

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.Doi Number

Aircraft Engines Remaining Useful Life Prediction Based on A Hybrid Model of Autoencoder and Deep Belief Network

Huthaifa AL-Khazraji¹, Ahmed R. Nasser¹, Ahmed M. Hasan¹, Ammar K. Al Mhdawi², Hamed Al-Raweshidy³, Amjad J. Humaidi¹

¹Control and Systems Eng. Dept., University of Technology- Iraq

²Department of Computer Science, Edge Hill University, UK

³Department of Electrical and Electronic Engineering, Brunel University London, UK

Corresponding author: Ammar K. Mhdawi (e-mail: Al-Mhdaa@edgehill.ac.uk).

ABSTRACT Remaining Useful Life (RUL) is used to provide an early indication of failures that required performing maintenance and/or replacement of the system in advance. Accurate RUL prediction offers cost-effective operation for decision-makers in the industry. The availability of data using intelligence sensors leverages the power of data-driven methods for RUL estimation. Deep Learning is one example of a data-driven method that has a lot of applications in the industry. One of these applications is the RUL prediction where DL algorithms achieved good results. This paper presents an Autoencoder-based Deep Belief Network (AE-DBN) model for Aircraft engines' RUL estimation. The AE-DBN DL model is utilized the feature extraction characteristic of AE and superiority in learning long-range dependencies of DBN. The efficiency of the proposed DL algorithm is evaluated by comparison between the proposed AE-DBN and the state-of-the-art related method for RUL prediction for four datasets. Based on the Root Mean Square Error (RMSE) and Score indices, the outcomes reveal that the AE-DBN RUL prediction model is superior to other DL approaches.

INDEX TERMS Artificial Intelligence, Deep Learning, Remaining Useful Life, Autoencoder, Deep Belief Network, Aircraft engine.

I. INTRODUCTION

Remaining Useful Life (RUL) is used to ensure the safety and reliability of the aircraft equipment. It is used to decide whether to perform maintenance or not and how many spare parts should be ordered such as the overall maintenance cost is reduced [1]. If the lifetime of the aircraft equipment cannot be known with certainty, it is recommended to keep monitoring the life and state of operating equipment. Hence, monitoring of the equipment gives the required information on the current working age and state by measuring certain variables that may affect its operation work [2].

Existing research on the methods of the RUL estimation of the aircraft equipment can be generally grouped into physics-based methods and data-driven methods. In the physics-based approach, the RUL estimation model is developed using damage propagation to identify potential failure for the

equipment. However, the complexity of the damage propagation as well as the uncertainty of the operating environment makes it extremely difficult to identify the potential failure mechanism for many components and systems. On the other side, data-driven methods employ data collected from sensors that are integrated with equipment to develop the RUL prediction model [3]. Aircraft are now fully integrated with advanced sensors that continually collect information regarding the operation condition of the aircraft. This data helps for the transition from physics-based methods to data-driven methods for RUL estimation of the aircraft equipment [4]. The power of data-driven methods for RUL estimation is arisen due to the rapid development of Internet-of-Things (IoT) and cyber physic technique which provides a massive amount of data. This data allows possibilities for Artificial Neural Network (ANN) methods such as Deep

Learning (DL) to be applied [5]. In this aspect, the RUL prediction can be considered as time series regression problem.

DL utilizes multiple layers of neurons to learn complex models. Recently, DL has had a lot of applications in the industry [6] [7]. One of these applications is the RUL prediction where DL algorithms achieved good results. For example, authors in [8] developed an extended recurrent neural network (ERNN) algorithm for the RUL prediction model using vibration data collected from a gearbox experimental system. Deutsch and He in [9] proposed Restricted Boltzman Machine (RBM) algorithm for the RUL prediction model using data collected from bearing run-to-failure tests. Authors in [10] employed Long Short-Term Memory (LSTM) with Recurrent Neural Network (RNN) (LSTM-RNN) method for the RUL prediction model of the capacity degradation trajectories of lithium-ion batteries. As shown above, DL methods have received a huge interest in RUL prediction for various applications. However, the prediction performance of these methods depends on the features, dependencies, and dimensionality of the input time-series data that is used to impellent the DL models for RUL prediction.

One of the advantages of Autoencoder (AE) DL architecture is its ability to capture and extract useful features from the input data. On the other hand, Deep Belief Network (DBN) architecture has the capabilities of deep hierarchical learning which allows it to capture the long-range dependence characteristic to learn sophisticated features from the input data. These features of DBN compared with the traditional shallow learning approaches yield improvement in prediction accuracy while keeping the prediction system simple.

Owing to feature extraction characteristic of AE and superiority in learning long-range dependencies of DBN, this paper presents a combined Autoencoder and Deep Belief Network (AE-DBN) based model for RUL estimation. In the proposed architecture the AE model is used to extract the abstract features from the input data while a Deep Belief Network model is used as a predictor for the time series RUL of the aircraft engine.

The remainder of this paper is organized according to the following. Section 2 investigates the related works regarding DL methods that are applied to perform an RUL prediction for aircraft. The proposed hybrid DL approach for RUL estimation is described in Section 3. The experiment study and comparisons with other DL methods using the C-MAPSS dataset are provided in Section 4. Finally, the paper concludes in Section 5.

II. RELATED WORKS

With the help of advanced sensors that are integrated with the aircraft equipment, the data that represents the status of the aircraft becomes available. This data allows accurate RUL data-driven prediction models for an aircraft based on DL methods to be applied [4]. This section aims to review the

related work of DL methods that are applied to aircraft for RUL prediction.

The early promising results to apply DL algorithms into RUL prediction of the engine of the aircraft were found by [11]. The Authors proposed Convolutional Neural Network (CNN) based regression approach for estimating the RUL. In this work, the structure of the CNN incorporates an automated feature learning from raw sensor signals. Later, the authors of [12] developed long short-term memory (LSTM) for estimating the RUL of the engine aircraft data. By comparing the result obtained from the LSTM with the results obtained from CNN, LSTM shows better performance.

The authors in [13] introduced a data-driven approach for RUL prediction based on Deep Convolution Neural Networks (DCNN). Raw collected data for the aero-engine unit is used to show the effectiveness of the proposed approach. The work in [14] used a stacked Sparse Autoencoder (SAE) for RUL prediction. The hyperparameters of the SSAE were determined based on the grid search method. The authors in [15] presented a combined DL algorithm for RUL prediction to predict multiple multivariate time series of the data collected from aircraft sensors was performed by LSTM based recurrent network. Moreover, Deep Belief Network (DBN) was utilized to assess system working conditions and categorize faults of aircraft equipment. In the same direction of using hybrid methods, the authors in [16] proposed a directed acyclic graph (DAG) network. This model combines long short term memory (LSTM) with the convolutional neural network (CNN) to predict the RUL. Based on the features of the training data collected from aircraft sensors, the method proposed in [17] utilized a modified Denoising Autoencoder (DAE) for robust feature extraction. The authors integrated an Updated Selection Strategy (USS) to ensure that the valuable data is passed through the training process. In terms of trucking the new rising data and forgetting gradually the old ones, Online Sequential Extreme Learning Machine (OS-ELM) was proposed with double dynamic forgetting factors (DDFF). Unlike the methods presented in [17], authors in [18] used LSTM to learn sequential features collected from aircraft sensors. Then, an attention mechanism with a feature fusion method is proposed for RUL estimation. A combined CNN with Bi-directional Long Short-Term Memory (BDLSTM) networks is presented in [19] for RUL prediction for aircraft. The CNN was used to extract spatial features while BDLSTM was utilized to extract temporal features. To improve the capability of the combination of CNN with LSTM, the work presents in [20] proposed a double-channel hybrid prediction model based on the CNN and a bidirectional LSTM network (CNN +BiLSTM). Besides, the work in [21] presented a new model combining between the parallel branches of CNN in series with LSTM. The model named multi-head CNN-LSTM. Recently, the authors in [22] proposed a new deep learning model combines between the

transformer encoder and temporal convolution neural network (Trans.+TCNN). The transformer encoder is used a scaled dot-product attention to extract dependencies across distances in time series, whereas the temporal convolution neural network is built to fix the insensitivity of the self-attention mechanism to local features.

Owing to feature extraction characteristic of AE and superiority in learning long-range dependencies of DBN, this paper presents a combined Autoencoder and Deep Belief Network (AE-DBN) based model for RUL estimation. In the proposed architecture the AE model is used to extract the abstract features from the input data while a Deep Belief Network model is used as a predictor for time series RUL of aircraft engine.

To improve the prediction accuracy and the performance of aircraft engines RUL prediction a combined deep learning architecture based on AE for feature extraction and DBN for RUL time series prediction is proposed in this work.

III. PROPOSED DEEP LEARNING ARCHITECTURE

This section presents the description of the proposed Autoencoder-based Deep Belief Network (AE-DBN).

A. ADEEP BELIEF NETWORKS

The Deep Belief Networks (DBN) can be viewed as a combination of simple unsupervised networks such as the Restricted Boltzmann Machine (RBM), which act as the hidden layer of each subnet and the visible layer of the next layer. The layer in the DBN structure has an efficient layer-by-layer procedure that determines how variables depend on variables in the layer above [23].

The developed DBN network consists of multiple visible and hidden RBM layers and logistic regression for classification in the last layer. In the first step of the development process, different feature spaces of the vectors are mapped, after that each layer of the RBM network is trained in an unsupervised way, respectively, to preserve the feature information of all. In the second step, fine adjustments are made. In the last step, the output feature vector of the RBM is taken as the input feature vector for the next RBM. The architecture of DBN is shown in Figure 1 [24].

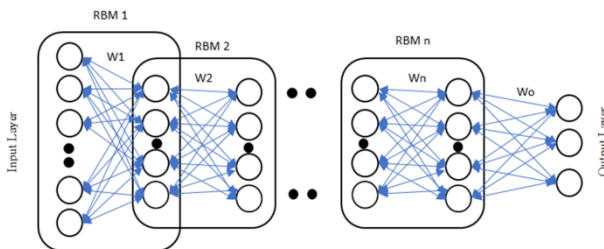


FIGURE 1. DBN architecture.

In the RBM model, the v_i in the visible layer and each hidden layer are represented by h_i . Weights between v_i and

h_j are guided and displayed with w_{ij} . Visible and hidden nodes have biases represented by vectors c and b . The b_i , c_i and w_{ij} values of all RBMs in the model form the θ parameter in the DBN. This parameter θ appears in the model with the probability of common states of the hidden layers and an energy function [25]. This energy function is given as in Eq.1.

$$E(\theta, v, h) = -\sum_{i=1}^m v_i c_i - \sum_{j=1}^n h_j b_j - \sum_{i=1}^m \sum_{j=1}^n v_i h_j w_{ij} \quad (1)$$

Since there is no inter-layer connection in the DBN network, the probability distributions of the visible and hidden layers are calculated as given in Eq. 2 and Eq. 3

$$P(v_i = 1|h) = 1/1 + e^{-b_i \sum h_j w_{ij}} \quad (2)$$

$$P(h_i = 1|v) = 1/1 + e^{-c_i \sum v_j w_{ij}} \quad (3)$$

After the weight calculations are completed, the reconstructed data is returned and can be determined by the $p(v|h)$ calculation. Output σ occurs when data is transmitted back to the hidden layer. Here, the logistic function σ is defined as in Eq. 4.

$$\sigma(x) = (1 + e^{-x})^{-1} \quad (4)$$

Likewise, in the case of $v_i = 1$, the conditional probability of v_i is calculated as given in Eq.5.

$$P(v_i = 1|v) = \sigma(a_i + \sum_{j=1}^n W_{ij} h_j) \quad (5)$$

B. DEEP AUTOENCODERS

Autoencoder (AE) is a three-layer unsupervised neural network. It is considered the most basic form of neural networks, which is used for representation learning such as feature selection or size reduction and tries to reconstruct the input patterns in the output layer [26].

The general feature of AE is that the input and output layer size is the same as with symmetrical architecture. The hidden layer in the network model typically contains fewer neurons compared to the visible layer. By using a small number of neurons, an attempt is made to encode or represent the input in a more compact way, capturing the meaningful properties of the input vectors. Figure 2 shows a basic AE and deep AE (DAE) network architecture [27].

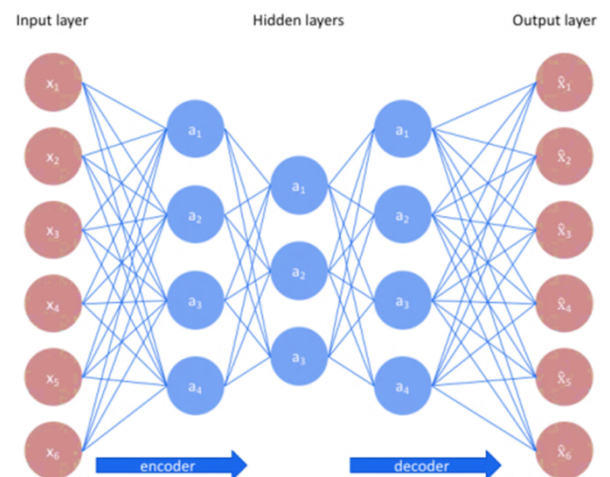


FIGURE 2. AE architecture.

Training of AE is carried out by applying a backpropagation algorithm as in feedforward neural networks based on Mean-Squared Error (MSE) loss function. The training process consists of two stages, coding, and decoding. In the coding phase, while trying to encode the inputs in a hidden representation by using the weight criteria of the lower half layer. In the decoding phase, the same input is tried to be reconstructed from the coding representation using the criteria of the upper half layer. Therefore, the encoding and decoding weights are forced to be transposed of each other. Consider X is data with n samples and m features. Y is the output of the encoder (i.e. reduced representation of X). The mathematical representations of the encoding and decoding operations for a basic AE are given in Eq. 6 and Eq.7, respectively [28].

$$Y = f(wX + b) \tag{6}$$

$$\hat{X} = g(\hat{w}Y + b) \tag{7}$$

In Eq. 6 and Eq. 7, w and b represent the adjustable parameters, f and g are the activation function, \hat{w} is the weights (w) transpose, and \hat{X} is the reconstructed input vector in the output layer.

Training an autoencoder involves finding parameters w and b that minimizes the error between the input data X and the reconstruction data \hat{X} .

C. THE PROPOSED AE-DBN ARCHITECTURE FOR RUL PREDICTION

The proposed AE-DBN architecture is shown in Figure 3. It is consisting of two main parts. In the first part, the AE is used as a deep learning model for feature extraction from the input data. The second part is represented by a deep learning model based on DBN for predicting the RUL.

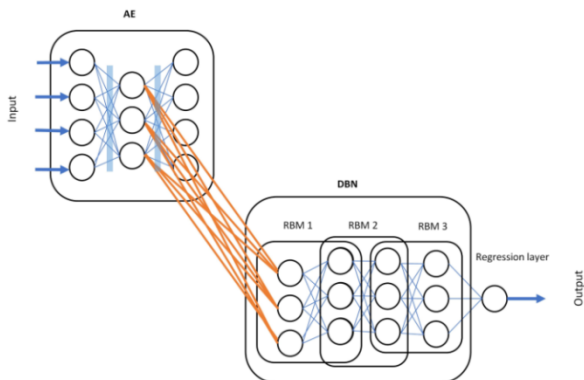


FIGURE 3. AE-DBN architecture

The encoder part of the AE is responsible for extracting the features which represent the characteristics of the input data. These extracted features are feed to the DBN part of the proposed architecture which is used to predict the RUL. Initially, the AE is trained separately to obtain the weight matrix before training the DBN RUL prediction model. The decoder part of AE is used to verify the extracted features' validity to reconstruct the original data. Then the obtained

weight matrix for AE is combined with the DBN model which is finally trained using the input data for RUL prediction.

IV. EXPERIMENTAL STUDY

A. DATASET

The widely used dataset for predictive maintenance in aircraft health systems in the literature is NASA Turbofan Engine Corruption Simulation dataset [29]. The dataset was created by NASA engineers using commercial simulation software called C-MAPSS. A turbofan engine with a capability of producing 90,000 pounds of power is simulated at altitudes of 0-40,000 ft., Mach number range of 0-0.9 and ambient temperatures range of -60 to 103 degrees F. The attributes of aircraft engines used for experimentation were the engine core speed, engine fan speed, pressure at fan inlet, High-Pressure Turbine (HPT) exit temperature, pressure at the High-Pressure Compressor (HPC), and engine-pressure ratio. To monitor these engine's attributes, a total of 21 onboard sensors monitoring temperature, pressure, and speed are distributed at various locations as shown in Figure 4.

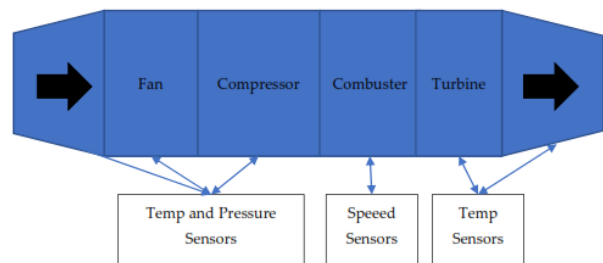


FIGURE 4. AIRCRAFT ENGINE AND SENSORS

The C-MAPSS software also has various regulators and limiters that prevent the engine from being taken out of the operating range specified by the manufacturer.

During creating the dataset, the engine was operated together with the control system. When the health index of the motor dropped to zero, the simulation was stopped and the obtained sensor data were recorded as a time series. The engine health index is determined to be 1 for each of the parts that make up the engine, and 0 when it goes out of the specified operating conditions. While the training data was continued until the motor health index was 0, the test and verification data were terminated before the motor broke down in order to measure RUL. The difference between the cycles in which the engine health index falls to zero at any given moment gives the RUL value of the engine.

Four different datasets are considered which are prepared for different operating conditions and different scenarios in the dataset. The FD001 and FD003 contain a single operation condition and fault type for 100 engines. The FD002 includes six operation conditions and 260 engines. The FD004

includes six operation conditions and 249 engines. The datasets samples contain the records of sensors data at each run to fail cycles for each engine which is used to train the model for predicting the RUL. The details of each dataset used to train and evaluate the proposed method for RUL prediction are shown in Table I below.

TABLE I. EVALUATION DATASETS DESCRIPTION

Dataset	FD001	FD002	FD003	FD004
Operating conditions	1	6	1	6
Fault modes	1	1	2	2
Train samples (number of engines)	100	260	100	248
Test samples (number of engines)	100	259	100	249

B. MODEL DESIGN

The proposed AE-DBN architecture contains different key structure variables such as the number of input nodes, the number of hidden layers for the AE model, the number of RBM stages for the DBN model, the number of hidden layers for the DNB model, and the model's hyperparameters. The key structure features that are considered for designing the proposed model are obtained through experimental trials where the objective is achieving the best RUL prediction performance. The parameters for each model are summarized in Table II and Table III.

TABLE II. AE MODELS PARAMETERS

Hyperparameter	Value
Hidden layers	5
Neurons	21,10,10,10,21
Number of Epochs	500
Loss function	MSE
Activation function	ReLU

TABLE III. DBN MODELS PARAMETERS

Hyperparameter	Value
Number of layers	3 RBMs
Hidden layers per RBM	10
Batch size	64
Window with	50 samples
Optimizer	SGD
Epochs	250
Learning rate	0.001

C. PERFORMANCE EVALUATION

In order to verify the effectiveness of the proposed AE-DBRN method for RUL prediction, the model is trained and evaluated using the four datasets (FD001, FD002, FD003 and FD004) explained in the previous section. To avoid model overfitting the datasets are split into 60% for training and 40% for testing samples. A comparison is carried out between the proposed AE-DBRN and the state-of-the-art

related method for RUL prediction such as DBN, CNN, LSTM, and attention-based LSTM. Commonly used evaluation criteria such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE), R^2 and Score are considered to evaluate the model performance in terms of RUL prediction accuracy [30]. These metrics are calculated as given in Eq.8, Eq.9, Eq. 10 and Eq.11 respectively.

$$RMSE = \sqrt{\sum_1^N (a_i - p_i)^2 / N} \quad (8)$$

$$MAE = \sum_1^N |p_i - a_i| / N \quad (9)$$

$$R^2 = 1 - \frac{\sum_1^N (a_i - p_i)^2}{\sum_1^N (a_i - p_{mean})^2} \quad (10)$$

$$Score = \begin{cases} \sum_{i=1}^N \left(e^{-\frac{a_i - p_i}{13}} - 1 \right) & \text{for } a_i < p_i \\ \sum_{i=1}^N \left(e^{\frac{a_i - p_i}{10}} - 1 \right) & \text{for } a_i \geq p_i \end{cases} \quad (11)$$

where N is the number of samples, a_i is the actual value and p_i is the predicted value and p_{mean} is the mean of the predicted values.

The proposed AE-DBN model is trained and evaluated using FD001, FD002, FD003 and FD004 datasets separately using the parameters explained in the model design section. Based on the selected window length the model predicts the RUL of a particular engine based on the previous 50 samples of sensors readings. To demonstrate the effectiveness of AE for extracting the features that provide a better representation of the data, the proposed AE-DBN is compared with the regular DBN method that feeds directly with the original data. The evaluation results are shown in Table IV.

Based on the achieved results for the proposed AE-DBN model and the regular DBN model it was found that adding the AE model for features extraction enhanced the overall model performance in terms of accuracy since it was used to reduce the original data features to extract the best respective features. The proposed AE-DBN model achieved better performance using the FD001 and FD003 dataset compared to the FD002 and FD004 dataset since it contains more samples and includes multiple operation conditions which can strain the model in terms of learning.

TABLE IV. EXPERIMENTAL RESULTS USING FD001 AND FD004 DATASETS

Dataset	Method	RMSE	MAE	Score	R^2
FD001	AE-DBN	11.2700	11.9100	219	0.9545
	DBN	13.4500	14.1900	228	0.9405
FD002	AE-DBN	14.2400	14.8500	1255	0.9411
	DBN	17.5500	19.1500	1379	0.9120
FD003	AE-DBN	11.1300	11.4800	264	0.9513
	DBN	12.3200	13.2500	285	0.9452
FD004	AE-DBN	26.8508	27.3347	2135	0.8999
	DBN	28.5444	28.9798	2147	0.8896

The predicted and the actual RUL using the AE-DBN model on the four datasets for each engine using the previous 50 samples of sensors readings are shown in Figures (5-8).

Based on results shown in Figure (5-8) of the predicted RUL for each engine matches very well with the actual

RUL, which indicates the viability of the proposed AE-DBN model for the RUL prediction.

For example based on the FD001 dataset, it can be noticed that the RMSE, MAE and Score are decreased from 13.45, 14.19 and 228 in the case of the standard DBN to 11.27, 11.91 and 219 for the proposed AE-DBN whereas the R^2 is increased from 0.9405 in the case of the standard DBN to 0.9545 for the proposed AE-DBN. For FD002 dataset, it can be noticed that the RMSE, MAE and Score are decreased from 17.55, 19.15 and 1379 in the case of the standard DBN to 14.24, 14.85 and 1255 for the proposed AE-DBN whereas the R^2 is increased from 0.9120 in the case of the standard DBN to 0.9411 for the proposed AE-DBN. For FD003 dataset, it can be noticed that the RMSE, MAE and Score are decreased from 12.32, 13.25 and 285 in the case of the standard DBN to 11.13, 11.48 and 264 for the proposed AE-DBN whereas the R^2 is increased from 0.9452 in the case of the standard DBN to 0.9513 for the proposed AE-DBN.

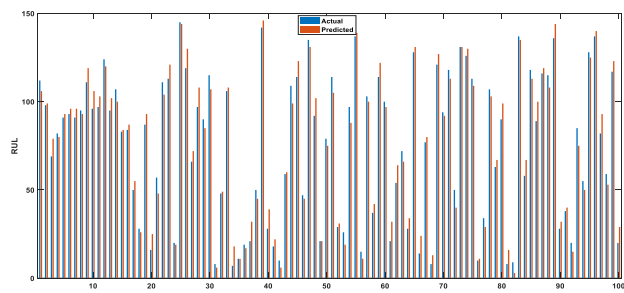


FIGURE 5. Actual RUL and the predicted RUL for FD001 dataset.

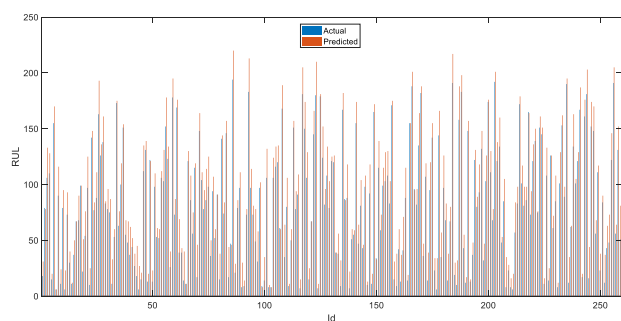


FIGURE 6. Actual RUL and the predicted RUL for FD002 dataset.

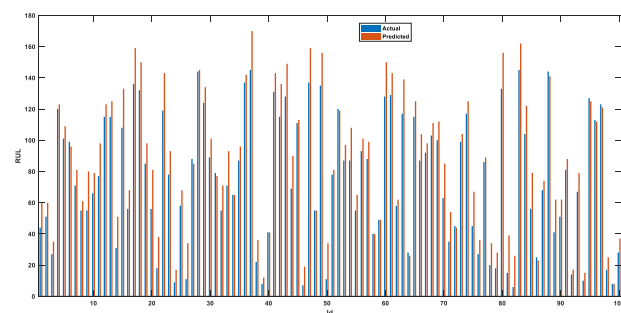


FIGURE 7. Actual RUL and the predicted RUL for FD003 dataset.

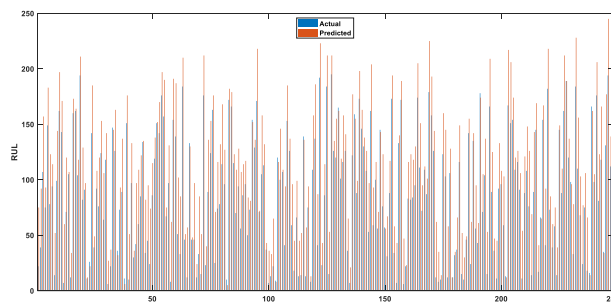


FIGURE 8. Actual RUL and the predicted RUL for FD004 dataset.

In the same way for the FD004 dataset, it can be noticed that the RMSE, MAE and Score are decreased from 28.5444, 28.9798 and 2147 in the case of the standard DBN to 26.8508, 27.3347 and 2135 for the proposed AE-DBN whereas the R^2 is increased from 0.8896 in the case of the standard DBN to 0.8999 for the proposed AE-DBN.

In addition, the proposed AE-DBN performance in terms of RMSE for RUL prediction is compared with a similar state-of-the-art method in the literature that used the FD001, FD002, FD003 and FD004 dataset as shown in Table V.

TABLE V. RMSE COMPARISON BETWEEN THE PROPOSED METHOD AND RELATED WORKS USING FD001, FD002, FD003 AND FD004 DATASET

Methods	RMSE			
	FD001	FD002	FD003	FD004
CNN [11]	18.45	30.29	19.82	29.16
LSTM [12]	16.14	24.49	16.18	28.17
DAG network [16]	11.96	20.34	12.466	22.66
Attention-based LSTM [18]	14.53	-	-	27.08
CNN +BiLSTM [20]	12.5	19.34	12.1	20.03
MHCNN-LSTM [21]	13.27	19.49	13.21	23.89
Trans. + TCNN [22]	12.31	15.35	12.32	18.35
AE-DBN (proposed method)	11.27	14.24	11.13	26.85

Based on the comparison results in Table V for the FD001 dataset, it can be seen that the proposed model reduces the RMSE from 18.45 in the case of regular CNN, 16.14 in the case of regular LSTM, 14.53 in the case of Attention-based LSTM, 13.27 in the case of MHCNN + LSTM, 12.5 in the case of BiLSTM + MSCNN, 12.31 in the case of Trans. + TCNN, 11.96 in the case of DAG network to 11.27 for the AE-DBN DL model. For FD002, the RMSE reduces from 30.29 in the case of CNN, 24.49 in the case of LSTM, 20 in the case of DAG network, 19.49 in the case of MHCNN+LSTM, 19.34 in the case of BiLSTM+MSCNN, 15.35 in the case of Trans.+TCNN to 14.24 for the proposed AE-DBN DL model.

For FD003, the RMSE reduces from 19.82 in the case of CNN, 16.18 in the case of LSTM, 13.21 in the case of MHCNN+LSTM, 12.466 in the case of DAG network,

12.32 in the case of Trans.+TCNN, 12.1 in the case of BiLSTM+MSCNN to 11.13 for the proposed AE-DBN DL model.

For the FD004 dataset, it can be seen that the proposed model reduces the RMSE from 29.16 in the case of the regular CNN, 27.08 in the case of Attention-based LSTM to 26.85 for the proposed AE-DBN DL model. However, DAG network, BiLSTM+MSCNN, MHCNN+LSTM, Trans.+TCNN have shown better result than the proposed model.

To end this, the proposed AE-DBN performance in terms of Score for RUL prediction is compared with a similar state-of-the-art method in the literature that used the FD001, FD002, FD003 and FD004 dataset as shown in Table VI.

TABLE VI. SCORE COMPARISON BETWEEN THE PROPOSED METHOD AND RELATED WORKS USING FD001, FD002, FD003 AND FD004 DATASET

Methods	Score			
	FD001	FD002	FD003	FD004
CNN [11]	1286	13570	15962	7886
LSTM [12]	338	4450	852	5550
DAG network [16]	229	2730	535	3370
Attention-based LSTM [18]	322	-	-	5649
CNN +BiLSTM [20]	231	2650	257	3400
MHCNN+LSTM [21]	259	4350	343	4340
Trans. + TCNN [22]	252	1267	296	2120
AE-DBN (proposed method)	219	1255	264	2135

Based on the comparison results in Table VI for the FD001 dataset, it can be seen that the proposed model reduces the Score from 1286 in the case of regular CNN, 338 in the case of regular LSTM, 322 in the case of Attention-based LSTM, 259 in the case of MHCNN + LSTM, 231 in the case of CNN + BiLSTM, 252 in the case of Trans. + TCNN, 229 in the case of DAG network to 219 for the AE-DBN DL model. For FD002, the Score reduces form 13570 in the case of CNN, 4450 in the case of LSTM, 2730 in the case of DAG network, 4350 in the case of MHCNN+LSTM, 2650 in the case of CNN + BiLSTM, 1267 in the case of Trans.+TCNN to 1255 for the proposed AE-DBN DL model.

For the FD003 dataset, it can be seen that the proposed model reduces the Score from 15962 in the case of regular CNN, 852 in the case of regular LSTM, 535 in the case of DAG network, 343 in the case of MHCNN + LSTM, 296 in the case of Trans. + TCNN to 264 for the AE-DBN DL model. However, CNN + BiLSTM have shown better result than the proposed model with a Score of 257. For the FD004 dataset, it can be seen that the proposed model reduces the Score from 7886 in the case of regular CNN, 5550 in the case of regular LSTM, 4340 in the case of MHCNN + LSTM, 3400 in the case of CNN + BiLSTM, 3370 in the case of DAG network, to 2135 for the AE-DBN

DL model. However, Trans. + TCNN have shown better result than the proposed model with a Score of 2120.

In general, this comparison proved that the proposed AE-DBN outperforms the similar related works based on the three dataset out of four based on the RMSE and two out of four based on the Score index.

V. CONCLUSION

Deep Learning (DL) based methods have been proven to be very promising for the Remaining Useful Life (RUL) prediction of equipment. This paper proposed an Autoencoder-based Deep Belief Network (AE-DBN) based model for RUL prediction for aircraft engines. An experimental study was conducted using a published dataset of aircraft engines data to evaluate the effectiveness of the proposed RUL prediction model. To investigate the estimation performance, the AE-based DBN is compared with the standard DBN model. The results show a considerable improvement of the AE-DBN in comparison with standard DBN in terms of RMSE, MAE, R² and Score. Moreover, the results are also compared with other DL algorithms. The results show that a three out of four dataset (FD001, FD002 and FD003), the RMSE of the proposed AE-DBN model is less than other state-of-the-art related method for RUL prediction. Besides, two out of four (FD001 and FD002), the Score of the proposed AE-DBN model is less than other state-of-the-art related method for RUL prediction. The overall results reveal that the AE-DBN RUL prediction model outperforms the state-of-the-art works and the standard DBN RUL prediction model. As future research, this work can be extended into two directions. The first one is to explore the ability to utilize the swarm-based optimization techniques to determine the optimal hyperparameters of the DL model to achieve higher accuracy and less implementation complexity. On other hand, hybrid with other DL algorithms to boost the performance of the RUL prediction can be considered as another future work.

REFERENCES

- [1] C. HU, H. PEI, Z. WANG, X. SI, and Z. Zhang, 2018. A new remaining useful life estimation method for equipment subjected to intervention of imperfect maintenance activities. *Chinese Journal of Aeronautics*, 31(3), pp.514-528.
- [2] D. Banjevic, 2009. Remaining useful life in theory and practice. *Metrika*, 69(2-3), pp.337-349.
- [3] Z. Tian, 2012. An artificial neural network method for remaining useful life prediction of equipment subject to condition monitoring. *Journal of Intelligent Manufacturing*, 23(2), pp.227-237.
- [4] D. Rengasamy, H.P. Morvan, and G.P. Figueredo, 2018, November. Deep learning approaches to aircraft maintenance, repair and overhaul: a review. In 2018 21st International Conference on Intelligent Transportation Systems (ITSC) (pp. 150-156). IEEE.
- [5] L. Ren, J. Cui, Y. Sun, and X. Cheng, 2017. Multi-bearing remaining useful life collaborative prediction: A deep learning approach. *Journal of Manufacturing Systems*, 43, pp.248-256.
- [6] Q. Wang, S. Zheng, A. Farahat, S. Serita, and C. Gupta, 2019, June. Remaining useful life estimation using functional data

- analysis. In 2019 IEEE International Conference on Prognostics and Health Management (ICPHM) (pp. 1-8). IEEE.
- [7] A. Nasser and H. Al-Khazraji, 2022. A hybrid of convolutional neural network and long short-term memory network approach to predictive maintenance. *International Journal of Electrical & Computer Engineering*, 12(1), pp. 2088-8708
- [8] Z. Tian, and M. J. Zuo, 2010. Health condition prediction of gears using a recurrent neural network approach. *IEEE Transactions on Reliability*, 59(4), pp.700-705.
- [9] J. Deutsch and D. He, 2016. Using deep learning based approaches for bearing remaining useful life prediction. In *Annual Conference of the PHM Society (Vol. 8, No. 1)*.
- [10] Y. Zhang, R. Xiong, H. He, and Z. Liu, 2017, July. A LSTM-RNN method for the lithium-ion battery remaining useful life prediction. In *2017 Prognostics and System Health Management Conference (PHM-Harbin)* (pp. 1-4). IEEE.
- [11] G. S. Babu, P. Zhao, and X. L. Li, 2016, April. Deep convolutional neural network based regression approach for estimation of remaining useful life. In *International conference on database systems for advanced applications* (pp. 214-228). Springer, Cham.
- [12] S. Zheng, K. Ristovski, A. Farahat, and C. Gupta, 2017, June. Long short-term memory network for remaining useful life estimation. In *2017 IEEE International Conference on Prognostics and Health Management (ICPHM)* (pp. 88-95). IEEE.
- [13] X. Li, Q. Ding, and J. Q. Sun, 2018. Remaining useful life estimation in prognostics using deep convolution neural networks. *Reliability Engineering & System Safety*, 172, pp.1-11.
- [14] J. Ma, H. Su, W. L. Zhao, and B. Liu, 2018. Predicting the remaining useful life of an aircraft engine using a stacked sparse autoencoder with multilayer self-learning. *Complexity*, 2018.
- [15] C. Che, H. Wang, Q. Fu, and X. Ni, 2019. Combining multiple deep learning algorithms for prognostic and health management of aircraft. *Aerospace Science and Technology*, 94, p.105423.
- [16] J. Li, X. Li, and D. He, 2019. A directed acyclic graph network combined with CNN and LSTM for remaining useful life prediction. *IEEE Access*, 7, pp.75464-75475.
- [17] T. Berghout, L. H. Mouss, O. Kadri, L. Saïdi, and M. Benbouzid, 2020. Aircraft engines remaining useful life prediction with an adaptive denoising online sequential extreme learning machine. *Engineering Applications of Artificial Intelligence*, 96, p.103936.
- [18] Z. Chen, M. Wu, R. Zhao, F. Guretno, R. Yan, and X. Li, 2020. Machine remaining useful life prediction via an attention-based deep learning approach. *IEEE Transactions on Industrial Electronics*, 68(3), pp.2521-2531.
- [19] I. Remadna, S. L. Terrissa, R. Zemouri, S. Ayad, and N. Zerhouni, 2020, May. Leveraging the Power of the Combination of CNN and Bi-directional Lstm Networks for Aircraft Engine RUL Estimation. In *2020 Prognostics and Health Management Conference (PHM-Besançon)* (pp. 116-121). IEEE.
- [20] C. Zhao, X. Huang, Y. Li, and M. Yousaf Iqbal, 2020. A double-channel hybrid deep neural network based on CNN and BiLSTM for remaining useful life prediction. *Sensors*, 20(24), p.7109.
- [21] H. Mo, F. Lucca, J. Malacarne, and G. Iacca, 2020, September. Multi-Head CNN-LSTM with prediction error analysis for remaining useful life prediction. In *2020 27th Conference of Open Innovations Association (FRUCT)* (pp. 164-171). IEEE.
- [22] H.K. Wang, Y. Cheng, and K. Song, 2021. Remaining Useful Life Estimation of Aircraft Engines Using a Joint Deep Learning Model Based on TCNN and Transformer. *Computational Intelligence and Neuroscience*, 2021.
- [23] H. B. Huang, X. R. Huang, R. X. Li, T. C. Lim, and W. P. Ding, 2016. Sound quality prediction of vehicle interior noise using deep belief networks. *Applied Acoustics*, 113, pp.149-161.
- [24] K. P. Lin, P. F. Pai, and Y. J. Ting, 2019. Deep belief networks with genetic algorithms in forecasting wind speed. *IEEE Access*, 7, pp.99244-99253.
- [25] M. Jiang, Y. Liang, X. Feng, X. Fan, Z. Pei, Y. Xue, and R. Guan, 2018. Text classification based on deep belief network and softmax regression. *Neural Computing and Applications*, 29(1), pp.61-70.
- [26] C. Zhou, and R. C. Paffenroth, 2017, August. Anomaly detection with robust deep autoencoders. In *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 665-674).
- [27] T. Amarbayasgalan, B. Jargalsaikhan, and K. H. Ryu, 2018. Unsupervised novelty detection using deep autoencoders with density based clustering. *Applied Sciences*, 8(9), p.1468.
- [28] Q. Meng, D. Catchpoole, D. Skillicom, and P. J. Kennedy, 2017, May. Relational autoencoder for feature extraction. In *2017 International Joint Conference on Neural Networks (IJCNN)* (pp. 364-371). IEEE.
- [29] A. Saxena, K. Goebel, D. Simon, and N. Eklund, 2008, October. Damage propagation modeling for aircraft engine run-to-failure simulation. In *2008 international conference on prognostics and health management* (pp. 1-9). IEEE.
- [30] D. Chicco, M. J. Warrens, and G. Jurman, 2021. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, 7, p.e623.