From Experiments to In-Service: Learning Embeddings for Damage Identification and Interpretation in Structural Health Monitoring

Kyle L. Hom[1], Homayoon Beigi[2,3,4], and Raimondo Betti[1]

Recognition Technologies Technical Report: RTI-20250304-01
DOI: 10.13140/RG.2.2.35145.04969

[1]Department of Civil Engineering and Engineering Mechanics, Columbia University, New York, NY
[2]Department of Mechanical Engineering, Columbia University, New York, NY
[3]Department of Electrical Engineering, Columbia University, New York, NY
[4]Recognition Technologies, Inc., South Salem, NY

***Abstract*—Structural tests in experimental environments provide invaluable insight into structural behavior under service conditions, but often are not directly applicable to damage detection in-service. We address the generalization gap between experimental and in-service structures through (i) development of a neural network damage model learned from various experimental structures and (ii) identification of damage progression using this model for an in-service structure. We generate cepstral coefficients from accelerometer time histories across experimental structures and damage scenarios, which are classified by a time-delay neural network (TDNN) architecture to determine feature patterns which describe particular damage expressions. Classification of various damage classes over a large experimental structure corpus (the LANL SHM dataset, the IASC-ASCE benchmark, the U. Nevada & UCSD Seismic Bridge Columns test, the U. Houston NEESR Reinforced Columns test, the U. Illinois Urbana-Champaign 144-DoF Truss, and the Purdue Truss) provides the model a diversity of scenarios to cover the span of damage phenomena.**

Generalization of the damage model to in-service structures is performed through clustering of embeddings from this TDNN, which are more-general representations of damage in the domain learned only from the experimental structures. A model of the priors for some labeled damage scenarios with these in-service structure embeddings is generator via probabilistic linear discriminant analysis (PLDA). We assess these techniques on the Z24 Bridge Benchmark and obtain strong results for both ambient and forced vibration tests, in (i) an a-posteriori diagnosis of previous in-service damage to demonstrate the extent of separability through supervised clustering and (ii) hypothesis testing of an unseen queried condition's deviation from baseline condition in the PLDA spanned purely by experimental damage classes. Explainability is also investigated via association of in-service damage with seen experimental structure damage conditions, demonstrating feasibility for interpretable output-only structural health assessment.

I. Introduction

A key challenge in structural health monitoring is the observation and identification of unseen damage indicators. Much work has been done with constitutive methods to describe damage mechanisms through first-principles, from variations of the finite element method to system identification. These models can describe the best-fit trend of behavior or the stochastic properties of a damage mechanism, but certain nonlinearities are impossible to fully represent in a closed-form fashion, as the precise dynamic behaviors of damage can vary down to the atomic level and span a vast array of physical mechanisms (e.g. from fracture to hystresis to rust).

As a first-pass assessment, listening to a structural vibration adjusted to human-audible frequency range instead yields valuable context to make a spot decision for damage assessment. The human ear has been the primary diagnostic tool in traditional acoustics fields, such as audio engineering, machine fault detection, and sonar surveillance. The motivation is simple and intuitive, then, to apply techniques drawn from the audio domain and apply them in the structural domain for structural health monitoring. Artificial neural networks (ANNs) used broadly in speech ([1, 2, 3]) have already shown strong performance in capturing the nonlinear behaviors indicative of damage [4, 5], with approaches spanning damage classification to physical simulation to anomaly detection. Several benchmark structural testing datasets, such as the LANL SHM Toolbox datasets [6, 7], the Z24 Bridge Benchmark [8], and the IASC-ASCE benchmark structure [9] have been used to assess such data-driven approaches for uncovering nonlinear damage behaviors in real structures. Though traditional constitutive methods can identify a structure's degradation over its lifespan [10, 11], damage mechanisms usually encompass deviations from ideal linear models, such as cracks, corrosion, and fatigue. Discovery of the uncaptured nonlinear phenomena for damage detection (despite the inherent black-box nature of flexible data-driven models such as ANNs) motivates the use of data-driven techniques in structural health monitoring.

As the availability of SHM datasets is limited, generalization from one structural domain to another is of great interest. Recent efforts in transfer learning from a set of source structures to a target structure have shown the feasibility of a damage detection system with no foreknowledge of target-domain damage (or "zero-shot learning") and minimal cross-domain preprocessing. A generative adversarial network (GAN) approach from Soleimani-Babakamali et al. [12] adapts the approach from Soleimani-Babakamali et al. [13] to train on the Z24

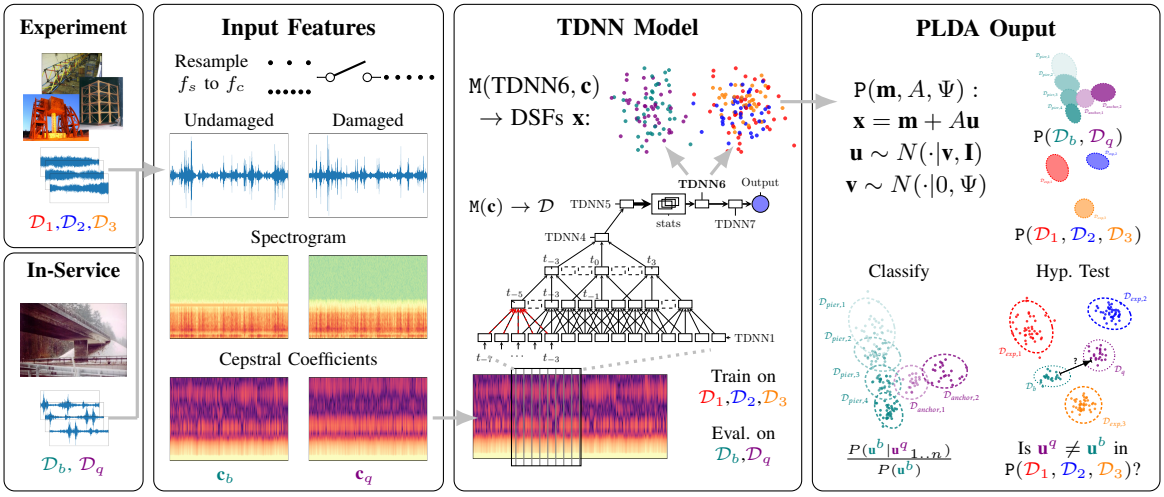


Fig. 1: Flowchart of TDNN-PLDA embedding generation and generalization approach. Accelerometer measurements taken from experimental structures conditions $\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3$ are transformed into LFCCs to train a TDNN as an embedding generator M. Embeddings $\mathbf{x} = \mathtt{M}(\text{TDNN6}, \mathbf{c})$ are used to train a PLDA model P. If preexisting in-service surveys $\mathcal{D}_q^{1\ldots Q}$, $\mathcal{D}_b^{1\ldots B}$ are available, a P developed on their corresponding embeddings can be used to classify and diagnose damage (Alg. 3). If a queried condition from an in-service structure $\mathcal{D}_q$ is being compared against a baseline condition $\mathcal{D}_b$, we can test whether these conditions are different in the latent space of $\mathtt{P}(\mathcal{D}_1, \mathcal{D}_2, \mathcal{D}_3)$ (Alg. 4).

Bridge, the QUDS structure, and the IASC-ASCE structure benchmarks individually to solve a binary undamaged/damaged classification problem per damage case, and a consensus of the GAN discriminators (calibrated by the target structure's baseline measurements) detects if a held-out target structure deviates from its baseline condition. From a population-based SHM approach, several feature transformation algorithms for domain adaptation were applied by Gardner et al. [14] to transfer a damage detection process between two model experimental structures and by Figueiredo et al. [15] for FEA to the Z24 in-service bridge. An innovative and unusual approach was taken by Tronci et al. [16, 17] to address the SHM data limitation challenge by leveraging the speech domain, via applying an x-vector time-delay neural network architecture trained on the speech VoxCeleb dataset (using Mel-Frequency Cepstral Coefficients as the preprocessed input features) and a probabilistic LDA classifier trained to distinguish damage cases for the Z24 Bridge.

Though these approaches provide robust damage detection processes with minimal prior knowledge or modeling of damage, insight into damage mechanisms associated with the detected damage is not well-explored. Furthermore, these studies mostly examine the binary problem of damage presence or absence for in-service structures. Extending detection to identification is a natural step to leverage the fitting capabilities of ANNs for damage mechanisms, which have so far been limited due to the availability of SHM datasets. Expanding the source domain to a wider corpus of structures is one possible path towards providing some attribution of the detected damage in the target domain, but has not been explored to date.

Learning damage patterns across a broad set of examples under a dimensionality-reducing framework should find a latent damage-aligned space to provide this insight. This perspective has been applied in speaker recognition, where learning underlying acoustic patterns for a vast corpus of speakers using a time-delay neural network (TDNN) architecture [18, 19] provides a speaker-feature backbone to generate a general representation of speakers from intermediate outputs of the TDNN. These low-dimensional representations, or embeddings, encompass a mapping to a latent space aligned with the speaker-recognition task in general. Recognition of unseen speakers and environments in this space is hence possible through less-demanding clustering or classifying algorithms. In particular, the x-vector TDNN architecture has shown state-of-the-art performance in speaker recognition [20, 21], using a probabilistic linear-discriminant analysis (PLDA) backend to cluster and classify the speaker embeddings [22, 23, 24]. Adaptation of the TDNN architecture and embeddings as damage-sensitive features (DSFs) has been explored for SHM transfer learning from a rich in-service vibration domain to a structural target domain, from the Z24 Bridge to the LANL UCI Column dataset in Hom et al. [25, 26] and the aforementioned speech to structures x-vector TDNN in Tronci et al. [16, 17].

This work further explores the suitability of TDNN embeddings as DSFs, using several SHM experimental structures to define a damage space mapped via a TDNN, and applying several PLDA classifier variants to infer changes in damage condition for the Z24 Bridge. Hence, our **source domain** is this library of experimental structures, and our **target domain** is an in-service structure. We extend the work from Tronci et al. [16, 17], Hom et al. [25, 26] in using a proven architecture from Snyder et al. [20] to show the efficacy of speaker recognition techniques on both in-service and experimental structures, and the general form of structural degradation learned can be applied across unseen structures. We outline our

TDNN-PLDA approach in Figure 1.

In brief, the contributions of this work are as follows:

1) We process structural vibration measurements into cepstral features to describe a potential damage-sensitive latent space.
2) We explore permutations of an established time-delay neural-network (TDNN) architecture to learn this damage-sensitive latent space from a broad corpus of experimental structures as our source domain.
3) Extracting intermediate outputs from our neural network as damage-sensitive features, we model damage mechanisms from experimental and in-service structures through probabilistic linear discriminant analysis (PLDA), to generate Gaussian distributions under a linear transformation across the damage classes.
4) We show that robust, explainable damage detection and identification on the Z24 Bridge example of in-service target domain can be achieved via the PLDA model of the TDNN DSFs learned from *only* experimental structures and damage.

II. Experimental and In-Service Structures

We provide a brief overview of each structure used, separated by experimental or in-service classification. Further information about these structures can be found in their reference documentation.

A. Experimental Testing Structures

We leverage several experimental structures from the following datasets:

1) Los Alamos National Laboratory SHMTools
   a) Bookshelf Frame Structure
   b) University of California, Irvine Bridge Column
2) U. Nevada & UCSD Seismic Performance of Bridge Columns
3) NEESR Payload Reinforced Concrete Columns
4) IASC-ASCE Building Benchmark
5) Purdue University Laboratory Truss SHM
6) University of Illinois at Urbana-Champaign 14-bay Truss

These experimental structure datasets were obtained from the NEES SHM challenge problem database, and offer a wide variety of excitations, materials, and designs to profile structural damage. To the best of the authors' knowledge, this is the broadest study over disparate structures to date (if only for model development and not strictly inference).

The *sole use* of experimental test structures is chosen specifically to illustrate the generalization utility of the TDNN-embedding method, as 1. no full in-service bridge dynamics are present in the training data 2. if successful in detecting in-service damage scenarios, experiments *alone* can yield a sufficient basis for damage detection for in-service structures. The collection of datasets is restricted to those accessible within a data-sparse domain (with respect to the number of unique structures), and additional structures would be of benefit for broader damage encapsulation. Hence, we describe in short detail the experimental structures employed, and provide a brief overview of the of experiments performed on these structures in Table I.

To present these damage labels for supervised training of our TDNN, we divide the damage types into five families of damage (Bolt Loosened, Stiffener Cut, Stiffener Removed, Hysteresis, and Earthquake) and collate overlapping damage scenarios in Table VIII and Table IX. Some engineering judgment is used to decide on the proper association between damage scenarios across structures, as we refer to the reference documentation to establish commonality between tests (i.e. points along the force-displacement hysteresis curves, quantity of bolts removed or cuts induced, etc.). We look to further developments from the SHM community and regulatory bodies to expand damage labeling standards.

1) LANL Bookshelf Frame Structure: The LANL Bookshelf problem [27] is a small model of a three-story building constructed of aluminum plates, struts, brackets, and steel bolts (Figure 2a). Damage is induced by loosening bolts to hand-tight, 5 ft-lbs, 10 ft-lbs, or removing bolts or brackets at various locations. Excitation is provided by a shaker at three levels of voltage (3, 5, and 7V) to shake the structure at an approximately 45 degree angle (to excite both bending and torsion modes of the structure). 24 accelerometers are placed along the corners of each plate to measure the lateral accelerations from the banded white-noise excitation. Each measurement is taken for 5.12 seconds at a sampling rate of 1600 Hz.

2) LANL UCI Bridge Column: The LANL UCI Bridge column dataset [7, 27] contains accelerometer measurements from two concrete bridge columns retrofitted with additional concrete jackets (cast-in-place and shotcrete). The columns are loaded along a hysteresis force-displacement curve to cyclically induce damage, resulting in six damage cases per column. A shaker mounted at the top of the column is used to excite the column with white noise. 40 accelerometers are mounted along the height of the column to measure lateral and vertical acceleration. Each measurement is taken for 8 seconds at a sampling rate of 1024 Hz.

3) U. Nevada & UCSD Seismic Performance of Bridge Columns: The U. Nevada & UCSD Seismic Performance of Bridge Columns dataset [28] examines the degradation of a full-scale bridge column after successive earthquake excitations (Figure 2c). Earthquake motion records from the 1989 Loma Prieta and 1995 Kobe earthquakes are increased in magnitude per test case to generate a progressive damage sequence over 10 different levels, and applied via a shake table. 23 accelerometers measuring horizontal and vertical accelerations are placed the height of the column. Additional strain gauges, linear voltage displacement transducers, spring potentiometers, and GPS sensors are placed on the column but are not used in this study (as vibrations captured through accelerations are the common modality we wish to use in damage identification). All sensors are acquired at a sampling rate of 240Hz over 160 seconds.

4) NEESR Payload Reinforced Bridge Columns: The NEESR Payload Reinforced Bridge Column study [29] was a series of tests on scale concrete bridge columns (e.g. Figure 2d) to evaluate "smart aggregate" sensors,

Name	Structure Type	Material	# Sensors (DoFs)	Sample Rate (Hz)	Time (s)	# Damage Cases	Damage Types
Z24 Bridge Benchmark	Bridge	Concrete	263	100	81.92	17	Various
LANL Bookshelf	Model	Aluminum	24	1600	5.12	4	Bolt Loosened
LANL UCI Bridge Column	Column	Concrete	40	1024	8.00	12	Hysteresis
Purdue Truss	Sign	Steel	23	280	320.00	8	Stiffener Cut
Seismic Bridge Columns	Column	Concrete	46	240	160.00	11	Earthquake
Reinforced Bridge Columns	Column	Concrete	36	20000	5.00	78	Hysteresis, Earthquake
Urbana Champaign Truss	Bridge	Steel	144	512	2.00	2	Stiffener Removed
IASC-ASCE Building Benchmark	Building	Steel	15	200	15.00	9	Bolt Loosened, Stiffener Removed

TABLE I: Summary of SHM Datasets Used

or piezoceramic transducers which generate an electric charge when stressed and a stress when excited by an electric charge. Hence, these smart aggregates could be used to excite or sense a concrete column to profile it for damage. Reinforced columns instrumented with these smart aggregates and accelerometers were subjected to hysteretic load cycles to induce progressive damage and failure . After each cycle, the column was excited by white noise on a shake table, and 9 smart aggregate outputs are recorded. An alternate round of progressive damage texting was also performed, replacing the load cycles with seismic excitations of increasing magnitude and excited by white noise afterwards, with 12 smart aggregates recording at a time. The smart aggregates are placed within the concrete and spaced evenly at the base, midpoint, and top of the column. Each waveform is taken at a sampling rate of 20kHz over 5 seconds.

5) IASC-ASCE Building Benchmark: The IASC-ASCE Building Benchmark [9] is a steel test structure designed to replicate a four-story building (Figure 2e). 9 different damage scenarios are induced, consisting of removal of braces between floors and bolt loosening for the connecting beams. Ambient, white-noise, and force-hammer excitations are used for each damage scenario. 15 accelerometers are placed along the lateral axes of the structure, and are recorded at a sample rate of 200Hz.

6) University of Illinois Urbana-Champaign 144-DoF Truss: The UIUC Truss [30] is a steel model truss bridge with a large number (144) of joints, shown in Figure 2f. The design was chosen to support damage localization techniques, and so damage scenarios are represented by replacement of a truss member with one of reduced cross-sectional area, resulting in loss of stiffness over various locations. This yields 11 different damage scenarios, though the type of damage remains the same across all cases. The truss is vertically excited by banded white-noise up to 200Hz in the midspan of the bridge. 8 accelerometers are placed vertically at 13 different, equally-distributed locations over 2 experiments per damage scenario. The waveforms are recorded at 512Hz over 2 seconds.

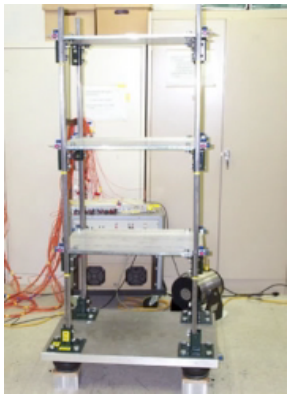
Damage Case	Assigned Location	Damage Scenario
1	None	Undamaged condition
2	HH4	Installation of pier settlement system
3	HH4	Lowering of pier, 20 mm
4	HH4	Lowering of pier, 40 mm
5	HH4	Lowering of pier, 80 mm
6	HH4	Lowering of pier, 95 mm
7	None	Lifting of pier, tilt of foundation
8	None	New reference condition
9	HH1,2,3	Spalling of concrete at soffit, 12 m2
10	HH1,2,3	Spalling of concrete at soffit, 24 m2
11	Setup1,9	Landslide of 1 m at abutment
12	Setup1,9	Failure of concrete hinge
13	Setup1,9	Failure of 2 anchor heads
14	Setup1,9	Failure of 4 anchor heads
15	HH1,3	Rupture of 2 out of 16 tendons
16	HH1,3	Rupture of 4 out of 16 tendons
17	HH1,3	Rupture of 6 out of 16 tendons

TABLE II: Z24 Progressive Damage Tests

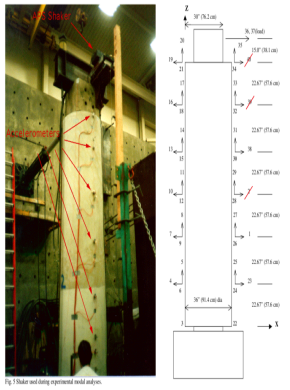
7) Purdue Highway Sign Support Truss: The Purdue Highway Sign Support Truss [31] is a full-scale aluminum highway sign truss taken from the I-29 interstate highway near Sioux City, Iowa (Figure 2g). Two single-damage cuts and one multidamage case were induced, involving partial cuts to the truss at the joint and midpoint locations of truss members. Hammer tests and shaker excitation of band-limited white noise from 0-200Hz were applied. To fully profile the damage over the entire structure, each excitation was applied for each damage case over 7 tests, to assemble a 144-DoF accelerometer waveform dataset. The 7 tests contain at most 23 acceleration waveforms per test, acquired at 280Hz over 320 seconds.

B. In-Service Structures

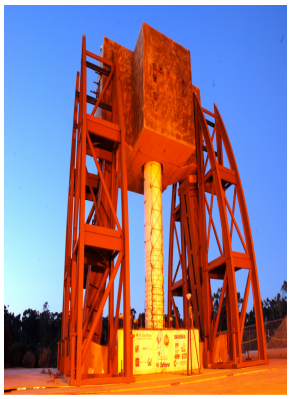
1) Z24: Experimental structural data for the Z24 Bridge Benchmark is provided from KU Leuven Structural Mechanics Section [32, 8]. An Environmental Monitoring System (EMS) dataset and a Progressive Damage Test



(a) LANL Bookshelf



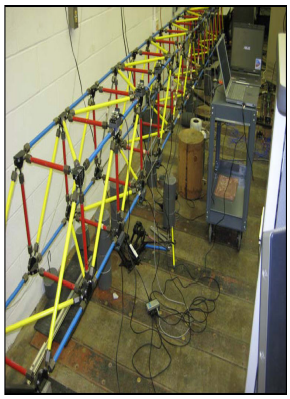
(b) LANL Column



(c) U. Nevada Seismic Column Bookshelf

(d) NEESR Reinforced Column

(e) IASC-ASCE Building



(f) UIUC Truss

(g) Purdue Highway Truss

Fig. 2: Experimental Structures Training Set

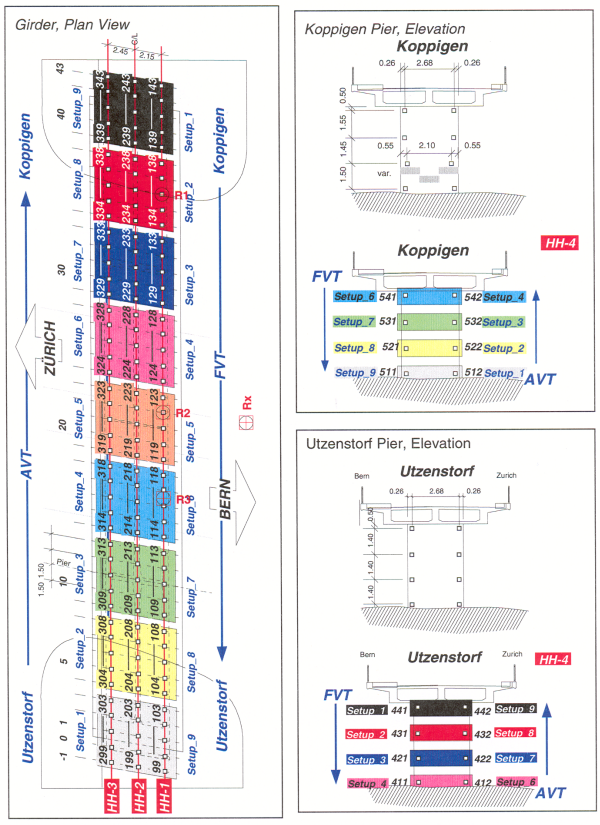


Fig. 3: Front and top view of the Z24 bridge. 3x45 accelerometers are placed along top and 2x8 along columns. Two shakers are placed at the middle span and one at the Koppigen side span [8].

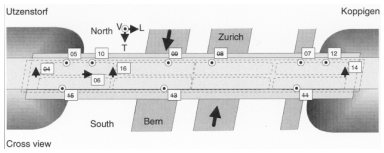
(PDT) were performed on the Z24 Bridge. The PDT dataset has 17 cases of applied damage Table II for the instrumentation in Figure 3 with ambient and forced vibration testing. In total, 263 measurement waveforms over the 3x45 span sensor grid and 2x8 pier sensor arrays are available per damage scenario, absent missing sensor records. As the number of channels which could be recorded was limited, each damage case is split into nine setups recorded at different times, indicated in the various colored blocks in Figure 3. The sensor measurements are sampled at 100Hz for 8.192 seconds. Anomalies in the sensor data are recorded in the test documentation, and are removed or addressed with the recommended processing from the documentation.

To evaluate our damage-detection system, we must choose a subset of accelerometers from the Z24 Bridge, as most typical SHM instrumentation will not be distributed to the spatial fidelity that the PDT is performed at. We use the Z24 Environmental Monitoring System test locations in Figure 4 to define a subset of 15 accelerometers from the PDT dataset for evaluation, as listed in Table III.

III. Input Features

A. Feature Selection and Processing

We choose cepstral coefficients as our vibration features, as they capture patterns of harmonic behavior and have shown strong effect in damage detection systems [33, 17].



(a) Z24 EMS Accelerometers on Span (b) Z24 EMS Accelerometers on Pier

Fig. 4: Z24 Bridge Environmental Monitoring System subset

Accelerometers				
304V	308V	317V	323V	334V
338V	104V	117V	134V	200T
211T	242T	209L	532L	522L

TABLE III: Z24 PDT Accelerometer Locations Corresponding to EMS Locations

Algorithm 1: Generate Linear Frequency Cepstral Coefficients from structural accelerometer waveforms (Input Features)

Input: structural acclerometer waveform $s \in \mathcal{S}$ at sampling rate f_s
Parameter: frequency cutoff f_c, frame width T_{frame}, frame overlap $T_{overlap}$, triangular filterbank $\mathcal{M}_{mk}$, cepstral coefficient resolution D
Output: structural LFCCs c_s

$\overline{\overline{s}} \leftarrow$ Resample s from f_s to f_c // 1

$[\overline{\overline{s}}\,[i]] \leftarrow$ Divide $\overline{\overline{s}}$ into frames $\overline{\overline{s}}\,[i]$ of T_{frame} duration and $T_{overlap}$ overlap

for $\overline{\overline{s}}\,[i] \in [\overline{\overline{s}}\,[i]]$ **do**
  $H_k \leftarrow$ Apply window function w, take Discrete Fourier Transform of $\overline{\overline{s}}\,[i]$ // 2
  $C_m \leftarrow$ Apply filterbank $\mathcal{M}_{mk}$, find log-spectrum of H_k // 3
  for $d \in D$ **do**
    $c_d \leftarrow$ Calculate d'th Linear-Frequency Cepstral Coefficient via Inverse Discrete Cosine Transform of C_m // 4
    $\tilde{c}_d \leftarrow$ Normalize via $\frac{c_d - \bar{c}_d}{\sigma_{c_d}}$
    Append $\tilde{c}_d$ to feature frame sequence $c_s[i]$
  end
end

From each waveform, we generate cepstral coefficients over a sliding sequence of frames, which will be provided to our TDNN damage classifier. We outline the signal processing pipeline in Algorithm 1.

1) Signal Processing: The sensor measurements from various structures all differ in number, length, location, sampling, and noise. As we want to minimize the prerequisites for transfer learning, we only consider sampling rate to standardize for our domain adaptation. Given multiple waveforms processed at different sampling rates, we must choose a content cutoff frequency f_c to standardize feature generation and resample all waveforms from f_s to f_c in our source and target structures to this frequency. For a signal s acquired at sampling frequency f_s, we resample s to $\overline{\overline{s}}$ with antialiasing FIR filter h_{aa} at frequency content cutoff f_c, which is chosen to be 250Hz (as the maximum frequency content in our experimental structures does not exceed this frequency):

$$\begin{aligned} \frac{p}{q} &= \text{rational}(\frac{f_c}{f_s}) \\ s_{\uparrow}[n_{up}] &= \sum_j s[j] h_{aa}[n_{up} - jp] \\ s_{\downarrow}[n_{down}] &= s_{\uparrow}[q n_{down}] \\ \overline{\overline{s}} &= s_{\downarrow} \end{aligned} \tag{1}$$

resulting in a signal s with N samples being resampled to have $\frac{p}{q}N$ samples (which, as enforced by the rational$(\cdot)$ operator, will be an integer-valued number of samples).

In order to track the evolution of damage patterns and test excitations, we must section the signal into a sequence of frames $[\overline{\overline{s}}\,[i]]$, where i is the index of the frame. We select frames of duration 1.6s with 0.64s overlap in the structural domain, which is sufficient for capturing several periods of low-frequency content. Each frame is used to generate our cepstral coefficient feature vector, which we detail below.

2) Cepstral Coefficients: Cepstral coefficients represent periodic characteristics within frequency spectra, such as the rate of repetitions of an echo, or the transformation of an excitation after reflection in a reverberant environment. The dynamics of a spectrum's changes over time is assumed to be of similar utility in the structural domain as it is in the speech domain. It is well-understood that structural modes may shift when stiffness is lost under a damaged condition. However, nonlinear damage phenomena such as fracture, snapping, and shearing manifest acoustically as grinding or rattling. Given that we want to identify changes in structural condition under a dynamic excitation, it is important to have features which separate the excitation's behavior from the structure's behavior and capture the identifying characteristics of nonlinear changes within the structure. We present the calculation of cepstral coefficients below to observe the benefits they can provide in this task.

Taking the i'th signal frame $\overline{\overline{s}}\,[i]$ and indexing by n into the frame, we calculate the cepstral coefficients following Appendix A in Beigi [34]. We begin by transforming the signal with the Discrete Fourier Transform (DFT) with window $w(n)$:

$$\begin{aligned} \bar{s}_n &= \overline{\overline{s}}\,[i][n] w(n) \\ H_k &= \sum_{n=0}^{N-1} \bar{s}_n e^{-i\frac{2\pi k n}{N}} \end{aligned} \tag{2}$$

We use the Hanning window as $w(n)$ for suppression of truncation transients, and the Fast Fourier Transform to perform the DFT.

The power spectrum is formed from the magnitude of the DFT, and a filterbank $\mathcal{M}_{mk}$ is applied to suppress

noise across frequency bands and reduce dimensionality of a high-resolution spectrum. Furthermore, the logarithm of this resulting filtered spectrum is taken, to reduce the dynamic range between high and low frequency content.

$$\begin{aligned} |H_k| &= \sqrt{\mathcal{R}e(H_k)^2 + \mathcal{I}m(H_k)^2} \\ \breve{H}_m &= \mathcal{M}_{mk}|H_k| \\ C_m &= \log\left(|\breve{H}_m|^2\right) \end{aligned} \tag{3}$$

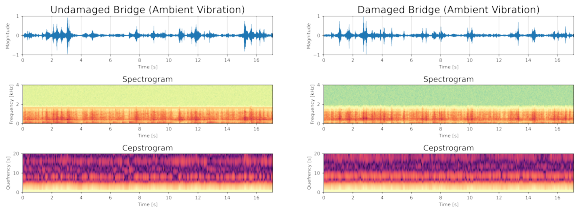
The filterbank is akin to a windowing function along the frequency axis, and assists in suppression of spurious frequency-transient effects when a cosine kernel function is applied.

Typically, the filterbank $\mathcal{M}_{mk}$ is spaced along Mel-frequency bins, which are logarithmically-spaced based on perception from a human ear However, as we are in the structural vibration domain, we expect structural changes to be best-identified in linear spacing of frequency, since both the plant and sensor behave linearly. Hence, we use linear-frequency triangular bins to construct $\mathcal{M}_{mk}$.

Many definitions of a cepstrum from a log-power-spectrum exist, as several basis functions can suffice to identify spectral patterns, but here we use the Inverse Discrete Fourier Transform (IDFT). As C_m is entirely real-valued, we can safely rely on the cosine component to isolate relevant log-spectral periodic behavior, which is more computationally efficient than the full IDFT. Applying the Inverse Discrete Cosine Transform (IDCT) of C_m gives us cepstral coefficients c_d:

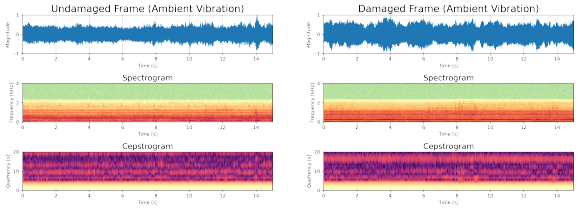
$$c_d = \sum_{m=0}^{M-1} a_m C_m \cos\left(\frac{\pi(2d+1)m}{2M}\right) \tag{4}$$

$$a_m = \begin{cases} \frac{1}{M} & \text{for } m = 0 \\ \frac{2}{M} & \forall\, m > 0 \end{cases} \tag{5}$$



(a) Undamaged Z24 Bridge under ambient vibration.

(b) Damaged Z24 Bridge under ambient vibration.



(c) Undamaged IASC-ASCE Frame under ambient vibration.

(d) Damaged IASC-ASCE Frame under ambient vibration.

Fig. 5: Waveform, spectrogram, and cepstrogram for structural vibrations from the Z24 Bridge and IASC-ASCE Frame. Even under ambient vibration, changes in structural dynamics are evident in the cepstral trajectories (and arguably more so than in spectral profiles).

These cepstral features are calculated for each frame, and are mean- and variance-normalized ($\frac{c_d - \bar{c}_d}{\sigma_{c_d}}$) over a sliding window up to three seconds in length to reduce transient variability across the entire duration of the measurement. This approach mirrors that taken in Snyder et al. [20], as the structural measurements have been transformed in the previous section to replicate the nominal acoustic characteristics of speech.

B. Data Augmentation

1) Corpus Expansion: The available sensor data for training the neural network model can be increased, or augmented, through speed perturbation and addition of foreground and background noise. Inclusion of these corrupted data will improve the network's tolerance for variations in structure and excitation. Speed variation of 0.9x, 1.0x, and 1.1x through resampling is performed before noise corruption, which is analogous to shifting the modes of the structure. The augmentation is applied in with three independent noise sources, of which half of the total augmentation is randomly selected for inclusion in the training set. In total, this augmentation expands the dataset by approximately eightfold.

- **foreground noise:** 22 white and brown noise sources are randomly selected and added to the original sensor waveform (15-25dB SNR)
- **background noise:** 22 white and brown noise sources are randomly selected and convolved and added to the original sensor waveform (15-25dB SNR)
- **background traffic:** One to three of 26 traffic audio recordings from pedestrian and vehicle throughput from soundjay.com are summed, then convolved and added to the original sensor waveform (10-20dB SNR)

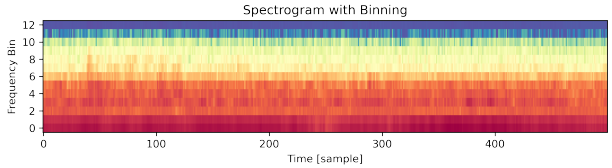
2) Spectral Augmentation: We can apply an IDCT to the LFCC features for an approximate reconstruction of the spectral content of the signal. This is generally performed after cepstral mean and variance normalization for a frequency-domain representation of processed cepstral features. In this domain, we can drop-out select sections of the reconstructed spectrum and attempt to classify this modified feature, with an example in Figure 6. By denying the full feature's information to the classifier, we can increase robustness of the classifier to spectral variations in characteristic cepstral aspects of damage-specific behaviors.

IV. Neural Network Vibration Model

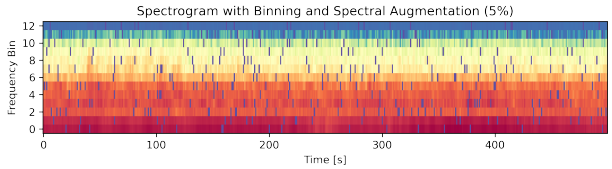
We provide a brief summary of kernel mechanisms and architecture for time delay neural networks for our vibration model. Far more detail is provided in general deep learning literature, but we direct some intuition for the utility of each technique towards nonlinear behaviors in physical systems (e.g. damage mechanisms in structures).

A. Kernel Mechanisms

1) Linear (Fully-Connected) Layer: The standard fully-connected layer applies a transformation matrix W and a



(a) Spectrogram after IDCT of LFCC.



(b) Spectrogram after IDCT of LFCC with spectral augmentation.

Fig. 6: Spectral augmentation drops-out frequency features over time.

bias $\mathbf{b}$ to map the input from a previous layer $\mathbf{h}^l$ to the next layer $\mathbf{h}^l$, within a nonlinearly-transformed space $\sigma(\cdot)$.

$$\mathbf{h}^{l+1} = \sigma(W^l \mathbf{h}^l + \mathbf{b}^l)$$
$$\text{where } \mathbf{h}^l = \sigma(W^{l-1} \mathbf{h}^{l-1} + \mathbf{b}^{l-1})$$

All elements of $\mathbf{h}^l$ are directly related (or "fully-connected") to $\mathbf{h}^{l+1}$. If we construct a network with several of fully-connected layers, the resulting nonlinear map is a sequence of affine transformations of nonlinear features, which is appropriate for modeling discriminating features as linearly-separable.

The selection of the nonlinearity $\sigma(\cdot)$, or activation function, is hence important for creating the nonlinear space to extract feature patterns from. Typical activations include the tanh, sigmoid, regularized linear unit (ReLU), and swish functions, and much work continues in finding task-appropriate activation functions across deep learning. We choose ReLU for our TDNN backbone and neck layers, and softmax for our TDNN classification head.

2) Convolutional Layer: In more-conventional/classical filtering contexts, a phenomena is often assumed to be decomposable into a basis function (or kernel) and an excitation, which are mixed (or "convolved") together via integrating the product of the two kernels over time. The resulting integrated energy from the stimulation/convolution of a system $g(t)$ via an excitation $f(t)$ gives us the total system response $f \circledast g(t)$:

$$f \circledast g(t) := \int_{-\infty}^{\infty} f(\tau) g(t-\tau) d\tau$$
$$:= \int_{-\infty}^{\infty} f(t-\tau) g(\tau) d\tau$$

We can extend this along a two dimensions in x, y:

$$f \circledast g(x,y) := \iint_{-\infty}^{\infty} f(\tau_x, \tau_y) g(x-\tau_x, y-\tau_y) d\tau_x d\tau_y$$

where we seek local cross-correlational structure in both x and y which are associated with the phenomenon of interest. Following this motivation, we use 2D convolutional layers to learn time-frequency relationships across a spectrogram or cepstrogram which are indicative of damage mechanisms, where we use the discrete 2D convolution for a convolutional $M \times N$ image layer output $h_{c_{out}}^{l+1}$:

$$h_{c_{out}}^{l+1}[m_{out}, n_{out}] = \sum_{c_{in}=0}^{C_{in}-1} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} W_{c_{out},c_{in}}[n,m] h_k^l[m_{in}-m, n_{in}-n]$$

In application to time-varying features, such as spectra or cepstra over frames, the convolutional layer serves to capture local variations of these features in kernels associated with a particular damage mechanism. This inherently enforces a sequential structure to the patterns learned to be representative of damage with fewer parameters, whereas the fully-connected layer relates all inputs of the previous layer to all outputs of the next layer without requiring sequential relationships in its design. Most modern convolutional networks decompose kernel functions into a 3-element convolution filter, as this is the most-atomic filter which can be used to identify sequential patterns.

3) Attention Layer: The attention mechanism from Vaswani et al. [35] is well-explored in literature. We cover the core motivation and concepts here briefly for use in damage detection. Multihead dot-product attention with positional encoding is used for finding patterns across long contexts, through a learned weighted average of dot-product similarity between an input 'key' and a query bank. In comparison with convolution layers, an identifying key element which only occurs over a long period would require an N-sample kernel width, resulting in large correlation maps for this kernel. Instead of generating a sequence of long correlation maps per layer, we would want to associate correlation outputs across long timespans (or large "context") without needing wide receptive fields via long kernels or stacking convolutional layers.

Hence, we follow the approach from Povey et al. [36] for time-restricted attention. We can restrict query, key, and value matrices Q, K, V to be in a time window $L + t + R < N$ (at time t for left and right context size of L, R samples respectively), and evaluate attention across this window slid over our N samples. With appropriate padding to the left and right, we obtain an array of time-restricted attention values $\text{Attn}_{tr} \in \mathbb{R}^{N \times (L+1+R)}$.

$$\text{Attn}(\mathbf{q_t}, \mathbf{k_t}, \mathbf{v_t}) = \sum_{i=0}^{N} softmax(\mathbf{q_i} \cdot \mathbf{k_i}) \mathbf{v_i}$$
$$\text{Attn}_{\text{tr}}(\mathbf{q_t}, \mathbf{k_t}, \mathbf{v_t}) = \sum_{\tau=t-L}^{t+R} softmax(\mathbf{q_t} \cdot \mathbf{k}_\tau) \mathbf{v_t}$$

and encode position of the attention in the time-window via appending a one-hot vector as follows:

$$\mathbf{v_t} = \text{extend}(\mathbf{v_t}, \mathbb{1}(\tau + L - t))$$
$$\mathbf{k_t} = \text{extend}(\mathbf{k_t}, \mathbb{1}(\tau + L - t))$$

where τ indicates the relative position of the value or query with respect to the current time t and the leftmost context L, and $\mathbb{1}(\tau + L - t)$ is the indicator function which is 1 at $\tau + L - t$ and 0 everywhere else. This explicitly indicates which key is associated with a query in the sequence's order. The time-restricted multihead attention mechanism,

for some learnable weight matrices W_i^Q, W_i^K, W_i^V is expressed as:

$$\text{Multihead}_{\text{tr}}(\mathbf{q_t}, \mathbf{k_t}, \mathbf{v_t}) = \text{Concat}(\text{head}_1, ..., \text{head}_h)W^O$$
$$\text{where head}_i = \text{Attn}_{\text{tr}}(\mathbf{x_t}W_i^Q, \mathbf{x_t}W_i^K, \mathbf{x_t}W_i^V)$$

This contraction of conventional attention trades-off the full context of the desired signal when compared to conventional dot-product attention; by restricting the context, we sacrifice potential key-query associations beyond the window of interest.

However, detection of long-delayed patterns can be recovered via the positional encoding, which addresses our motivation to capture long-timespan patterns which may be indicative of damage.

B. TDNN Architecture

We use a time-delay neural network with the same structure presented in the x-vector formulation [20] to find distinguishing structural damage characteristics across the experimental structure damage scenarios. The TDNN captures vibration dynamics by learning relationships between features over a sequence of frames (allocated for each accelerometer signal in Section III-A) corresponding to the damage scenario. Exposure to past or future features from a given point in the sequence is accomplished by defining a range of frames, or context, that are connected in a layer. As the input layer is restricted to an ordered sequence of time-dependent features, the outputs of the following hidden layer can be considered as an compressed (or abstracted) representation of structural damage dynamics over the defined context.

As shown in Figure 7, the first four layers (TDNN1-TDNN4) collect several frames of context from the sensor waveform before and after the frame being assessed. To reduce sensitivity to the selection of frames when segmenting the signal, a statistics pooling layer is inserted after sufficient frame-level representations are collected. This statistics pooling layer collects (or 'pools') all of the TDNN4 outputs in the segment of the measurement input and returns the mean and standard deviation for the segment, compressing the context of the following layers into segment-level representations. The following layers (TDNN6, TDNN7) fit the statistics from the T sequences to the corresponding damage scenarios.

We show the original x-vector TDNN architecture in Table IV. We also experiment with the kernel mechanisms described above, in the configurations listed in Table V. We maintain the total context seen (15 frames) across all architectures, by enforcing the convolutional kernels to have a kernel context width of 5 in TDNN1 and 3 (with dilation of 2) in TDNN2, and the attention context to be 5 in TDNN2. These architectural studies are by no means exhaustive, but serve as a preliminary modification to the TDNN x-vector architecture used in previous structural damage detection efforts [16, 17, 25, 26].

To prevent overfitting, we apply dropout and batch-normalization for each layer, which reduces network on neurons and training data respectively.

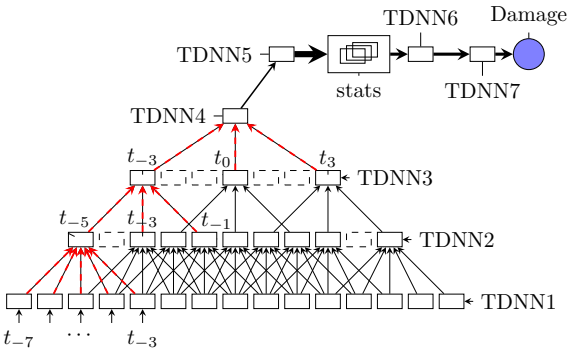


Fig. 7: The contextual structure of the TDNN, with labels t_n for the boxes, or layers, centered at that particular frame from the input sequence. Each 'box' in each layer (TDNN#) represents a network kernel mechanism described in Table IV and Table V.

1-2: 15 frames ($t_{-7} \leq t_0 \leq t_7$) are provided to the input layer (TDNN1). Following the red arrows from TDNN1 to TDNN2, TDNN2's t_{-5} node receives five frames (with context defined as [-2,-1,0,1,2]) as inputs (which then corresponds to the range $t_{-7} \leq t_{-5} \leq t_{-3}$).

2-3: As TDNN3 has context defined as [-2,0,2], TDNN3's t_{-3} node receives the outputs from TDNN2's t_{-5}, t_{-3}, t_{-1} nodes.

3-4: Similarly, TDNN4 has context defined as [-3,0,3], and receives the outputs from TDNN3's t_{-3}, t_0, t_3 nodes.

4-5: TDNN5 does *not* use any additional temporal context, and so is connected to TDNN4 only between neurons (there are only 512×1500 connections between the 512 neurons in TDNN4 and 1500 in TDNN5).

stats: Once the frame-level relationships have been established in layers TDNN1-5, the mean and standard deviation for each of the TDNN5 outputs across the entire segment (or 'chunk' of accelerometer signal determined from a segment duration hyperparameter) are pooled together and concatenated to represent the segment (taking the 1500 outputs of the TDNN5 layer, calculating mean and standard deviation across the T sequences in the chunk, and producing 3000 outputs, which are the 1500 means and 1500 standard deviations concatenated together).

6-Out: This pooled statistic output, representative of the acoustic behavior over the entire chunk of the accelerometer signal, is propagated through TDNN6, TDNN7, and the final softmax classification layer to fit segment-level behavior to the damage scenario.

TABLE IV: x-vector TDNN Layers

Layer	Layer Type	Layer Context	Total Context	Input × Output
TDNN1	Fully-Connected	[-2,-1,0,1,2]	5	(40×5)×512
TDNN2	Fully-Connected	[-2,0,2]	9	(512×3)×512
TDNN3	Fully-Connected	[-3,0,3]	15	(512×3)×512
TDNN4	Fully-Connected	[0]	15	(512×1)×512
TDNN5	Fully-Connected	[0]	15	512×1500
stats	μ, σ	[0:T)	T	(1500×T) →(2*1500)
TDNN6	Fully-Connected	[0]	T	3000×512
TDNN7	Fully-Connected	[0]	T	512×512
Damage	Fully-Connected	[0]	T	512×$n_{classes}$

TABLE V: TDNN Layer Permutations

Experiment	Layer Name	Layer Type	Layer Context	Total Context	Input × Output
Fully-Connected	TDNN1	Fully-Connected	[-2:2]	5	(40×5)×512
	TDNN2	Fully-Connected	[-2,0,2]	9	(512×3)×512
	TDNN3	Fully-Connected	[-3,0,3]	15	(512×3)×512
Convolution	TDNN1	Convolution	[-2:2]	5	(40×5) ⊛ (5×3*64)
	TDNN2	Convolution	[-2,0,2]	9	(15×20*64) ⊛ (3×3*128)
	TDNN3	Fully-Connected	[-3,0,3]	15	(11×10*128) × 512
Attention	TDNN1	Fully-Connected	[-2:2]	5	(40×5) × 512
	TDNN2	Attention	[-2:2]	9	(512×5) × (5×(60+5))
	TDNN3	Fully-Connected	[-3,0,3]	15	5×(60+5×3) ×512

C. Training Task

Using the labels from our training data, we attempt supervised classification of different damage scenarios across several structures. Our objective function is multiclass cross-entropy, to determine the likelihood $P(y_k|\mathbf{x}_{nk})$ of damage class y_k given a corresponding feature set $\mathbf{x}_{nk}$, over N frames of the waveform and K damage classes:

$$E = \sum_{n=1}^{N}\sum_{k=1}^{K} -\mathbb{1}_{y_k}(n)\log(P(y_k|\mathbf{x}_{nk}))$$

Potential class-imbalance is handled via randomly selecting a segment of each experimental structure waveform and distributing these segments evenly by class and duration per batch, with the augmentations of speedup and noise applied beforehand. The TDNN accordingly is presented with each class evenly, with sufficient batches such that all segments of the experimental structure waveforms are seen per epoch.

Minimization of the cross-entropy loss is performed by updating the neural network weights through backpropagation. We use natural gradient descent and averaging of network parameters to update our TDNN weights [37].

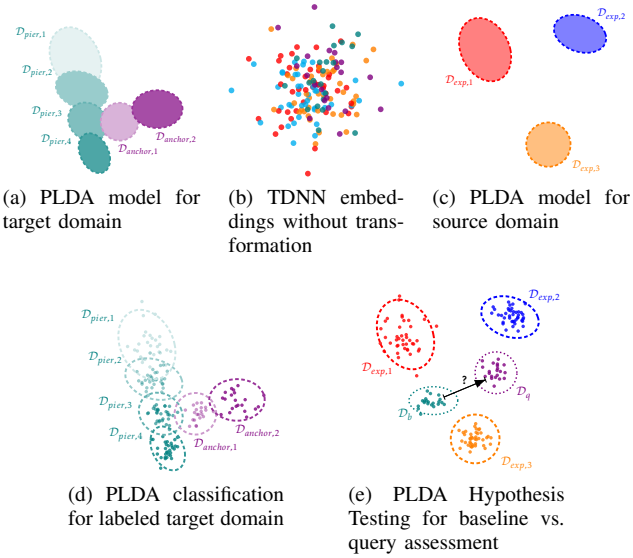
D. Embedding Extraction

The damage-sensitive features we wish to classify for each accelerometer signal are embeddings obtained from the outputs of TDNN6, which is a layer with 512×1 neurons. We choose this layer as it corresponds to the x-vector extraction layer in Snyder et al. [20]. Hence, each embedding is a 512×1 dimensional vector, which present more degrees to attempt classification on.

We expect that extracting outputs at a layer close, but not at the final output layer may yield better class representations and reduce the effect of overfitting the network to particular damage scenarios. As these embeddings from TDNN6 are not restricted to the dimension of the training classes, we treat them as damage-sensitive features for any structure in our source or target domain. The translation of embeddings into the target domain requires a post-processing step to discriminate between target domain damage scenarios, which we lay out in the next section.

V. Classification and Hypothesis Testing

A. PLDA Modeling



(a) PLDA model for target domain

(b) TDNN embeddings without transformation

(c) PLDA model for source domain

(d) PLDA classification for labeled target domain

(e) PLDA Hypothesis Testing for baseline vs. query assessment

Fig. 8: PLDA model and assessment approaches. As embeddings from the TDNN are not generally aligned with the target domain like in (b), they must be transformed to support detection in the target domain (e.g. (a) for labeled target-domain data, (c) for labeled source-domain data). If explicit labels are provided, then classification can be performed via the Gaussian models of the labeled target domain in the PLDA latent space (d). If no explicit labels are provided, then an existing PLDA from the source domain can be used to support modeling of baseline and query embeddings drawn from the target domain (e).

1) Model Development: We first use Linear Discriminant Analysis (LDA) as a dimensionality reduction technique, to reduce occlusion from overlapping features between class clusters. The LDA technique is very mature, and we will not go through the full derivation here.

Algorithm 2: Probabilistic Linear Discriminant Analysis model from Time-Delay Neural Network structural waveform embeddings (PLDA Development)

Input: LFCCs $\mathbf{c}_{\mathcal{S},\mathcal{D}}$ for structures $\mathcal{S}$ and damage $\mathcal{D}$

Parameter: TDNN `M`, TDNN embedding layer number l, dimension reduction d_{LDA}

Output: PLDA model $\mathtt{P}(\mathbf{m}, A, \Psi)$

$x_{\mathcal{S},\mathcal{D}} \leftarrow$ extract embeddings at l from $\mathtt{M}(l, \mathbf{c}_{\mathcal{S},\mathcal{D}})$

$T, S_W, S_B \leftarrow$ calculate LDA transform and between- and within-class scatter matrices for $x_{\mathcal{S},\mathcal{D}}$ over $\mathcal{D}$ //[34, 38]

$\Lambda_W, \Lambda_B \leftarrow$ apply dimensionality reduction $T(d_{LDA})$ to S_W, S_B // 6

$\mathbf{m}, A, \Psi \leftarrow$ calculate P's parameters from $\mathbf{x}_{\mathcal{S},\mathcal{D}}, \Lambda_W, \Lambda_B, T(d_{LDA})$ via max likelihood // 8,[22]

Algorithm 3: Probabilistic Linear Discriminant Analysis scoring from Time-Delay Neural Network structural waveform embeddings (PLDA Classification)

Input: LFCCs $\mathbf{c}_{\mathcal{S},\mathcal{D}_g}, \mathbf{c}_{\mathcal{S},\mathcal{D}_p}$ for structures $\mathcal{S}$, gallery damage conditions $\mathcal{D}_g$, probes from damage condition(s) $\mathcal{D}_p$

Parameter: TDNN embedding extractor $\mathtt{M}(l, \mathbf{c})$, LDA transform $T(d_{LDA})$, PLDA model $\mathtt{P}(\mathbf{m}, A, \Psi, \mathbf{x})$

Output: Log-Likelihood Ratio score $LLR_{\mathcal{D}}$ for $\mathbf{u}^p \in \mathcal{D}_g$

$\mathbf{x}_{\mathcal{S},\mathcal{D}_g}, \mathbf{x}_{\mathcal{S},\mathcal{D}_p} \leftarrow$ extract embeddings at l from $\mathtt{M}(l, \mathbf{c}_{\mathcal{S},\mathcal{D}_g})$, $\mathtt{M}(l, \mathbf{c}_{\mathcal{S},\mathcal{D}_g})$

$\mathbf{u}^g, \mathbf{u}^p \leftarrow$ apply $T(d_{LDA})$ to embeddings $\mathbf{x}_{\mathcal{S},\mathcal{D}_g}, \mathbf{x}_{\mathcal{S},\mathcal{D}_p}$

$LLR_{\mathcal{D}} \leftarrow$ compare probe $\mathbf{u}^p$ to $\frac{P(\mathbf{u}^p|\mathbf{u}^g_{1..n})}{P(\mathbf{u}^p)}$ from the PLDA model P // 9

The LDA problem aims to find some transformation T which maximizes the scatter between classes S_B while minimizing the scatter within classes S_W. T can be found via the solution to the generalized eigenvalue problem [34, p.402-404], [38] The dimensionality reduction occurs in keeping the d_{LDA} largest eigenvalues (and corresponding eigenvectors in T) to generate the LDA transformation $T_{d_{LDA}}$:

$$\begin{aligned} T_{d_{LDA}} &= [\mathbf{t}_1, \mathbf{t}_2, \dots \mathbf{t}_{d_{LDA}}] \\ \mathbf{x} &\leftarrow T^T_{d_{LDA}} \mathbf{x} \\ \Lambda_b &= T^T_{d_{LDA}} S_b T_{d_{LDA}} \\ \Lambda_w &= T^T_{d_{LDA}} S_w T_{d_{LDA}} \end{aligned} \tag{6}$$

As classification is inherently a distribution-separation problem, we can leverage the probabilistic extension of LDA (appropriately called Probabilistic LDA, or PLDA) from Ioffe [22], to score the embeddings for their likelihood of membership in the damage scenarios. The PLDA model assumes the following structure:

$$\begin{aligned} \mathbf{x} &= \mathbf{m} + A\mathbf{u} \\ \mathbf{u} &\sim N(\cdot|\mathbf{v}, \mathbf{I}) \\ \mathbf{v} &\sim N(\cdot|0, \Psi) \end{aligned} \tag{7}$$

Our modeled example $\mathbf{x}$ is drawn from a distribution centered around a mean $\mathbf{m}$ and a latent space representation $\mathbf{u}$ projected to the native space by A. $\mathbf{u}$ is centered about class centers $\mathbf{v}$ with unit within-class covariance I, and each class center $\mathbf{v}$ is drawn from a normal distribution centered at $\mathbf{0}$ with between-class covariance Ψ. In other words, all class centers $\mathbf{v}$ live within the covariance ellipsoid described by $N(\mathbf{0}, \Psi)$, within which all class examples u live within the covariance ellipsoids described by $N(\mathbf{v}, \mathbf{I})$.

The PLDA model parameters $\mathbf{m}, A, \Psi$ are obtained by solving a maximum likelihood problem. Following the procedure laid out in Ioffe [22], we obtain the following parameters $\mathbf{m}, A, \Psi$ from the LDA transformation to specify the Gaussian distributions about the classes, where $\mathbf{x}$ is represented by mean $\mathbf{m}$ and distributed about $\mathbf{u}$ in the latent space.

$$\begin{aligned} \mathbf{x} &= \mathbf{m} + A\mathbf{u} \\ \mathbf{m} &= \frac{1}{N}\sum_{i=1}^{N} \mathbf{x}_i \\ A &= T^T(\frac{n}{n-1}\Lambda_w)^{1/2} \\ \Psi &= \max(\frac{n-1}{n}(\Lambda_b/\Lambda_w) - \frac{1}{n}) \end{aligned} \tag{8}$$

2) Classification Scoring with PLDA: Classification hence is performed in the LDA-transformed (or 'latent') space. In Ioffe [22] and the corresponding implementation in Povey et al. [39], the scoring of a given "probe" example $\mathbf{u}^p$ to be classified in the PLDA model can be determined from the ratio of the likelihood of the probe coming from the "gallery" population of a particular class $\mathbf{u}^g_{1\dots n}$ (i.e. the data samples enrolled into the PLDA classifier for a damage case) against the probability of that probe example occurring in $\mathcal{X}$. This is expressed as a likelihood ratio $P(\mathbf{u}^p|\mathbf{u}^g_{1..n})/P(\mathbf{u}^p)$, where the numerator is given by the following equation in Ioffe [22]:

$$\begin{aligned} P(\mathbf{u}^p|\mathbf{u}^g_{1\dots n}) &= N(\mathbf{u}^p|\frac{n\Psi}{n\Psi + I}\bar{\mathbf{u}}^g, I + \frac{\Psi}{n\Psi + I}) \\ \bar{\mathbf{u}}^g &= \sum_{i=1}^{n} \mathbf{u}^g_i \end{aligned}$$

If $n = 0$, we obtain the denominator corresponding to the probe belonging to none of the gallery classes in the PLDA model:

$$P(\mathbf{u}^p) = N(\mathbf{u}^p|\mathbf{0}, I + \Psi)$$

And so our likelihood ratio (and log-likelihood ratio) are

given by

$$\begin{aligned}
\frac{P(\mathbf{u}^p|\mathbf{u}^g_{1..n})}{P(\mathbf{u}^p)} =& \frac{N(\mathbf{u}^p|\frac{n\Psi}{n\Psi+I}\bar{\mathbf{u}}^g, I+\frac{\Psi}{n\Psi+I})}{N(\mathbf{u}^p|\mathbf{0}, I+\Psi)} \\
\log\frac{P(\mathbf{u}^p|\mathbf{u}^g_{1..n})}{P(\mathbf{u}^p)} =& \\
-\frac{1}{2}\Bigg[&(\mathbf{u}^p-\mathbf{m})(I+\frac{\Psi}{n\Psi+I})^{-1}(\mathbf{u}^p-\mathbf{m})+ \\
&\log\det(I+\frac{\Psi}{n\Psi+I})\Bigg]- \\
&\left(-\frac{1}{2}\left[\mathbf{u}^p(I+\Psi)\mathbf{u}^p+\log\det(I+\Psi)\right]\right) \\
\mathbf{m} =& \frac{n\Psi}{n\Psi+I}\bar{\mathbf{u}}^g
\end{aligned} \tag{9}$$

Hence, we use the PLDA model to generate a log-likelihood score from Equation (9), and whichever class the log-likelihood ratio is highest for signifies that the probe accelerometer signal comes from that damage class. Much more could be performed using the PLDA approach (e.g. attempting to identify what damage scenario a held-out scenario is most similar to), but we have not investigated further applications yet. We also do not currently leverage the ability of the PLDA to model variations in the discriminants, to identify new, unseen classes as a combination of the classes seen in the LDA transformation. These remain for future efforts.

3) Hypothesis Testing with PLDA: If we only have a baseline condition to generate embeddings from the target domain, we can still calibrate a PLDA classifier and assess whether a new series of measurements taken at a later time has degraded from the taken baseline. This is akin to an anomaly detection approach, where only an initial sample of the target domain data is enrolled, and deviations over further samples are assessed as degradation from the nominal condition. In our approach, this is performed through viewing PLDA as a hypothesis testing task, where the model of classes as normal distributions in the latent space can be used to determine the likelihood of an unseen scenario belonging to another unseen scenario within this latent space.

Given our source structures' embeddings and our target structure baseline embeddings, we desire to generate a damage space which best-separates a new potential damage scenario from the baseline case. Association of the target baseline to the appropriate source baselines is not necessarily straightforward, as the desired target domain contains in-service structures, whereas the source domain contains experimental structures (which are generally scale models and components of these in-service structures). We do not always know the underlying pre-existing condition when we begin monitoring of a structure, and often have to assume some initial damage state. For this work, we assume that an in-service structure first queried for assessment is in good operational status, and is appropriate for association with the other undamaged baseline states from our source domain.

If we include the target domain's current condition with our source domain classes, we can then take advantage of the damage progression done across the experimental

Algorithm 4: Hypothesis Testing for unseen damage via pretrained TDNN-PLDA damage classifier (PLDA Hypothesis Testing)

Input: LFCCs $\mathbf{c}_{\mathcal{S},\mathcal{D}_b}, \mathbf{c}_{\mathcal{S},\mathcal{D}_q}$ for structure $\mathcal{S}$, baseline condition $\mathcal{D}_b$, unseen queried damage condition $\mathcal{D}_q$
Parameter: TDNN embedding extractor $\mathtt{M}(l, \mathbf{c})$, LDA transform $T(d_{LDA})$, PLDA model $\mathtt{P}(\mathbf{m}, A, \Psi, \mathbf{x})$, similarity prior for baseline and damage condition π_{same}, discrepancy prior for baseline and damage condition π_{diff}
Output: Damage Decision $D^?$ for $\mathcal{D}_q \in \mathcal{D}_b$
$\mathbf{x}_{\mathcal{S},\mathcal{D}_b}, \mathbf{x}_{\mathcal{S},\mathcal{D}_q} \leftarrow$ extract embeddings at l from $\mathtt{M}(l, \mathbf{c}_{\mathcal{S},\mathcal{D}_b}), \mathtt{M}(l, \mathbf{c}_{\mathcal{S},\mathcal{D}_q})$
$\mathbf{u}_{\mathcal{S},\mathcal{D}_b}, \mathbf{u}_{\mathcal{S},\mathcal{D}_q} \leftarrow$ project into LDA-transformed space via $T(d_{LDA})$
$\mathcal{P}_{same}(\mathbf{u}_{\mathcal{S},\mathcal{D}_b}, \mathbf{u}_{\mathcal{S},\mathcal{D}_q}) \leftarrow$ calculate likelihood from concat($\mathbf{x}_{\mathcal{S},\mathcal{D}_b}, \mathbf{x}_{\mathcal{S},\mathcal{D}_q}$), Ψ `// 11`
$\mathcal{P}(\mathbf{u}_{\mathcal{S},\mathcal{D}_b}) \leftarrow$ calculate likelihood from $\mathbf{x}_{\mathcal{S},\mathcal{D}_b}$, Ψ `// 11`
$\mathcal{P}(\mathbf{u}_{\mathcal{S},\mathcal{D}_q}) \leftarrow$ calculate likelihood from $\mathbf{x}_{\mathcal{S},\mathcal{D}_q}$, Ψ `// 11`
$\mathcal{P}_{diff}(\mathbf{u}_{\mathcal{S},\mathcal{D}_b}, \mathbf{u}_{\mathcal{S},\mathcal{D}_q}) \leftarrow \mathcal{P}(\mathbf{u}_{\mathcal{S},\mathcal{D}_b})\mathcal{P}(\mathbf{u}_{\mathcal{S},\mathcal{D}_q})$
$R^{-1} \leftarrow \frac{\mathcal{P}_{diff}(\mathbf{u}_{\mathcal{S},\mathcal{D}_b}, \mathbf{u}_{\mathcal{S},\mathcal{D}_q})}{\mathcal{P}_{same}(\mathbf{u}_{\mathcal{S},\mathcal{D}_b}, \mathbf{u}_{\mathcal{S},\mathcal{D}_q})}$ `// 10`
if $R^{-1} > \frac{\pi_{diff}}{\pi_{same}}$ **then**
  $D^? \leftarrow$ True
else
  $D^? \leftarrow$ False
end

structure tests to find the latent space with axes aligned with the types and severity of damage progression. As the damage class centers in this latent space have been diagonalized in Ψ, we can easily construct a likelihood ratio to test whether a queried set of embeddings belongs to a baseline condition set of embeddings [22]:

$$\begin{aligned}
R(\mathbf{u}^{baseline}_{1\ldots m}, \mathbf{u}^{query}_{1\ldots n}) &= \frac{\mathcal{P}(\mathbf{u}^{query} = \mathbf{u}^{baseline})}{\mathcal{P}(\mathbf{u}^{query} \neq \mathbf{u}^{baseline})} \\
R^{-1}(\mathbf{u}^{baseline}_{1\ldots m}, \mathbf{u}^{query}_{1\ldots n}) &= \frac{\mathcal{P}(\mathbf{u}^{query} \neq \mathbf{u}^{baseline})}{\mathcal{P}(\mathbf{u}^{query} = \mathbf{u}^{baseline})} \\
&= \frac{\mathcal{P}(\mathbf{u}^{baseline}_{1\ldots m})\mathcal{P}(\mathbf{u}^{query}_{1\ldots n})}{\mathcal{P}(\mathbf{u}^{baseline}_{1\ldots m}, \mathbf{u}^{query}_{1\ldots n})}
\end{aligned} \tag{10}$$

With $\mathcal{P}(\mathbf{u})$ as

$$\mathcal{P}(\mathbf{u}) = \prod_{t=1}^{d} \frac{\exp(-\frac{(\bar{u}_t)^2}{2(\psi_t+\frac{1}{n})} - \frac{\sum_{i=1}^{n}(u^i_t-\bar{u})^2}{2})}{(2\pi)^{n/2}(\psi_t+\frac{1}{n})^{n/2}} \tag{11}$$

Where ψ_t is the variance for the t'th feature along Ψ's diagonal, and $\bar{\mathbf{u}}_t$ is the mean over n at the t'th feature.

Note that for our purposes, we invert R to obtain R^{-1}, which is the likelihood ratio that the queried survey had changed from baseline (vs. the queried survey remained the same as baseline with R). As we want to detect a structural degradation, vs. identifying that a human speaker is the same as from his/her enrolled speech

recordings in a PLDA model, this change in perspective towards damage detection is more appropriate. Hence, given priors π_{diff}, π_{same} for whether u_{query} is in the same or a different class for $u_{baseline}$, the likelihood ratio $R^{-1} > \pi_{diff}/\pi_{same}$ indicates that u^{query} is different (i.e. a different damage state) from $u^{baseline}$. Even if an unseen damage mechanism for the source domain is present in the target domain, this hypothesis test between the query set u^{query} and the baseline set $u^{baseline}$ can still be performed so long as the latent space separating experimentally-observed damage is sufficient for discrimination.

VI. Results

We evaluate the approach for the following experiments:

1) a) Train TDNN on experimental structures, train PLDA on experimental structures, supervised classification of experimental structure damage Algorithm 2
   b) Train TDNN on experimental structures, train PLDA on Z24, supervised classification of Z24 damage Algorithm 2
2) Train TDNN on experimental structures, train PLDA on experimental structures, unsupervised comparison of Z24 baseline and damage cases Algorithm 4

We aim to illustrate the efficacy of the TDNN embeddings as damage-identifying features, for both in-domain and out-of-domain classification. 1.(a) demonstrates the capacity of the classifier to identify damage across the library of experimental structures. 1.(b) and 2. illustrate generalization of the TDNN trained on experimental data to detection and identification of damage within an in-service structure, where in 1.(b) a progression of potential damage cases are already available and in 2. a deviation from a reference baseline case is being assessed. Though we do not provide a completely unsupervised methodology for damage identification with *no* prior knowledge of the target structure, we illustrate potential approaches enabled by the TDNN embeddings towards robust and interpretable damage inspection.

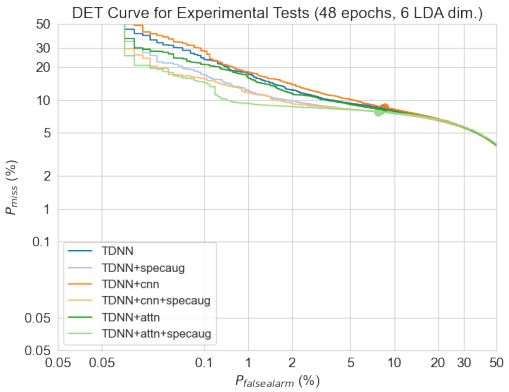
We train our TDNN for 48 epochs with a batchsize of 64 using the Kaldi Speech Recognition toolkit [39], and a learning rate schedule linearly decaying from 0.01 to 0.00001. We use a computer with an NVIDIA GeForce RTX 2080 TI GPU and an i7-6700K CPU for training the TDNN and PLDA, and inference with the PLDA. The PLDA training and inference is only done on the CPU, and is much shorter to run than training the TDNN.

A. Performance Metrics

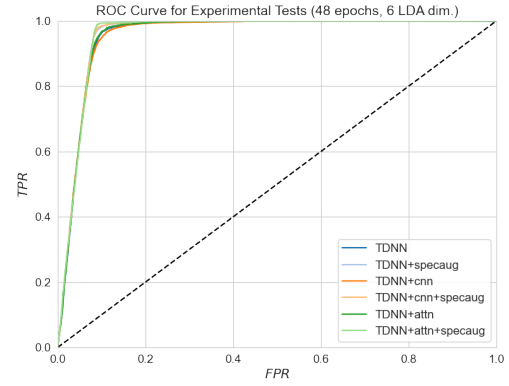
We use equal-error-rate (EER), area-under-the-curve (AUC), F1 score, and accuracy to assess performance of the classifier, similar to the metrics used by Tronci et al. [17], Snyder et al. [21]. We map the LLR scores from our TDNN-PLDA classifier (Equation (9)) for each accelerometer measurement to a one-hot vector indicating the ground truth target and nontarget classes. By varying the threshold for which a score is attributed to a class category, we can profile how well our classifier assigns an input feature set to its damage scenario from this ground-truth one-hot reference.

B. Supervised Damage Assessment

1) Experimental Structures: We first show the DET and ROC curves for the TDNN architectures on the test set of experimental structure measurements via supervised training of a PLDA backend in Figure 9. We observe that all TDNNs perform well on the test set, albeit with some measurements not being classified correctly. This is likely due to the overlap of our class labels in Table VIII and Table IX. Since damage testing is progressive for the majority of structures, the effect of previous labeled classes in the structural test are included in the progressively-damaged labels, which can confuse the TDNN classifier across severity of certain mechanisms. As the corpus of experimental structure measurements is very large, we will leave a diagnosis of the performance on experimental structures for future investigation. From the aforementioned shortcomings of constructing the experimental structure training set, we believe further labeling resolution or a contrastive training approach may mitigate these issues.



(a) DET for Z24



(b) ROC for Z24

Fig. 9: DET and ROC curves over TDNN architecture permutations for the test set of experimental structures, using PLDA supervised classification. Though not all cases from the training dataset are able to be classified correctly, we expect enough damage mechanisms are represented to apply towards unseen in-service damage.

We see in Table VI that the best-performing architecture is the fully-connected TDNN with spectral augmentation, followed closely by the TDNN with attention and spectral

TABLE VI: Experimental Supervised Classification Results

Architecture	EER (%)	AUC	Accuracy (%)	F1 Score (%)
TDNN	8.390	0.955	91.610	34.207
TDNN+specaug	**7.747**	**0.958**	**92.253**	**36.185**
TDNN+cnn	8.575	0.954	91.425	33.674
TDNN+cnn+specaug	7.814	0.958	92.186	35.971
TDNN+attn	8.193	0.956	91.807	34.794
TDNN+attn+specaug	7.785	0.958	92.215	36.064

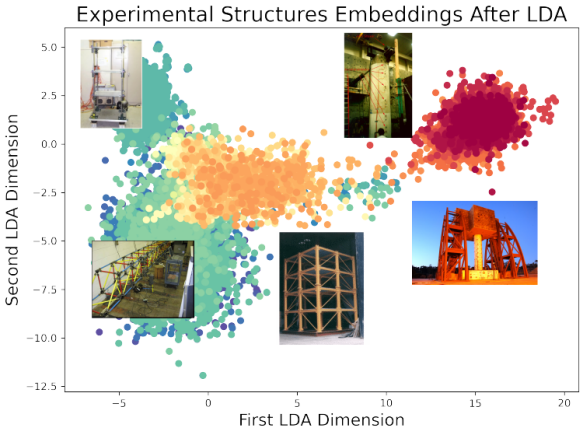
augmentation, then by the TDNN with convolutional layers and spectral augmentation. This validates the efficacy of spectral augmentation for structural vibrations, but appears to somewhat contradict the proposed utility of the convolution and attention layers. We will see that attention provides benefit in the generalization task to in-service structures.

2) Z24 Bridge: As the TDNNs are able to classify the large majority of experimental damage cases correctly, we apply the trained TDNNs as generators of DSFs, and train PLDA classifiers for each TDNN architecture for classifying Z24 damage cases. Inspection of the clustering obtained from LDA transformation in Figure 10 demonstrates the motivation for attempting to classify in the LDA latent space.

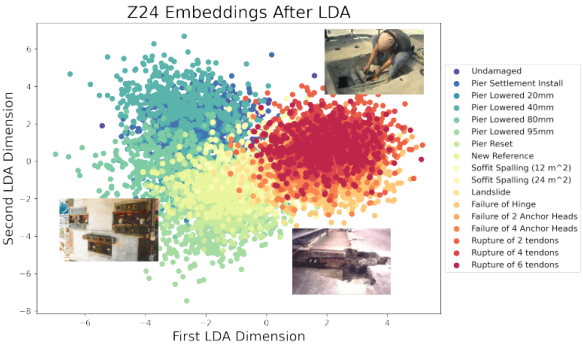
We show the DET and ROC curves for generalizing the TDNNs trained on experimental structures to the Z24 damage scenarios via supervised training of a PLDA backend in Figure 11 and Figure 12. We obtain strong classification performance for both forced and ambient vibration testing, with near-perfect assignment of each accelerometer on the Z24 bridge to their respective damage case. We observe that the TDNN with attention and spectral augmentation performs best in the Z24 damage multiple-classification task, generalizing well for both forced and ambient vibration in Table VII. This is likely due to the multiheaded-attention mechanism's ability to learn long-range patterns beyond the TDNN's immediate context, which can help keep track of unseen damage mechanisms over a long timespan. The strong performance for the unseen Z24 dataset provides empirical evidence that damage patterns learned from experimental data can be extended to an in-service structure, where if perfect knowledge of the damage scenarios is provided, the DSFs from the TDNN embeddings are able to strongly separate damage conditions.

To profile the amount of training and embedding LDA dimensionality on experimental, out-of-domain data needed to identify in-service Z24 damage cases, we plot EER over both number of training epochs and LDA dimension in Figure 13

In all cases, we observe that LDA dimensionality of 20-25 yields the best performance, and aligns with previous LDA dimensionality for the Z24 Bridge in Hom et al. [26], Tronci et al. [16, 17]. Some possibility of overfit is present the full training of the TDNNs, where we obtain better EER in epochs 10-18 vs. epoch 48 for the fully-connected, CNN with spectral augmentation, and attention variants of the TDNN architectures. However, as epoch 48 yielded the best validation results on our training experimental structures dataset, we use epoch 48 in all of our results.



(a) Separation of experimental structure testing along structure types and damage with first and second LDA dimensions

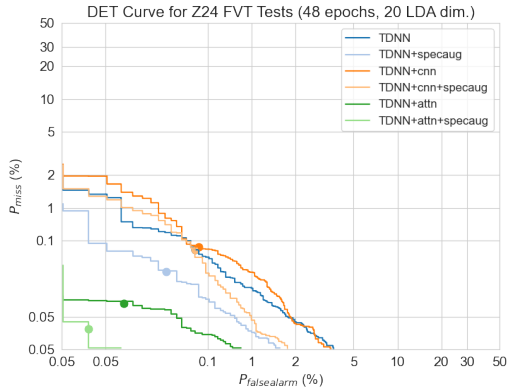


(b) Separation of Z24 damage scenarios along severity for the first LDA dimension and pier lowering for the second

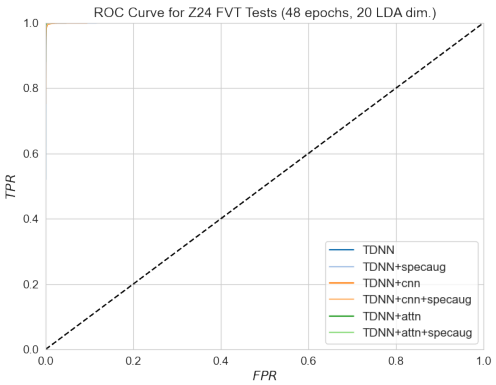
Fig. 10: Visualization of the LDA-transformed embeddings from the first two LDA dimensions with the largest eigenvalues for experimental structures embeddings in (a) and Z24 embeddings in (b), demonstrating separability of the embeddings. Separation appears to follow intuitive groupings of the damage scenarios and order of importance for discrimination.

TABLE VII: Z24 Supervised Classification Results

Architecture	Forced Vibration			
	EER (%)	AUC	Acc. (%)	F1 (%)
TDNN	0.267	1.000	99.616	99.617
TDNN+specaug	0.158	1.000	99.616	99.616
TDNN+cnn	0.333	1.000	99.437	99.438
TDNN+cnn+specaug	0.211	1.000	99.744	99.744
TDNN+attn	0.083	1.000	99.770	99.770
TDNN+attn+specaug	**0.011**	**1.000**	**99.872**	**99.872**
Architecture	Ambient Vibration			
	EER (%)	AUC	Acc. (%)	F1 (%)
TDNN	4.031	0.993	91.841	91.839
TDNN+specaug	3.495	0.995	93.913	93.913
TDNN+cnn	5.328	0.989	88.951	88.928
TDNN+cnn+specaug	3.881	0.993	93.018	93.000
TDNN+attn	3.036	0.996	94.731	94.733
TDNN+attn+specaug	**2.068**	**0.998**	**96.829**	**96.829**

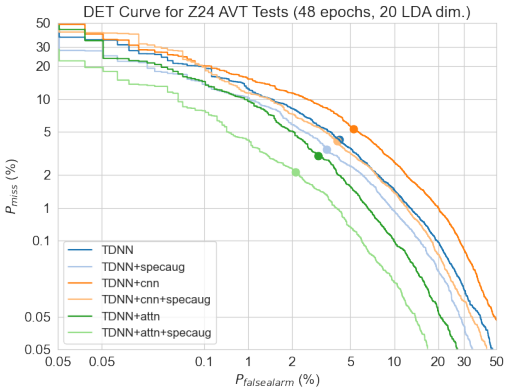


(a) DET for Z24 Forced Vibration

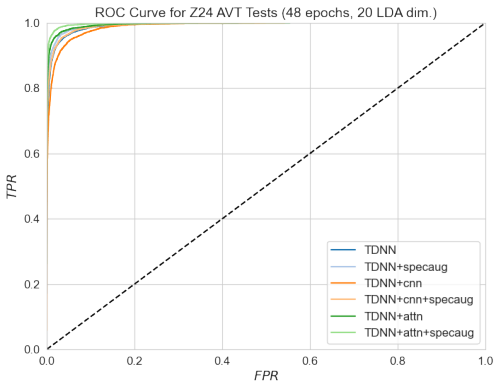


(b) ROC for Z24 Forced Vibration

Fig. 11: DET and ROC curves over TDNN architecture permutations for Z24 forced vibration, using PLDA supervised classification. Forced vibration testing is nearly perfectly identified between each damage case.



(a) DET for Z24 Ambient Vibration



(b) ROC for Z24 Ambient Vibration

Fig. 12: DET and ROC curves using PLDA supervised classification for Z24 ambient vibration. Ambient vibration testing suffers a small performance loss when compared to Z24 forced vibration.

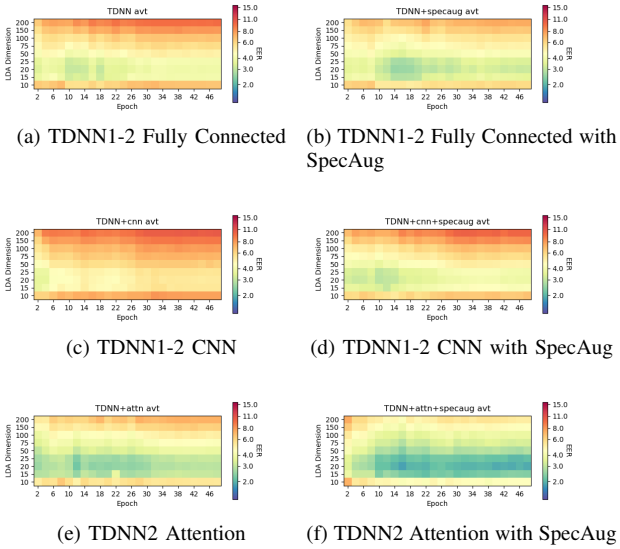
C. Unsupervised Damage Assessment

The above classification of existing damage cases is useful to profile a damage detector (and, operationally, assess significant changes in a structure's health condition from previously-recorded vibration history). However, we also are interested in the monitoring problem, where we have no immediate information about the change in damage state, but have a reference condition from which we want to ascertain a potential deviation against. A rudimentary assessment of a health monitor entails 1. detection of a change in structural condition 2. identification of the condition change as some damage mechanism. From a damage sensitive feature standpoint, a structural health monitor would discriminate against the magnitude and direction of the DSF in the damage latent space. We explore the efficacy of the TDNN embeddings in health monitoring for the Z24 bridge below.

1) Damage Detection via PLDA Hypothesis Testing: We leverage the hypothesis-testing framework from Section V-A3 to detect whether an in-service structure's condition has changed from one survey time to another. Within the PLDA latent space spanned *only* by experimental structures and/or the in-service structure's baseline (i.e. we only develop a PLDA on experimental and Z24 undamaged data), we score the in-service structure's TDNN embeddings from a given set of sensors over the reference and queried survey times, and answer the hypothesis that the queried survey belongs to the reference survey condition according to Algorithm 4. If the queried survey is more likely different from than similar to the reference survey, then a change in structural condition has occurred, and identification of damage can be performed.

Since our source domain spans multiple structures and damage conditions, we can perform an unsupervised damage assessment, where no information about the target domain is provided until test time, and a diagnosis of the structure's condition is made from an initial sample of the target structure's measurements and associated with the damage scenarios from our source domain. Due to the nature of SHM in practice, this sort of diagnosis is subject to local effects best-represented by constitutive methods or where location and geometry is explicit in the model (i.e. population-based SHM [40]). However, our approach assumes that damage is a global phenomenon which presents itself throughout the structure, and the consensus of sensors across the structure should capture this damage state's presence.

For brevity, we use the TDNN with attention and spectral augmentation as our embedding generator, as it

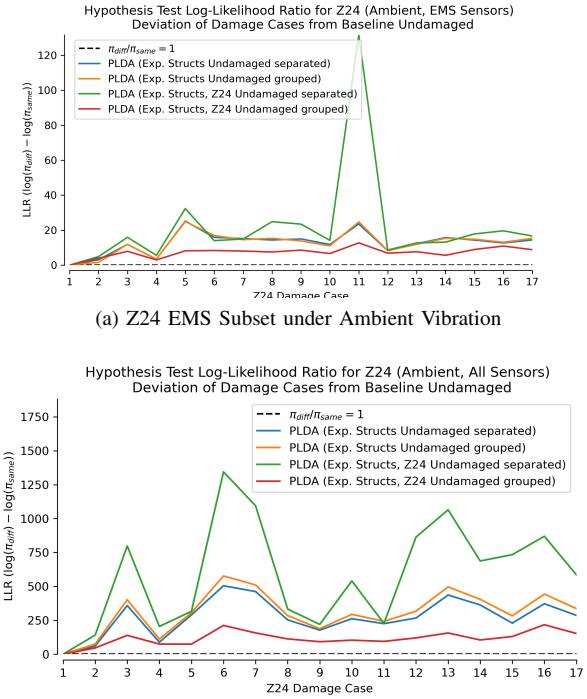


(a) TDNN1-2 Fully Connected (b) TDNN1-2 Fully Connected with SpecAug

(c) TDNN1-2 CNN (d) TDNN1-2 CNN with SpecAug

(e) TDNN2 Attention (f) TDNN2 Attention with SpecAug

Fig. 13: EER plotted over epochs vs. LDA dimension for PLDA classifiers trained to classify Z24 damage using TDNN embeddings trained on experimental structures. We can see that the best-possible generalization for the Z24 bridge occurs when the TDNN is trained 14-22 epochs on experimental data, with a LDA dimensionality of 20-25.



(a) Z24 EMS Subset under Ambient Vibration

(b) Z24 Full Sensor Set under Ambient Vibration

Fig. 14: PLDA Hypothesis Testing for Z24 Bridge damage cases against initial baseline, over all sensors and a sensor subset (at EMS sensor locations and orientations) using only ambient vibration. The only in-service structure information seen beforehand is the baseline survey, from which any future deviations would be considered damage.

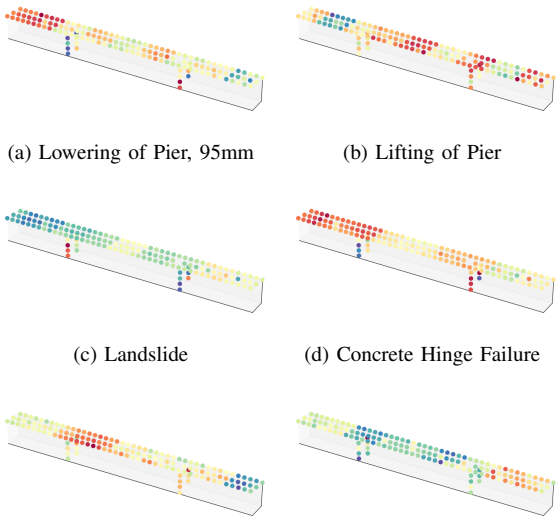
exhibits strong performance for both the experimental and Z24 classification task. In Figure 14, we show the efficacy of the hypothesis test for baseline versus damaged Z24 conditions, with and without calibrating the PLDA source-domain classifier on Z24 data (via adding the baseline undamaged Z24 condition as a class for PLDA training). Furthermore, we assess whether grouping together undamaged conditions across structures provide any benefit in the Z24 hypothesis test. We perform this test for Z24's undamaged baseline case against all 16 damage cases, using all sensors and the EMS subset defined in the previous section. We use log-likelihood ratio for convenience, to better distinguish the variations across experiments.

Given the assumption that $\pi_{diff} = \pi_{same}$ for each comparison of a queried survey of the Z24 Bridge, which coincides with an induced damage case, all queries are more likely to belong to the a damaged case than the baseline case, and so every damage case is detected as deviation from the baseline. We observe that adding the Z24 baseline to the PLDA classifier improves LLR separation of damage cases from the baseline, but grouping undamaged cases into one class does not increase performance.

Notably, in both the Z24 EMS and full sensor sets, the landslide damage case (case 11) is found to be highly separable from the baseline condition. As landslides are not in our experimental structure dataset, it appears unintuitive that such a damage scenario would be found to be particularly highly-separable from the baseline undamaged case. Tronci et al. [17] found that the TDNN-PLDA approach for damage detection is generally agnostic to changes in structural conditions unrelated to damage (i.e. temperature and humidity effects). Intuitively, the impact of a landslide is observable within the vibrational domain, and should also be observable in the damage domain. Though landslides are not observed in the experimental dataset, some composition of the enrolled damage mechanisms should span a space which captures the effect of the landslide. We will observe this effect in Figure 16.

For the full sensor set, pier lowering of 95mm and re-referencing (case 6-7), and concrete hinge failure, anchor cutting, and tendon snapping (case 13-17) are found to be very different from the baseline undamaged case. Except for the new reference condition, these damage mechanisms are likely very significant levels of damage for the Z24 bridge. However, as damage is generally an always-degrading state, and the re-referencing was solely the raising of the piers back to their nominal level (i.e. the fractures induced by the pier lowering were not repaired), it is likely that the new reference condition still carried the damage effects of the progressive pier lowering. We will see in Figure 16 and Figure 17 that this root-cause analysis bears out in comparison to the experimental structures modeled by the PLDA

As damage is a holistic assessment, ROC and DET curves based on per-sensor scores is less appropriate than in the supervised classification case (which served to demonstrate the maximum possible separation between Z24 damage classes using the TDNN-PLDA approach). Additional profiling of this monitoring approach will require assessment on continuous SHM instrumentation, such as the EMS portion of the Z24 Bridge Benchmark.



(a) Lowering of Pier, 95mm (b) Lifting of Pier

(c) Landslide (d) Concrete Hinge Failure

(e) Failure of 2 Anchor Heads (f) Tendon Rupture (2 of 16)

Fig. 15: Changes in Hypothesis Testing scores from damage scenario to scenario over Z24 vertical accelerometer locations, for ambient vibration and TDNN+attn+specaug.

Since we have a full array of accelerometers to assess, we can attempt damage localization through inspection. In Figure 15, we show the changes in hypothesis testing scores between damage cases over all vertical accelerometers for the Z24 ambient vibration testing. As documentation of where damage precisely was induced is not available, we must evaluate these local score distributions by inspection, to see whether the actual damage mechanism corresponds to the sensors contributing the most change in score.

- Pier lowering and raising scores (a,b) appear high at the appropriate column (rightmost), and score concentration on the span appears to track a higher-order bending mode.
- Landslide score concentration (c) appears to be on the Utzenstorf pier, and then on Koppigen (rightmost) span. Documentation does not state where the landslide was caused, but we would suspect either impact on the Utzendorf pier or under the Koppigen span.
- Concrete hinge failure is likely on the Koppigen side, which has some score concentration in (d), but it is possible that such extensive damage manifests in a lower-order bending mode creating a score distribution on the Utzendorf side.
- Anchor failure (e) appears to trigger a torsion mode, which corresponds to the reduction in cable tension on the left side of the span.
- Tendon snapping (f) appears to be correctly located at the Koppigen side of the span, where tendons were cut.

2) Damage Identification via PLDA Class Centers: We note that without direct control of the structural testing, attribution of damage to a particular mechanism is an inexact science, and can be highly colored by the detector and the availability of the mechanisms represented. However, we can still draw some useful inferences from associating Z24 damage with the observed mechanisms present in the experimental structures dataset.

In the PLDA framework, we model damage cases as normal distributions centered about a mean $\mathbf{m}$. Damage identification can be performed via association of the enrolled experimental damage case centers with the assessed in-service structural damage case centers in the PLDA latent space. This simple model of an in-service structure's damage state (e.g. the mean of the Z24 embeddings after LDA transformation) allows us to use simple geometric conceptualization for association with other damage mechanisms. If a change in center from the Z24 baseline to Z24 damage is aligned with one or more of the experimental damage case centers, we can attribute this shift to those experimental damage mechanisms. If two Z24 damage conditions are assessed against a baseline case, that change in alignment should be described by a change in mechanisms.

A simple metric to assess alignment of two vectors is the normalized dot product, or cosine distance:

$$\phi(\mathbf{a}, \mathbf{b}) = \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\|\|\mathbf{b}\|}$$

which we can use to find how much the change between Z24 conditions aligns within the experimental damage class clusters.

The change in Z24 baseline and damage PLDA centers $\bar{\mathbf{u}}_{Z24,b}$, $\bar{\mathbf{u}}_{Z24,d_1}$ (and experimental damage PLDA center $\bar{\mathbf{u}}_{exp,d_e}$ for experimental damage mechanism d_e in $\mathcal{D}_{exp}$) is given by the simple difference between two vectors. The cosine distance between these vectors (ϕ_{d_1}) gives us how much the changes originating at $\bar{\mathbf{u}}_{Z24,b}$ towards $\bar{\mathbf{u}}_{Z24,d_1}$ and $\bar{\mathbf{u}}_{exp,d_e}$ are aligned.

$$\begin{aligned}
\Delta\bar{\mathbf{u}}_{Z24,b}^{Z24,d_1} &= \bar{\mathbf{u}}_{Z24,d_1} - \bar{\mathbf{u}}_{Z24,b} \\
\Delta\bar{\mathbf{u}}_{Z24,b}^{exp,d_e} &= \bar{\mathbf{u}}_{exp,d_e} - \bar{\mathbf{u}}_{Z24,b} \\
\phi_{d_1} &= \phi(\Delta\bar{\mathbf{u}}_{Z24,b}^{Z24,d_1}, \Delta\bar{\mathbf{u}}_{Z24,b}^{exp,d_e})
\end{aligned}$$

The alignments for select damage cases are shown in Figure 16.

As damage cases are progressive, the effects of previous damage cases can make their cumulative effects difficult to distinguish. We can calculate the change in alignment $\Delta\phi$ to observe how much the evolution of a new Z24 damage case center deviates from the previous trajectory of Z24 damage:

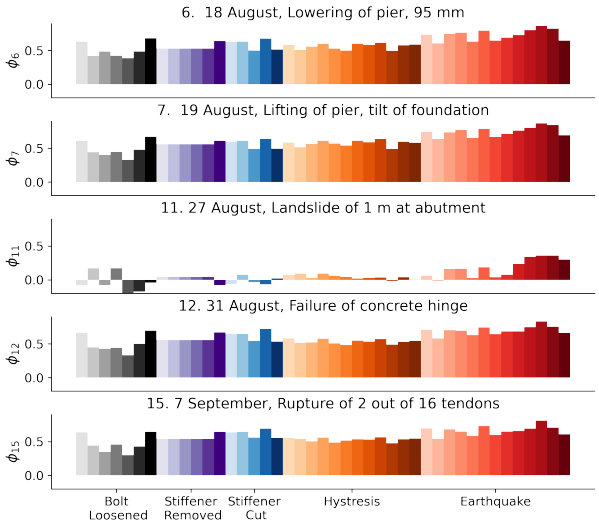
$$\begin{aligned}
\phi_{d_2} &= \phi(\Delta\bar{\mathbf{u}}_{Z24,b}^{Z24,d_2}, \Delta\bar{\mathbf{u}}_{Z24,b}^{exp,d_e}) \\
\Delta\phi_{d_1 \to d_2} &= \phi_{d_2} - \phi_{d_1}
\end{aligned}$$

Finally, we can calculate $\boldsymbol{\phi}, \Delta\boldsymbol{\phi}$ for all experimental damage mechanisms, which are vectors with size equal to the number of experimental damage cases:

$$\begin{aligned}
\boldsymbol{\phi}_{d_1} &= \phi(\Delta\bar{u}_{Z24,b}^{Z24,d_1}, \Delta\bar{u}_{Z24,b}^{exp,d_e}) \text{ for all } d_e \in \mathcal{D}_{exp} \\
\boldsymbol{\phi}_{d_2} &= \phi(\Delta\bar{u}_{Z24,b}^{Z24,d_2}, \Delta\bar{u}_{Z24,b}^{exp,d_e}) \text{ for all } d_e \in \mathcal{D}_{exp} \\
\Delta\boldsymbol{\phi}_{d_1 \to d_2} &= \boldsymbol{\phi}_{d_2} - \boldsymbol{\phi}_{d_1}
\end{aligned}$$

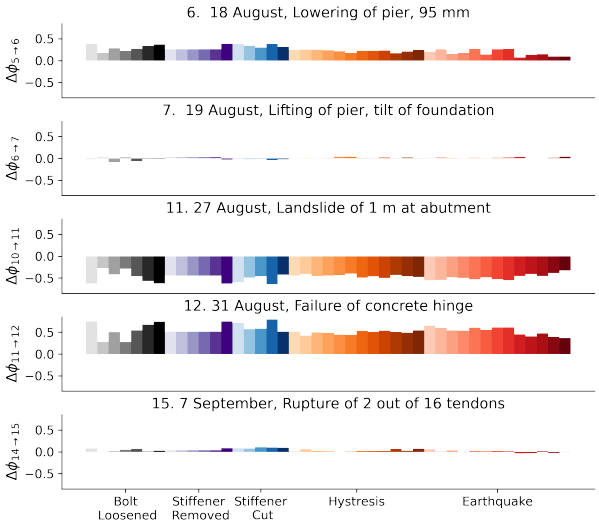
which we observe in Figure 17.

Recalling the previous hypothesis test results, we observe that the change in damage from the landslide case



(a) Cosine Distance (ϕ_d) for Z24 Damage Cases

Fig. 16: Cosine distance alignment (ϕ_d) of Z24 Bridge damage cases against initial Z24 baseline and damage ($\Delta\bar{\mathbf{u}}^{Z24,d}_{Z24,b}$) to experimental damage ($\Delta\bar{\mathbf{u}}^{exp,d_e}_{Z24,b}$). Except for case 11 (landslide), ϕ across all damage cases maintains some alignment with all damage clusters (as the bridge-damage dimension is likely a dense combination of all individual experimental damage components).



(a) Change in Cosine Distance ($\Delta\phi_{d_1\to d_2}$) over Z24 Damage Cases

Fig. 17: Change of alignment ($\Delta\phi_{d_1\to d_2}$) of Z24 Bridge damage cases against initial Z24 baseline and damage ($\Delta\bar{\mathbf{u}}^{Z24,d}_{Z24,b}$) to experimental damage ($\Delta\bar{\mathbf{u}}^{exp,d_e}_{Z24,b}$). We assess $\Delta\phi_{d_1\to d_2}$ with d_1 being the damage case immediately preceding the assessed damage case d_2. From case 7, we see minimal $\Delta\phi_{d_6\to d_7}$, indicating the pier lowering damage still persists even after pier raising. From case 11, we see that the landslide is unlike any of the experimental damage mechanisms as $\Delta\phi_{d_{10}\to d_{11}}$ is negative across all experimental damage mechanisms. From case 15, the tendon rupture shows $\Delta\phi_{d_{14}\to d_{15}}$ most attributable to stiffener cutting and bolt loosening.

and its predecessor case is almost entirely aligned away from all of the experimental damage mechanisms in Figure 17. This intuitively makes sense, as no particular damage scenario is descriptive of the landslide effect. For the change from pier lowering of 95mm to the lifting of pier and re-referencing, we see that few dimensions changed for the change in alignment, indicating that the pier lifting did not mitigate much in terms of damage.

The resulting interpretability is likely both dependent on source dataset and target structure, if we restrict ourselves to applying learned behaviors from sensor measurements. As there is little direct information on the precise locations or severity of damage, we must rely on engineering judgment to determine the accuracy of the attribution of an in-service structure's damage to the mechanisms identified in the training corpus.

VII. Conclusions

In this work, we examine the viability of learning damage-sensitive features from experimental structures for use in damage detection and identification. To the best of our knowledge, we use the broadest population of experimental structures to date as our training dataset. By training of a TDNN as a damage classifier of these structures' damage mechanisms, we generate damage-sensitive features via the embeddings from an intermediate pre-final layer. Generalization is investigated and achieved through a PLDA backend, which classifies the embeddings drawn from an in-service bridge for supervised damage classification, or assesses whether a queried survey condition has changed against a baseline condition via an unsupervised hypothesis test. We contribute a simple but novel application of the PLDA model for attribution of in-service damage evolution to our training experimental structure damage mechanisms, using the cosine-distance as a metric for alignment of PLDA centers. These approaches illustrate the applicability of learned damage mechanisms solely from experimental structures to inference on in-service structures.

We evaluate the performance of these approaches on the Z24 Bridge Benchmark. Given supervised a-priori labels, we achieve near-perfect separation across disparate damage cases over accelerometers from the Z24 Bridge. Without prior knowledge about the target structure, we can detect damage cases against the Z24 baseline undamaged condition via unsupervised hypothesis-testing with the PLDA spanning *only* experimental structures. Furthermore, the unsupervised hypothesis-testing can be made interpretable for Z24 condition changes within the experimental structure PLDA space, through cosine-distance alignment with the observed damage mechanism PLDA centers. Though these findings, we demonstrate (on a real in-service structure) the feasibility of the TDNN-PLDA approach for zero-shot learning and interpretable damage identification, and aim to enable greater explainability of damage assessments with minimal domain adaptation towards the target domain.

Several avenues to improve this TDNN-PLDA generalization approach present themselves. To account for damage progression and ordering, we look to investigate more-principled interpretation of enrolled damage mechanisms to account for sequence and severity of unseen damage,

and the nonlinear combinations thereof. Expansion of the experimental structures corpus with simulated data, such as from finite-element analysis and constitutive models of damage, will help broaden this approach's perspective into in-service structures and unseen/un-measured damage mechanisms. Evaluating more in-service structures for the hypothesis-testing and alignment interpretation approach will yield more insight into both efficacy and requisite additional damage mechanisms to better span the in-service damage space. Notably, no high-fidelity SHM test exists to track damage location and specific isolated mechanisms, which would address the inability to rigorously score our hypothesis-testing or cosine-distance damage attributions on the Z24 Bridge. Correspondingly, we seek to develop higher-resolution modeling capability to backtest attribution of damage mechanisms in a damage identifier's decisions.

VIII. Acknowledgements

We express our gratitude to the KU Leuven Structural Mechanics Section for providing the Z24 Bridge Benchmark, as the current best-tested and best-instrumented in-service SHM dataset. We thank Los Alamos National Laboratories for their SHM Dataset, and the Network for Earthquake Engineering Simulation at UCSB for their curated repository of SHM datasets.

Appendix

Table VIII and Table IX show our labeling for the SHM datasets. As some damage cases are shared across structures, we use the information provided in the provided documentation to aggregate damage cases together. Some engineering intuition is used to determine common damage case metrics Section II-A, such as ductility calculation from force-displacement curves and amount of stiffener cutting for different diameter steel tubing. Notably, most overlap occurs for concrete structures in Table IX, as 1. force-displacement curves are common in profiling a column's damage trajectory and 2. empirical validation of SHM techniques for concrete columns is very important for eventual application to in-service structures.

Label	Damage Type			Structures			
	Bolts Loosened	Stiffeners Removed	Stiffeners Cut	LANL Bookshelf	IASC-ASCE Building Benchmark	Purdue Truss	Urbana Champaign Truss
1	0			1	7		
2	0.25			2			
3	0.5			3			
4	0.75			4			
5	1			5			
6	2			6	9		
7	4				8		
8		0		1	1		1
9		0.5					2 to 14
10		1		7	5		
11		2		8	4, 6		
12		4			3		
13		8			2		
14			0			1	
15			0.25			2	
16			0.5			3	
17			1			4, 5	
18			2			6, 7	

TABLE VIII: Metal Structure Training Labels

Label	Damage Type		Structures						
	Ductility (Hysteresis)	Ductility (Earthquake)	LANL UCI Bridge Column (Cast)	LANL UCI Bridge Column (Shotcrete)	Reinforced Bridge Columns (Exp1)	Reinforced Bridge Columns (Exp2)	Seismic Bridge Columns	Reinforced Bridge Columns (Exp3, Exp5)	Reinforced Bridge Columns (Exp4, Exp6)
19	0		0	0	1 to 3	1			
20	0.5				3 to 6	2 to 5			
21	1		1	1	7 to 12	6 to 8			
22	1.5		2	2	13 to 15	9			
23	2				16 to 18	10 to 11			
24	2.5		3	3					
25	3				19 to 21	15			
26	4		4	4	22 to 24	17			
27	4.5				25 to 27				
28	5				28 to 30	20			
29	6				31 to 33				
30	7		5	5	34 to 36				
31		0					0	0	0
32		0.5					1	1, 2	1, 2
33		1						3, 4	3, 4
34		1.5					2	5	5
35		2					4	6	6
36		3						7	7
37		4					3	9	9
38		4.5						10	10
39		5						11	11
40		5.5					6		
41		6					7		
42		7					8, 9		
43		8.5					10		

TABLE IX: Concrete Structure Training Labels

References

[1] A. van den Oord, S. Dieleman, H. Zen, K. Simonyan, O. Vinyals, A. Graves, N. Kalchbrenner, A. Senior, and K. Kavukcuoglu, "Wavenet: A generative model for raw audio," in *Arxiv*, 2016. [Online]. Available: https://arxiv.org/abs/1609.03499

[2] Y. Zhang, W. Han, J. Qin, Y. Wang, A. Bapna, Z. Chen, N. Chen, B. Li, V. Axelrod, G. Wang, Z. Meng, K. Hu, A. Rosenberg, R. Prabhavalkar, D. S. Park, P. Haghani, J. Riesa, G. Perng, H. Soltau, T. Strohman, B. Ramabhadran, T. Sainath, P. Moreno, C.-C. Chiu, J. Schalkwyk, F. Beaufays, and Y. Wu, "Google usm: Scaling automatic speech recognition beyond 100 languages," 2023. [Online]. Available: https://arxiv.org/abs/2303.01037

[3] C. Wang, S. Chen, Y. Wu, Z. Zhang, L. Zhou, S. Liu, Z. Chen, Y. Liu, H. Wang, J. Li, L. He, S. Zhao, and F. Wei, "Neural codec language models are zero-shot text to speech synthesizers," 2023. [Online]. Available: https://arxiv.org/abs/2301.02111

[4] M. Azimi, A. D. Eslamlou, and G. Pekcan, "Data-driven structural health monitoring and damage detection through deep learning: State-of-the-art review," *Sensors*, vol. 20, no. 10, 2020. [Online]. Available: https://www.mdpi.com/1424-8220/20/10/2778

[5] A. Dadras Eslamlou and S. Huang, "Artificial-neural-network-based surrogate models for structural health monitoring of civil structures: a literature review," *Buildings*, vol. 12, no. 12, p. 2067, 2022.

[6] C. Farrar, P. Cornwell, S. Doebling, and M. Prime, "Structural health monitoring studies of the alamosa canyon and i-40 bridges," 07 2000.

[7] C. Farrar, D. Nix, T. Duffey, P. Cornwell, and G. Pardoen, "Damage identification with linear discriminant operators," 07 2003.

[8] E. Reynders and G. De Roeck, *Continuous Vibration Monitoring and Progressive Damage Testing on the Z24 Bridge*. John Wiley and Sons, 2009.

[9] S. Dyke, D. Bernal, J. Beck, and C. Ventura, "Experimental phase ii of the structural health monitoring benchmark problem," 07 2003.

[10] G. De Roeck, B. Peeters, and J. Maeck, "Dynamic monitoring of civil engineering structures," 2000.

[11] S. W. Doebling, C. R. Farrar, M. B. Prime, and D. W. Shevitz, "Damage identification and health monitoring of structural and mechanical systems from changes in their vibration characteristics: a literature review," Los Alamos National Lab., NM (United States), Tech. Rep., 1996.

[12] M. H. Soleimani-Babakamali, R. Soleimani-Babakamali, A. Kashfi-Yeganeh, K. Nasrollahzadeh, O. Avci, S. Kiranyaz, and E. Taciroglu, "Multi-source transfer learning for zero-shot structural damage detection," *Applied Soft Computing*, p. 112519, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1568494624012936

[13] M. H. Soleimani-Babakamali, R. Soleimani-Babakamali, K. Nasrollahzadeh, O. Avci, S. Kiranyaz, and E. Taciroglu, "Zero-shot transfer learning for structural health monitoring using generative adversarial networks and spectral mapping," *Mechanical Systems and Signal Processing*, vol. 198, p. 110404, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0888327023003114

[14] P. Gardner, X. Liu, and K. Worden, "On the application of domain adaptation in structural health monitoring," *Mechanical Systems and Signal Processing*, vol. 138, p. 106550, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S088832701930771X

[15] E. Figueiredo, M. O. Yano, S. da Silva, I. Moldovan, and M. A. Bud, "Transfer learning to enhance the damage detection performance in bridges when using numerical models," *Journal of Bridge Engineering*, vol. 28, no. 1, p. 04022134, 2023. [Online]. Available: https://ascelibrary.org/doi/abs/10.1061/%28ASCE%29BE.1943-5592.0001979

[16] E. M. Tronci, H. Beigi, M. Q. Feng, and R. Betti, "A transfer learning shm strategy for bridges enriched by the use of speaker recognition x-vectors," *Journal of Civil Structural Health Monitoring*, pp. 1285–1298, 2022.

[17] E. M. Tronci, H. Beigi, R. Betti, and M. Q. Feng, "A damage assessment methodology for structural systems using transfer learning from the audio domain," *Mechanical Systems and Signal Processing*, vol. 195, p. 110286, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0888327023001930

[18] A. H. Waibel, T. Hanazawa, G. E. Hinton, K. Shikano, and K. Lang, "Phoneme recognition using time-delay neural networks," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 37, pp. 328–339, 1989.

[19] D. Snyder, D. Garcia-Romero, D. Povey, and S. Khudanpur, "Deep neural network embeddings for text-independent speaker verification," in *Proc. Interspeech 2017*, 2017, pp. 999–1003. [Online]. Available: http://dx.doi.org/10.21437/Interspeech.2017-620

[20] D. Snyder, D. Garcia-Romero, G. Sell, D. Povey, and S. Khudanpur, "X-vectors: Robust dnn embeddings for speaker recognition," 04 2018, pp. 5329–5333.

[21] D. Snyder, D. Garcia-Romero, A. McCree, G. Sell, D. Povey, and S. Khudanpur, "Spoken language recognition using x-vectors," 06 2018, pp. 105–111.

[22] S. Ioffe, "Probabilistic linear discriminant analysis," in *Computer Vision – ECCV 2006*, A. Leonardis, H. Bischof, and A. Pinz, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 531–542.

[23] D. Garcia-Romero, X. Zhang, A. McCree, and D. Povey, "Improving speaker recognition performance in the domain adaptation challenge using deep neural networks," in *2014 IEEE Spoken Language Technology Workshop (SLT)*, 2014, pp. 378–383.

[24] R. Li, W. Zhang, and D. Chen, "The coral++ algorithm for unsupervised domain adaptation of speaker recognition," in *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal*

Processing (ICASSP), 2022, pp. 7172–7176.

[25] K. L. Hom, H. Beigi, and R. Betti, "Application of speaker recognition x-vectors to structural health monitoring," in *Model Validation and Uncertainty Quantification, Volume 3*, Z. Mao, Ed. Cham: Springer International Publishing, 2022, pp. 139–148.

[26] K. Hom, H. Beigi, and R. Betti, "An exploration of x-vectors for damage detection and identification," ser. ASME International Mechanical Engineering Congress and Exposition, vol. Volume 1: Acoustics, Vibration, and Phononics, 11 2021, p. V001T01A008. [Online]. Available: https://doi.org/10.1115/IMECE2021-73324

[27] E. Figueiredo, G. Park, J. Figueiras, C. Farrar, and K. Worden, "Structural health monitoring algorithm comparisons using standard data sets," United States, 2009.

[28] M. J. Schoettler, J. I. Restrepo, G. Guerrini, H. Gu, D. E. Duck, and F. Carrea, "A full-scale, single-column bridge bent tested by shake-table excitation," United States, 2012.

[29] Y. L. Mo, G. Song, Y. Moslehy, H. Gu, D. H. Sangers, and D. Belarbi, "Neesr payload: Damage detection of reinforced concrete columns subjected to combined actions," United States, 2011.

[30] Y. Gao and B. F. Spencer Jr., "Structural health monitoring strategies for smart sensor networks," United States, 2005.

[31] S. S. Krishnan, "Establishing a baseline damage index for reliable damage detection: Full scale validation," Ph.D. dissertation, 2012.

[32] J. Maeck and G. De Roeck, "Description of z24 benchmark," *Mechanical Systems and Signal Processing - MECH SYST SIGNAL PROCESS*, vol. 17, pp. 127–131, 01 2003.

[33] L. Balsamo, R. Betti, and H. Beigi, "A structural health monitoring strategy using cepstral features," *Journal of Sound and Vibration*, vol. 333, p. 4526–4542, 09 2014.

[34] H. Beigi, *Fundamentals of Speaker Recognition*. New York: Springer, 12 2011.

[35] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," 2023. [Online]. Available: https://arxiv.org/abs/1706.03762

[36] D. Povey, H. Hadian, P. Ghahremani, K. Li, and S. Khudanpur, "A time-restricted self-attention layer for asr," 04 2018, pp. 5874–5878.

[37] D. Povey, X. Zhang, and S. Khudanpur, "Parallel training of deep neural networks with natural gradient and parameter averaging," *CoRR*, vol. abs/1410.7455, 2015.

[38] B. Ghojogh, F. Karray, and M. Crowley, "Eigenvalue and generalized eigenvalue problems: Tutorial," 2023. [Online]. Available: https://arxiv.org/abs/1903.11240

[39] D. Povey, A. Ghoshal, G. Boulianne, L. Burget, O. Glembek, N. Goel, M. Hannemann, P. Motlicek, Y. Qian, P. Schwarz, J. Silovsky, G. Stemmer, and K. Vesely, "The kaldi speech recognition toolkit," in *IEEE 2011 Workshop on Automatic Speech Recognition and Understanding*. IEEE Signal Processing Society, Dec. 2011, iEEE Catalog No.: CFP11SRW-USB.

[40] L. Bull, P. Gardner, J. Gosliga, T. Rogers, N. Dervilis, E. Cross, E. Papatheou, A. Maguire, C. Campos, and K. Worden, "Foundations of population-based shm, part i: Homogeneous populations and forms," *Mechanical Systems and Signal Processing*, vol. 148, p. 107141, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0888327020305276