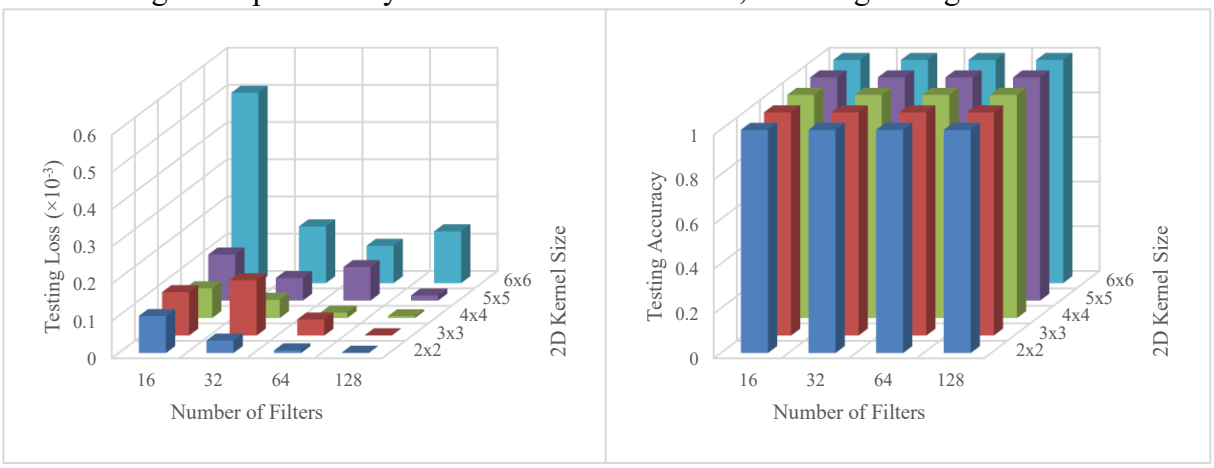

Figure 7: Average testing loss & accuracy for volumetric data across number of filters and kernel size over 100 runs. To experimental accuracy, there was essentially no difference in the performance – all the combinations gave 100% accuracy and infinitesimal loss.

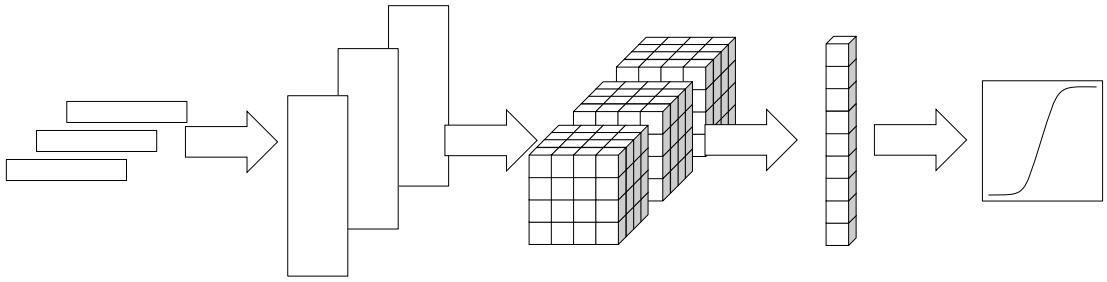

Figure 8: Final structure of network for microwire data. There is only one set of a 1D convolutional and a batch normalisation layers, followed by a maximum pooling layer, a flattening layer, and a softmax layer for output.

is necessary. For the microwire data, where the filter is a vector, the size was varied from 2 × 1 through 6 × 1, whereas for the volumetric data, where the filter is a square matrix, the size was varied from 2 × 2 through 6 × 6.

As the number of filters and the kernel size both represent the physical functions learnt, it was reasonable to optimise them together. Thus, for each combination of a number of filters and a kernel size, the network was trained and validated 100 times for both microwire (**Figure 6**) and volumetric data (**Figure 7**), and the testing loss and accuracy were compared. For the microwire data, the accuracy for all cases was almost 100%, while the loss decreased almost monotonically with increasing number of filters and increasing kernel size (the variations from monotonicity can be attributed to stochastic error). For the volumetric data, to experimental accuracy, there was essentially no difference in the performance – all the combinations gave 100% accuracy and infinitesimal loss.

This performance is attributed to the nature of simulation data, and further tuning of the number of filters and kernel size needs to be done with experimental data.

7. FINALISED STRUCTURES

The finalised structures for the networks are shown in **Figure 8** for the microwire data and in **Figure 9** for the volumetric data. The average training times for the networks were almost constant at ~5 seconds for the microwire data and from 0.5 to 1.5 minutes for the volumetric

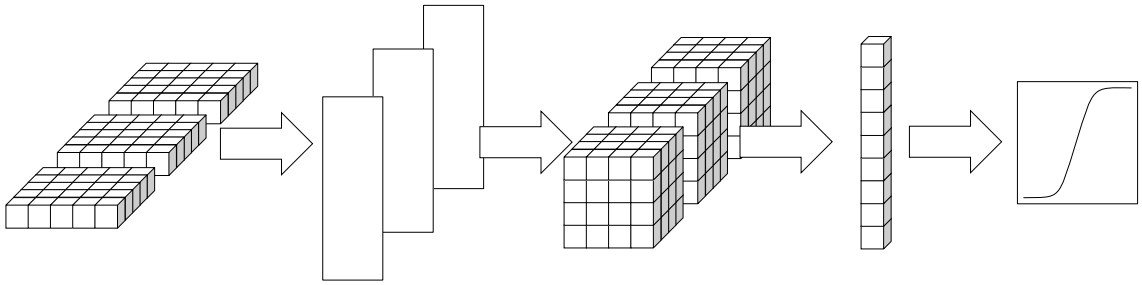

Figure 9: Final structure of network for volumetric data. There is only one set of a 2D convolutional and a batch normalisation layers, followed by a maximum pooling layer, a flattening layer, and a softmax layer for output.

sdata (increasing with increasing number of filters and kernel size).

8. CONCLUSION & FUTURE WORK

This work demonstrates a pathway of training convolutional neural networks for classifying physics-based damages in composites in a robust and explainable way in the absence of experimental data. While the fine-tuning will need to be done with experimental data, it is possible to perform the initial hyperparameter optimisation using finite-element simulations.

The next steps will involve creating and following a test plan for the project, generating FEM simulation data for the strains and experimental data for the electromagnetic response. The test plan will also include more faults with increasing extents, and the full dataset will be used to fine-tune the network for detecting, identifying, and quantifying damages in sensorised composite coupons.

Outside the scope of the INFINITE project, results will be extended to composite aerospace structures to enable condition-based proactive maintenance and reducing the cost of reactive maintenance.

9. REFERENCES

[1] US Bureau of Transportation Statistics, US Dept. of Transportation, “Transportation Statistics Annual Report US DOT Form 41,” 2016.

[2] International Air Transport Association (IATA) Maintenance Cost Task Force (MCTF), “Airline Maintenance Cost Executive Commentary—An Exclusive Benchmark Analysis of Maintenance Cost Task Force (MCTF) FY 2013,” 2014.

[3] D. Evans, “Aftermarket Outlook,” May 2006. [Online]. Available: http://www.aviationtoday.com/am/categories/rotocraft/273.html.

[4] Warranty Week, “Aerospace Warranty Expense Report,” 13 April 2017. [Online]. Available: http://www.warrantyweek.com/archive/ww20170413.html.

[5] L. D. Vijay Anand, D. Hepsiba, S. Palaniappan, B. Sumathy, P. Vijayakumar and S. Sheeba Rani, “Automatic strain sensing measurement on steel beam using strain gauge,” 2021.

[6] O. G. Morozov, I. I. Nureev, A. Z. Sakhabutdinov, R. S. Misbakhov and A. A. Kuznetcov, “Advanced Microwave Photonics Sensor Systems: Address FBG Sensors, Interrogation and Calibration,” in *2022 Systems of Signal Synchronization, Generating and Processing in Telecommunications (SYNCHROINFO)*, 2022.

[7] M. Al Ali, P. Platko, V. Bajzecerova, S. Kusnir, S. Kmet, S. Nalevanko, A. Spegarova, L. Galdun and R. Varga, “Application of bistable glass-coated microwire for monitoring and measuring the deformations of metal structural members,” *Measurement,* vol. 208, 2023.

[8] D. Praslicka, J. Blazek, M. Smelko, J. Hudak, A. Cverha, I. Mikita, R. Varga and A. Zhukov, “Possibilities of Measuring Stress and Health Monitoring in Materials Using Contact-Less Sensor Based on Magnetic Microwires,” *IEEE Transactions on Magnetics,* vol. 49, no. 1, p. 128–131, 2013.

[9] A. P. Saraf, K. Chan, M. Popish, J. Browder and J. Schade, "Explainable Artificial Intelligence for Aviation Safety Applications," in *AIAA Aviation 2020 Forum*.

[10] B. Shukla, I.-S. Fan and I. Jennions, "Opportunities for Explainable Artificial Intelligence in Aerospace Predictive Maintenance," in *PHM Society European Conference*, 2020.

[11] I. Goodfellow, Y. Bengio and A. Courville, Deep Learning, MIT Press, 2016.

[12] T. Wiatowski and H. Bölcskei, "A Mathematical Theory of Deep Convolutional Neural Networks for Feature Extraction," *IEEE Transactions on Information Theory,* vol. 64, no. 3, p. 1845–1866, 2018.

[13] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Mathematics of Control, Signals, and Systems,* vol. 2, no. 4, p. 303–314, 1989.

[14] K. Hornik, M. Stinchcombe and H. White, "Multilayer feedforward networks are universal approximators," *Neural Networks,* vol. 2, no. 5, p. 359–366, 1989.

[15] K.-I. Funahashi, "On the approximate realization of continuous mappings by neural networks," *Neural Networks,* vol. 2, no. 3, p. 183–192, 1989.

[16] V. Zhukova, P. Corte-León, J. M. Blanco, A. Allue, M. Ipatov, A. Zhukov and K. Gondra, "Applications of Co-rich Amorphous Glass-coated Microwires for Monitoring the Matrix Polymerization," in *2024 IEEE Applied Sensing Conference (APSCON)*, 2024.

[17] E. M. Dogo, O. . J. Afolabi, N. I. Nwulu, B. Twala and C. O. Aigbavboa, "A Comparative Analysis of Gradient Descent-Based Optimization Algorithms on Convolutional Neural Networks," in *2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS)*, 2018.

[18] R. Mahima, M. Maheswari, S. Roshana, E. Priyanka, N. Mohanan and N. Nandhini, "A Comparative Analysis of the Most Commonly Used Activation Functions in Deep Neural Network," in *4th International Conference on Electronics and Sustainable Communication Systems (ICESC)*, 2023.