A hybrid of convolutional neural network and long short-term memory network approach to predictive maintenance

Ahmed Nasser, Huthaifa Al-Khazraji
Department of Control and System Engineering, University of Technology Iraq, Baghdad, Iraq

Article Info	ABSTRACT
	Predictive maintenance (PdM) is a successful strategy used to reduce cost by minimizing the breakdown stoppages and production loss. The massive amount of data that results from the integration between the physical and digital systems of the production process makes it possible for deep learning (DL) algorithms to be applied and utilized for fault prediction and diagnosis. This paper presents a hybrid convolutional neural network based and long short-term memory network (CNN-LSTM) approach to a predictive maintenance problem. The proposed CNN-LSTM approach enhances the predictive accuracy and also reduces the complexity of the model. To evaluate the proposed model, two comparisons with regular LSTM and gradient boosting decision tree (GBDT) methods using a freely available dataset have been made. The PdM model based on CNN-LSTM method demonstrates better prediction accuracy compared to the regular LSTM, where the average F-Score increases form 93.34% in the case of regular LSTM to 97.48% for the proposed CNN-LSTM. Compared to the related works the proposed hybrid CNN-LSTM PdM approach achieved better results in term of accuracy.

Corresponding Author:

Ahmed Nasser
Department of Control and System Engineering, University of Technology Iraq
Baghdad, Iraq
Email: ahmed.r.nasser@uotechnology.edu.iq

1. INTRODUCTION

Maintenance is an essential part of modern manufacturing systems [1]. The complexity of manufacturing machines that are consisted of multiple interdependent components is required an advanced maintenance model [2]. The massive amount of data that results of the integration between the physical and digital systems of the production process makes it possible for deep learning (DL) algorithms to be applied for fault prediction and diagnosis. Machines are now equipped with sensors that continuously collect information about their status. Predictive maintenance (PdM) is one type of maintenance strategy that is applied to production equipment based on an estimate of the status of that piece of equipment [3]. The estimation aims to avoid breakdowns and maximizing the service life of the equipment [4]. By connecting the devices with the sensors via the internet of things (IoT), the data of the devices can be used to identify patterns that lead to discovering failures before it happens [5]. Hu *et al.* [6] applied wireless sensor networks and IoT for online monitoring the industrial equipment and utilizes it for PdM.

Deep learning (DL) is a subclass of artificial intelligence such as machine learning (ML) and a branch of artificial neural networks (ANN) that significantly affects human's life. DL in associated with advanced technologies (i.e. IoT, big data, cloud computing, and 3D printing) are formulated Industrial 4.0 [7]. It is considered a data-driven method. Research on the applications of DL has become an essential tool to the development of systems that monitor the current state of machines.

There is a reasonably of related works on the application of DL and ML based models to PdM. For example, Sarkar *et al.* [8] formulated deep autoencoder (DAE) to describe the damage in the form of cracks on a composite material. Chen *et al.* [9] proposed a recurrent neural network (RNN) approach to design effective mechanical state prediction model systems. Tian *et al.* [10] developed a condition-based maintenance (CBM) model based on an extended RNN (ERNN) to gearboxes using the vibration data that is collected from a gearbox experimental system. Huuhtanen and Jung [11] applied a convolutional neural network (CNN) for monitoring the process of photovoltaic panels. Abbasi *et al.* [12] introduced DL algorithm based on RNN using long short-term memory (LSTM) to design a predictive maintenance of air booster compressor motor.

Hrnjica *et al.* [5] presented a PdM model using machine learning technique named gradient boosting decision tree GBD to predict the failing component in multicomponent mechanical systems. The model is built and evaluated using Microsoft PdM dataset and the results shows the model obtained and 94.56% of average accuracy. Rivas *et al.* [13] presented a deep learning model for estimating remaining useful life (RUL) for industrial equipment and using it for PdM. The model is based on recurrent neural networks RNN. The model is evaluate based on historical sensors data obtained from engines. The model is produced an accuracy results of 86% of F-score. Ruiz-Gonzalez *et al.* [14] proposed a machine learning approach based on support vector machine (SVM) to predict the failures in rotating machines based on vibration signals. The model is used for Agro-Industrial Machinery and evaluated using vibration signal the model is achieved an f-score accuracy of 85%.

Hwang *et al.* [15] used machine learning model based on SVM to identifies the machines state whether it work normally or abnormally via classifying the sensor data collected form the machines using IoT. For testing scenarios, they used data collected according to the movement time of the crane. The model reached a result of 81% in term of accuracy. Rahhal *et al.* [16] presented a system for PdM based on IoT and deep learning to predict the failure time of equipment. The PdM model is built using LSTM and RNN deep neural networks and used to predict the RUL for light bulbs with a minimum error rate of 0.79%. Bampoula *et al.* [17] utilized LSTM autoencoders deep learning method to build a model for planning maintenance in Cyber-Physical Production Systems. The autoencoder deplaning model is used to classify classifying real-world machine status based on the machine's sensor data. The model is evaluated using data collected from steel industry production process. The model obtained an average accuracy of 94.2%.

Deep learning methods such RNN have demonstrated its effectiveness in various time series classification applications. LSTM method have the advantage over regular RNN method when working with time series data having long-term dependencies due the LSTM memory mechanism [18]. CNN deep learning method is mainly used for image classification tasks. The main characteristic of CNN is the multiple stacked layer construction structure, which leads to effective representation of the input data features. Therefore, CNN method has the advantages of capturing and extracting the features form the data more effectively [19].

Deep learning methods can be combined together in hybrid form to utilize the strength of each method and these types of hybrid deep learning method are receiving increasing interest in different machine learning application, due to higher actuary can be achieved when combining different deep learning together [20]. The advantages of CNN for features extraction and the effectiveness of LSTM for time series data classification, motivate us to propose a hybrid convolutional neural network and long short-term memory network (CNN-LSTM) approach which is based on our knowledge have not been applied for predictive maintenance problem. The model utilizes a published dataset of historical sensor data and machine status in order to perform predictive maintenance of production equipment. The proposed hybrid CNN-LSTM PdM model is evaluated and compared with regular LSTM PdM model and the related works in term of prediction accuracy. The evaluation results show the effectiveness of the proposed hybrid CNN-LSTM model for PdM due to its higher perdition accuracy.

The rest of the paper is arranged as follows: in section 2 the details of PdM is described. In section 3 the proposed architecture for PdM is presented. A case study is performed in section 4 to evaluate the performance of the proposed PdM architecture. Finally, section 5 illustrates the conclusion of the paper.

2.   PREDICTIVE MAINTENANCE

A maintenance policy is a policy that is used to maintain production machines in an acceptable operating condition. There are mainly three types of maintenance policies: corrective (breakdown) maintenance, preventive (scheduled) maintenance, and predictive (condition-based) maintenance. Corrective maintenance is performed only after the occurrence of failures, whereas preventive maintenance are carried out based on a planned schedule time. On other hand, unlike the preventive maintenance approach which can be considered as a time-based approach, predictive maintenance is a condition-based approach performed based on an estimate of the working status of the production machine [3]. Predictive maintenance (PdM) is

considered as one of the most powerful and widely-used maintenance policy. According to Hashemian [21] PdM is selected as maintenance policy in 89% of the industrial cases in comparison to 11% of cases for other time-based maintenance policies.

It is evolved due to advanced technology such as advanced sensors, advances in computing, IoT, ANN, and data-driven modelling. Traditionally, eyes, ears, and noses were used as signs that the equipment begins to fail. However, advanced sensors are now available to identify equipment degradations and failures [21]. This leads to reducing the need for frequent maintenance such as periodic and preventive maintenance [12]. A predictive maintenance model for a multi-component production system should take into account the known state and the degradation threshold of each component [2].

3.   PROPOSED DEEP LEARNING ARCHITECTURE

Deep learning (DL), as a subclass of artificial intelligence, has emerged as a powerful tool for developing intelligent algorithms in many applications [22]. DL is a technique inspired by the human nervous system and the structure of the brain [23]. With the ability to handle high dimensional and multivariate data, DL becomes an attractive methodology by practitioners for PdM applications. With the number of layers and neurons increased, the ability to unsupervised learning of more complex problems is increased [24]. Examples of DL algorithms are deep neural networks (DNN), convolutional neural network (CNN), deep belief networks (DBN), and recurrent neural networks (RNN). However, the performance of DL algorithms depends on the appropriate choice of the DL technique for a given problem. Therefore, this section presents a model based on hybrid convolutional neural network and long short-term memory network (CNN-LSTM) approach to identify possible failures. The proposed CNN-LSTM approach improves the predictive accuracy from raw data and also reduces the complexity of the model [25], [26]. The CNN and LSTM networks have been studied separately. This paper aims to combine the advantages of the two approaches and applied to PdM problem.

3.1. Convolutional neural networks (CNN)

NN is a popular deep neural network (DNN) that takes this name based on the mathematical linear operation between matrixes called convolution. CNN is multiple layers fully connected layers. The most beneficial aspect of CNN is reducing the number of parameters [27], [28]. It was first introduced by LeCun with the LeNet-5 architecture in the early 1980s [29]. This architecture consists of an input layer, several convolutional layers, pooling, and output layers. The output layer of the architecture can be bonded to fully connected layers or classifier layers such as the sigmoid layer. Since it may be multimedia data such as image, sound, video, it is preferred by researchers working in many signal processing fields because of its high performance. The general structure of CNN is shown in Figure 1.

To reduce the margin of error, a strategy by using a backpropagation algorithm is implemented. This strategy adjusts the CNN architectures to update their learning weights with a margin of error throughout the training process [30]. The details of the CNN algorithm used in this work are described in Algorithm 1.

Algorithm 1: CNN model
```
Input: x input features vector
F filter with size k x d
Output: ć output features vector
1 For i=1 to N
```
2 $w_i = [x_i, x_{i+1}, \ldots x_{i+k-1}]$
3 $c_i = ReLU(w_i \odot F)$
```
4 End
```
5 $ć = Maxpool(c)$

3.2. Long short-term memory (LSTM)

LSTM is a type of RNN network for sequence learning tasks. It does not have long-term time dependency problems where due to the sequential nature of sensor data, information can be reminded for long periods of time [31]. It can be defined as an iterative neural network algorithm that can learn long-term dependencies. Standard iterative algorithms such as neural network consisting of long-term information and is designed to avoid dependency problems remembering at this time. One of the main problems in RNN structure comes from vanishing gradients over time. Since the learning network created during the training becomes complex, and the backward weight values of the network are updated as a result of zero or close to zero values, an update cannot be performed and training may stop. This problem of backward coherence in RNN structure is presented as a solution by accompanying a memory cell RNN structure in LSTM structure. With this memory cell, information from the previous time can be taken and transferred to the next [32], [33]. These units in the LSTM network remember long or short time periods. The values kept to be reminded in

these units do not interact in any way or experience change and disappear. Figure 2 shows the structure of the LSTM unit. In the LSTM unit structure shown in the Figure 2: the input X(t) takes the current input value, h(t-1) takes the previous hidden state and c(t-1) takes the previous memory state values. The output h(t) generates the current latent state and c(t) the current memory state. Algorithm 2 describes the LSTM algorithm used in this work.
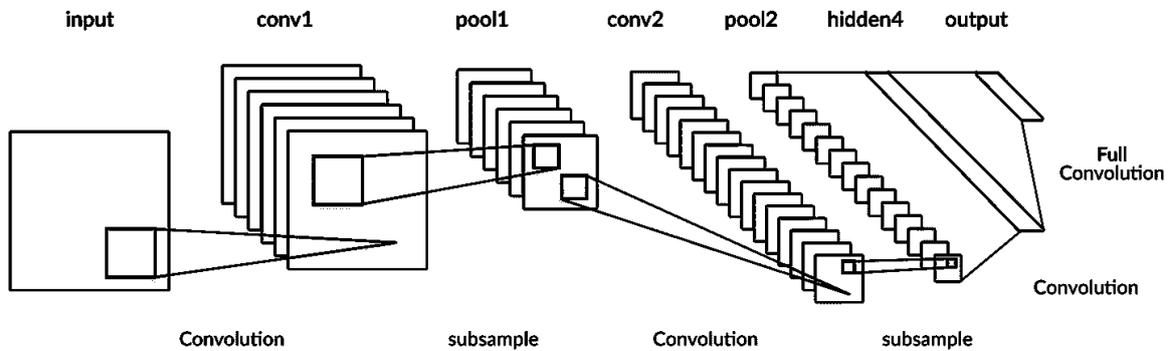
Figure 1. CNN structure
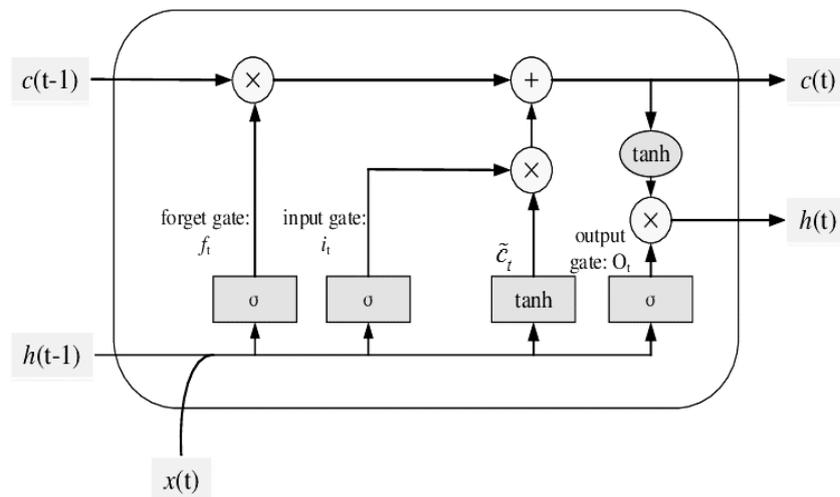
Figure 2. LSTM structure

Algorithm 2: LSTM model
```
Input: features vector ć
Output: h vector
1 For j=1 to t
2 i_j = σ(W_i.[h_{t-1} + b_i])
3 f_j = σ(W_f.[h_{t-1}, x_t] + b_f)
4 q_j = tanh(W_q[h_{t-1}, x_t] + b_q)
5 o_j = σ(W_o[h_{t-1}, x_t] + b_o)
6 c_j = f_j⊙c_{j-1} + i_j⊙q_j
7 h_jo_j = tanh⊙(c_j)
8 End
```

4. CASE STUDY

In this section, the explanation of applying the CNN-LSTM into published real-world industrial machines is presented. The description of the dataset, data preparation required for running machine learning algorithms, and model evaluation is given in the following subsections.

4.1. Dataset

The data that is used to build and evaluate the proposed hybrid deep learning model for predictive maintenance is obtained from Microsoft's GitHub repository [5]. This time-series data was generated by recording several events of different 100 machines based on an hourly per day for one year period. These events comprise the historical data regarding telemetry, machines, errors, and failures, which are described as follows:
− Telemetry: consist form data obtained by logging real-time machines sensors reading such as voltage, and rotation speed
− Machines: describe the machine's basic information such as model and age
− Errors: consist of machine error data that occurred before the machine failure
− Failures: contains information of the machine components that were replaced due to a failure.

4.2. Data preparation

Figure 3 shows the process of preparing the datasets that are used for building and testing the proposed Pdm model. The training and testing datasets are built as follows. The four time-series data files obtained from Microsoft's GitHub repository telemetry, machines errors, and failure data. These time series are combined together based on DateTime and machine_id attributes. Another attribute is calculated and added to represent the number of machine errors before the failure occurrence. Data prepossessing techniques such as data cleaning and normalization are applied. The target attribute for the Pdm model decision is chosen to be the failed component. The target attribute contains five categories that are used to describe the machine state for the next hour (four of them used to describe the component failure and others used to describe the normal state).

The obtained dataset consists of 876445 instances where each instance denotes the machine events, statutes, and failure state for one hour. The total number of dataset instances represents the operating period of one year for the machines. The dataset is split into two unique datasets. The first one is used to train the dataset which consists of 70% of the total 876445 instances. The second one is used to test the dataset which consists of the remaining 30% of the total instances. Table 1 shows the attributes and their descriptions for the generated training and dataset.
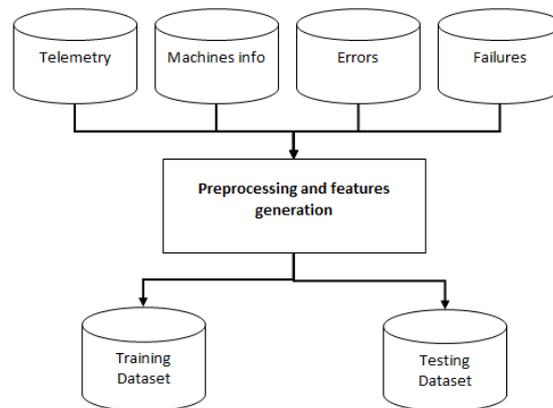
Figure 3. Dataset preparation

Table 1. Features of the generated dataset

Feature	Description
1	Time stamp (Date Hour)
2	Machine ID (1-100)
3-7	Sensors values (voltage, rotation speed, pressure and vibration)
8-9	Machine information (Model and age)
10	Error ID (Error1, Error2, Error3, Error4, Error5)
11	Number of errors before last failure (number)
12	Failure ID (Normal, Comp1, Comp2, Comp3, Comp4)

4.3. PdM hybrid CNN-LSTM model construction

The general structure of the proposed hybrid CNN-LSTM model for PdM is shown in Figure 4. This structure takes the advantages of CNN and LSTM by combining them together. CNN deep learning method is well-known for its robustness and effectiveness for extracting features from the data, while LSTM method

is effective and powerful for classifying time series. The structure of the hybrid CNN-LSTM model is described as follows: The model accepts 9 input features shown in Table 1 and output decision for the machine failure state for the next hour based on the machine state for the last 3 hours. The algorithm of the proposed Hybrid CNN-LSTM model for PdM is described in Algorithm 3.
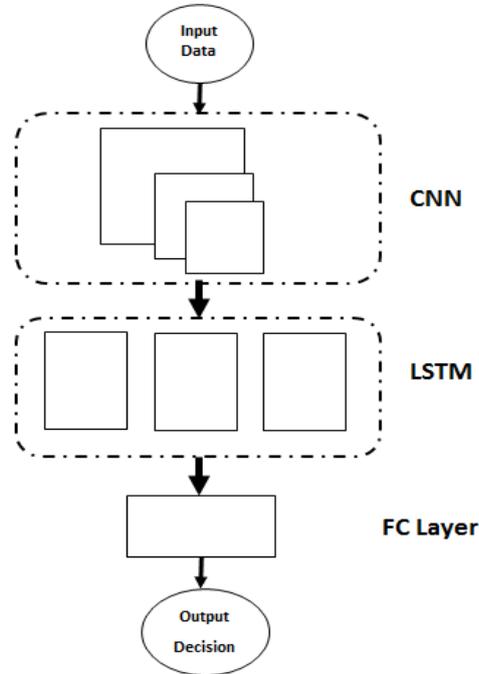
Figure 4. Proposed architecture of a hybrid CNN-LSTM for predicative maintenance

Algorithm 3: Hybrid CNN-LSTM model

```
Input X set x features
Output: Y prediction
1 Foreach x in X
2 Cₓ=CNN(Xₓ)
3 End
4 Foreach x in C
5 Oₓ=LSTM(Cₓ)
6 End
7 Foreach x in O
8 Yₓ=sigmoid(Oₓ)
9 End
```

For CNN model the 2-dimensional convolution layer is used. This layer has a filter size of 64 and kernel size of 3 and uses rectified linear unit (ReLU) as an activation function. This layer is followed by a 2-dimensional maxpooling layer with a pool size equal to 2, and then a dropout layer is added to overcome overfitting issues. The output from this layer is projected and feed to the LSTM model. On other hand, LSTM model is consisting of 100 hidden layer uses ReLU as activation function window size of 3 and categorical_crossentropy as loss function. The output of the LSTM model is feed into a fully connected layer with Softmax activation. This layer is responsible for providing the final classification decision for a given input to the model. The details of the proposed hybrid CNN-LSTM model for PdM are shown in Figure 5.

For CNN model the 2-dimensional convolution layer is used. This layer has filter size of 64 and kernel size of 3 and used rectified linear unit (ReLU) as activation function. This layer is followed by 2-dimensional maxpooling layer with pool size equal to 2, and then a dropout layer is added to overcome overfitting issues. The output from this layer is projected and feed to the LSTM model. On other hand, LSTM model is consisting from 100 hidden layer uses ReLU as activation function window size off 3 and categorical_crossentropy as loss function. The output of the LSTM model is feed into a full connected layer with Softmax activation. This layer is responsible of providing the final classification decision for a given input to the model.
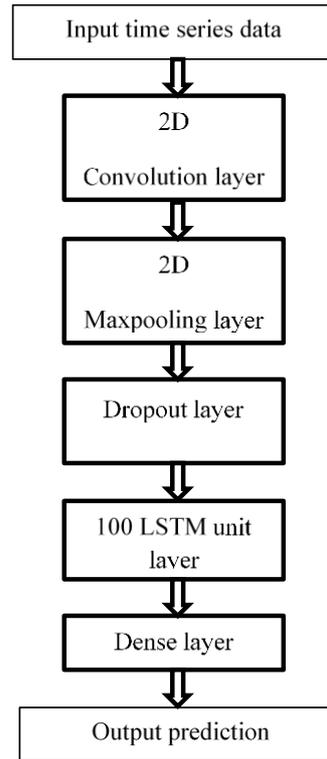
Figure 5. The details of the proposed hybrid CNN-LSTM model

4.4. CNN-LSTM PdM model training and evaluating

The proposed CNN-LSTM model for PdM is trained using the training dataset generated in the previous steps. The model is implemented in the TensorFlow platform [34] with GPU acceleration is active. The hyper-parameters for the model are used based on the values shown in Table 2. After the training of CNN-LSTM model is completed, the model is evaluated using the testing dataset prepared earlier. Since the training data is unbalanced, the performance of the model is evaluated based on the model accuracy metrics precision, recall, and F-score as given in (1), (2) and (3). These metrics are calculated based on the confusion matrix shown in Table 3 as follows:

Precision measures the exactness of the model and can be calculated by (1):

$$Pression \ = \ TP \ / \ (TP + FP) \qquad\qquad (1)$$

Recall measures the completeness of the model and can be calculated as shown in (2):

$$Recall \ = \ TP \ / \ (TP + FN) \qquad\qquad (2)$$

F-score is the weighted average of Recall and Precision. It's wildly used to measure the model accuracy in where imbalanced data used for training. F-score can be calculated as in (3):

$$F - score \ = \ (2 * Pression * Recall) / \ (Pression + Recall) \qquad\qquad (3)$$

Table 2. Model hyperparameters

Hyperparameter	Value
Learning rate	0.01
Hidden layers	100
Activation function	ReLU/softmax
Loss function	categorical_crossentropy
Dropout	0.2
Epoch	10
Batch size	128
Optimizer	SGD

Table 3. Model confusion matrix

	Predicted as True	Predicted as False
Actual True	TP	FN
Actual false	FP	TN

The evaluation results of the CNN-LSTM PdM model based on the performance metrics are shown in Table 4 which are calculated using test dataset. Since the dataset used for training the model consist of multiple classes for prediction, the evaluation metrics were calculated for each class separately then the overall model performance is calculated by averaging all results for the classes. To investigate the effectiveness of the proposed deep learning method for PdM, the regular LSTM deep learning method is trained and evaluated using the same dataset used for the CNN-LSTM model. The evaluation results of LSTM PdM model are shown in Table 5.

Table 4. Evaluation results of the PdM CNN-LSTM model

Failure Type	Precision	Recall	F-Score (%)
Comp1	0.97	0.93	94.96
Comp2	0.95	0.98	96.48
Comp3	0.97	0.99	97.99
Comp4	0.99	0.98	98.49
Normal	1.00	0.99	99.47
Average	0.976	0.974	97.48

Table 5. Evaluation results of the PdM LSTM model

Failure Type	Precision	Recall	F-Score (%)
Comp1	0.92	0.93	92.5
Comp2	0.91	0.94	92.48
Comp3	0.94	0.91	92.48
Comp4	0.93	0.92	96.00
Normal	0.98	0.96	96.99
Average	0.936	0.932	93.34

Based on the evaluation results of the PdM CNN-LSTM model, it can be observed that the model is able to predict the machine state in terms of specifying which component is going to be failed or whether the machine will work without any failure for the next hour with high accuracy of 97.48%. Comparing the results of the hybrid CNN-LSTM model to LSTM model for PdM, it shown that using the hybrid model increases the average prediction accuracy by 4.44%. The results of the LSTM and hybrid CNN-LSTM are compared to the results of related works in term of accuarcy. The results of each model are shown in Table 6. The results shows that the proposed hybrid CNN-LSTM achieved higher prediction accuracy compared to the other realed PdM works.

Table 6. Compaction between the CNN-LSTM, LSTM and related works for PdM

Method	Average F-Score (%)
Gradient boosting decision tree (GBDT) [5]	94.59
Recurrent neural networks (RNN) [13]	86.00
Support vector machine (SVM) [14]	85.00
Support vector machine (SVM) [15]	81.00
LSTM autoencoders [17]	94.20
LSTM	93.34
The proposed hybrid CNN-LSTM	97.48

5. CONCLUSION

The use of sensor technology to gather information regarding production equipment's status has allowed data-driven solutions such as DL algorithms to be applied in the field of PdM. In this study, a hybrid CNN-LSTM DL algorithm is presented to develop a PdM model of multiple interdependent components production systems. The proposed combines the robustness of a CNN network and the time series forecasting and the classification of the LSTM. To evaluate CNN-LSTM DL model, a published real-world industrial machine data is used. The results show that the proposed model achieves an average of 97.6% Precision, 97.4% Recall, and 97.48% F-Score on the testing dataset. The hybrid CNN-LSTM model shows improvement in prediction accuracy of 4.44% compared to the regular LSTM model. The hybrid model accuracy results also outperformed the result of the realed PdM works.

REFERENCES
[1]   C. Chen, Y. Liu, X. Sun, C. D. Cairano-Gilfedder, and S. Titmus, "Automobile maintenance prediction using deep learning with GIS data," *Procedia CIRP*, vol. 81, pp. 447-452, 2019, doi: 10.1016/j.procir.2019.03.077.

[2]   A. Van Horenbeek, and L. Pintelon, "A dynamic predictive maintenance policy for complex multi-component systems," *Reliability Engineering & System Safety,* vol. 120, pp. 39-50, 2013, doi: 10.1016/j.ress.2013.02.029.

[3]   G. A. Susto, A. Schirru, S. Pampuri, S. McLoone, and A. Beghi, "Machine learning for predictive maintenance: A multiple classifier approach," *IEEE Transactions on Industrial Informatics,* vol. 11, no. 3, pp. 812-820, 2014, doi: 10.1109/TII.2014.2349359.

[4]   T. Rieger, S. Regier, I. Stengel, and N. L. Clarke, "Fast Predictive Maintenance in Industrial Internet of Things (IIoT) with Deep Learning (DL): A Review," *CERC,* 2019, pp. 69-80.

[5]   B. Hrnjica, and S. Softic, "Explainable AI in Manufacturing: A Predictive Maintenance Case Study," *IFIP International Conference on Advances in Production Management System*s, 2020, pp. 66-73, doi: 10.1007/978-3-030-57997-5_8.

[6]   J. Hu, "Applying wireless sensor networks in an online monitoring and energy management system for industrial motors," *International Journal of Intelligent Engineering and Systems*, vol. 2, no. 3, pp. 27-35, 2009.

[7]   K. Wang, and Y. Wang, "How AI affects the future predictive maintenance: a primer of deep learning," *International Workshop of Advanced Manufacturing and* Automation, vol. 451, 2017, pp. 1-9, doi: 10.1007/978-981-10-5768-7_1.

[8]   S. Sarkar, K. K. Reddy, M. Giering, M. R. Gurvich, "Deep learning for structural health monitoring: A damage characterization application," *Annual conference of the prognostics and health management society*, 2016, pp. 176-182.

[9]   Z. Chen, Y. Liu, and S. Liu, "Mechanical state prediction based on LSTM neural netwok," *2017 36th Chinese Control Conference (CCC),* 2017, pp. 3876-3881, doi: 10.23919/ChiCC.2017.8027963.

[10]   Z. Tian, and M. J. Zuo, "Health condition prediction of gears using a recurrent neural network approach," *IEEE transactions on reliability*, vol. 59, no. 4, pp. 700-705, 2010, doi: 10.1109/TR.2010.2083231.

[11]   T. Huuhtanen, and A. Jung, "Predictive maintenance of photovoltaic panels via deep learning," *2018 IEEE Data Science Workshop (DSW),* 2018, pp. 66-70, doi: 10.1109/DSW.2018.8439898.

[12]   T. Abbasi, K. H. Lim, and K. S. Yam,"Predictive maintenance of oil and gas equipment using recurrent neural network," *Iop conference series: Materials science and engineering,* vol. 495, no. 1, 2019, Art. no. 012067.

[13]   A. Rivas, J. M. Fraile, P. Chamoso, A. González-Briones, I. Sittón, and J. M. Corchado, "A predictive maintenance model using recurrent neural networks," *International Workshop on Soft Computing Models in Industrial and Environmental Applications*, 2019, pp. 261-270, doi: 10.1007/978-3-030-20055-8_25.

[14]   R. R.-Gonzalez, J. G.-Gil, F. J. Gomez-Gil, and V. Martínez-Martínez, "An SVM-based classifier for estimating the state of various rotating components in agro-industrial machinery with a vibration signal acquired from a single point on the machine chassis," *Sensors,* vol. 14, no. 11, pp. 20713-20735, 2014, doi: 10.3390/s141120713.

[15]   S. Hwang, J. Jeong, and Y. Kang, "SVM-RBM based Predictive Maintenance Scheme for IoT-enabled Smart Factory," *2018 Thirteenth International Conference on Digital Information Management (ICDIM),* 2018, pp. 162-167, doi: 10.1109/ICDIM.2018.8847132.

[16]   J. S. Rahhal, and D. Abualnadi, "IOT Based Predictive Maintenance Using LSTM RNN Estimator," *2020 International Conference on Electrical, Communication, and Computer Engineering (ICECCE),* 2020, pp. 1-5, doi: 10.1109/ICECCE49384.2020.9179459.

[17]   X. Bampoula, G. Siaterlis, N. Nikolakis, K. Alexopoulos, "A Deep Learning Model for Predictive Maintenance in Cyber-Physical Production Systems Using LSTM Autoencoders," *Sensors,* vol. 21, no. 3, 2021, Art. 972, doi: 10.3390/s21030972.

[18]   T. N. Sainath, O. Vinyals, A. Senior, H. Sak, "Convolutional, long short-term memory, fully connected deep neural networks," *2015 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, 2015, pp. 4580-4584, doi: 10.1109/ICASSP.2015.7178838.

[19]   J. Yang, M. N. Nguyen, P. P. San, X. Li, S. Krishnaswamy, "Deep convolutional neural networks on multichannel time series for human activity recognition," *Ijcai*, vol. 15, pp. 3995-4001, 2015.

[20]   N. Xue, I. Tiguero, G. P. Figureueredo, D. Landa-Silva, "Evolving deep CNN-LSTMs for inventory time series prediction," *2019 IEEE Congress on Evolutionary Computation (CEC)*, 2019, pp. 1517-1524, doi: 10.1109/CEC.2019.8789957.

[21]   H. M. Hasemian, "State-of-the-art predictive maintenance techniques," *IEEE Transactions on Instrumentation and measurement*, vol. 60, no. 1, pp. 226-236, doi: 10.1109/TIM.2009.2036347.

[22]   A. Shrestha, A. Mahmood, "Review of deep learning algorithms and architectures," *IEEE Access*, vol. 7, pp. 53040-53065, 2019, doi: 10.1109/ACCESS.2019.2912200.

[23]   X. K. Dang, X.K., Truong, H.N., Nguyen, V.C. and Pham, T.D.A., "Applying convolutional neural networks for limited-memory application," *TELKOMNIKA Telecommunication, Computing, Electronics and Control*, vol. 19, no. 1, pp. 244-251, doi: 10.12928/telkomnika.v19i1.16232.

[24]   Y. Ran, X. Zhou, P. Lin, Y. Wen, R. Deng, "A survey of predictive maintenance: Systems, purposes and approaches," 2019, *arXiv preprint arXiv:1912.07383*.

[25]   K. Raed and G. Methaq, "A hybrid of CNN and LSTM methods for securing web application against cross-site scripting attack," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS),* vol. 21, pp. 1022-1029, 2021, doi: 10.11591/ijeecs.v21.i2.pp1022-1029.

[26]   M. M. B. Ismail, "Insult detection using a partitional CNN-LSTM model," *Computer Science and Information Technologies,* vol. 1, no. 2, pp. 84-92, 2020, doi: 10.11591/csit.v1i2.p84-92.

[27]   S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," *2017 International Conference on Engineering and Technology (ICET)*, 2017, pp. 1-6, doi: 10.1109/ICEngTechnol.2017.8308186.

[28]   J. S. Ashwin, and N. Manoharan, "Convolutional neural network based target recognition for marine search," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 8, no. 2, pp. 561-563, 2017, doi: 10.11591/ijeecs.v8.i2.pp561-563.

[29]   Y. L. Cun *et al.,* "Handwritten digit recognition with a back-propagation network," *Proceedings of the 2nd International Conference on Neural Information Processing Systems*, 1989, pp. 396-404.

[30]   A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems,* vol. 25, pp. 1097-1105, 2012.

[31]   S. Zheng, K. Ristovski, A. Farahat, and C. Gupta, "Long short-term memory network for remaining useful life estimation," *2017 IEEE international conference on prognostics and health management (ICPHM),* 2017, pp. 88-95, doi: 10.1109/ICPHM.2017.7998311.

[32]   S. Hochreiter, and J. Schmidhuber, "Long short-term memory," *Neural computation,* vol. 9, no. 8, pp. 1735-1780, 1997, doi: 10.1162/neco.1997.9.8.1735.

[33]   C. Hwang, K. Lee, and H. Jung, "Improving data quality using a deep learning network," *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*, vol. 20, no. 1, pp. 306-312, doi: 10.11591/ijeecs.v20.i1.pp306-312.

[34] M. Abadi *et al.,* "Tensorflow: A system for large-scale machine learning," *12th symposium on operating systems design and implementation*, 2016, pp. 265-283.

BIOGRAPHIES OF AUTHORS

Ahmed Nasser 🆔 🔳 SC Ⓟ was born in Bagdad, Iraq in 1984. He got his BSc degree from University of Technology Control and Systems Eng. Faculty, Baghdad Iraq in 2006, MSc degree in Computer Eng. from Istanbul University, Istanbul Turkey in 2012, PhD degree from Hacettepe University, Ankara Turkey in Computer Engineering in 2018. Currently he is Lecturer at Control and Systems Engineering Department in the University of Technology. His current interest on "Data Mining", "Natural Language Processing" and "Machine and Deep Learning". He can be contacted at email: ahmed.r.nasser@uotechnology.edu.iq.

Huthaifa Al-Khazraji 🆔 🔳 SC Ⓟ was born in Baghdad, Iraq in 1984. He recive his B.S degree in Control and Systems Engineering Department at the University of Technology, Iraq in 2006. And His M.S Degee in Industrial Engineering and Management at Politecnico di Torino University, Italy in 2010. From 2010 to 2014 he was Assistant Lecturere at Control and Systems Engineering Department at the University of Technology, Iraq. He was accepted in Central Queensland University in Australia as a PhD student in March 2015. He received his PhD degree in 2019. He received his scientific promotion to be a University Lecturer at Control and Systems Engineering Department in the University of Technology, Iraq in 2019. He writes severals papers in Modeling, Simulation, Control System, Inventory Control, Process Capability Index, Overall Equipment Effectiveness and Assemble Line Balancing Problem. His area of interest: Control Engineering, Industrial Engineering and Optimization. He can be contacted at email: 60141@uotechnology.edu.iq.