REAL-TIME AI-DRIVEN INTERPRETATION OF ULTRASONIC DATA FROM RESISTANCE SPOT WELD PROCESS MONITORING FOR ADAPTIVE WELDING

RYAN SCOTT*†‡, DANILO STOCCO*†, ANDRIY CHERTOV*†, AND ROMAN GR. MAEV*†

ABSTRACT

Adaptive resistance spot welding systems typically rely on real-time analysis of dynamic resistance curves and other indirect measurements to estimate weld progress and guide adaptive weld control algorithms. Though efficient, these approaches are not always reliable, and consequently there is a need for improved feedback systems to drive adaptive welding algorithms. As an alternative, an advanced in-line integrated ultrasonic monitoring system is proposed, with real-time weld process characterization driven by artificial intelligence (AI) to create actionable feedback for the weld controller. Such a system would require real-time ultrasonic data interpretation, and for this a solution using deep learning was investigated. The proposed solution monitors the ultrasonic data for key process events and estimates the vertical size of the weld nugget proportional to the stack size throughout the welding process. This study shows that adaptive welding using ultrasonic process monitoring backed by AI-based data interpretation has immense potential. This research highlights the importance of nondestructive evaluation (NDE) in the zero-defect manufacturing paradigm.

KEYWORDS: resistance spot welding, ultrasound, artificial intelligence, deep learning, NDE 4.0

Introduction

Zero-defect manufacturing (ZDM) has been a dream for decades (Psarommatis et al. 2022, 2023). With respect to many manufacturing processes, this dream is considered within reach given the novel technologies that should be ubiquitous in an idealized Industry 4.0. Unfortunately, however, Industry 4.0 is not yet fully realized and thus the realization of ZDM suffers as well (Psarommatis et al. 2022). Though some requirements of Industry 4.0 are increasingly fulfilled (e.g., big data production, storage, and analytics; increased connectivity and Internet of Things; industrial automation), its full realization requires NDE 4.0 (Meyendorf et al. 2017). For example, NDE 4.0 is a prerequisite for Industry 4.0's widely unfulfilled key requirement of decentralized and autonomous decision-making (Escobar et al. 2021). Fulfillment of these requirements with respect to manufactured products and joining processes is promised by NDE 4.0 through (a) the automation of nondestructive inspections; (b) the automated, consistent, generalized, and accurate interpretation of inspection data; and (c) the resultant characterization of manufactured products, which would be used to inform downstream decision-making without human intervention.

Resistance spot welding (RSW) is a manufacturing process for which the ZDM dream is potentially within reach. Many industries heavily rely on RSW joints including automotive, aerospace, rail, and military. RSW is a favorable joining method in many cases because it is inexpensive to perform, has a fast cycle time, maintains integrity of the joined sheets, has minimal added weight and volume, is highly adaptable, is robust, and is generally amenable to nondestructive evaluation (NDE) (El-Banna 2006). However, across all industries, novel materials are increasingly being developed and incorporated into manufactured products (Perez-Regalado et al. 2013). For example, in the automotive industry—which uses RSW approximately 5000 to 7000 times per vehicle—increasing vehicle electrification imposes new engineering challenges with respect to safety, lightweighting, and weight distribution (Dugmore 2021). Consequently, there is an increasing use of novel lightweight and high-strength materials (e.g., advanced high-strength steels and aluminum alloys), as well as dissimilar-material joints, which pose new challenges for RSW (Dugmore 2021). Thus, there is an increasing demand for solutions that enable ZDM of RSW.

* The Institute for Diagnostic Imaging Research, University of Windsor, Canada
† Tessonics Inc., Windsor, Canada
‡ Corresponding author: rscott@uwindsor.ca

Materials Evaluation 81 (7): 61–70
https://doi.org/10.32548/2023.me-04344

There have been several attempts to support the realization of ZDM in RSW through the use of adaptive welding systems. Conceptually, modern adaptive welding systems monitor one or more indirect proxies of weld progress (e.g., dynamic resistance curves, current, voltage, force, tip displacement [El-Banna 2006; Neugebauer et al. 2013; Reis et al. 2016]), process these monitored features in real time to create feedback, and serve the feedback to an algorithm that adapts weld process parameters (e.g., weld time, force, and current) accordingly. In practice, these proxies do not produce sufficiently reliable and consistent feedback for adaptive weld controllers, so these systems generally fail to meet expectations and consequently many users revert to fixed schedules with adaptive capabilities disabled.

RSW is well-positioned to simultaneously meet the requirements of NDE 4.0 and achieve a breakthrough in ZDM, largely due to recent advancements in RSW NDE research. RSW NDE can be conducted either in-process (during the weld) or post-process (after the weld) using a variety of NDE modalities (Runnemalm and Appelgren 2012; Summerville et al. 2019). One of the most prevalent modalities is ultrasound (Chertov and Maev 2004; Denisov et al. 2004; Ouellette et al. 2013; Maev et al. 2014, 2016; Sung Hoon et al. 2020). Ultrasonic inspection has important advantages in inspection speed, insensitivity to sample thickness, adaptability, and the ability to directly inspect the internal geometric properties of the joint. The current state of the art in ultrasonic NDE for RSW consists of post-process offline inspection via portable ultrasonic systems with 2D matrix probes (e.g., Denisov et al. 2004; Maev et al. 2005), post-process robotized in-line systems with a similar ultrasonic configuration, and in-line real-time process monitoring systems using single-element probes (e.g., Chertov and Maev 2004; Ouellette et al. 2013; Maev et al. 2013, 2014; Sung Hoon et al. 2020). In any case, many NDE 4.0 requirements are already being met for such inspection systems, but only the in-line approach can provide real-time process monitoring and NDE data with 100% joint coverage, which is actionable in the context of an adaptive welding system that facilitates ZDM.

In its current form, the in-line inspection approach involves embedding a single-element ultrasonic transducer into a welding electrode (Chertov and Maev 2004; Ouellette et al. 2013; Maev et al. 2014, 2016; Sung Hoon et al. 2020). The transducer is immersed in flowing water, which both cools the transducer and provides coupling. The copper electrode caps focus the ultrasonic waves into the heat-affected zone of the workpiece and provide coupling against the stackup due to the application of intense force during welding (Maev and Chertov 2010). Throughout the welding process, A-scans are sampled every millisecond in pulse-echo mode, aiming through the center of the weld region between the electrodes. An M-scan—a 2D ultrasonic signature of the weld process—is then formed by horizontally stacking A-scans, and currently only post-process interpretation of the ultrasonic signature is conducted for quality control (Maev et al. 2021). Therefore, toward adaptive welding, a major missing piece in existing in-line ultrasonic systems is real-time interpretation of the sequence of A-scan signals as they are collected.

Classically, ultrasonic NDE data interpretation may involve signal/image processing, statistical analyses, search algorithms, model fitting, and hand-coded rules for decision-making. In some cases, these classical approaches are sufficient. However, in many application domains, such as RSW inspection—due to the many potential geometries, material combinations, and weld parameterizations, which can be encountered in a production environment—these approaches fail to meet the required performance, inference speed, and generality. Recently, deep learning approaches have been increasingly applied, to great effect, to a variety of problems in ultrasonic NDE data interpretation spanning essentially all use cases and specific tasks (e.g., defect detection and characterization, measurement automation, and so on [Cantero-Chinchilla et al. 2022; Taheri et al. 2022]). For example, Virkkunen et al. (2021) used a convolutional neural network (CNN) for crack detection in ultrasonic inspection data from butt-fused stainless steel pipes. Similarly, Shafiei Alavijeh et al. (2020) developed an ultrasonic inspection approach using a chord transducer for butt-fused plastic pipe joints. In this case, they used an autoencoder to conduct outlier detection on A-scans. Subsequently, the group developed an approach that classified A-scans in terms of defect presence/absence and according to defect type when a defect is detected (Shafiei Alavijeh et al. 2021). They compared several classical machine learning algorithms to four deep neural network architectures and determined that a CNN generally achieved the best performance on this task. Guo et al. (2019) combined CNN with recurrent neural networks (gated recurrent unit [GRU] and long short-term memory [LSTM]) to achieve high-performance debonding defect detection in ultrasonic C-scans of braided composite materials. They subsequently refined the approach in later works by instead framing the problem as semantic segmentation (Guo et al. 2023). Huang et al. (2022) also combined CNN and LSTM to detect defects in copper pipes in data from laser ultrasonic scanning. Maev et al. (2021) used an object detection approach with YOLOv3 (the "you-only-look-once" v3 object detector) to conduct post-process characterization of ultrasonic weld process signatures by identifying expulsions (discharge of molten material from the stackup due to intense pressure and rapid heating), while also identifying discrete weld-process events and measuring the position of the nugget at its maximum vertical size within the welded stackup. A more recent study by Zamiela et al. (2023) combined infrared with ultrasonic imaging data and developed a two-branch U-Net, which conducts semantic segmentation on the aligned images simultaneously to identify and characterize pores in metal structures in a single, unified output map. Deep learning has been proven to outperform classical computational NDE data interpretation approaches in terms of performance, inference speed, and generality; thus, it is a promising potential solution for time-sensitive contexts such as real-time inspection and adaptive welding.

The purpose of this study was first to investigate the feasibility of a deep learning approach for real-time interpretation of ultrasonic NDE data from RSW process monitoring. Subsequently, a novel approach was developed and evaluated for real-time characterization of ultrasonic data from the RSW process, which adheres to the needs of an adaptive weld controller that requires either continuous or discrete feedback.

Methodology

A summary of the experimental work is provided here, and a detailed description of the methodology follows. First, a large dataset of weld samples was developed with ultrasonic process monitoring enabled. Subsequently, a machine learning task was devised, the outputs of which could be used as actionable feedback for welding. The ultrasonic M-scan data were labeled to produce a machine learning dataset accordingly. A feasibility study was conducted to identify neural network architectures to perform the task within time constraints, and finally a feasible neural network was trained and evaluated.

Ultrasonic Configuration

Ultrasonic data were collected using an in-line real-time ultrasonic process monitoring system for spot welding (Figure 1). In this system, a 12 MHz single-element transducer is embedded into a welding electrode and immersed in water to keep the transducer cool; the water also acts as an ultrasound couplant. The transducer aims through the center of the weld region, between the electrodes. Ultrasonic A-scans are obtained every 1 ms in pulse-echo mode with a sampling frequency of 125 MHz. The number of samples in a given A-scan was set such that the entire stackup was visible in the A-scan throughout welding. Resultant M-scans were composed by horizontally stacking the A-scans.

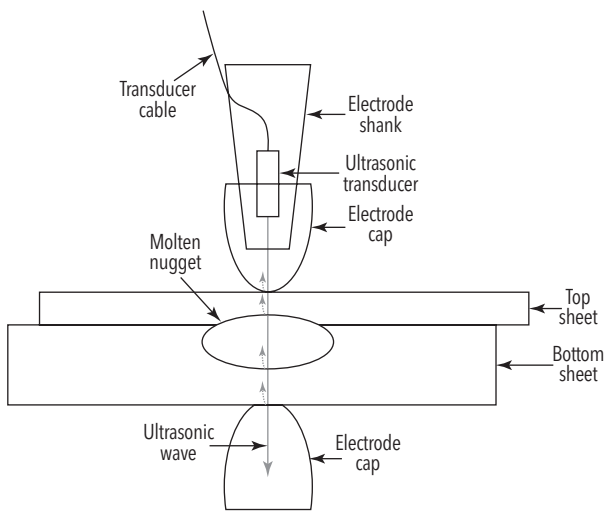

Figure 1. Ultrasonic configuration schematic for in-line real-time ultrasonic process monitoring system for spot welding. Ultrasonic waves (gray) are transmitted into the welded stack every 1 ms throughout welding process, and the transducer receives reflected waves as A-scan signals. The graphic shows an instant of an asymmetrical two-sheet weld already in progress (i.e., the molten nugget has been formed).

Dataset Development

A weld dataset of 18 223 RSWs, with the ultrasonic process monitoring system enabled, was developed. The welds were designed to cover a wide variety of weld geometries, materials, and parameters that are observed in automotive assembly. The dataset covered more than 80 sheet thickness combinations including 2- and 3-sheet similar- and dissimilar-material welds of 0.65–2.0 mm sheets made of mild and high-strength steels. Weld times varied from 75 to 400 ms, force ranged from 300 to 1000 kN, and current from 6 to 13 kA. Resultant welds were fabricated with diameters ranging from 0 to ~10 mm, with vertical maximum nugget size proportional to stack of ~0.0–0.8, and in some cases expulsions were purposely induced. Resultant M-scans were generally 75–400 pixels wide (based on weld time) and 100–400 pixels high (based on A-scan length, which varies by stack size).

For each weld, alongside its M-scan (Figure 2a), various metadata were captured, such as current-on timing, current-off timing, sheet thicknesses, and so forth. Metadata were used in the data preprocessing and augmentation stages to compute

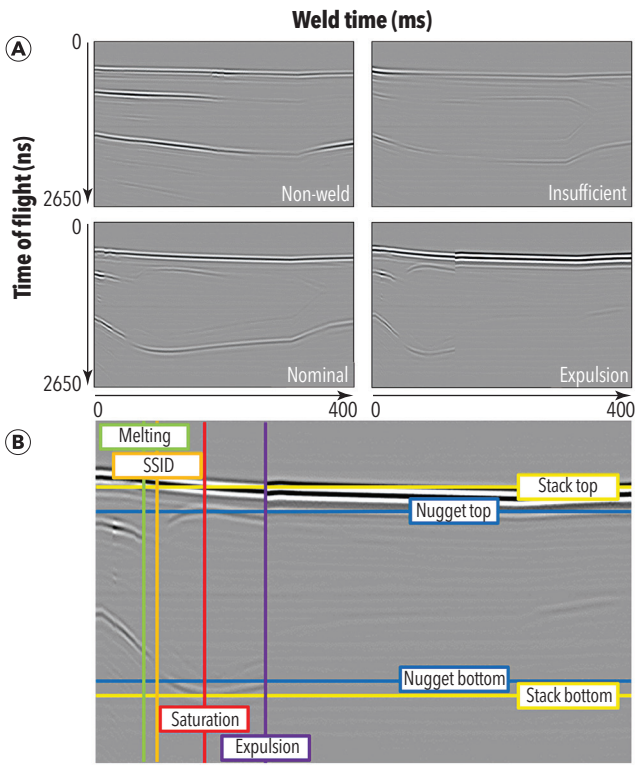

Figure 2. Ultrasonic M-scan samples from welds of same stackup but varying quality: (a) ultrasonic time of flight (Y axis) given the weld time (X axis). The top-left weld fails to breach the steel-steel interface, top-right weld insufficiently penetrates top sheet, bottom-left weld is ideal, and bottom-right weld features an expulsion (abrupt discontinuity in weld process). (b) Ultrasonic M-scan sample labeled for deep learning model development.

the regions of interest in ultrasonic scans and facilitate image cropping (see the Model Training and Performance Evaluation Section). The M-scans were labeled (Figure 2b) for the timing of four types of key events (see the cumulative distribution function plots in Figure 3): melting (the moment at which the molten nugget was first visible; 17 102 events); steel-steel interface disappearance (SSID; the moment at which all steel-steel interfaces appeared to have been breached by the molten nugget; 16 907 events), saturation (the moment at which the molten nugget appeared to stop growing vertically; 14 500 events); and expulsions (all first moments of discontinuity in the M-scan, which were suspected to be due to expulsion; 6375 events). In addition, the top and bottom of the nugget as well as the top and bottom of the stack were labeled, relative to the ultrasonic M-scan, at the moment of saturation. As observed in the M-scan, historical vertical maximum nugget size (MNS) proportional to the stack throughout the weld was then derived from these labels. To derive MNS, a linear interpolation l between melting-event timestamp to saturation-event timestamp (horizontally) and zero to maximum overall nugget size proportional to the stack (vertically) was first computed. Given melting timestamp m, saturation timestamp s, maximum overall nugget size proportional to the stack n, and weld timestep t:

$$l = 0, \text{ if } t < m$$

$$l = \frac{t-m}{s-m} \times n, \text{ if } m \leq t \leq s$$

$$l = n, \text{ otherwise}$$

Then, a sigmoidal function of the following form was fitted to l using the SciPy software package (Virtanen et al. 2020):

$$y = \frac{n}{1 + e^{-a \times (t-b)}}$$

where

a is a free parameter that controls the nugget growth rate, and

b is a bias that shifts the nugget growth in time.

Finally, a blend between l and y was computed such that the curve began fully linear at the melting point (i.e., the weight of $l = 1$, weight of $y = 0$) and ends fully sigmoidal at the saturation point (i.e., the weight of $l = 0$, weight of $y = 1$). MNS was the resultant blended curve.

Consequently, at each time step of the welding process, the model was tasked with binary classification for the first occurrence of each of the key events: melting, SSID, saturation, and expulsion. That is, for each event, the model was tasked to output zero for every time step prior to the first occurrence of the event and one for every time step thereafter. The model was also tasked with regression of MNS.

Model Design, Training, and Performance Evaluation

The machine learning task defined previously is essentially many-to-many sequence processing. Many-to-many sequence processing produces any number of sequential outputs given any number of sequential inputs; here, for every A-scan input the model is tasked with producing a corresponding output that describes the occurrence of events and MNS. All outputs are real numbers in the range of zero to one. A 1 ms per A-scan processing time constraint was imposed due to the required temporal resolution and response time such that feedback to a weld controller is actionable, as well as the rate of data acquisition such that the AI system does not accumulate latency throughout the course of a weld. Due to the severe computational time constraint of <1 ms per A-scan, the aim to maximize performance, and the sequential nature of the ultrasonic data, a recurrent neural network approach was investigated. In particular, to exploit the spatial information in each A-scan and

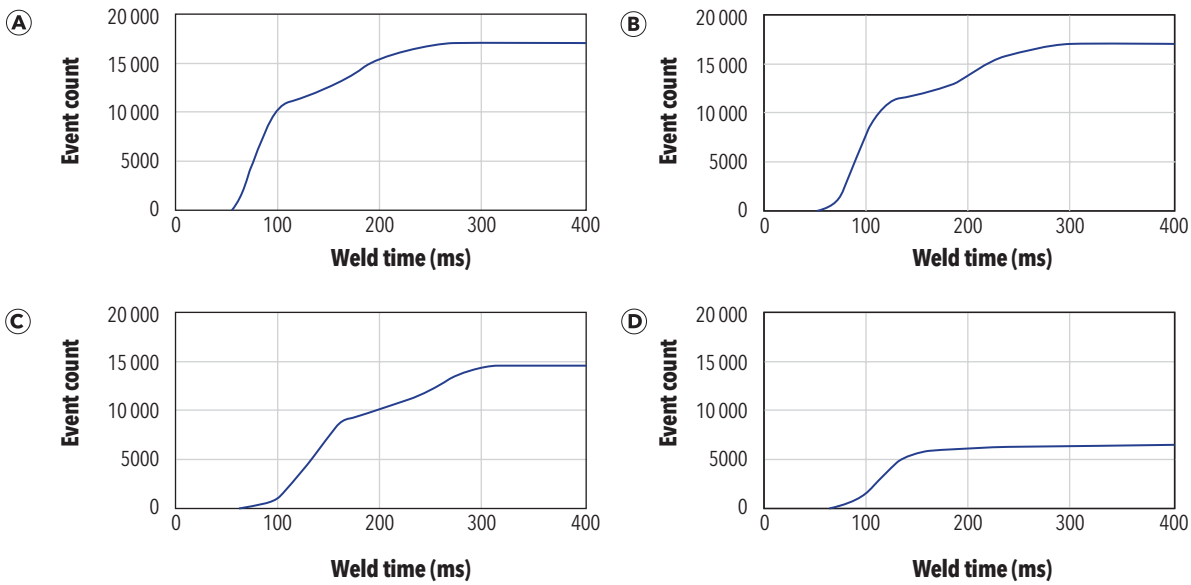

Figure 3. Cumulative distribution functions for the four events over the entire dataset: (a) melting; (b) SSID; (c) saturation; (d) expulsion.

long-term dependencies in weld sequences, convolutional long short-term memory (ConvLSTM) architectures (Shi et al. 2015) were explored. All investigations were conducted with Tensorflow (Abadi et al. 2015) and Keras (Chollet et al. 2015) for Python. Due to the sequential processing of ConvLSTMs and relatively small sequence items (resized A-scans; see the Model Training and Performance Evaluation Section), the amenability of processing to parallelization is heavily reduced for this task, and it was consequently found in preliminary work that CPUs were faster for both training and inference. Sequence processing using ConvLSTM differs from, for example, a pure CNN or transformer architecture, which is highly parallelizable and benefits greatly from computing on GPU. Thus, all computations were performed using an Intel® Core™ i7 CPU.

FEASIBILITY STUDY

The feasibility of a ConvLSTM-based architecture was investigated with input M-scans (i.e., arbitrary-length sequences of A-scans) resized vertically to 128 pixels. This investigation was designed to estimate the upper limit on the number of filters per layer and the number of layers based on the processing time requirement of <1 ms per A-scan in a production environment, which includes input preprocessing and potential communications overhead. Preliminary tests determined that the production environment ran inference approximately 35–45% faster than the development environment due to, for example, the removal of training overhead in the exported network graph, differences in Tensorflow compilation, differences in programming language, and so on. Accounting for the speedup in the production environment, along with overhead from preprocessing and so forth, a cutoff of 1.1 ms per A-scan was imposed. An overarching architecture (Figure 4) was designed with one ConvLSTM module and one max pooling operation per layer; variants were tested having 1–5 layers and an initial layer with 8–32 filters, with number of filters doubling per layer. Following a flattening, the last layer was a time-distributed (i.e., shared across all time steps) fully connected layer with five outputs. To guard against extra computational overhead from initial resource allocation, one M-scan was fully processed prior to recording inference times. Subsequently, 10 randomly selected, arbitrary-length M-scans, comprising of a total of 2589 A-scans, were processed, during which inference times were recorded. As M-scan length has no impact on mean inference time per A-scan, though it may subtly impact variance of inference times, the selected M-scans were held constant through all trials so that the same exact A-scans were processed in each trial. The largest feasible model was used for further training and evaluation.

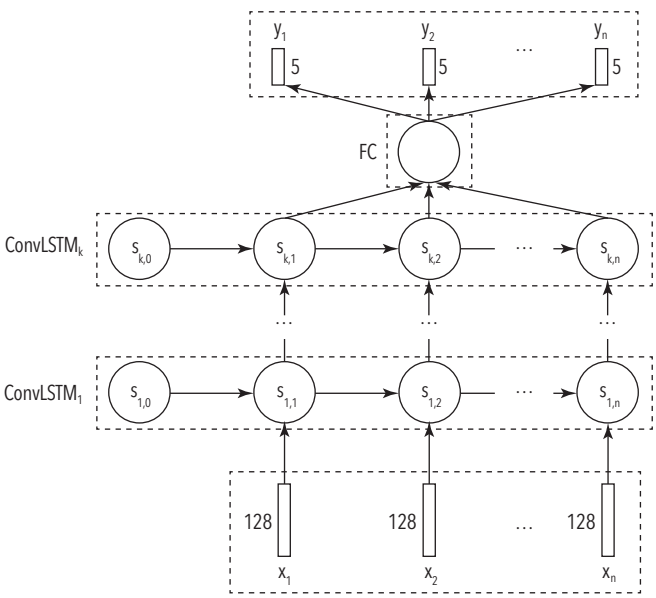

Figure 4. An unrolled schematic diagram of the ConvLSTM architecture used in this study. Data flow over depth of network is from bottom to top. Data flow over weld time is from left to right. Input A-scans (x_t = 1…n) are fed to the network. ConvLSTM layer (1…k) states (denoted s; composed of C and H states observed in standard LSTMs) are initially zeros and modified over time given previous states and new inputs from previous layer. A max pooling operation follows each ConvLSTM. Outputs of the last ConvLSTM layer are fed into the time-distributed (i.e., shared across all time steps) fully connected decision-making layer (denoted FC in the figure). Input and output dimensionalities are depicted.

MODEL TRAINING AND PERFORMANCE EVALUATION

At both training and testing time, M-scan images were cropped vertically to tightly focus on the welded stackup, resized to a height of 128 pixels, and cropped horizontally starting at the current-on timing until the end of the weld process. Data augmentation was conducted only at training time and was designed to desensitize the model to various potential situations that could occur in a production system (e.g., electromagnetic interference, slight misreporting of current-on timing, gain and contrast variance, shift in A-scan gating, etc.). Thus, augmentation involved some typical image augmentation steps such as random vertical shifts of both top and bottom image cropping positions prior to resizing vertically to 128 pixels, random horizontal shift of current-on (image left edge) position, addition of artificial noise, and random contrast adjustments. In addition, random horizontal resizing of M-scans to uniformly distributed randomly-selected widths from 75–400 pixels was conducted to desensitize the model to the weld timing distribution of the training data, with the aim of producing a more robust model such that it can correctly interpret data from welds having weld times vastly different from those typically observed in the training data. Key event timings and MNS curves were adjusted according to any augmentations performed. Due to the random horizontal resizing, inputs and targets were zero-padded after the end of the sequence.

Three models were trained using Monte-Carlo validation and evaluated on a held-out testing dataset. Of the 18 223 labeled M-scan samples, 16 400 were used for training, 1640 for validation, and 1823 for testing. Each model was trained using the Adam optimizer (Kingma and Ba 2015) for 400 epochs with

TABLE 1

Summary of feasibility study results

Architecture (filters per ConvLSTM layer)	Parameters	Inference time (ms)
8	3461	0.866 (0.193)
8-16	8133	1.024 (0.155)
8-16-32	26 693	1.060 (0.129)
8-16-32-64	100 677	1.221 (0.142)
8-16-32-64-128	396 101	1.428 (0.155)
16	8453	0.952 (0.304)
16-32	27 013	0.954 (0.160)
16-32-64	100 997	1.085 (0.151)
16-32-64-128	396 421	1.344 (0.132)
16-32-64-128-256	1 577 093	1.992 (0.154)
32	23 045	0.999 (0.268)
32-64	97 029	0.970 (0.108)
32-64-128	392 453	1.264 (0.126)
32-64-128-256	1 573 125	1.942 (0.144)
32-64-128-256-512	6 293 765	4.271 (0.189)

Note: Mean inference time per A-scan over 2589 A-scans, standard deviation in parentheses

a batch size of 32, with early stopping and learning rate reduction based on validation loss. Binary cross-entropy loss was used for the event outputs, while mean-squared error was used on the MNS regression output. Inputs and loss were masked such that models would skip the zero-vector A-scans during training and not backpropagate loss on those inputs; thus, the model did not learn from the zero-padded regions.

To evaluate performance, a number of different performance indicators were used for each task. With respect to event detection, performance was evaluated using sensitivity with respect to the absolute error of ground truth versus model predictions of event timings from 0–30 ms, overall specificity, and histograms of timing error for true positives. MNS regression performance was assessed using the percentage of A-scans that are correct within an absolute difference of 0.1.

Results

The results of the feasibility study and performance evaluation are discussed next.

Feasibility Study

The feasibility study results (Table 1) demonstrated that architectures starting with eight filters in the first layer were feasible up to three layers (Figure 5a). The three-layer architecture with eight filters in the first layer had an inference time of 1.06 ms (SD = 0.13 ms), whereas a four-layer architecture had an inference time of 1.22 ms (SD = 0.14 ms). With 16 filters (Figure 5b),

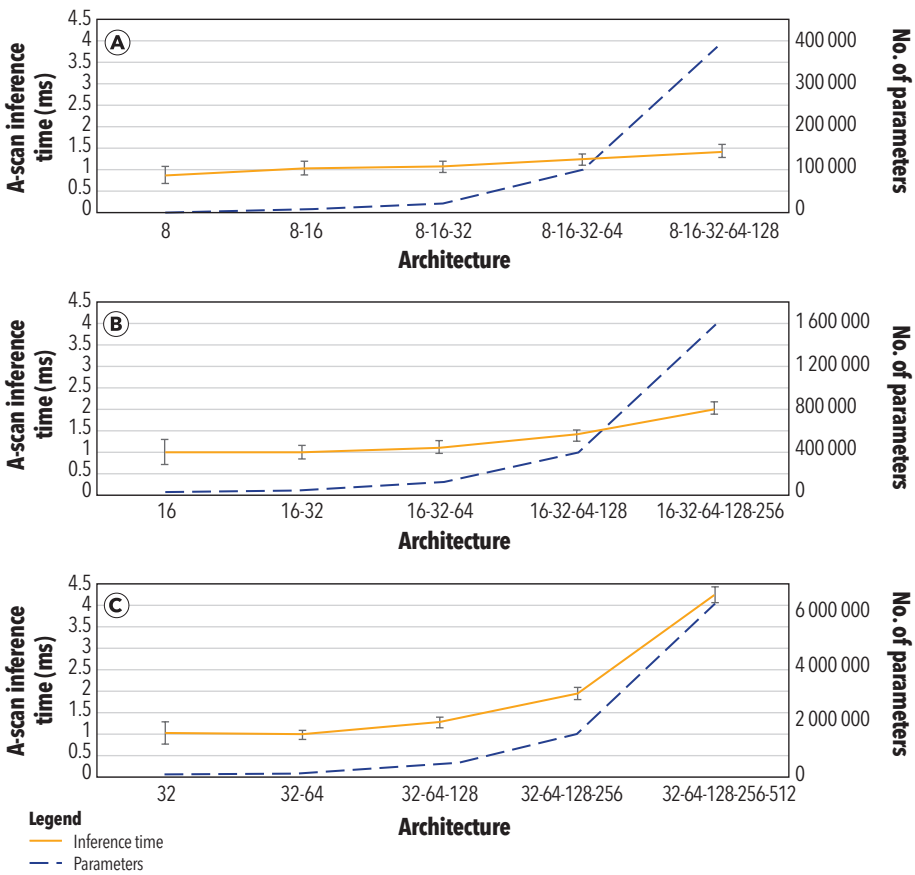

Figure 5. Inference time per A-scan in milliseconds (orange line) and parameter count (blue dashed line) per architecture (X axis; 1–5 layers), with first layer number of filters: (a) 8; (b) 16; and (c) 32. Inference time generally grows superlinearly with respect to both number of layers and parameters. Means across all 2589 A-scans are plotted with 1 standard deviation of mean shown with error bars.

TABLE 2
Summary of performance results

	Melting	SSID	Saturation	Expulsion
Sensitivity (within 10 ms)	0.870 (0.010)	0.886 (0.009)	0.742 (0.023)	0.935 (0.005)
Sensitivity (within 20 ms)	0.963 (0.006)	0.962 (0.007)	0.860 (0.014)	0.954 (0.006)
Sensitivity (within 30 ms)	0.984 (0.003)	0.979 (0.002)	0.898 (0.009)	0.977 (0.001)
Specificity	0.807 (0.025)	0.821 (0.031)	0.933 (0.018)	0.986 (0.002)
% A-scans correct (within 0.1)	90.5 (0.4)			

Note: Mean sensitivity for each event within 10, 20, and 30 ms of event ground truth, mean specificity for each event, and mean accuracy of MNS within 0.1. Standard deviation in parentheses.

a three-layer architecture was feasible with an inference time of 1.08 ms (SD = 0.15 ms), while a four-layer architecture was infeasible at 1.34 ms per A-scan (SD = 0.13 ms). Finally, with 32 filters (Figure 5c), architectures with up to two layers were feasible with 0.97 ms per A-scan (SD = 0.27 ms), while a three-layer architecture had an inference time of 1.26 ms (SD = 0.13 ms). In terms of parameter count, the largest feasible architecture was three layers with 16 filters in the first layer, yielding 100 997 parameters. Thus, this architecture was used for all subsequent experimentation.

Performance Evaluation

The architecture selected in the feasibility study demonstrated a very strong performance overall (Table 2). Within 10 ms of the event, the sensitivity of the four events ranged from 0.742 (SD = 0.023) to 0.935 (SD = 0.005). Within 30 ms, sensitivity reached 0.977 for all events but saturation, which peaked at 0.898 (SD = 0.009). Overall, specificity for expulsion events was highest at 0.986 (SD = 0.002), while the melting and SSID events had proportionally more false positives with specificity of 0.807 (SD = 0.025) and 0.821 (SD = 0.031), respectively. Event detectability curves were similar for melting and SSID, while curves for expulsion and saturation differed greatly and models were consistent with respect to each event across all timing error windows (Figure 6). Expulsion detection reached an asymptote at approximately the 5 ms error window and melting and SSID reached an asymptote at approximately 15 ms, while saturation reached an asymptote at approximately 30 ms of absolute error.

Example distributions of timing error for melting (Figure 7a) and SSID (Figure 7b) events showed very symmetrical distributions, centered at approximately zero with relatively mild variance. Saturation (Figure 7c), on the other hand, yielded a timing error distribution with slight negative skew and greater variance, and was centered slightly above zero. Expulsion timing error (Figure 7d) yielded an extremely tight distribution centered just above zero. All models yielded similar error distributions.

Model outputs plotted over time and compared against ground truth data (Figure 8) showed stability and smoothness on relatively clear M-scans, while output noise increased with decreasing M-scan quality. Overall, the models were insensitive

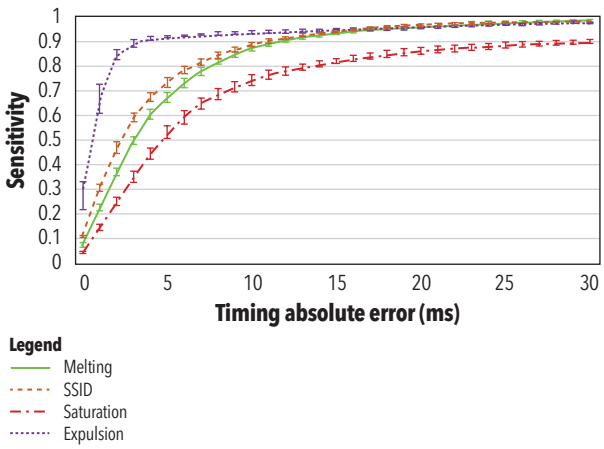

Figure 6. Event detection sensitivity per event given absolute error of model prediction of event timing versus ground truth event timestamp. Means across three models are plotted with 1 standard deviation of mean shown with error bars. Expulsion (purple line) was most easily detectable, SSID (orange line) and melting (green line) were similarly moderately detectable, and saturation (red line) was most difficult to detect.

to reasonable amounts of electromagnetic noise, weld time, stackup, and weld quality. Output curves for MNS were smooth and consistent with ground truth curves in terms of shape and position. In addition, welds without nugget formation (e.g., Figure 8a) were often correctly characterized. In general, welds with extremely late nugget formation (e.g., Figure 8b) were more difficult to characterize than those with earlier nugget formation (Figure 8c).

Discussion

A fast and performant approach was developed for real-time interpretation of data from ultrasonic RSW process monitoring, with the aim of creating actionable feedback to a weld controller using deep learning.

All events were reliably detected; over 95% of events were detected within 18 ms of ground truth for all events except for saturation, which was detected at a rate of 90% within 30 ms. It was expected that expulsions would be most reliably detected as they appear very clearly on M-scans as a discontinuity in which the stack bottom boundary abruptly moves upward

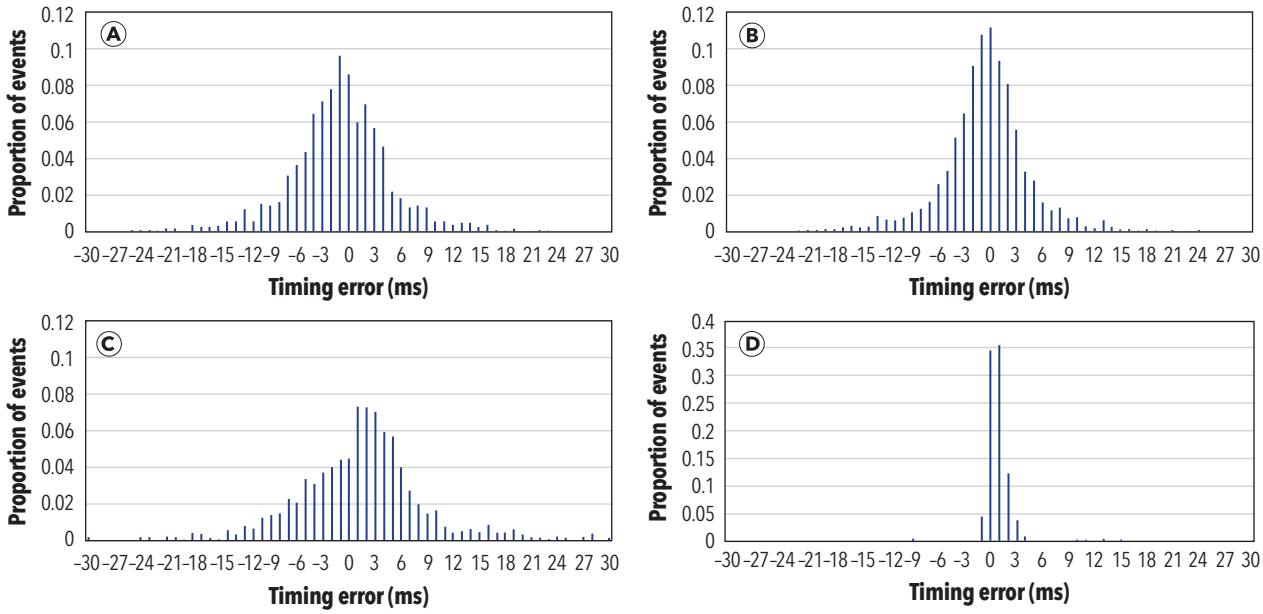

Figure 7. Timing error distributions for detected (a) melting; (b) SSID; (c) saturation; and (d) expulsion for the model with the best overall performance.

in the M-scan (Figure 8d). It was not expected, however, that they would be so rapidly detected with >90% of detections occurring within 4 ms of ground truth. Because they appear so clearly in M-scans, they are also the easiest event to label during dataset preparation. On the other hand, saturation is by far the most difficult event to label as the saturation point was defined as "the moment at which the molten nugget appeared to stop growing vertically," which is highly subjective without perfect nugget and stack boundary annotations. Similarly, but less so, melting and SSID are not always as apparent as expulsions. Thus, from our experience in reading these images, and considering the relative difficulty for a human to interpret ultrasonic M-scans and identify these events and the relative consistency of event annotations, we found that the relative detection rates of the four events completely align with expectations.

Relatedly, as the ground truth labels for event timing as well as the top and bottom labels for the nugget and stack were used to develop the curves for MNS, the subjectivity and consistency of labels affects the performance of the models on the regression task as well. In particular, stack boundaries are almost always reasonably visible aside from after expulsions, while nugget boundaries vary in visibility based on nugget pool size, stage of weld, and stack geometry. With the investigated

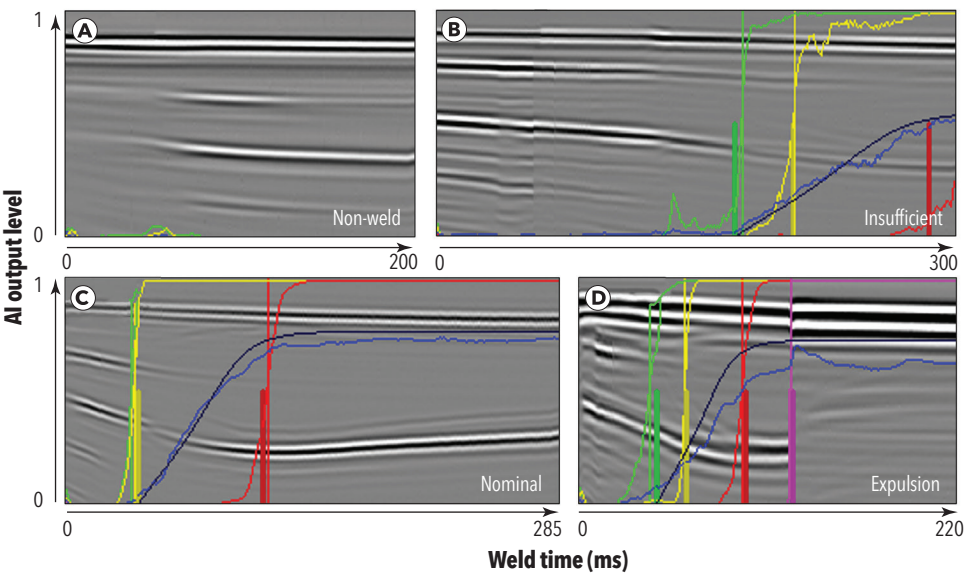

Figure 8. M-scan samples with ground truth markup and model outputs superimposed. Ground truth event timestamps are shown as dark thick vertical lines reaching the halfway mark vertically in images; ground truth MNS is darker blue curve. Model outputs (thin curves – unprocessed model outputs; thin vertical lines – event probability outputs thresholded at 0.5) are from most performant model. For event colors, green – melting; yellow – SSID; red – saturation; purple – expulsion. Blue indicates model output for MNS. All model outputs and MNS targets are superimposed on images with 0 = bottom of image, 1 = top of image. Images cover various stackups and weld outcomes.

approach, continuous feedback could be provided to an adaptive weld controller using the regression output of MNS, or other discrete events could be derived from it. In either case, optimal use of MNS output in a production implementation would likely require calibration welds for novel weld type (given, for example, sheet thicknesses, sheet materials, etc.) to determine an appropriate threshold or target value for MNS output. In future work, this ultrasound-based approach with AI-driven feedback will be rigorously compared against alternative feedback approaches (e.g., resistance-based feedback, etc.).

With respect to this AI-based approach, a more rigorous optimization of hyperparameters within the feasible space of architectures may yield better performance. In this study, the largest model possible (in terms of parameter count) was used based on feasibility study results; however, it is possible that better performance may be achieved by using various popular modules in the network (if feasible), for example, skip connections (He et al. 2016; Ronneberger et al. 2015), atrous spatial pyramid pooling (Chen et al. 2018), batch normalization (Ioffe and Szegedy 2015), attention mechanisms such as convolution block attention module (Woo et al. 2018), and so forth. Alternatively, other novel architectures, such as vision transformer (Liu et al. 2021), could be explored in future work. In addition, providing known welding parameters to the model as inputs (such as sheet thicknesses, sheet material encodings, force, welding cap face diameter, etc.) is another potential opportunity for improvement, which can be investigated in the future.

Precise and continuous annotations, for both the nugget and stack, at all times throughout the weld would be ideal in order to derive event timestamps and MNS curves. From the standpoint of dataset development, this would essentially be the same as labeling the M-scans for semantic segmentation of the nugget and stack boundaries, which is significantly more tedious and laborious than the proposed approach, and still subjective (though, perhaps less subjective as it is less abstract). One advantage of the proposed approach is that it required, at most, eight clicks per annotated M-scan (each of the four event timings, two nugget labels, two stack labels) during data annotation, whereas semantic segmentation would conservatively require 20 clicks per segmented region to delineate each polygon—40 clicks in total between nugget and stack regions—so the proposed approach yielded a five-fold reduction in data preparation time. That said, semantic segmentation of the ultrasonic data is still a natural next step for this work like in the case of Guo et al. (2023). Other works have demonstrated the potential for semantic segmentation in real-time ultrasonic inspection in both NDE and medical contexts (Fiorito et al. 2018; Hu et al. 2022; Shandiz and Tóth 2022). If it were found to be performant, generalizable, and still sufficiently fast for adaptive RSW (i.e., <1 ms per A-scan inference time in a production environment), semantic segmentation could yield more precise and continuous measurements, and consequently better feedback. This would be especially valuable if continuous feedback to an adaptive weld controller was preferred over discrete feedback, or perhaps necessitated for a particular adaptive welding algorithm.

Conclusion

The investigated approach is not limited to ultrasonic NDE nor resistance spot welding; such an approach could be applied to the interpretation of NDE data from a variety of other modalities for a variety of other joining methodologies. In all, the investigated approach is an exciting first step toward real-time interpretation of ultrasonic NDE data from RSW. It demonstrates the enormous potential of ultrasound-based process monitoring backed by real-time interpretation using deep learning, for real-time adaptive feedback systems in modern manufacturing. Such NDE 4.0 systems are integral to Industry 4.0 and the ZDM paradigm, and this work brings zero-defect RSW closer to reality.

ACKNOWLEDGMENTS

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Collaborative Research and Development (CRD) grant CRDPJ 508935-17. It was also supported by the National Research Council Canada (NRC) Industrial Research Assistance Program (IRAP). The authors would also like to thank the Institute for Diagnostic Imaging Research at the University of Windsor, Canada, as well as Tessonics Inc. in Windsor, Canada, for providing experimental facilities, equipment, and materials for this research.

REFERENCES

Abadi, M., A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, et al. 2015. "TensorFlow: Large-scale machine learning on heterogeneous systems." https://www.tensorflow.org.

Shafiei Alavijeh, M., R. Scott, F. Seviaryn, and R. Gr. Maev. 2020. "NDE 4.0 compatible ultrasound inspection of butt-fused joints of medium-density polyethylene gas pipes, using chord-type transducers supported by customized deep learning models." *Research in Nondestructive Evaluation* 31(5–6). https://doi.org/10.1080/09349847.2020.1841864.

Shafiei Alavijeh, M., R. Scott, F. Seviaryn, and R. Gr. Maev. 2021. "Using machine learning to automate ultrasound-based classification of butt-fused joints in medium-density polyethylene gas pipes." *Journal of the Acoustical Society of America* 150 (1): 561–72. https://doi.org/10.1121/10.0005656.

Cantero-Chinchilla, S., P. D. Wilcox, and A. J. Croxford. 2022. "Deep learning in automated ultrasonic NDE – Developments, axioms and opportunities." *NDT & E International* 131. https://doi.org/10.1016/j.ndteint.2022.102703.

Chen, L.-C., Y. Zhu, G. Papandreou, F. Schroff, and H. Adam. 2018. "Encoder-decoder with atrous separable convolution for semantic image segmentation." In *Computer Vision – ECCV 2018: 15th European Conference*: 833–851. https://doi.org/10.1007/978-3-030-01234-2_49.

Chertov, A., and R. Gr. Maev. 2004. "Determination of resistance spot weld quality in real time using reflected acoustic waves. Comparison with through-transmission mode." 16th World Conference on NDT. Montreal, Canada.

Chollet, F., et al. 2015. Keras. Software available at keras.io.

Denisov, A. A., C. M. Shakarji, B. B. Lawford, R. G. Maev, and J. M. Paille. 2004. "Spot weld analysis with 2D ultrasonic Arrays." *Journal of Research of the National Institute of Standards and Technology* 109 (2): 233–44. https://doi.org/10.6028/jres.109.015.

Dugmore, A. 2021. "New composites target EV applications." *SAE International*. https://www.sae.org/news/2021/08/new-composites-target-ev-applications.

El-Banna, M. 2006. "Dynamic resistance based intelligent resistance welding." Doctoral dissertation. Wayne State University.

Escobar, C. A., M. E. McGovern, and R. Morales-Menendez. 2021. "Quality 4.0: A review of big data challenges in manufacturing." *Journal of Intelligent Manufacturing* 32 (8): 2319–34. https://doi.org/10.1007/s10845-021-01765-4.

Fiorito, A. M., A. Østvik, E. Smistad, S. Leclerc, O. Bernard, and L. Lovstakken. 2018. "Detection of cardiac events in echocardiography using 3D convolutional recurrent neural networks." *2018 IEEE International Ultrasonics Symposium (IUS)*. Kobe, Japan: 1–4. https://doi.org/10.1109/ULTSYM.2018.8580137.

Psarommatis, F., F. Fraile, J. P. Mendonca, O. Meyer, O. Lazaro, and D. Kiritsis. 2023. "Zero defect manufacturing in the era of industry 4.0 for achieving sustainable and resilient manufacturing." Frontiers in Manufacturing Technology 3. https://doi.org/10.3389/fmtec.2023.1124624.

Psarommatis, F., J. Sousa, J. P. Mendonça, and D. Kiritsis. 2022. "Zero-defect manufacturing the approach for higher manufacturing sustainability in the era of industry 4.0: A position paper." *International Journal of Production Research* 60 (1): 73–91. https://doi.org/10.1080/00207543.2021.1987551.

Guo, Y., Z. Xiao, L. Geng, J. Wu, F. Zhang, Y. Liu, and W. Wang. 2019. "Fully convolutional neural network with GRU for 3D braided composite material flaw Detection." *IEEE Access: Practical Innovations, Open Solutions* 7:151180–88. https://doi.org/10.1109/ACCESS.2019.2946447.

Guo, Y., Z. Xiao, and L. Geng. 2023. "Defect detection of 3D braided composites based on semantic segmentation." *Journal of the Textile Institute*. https://doi.org/10.1080/00405000.2022.2054103.

He, K., X. Zhang, S. Ren, and J. Sun. 2016. "Deep residual learning for image recognition." *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, NV: 770–778. https://doi.org/10.1109/CVPR.2016.90.

Hu, J., E. Smistad, I. M. Salte, H. Dalen, and L. Lovstakken. 2022. "Exploiting temporal information in echocardiography for improved image segmentation." *2022 IEEE International Ultrasonics Symposium (IUS)*. Venice, Italy: 1–4. https://doi.org/10.1109/IUS54386.2022.9958670.

Huang, L., X. Hong, Z. Yang, Y. Liu, and B. Zhang. 2022. "CNN-LSTM network-based damage detection approach for copper pipeline using laser ultrasonic scanning." *Ultrasonics* 121:106685. https://doi.org/10.1016/j.ultras.2022.106685.

Ioffe, S. and C. Szegedy. 2015. "Batch normalization: accelerating deep network training by reducing internal covariate shift." arXiv:1502.03167 [cs.LG]. https://doi.org/10.48550/ARXIV.1502.03167.

Kingma, D. P. and J. Ba. 2015. "Adam: A method for stochastic optimization." 3rd International Conference for Learning Representations, San Diego, CA.

Liu, Z., Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, and B. Guo. 2021. "Swin transformer: hierarchical vision transformer using shifted windows." 2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal, QC, Canada: 9992–10002. https://doi.org/10.1109/ICCV48922.2021.00986.

Maev, R., F. Ewasyshyn, S. Titov, J. Paille, E. Maeva, A. Denisov, and F. Seviaryn. 2005. Method and apparatus for assessing the quality of spot welds. US Patent 7,775,415 B2, filed 14 June 2005, and issued 17 August 2010.

Maev, R. Gr., A. M. Chertov, J. M. Paille, and F. J. Ewasyshyn. 2013. Ultrasonic in-process monitoring and feedback of resistance spot weld quality. US Patent 9,296,062 B2, filed 10 June 2013, and issued 29 March 2016.

Maev, R. Gr., A. M. Chertov, W. Perez-Regalado, A. Karloff, A. Tchipilko, P. Lichaa, D. Clement, and T. Phan. 2014. "In-line inspection of resistance spot welds for sheet metal assembly." *Welding Journal* 93: 58-62.

Maev, R. Gr., and A. M. Chertov. 2010. Electrode cap for ultrasonic testing, US Patent 8,381,591 B2, filed 18 March 2010, and issued 26 February 2013.

Maev, R. Gr., A. Chertov, R. Scott, D. Stocco, A. Ouellette, A. Denisov, A., and Y, Oberdoerfer. 2021. "NDE in the automotive sector." in *Handbook of Nondestructive Evaluation 4.0*. Springer Nature Switzerland AG.

Meyendorf, N., L. Bond, J. Curtis-Beard, S. Heilmann, S. Pal, R. Schallert, H. Scholz, and C. Wunderlich. 2017. "NDE 4.0—NDE for the 21st century—the internet of things and cyber physical systems will revolutionize NDE." 15th Asia Pacific Conference for Non-Destructive Testing (APCNDT 2017), Singapore.

Neugebauer, R., T. Wiener, and A. Zösch. 2013. "Quality control of resistance spot welding of high strength steels." *Procedia CIRP* 12:139–44. https://doi.org/10.1016/j.procir.2013.09.025.

Ouellette, A., A. C. Karloff, W. Perez-Regalado, A. M. Chertov, R. G. Maev, and P. Lichaa. 2013. "Real-time ultrasonic quality control monitoring in resistance spot welding: Today and tomorrow." *Materials Evaluation* 71 (7).

Perez-Regalado, W., A. Ouellette, A. M. Chertov, V. Leshchynsky, and R. G. Maev. 2013. "Joining dissimilar metals: A novel two-step process with ultrasonic monitoring." *Materials Evaluation* 71 (7): 828–33.

Reis, F.F., V. Furlanetto, and G. F. Batalha. 2016. "Resistance spot weld in vehicle structures using dynamic resistance adaptive control." SAE Technical Paper, 2016-36-0303. https://doi.org/10.4271/2016-36-0303.

Ronneberger, O., P. Fischer, and T. Brox. 2015. "U-Net: Convolutional networks for biomedical image segmentation." *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*: 234–241. https://doi.org/10.1007/978-3-319-24574-4_28.

Runnemalm, A., and A. Appelgren. 2012. "Evaluation of non-destructive testing methods for automatic quality checking of spot welds." Report. Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:se:hv:diva-5578.

Shandiz, A. H., and L. Tóth. 2022. "Improved processing of ultrasound tongue videos by combining convLSTM and 3D convolutional networks." *Advances and Trends in Artificial Intelligence. Theory and Practices in Artificial Intelligence:* 265–274. https://doi.org/10.1007/978-3-031-08530-7_22.

Shi, X., Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-C. Woo. 2015. "Convolutional LSTM Network: a machine learning approach for precipitation nowcasting." In *Proceedings of the 28th International Conference on Neural Information Processing Systems*: 802–810.

Summerville, C., P. Compston, and M. Doolan. 2019. "A comparison of resistance spot weld quality assessment techniques." *Procedia Manufacturing* 29:305–12. https://doi.org/10.1016/j.promfg.2019.02.142.

Sung Hoon, J., N. Yang Woo, Y. Sanghyun, K. Si Eun, R. Gr. Maev, A. M. Chertov, D. R. Scott, and D. Stocco. 2020. System and method for resistance spot welding control. Korea Patent 10-2166234-0000. Korean Intellectual Property Office. Filed 28 January 2020, and issued 25 August 2020.

Taheri, H., M. Gonzalez Bocanegra, and M. Taheri. 2022. "Artificial intelligence, machine learning and smart technologies for nondestructive evaluation." *Sensors (Basel)* 22 (11): 4055. https://doi.org/10.3390/s22114055.

Virkkunen, I., T. Koskinen, O. Jessen-Juhler, and J. Rinta-aho. 2021. "Augmented ultrasonic data for machine learning." *Journal of Nondestructive Evaluation* 40 (1): 4. https://doi.org/10.1007/s10921-020-00739-5.

Virtanen, V., R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, et al. 2020. "SciPy 1.0: Fundamental algorithms for scientific computing in python." *Nature Methods* 17: 261–72. https://doi.org/10.1038/s41592-019-0686-2.

Woo, S., J. Park, J-Y. Lee, and I.S. Kweon. 2018. "CBAM: Convolutional block attention module." *Proceedings of the European Conference on Computer Vision (ECCV):* 3–19. https://doi.org/10.48550/arXiv.1807.06521.

Zamiela, C., Z. Jiang, R. Stokes, Z. Tian, A. Netchaev, C. Dickerson, W. Tian, and L. Bian. 2023. "Deep multi-modal U-Net fusion methodology of thermal and ultrasonic images for porosity." *Journal of Manufacturing Science and Engineering* 145 (6). https://doi.org/10.1115/1.4056873.

ACOUSTIC EMISSION SOURCE LOCALIZATION USING DEEP TRANSFER LEARNING AND FINITE ELEMENT MODELING–BASED KNOWLEDGE TRANSFER

XUHUI HUANG*, OBAID ELSHAFIEY*, KARIM FARZIA†, LALITA UDPA*, MING HAN*, AND YIMING DENG*‡

ABSTRACT

This paper presents a novel data-driven approach to localize two types of acoustic emission sources in an aluminum plate, namely a Hsu-Nielsen source, which simulates a crack-like source, and steel ball impacts of varying diameters acting as the impact source. While deep neural networks have shown promise in previous studies, achieving high accuracy requires a large amount of training data, which may not always be feasible. To address this challenge, we investigated the applicability of transfer learning to address the issue of limited training data. Our approach involves transferring knowledge learned from numerical modeling to the experimental domain to localize nine different source locations. In the process, we evaluated six deep learning architectures using tenfold cross-validation and demonstrated the potential of transfer learning for efficient acoustic emission source localization, even with limited experimental data. This study contributes to the growing demand for running deep learning models with limited capacity and training time and highlights the promise of transfer learning methods such as fine-tuning pretrained models on large semi-related datasets.

KEYWORDS: acoustic emission, deep neural network, finite element modeling, transfer learning, fiber optics, source localization

Introduction

Acoustic emission source localization is crucial in structural health monitoring (SHM) and proactive maintenance of metallic structures. The constraints in deploying acoustic emission testing (AE) sensor arrays in real-world structures necessitate a shift toward intelligent, automated single-sensor approaches. Holford et al. (2001) pioneered the application of AE for damage location in steel bridges, establishing its importance in SHM. Ebrahimkhanlou and Salamone (2017) further examined acoustic source localization and its significance in determining the origin of acoustic emission waves and assessing damage severity. Cheng et al. (2021) developed an acoustic emission source localization method using Lamb wave propagation simulation and artificial neural networks, proving effective in I-shaped steel girder inspections. Ai et al. (2021) studied source localization on large-scale canisters used for nuclear fuel storage, addressing the need for optimal AE sensor deployment. Ciampa and Meo (2010) proposed an approach using wavelet analysis and a Newton-based optimization technique for acoustic emission source localization and velocity determination, contributing to the broader understanding of acoustic emission wave propagation and source detection.

Significant progress has been achieved in acoustic emission source localization through the application of deep learning, demonstrating its promise in localizing acoustic emission signals (LeCun et al. 2015). Ebrahimkhanlou and Salamone (2018) proposed a deep learning approach for localizing acoustic emission sources using a single sensor in plate-like structures. This was further advanced by Ebrahimkhanlou et al. (2019), who introduced a deep learning–based framework for localizing and characterizing acoustic emission sources in metallic panels using only one sensor. Garrett et al. (2022) utilized artificial intelligence for estimating fatigue crack length from acoustic emission waves, a significant step forward in damage localization and quantification. Despite the challenge of false positives, the fusion of artificial intelligence and AE holds promising opportunities for enhancing SHM (Verstrynge et al. 2021; Hassan et al. 2021).

A key challenge in using supervised learning algorithms for acoustic emission source localization is the difficulty in accessing labeled acoustic emission signals for existing structures. Transfer learning is a strategy that assists the supervised learning task when available training data is limited

* Department of Electrical and Computer Engineering, Michigan State University, East Lansing, MI
† Nikon Inc., 9453 Innovation Dr., Manassas, VA
‡ Corresponding author: dengyimi@egr.msu.edu

Materials Evaluation 81 (7): 71–84
https://doi.org/10.32548/2023.me-04348