

**O‘ZBEKISTON RESPUBLIKASI OLIY TA’LIM, FAN VA
INNOVATSIYALAR VAZIRLIGI**

**MIRZO ULUG‘BEK NOMIDAGI O‘ZBEKISTON MILLIY
UNIVERSITETINING JIZZAX FILIALI**



KOMPYUTER ILMLARI VA MUHANDISLIK TEXNOLOGIYALARI
mavzusidagi Xalqaro ilmiy-texnik anjuman materiallari to‘plami
2026-yil 25-26-sentabr

(1-qism)

JIZZAX-2026

Kompyuter ilmlari va muhandislik texnologiyalari. Xalqaro ilmiy-texnik anjuman materiallari to'plami – Jizzax: O'zMU Jizzax filiali, 2026-yil 25-26-sentabr. 507-bet.

Xalqaro miqyosidagi ilmiy-texnik anjuman materiallarida zamonaviy kompyuter ilmlari va muhandislik texnologiyalari sohasidagi innovatsion tadqiqotlar aks etgan.

Globalashuv sharoitida davlatimizni yanada barqaror va jadal sur'atlar bilan rivojlantirish bo'yicha amalga oshirilayotgan islohotlar samarasini yaxshilash sohasidagi ilmiy-tadqiqot ishlariga alohida e'tibor qaratilgan. Zero iqtisodiyotning, ijtimoiy sohalarni qamrab olgan modernizatsiya jarayonlari, hayotning barcha sohalarini liberallashtirishni talab qilmoqda.

Ushbu ilmiy ma'ruza tezislari to'plamida mamlakatimiz va xorijlik turli yo'nalishlarda faoliyat olib borayotgan mutaxassislar, olimlar, professor-o'qituvchilar, ilmiy tadqiqot institutlari va markazlarining ilmiy xodimlari, tadqiqotchilari, magistr va talabalarining ilmiy-tadqiqot ishlari natijalari mujassamlashgan.

Mas'ul muharrirlar: DSc.prof. Turakulov O.X., t.f.n., dots. Baboyev A.M.

Tahrir hay'ati a'zolari: p.f.d.(DSc), prof. Turakulov O.X., t.f.n., dots. Baboyev A.M., t.f.f.d.(PhD), prof. Abduraxmanov R.A., p.f.f.d.(PhD) Eshankulov B.S., p.f.n., dots. Aliqulov S.T., p.f.n., dots. Alimov N.N., p.f.f.d.(PhD), dots. Alibayev S.X., t.f.f.d.(PhD), dots. Abdumalikov A.A., p.f.f.d.(PhD) Hafizov E.A., f.f.f.d.(PhD) dots., Norbekova B.Sh. p.f.f.d.(PhD) dots. Sindorov L.K., t.f.f.d.(PhD), dots. Nasirov B.U., b.f.f.d. (PhD) O'ralov A.I., p.f.n., dots. Aliqulov S.T., t.f.f.d.(PhD) Kuvandikov J.T., i.f.n., dots. Tsoy M.P., Sharipova S.F., Jo'rayev M.M.

Mazkur to'plamga kiritilgan ma'ruza tezislarining mazmuni, undagi statistik ma'lumotlar va me'yoriy hujjatlarning to'g'riligi hamda tanqidiy fikr-mulohazalar, keltirilgan takliflarga mualliflarning o'zlari mas'uldirlar.

7. B. Remeseiro and V. Bolón-Canedo, “A review of feature selection methods in medical applications,” *Computers in Biology and Medicine*, vol. 112, Art. no. 103375, 2019, doi: 10.1016/j.combiomed.2019.103375.

DATA ANOMALY DETECTION IN STRUCTURAL HEALTH MONITORING USING A MODIFIED TRANSFORMER ENCODER WITH 1D-CNN LAYERS

Sirojiddin Nuriev

TassVision nuriyev14051989@gmail.com

Abstract—Structural health monitoring (SHM) systems produce continuous, high-volume sensor streams in which environmental effects, transmission faults and sensor malfunctions create data anomalies that corrupt downstream structural assessment. We present a lightweight classifier that assigns each one-hour acceleration record to one of seven patterns: normal, missing, minor, outlier, square, trend and drift. Each record is split into 30-s windows, and four statistical and 20 frequency-domain features are extracted per window, giving a 120×24 feature sequence. The sequence is classified by a Transformer encoder in which the position-wise feed-forward network is replaced by two 1×1 one-dimensional convolutional (1D-CNN) layers and positional encoding is removed. On data from 38 accelerometers on a cable-stayed bridge, the model reaches 97.1% test accuracy with about 51K parameters, against 95.8% for a standard encoder with 446K parameters, and exceeds the accuracies reported for image-based DNN/CNN and shapelet-based methods on data from the same bridge.

Index Terms—structural health monitoring, data anomaly detection, Transformer, 1D-CNN, time-series classification.

I. INTRODUCTION

SHM systems on long-span bridges and other civil structures generate data volumes that grow continuously as sensor networks expand. These records frequently contain anomalies caused by environmental factors, system disruptions or sensor faults; left unflagged, they bias modal identification, damage detection and any assessment of the structure’s condition. Conventional detectors such as statistical thresholding and distance-, rule- and clusteringbased methods require per-channel tuning and become less effective and more expensive as the data flow grows.

Deep learning recasts the task as multiclass classification of fixed-length segments. Bao *et al.* [1] rendered each hour of acceleration as time- and frequency-domain images and trained a deep neural network (DNN) classifier, reaching 87.0% accuracy. Tang *et al.* [2] fused time and frequency plots into dual-channel images for a CNN (94.1%), and Arul and Kareem [3] classified shapelet-transform features with a random forest (92.97%). Image-based pipelines add rendering cost and large inputs, and all three methods struggle with rare patterns such as outliers and drift.

Transformers [4] model long-range dependencies through self-attention, but the standard encoder is parameter-heavy for an SHM setting with many channels and a continuous stream. This paper, which summarizes [5], makes three contributions:

(i) a compact representation of 24 statistical and spectral features per window instead of images; (ii) a

modified encoder that replaces the feed-forward network with two 1D-CNN layers and drops positional encoding, using $8.7\times$ fewer parameters; and (iii) an ablation and per-class evaluation on 38-channel bridge data, where the model reaches 97.1% test accuracy.

II. PROPOSED METHOD

A. Feature Extraction

Each channel's one-hour record is normalized and divided into n non-overlapping subsequences $x^k \in \mathbb{R}^m$, $k = 1, \dots, n$.

For every subsequence we compute four statistics, $\min^k$, $\max^k$, $\mu^k = (1/m) \sum_i x_i^k$, $\text{var}^k = (1/m) \sum_i (x_i^k - \mu^k)^2$, (1)

and take the FFT magnitude spectrum of x^k , from which $w = 20$ frequency-domain features are retained. Stacking the windows gives $X \in \mathbb{R}^{n \times (w+4)}$. We compare windows of 30 s, 1 min, 5 min and 10 min, i.e. X of size 120×24 , 60×24 , 12×24 and 6×24 .

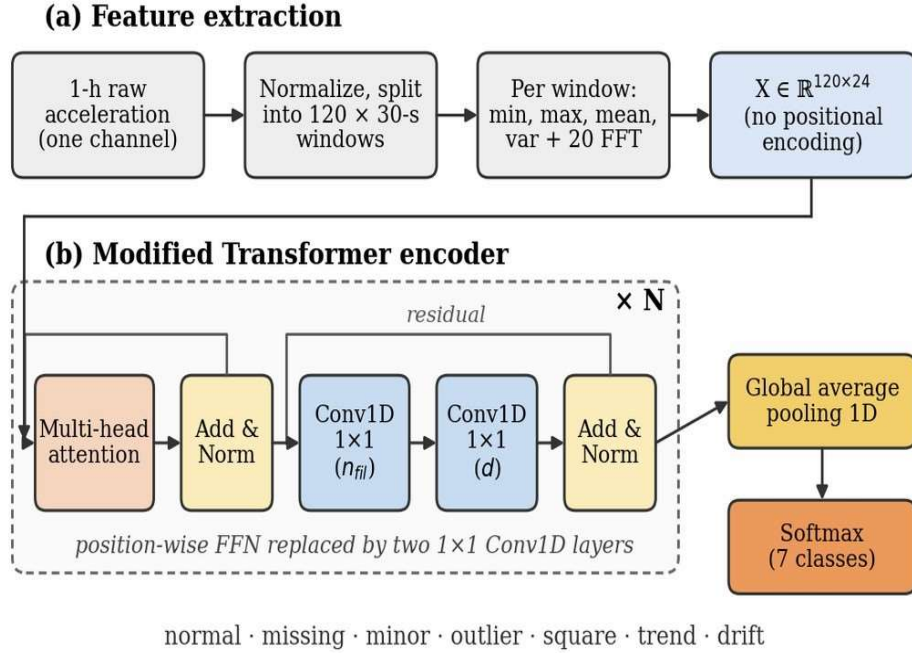


Fig. 1. Proposed pipeline. (a) Each one-hour channel record becomes a sequence of 120 window-level feature vectors. (b) Encoder layer with the feed-forward block replaced by two 1×1 Conv1D layers, followed by global average pooling and a 7-way softmax.

B. Modified Transformer Encoder

As shown in Fig. 1(b), each encoder layer applies multihead attention (MHA) with h heads, key dimension d_k and value dimension d_v , followed by two 1×1 Conv1D layers with n_{fil} and d filters:

$Z_2 = \text{LN}(X + \text{MHA}(X))$, $Z_5 = \text{LN}(Z_2 + C_d(C_{n_{\text{fil}}}(Z_2)))$, (2) where LN is layer normalization and C_k a 1×1 convolution with k filters. Positional encoding is omitted, so each window is treated as a token whose content, rather than its absolute position within the hour, drives the classification. After N layers, global average

pooling over the n windows and a softmax give the seven class probabilities. The number of learnable parameters is

$$N_p = N[h(d+1)(2d_k+d_v) + d(hd_v+1)(3) + n_{fil}(d+1) + d(n_{fil}+1)].$$

III. EXPERIMENTS

A. Dataset and Setup

We use acceleration data from 38 channels installed on the deck and towers of a cable-stayed bridge, recorded in January and February 2012 and segmented into non-overlapping onehour samples, each labeled with one of the seven patterns. Data are split chronologically (Table I); the class distribution is highly imbalanced, with outliers and drift under 3% of samples.

TABLE I
NUMBER OF ONE-HOUR SAMPLES PER SPLIT

Class	Train			Validation	Test
Period	Jan 1–24			Jan 25–31	Feb
Normal	10,471	3,104	13,575		
Missing	2,419	523	2,942		
Minor	1,378	397	1,775		
Outlier	356	171	527		
Square	2,300	696	2,996		
Trend	4,426	1,352	3,778		
Drift	538	141	679		
Total	21,888	6,384	26,272		

Hyperparameters were searched over $N \in \{1, 2, 3, 4, 6, 8\}$, $h \in \{1, 2, 4, 8, 10, 12, 14, 16\}$, $d_k, d_v \in \{16, 24, 32, 64, 128\}$ and $n_{fil} \in \{8, 16, 32, 64, 128\}$, with configurations grouped by parameter budget ($\approx 25K$, $50K$ and $100K$). Longer windows (1–5 min) reached slightly higher training accuracy but lower validation accuracy; 30-s windows gave the best validation accuracy ($\approx 96.3\%$), which saturated beyond $\approx 50K$ parameters. We therefore select 30-s windows and a $\approx 51K$ parameter model. Performance is reported as accuracy, precision, recall and F_1 score.

B. Results

The selected model reaches 98.1% training, 96.3% validation and 97.1% test accuracy. Table II isolates the two modifications. Replacing the feed-forward network with 1DCNN layers cuts the parameter count from 446K to 51K and raises test accuracy by 0.9–1.3 points; the full-size encoders fit the training set best (up to 99.4%) but generalize worst, indicating overfitting. Removing positional encoding improves test accuracy in both configurations (+0.5 and +0.9 points). The proposed model also exceeds the accuracies reported in [1]–[3].

TABLE II
ACCURACY OF ENCODER VARIANTS AND PRIOR METHODS

Encoder variant / method	Params	Train	Val.	Test
FFN + PE	446K	99.4	94.2	95.3
FFN, no PE	446K	98.9	94.5	95.8
1D-CNN + PE	51K	98.5	95.1	96.2
1D-CNN, no PE (ours)	51K	98.1	96.3	97.1

Image-based DNN [1] — — — 87.0* Dual-image CNN [2] —
 — — 94.1*
 Shapelet + RF [3] — — — 92.97*

FFN: feed-forward network; PE: positional encoding. Accuracies in %. *As reported by the original authors; data periods and splits differ from ours.

Per-class results (Table III) show near-perfect separation of missing and square patterns and F_1 above 91% for normal, minor and trend. The weakest classes are outlier ($F_1 = 75.3\%$), whose errors are mostly predicted as normal, and drift (82.8%), which is mainly confused with trend; both are also the rarest classes in training.

TABLE III
PER-CLASS TEST PRECISION,

Class	Precision	Recall	F_1	RECALL AND F_1 (%)
Normal	97.7	99.4	98.5	We presented a compact SHM data-anomaly classifier that combines window-level statistical and spectral features with a Transformer encoder whose feed-forward block is replaced by 1D-CNN layers and whose positional encoding is removed. With about 51K parameters, it reaches 97.1% test accuracy on 38-channel bridge data, outperforming a
Missing	99.9	99.8	99.8	
Minor	93.3	89.8	91.5	
Outlier	79.1	71.9	75.3	
Square	99.9	97.3	98.6	
Trend	97.4	95.5	96.4	
Drift	80.9	84.8	82.8	
IV. CONCLUSION				

standard 446K-parameter encoder. Future work will address the minority outlier and drift classes through class-balanced training and augmentation, and evaluate the model on other structures and sensor types.

REFERENCES

- [1] Y. Bao, Z. Tang, H. Li, and Y. Zhang, “Computer vision and deep learning–based data anomaly detection method for structural health monitoring,” *Struct. Health Monit.*, vol. 18, no. 2, pp. 401–421, 2019.
- [2] Z. Tang, Z. Chen, Y. Bao, and H. Li, “Convolutional neural networkbased data anomaly detection method using multiple information for structural health monitoring,” *Struct. Control Health Monit.*, vol. 26, no. 1, e2296, 2019.
- [3] M. Arul and A. Kareem, “Data anomaly detection for structural health monitoring of bridges using shapelet transform,” arXiv:2009.00470, 2020.
- [4] A. Vaswani *et al.*, “Attention is all you need,” in *Advances in Neural Information Processing Systems*, vol. 30, 2017.
- [5] S. Nuriev *et al.*, “Data anomaly detection in structural health monitoring using modified transformer encoders with 1D-CNN layers,” *Smart Struct. Syst.*, vol. 35, no. 6, p. 337, 2025.