# A Convolutional Autoencoder for Fast Compressive Sensing Reconstruction of Vibration Signals

Imen Tounsi[1], Fadi Karkafi[2], Mohammed El Badaoui[3], and François Guillet[4]

[1,2,3] *Safran Tech, Rue des Jeunes Bois – Châteaufort 78772 Magny–les–Hameaux, France*
*imen.tounsi@safrangroup.com*
*fadi.karkafi@safrangroup.com*
*mohammed.el-badaoui@safrangroup.com*

[1,3,4] *UJM-St-Etienne, LASPI, UR3059, F-42023, Saint-Etienne, France*
*guillet@univ-st-etienne.fr*

## ABSTRACT

In many health monitoring applications, large volumes of high-frequency measurement data must be acquired and processed to extract reliable health indicators for fault detection and identification. Compressive sensing (CS) provides an effective framework to reduce data dimensionality at the acquisition stage by exploiting signal sparsity, enabling sub-Nyquist sampling and lowering storage and transmission requirements. However, practical CS deployment is often limited by the reconstruction step, which typically relies on iterative optimization algorithms that are computationally expensive and difficult to implement in real-time monitoring systems. This work proposes a learned reconstruction strategy that replaces conventional CS solvers with a convolutional autoencoder based approach. The sensing process follows the standard CS formulation, where the original signal is projected onto a lower-dimensional measurement space using a fixed random sensing matrix. During training, the autoencoder is constrained so that its encoder reproduces the measurement operation, while the decoder learns a data-driven inverse mapping to reconstruct the original signal from compressed measurements. At inference time, compressed measurements are directly fed into the decoder, eliminating iterative reconstruction. Experimental results obtained on simulated gearbox signals and real vibration measurements demonstrate that the proposed method significantly reduces reconstruction time compared with classical CS algorithms while preserving diagnostically relevant information for fault detection.



## 1. INTRODUCTION

Vibration monitoring plays an important role in the health assessment and fault diagnosis of mechanical and structural systems (Braun, 1988). Modern sensing technologies enable continuous, high-frequency measurements in applications such as rotating machinery, turbines, and aerospace components, providing rich datasets that capture essential diagnostic information (Karkafi et al., 2024). However, capturing these signals often requires high sampling rates, resulting in large volumes of data. This creates significant challenges in terms of data storage, transmission, and real-time processing, particularly in long-term monitoring and distributed sensing applications.

Compressive sensing (CS) addresses this issue by enabling signal acquisition below the Nyquist sampling rate (Donoho, 2006). Instead of sampling the full signal, CS collects a reduced set of compressed measurements through linear projections. By exploiting signal sparsity in a suitable transform domain, the original signal can be reconstructed while reducing data acquisition and transmission requirements. However, a key challenge lies in the reconstruction stage. Classical algorithms such as Orthogonal Matching Pursuit (OMP) (Tropp & Gilbert, 2007), Matching Pursuit (MP) (Mallat & Zhang, 1993), and Compressive Sampling Matching Pursuit (CoSaMP) (Needell & Tropp, 2009) recover signals by solving optimization problems through iterative procedures involving repeated matrix multiplications and thresholding. As a result, reconstruction can be computationally expensive and time-consuming, limiting the use of classical CS methods in real-time vibration monitoring.

In recent years, deep learning has been increasingly explored as an alternative approach for compressive sensing reconstruction (Machidon & Pejovic, 2023). Instead of solving an optimization problem for each signal, neural networks can learn a direct mapping from compressed measurements to

the original signal using training data, enabling significantly faster reconstruction. Several studies have demonstrated the effectiveness of deep learning for this task. For example, ReconNet is an early convolutional neural network designed for fast reconstruction of compressively sensed images (Kulkarni et al., 2016) (Kulkarni et al., 2016). Metzler et al. (Metzler et al., 2017) introduced LDAMP, which integrates deep learning with the denoising based approximate message passing algorithm, while Zhang and Ghanem (Zhang & Ghanem, 2018) proposed ISTA-Net by unfolding the classical ISTA algorithm into a trainable deep architecture. Similarly, Yang et al. (Yang et al., 2016) (Yang et al., 2020) developed ADMM-Net and ADMM-CSNet by incorporating the ADMM optimization framework into deep neural networks. Autoencoder based approaches have also been explored to model compressive sensing as an encoder–decoder process. For instance, Mousavi et al. (Mousavi et al., 2015) proposed a stacked denoising autoencoder for CS reconstruction, demonstrating that neural networks can efficiently recover signals from compressed measurements while significantly reducing reconstruction time.

However, existing deep learning based reconstruction methods still present several limitations. Many approaches treat the reconstruction problem as a purely data-driven mapping from compressed measurements to signals, without explicitly enforcing consistency with the measurement process. As a result, the reconstructed signals may deviate from the physical measurement model. In addition, some methods rely on complex architectures such as generative adversarial networks (GANs) (Machidon & Pejovic, 2023), which can suffer from unstable training and require large amounts of training data. Furthermore, certain deep learning frameworks involve multi-stage reconstruction pipelines or preprocessing steps, increasing model complexity and reducing deployment efficiency. Moreover, existing deep compressed sensing methods often exhibit limited generalization when trained on simulated datasets and evaluated on real vibration signals.

To address the limitations of existing approaches, this paper proposes a measurement-consistent reconstruction framework based on a convolutional autoencoder. During training, the autoencoder is optimized not only to reconstruct vibration signals but also to encourage the encoder to approximate the measurement process, while the decoder learns a corresponding reconstruction mapping, effectively embedding the sensing mechanism into the network. This is achieved through a joint loss function that encourages the latent representation to match the compressed measurements. During deployment, only the decoder is used to reconstruct the signal directly from the compressed measurements, eliminating iterative reconstruction and reducing computational complexity. As a result, the proposed approach enables fast and efficient signal recovery, making it well suited for real-time monitoring applications. Furthermore, the method is comprehensively evaluated on both simulated and real gearbox vibration datasets, demonstrating its effectiveness in terms of reconstruction accuracy, computational efficiency, and preservation of diagnostic information.

## 2. THEORETICAL BACKGROUND

### 2.1. Compressive sensing model

Compressive Sensing (CS) offers a robust theoretical and computational framework for efficient signal acquisition and reconstruction. In contrast to conventional Nyquist-rate sampling, CS exploits the inherent sparsity of signals in a suitable transform domain, enabling accurate recovery from a significantly reduced number of linear measurements. The reconstruction stage is formulated as a constrained optimization problem in which sparsity is enforced as a regularization criterion while simultaneously maintaining fidelity to the acquired measurements (Brunton & Kutz, 2019). As illustrated in Figure 1, the CS framework consists of three principal components: sparse representation $s$ of the signal in an appropriate basis, compressed measurements $y$ obtained through a linear sensing process, and signal reconstruction, achieved by solving a sparsity-constrained inverse problem.

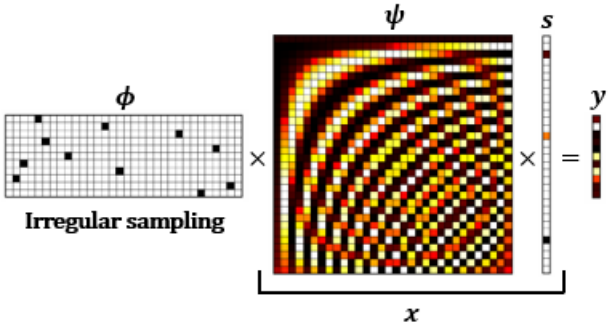


Figure 1. Schematic diagram of the compressive sensing algorithm.

A signal $x \in \mathbb{R}^N$ is said to be $k$-sparse if it can be represented as:

$$x = \psi s \tag{1}$$

where $\psi \in \mathbb{R}^{N \times N}$ is a sparsifying basis and $s$ is the sparse representation of the original signal that contains only $k$ nonzero coefficients in the selected transform domain. CS reduces the number of required samples by acquiring a compact set of linear projections that preserve the essential information of the original signal. Instead of sampling at the Nyquist rate, CS employs a properly designed sensing matrix $\phi \in \mathbb{R}^{M \times N}$, where $M \ll N$, to directly obtain compressed measurements. Here, $M$ denotes the number of acquired measurements. The compressed measurement vector $y$ is then obtained through linear projection:

$$y = \phi x = \phi \psi s \tag{2}$$

The compression ratio (CR) represents the percentage of

measurements retained after compression and is defined as:

$$CR = \frac{M}{N} \times 100 \tag{3}$$

Given the compressed measurement vector $y$, the objective is to reconstruct the original signal $x$ from an underdetermined linear system. Accordingly, the reconstruction problem is formulated as the optimization problem in (4), which seeks a solution that satisfies the measurement constraints while enforcing sparsity priors on the signal.

$$\min_{s} \|s\|_1 \quad \text{s.t.} \quad \boldsymbol{\phi\psi} s = y \tag{4}$$

The recovered signal is then obtained as: $\hat{x} = \boldsymbol{\psi} s$.
Signal reconstruction from undersampled measurements is typically addressed using convex optimization, greedy methods, or non-convex minimization algorithms. A common drawback of these approaches is their high computational cost, as they rely on iterative calculations. For CS to be practical in real-world ubiquitous computing applications, faster and more efficient reconstruction methods are required. Deep learning has emerged as a promising alternative. Although training deep networks demands substantial computational resources and memory due to the large number of parameters, most of this cost occurs during the training phase. Once trained, inference is significantly faster than conventional iterative reconstruction algorithms.

### 2.2. Deep learning for compressive sensing reconstruction

Deep learning has emerged as an efficient approach for compressive sensing reconstruction by learning a direct mapping from compressed measurements to the original signal (Machidon & Pejovic, 2023). Unlike traditional reconstruction methods that rely on iterative optimization, neural networks approximate the inverse mapping directly from data.
Given the vibration signal $x$, the compressed measurements are obtained using the measurement matrix $\Phi$, as expressed in (2). Then, instead of solving a sparsity-constrained optimization problem for each signal instance, a parametric nonlinear mapping is learned to reconstruct $x$ directly from its compressed representation $y$. Specifically, the reconstruction problem can be formulated as learning a nonlinear function $f_W(\cdot)$, parameterized by network weights $W$, such that

$$\hat{x} = f_W(y) \tag{5}$$

where $\hat{x}$ denotes the reconstructed signal.
To learn this mapping, the network parameters are optimized by minimizing a reconstruction loss defined as

$$\mathcal{L}_{\text{rec}} = \|x - \hat{x}\|_2^2 \tag{6}$$

During training, this loss is minimized over a dataset of training samples. Once training is completed, the learned model can be used for signal reconstruction during inference. In this phase, the compressed measurements are directly provided as input to the trained network, which reconstructs the signal through a single forward pass. This non-iterative reconstruction process eliminates the need for optimization based solvers and significantly reduces the computational cost.

## 3. PROPOSED METHOD

Following the direct-mapping approach described in the previous section, a data-driven reconstruction framework is proposed to learn a nonlinear mapping from compressed measurements to the original signal.
As illustrated in Figure 2, the framework consists of two distinct phases:

1. Training Phase: an autoencoder is trained to reconstruct the original signal while encouraging alignment between the latent space and the compressed measurements.

2. Testing Phase: the encoder is discarded, and only the trained decoder is used to reconstruct signals directly from real compressed measurements.

The proposed model is based on an autoencoder structure composed of: an encoder $E_\theta(\cdot)$ and a decoder $D_W(\cdot)$, where $\theta$ and $W$ denote the trainable parameters. The encoder maps the original signal $x$ to a latent representation: $z = E_\theta(x)$, where $z \in \mathbb{R}^M$. To ensure physical consistency with the compressive sensing acquisition process, the latent representation is encouraged to match the compressed measurement:

$$z \approx y = \boldsymbol{\phi} x. \tag{7}$$

Finally, the decoder reconstructs the signal from the latent representation: $\hat{x} = D_W(z)$.
The network is trained using a composite loss function composed of two terms. The first term represents the reconstruction loss, defined in (6), which ensures signal reconstruction. The second term is the measurement consistency loss, given by

$$\mathcal{L}_{\text{meas}} = \|E_\theta(x) - \boldsymbol{\phi} x\|_2^2 \tag{8}$$

which encourages alignment between the latent representation and the compressed measurements.
The overall training objective is therefore formulated as

$$\mathcal{L} = \mathcal{L}_{\text{rec}} + \lambda \mathcal{L}_{\text{meas}} \tag{9}$$

where $\lambda$ controls the trade-off between reconstruction accuracy and measurement consistency. This training strategy encourages the decoder to learn a stable reconstruction mapping that remains consistent with the measurement process.

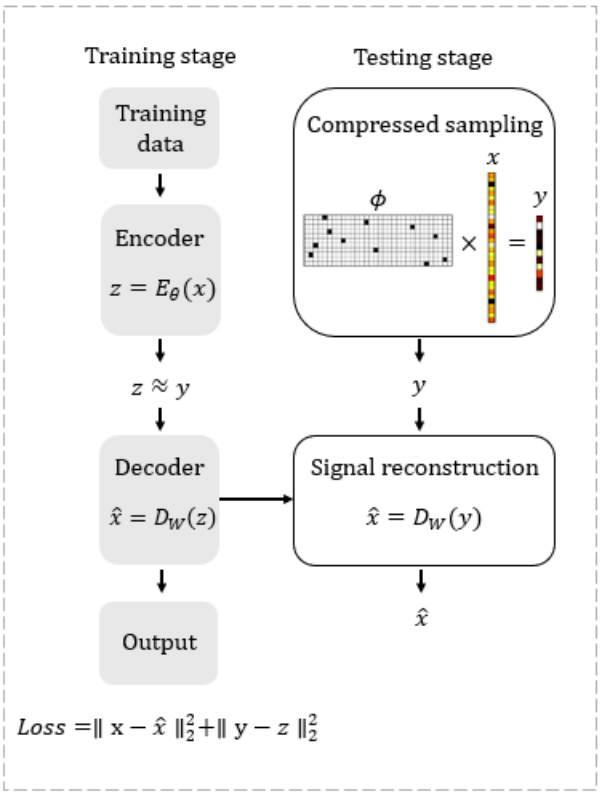


Figure 2. Overview of the proposed framework for compressed sensing acquisition and signal reconstruction during training and testing.

## 4. EXPERIMENTAL EVALUATION AND PERFORMANCE ANALYSIS

### 4.1. Dataset description

To train the proposed model, a dataset of simulated vibration signals was generated based on an empirical dynamic model of a gearbox. Each signal is modeled as an amplitude-modulated sinusoidal carrier representing gear meshing, combined with lower-frequency modulation components associated with shaft dynamics:

$$s(t) = s_1(t)\left(1 + s_2(t) + s_3(t)\right) \tag{10}$$

where $s_1(t)$ denotes the high-frequency carrier component (gear meshing), $s_2(t)$ represents the modulation due to the first gear, and $s_3(t)$ represents the modulation due to the second gear.

Each component $s_i(t)$ is modeled as a sinusoidal signal:

$$s_i(t) = A_i \sin\left(2\pi f_i t\right), \quad i = 1, 2, 3 \tag{11}$$

where $A_i$ and $f_i$ denote the amplitude and frequency of the $i$-th component, respectively.

The measured signal is obtained as

$$x(t) = s(t) + \nu(t), \quad \nu(t) \sim \mathcal{N}(0, \sigma^2) \tag{12}$$

where $\nu(t)$ is additive Gaussian noise.

The signal parameters (amplitudes, frequencies and noise level) are randomly sampled within predefined ranges, allowing the generation of a large number of signals with controlled variability.

For evaluation, the trained model was tested on real vibration measurements obtained from the CETIM gearbox dataset. The data were acquired using the gearbox test bench illustrated in Figure 3 (El Badaoui, Cahouet, Guillet, Danière, & Velex, 2001). This dataset contains vibration signals collected from a gearbox subjected to a progressive degradation process. It consists of twelve vibration recordings, each corresponding to a different day of acquisition, capturing the gradual transition from a healthy operating condition to a faulty state. The vibration signals were sampled at 20 kHz over a duration of 3 seconds per acquisition. The dominant component in the vibration spectrum corresponds to the gear meshing frequency, which varies between approximately 330 Hz and 346 Hz depending on the rotational speed. Compared to the simulated training data, the CETIM dataset presents more complex vibration patterns and higher noise levels, providing a more challenging scenario to evaluate the robustness and generalization capability of the proposed model.

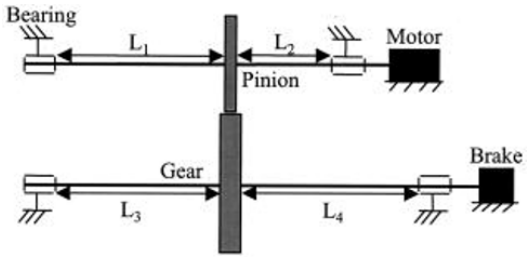


Figure 3. Schematic diagram of the CETIM gearbox test bench used for vibration signal acquisition.

### 4.2. Experimental setup

In the experiments, vibration signals are processed in segments of length $N$ =10000 samples. Simulated signals are generated at this length, whereas the CETIM dataset divides each 3-second acquisition into six non-overlapping segments, enabling multiple samples per acquisition and ensuring compatibility with the training data. Compressed measurements are generated according to the standard compressive sensing model described in (2). Five measurement sizes are evaluated, $M = \{5000, 4000, 3000, 2500, 2000\}$, corresponding to compression ratios of $50\%$, $40\%$, $30\%$, $25\%$, and $20\%$.

The sensing matrix is constructed using a structured random sampling strategy. The signal is partitioned into $M$ approximately equal intervals, and one sample is randomly selected

within each interval. Each row of $\Phi$ therefore contains a single non-zero entry equal to 1, producing a sparse binary measurement matrix that ensures a uniform coverage of the signal domain. The sensing matrix is generated once with a fixed random seed and kept constant during both training and testing.

The proposed reconstruction model is implemented as a convolutional autoencoder composed of an encoder and a decoder. The encoder maps the input vibration signal to a latent representation corresponding to the compressed measurements, while the decoder reconstructs the signal from this latent space. Table 1 presents the architecture used for a compression ratio of 25%. The encoder progressively reduces the signal length using convolutional layers and max-pooling operations until reaching a latent representation of length 2500. The decoder follows a symmetric structure that reconstructs the signal through convolutional layers and upsampling operations. For other compression ratios, the encoder–decoder architecture is reconfigured so that the latent representation matches the number of compressed measurements. Both the network structure and the dimensionality of intermediate layers are adjusted accordingly to ensure consistent reconstruction performance.

Table 1. Architecture of the proposed convolutional autoencoder for CR = 25%.

| Layer | Operation | Output size |
|---|---|---|
| Input | Vibration signal | (1, 10000) |
| | *Encoder* | |
| Conv1D + ReLU | 16 filters, kernel = 9 | (16, 10000) |
| MaxPool1D | stride = 2 | (16, 5000) |
| Conv1D + ReLU | 32 filters, kernel = 9 | (32, 5000) |
| MaxPool1D | stride = 2 | (32, 2500) |
| Conv1D | 1 filter, kernel = 9 | (1, 2500) |
| | *Decoder* | |
| Conv1D + ReLU | 32 filters, kernel = 9 | (32, 2500) |
| Upsample | factor = 2 | (32, 5000) |
| Conv1D + ReLU | 16 filters, kernel = 9 | (16, 5000) |
| Upsample | factor = 2 | (16, 10000) |
| Conv1D | 1 filter, kernel = 9 | (1, 10000) |
| Output | Reconstructed signal | (1, 10000) |

The proposed network is implemented in PyTorch and trained using the Adam optimizer with a learning rate of $10^{-3}$. The training objective combines two mean squared error (MSE) terms: a reconstruction loss between the original signal $x$ and the reconstructed signal $\hat{x}$, and a latent consistency loss between the encoder output and the corresponding compressed measurements. The weighting parameter $\lambda$ was selected through a grid search over $\{1, 0.1, 0.01, 0.001, 0.0001\}$, and $\lambda = 1$ was retained for all compression ratios. Training is performed for up to 200 epochs using early stopping with a patience of 10 epochs and a minimum improvement threshold of $10^{-5}$.

### 4.3. Reconstruction results and discussion

Reconstruction performance is evaluated using quantitative metrics that measure both signal fidelity and computational efficiency. The reconstruction quality is assessed using the normalized mean squared error (NMSE) and the cosine similarity index (CSI). NMSE measures the normalized difference between the original signal $x[n]$ and the reconstructed signal $\hat{x}[n]$:

$$\mathrm{NMSE} = \frac{\sum_{n=1}^{N}(x[n] - \hat{x}[n])^2}{\sum_{n=1}^{N} x[n]^2} \tag{13}$$

CSI evaluates spectral similarity by computing the cosine of the angle between the original spectrum $X[k]$ and the reconstructed spectrum $\hat{X}[k]$:

$$CSI = \frac{\sum_{k=1}^{K} X[k]\hat{X}[k]}{\sqrt{\sum_{k=1}^{K} |X[k]|^2}\sqrt{\sum_{k=1}^{K} |\hat{X}[k]|^2}} \times 100 \tag{14}$$

In addition, computational efficiency is evaluated by measuring the average reconstruction time per signal for both classical iterative methods and the proposed decoder based approach.

The proposed method is evaluated at five compression ratios (50%, 40%, 30%, 25% and 20%). Its performance is compared with greedy reconstruction algorithms, which are among the fastest sparse recovery methods, namely OMP, MP, and CoSaMP, as well as with a CNN baseline that uses the same architecture as the decoder of the proposed method and performs a direct mapping from compressed measurements to the reconstructed signal. All methods use the same measurement matrix and compression ratios to ensure a fair comparison.

Table 2 reports the NMSE values obtained by the different reconstruction methods. Lower NMSE indicates better reconstruction quality.

Table 2. NMSE between original and reconstructed signals for different compression ratios.

| | **50%** | **40%** | **30%** | **25%** | **20%** |
|---|---|---|---|---|---|
| OMP | 0.197 | 0.201 | 0.232 | **0.251** | **0.262** |
| MP | 0.196 | 0.2 | 0.205 | 0.253 | 0.265 |
| CoSaMP | 0.194 | 0.199 | 0.204 | 0.259 | 0.269 |
| CNN | 0.103 | 0.22 | 0.326 | 0.592 | 0.627 |
| Proposed method | **0.068** | **0.091** | **0.192** | 0.356 | 0.501 |

While NMSE evaluates time-domain reconstruction accuracy, vibration monitoring applications also require preservation of spectral characteristics. Table 3 reports the CSI between the spectra of the original and reconstructed signals.

Table 3. CSI (%) between the spectra of the original and reconstructed signals for different compression ratios.

| | **50%** | **40%** | **30%** | **25%** | **20%** |
|---|---|---|---|---|---|
| OMP | 89.62 | 89.83 | 89.39 | 89.09 | **88.95** |
| MP | 89.42 | 89.39 | 89.32 | **89.21** | 88.93 |
| CoSaMP | 89.57 | 89.42 | 85.99 | 89.05 | 88.86 |
| CNN | 96.47 | 93.97 | 92.47 | 84.43 | 78.4 |
| Proposed method | **98.08** | **97.42** | **93.56** | 87.95 | 82.95 |

At a CR of 50%, the proposed method clearly outperforms all competing approaches, achieving the lowest NMSE (0.068) and the highest CSI (98.08%), indicating more accurate reconstruction of the original signals compared with OMP, MP, CoSaMP, and CNN. The proposed method continues to provide the best performance at 40% and 30% CRs, maintaining the lowest NMSE values (0.091 and 0.192, respectively) and the highest CSI (97.42% and 93.56%). For 25% and 20% CRs, OMP and MP achieve the best performance in terms of reconstruction accuracy. However, the proposed method still remains competitive and consistently outperforms CNN in both NMSE and CSI metrics. This behavior can be explained by the fact that as the CR decreases, the dimensionality of the latent space becomes smaller, which limits the diversity of compressed measurements available to the decoder during reconstruction. Consequently, the decoder may not have sufficiently representative examples to accurately reconstruct the original signal. Nevertheless, the proposed method demonstrates strong reconstruction capabilities for all CRs and especially in the moderate-to-high compression regimes (30–50%), where it achieves the best overall performance. Moreover, although optimization algorithms may outperform the proposed method in certain cases, the latter provides additional advantages, particularly in terms of reconstruction time, as will be further analyzed in Table 4.

Table 4. Reconstruction time (s) for different compression ratios.

| | **50%** | **40%** | **30%** | **25%** | **20%** |
|---|---|---|---|---|---|
| OMP | 13.88 | 10.64 | 9.04 | 8.05 | 6.32 |
| MP | 14.13 | 10.72 | 8.48 | 7.83 | 6.12 |
| CoSaMP | 10.47 | 9.18 | 7.15 | 6.9 | 5.44 |
| CNN | 0.005 | 0.006 | **0.002** | **0.007** | **0.003** |
| Proposed method | **0.004** | **0.004** | 0.003 | **0.007** | 0.008 |

Specifically, OMP, MP, and CoSaMP require several seconds for reconstruction depending on the compression ratio, making them unsuitable for real-time monitoring scenarios. OMP requires between 6.32 and 13.88 seconds, MP between 6.12 and 14.13 seconds, and CoSaMP between 5.44 and 10.47 seconds per signal. In contrast, the proposed method achieves significantly faster reconstruction. It requires only 0.003–0.008 seconds per signal, which is several orders of magnitude faster than the traditional iterative sparse recovery algorithms. This improvement is achieved because reconstruction only involves a single forward pass through the trained decoder without requiring any iterative optimization procedure.

To further illustrate the reconstruction quality of the proposed method, Figure 4 presents an example of reconstructed vibration signals and their corresponding spectra at a compression ratio of 25% for day 1 of the CETIM dataset, corresponding to the healthy operating condition. The reconstructed signal closely follows the original signal in the time domain, while the dominant spectral components, including the gear meshing frequency and its harmonics, are well preserved in the frequency domain.

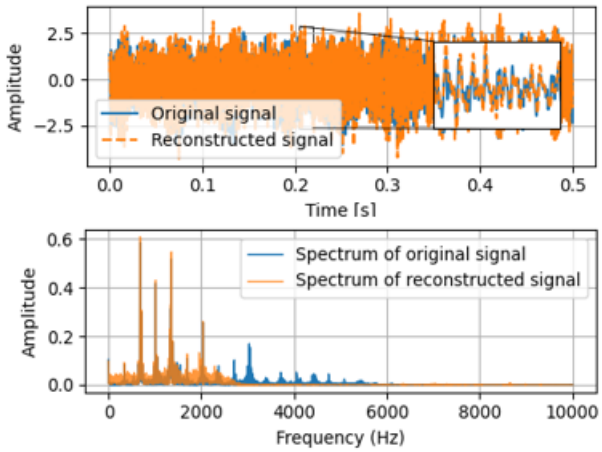


Figure 4. Comparison of original and reconstructed signals from day 1 in time (top) and frequency (bottom) domains.

Figure 5 shows a second example corresponding to day 12 of the dataset, representing the faulty state of the gearbox. Despite the more complex vibration patterns, the reconstructed signal remains consistent with the original signal in both time and frequency domains. A slight low-pass filtering effect is observed in the reconstructed spectra for both days, which is common in autoencoder based reconstruction. Nevertheless, the main spectral components are preserved, demonstrating that the proposed method maintains reliable reconstruction performance even under degraded operating conditions.

In a second step, the preservation of diagnostic information after the reconstruction process is evaluated. For this purpose, a cepstrum-based gear health indicator, originally proposed in (El Badaoui et al., 2004), is employed. In gearbox vibration signals, cepstral analysis reveals periodic structures associated with rotating components, where each gear generates a comb of peaks whose spacing corresponds to its rotational period. The amplitude of the first peak of each comb is proportional to the energy emitted by the corresponding gear

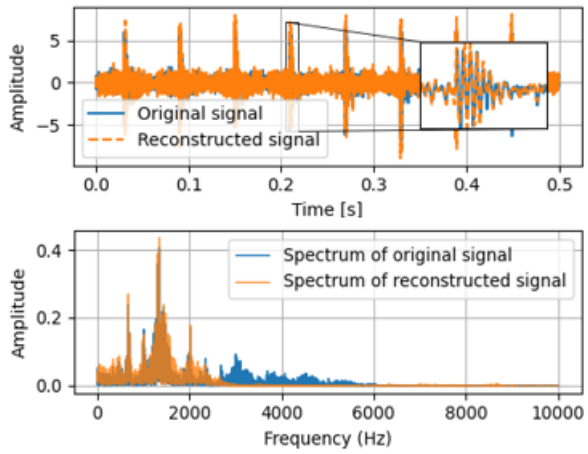


Figure 5. Comparison of original and reconstructed signals from day 12 in time (top) and frequency (bottom) domains.

during one rotation. A defect on a gear causes an increase in its associated cepstral peak and a decrease in the other. Based on this principle, a normalized fault indicator is defined as

$$d(t) = \frac{A_p(t) - A_r(t)}{A_p(t) + A_r(t)} \tag{15}$$

where $A_p(t)$ and $A_r(t)$ denote the amplitudes (or areas) of the first cepstral peaks associated with the pinion and the wheel, respectively. This indicator varies between $-1$ and 1 and reflects the redistribution of vibrational energy between the two gears.

Figure 6 shows the evolution of this health indicator computed from the original signals and from the reconstructed signals obtained using the different reconstruction methods.

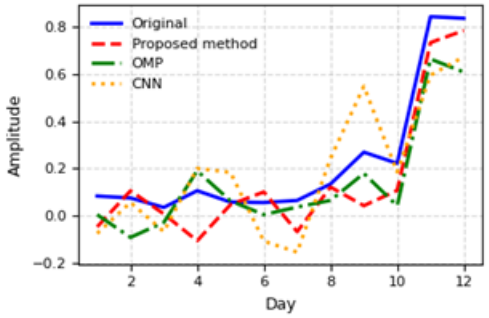


Figure 6. Evolution of the cepstrum based gear health indicator over the monitoring period for the original and reconstructed signals using different reconstruction methods.

During the healthy operating period (days 1–10), the proposed method generally follows the trend of the original indicator, although some fluctuations can be observed. A similar behavior is observed for OMP, which also reproduces the overall evolution of the indicator and performs better than the CNN baseline. Despite these small variations, the indicator clearly exhibits a noticeable increase during the last two days (days 11 and 12), corresponding to the faulty gearbox condition. This confirms that the main diagnostic information is preserved after reconstruction. While all reconstruction methods preserve the clear increase of the health indicator during the faulty stage, some earlier variations visible in the original trend appear attenuated after reconstruction. This behavior may result from the smoothing effect introduced by compression and reconstruction, which can reduce the visibility of weak fault-related features. Nevertheless, the proposed method preserves the overall evolution of the indicator and successfully identifies the transition from healthy to faulty operation.

Although OMP achieves slightly better performance in terms of NMSE and CSI, the proposed method remains competitive in preserving the fault-related information required for condition monitoring. Overall, it provides a favorable compromise between reconstruction accuracy, preservation of diagnostic indicators, and reconstruction time.

Beyond the comparison with traditional sparse recovery algorithms, it is also relevant to position the proposed approach with respect to recent deep learning based reconstruction methods reported in the literature, including ReconNet, ISTA-Net, ADMM-Net, and ADMM-CSNet. These approaches have demonstrated excellent reconstruction performance, particularly in image reconstruction tasks. However, many of them rely on optimization-inspired architectures, iterative unfolding strategies, or multi-stage reconstruction pipelines, which increase architectural complexity and often require application-specific adaptation. In contrast, the proposed framework is based on a lightweight convolutional autoencoder with an explicit measurement-consistency constraint that directly links the latent representation to the compressed measurements. This design enables straightforward deployment, low inference complexity, and direct reconstruction through a single forward pass.

## 5. CONCLUSION

This paper introduced a measurement-consistent deep learning framework for the fast reconstruction of compressively sensed vibration signals. By embedding the sensing model into a convolutional autoencoder, the proposed approach enables direct and non-iterative reconstruction while preserving physical consistency with the acquisition process. Experimental results on both simulated gearbox signals and real vibration data from the CETIM dataset demonstrate that the method achieves a favorable trade-off between reconstruction accuracy, computational efficiency, and preservation of diagnostic information. In particular, the proposed model outperforms classical sparse recovery algorithms in moderate compression regimes while reducing reconstruction time to the millisecond scale, making it suitable for real-time monitoring applications. Future work will focus on several direc-

tions. First, improving reconstruction performance at very low compression ratios by enhancing the robustness of the latent representation remains an important challenge. Second, extending the framework to more complex and diverse real-world datasets will be essential to improve generalization. Finally, integrating the proposed method into end-to-end diagnostic pipelines operating directly on compressed measurements represents a promising direction for efficient health monitoring systems.

## ACKNOWLEDGMENT

Imen Tounsi gratefully acknowledges the European Commission for its support of the Marie Sklodowska Curie program through the Horizon Europe DN PATRON project (GA 101120172).

## REFERENCES

Braun, S. (1988). Mechanical signature analysis—theory and applications. *Journal of Vibration, Acoustics, Stress, and Reliability in Design*, *110*(3), 418–419. doi: 10.1115/1.3269435

Brunton, S. L., & Kutz, J. N. (2019). Singular value decomposition (svd). In *Data-driven science and engineering: Machine learning, dynamical systems, and control* (pp. 3–46). Cambridge University Press. doi: 10.1017/9781108380690.002

Donoho, D. L. (2006). Compressed sensing. *IEEE Transactions on Information Theory*, *52*(4), 1289–1306.

El Badaoui, M., Cahouet, V., Guillet, F., Danière, J., & Velex, P. (2001). Modeling and detection of localized tooth defects in geared systems. *Journal of Mechanical Design*, *123*(3), 422–430. doi: 10.1115/1.1349420

El Badaoui, M., Guillet, F., & Daniere, J. (2004). New applications of the real cepstrum to gear signals, including definition of a robust fault indicator. *Mechanical Systems and Signal Processing*, *18*(5), 1031–1046. doi: 10.1016/j.ymssp.2004.01.005

Karkafi, F., Abboud, D., Leclere, Q., Antoni, J., & El Badaoui, M. (2024). Separation of vibratory components in complex systems for condition monitoring. *International Journal of COMADEM*, *27*(2), 21–29. Retrieved from https://hal.science/hal-04619794

Kulkarni, K., Lohit, S., Turaga, P., Kerviche, R., & Ashok, A. (2016). Reconnet: Non-iterative reconstruction of images from compressively sensed measurements. In *Proceedings of the ieee conference on computer vision and pattern recognition (cvpr)* (pp. 449–458).

Machidon, A. L., & Pejovic, V. (2023). Deep learning for compressive sensing: a ubiquitous systems perspective. *Artificial Intelligence Review*, *56*, 3619–3658. doi: 10.1007/s10462-022-10259-5

Mallat, S. G., & Zhang, Z. (1993). Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Processing*, *41*(12), 3397–3415.

Metzler, C. A., Maleki, A., & Baraniuk, R. G. (2017). Learned d-amp: Principled neural network based compressive image recovery. In *Advances in neural information processing systems (neurips)* (pp. 1772–1783).

Mousavi, A., Patel, A. B., & Baraniuk, R. G. (2015). A deep learning approach to structured signal recovery. In *Proceedings of the annual allerton conference on communication, control, and computing* (pp. 1336–1343). IEEE.

Needell, D., & Tropp, J. A. (2009). Cosamp: Iterative signal recovery from incomplete and inaccurate samples. *Applied and Computational Harmonic Analysis*, *26*(3), 301–321.

Tropp, J. A., & Gilbert, A. C. (2007). Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Transactions on Information Theory*, *53*(12), 4655–4666.

Yang, Y., Sun, J., Li, H., & Xu, Z. (2016). Deep admm-net for compressive sensing mri. In *Advances in neural information processing systems (neurips)* (Vol. 29, pp. 10–18).

Yang, Y., Sun, J., Li, H., & Xu, Z. (2020). Admm-csnet: A deep learning approach for image compressive sensing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *42*(3), 521–538. doi: 10.1109/TPAMI.2018.2883941

Zhang, J., & Ghanem, B. (2018). Ista-net: Interpretable optimization-inspired deep network for image compressive sensing. In *Proceedings of the ieee conference on computer vision and pattern recognition (cvpr)* (pp. 1828–1837).