Logic Meets Attention: A Neuro-symbolic Approach to Vibration Fault Detection

Darian M. Onchiș[a,*] and Eduard-Florin Hogea[a]

[a]Faculty of Mathematics and Computer Science, Department of Computer Science, West University of Timişoara, Vasile Parvan Avenue, Nr. 4, Timișoara, 300223, Romania
ORCID (Darian M. Onchiș): https://orcid.org/0000-0003-4846-3752, ORCID (Eduard-Florin Hogea): https://orcid.org/0000-0002-5814-992X

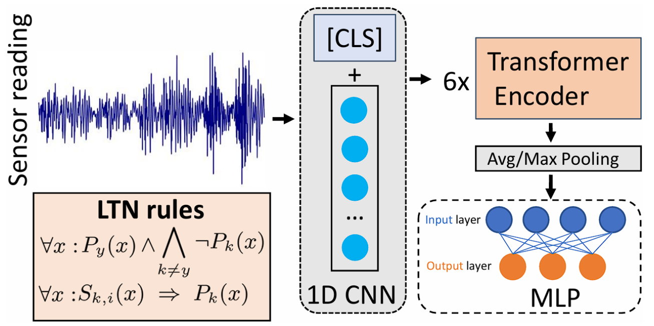

Figure 1. 1D Transformer pipeline with Logic Tensor Networks enforcing per-sample label consistency and dynamic similarity constraints.

Abstract. The early detection and classification of mechanical faults in rotating machinery is essential for predictive maintenance. We propose a novel neuro symbolic framework that integrates a one-dimensional Transformer encoder with Logic Tensor Networks and a dynamic rule generation module. The Transformer extracts temporal and spectral features from raw vibration segments via multi-head attention, while logic rules enforce label consistency and similarity constraints that adapt to evolving cluster patterns. We tested our approach on two benchmark datasets: on the OEDI recorded by using SpectraQuest's Gearbox Fault Diagnostics Simulator, where it achieves an F1 score of 0.992, and on the nine-class UoC gear fault data it reaches 0.899 versus 0.756 for a Transformer alone, thus delivering accurate and interpretable fault classification.

1 Introduction

Rotating machines are used in many industrial settings, and hidden faults have been shown to cause costly shutdowns or safety risks [12, 25]. Older vibration methods pick out spectral peaks, envelope signals, or wavelet features and feed them into rule-based or fuzzy systems [13, 21, 2, 4]. However, these hand-crafted steps take a lot of work and often break down when the data get large or the operating conditions change.

Deep end-to-end models. Convolutional, recurrent and autoencoder networks now learn features directly from raw signals, outperforming classical methods on CWRU, IMS and MFPT benchmarks and adapting through augmentation or domain adversaries [11, 15, 7, 26, 27, 23, 28]. Surveys published summarise these gains but note two gaps: large labelled datasets remain necessary and explanations are opaque [29, 19, 24].

Neuro-symbolic promise and objective. Integrating neural representation learning with symbolic reasoning can inject domain rules, boost data efficiency, and produce human-readable explanations [5, 6, 18]. We therefore propose a diagnosis model that couples a Transformer with a Logic Tensor Network (LTN) [22]. Self-attention captures long-range temporal patterns, while the LTN layer enforces first-order rules linking observed cues to fault modes.

2 Related Work

Transformer architectures have recently become a backbone for machinery diagnosis. The Time-Series Transformer (TST) lifted CWRU accuracy to 99.1%, four points above CNN/LSTM baselines [14]. Existing improvements include CNN tokenisers [16] and works which prune redundant attention and halves the number of floating operations without sacrificing accuracy [10]. Vision-style patching further enables a Siamese ViT to reach state-of-the-art performance with only 20% labeled data [8].

Efforts to inject expert knowledge have given rise to several neuro-symbolic approaches. LTNs embedded in LSTM backbones enhance generalisation when labels are limited [9]. DeepProbLog combines neural perception with probabilistic logic programming, and forces models to predict intermediate engineering quantities, enabling subsequent auditability [17].

Despite these advances, Transformer-based models rarely encode formal knowledge, whereas LTN systems rely on legacy feature extractors. To date, no study has combined a Transformer backbone with a neuro-symbolic reasoning layer. Our work unifies self-attention representation learning and differentiable first-order logic in a single end-to-end framework aimed at improving accuracy, data efficiency and transparency.

3 Methodology

We fuse first-order reasoning with a lightweight 1-D Transformer to classify sample vibration segments(best results with 20) under explicit consistency constraints. Symbolic rules are encoded with Logic Tensor Networks (LTNs) [1] their penalisation signals update the network, letting prior knowledge shape the learned features.

* Corresponding Author. Email: darian.onchis@e-uvt.ro

3.1 LTN Optimisation

An LTN grounds every formula φ in a truth degree $\mathcal{G}_\theta(\varphi) \in [0,1]$. Model parameters θ are learned by maximising the aggregated satisfiability of a rule set $\mathcal{K}$:

$$\theta^\star = \arg\max_\theta \mathrm{Agg}_{\varphi\in\mathcal{K}} \mathcal{G}_\theta(\varphi). \tag{1}$$

We adopt the *aggregated p-mean error* (ApME),

$$A_{p\mathrm{ME}}(x_{1:n}) = 1 - \left(\tfrac{1}{n}\sum_i (1-x_i)^p\right)^{1/p}, \tag{2}$$

with p=2; low-valued clauses thus receive larger gradients.

Each class k is a fuzzy predicate $P_k(\mathbf{x}) = \hat{p}_k$, i.e. the network softmax output. For a labelled segment $\mathbf{x}$ with ground truth y we impose

$$P_y(\mathbf{x}), \qquad \neg P_k(\mathbf{x})\ (k \neq y), \tag{3}$$

using product t- and s-norms plus the Goguen implication; run-time overhead is negligible.

3.2 Hybrid Encoder meets LTN

Each vibration segment is a length-20 time series $\mathbf{x} \in \mathbb{R}^{1\times 20}$ with a categorical label $y \in \{1,\dots,C\}$, where C is the number of classes in the current dataset. The proposed model attaches a lightweight one-dimensional Transformer to the LTN layer, so that symbolic constraints are applied directly to the learned representation.

Transformer encoder. The raw signal is split into *three* overlapping patches of ten samples (stride = 5). A 1×10 convolution projects each patch to a 64-dimensional token $\mathbf{u}_i$. A learnable class token $\mathbf{z}_{\mathrm{cls}}$ is *prepended* to these tokens and sinusoidal positional encodings are added, yielding the input sequence $\mathbf{Z}_0 = \big[\mathbf{z}_{\mathrm{cls}} \parallel \mathbf{u}_1 \parallel \mathbf{u}_2 \parallel \mathbf{u}_3\big]$. Six pre-norm Transformer layers, each with eight attention heads, refine the sequence. With only four tokens (one CLS + three patches) the memory footprint remains low while self-attention still captures long-range dependencies. The updated CLS token provides a global summary, whereas the three patch tokens are average– and max-pooled, concatenated, and fed to a two-layer MLP that outputs class probabilities $\hat{\mathbf{p}}(\mathbf{x}) = (\hat{p}_1,\dots,\hat{p}_C)$. These probabilities ground the LTN predicates $P_k(\mathbf{x}) = \hat{p}_k$, allowing the logic rules to influence the entire encoder.

Similarity rules. To regularise the embedding space, *similarity rules* are refreshed after every training epoch. For each class k we run k-means on the current embeddings and retain two centroids $\{\boldsymbol{\mu}_{k,1}, \boldsymbol{\mu}_{k,2}\}$, sufficient to distinguish the low- and high-load regimes observed in practice. Proximity of $\mathbf{x}$ to centroid i is measured by a Gaussian kernel

$$S_{k,i}(\mathbf{x}) \;=\; \exp\!\Big[-\tfrac{1}{2}\big\|\mathbf{f}(\mathbf{x}) - \boldsymbol{\mu}_{k,i}\big\|_2\Big],$$

which yields the soft implication

$$S_{k,i}(\mathbf{x}) \;\Rightarrow\; P_k(\mathbf{x}). \tag{4}$$

Rule (4) encourages any sample that lies close to a class centroid to be assigned that class, yet it never conflicts with the primary label rules in (3).

Training objective. Let $\{\alpha_r\}_{r=1}^R$ be the truth values of all instantiated rules. We define the satisfiability aggregation

$$\mathrm{SatAgg} = 1 - \sqrt{\frac{1}{R}\sum_{r=1}^{R}\big(1-\alpha_r\big)^2}\,, \tag{5}$$

which attains its maximum value of 1 when every rule is fully satisfied. The overall loss is then

$$\mathcal{L} = 1 - \mathrm{SatAgg} + \beta\|\Theta\|_2^2\,, \tag{6}$$

where Θ comprises all learnable parameters and $\beta = 10^{-3}$ is the weight-decay coefficient.

4 Experimental Evaluation

All experiments ran on a 16 GiB Linux workstation with an AMD EPYC CPU. Datasets are **UoC gearbox** [3] (nine fault modes, 20 kHz): HEA (healthy), CTF (chipped-tooth PGB), MTF (missing-tooth PGB), RCF (root-crack PGB), SWF (surface-wear PGB), BWF (ball-wear bearing), CWF (composite-wear races), IRF (inner-race bearing), ORF (outer-race bearing); and data from OEDI [20] recorded by using SpectraQuest's Gearbox Fault Diagnostics Simulator, refered to as **SpectraSimulator** (healthy, broken-tooth).

Pre-processing Continuous vibration signals are segmented into windows of 20 samples with a stride of 10 (yielding 1×20 inputs). Signals are standardised; faulty frames are discarded. Class balance is enforced by uniform sampling. Splits use a stratified 80/20 train–test ratio, and results are averaged over two folds.

Training Details Training uses Adam with a learning rate of 10^{-4} and ℓ_2 weight-decay 10^{-3}. When training the vanilla variant, training has been kept identical.

4.1 Results

UoC. Our hybrid model attains 89.9 % accuracy and macro-F1 = 0.900 as in Table 1, outperforming the vanilla Transformer with an F1 = 0.756. Gains are largest for data-sparse classes such as HEA and CTF, highlighting the benefit of logical regularisation.

Table 1. Results comparison on the UoC dataset: Transformer–LTN vs. vanilla Transformer (the transformer itself has been kept identical).

	Hybrid Transformer–LTN			**Vanilla Transformer**		
Class	**Prec.**	**Rec.**	**F1**	**Prec.**	**Rec.**	**F1**
HEA	0.830	0.940	0.881	0.611	0.795	0.691
CTF	0.895	0.819	0.855	0.600	0.723	0.656
MTF	0.921	0.843	0.881	0.919	0.687	0.786
RCF	1.000	0.964	0.982	0.976	0.952	0.964
SWF	0.928	0.928	0.928	0.769	0.843	0.805
BWF	0.860	0.892	0.876	0.790	0.590	0.676
CWF	0.906	0.917	0.911	0.716	0.810	0.760
IRF	0.847	0.867	0.857	0.720	0.651	0.684
ORF	0.916	0.916	0.916	0.824	0.735	0.777
AVG	0.900	0.898	0.899	0.770	0.754	0.756

SpectraSimulator. With only two classes, our model reaches macro-F1 = 0.992 (Table 2), confirming a strong generalization.

Table 2. 1-D Transformer with LTN model results on SpectraSimulator.

Class	Prec.	Rec.	F1
Healthy	0.987	0.997	0.992
Broken tooth	0.997	0.987	0.992

5 Conclusion

We introduced a compact neuro-symbolic pipeline that unites a 1-D Transformer with LTNs, enabling end-to-end learning under first-order constraints. Experiments on real and simulated gear faults show that logical supervision raises both accuracy and class balance without extra computation, while every prediction remains traceable to interpretable rules.

6 Acknowledgements

This research is partially supported by the project "Romanian Hub for Artificial Intelligence - HRIA", Smart Growth, Digitization and Financial Instruments Program, 2021-2027, MySMIS no. 334906.

References

[1] S. Badreddine, A. d. Garcez, L. Serafini, and M. Spranger. Logic tensor networks. *Artificial Intelligence*, 303:103649, 2022.

[2] M. E. H. Benbouzid. A review of induction motors signature analysis as a medium for fault detection. *IEEE Transactions on Industrial Electronics*, 47(5):984–993, 2000.

[3] P. Cao, S. Zhang, and J. Tang. Gear fault data. *https://doi.org/10.6084/m9.figshare*, 6127874:v1, 2018.

[4] J. Chen and H. Hashimoto. An experimental study of fuzzy logic approach for machine fault diagnosis. *ISA Transactions*, 34(1):79–90, 1995.

[5] A. S. d'Avila Garcez, L. C. Lamb, and D. M. Gabbay. *Neural-Symbolic Cognitive Reasoning*. Springer, 2009.

[6] I. Donadello, L. Serafini, and A. d'Avila Garcez. Logic tensor networks for semantic image interpretation. In *Proc. of the 26th Int. Joint Conf. on Artificial Intelligence (IJCAI)*, pages 1596–1602, 2017. doi: 10.24963/ijcai.2017/221.

[7] M. Dong and D. Wang. Fault diagnosis based on deep feature learning with improved autoencoder. *Sensors*, 16(10):1654, 2016.

[8] Q. He, S. Li, Q. Bai, A. Zhang, J. Yang, and M. Shen. A siamese vision transformer for bearings fault diagnosis. *Micromachines*, 13(10):1656, 2022.

[9] E. Hogea, D. M. Onchiş, R. Yan, and Z. Zhou. LogicLSTM: Logically-driven long short-term memory model for fault diagnosis in gearboxes. *Journal of Manufacturing Systems*, 77(11):892–902, 2024. doi: 10.1016/j.jmsy.2024.10.003.

[10] Y. Hou, J. Wang, Z. Chen, J. Ma, and T. Li. Diagnosisformer: An efficient rolling bearing fault diagnosis method based on improved Transformer. *Engineering Applications of Artificial Intelligence*, 124:106507, 2023. doi: 10.1016/j.engappai.2023.106507.

[11] T. Ince, S. Kiranyaz, L. Eren, M. Askar, and M. Gabbouj. Real-time motor fault detection by 1-d convolutional neural networks. *IEEE Transactions on Industrial Electronics*, 63(11):7067–7075, 2016.

[12] R. Isermann. *Fault-Diagnosis Systems: An Introduction from Fault Detection to Fault Tolerance*. Springer, 2006.

[13] A. K. S. Jardine, D. Lin, and D. Banjevic. A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing*, 20(7):1483–1510, 2006.

[14] Y. Jin, L. Hou, and Y. Chen. A time series transformer based method for the rotating machinery fault diagnosis. *Neurocomputing*, 494:379–395, 2022.

[15] Y. Lei, F. Jia, J. Lin, S. Xing, and S. Ding. An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data. *IEEE Transactions on Industrial Electronics*, 63(5):3137–3147, 2016.

[16] Z. Lu, L. Liang, J. Zhu, W. Zou, and L. Mao. Rotating Machinery Fault Diagnosis Under Multiple Working Conditions via a Time-Series Transformer Enhanced by Convolutional Neural Network. *IEEE Transactions on Instrumentation and Measurement*, 72:1–11, 2023. doi: 10.1109/TIM.2023.3318707.

[17] R. Manhaeve, S. Verbeke, G. Marra, D. Fierens, and L. D. Raedt. Deepproblog: Neural probabilistic logic programming. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 31, pages 3749–3759, 2018.

[18] J. Mao, C. Gan, P. Kohli, J. B. Tenenbaum, and J. Wu. Neuro-symbolic concept learner: Interpreting scenes, words, and sentences with scene graphs. In *Proc. of the 2019 IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)*, pages 11135–11144, 2019.

[19] S. J. Mousavirad and M. Ishkakov. Review on deep learning classifiers for faults diagnosis of rotating machinery. *Artificial Intelligence Review*, 56:5809–5860, 2023.

[20] Y. Pandya. Gearbox fault diagnosis data, 06 2018. URL https://data.openei.org/submissions/623.

[21] R. B. Randall and J. Antoni. Rolling element bearing diagnostics—a tutorial. *Mechanical Systems and Signal Processing*, 25(2):485–520, 2011.

[22] L. Serafini and A. S. d'Avila Garcez. Learning and reasoning with logic tensor networks. In *Proc. of the 15th Int. Conf. of the Italian Association for Artificial Intelligence (AI*IA 2016)*, volume 10037 of *LNCS*, pages 334–348. Springer, 2016.

[23] S. Shao, S. McAleer, R. Yan, and P. Baldi. Highly accurate machine fault diagnosis using deep transfer learning. *IEEE Transactions on Industrial Informatics*, 15(4):2446–2455, 2019.

[24] B. A. Tama, M. Vania, S. Lee, and S. Lim. Recent advances in the application of deep learning for fault diagnosis of rotating machinery using vibration signals. *Artificial Intelligence Review*, 56:4667–4709, 2023.

[25] N. Tandon and A. Choudhury. A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings. *Tribology International*, 32(8):469–480, 1999.

[26] D. Verstraete, A. Abdelguerfi, and A. E. Smith. Deep learning for fault detection in rotating machinery. *Shock and Vibration*, 2017:Article ID 3901549, 2017.

[27] L. Wen, X. Li, L. Gao, and Y. Zhang. A new deep transfer learning method for intelligent fault diagnosis. *IEEE Transactions on Industrial Electronics*, 64(3):2307–2316, 2017.

[28] W. Zhang, G. Peng, C. Li, Y. Chen, and Z. Zhang. A new deep learning model for fault diagnosis of rotating machinery under varying speed conditions. *Neurocomputing*, 290:219–242, 2018.

[29] Z. Zhu, Y. Lei, G. Qi, Y. Chai, N. S. Mazur, Y. An, and X. Huang. A review of the application of deep learning in intelligent fault diagnosis of rotating machinery. *Measurement*, 206:112346, 2023.