

Hybrid-BiLSTM: An Attention-Enhanced Deep Learning Framework for Bearing Remaining Useful Life Prediction

Urvashi Singh Jadaun^a, Shailendra Kumar Shriwastava^b, Dr. Poonam Bhartiya^c

^aComputer Science Engineering Department, BTIRT, Sagar, MP, India

^bAssistant Professor (Guide) Computer Science Engineering Department, BTIRT, Sagar, MP, India

^cAssistant Professor (Co-Guide) Computer Science Engineering Department, BTIRT, Sagar, MP, India

^aurvashi.jadaun@gmail.com, ^bshriwastava.shailendra@gmail.com, ^cpoonambhartiya1989@gmail.com,

ABSTRACT

Rolling element bearings are among the most crucial parts of rotating machinery and the most failure-prone components, so estimating their remaining useful life before failure is essential for planning maintenance rather than reacting to breakdowns. The proposed work develops a system to predict the Remaining Useful Life (RUL) of high-speed bearings that combines the conventional machine learning methods with a custom deep learning architecture. For the Experiments, the PRONOSTIA dataset from the FEMTO-ST Institute was used, which contains vibration signals recorded at 25.6 kHz under varying speeds and loads. Time-domain and frequency-domain features are calculated from the Raw signals that RE splits into overlapping windows. The features are standardized and arranged into fixed-length sequences to preserve temporal patterns. Baseline models SVM, Random Forest, XGBoost, LSTM, and BiLSTM are trained for comparison. The proposed Hybrid-BiLSTM model builds on these by combining a multi-scale convolutional layer, dilated Temporal Convolutional Network blocks, bidirectional LSTM layers, and multi-head attention, followed by pooling and a regression layer. Trained with a Huber loss and early stopping. The MAE, RMSE, and R^2 , are effectively evaluated for estimating model performance and for RUL estimation.

Keywords: - *Remaining useful life (RUL), Predictive Maintenance, Vibration Signal Analysis, Deep Learning, Time-Series Prediction, signal processing.*

1. Introduction

A bearing is a rotating component that is used to reduce friction between moving parts and enable smooth motion under heavy load. Bearings are widely used in almost all industries, including the automotive, aerospace, wind turbine, rail transport, and manufacturing. Bearing failure is one of the major causes of unplanned downtime, making it necessary to accurately determine the Remaining Useful Life (RUL), which is important in Prognostics and Health Management (PHM) in general, where predictive and condition-based maintenance are widely used [1].

Past prognostic approaches were primarily physics-based. Yanfei Lu et al. [4] integrated realized volatility (RV) with a wavelet neural network (WNN) to forecast bearing RUL, using an AR filter to extract fault and energy ratios to identify irregularities on vibration signals. Meanwhile, these techniques scarcely generalize with the environment in which they are implemented, which leads to the development of AI techniques that are more proficient in capturing nonlinear dependencies between condition-monitoring parameters and RUL. Various Supervised methods were used for solving this task, such as Artificial Neural Networks (ANN)

[2], Support Vector Machines (SVM) [3], Relevance Vector Machines (RVM) [5], Random Forest (RF) [6], gradient boosting methods such as XGBoost, LightGBM, and CatBoost, and k-Nearest Neighbors (k-NN) [7], all of which learn the relationship between extracted features for forecasting and RUL variables. Li et al. [8] stated that SVM is quite efficient for estimating the RUL of bearings because it can capture nonlinear degradation processes using kernel functions; however, it requires careful preprocessing and hyperparameter tuning [9]. Yan et al. [10] demonstrated consistently strong SVM classification performance for bearing fault detection. The k-NN algorithm offers a flexible alternative, estimating RUL by identifying the 'k' nearest data points in the training set and aggregating their values, as described by Nistane et al. [3].

Deep learning methods have since gained prominence for automatically learning discriminative features from raw signals. which were designed primarily for image data, have been extensively utilized for fault diagnosis of rolling bearings and for estimating RUL, because of their strong feature extraction capabilities [11]. In the case of sequential data, LSTM networks are effective in dealing with long term dependencies, while GRUs provide comparable performances with less parameters [12]. In contrast to the LSTM, the LSTM units themselves are connected to each other in time, which makes them appropriate for time-series data. Yiwei Cheng et al. [13] extended the Bidirectional LSTM (BiLSTM) by processing sequences in both directions, capturing hidden long-term dependencies more effectively than unidirectional models.

More recent studies combine BiLSTM with attention-based, multi-scale, and hybrid feature-learning mechanisms to further improve accuracy. These include residual attention with multi-scale feature extraction [14], convolutional autoencoders with temporal pattern attention [15], and wavelet time-frequency representations with transformer-based multi-head attention [16]. Magadán et al. (2024) [17] proposed a stacked variational denoising autoencoder (SVDAE) combined with a BiLSTM for feature extraction and RUL estimation and validated it on the IMS bearing dataset. Zhang et al. (2025) [18] introduced a two-stage framework that integrates unsupervised anomaly detection for degradation-onset detection with a Spatiotemporal Attention BiLSTM (STA-BiLSTM) network for RUL assessment. Moreover, approaches with dual-stage attention GRU architectures have been proposed for probabilistic RUL estimation under uncertainty [19]. Furthermore, the work by Ali Saeed et al. (2025) [20] introduced a hybrid forecasting framework that employed modified multiscale permutation entropy (MMPE) to construct health indicators, jointly with a Regression Transformer, to estimate degradation levels. The experiments are validated on FEMTO and CWRU bearing datasets. Collectively, the proposed methods highlight the general concept that a combination of convolutional, recurrent, attention-based, and transformer-type architectures yields better RUL performance than any single architecture, a fact already demonstrated.

Despite these advances, there are still existing challenges. The traditional ML approaches rely on features that are manually generated and they have difficulty in capturing patterns occurring in long periods. CNNs rely on capturing local features instead of long-range dependencies. Existing models like LSTM and BiLSTM enable strong performance in sequence modeling, but they consist of only two components or a single level of convolution. To address this, this thesis proposes a Hybrid-BiLSTM framework for bearing RUL prediction. Vibration signals are segmented into overlapping windows, from which time-domain and

frequency-domain features are extracted and arranged into fixed-length sequences. The model combines a multi-scale convolutional front-end, dilated Temporal Convolutional Network (TCN) blocks for long-range causal dependencies, a Bidirectional LSTM for degradation trends in both directions, and multi-head self-attention with residual connections to emphasize the most relevant time steps. Dual global pooling (average and max) feeds a regression head that predicts RUL, trained with Huber loss, adaptive learning-rate reduction, and early stopping.

By integrating these components into a single architecture, the proposed framework aims to capture both short- and long-term degradation trends more comprehensively than existing standalone or partially combined models, providing a more reliable basis for RUL estimation in high-speed rolling element bearings.

The remainder of this thesis is organized as follows: Section 2 presents the Methodology, detailing the feature extraction process and the proposed Hybrid-BiLSTM architecture; Section 3 describes the Experimental Setup and Dataset used for evaluation; Section 4 presents the Results and Discussion, comparing the proposed model against baseline methods; and Section 5 concludes the thesis with a summary of findings and directions for future work

2. Methodology: -

As shown in Figure 1, bearing prognosis generally follows six steps: data acquisition and preprocessing, signal processing, feature extraction, feature selection, model construction, and RUL prediction. Rathore et al. [35] note that preprocessing the vibration signal improves data quality and prediction accuracy. The proposed Hybrid-BiLSTM framework follows this structure through three stages: data pipeline, model architecture, and training and evaluation.

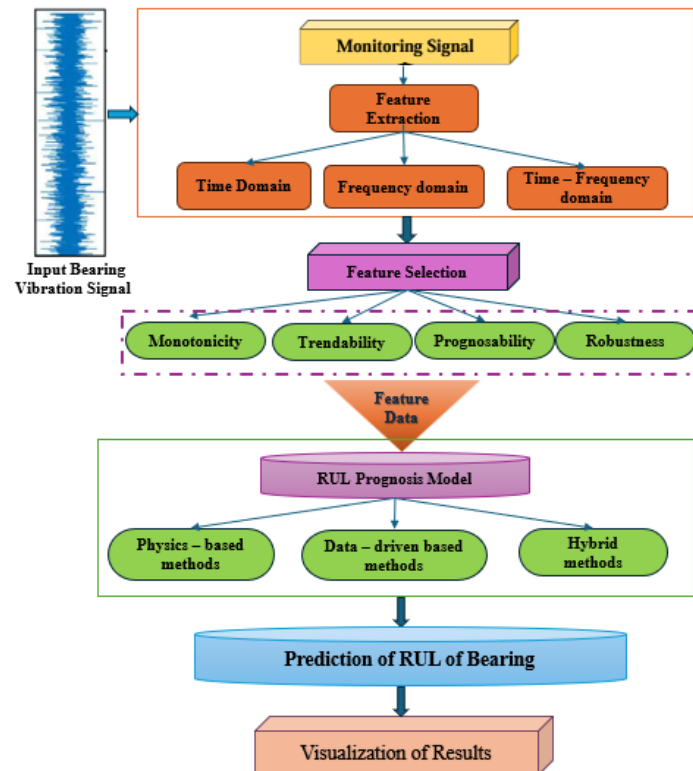


Fig. 1. RUL prediction Model

In the data pipeline stage, the raw horizontal vibration signal is segmented into overlapping windows of 2560 samples with 50% overlap. From each window, 16 hybrid features are extracted: time-domain statistics (RMS, kurtosis, skewness, peak-to-peak, crest factor), frequency-domain descriptors computed via FFT (dominant frequency, spectral kurtosis, centroid, roll-off, entropy), and bearing fault-characteristic band energies (BPFO, BPFI, FTF) with their ratios, a composite time-frequency feature, and a normalized life index. RUL is assigned as a linear countdown from the total window count. Features are standardized, arranged into sequences of length 40, and split 80/20 into training and test sets.

In the model architecture stage, illustrated in Figure 2, the sequences pass through four components. First, a multi-scale CNN front-end (kernel sizes 3, 5, 7) for local pattern extraction, second, the three dilated TCN blocks (dilation rates 1, 2, 4) for medium-range temporal dependencies, third, two stacked BiLSTM layers for bidirectional sequence modeling, and fourth, a four-head self-attention mechanism with residual connections for global context weighting. The output is aggregated via dual global (average and max) pooling and passed through a fully connected regression head, producing the final RUL prediction framework.

In the training and evaluation stage, the model is trained end-to-end using Huber loss and the Adam optimizer, with early stopping and adaptive learning-rate reduction to prevent overfitting. Performance is evaluated using MAE, RMSE, and R^2 , and compared against baseline models, SVM, Random Forest, XGBoost, LSTM, and BiLSTM. Huber loss is used to limit the influence of outliers during the stable early degradation phase, while Adam provides stable convergence despite the variable gradients arising from the model's multiple components. Combining average and max pooling preserves both the mean degradation trend and localized peak fault responses in the final feature representation. Together, these design choices make the Hybrid-BiLSTM framework a robust and computationally efficient solution for bearing RUL prediction.

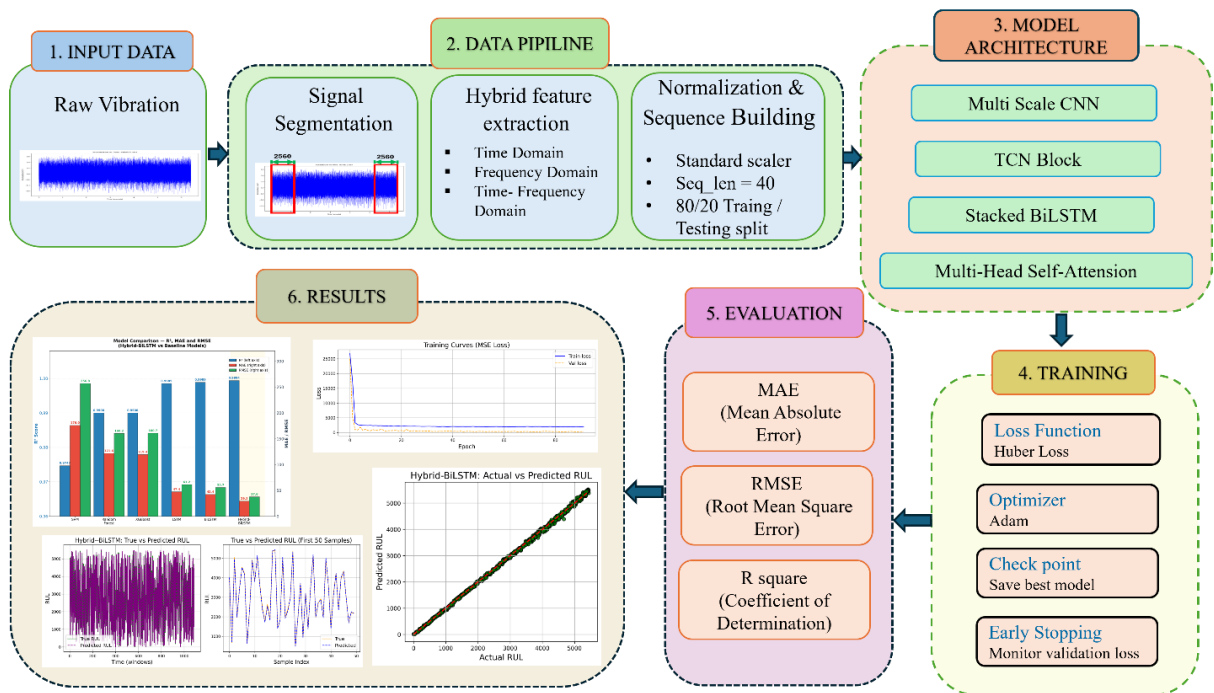


Fig.2. Proposed Hybrid-BiLSTM framework for bearing RUL prediction

3. Experimental Dataset and Setup

The dataset used in this study originates from the 2012 IEEE PHM prognostics challenge, jointly organized by the IEEE Reliability Society and the FEMTO-ST Institute, using the PRONOSTIA experimental platform (Figure 4). The platform consists of a motor-driven rotating shaft with a pneumatic load actuation system that applies radial load to test bearings, generating accelerated degradation data from initial operation through to failure. Two accelerometers (horizontal and vertical) and a temperature sensor continuously monitor each bearing during operation, with vibration data sampled at 25.6 kHz, 2560 points collected every 0.1 seconds at 10-second intervals. The dataset is publicly available on NASA's website [8] and is widely used as a benchmark for machine learning-based prognostics research.

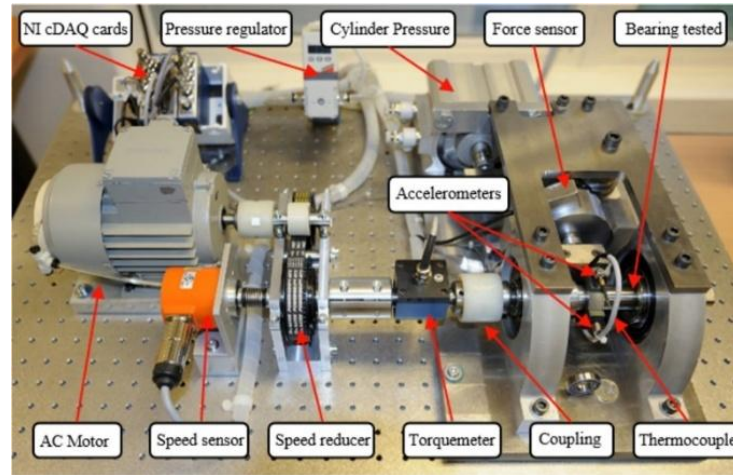


Fig. 3.. FEMTO-ST bearing test rig [21]

Bearings were tested under three load–speed conditions (1800 N–4000 rpm, 4200 N–1650 rpm, and 5000 N–1500 rpm). This study uses the Bearing1_1 dataset (4000 rpm, 1800 N), which represents accelerated degradation conditions. Since horizontal vibration signals are reported to carry more degradation information than vertical signals, only horizontal signals were used for feature extraction and model training. The bearings tested are deep groove ball bearings (type 6205); specifications are given in Table 4.1.

Table 1: Bearing Specifications [21]

Parameter	Specification
Bearing Type	Deep Groove Ball Bearing
Bearing Model	6205
Inner Diameter	25 mm
Outer Diameter	52 mm
Bearing Width	15 mm
Number of rolling elements	9 balls
Ball diameter	7.94 mm
Pitch diameter	39 mm
Contact angle	0°

4. Result and Discussion

4.1 Evaluation Parameters:

Classification metrics are employed to assess the performance of machine learning models that are developed for binary and multi-class classification tasks. These metrics quantify the model's ability to correctly assign samples to their respective class. Some key classification metrics include [35]:

(a) Mean Absolute Error (MAE):

$$MAE = \frac{1}{N} \sum_{i=1}^N |\hat{y}_i - y_i| \quad (1)$$

(b) Root Mean Square Error (RMSE):

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (2)$$

(c) Mean Squared Error (MSE):

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (3)$$

(d) R² (Coefficient of Determination):

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (4)$$

Where y_i is the actual value

$\hat{y}_i$ is the predicted value

N is the total number of data points

4.2 Evaluation parameters result:

Table 5.1 summarizes the RUL prediction performance of six models, namely SVM, Random Forest Regression, XGBoost, LSTM, BiLSTM, and the proposed Hybrid-BiLSTM, evaluated on the test set using R², MAE, and RMSE.

Among the conventional machine learning models, XGBoost achieved the best performance, with an R² of 0.99, an MAE of 119.40 (0.21%), and an RMSE of 160.74 (0.28%). Its predictions seem to be in line with the reference line of $y = x$, indicating accurate modeling of the degradation process based on the input of the hand-crafted features, instead of the learned temporal representations. Also, the sequential techniques showed evident advantage over tree-based and kernel-based techniques due to the ability to model time-dependent nature of the degradation process properly. Thus, LSTM model provided R² equal to 0.9985, while BiLSTM improved the results further to R² of 0.9989..

The proposed Hybrid-BiLSTM model beat every model used as a baseline in the experiment with an R² of 0.9994 with the lowest MAE (29.34, 0.053%) and RMSE (37.64, 0.07%) recorded across all six models. The higher performance of the model was due to its applying a

combination of multi-scale convolutional feature extraction, dilated temporal convolution for long-range dependency modeling, bidirectional sequence learning, and attention-based degradation-relevant time steps. The reduction in both MAE and RMSE relative to BiLSTM indicates that these additional architectural components reduce prediction error consistently across the test sequence rather than only on specific subsets of samples. These comparative results are also presented visually in Figure 6, which shows the R^2 , MAE, and RMSE values for all six models in bar-graph form for ease of comparison.

Table 2.: Comparison of performance metrics for the regression model

Sr. No.	Regression Model	R^2 Value	MAE		RMSE	
			E_{window}	$E\%$	E_{window}	$E\%$
1	SVM	97.47	175.96	0.31%	256.75	0.46%
2	Random Forest	99	121.64	0.21%	160.25	0.28%
3	XG boost	99	119.40	0.21%	160.74	0.28%
4	LSTM	99.85	47.35	0.08%	61.18	0.11%
5	BiLSTM	99.88	42.37	0.07%	55.71	0.09%
6	Hybrid-BiLSTM	99.94	29.34	0.053%	37.64	0.07%

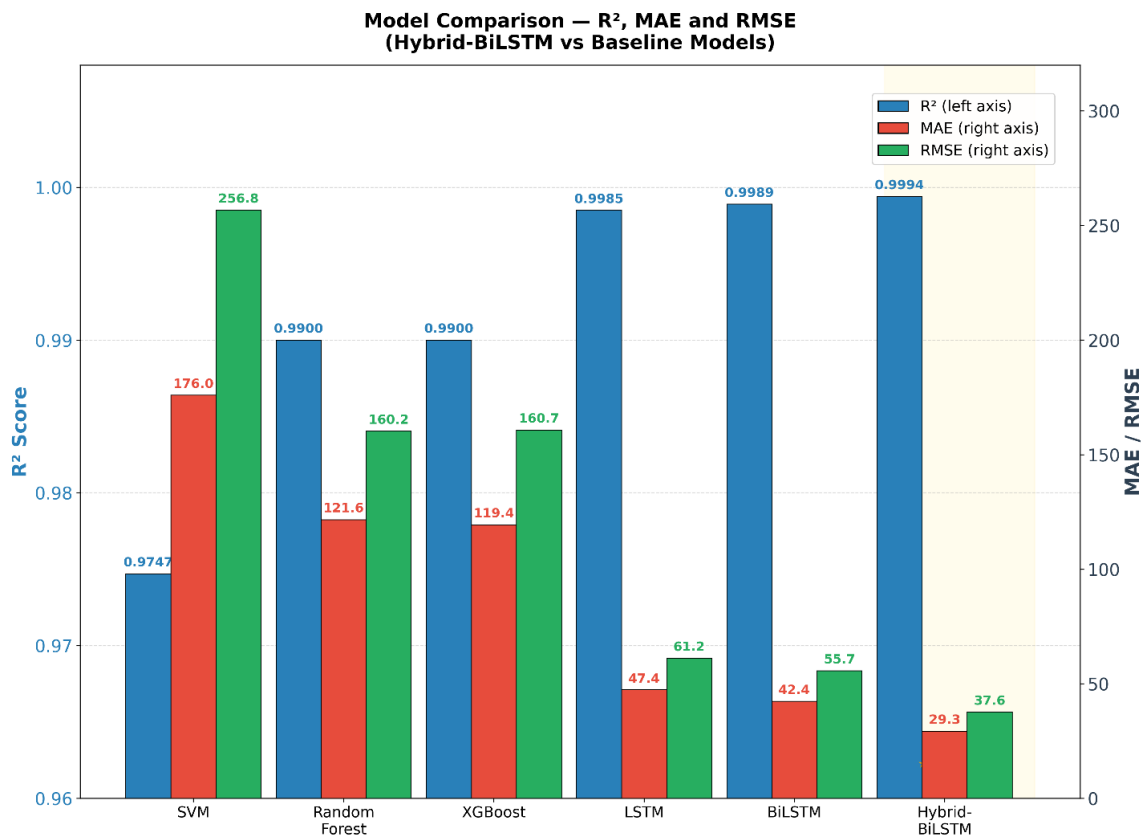


Fig. 4.. Bar plot for MAE and RMSE comparison for different modes

4.3 Remaining Useful Life Calculation (RUL)

RUL of Bearing:

$$\widehat{RUL}_{win} = N - i$$

Where:

- N = total number of windows until failure
- i = current window index
- Δt is step times (in seconds)

Then convert it to seconds:

$$\widehat{RUL}_{sec} = \widehat{RUL}_{win} \times \Delta t$$

Below are the parameters used for our training Model

- Frequency Samples (FS) = 2560 Hz
- WINDOW_SIZE = 2560 samples
- OVERLAP = 0.5
- SEQ_LEN = 30
- Life of Bearing 1_1 = 28030 seconds (from PRONOSTIA dataset)

1) First Step time (in seconds)

Step size (in samples)

$$\begin{aligned} \text{Step} &= \text{WINDOW_SIZE} \times (1 - \text{OVERLAP}) \\ &= 2560 \times (1 - 0.5) = 2560 \times 0.5 = 1280 \text{ samples} \end{aligned}$$

Step time (in seconds)

$$\Delta t = \frac{\text{step}}{FS} = \frac{1280}{2560} = 0.5 \text{ seconds}$$

- Bearing1_1 life = **28,030 seconds**
- Step time = **0.5 seconds**

2) **Second Step**: - calculate the total windows (N):

$$N = \frac{28030}{0.5} = 56060 \text{ windows (approximately)}$$

Now we want to know the remaining life before failure. For this, it is required to feed the current vibration data into the BiLSTM model, then let it decide: “This vibration pattern looks like window index X in training data. i.e X = 16,000”

Then:

$$\widehat{RUL}_{win} = N - i = 56060 - 16000 = 40000$$

Convert RUL to seconds

$$\widehat{RUL}_{sec} = 40000 \times 0.5 = 20000 \text{ seconds}$$

Convert RUL to hours

$$\widehat{RUL}_{hr} = \frac{20000}{3600} = 5.5556 \text{ hours}$$

That means **about 5 hours and 33 minutes remaining**.

4.3.1 Physical Interpretation of Prediction Errors in RUL Estimation:-

The Hybrid-BiLSTM model achieved a Mean Absolute Error (MAE) of 29.34, it is converted in seconds as

$$MAE_{sec} = 29.34 \times \Delta t = 29.34 \times 0.5 = \mathbf{14.67 \text{ seconds}}$$

Relative error:

$$\frac{14.67}{28030} \times 100 \approx 0.053\%$$

The obtained prediction error corresponds to less than 0.1% of the total bearing life, indicating excellent predictive performance. Physically, this implies that the average error in estimating the failure time of Bearing1_1 is approximately 16 seconds, which confirms the high accuracy of the proposed model for RUL prediction.

4.4 Actual vs. Predicted RUL Scatter Plot with Best Fit Line

The scatter plot of actual versus predicted Remaining Useful Life (RUL) provides a visual assessment of the prediction accuracy of the developed models. The actual-versus-predicted Remaining Useful Life (RUL) scatter plots shown in Fig. 5.5 evaluate the regression performance of the developed prognostic models. The horizontal axis represents the actual RUL values, while the vertical axis denotes the predicted RUL values. The diagonal reference line ($y = x$) represents the ideal prediction condition, where the predicted values perfectly match the actual RUL values. Figures 5.5 (a)–(f) present the prediction distributions obtained using SVM, Random Forest (RF), XGBoost, LSTM, BiLSTM, and the proposed Hybrid-BiLSTM models, respectively. Prediction accuracy is indicated by the concentration of data points around the reference line. Among all evaluated models, the proposed Hybrid-BiLSTM model exhibits the closest alignment with the ideal line and the least prediction dispersion, demonstrating superior regression accuracy and enhanced ability. Additionally, quantitative results indicate that this hybrid model has an R^2 value of 99.93%, the highest recorded among the models evaluated, and the lowest mean absolute error (MAE) and root mean square error (RMSE). This indicates that the hybrid model is the closest model relative to its ideal alignment with the reference line in terms of regression accuracy and learning complex temporal degradation features to learn complex temporal degradation characteristics

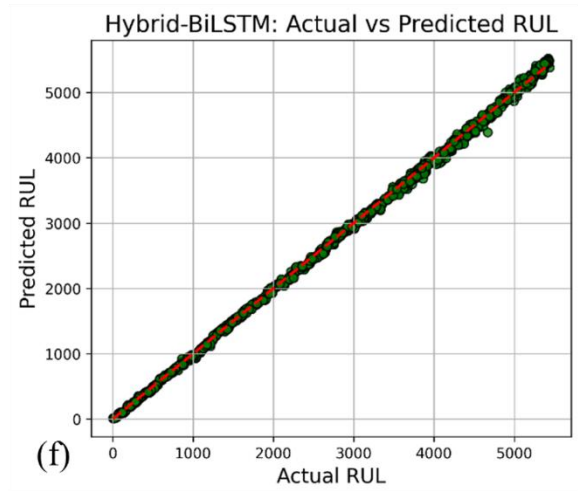
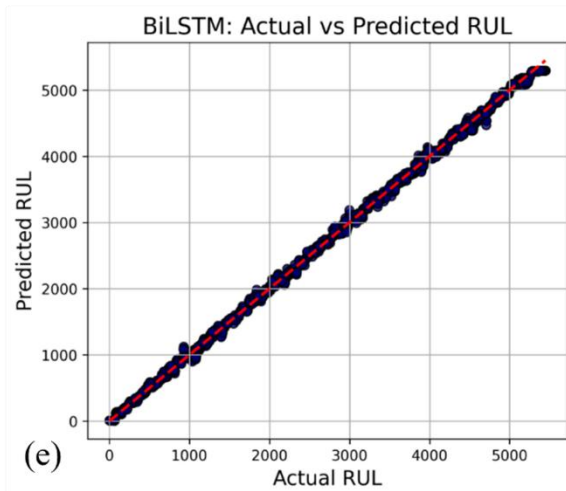
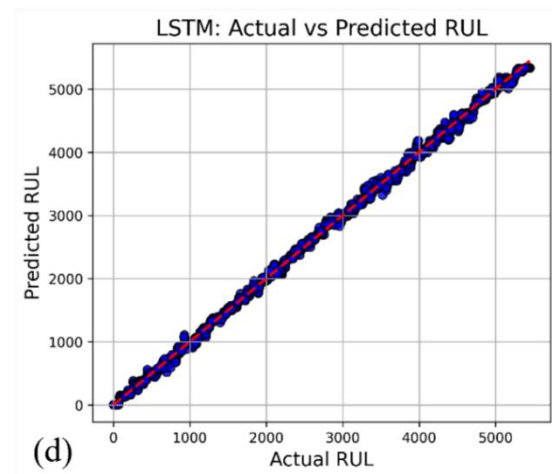
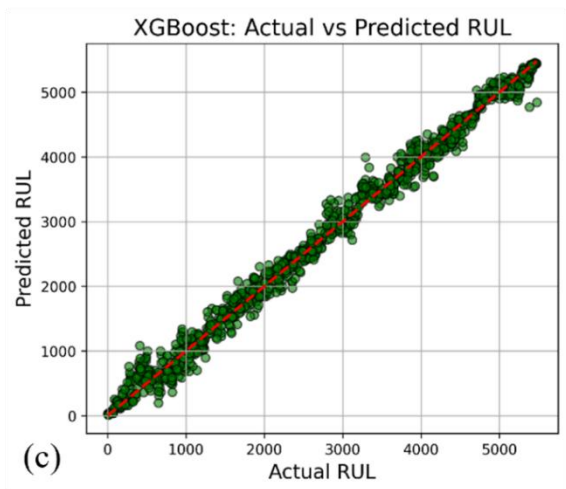
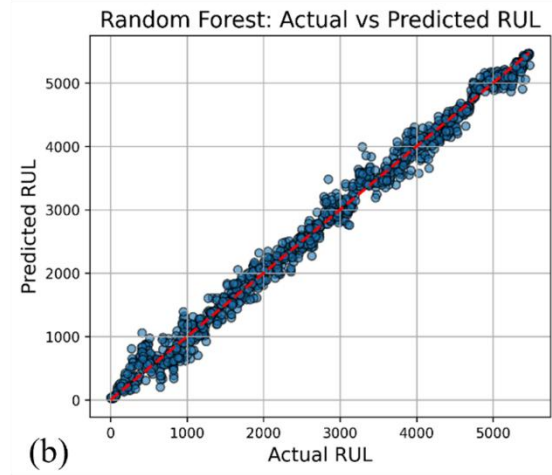
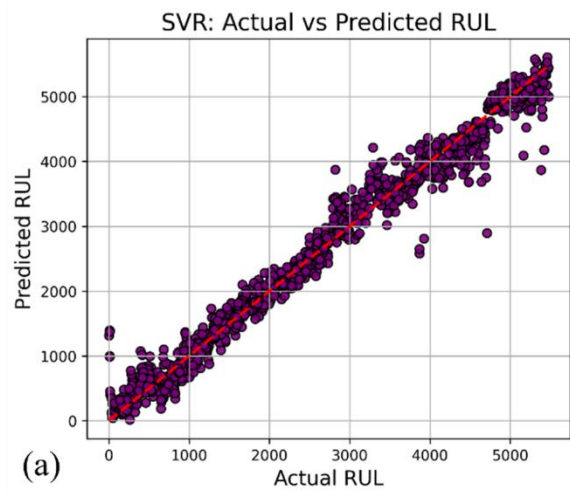


Fig. 5. Actual vs. Predicted RUL Scatter Plot with Best Fit Line (a) SVM (b) Random Forest (c) XGBoost (d) LSTM (e) BiLSTM (f) Hybrid-BiLSTM

4.5 Training MSE Curve vs Epoch for the BiLSTM Model

Further analysis of the model's learning behavior is presented through the Training MSE Curve versus Epoch for the Hybrid-BiLSTM model, as illustrated in Figure 8. This plot tracks both the training and validation MSE loss across successive epochs, offering insight into how effectively the model learns degradation-relevant features from the vibration sequences over the course of training. Both curves decline sharply during the initial epochs, reflecting rapid learning of discriminative degradation features as the model begins to capture the underlying temporal patterns in the data. As training progresses, the rate of decline gradually diminishes and the curves flatten, indicating that the model has converged and further training yields only marginal improvement. Throughout this process, the validation loss closely tracks the training loss, with no substantial divergence between the two curves at any stage.

This close correspondence between training and validation MSE indicates that the Hybrid-BiLSTM model generalizes well to unseen data rather than memorizing patterns specific to the training set, and that overfitting is minimal. The stable, low-gap convergence behavior observed in Figure 8 confirms that the proposed architecture effectively learns the temporal dynamics of degradation necessary for accurate and reliable RUL prediction.

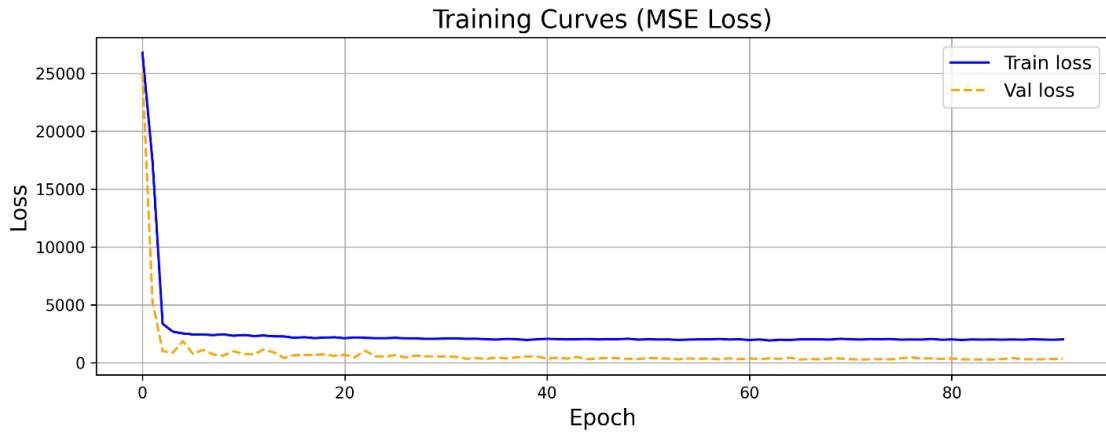


Fig. 6. Training MSE Curve vs Epoch for the Hybrid- BiLSTM Model

Conclusion

This paper presented a Hybrid-BiLSTM framework for Remaining Useful Life (RUL) prediction of rolling element bearings, integrating a multi-scale convolutional front-end, dilated Temporal Convolutional Network (TCN) blocks, bidirectional LSTM layers, and a multi-head self-attention mechanism within a single end-to-end architecture. The model was trained on hybrid time- and frequency-domain features, including bearing fault characteristic band energies, extracted from the PRONOSTIA dataset (FEMTO-ST Institute), and optimized using a Huber loss with adaptive learning-rate scheduling and early stopping.

Comparative evaluation against SVM, Random Forest, XGBoost, LSTM, and BiLSTM demonstrated that the proposed model achieves superior predictive performance, attaining the highest R^2 (0.9994) and the lowest MAE (29.34) and RMSE (37.64) among all evaluated

models. These results indicate that jointly integrating multi-scale spatial feature extraction, causal long-range temporal modelling, bidirectional sequence learning, and attention-based feature weighting yields measurable improvements in prediction accuracy over architectures that incorporate these components individually or in partial combination. The close alignment between training and validation loss curves further confirms that the model generalizes effectively without significant overfitting. Future work will focus on validating the proposed framework across multiple bearings and operating conditions within the PRONOSTIA dataset, and on evaluating its cross-domain generalization through transfer learning

References

- [1] A. Gabrielli, M. Battarra, E. Mucchi, and G. Dalpiaz, “Physics-based prognostics of rolling-element bearings: The equivalent damaged volume algorithm,” *Mech. Syst. Signal Process.*, vol. 215, Jun. 2024, doi: 10.1016/j.ymssp.2024.111435.
- [2] S. G. Kumbhar, R. G. Desavale, and N. V. Dharwadkar, “Fault size diagnosis of rolling element bearing using artificial neural network and dimension theory,” *Neural Comput. Appl.*, vol. 33, no. 23, pp. 16079–16093, Dec. 2021, doi: 10.1007/s00521-021-06228-8.
- [3] V. M. Nistane and S. P. Harsha, “Prognosis of degradation progress of ball bearings using supervised machine learning,” *Proceedings of the Institution of Mechanical Engineers, Part K: Journal of Multi-body Dynamics*, vol. 232, no. 2, pp. 183–198, Jun. 2018, doi: 10.1177/1464419317731046.
- [4] Y. Lu, Q. Li, and S. Y. Liang, “Physics-based intelligent prognosis for rolling bearing with fault feature extraction,” *International Journal of Advanced Manufacturing Technology*, vol. 97, no. 1–4, pp. 611–620, Jul. 2018, doi: 10.1007/s00170-018-1959-0.
- [5] K. Adamsab, “Machine learning algorithms for rotating machinery bearing fault diagnostics,” in *Materials Today: Proceedings*, Elsevier Ltd, 2020, pp. 4931–4933. doi: 10.1016/j.matpr.2020.12.050.
- [6] M. G. Alfarizi, B. Tajiani, J. Vatn, and S. Yin, “Optimized Random Forest Model for Remaining Useful Life Prediction of Experimental Bearings,” *IEEE Trans. Industr. Inform.*, vol. 19, no. 6, pp. 7771–7779, Jun. 2023, doi: 10.1109/TII.2022.3206339.
- [7] M. S. Rathore and S. P. Harsha, “Prognostic Analysis of High-Speed Cylindrical Roller Bearing Using Weibull Distribution and k-Nearest Neighbor,” *J. Nondestruct. Eval. Diagn. Progn. Eng. Syst.*, vol. 5, no. 1, Feb. 2022, doi: 10.1115/1.4051314.
- [8] L. Li, J. Xu, and J. Li, “Estimating remaining useful life of rotating machinery using relevance vector machine and deep learning network,” *Eng. Fail. Anal.*, vol. 146, Apr. 2023, doi: 10.1016/j.engfailanal.2023.107125.
- [9] A. Soualhi, K. Medjaher, and N. Zerhouni, “Bearing health monitoring based on hilbert-huang transform, support vector machine, and regression,” *IEEE Trans. Instrum. Meas.*, vol. 64, no. 1, pp. 52–62, Jan. 2015, doi: 10.1109/TIM.2014.2330494.

- [10] X. Yan and M. Jia, "A novel optimized SVM classification algorithm with multi-domain feature and its application to fault diagnosis of rolling bearing," *Neurocomputing*, vol. 313, pp. 47–64, Nov. 2018, doi: 10.1016/j.neucom.2018.05.002.
- [11] P. Ma, G. Li, H. Zhang, C. Wang, and X. Li, "Prediction of Remaining Useful Life of Rolling Bearings Based on Multiscale Efficient Channel Attention CNN and Bidirectional GRU," *IEEE Trans. Instrum. Meas.*, vol. 73, pp. 1–13, 2024, doi: 10.1109/TIM.2023.3347787.
- [12] B. Li, B. Tang, L. Deng, and M. Zhao, "Self-Attention ConvLSTM and Its Application in RUL Prediction of Rolling Bearings," *IEEE Trans. Instrum. Meas.*, vol. 70, 2021, doi: 10.1109/TIM.2021.3086906.
- [13] Y. Cheng, K. Hu, J. Wu, H. Zhu, and X. Shao, "A convolutional neural network based degradation indicator construction and health prognosis using bidirectional long short-term memory network for rolling bearings," *Advanced Engineering Informatics*, vol. 48, Apr. 2021, doi: 10.1016/j.aei.2021.101247.
- [14] J. Liu, Z. Yang, J. Xie, R. Wang, S. Liu, and D. Xi, "A Feature Fusion-Based Method for Remaining Useful Life Prediction of Rolling Bearings," *IEEE Trans. Instrum. Meas.*, vol. 72, 2023, doi: 10.1109/TIM.2023.3318706.
- [15] A. Rohani Bastami, A. Aasi, and H. A. Arghand, "Estimation of Remaining Useful Life of Rolling Element Bearings Using Wavelet Packet Decomposition and Artificial Neural Network," *Iranian Journal of Science and Technology - Transactions of Electrical Engineering*, vol. 43, pp. 233–245, Jul. 2019, doi: 10.1007/s40998-018-0108-y.
- [16] P. K. Sahu and R. N. Rai, "LSTM-based deep learning approach for remaining useful life prediction of rolling bearing using proposed C-MMPE feature," *Journal of Mechanical Science and Technology*, vol. 38, no. 5, pp. 2197–2209, May 2024, doi: 10.1007/s12206-024-0402-8.
- [17] L. Magadán, J. C. Granda, and F. J. Suárez, "Robust prediction of remaining useful lifetime of bearings using deep learning," *Eng. Appl. Artif. Intell.*, vol. 130, Apr. 2024, doi: 10.1016/j.engappai.2023.107690.
- [18] Z. Zhang, W. Zhang, Y. Zhao, Y. Ding, and J. Cai, "Remaining Life Prediction of Rolling Bearings Based on Unsupervised Anomaly Detection and STA-BiLSTM," *IEEE Sens. J.*, vol. 24, no. 24, pp. 41659–41674, 2024, doi: 10.1109/JSEN.2024.3487587.
- [19] Z. Cen, S. Hu, Y. Hou, Z. Chen, and Y. Ke, "Remaining useful life prediction of machinery based on improved Sample Convolution and Interaction Network," *Eng. Appl. Artif. Intell.*, vol. 135, Sep. 2024, doi: 10.1016/j.engappai.2024.108813.
- [20] A. Saeed *et al.*, "Hybrid framework for Remaining Useful Life (RUL) prediction of rolling bearing faults," *Array*, vol. 27, Sep. 2025, doi: 10.1016/j.array.2025.100498.
- [21] D. Yao, B. Li, H. Liu, J. Yang, and L. Jia, "Remaining useful life prediction of roller bearings based on improved 1D-CNN and simple recurrent unit," *Measurement (Lond.)*, vol. 175, Apr. 2021, doi: 10.1016/j.measurement.2021.109166.