Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Near-Sensor Damage Localization in Multi-Story Buildings with Level-Crossing Spike Encoding and a Compact Spiking Neural Network

Saher Elsayed
University of Pennsylvania
Mohamed Ali
Texas A&M University
Khairi Azhar Aziz
khairi.azhar@uniten.edu.my

Universiti Tenaga Nasional

Article

Keywords: structural health monitoring, damage localization, spiking neural networks, level-crossing sampling, neuromorphic sensing, energy-efficient edge inference

Posted Date: July 2nd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10162479/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Near-Sensor Damage Localization in Multi-Story Buildings with Level-Crossing Spike Encoding and a Compact Spiking Neural Network

Saher Elsayed[1], Mohamed Ali[2], Khairi Azhar Aziz[3*]

[1]School of Engineering and Applied Science, University of Pennsylvania, Philadelphia, PA, USA.
[2]College of Architecture, Texas A&M University, College Station, TX, USA.
[3*]College of Computing and Informatics, Universiti Tenaga Nasional (UNITEN), Kajang, Selangor, Malaysia.

*Corresponding author(s). E-mail(s): khairi.azhar@uniten.edu.my;
Contributing authors: selsayed@seas.upenn.edu;
mohameda99@tamu.edu;

Abstract

Wireless, battery-powered accelerometer nodes are widely used for vibration-based structural health monitoring (SHM), but battery life limits the energy for sampling, on-node inference, and radio transmission. We ask whether an event-driven spike front-end can localize stiffness-degrading damage in a multi-story building while making compute cost scale with signal content. Our pipeline places a small modal filterbank before a level-crossing (temporal-contrast) encoder whose single threshold $\boldsymbol{\theta}$ jointly controls accuracy, synaptic-operation count, and the data volume leaving the sensor; a tiny feed-forward spiking neural network (SNN) trained with a surrogate gradient then classifies the spikes. On a six-class task for a five-story shear building under ambient excitation with random damage severity and noise, the SNN reaches **87.2%** accuracy, within two points of a dense 1D convolutional baseline (**89.2%**), at an estimated **49** versus **318** nJ per inference (operation counts at **45** nm). Sweeping $\boldsymbol{\theta}$ traces an accuracy-energy Pareto frontier spanning an order of magnitude, and ablations identify the filterbank as the key enabler. A tuned classical modal-feature classifier is the most accurate method here, so we frame the pipeline's value as event-driven, low-data-rate sensing at competitive accuracy rather than an accuracy gain. Code, trained weights, and seeds are publicly available.

Keywords: structural health monitoring, damage localization, spiking neural networks, level-crossing sampling, neuromorphic sensing, energy-efficient edge inference

1 Introduction

Continuous monitoring of civil structures aims to detect and localize damage from operational vibration data before it threatens safety or serviceability [1, 2]. Dense networks of wireless accelerometers make pervasive instrumentation practical, but each node must operate for years on a small battery or a harvested energy budget. In such nodes the dominant costs are the analog-to-digital conversion of a continuously sampled signal, any on-node inference, and the radio transmission of data or features to a base station. Reducing all three is therefore central to deployable SHM.

Data-driven methods, and in particular one-dimensional convolutional neural networks (1D-CNNs) applied directly to raw acceleration, have become a standard and accurate tool for vibration-based damage detection and localization [3–5]. These networks perform a fixed and dense amount of computation per window regardless of whether the structure is excited or quiescent, and they presuppose a densely sampled multi-channel input. For an always-on sensor, this couples energy to wall-clock time rather than to information.

Spiking neural networks (SNNs), the "third generation" of neural models [6], offer a different cost structure. A spiking neuron communicates with sparse binary events and performs an accumulate operation only when a presynaptic neuron fires, so its compute cost scales with activity rather than with network size or time [7]. Paired with an event-driven sensor front-end, the pipeline can in principle become *activity-proportional*: when the structure is quiescent, few events are produced and little energy is spent. Asynchronous level-crossing (also called send-on-delta or temporal-contrast) sampling is the natural front-end for this regime, emitting an event only when a signal changes by a fixed amount [8, 9].

Despite this conceptual fit, prior application of SNNs to SHM has been narrow. The closest works perform binary anomaly detection on a single structure and treat the spike encoding as a fixed implementation detail rather than a control variable [10–12]. To our knowledge, no prior work (i) performs multi-story damage *localization* from spike-encoded vibration, (ii) characterizes how the encoding threshold trades accuracy against compute cost and sensor data rate, or (iii) connects the resulting spike activity back to structural mechanics.

This paper addresses exactly that gap with a deliberately focused study. The goal is a clear, mechanistic account of one simulated shear-building setting rather than a broad, field-validated benchmark. Our contributions are:

1. A compact, event-driven pipeline for near-sensor damage localization: a small modal filterbank followed by level-crossing encoding and a tiny feed-forward SNN. The single threshold θ is shown to be one design knob that controls accuracy, compute cost, and data rate together (Section 3).

2. A systematic accuracy-energy characterization on a six-class localization task. The proposed SNN matches a standard dense 1D-CNN to within 2 points of accuracy at 6.5× lower estimated FP32 compute energy, and sweeping θ yields a Pareto frontier along which the estimated energy varies by an order of magnitude (Section 5).
3. Ablations that isolate *why* the pipeline works: the modal filterbank front-end is essential, recurrence is unnecessary and costly once that front-end is present, and a dense CNN given the same front-end is more accurate only at 18.6× the estimated energy.
4. A result linking the learned behavior to mechanics: damage produces a spatially structured change in per-floor spike activity that the classifier can use.
5. A transparent discussion, including strong INT8 CNN and classical modal-feature baselines, of the regimes in which event-driven inference helps and a clear statement of what this simulation study does not demonstrate. The implementation, trained weights, and random seeds are publicly available.

2 Related Work

2.1 Vibration-based SHM and deep learning

Classical SHM extracts damage-sensitive features such as modal frequencies, mode shapes, and their derivatives, and flags deviations from a healthy baseline [1, 2]. Stiffness-degrading damage shifts natural frequencies and distorts mode shapes in a spatially dependent way, which is the physical basis for localization. Deep learning, and 1D-CNNs in particular, removed the need for hand-crafted features by learning directly from raw acceleration, achieving real-time damage detection and localization on benchmark structures [3]; subsequent surveys document the breadth of these methods [4, 5]. Our baseline is a compact 1D-CNN in this tradition.

2.2 Spiking neural networks and neuromorphic hardware

SNNs compute with sparse, asynchronous events and are naturally matched to event-driven sensors and to neuromorphic processors that exploit activity sparsity for compute efficiency [7, 13]. The central training difficulty, the non-differentiable spike, is now routinely handled by surrogate gradients that replace the spike derivative with a smooth function during the backward pass [14–16]. Recurrent spiking architectures add temporal memory [17]. Dedicated digital neuromorphic accelerators report per-synaptic-operation energies in the low picojoule range [18], and FPGA frameworks make compact SNN inference deployable at the edge [19].

2.3 Event-driven and level-crossing sensing

Level-crossing sampling emits a signed event whenever a signal departs from a reference by a fixed quantum, concentrating sampling activity where the signal is changing [8]. The same temporal-contrast principle underlies neuromorphic vision sensors [9]. For vibration sensing this means that events, and therefore compute cost, are produced in proportion to structural motion.

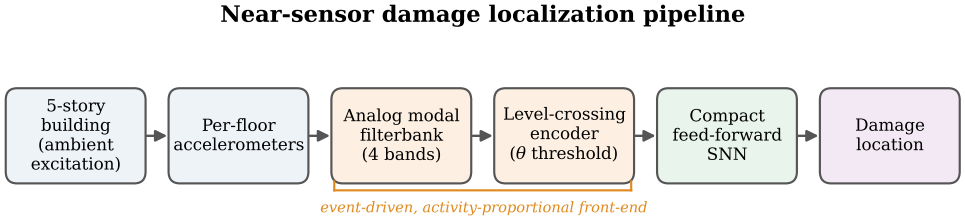


Fig. 1 Proposed near-sensor damage localization pipeline. The modal filterbank and the level-crossing converter form an event-driven front-end whose single threshold θ controls accuracy, compute cost, and data rate. Filterbank power, analog-to-digital conversion, and radio transmission are not included in the reported energy estimates.

2.4 SNNs for SHM

The most directly related works apply SNNs to SHM for near-sensor anomaly detection, using recurrent spiking models on a single instrumented structure with a fixed encoding [10–12]. Our work differs in three ways: it addresses multi-story *localization* rather than binary detection; it treats the encoding threshold as a first-class control variable that we characterize; and it connects spike activity to structural mechanics.

3 Methods

Figure 1 summarizes the pipeline. A structure is instrumented with one accelerometer per floor; each channel passes through a small modal filterbank; each sub-band is digitized by a level-crossing converter that emits spikes; and a compact feed-forward SNN maps the spike stream to a predicted damage location.

3.1 Structural model and damage scenarios

We model an $N = 5$ story building as a lumped-mass shear frame with floor masses m and inter-story lateral stiffnesses k_i. The equation of motion under base excitation is

$$\mathbf{M}\ddot{\mathbf{x}} + \mathbf{C}\dot{\mathbf{x}} + \mathbf{K}\mathbf{x} = -\mathbf{M}\mathbf{r}\,\ddot{u}_g(t), \tag{1}$$

where $\mathbf{x}$ are floor displacements relative to the ground, $\mathbf{r} = \mathbf{1}$ is the influence vector, and $\ddot{u}_g$ is the ground acceleration [20]. The mass matrix is diagonal and the stiffness matrix is tridiagonal in the inter-story stiffnesses. Damping is Rayleigh, proportional to $\mathbf{M}$ and $\mathbf{K}$ and matched to a modal ratio $\zeta = 0.02$ on the first two modes. Sensors measure absolute floor accelerations $\ddot{\mathbf{x}} + \mathbf{r}\,\ddot{u}_g$. Time histories are integrated with the unconditionally stable Newmark-β average-acceleration scheme [21].

Damage is modeled as a localized fractional reduction of one inter-story stiffness, the standard surrogate for cracking, connection loosening, or bracing loss [1]. The classification target has six classes: the healthy structure and damage localized at each of the five stories. Because stiffness loss at story i shifts the five natural frequencies

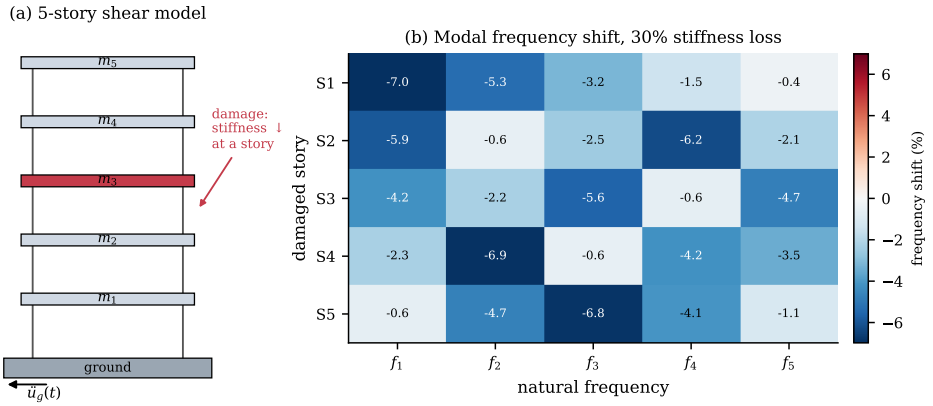


Fig. 2 (a) Five-story shear-building model with localized inter-story stiffness loss. (b) Percentage shift of each natural frequency $f_1, \ldots, f_5$ for a 30% stiffness loss at each story. The story-dependent pattern is the physical basis for localization.

and distorts the mode shapes in a story-dependent way, the location is in principle identifiable from the response. Figure 2(b) shows that a 30% stiffness loss produces a distinct frequency-shift fingerprint for each story.

Ambient excitation is band-limited stochastic ground acceleration (0.5 to 20 Hz), independently realized for every window so that no excitation is shared across samples. The sensor sampling rate is $f_s = 50$ Hz and each window spans $L = 128$ samples (2.56 s); the Nyquist limit of 25 Hz comfortably exceeds the highest mode (10.6 Hz). Damaged windows draw a severity uniformly in $[0.15, 0.40]$. Additive white measurement noise is applied per channel at a signal-to-noise ratio drawn uniformly in $[20, 40]$ dB.

3.2 Sensor front-end: modal filterbank and level-crossing encoding

Modal filterbank.

The discriminative information for localization lies in the structure's modal band. We therefore pass each accelerometer channel through a bank of $B = 4$ causal fourth-order band-pass filters, with pass-bands 0.8–3, 3–6, 6–8.5, and 8.5–14 Hz chosen to separate the five modes. This is the vibration analogue of a silicon cochlea: it exposes frequency content as distinct channels before spike encoding, with the relative amplitudes across floors carrying mode-shape information. The filterbank represents a realizable analog preprocessing stage, but its power, component tolerance, and phase response are not modeled here, and its contribution to node energy is excluded from all reported figures. A global (single-scale) normalization preserves these per-floor relative amplitudes. The N sensors thus produce $N \cdot B = 20$ sub-band channels.

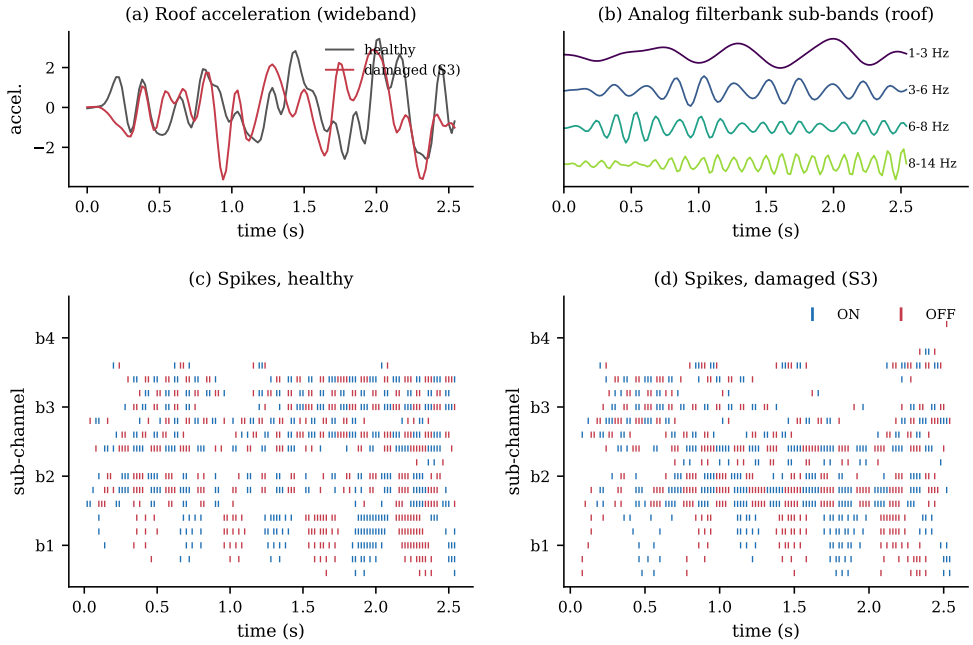


Fig. 3 Encoding pipeline on example windows. (a) Roof acceleration, healthy versus damage at story 3. (b) The four filterbank sub-bands for the roof sensor. (c, d) Level-crossing spike rasters (ON in blue, OFF in red) across the 20 sub-channels for a healthy and a damaged window at the operating threshold.

Level-crossing encoding.

Each sub-band channel is encoded by a level-crossing converter. For a normalized channel $x[t]$ with running reference R, an ON event is emitted and R is raised by θ whenever $x[t] - R \geq \theta$; an OFF event is emitted and R is lowered by θ whenever $x[t] - R \leq -\theta$. At most one event per polarity is emitted per step, which is hardware-realistic. This yields a binary spike tensor of width $2NB = 40$ (ON and OFF for each sub-band). The single threshold θ governs the trade-off the paper studies: a smaller θ produces more events, a finer representation, and higher compute cost; a larger θ produces fewer events, lower cost, and a coarser representation. Figure 3 illustrates the encoding on a healthy and a damaged window.

3.3 Spiking neural network classifier

The classifier is a deliberately small feed-forward SNN. The $2NB$ input spike channels project through a dense weight $\mathbf{W}_1$ to a layer of $H = 64$ leaky integrate-and-fire (LIF) neurons, which project through $\mathbf{W}_2$ to a leaky-integrator readout of $K = 6$ units. With subtractive reset, the hidden dynamics for a window of length T are

$$\mathbf{I}_1[t] = \mathbf{S}[t]\,\mathbf{W}_1 + \mathbf{b}_1, \tag{2}$$

$$\mathbf{U}_1[t] = \beta\,\mathbf{U}_1[t-1] + \mathbf{I}_1[t] - V_{\text{th}}\,\mathbf{s}_1[t-1], \tag{3}$$

$$\mathbf{s}_1[t] = \Theta(\mathbf{U}_1[t] - V_{\text{th}})\,, \tag{4}$$

where $\Theta(\cdot)$ is the Heaviside step, $\beta = 0.92$ is the membrane decay, and $V_{\text{th}} = 1$. The readout integrates its input current with the same leak, and the class scores are the time-averaged readout membrane potentials, $\mathbf{z} = \frac{1}{T}\sum_t \mathbf{U}_2[t]$, passed to a softmax cross-entropy loss. Time-averaging keeps the logits at a stable scale independent of T.

Training uses backpropagation through time with a fast-sigmoid surrogate gradient [14–16]: the derivative of the spike is replaced during the backward pass by $\big(1 + \text{slope}\,|\mathbf{U}_1 - V_{\text{th}}|\big)^{-2}$ with slope 10. The entire model, including the through-time gradient, is implemented from scratch in NumPy; its gradients are verified against numerical differentiation to a relative error below 10^{-9} for every parameter. A recurrent hidden variant, in which hidden spikes additionally feed back through a learned matrix $\mathbf{V}$, is implemented identically and used only as an ablation.

3.4 Energy and data-rate accounting

Scope of the energy estimates.

Every energy value in this paper is an operation-count estimate and should not be read as measured sensor-node or accelerator power. The estimates exclude analog filterbank power, sensing, analog-to-digital conversion, memory movement, control logic, and radio transmission, and they serve only to compare the algorithms relative to one another. We state this scope once here; for brevity the remainder of the paper writes "energy" for this estimated compute energy.

A dense network's cost is dominated by multiply-accumulate (MAC) operations, whereas a spiking network performs only accumulate (AC) operations, and only when a presynaptic neuron fires. Per-operation energies are taken from Horowitz at 45 nm CMOS, as reproduced by Sze et al.: $E_{\text{MAC}} = 4.6\,\text{pJ}$ and $E_{\text{AC}} = 0.9\,\text{pJ}$ in 32-bit floating point, with INT8 values of 0.2 pJ and 0.03 pJ respectively [22, 23]. The CNN energy is its (input-independent) MAC count times E_{MAC}. The SNN energy is its synaptic-operation (SOP) count times E_{AC}, where for the feed-forward model

$$\text{SOP} = (\text{input spikes}) \cdot H + (\text{hidden spikes}) \cdot K, \tag{5}$$

averaged over the test set. Because the SNN's advantage is driven by op-count sparsity, it is robust to the absolute precision assumption; we therefore report both FP32 and INT8 pairings. We also report a sensor data-rate reduction: the number of level-crossing events per window relative to dense Nyquist sampling of the five physical sensors, both as a token ratio and as a bit ratio (assuming 12-bit samples versus one polarity bit plus a sub-channel address per event). We make no claim of physical chip synthesis; a neuromorphic processor or an FPGA accelerator is the intended deployment target [18, 19].

4 Experimental Setup

The dataset contains 200 training, 40 validation, and 100 test windows per class (a total of 1,440 train, 240 validation, and 600 test windows), generated from disjoint

random-seed streams so that no excitation realization is shared across splits. The global normalizer is fit on healthy training windows only.

Deep learning baselines.

The 1D-CNN baseline operates on the five raw accelerometer channels after a single wideband (0.5–15 Hz) anti-alias filter and dense sampling; it has two convolutional blocks (8 and 16 channels, kernel 9, max-pool 4) followed by a fully connected classifier, following the established raw-acceleration approach of Abdeljaber et al. [3]. The proposed SNN operates on the filterbank level-crossing spikes. Both models are trained with Adam and identical budgets per family; the CNN uses a flat learning rate and the SNN uses cosine decay.

Classical SHM baselines.

We additionally include classical SHM baselines that extract, from each of the 20 filterbank sub-band channels, the root-mean-square energy and the spectral peak frequency, giving a 40-dimensional modal feature vector per window. These are classified with an RBF-kernel support vector machine (SVM) and a random forest (RF), each tuned by cross-validated grid search over its main hyperparameters. These baselines position the proposed SNN against conventional feature-based localization, not only against dense neural models, consistent with good practice in SHM evaluation [1, 2]. They use the same train, validation, and test splits, class labels, and random-seed partitioning as the neural models.

We report test accuracy, per-class confusion, estimated compute energy, data-rate reduction, robustness to noise and severity, and ablations. Both deep models are implemented and trained in the same framework-free codebase.

5 Results

5.1 Localization accuracy

The proposed feed-forward SNN at its operating threshold ($\theta = 0.6$) reaches 87.2% test accuracy, 2.0 points below the dense 1D-CNN baseline at 89.2%. A tuned classical modal-feature baseline is, in fact, a strong competitor on this task: an RBF support vector machine (SVM) on per-sub-band root-mean-square energy and spectral peak-frequency features reaches 94.3%, and a random forest (RF) reaches 86.3% (Table 1). That the SVM exceeds both deep models is unsurprising, since localization here is fundamentally a modal-discrimination problem that these hand-crafted features capture directly. This reframes the contribution of the proposed pipeline: its value is not higher accuracy but event-driven, activity-proportional operation on a sparse spike stream, whereas the SVM, the random forest, and the CNN all require the full densely sampled signal and centralized feature or inference computation. The SNN recovers most of the classical accuracy while reducing the sensor data rate, as quantified in Section 5 below. To confirm that the gap to the CNN is not an artifact of a single initialization, we retrained both deep models at the operating point from five random seeds: the SNN attains $86.6 \pm 1.0\%$ and the CNN $88.7 \pm 2.6\%$ (mean $\pm$ sample standard deviation). The 2.1-point mean gap lies well within the CNN's own run-to-run

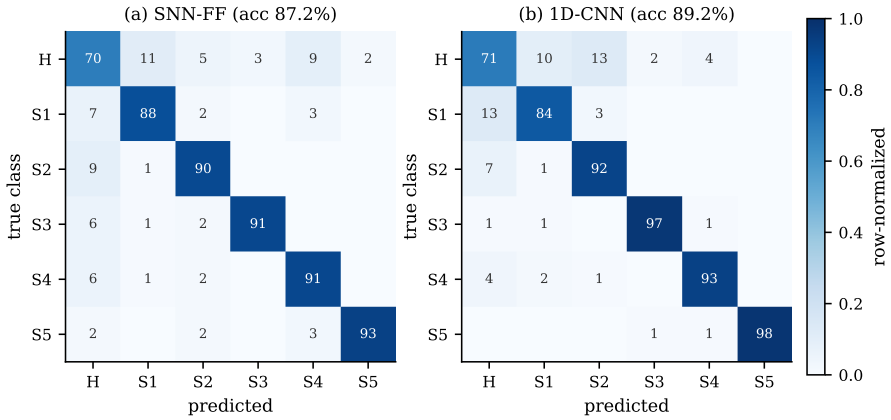


Fig. 4 Row-normalized confusion matrices on the 600-window test set. (a) Proposed feed-forward SNN. (b) Dense 1D-CNN baseline. H denotes healthy; S1–S5 denote damage at each story. Errors concentrate near the healthy/mild-damage boundary and between adjacent stories.

variability, and the SNN is in fact the more stable of the two. Figure 4 shows that both deep models confuse classes almost exclusively between the healthy state and mild damage, and between adjacent stories, which is consistent with the small modal changes such cases produce; clearly damaged cases are localized reliably.

5.2 Accuracy-energy Pareto frontier

The central result is Figure 5 and Table 1. Sweeping the single threshold θ traces a Pareto frontier: as θ increases, fewer spikes are produced, the estimated energy falls by an order of magnitude (from 92 to 9.5 nJ), and accuracy degrades gracefully from its peak. An intermediate θ is optimal, since too small a threshold over-encodes noise (more energy, lower accuracy) while too large a threshold discards signal. At the operating point the SNN uses an estimated 49 nJ per inference against the CNN's 318 nJ in FP32, a 6.5× reduction; at $\theta = 0.8$ it still reaches 86.0% at 35.5 nJ (8.9×). We note transparently that a quantized INT8 CNN is a strong baseline at an estimated 13.8 nJ; as discussed in Section 6, the proposed SNN's distinct advantages are its sparse event stream and activity-proportional operation count rather than a uniform compute-energy win across all regimes.

5.3 Ablations

Table 2 and Figure 6 isolate the sources of performance. Removing the modal filterbank and encoding the raw wideband signal collapses accuracy from 87.2% to 59.7%: the front-end, which exposes modal content as separate spike channels, is the key enabler. Adding recurrence to the hidden layer does not help once the filterbank is present (83.8% versus 87.2%) and adds 3.4× to the energy, because every hidden spike then drives an additional fan-out; for this task the temporal memory that recurrence provides is redundant with the explicit frequency channels. Finally, a dense CNN given

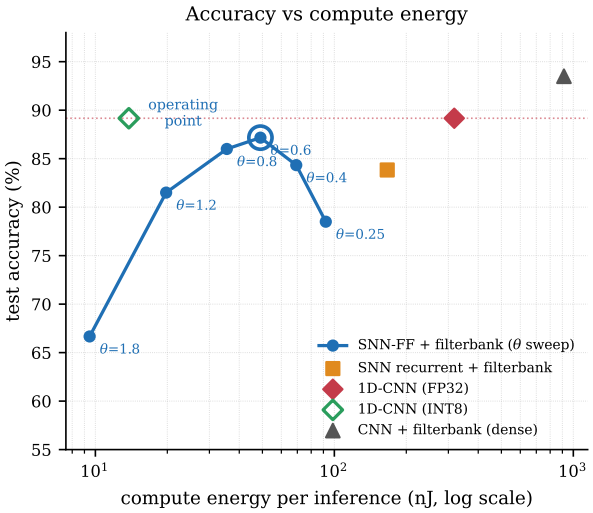


Fig. 5 Accuracy versus estimated compute energy per inference (log scale). The proposed feed-forward SNN with the filterbank front-end traces a Pareto frontier as the level-crossing threshold θ varies. Reference points show the FP32 and INT8 1D-CNN, the recurrent SNN ablation, and a dense CNN given the same filterbank front-end. All energy values are operation-count estimates; see Section 3.4 for scope.

Table 1 Main results. SNN variants use the modal filterbank front-end unless noted. Energy is estimated per inference at 45 nm using the Horowitz/Sze operation-count methodology and excludes filterbank power, sensing, memory, and radio. Data reduction is events versus dense Nyquist sampling of the five physical sensors (token / bit). The proposed operating point is in bold. †Classical baselines operate on the same modal-band features; their compute cost is negligible relative to the deep models and is not reported on the same energy scale.

Model	θ	Acc. (%)	MACs/SOPs	E_{FP32} (nJ)	E_{INT8} (nJ)	Data red.
Modal features + SVM†	–	94.3	–	–	–	–
Modal features + RF†	–	86.3	–	–	–	–
1D-CNN (raw, dense)	–	89.2	69,024 (MAC)	317.5	13.8	1.0×
SNN-FF (proposed)	**0.60**	**87.2**	**54,600 (SOP)**	**49.1**	**1.6**	**1.0/1.9×**
SNN-FF	0.80	86.0	39,400 (SOP)	35.5	1.2	1.4/2.8×
SNN-FF	1.20	81.5	22,000 (SOP)	19.8	0.7	2.8/5.5×
SNN-FF	1.80	66.7	10,600 (SOP)	9.5	0.3	6.8/13.7×

the same filterbank features reaches a higher 93.5%, confirming that the front-end carries strong information, but it does so at 914 nJ, 18.6× the proposed SNN's energy.

5.4 Robustness

Figure 7(a) varies the measurement SNR from 15 to 40 dB. The proposed SNN holds between 79% and 87% and tracks the CNN, which holds between 86% and 90%; the

Table 2 Ablations at the operating threshold. FB denotes the modal filterbank front-end. Energy values are operation-count estimates; see Section 3.4 for scope. The proposed configuration is in bold.

Configuration	Accuracy (%)	E_{FP32} (nJ)
SNN feed-forward, with FB (proposed)	**87.2**	**49.1**
SNN recurrent, with FB	83.8	166.5
SNN feed-forward, no FB (raw)	59.7	–
1D-CNN, raw wideband (baseline)	89.2	317.5
1D-CNN, with FB (dense control)	93.5	913.7

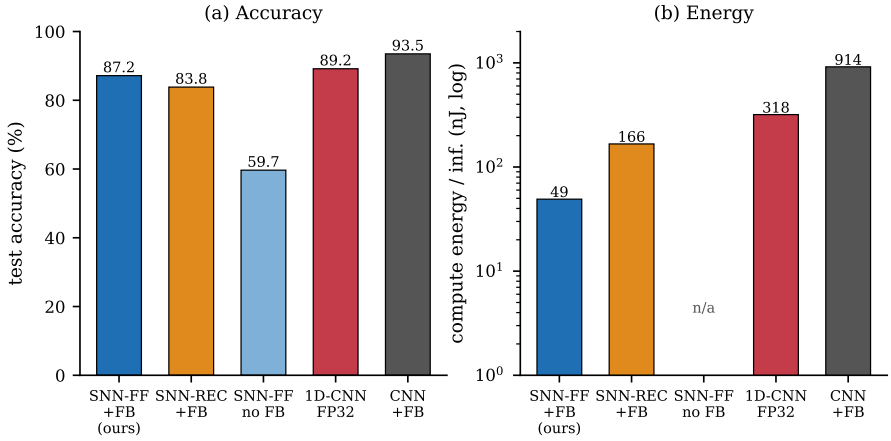


Fig. 6 Ablation summary. (a) Accuracy and (b) estimated compute energy (log scale). The filterbank front-end is essential for the SNN; recurrence is unnecessary and costly; the dense CNN on the filterbank is more accurate only at far higher energy.

level-crossing front-end degrades gracefully under noise. Figure 7(b) shows localization accuracy as a function of damage severity. As expected from mechanics, very mild damage (a 15–20% stiffness loss, which perturbs the modes only slightly) is the hardest case at 71%, while clearly developed damage is localized essentially perfectly, reaching 100% on the test subset for a 35–40% loss. The main limitation of the approach is therefore the detection-limited regime of incipient damage, which it shares with the dense baseline.

5.5 Interpretability

Figure 8 relates the learned representation to mechanics. We aggregate, per damage class, the mean number of spike events at each floor (summed over all sub-bands and polarities) and subtract the healthy baseline. Damage produces a spatially structured change in activity rather than a uniform one: the clearest effect is a localized increase in events near the damaged story against a background of reduced activity, with damage at the lowest story most strongly suppressing global activity, consistent with its

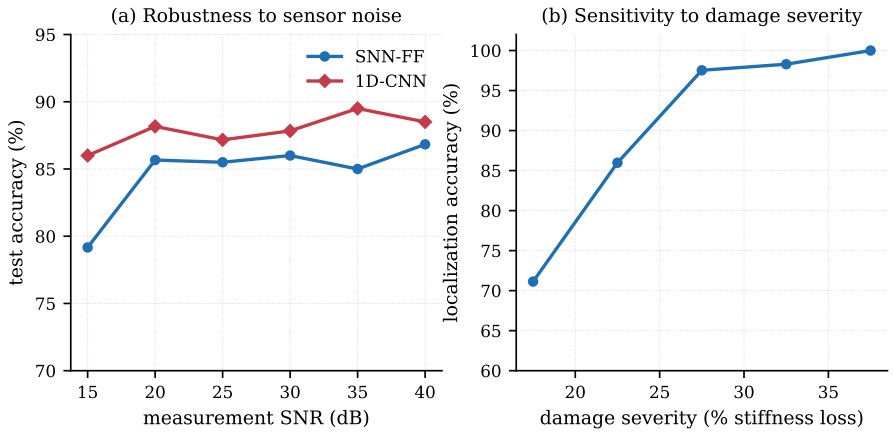


Fig. 7 Robustness. (a) Accuracy versus measurement SNR for the proposed SNN and the CNN. (b) Localization accuracy versus damage severity for the proposed SNN; incipient damage is the hardest case.

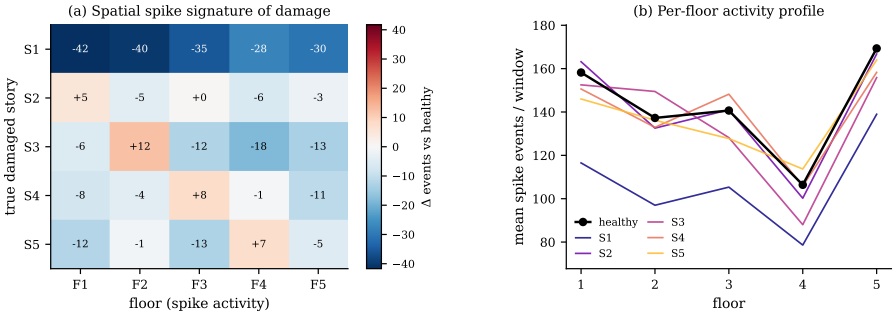


Fig. 8 Spatial spike signature of damage. (a) Change in mean per-floor spike events relative to healthy, for each damaged-story class. (b) Per-floor activity profiles. Damage shifts activity in a spatially structured, location-dependent way.

large effect on the fundamental mode. Because this signature follows the mechanics of story-level stiffness loss, it suggests that the classifier can rely on physically meaningful spatial activity patterns, providing a degree of transparency that pure black-box accuracy does not.

6 Discussion

When does the event-driven approach help?

Our results support a nuanced and deliberately humble answer. On accuracy, a tuned classical SVM on modal features is the strongest method here (94.3%), and its feature extraction is computationally light, so the deep models do not surpass it and the proposed SNN should not be read as an accuracy advance. The SNN's distinguishing

property is instead the sensing modality: every other method, classical or deep, consumes the full densely sampled multichannel signal, whereas the SNN operates directly on a sparse level-crossing event stream. Two consequences follow, both orthogonal to raw multiply-accumulate energy. First, the front-end emits far less data: at higher thresholds the volume leaving the sensor falls by up to 13.7× in bits, which matters where radio transmission, not local compute, dominates the energy budget, as is common in wireless SHM nodes. Second, the cost is activity-proportional: a dense pipeline spends the same energy on a quiescent structure as on an excited one, whereas the proposed pipeline spends energy only when informative motion produces events. For an always-on monitor that is quiescent most of the time, this proportionality, rather than peak accuracy or peak efficiency, is what governs battery life.

Why recurrence is unnecessary here.

A natural expectation is that recurrent spiking memory should help with temporal vibration data, and indeed prior SNN-SHM work is recurrent [10]. Our ablation shows the opposite once a modal filterbank is present: the filterbank already resolves the frequency content that distinguishes damage locations, so the network needs no learned temporal memory to recover it, and recurrence only adds synaptic operations. This is a useful negative result for designers of efficient SHM front-ends.

Limitations.

The evaluation is restricted to one simulated five-story shear frame and should be read as a controlled simulation study on that model, not as evidence of field-ready generalization. The method has not been validated on experimental SHM benchmarks such as IASC-ASCE or the Z24 bridge, nor on full-scale field data with temperature variability, operational variability, sensor drift, and changing boundary conditions; such validation is a necessary step before any deployment claim. Incipient damage remains hard, as Figure 7(b) shows. As stated in Section 3.4, all energy values are operation-count estimates rather than measured power, and the analog filterbank, while conceptually cheap, is an additional front-end component whose own power and tolerance must enter a full system budget. We do not claim real-time deployment; latency, memory footprint, and measured hardware energy remain open implementation tasks whose natural next step is a neuromorphic or FPGA realization [18, 19].

Outlook.

The single-knob trade-off we characterize suggests a path to adaptive monitoring: a node could raise θ during quiet periods to save energy and lower it when activity or a preliminary alarm warrants finer analysis, moving along the Pareto frontier at run time. Extending the approach to more realistic structures, to continuous severity estimation, to experimental benchmark validation on IASC-ASCE and Z24, and to on-device implementation are natural directions.

7 Conclusion

We presented a focused, controlled simulation study of near-sensor damage localization using a modal filterbank, level-crossing spike encoding, and a compact feed-forward spiking neural network. On a six-class localization task for a five-story building under band-limited stochastic ambient excitation, noise, and random severity, the proposed pipeline reached 87.2% accuracy, within two points of a standard dense 1D-CNN, while requiring 6.5× less estimated FP32 compute energy, and a single level-crossing threshold traced a clean accuracy-energy Pareto frontier spanning an order of magnitude. Ablations identified the filterbank front-end as the key enabler and recurrence as unnecessary and costly, and the per-floor spike activity carried a spatially structured signature consistent with the damaged story. We were explicit about where the approach does and does not win: a tuned classical modal-feature SVM is the most accurate method on this task (94.3%, with a random forest at 86.3%), and a quantized INT8 CNN is a strong compute-energy baseline, so the proposed pipeline's contribution is event-driven, low-data-rate sensing at competitive accuracy rather than accuracy or peak-efficiency superiority. Because all energy values are operation-count estimates, hardware validation and experimental benchmark testing remain important next steps. The work offers a transparent account of one idea at the intersection of edge AI and structural engineering, and points toward run-time-adaptive, communication-aware monitoring on neuromorphic and FPGA hardware.

Supplementary information. All structural simulation, encoding, network, training, and figure code is implemented from scratch in NumPy and SciPy, with no deep-learning framework dependency. Every gradient is verified against numerical differentiation. The implementation, trained weights, configuration files, random seeds, and metric scripts are available in the public repository linked under Data and code availability.

Acknowledgements. The authors gratefully acknowledge the support of Universiti Tenaga Nasional (UNITEN) under internal research funding, and thank colleagues who provided feedback on the neuromorphic-computing implementation.

Declarations

- **Funding.** This work was supported in part by Universiti Tenaga Nasional (UNITEN) under internal research funding.
- **Competing interests.** The authors declare no competing interests.
- **Author contributions.** S.E. conceived the study, designed and implemented the structural simulation, spike encoding, and learning pipeline, performed the experiments and analysis, prepared all figures, and wrote the manuscript. M.A. contributed to the structural and application framing and to manuscript review and editing. K.A.A. contributed to the study design and methodology, provided supervision and funding, and reviewed and edited the manuscript. All authors reviewed and approved the final manuscript.
- **Use of AI and AI-assisted technologies.** No artificial intelligence (AI) tools or AI-assisted technologies were used in the writing, editing, data generation, analysis,

coding, figure preparation, or bibliographic checking of this manuscript. All content, methods, results, code, figures, and references were prepared, reviewed, and verified by the authors, who take full responsibility for the integrity and accuracy of the work.

- **Data and code availability.** The complete implementation, including the structural simulation, spike encoding, the from-scratch SNN and CNN, the classical baselines, all configuration files and random seeds, and the scripts that reproduce every figure and reported metric, is available at https://github.com/saher-elsayed/snn-shm-localization.

References

[1] Farrar, C.R., Worden, K.: Structural Health Monitoring: A Machine Learning Perspective. John Wiley & Sons, Chichester, UK (2013). https://doi.org/10.1002/9781118443118

[2] Worden, K., Farrar, C.R., Manson, G., Park, G.: The fundamental axioms of structural health monitoring. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences **463**(2082), 1639–1664 (2007) https://doi.org/10.1098/rspa.2007.1834

[3] Abdeljaber, O., Avci, O., Kiranyaz, S., Gabbouj, M., Inman, D.J.: Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks. Journal of Sound and Vibration **388**, 154–170 (2017) https://doi.org/10.1016/j.jsv.2016.10.043

[4] Kiranyaz, S., Avci, O., Abdeljaber, O., Ince, T., Gabbouj, M., Inman, D.J.: 1d convolutional neural networks and applications: A survey. Mechanical Systems and Signal Processing **151**, 107398 (2021) https://doi.org/10.1016/j.ymssp.2020.107398

[5] Avci, O., Abdeljaber, O., Kiranyaz, S., Hussein, M., Gabbouj, M., Inman, D.J.: A review of vibration-based damage detection in civil structures: From traditional methods to machine learning and deep learning applications. Mechanical Systems and Signal Processing **147**, 107077 (2021) https://doi.org/10.1016/j.ymssp.2020.107077

[6] Maass, W.: Networks of spiking neurons: The third generation of neural network models. Neural Networks **10**(9), 1659–1671 (1997) https://doi.org/10.1016/S0893-6080(97)00011-7

[7] Roy, K., Jaiswal, A., Panda, P.: Towards spike-based machine intelligence with neuromorphic computing. Nature **575**(7784), 607–617 (2019) https://doi.org/10.1038/s41586-019-1677-2

[8] Miskowicz, M.: Send-on-delta concept: An event-based data reporting strategy. Sensors **6**(1), 49–63 (2006) https://doi.org/10.3390/s6010049

[9] Lichtsteiner, P., Posch, C., Delbruck, T.: A 128×128 120 db 15 μs latency asynchronous temporal contrast vision sensor. IEEE Journal of Solid-State Circuits **43**(2), 566–576 (2008) https://doi.org/10.1109/JSSC.2007.914337

[10] Barchi, F., Zanatta, L., Parisi, E., Burrello, A., Brunelli, D., Bartolini, A.,

Acquaviva, A.: Spiking neural network-based near-sensor computing for damage detection in structural health monitoring. Future Internet **13**(8), 219 (2021) https://doi.org/10.3390/fi13080219

[11] Zanatta, L., Barchi, F., Burrello, A., Bartolini, A., Brunelli, D., Acquaviva, A.: Damage detection in structural health monitoring with spiking neural networks. In: 2021 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd4.0&IoT), pp. 105–110 (2021). https://doi.org/10.1109/MetroInd4.0IoT51437.2021.9488476

[12] Joseph, G.V., Pakrashi, V.: Spiking neural networks for structural health monitoring. Sensors **22**(23), 9245 (2022) https://doi.org/10.3390/s22239245

[13] Davies, M., Srinivasa, N., Lin, T.-H., Chinya, G., Cao, Y., Choday, S.H., Dimou, G., Joshi, P., Imam, N., Jain, S., Liao, Y., Lin, C.-K., Lines, A., Liu, R., Mathaikutty, D., McCoy, S., Paul, A., Tse, J., Venkataramanan, G., Weng, Y.-H., Wild, A., Yang, Y., Wang, H.: Loihi: A neuromorphic manycore processor with on-chip learning. IEEE Micro **38**(1), 82–99 (2018) https://doi.org/10.1109/MM.2018.112130359

[14] Neftci, E.O., Mostafa, H., Zenke, F.: Surrogate gradient learning in spiking neural networks: Bringing the power of gradient-based optimization to spiking neural networks. IEEE Signal Processing Magazine **36**(6), 51–63 (2019) https://doi.org/10.1109/MSP.2019.2931595

[15] Zenke, F., Ganguli, S.: SuperSpike: Supervised learning in multilayer spiking neural networks. Neural Computation **30**(6), 1514–1541 (2018) https://doi.org/10.1162/neco_a_01086

[16] Eshraghian, J.K., Ward, M., Neftci, E.O., Wang, X., Lenz, G., Dwivedi, G., Bennamoun, M., Jeong, D.S., Lu, W.D.: Training spiking neural networks using lessons from deep learning. Proceedings of the IEEE **111**(9), 1016–1054 (2023) https://doi.org/10.1109/JPROC.2023.3308088

[17] Bellec, G., Salaj, D., Subramoney, A., Legenstein, R., Maass, W.: Long short-term memory and learning-to-learn in networks of spiking neurons. In: Advances in Neural Information Processing Systems (NeurIPS), vol. 31 (2018)

[18] Frenkel, C., Lefebvre, M., Legat, J.-D., Bol, D.: A 0.086-mm^2 12.7-pj/sop 64k-synapse 256-neuron online-learning digital spiking neuromorphic processor in 28-nm cmos. IEEE Transactions on Biomedical Circuits and Systems **13**(1), 145–158 (2019) https://doi.org/10.1109/TBCAS.2018.2880425

[19] Carpegna, A., Savino, A., Di Carlo, S.: Spiker+: A framework for the generation of efficient spiking neural networks FPGA accelerators for inference at the edge. IEEE Transactions on Emerging Topics in Computing **13**(3), 784–798 (2024) https://doi.org/10.1109/TETC.2024.3511676

[20] Chopra, A.K.: Dynamics of Structures: Theory and Applications to Earthquake Engineering, 5th edn. Pearson, Hoboken, NJ (2017)

[21] Newmark, N.M.: A method of computation for structural dynamics. Journal of the Engineering Mechanics Division, ASCE **85**(3), 67–94 (1959) https://doi.org/10.1061/JMCEA3.0000098

[22] Horowitz, M.: 1.1 computing's energy problem (and what we can do about it). In: 2014 IEEE International Solid-State Circuits Conference Digest of Technical

Papers (ISSCC), pp. 10–14 (2014). https://doi.org/10.1109/ISSCC.2014.6757323
[23] Sze, V., Chen, Y.-H., Yang, T.-J., Emer, J.S.: Efficient processing of deep neural networks: A tutorial and survey. Proceedings of the IEEE **105**(12), 2295–2329 (2017) https://doi.org/10.1109/JPROC.2017.2761740