Received January 28, 2018, accepted February 22, 2018, date of publication February 27, 2018, date of current version April 18, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2809681

Industrial Big Data Analytics for Prediction of Remaining Useful Life Based on Deep Learning

HEHUA YAN[1], JIAFU WAN[2], CHUNHUA ZHANG[2], SHENGLONG TANG[2], QINGSONG HUA[3], AND ZHONGREN WANG[4], (Member, IEEE)
[1]School of Electrical Engineering, Guangdong Mechanical and Electrical College, Guangzhou 510515, China
[2]School of Mechanical and Automotive Engineering, South China University of Technology, Guangzhou 510640, China
[3]School of Mechanical and Electrical Engineering, Qingdao University, Qingdao 266000, China
[4]School of Mechanical Engineering, Hubei University of Arts and Science, Xiangyang 441000, China

Corresponding author: Chunhua Zhang (chhzhang@scut.edu.cn)

This work was supported in part by the Natural Science Foundation of Guangdong Province, China, under Grant 2015A030313746 and Grant 2017B030311008, in part by the National Key Research and Development Project under Grant 2017YFE0101000, in part by the Major Projects for Numerical Control Machine under Grant 2015ZX04005001, in part by the Natural Science Foundation of Hubei Province, China, under Grant 2014CFB637, and in part by the Research Fund Program of Guangdong Provincial Key Laboratory of Computer Integrated Manufacturing under Grant CIMSOF2016004.

ABSTRACT Due to the recent development of cyber-physical systems, big data, cloud computing, and industrial wireless networks, a new era of industrial big data is introduced. Deep learning, which brought a revolutionary change in computer vision, natural language processing, and a variety of other applications, has significant potential for solutions providing in sophisticated industrial applications. In this paper, a concept of device electrocardiogram (DECG) is presented, and an algorithm based on deep denoising autoencoder (DDA) and regression operation is proposed for the prediction of the remaining useful life of industrial equipment. First, the concept of electrocardiogram is explained. Then, a problem statement based on manufacturing scenario is presented. Subsequently, the architecture of the proposed algorithm called integrated DDA and the algorithm workflow are provided. Moreover, DECG is compared with traditional factory information system, and the feasibility and effectiveness of the proposed algorithm are validated experimentally. The proposed concept and algorithm combine typical industrial scenario and advance artificial intelligence, which has great potential to accelerate the implementation of industry 4.0.

INDEX TERMS Cyber-physical systems, deep learning, device electrocardiogram, industrial big data, industry 4.0.

I. INTRODUCTION

In recent years, the development of Internet of Things (IoT) [1], [2] and cloud technologies [3], [4] caused deployment of sensors and line layout in manufacturing environment, which increased the cost of manufacturing maintenance. The predictive maintenance, as an important part in manufacturing maintenance, plays a vital role in scheduling, maintenance management, and quality improvement [5]. In general, the predictive maintenance can be categorized into experience-based models, physics-based models, and data-driven models [6].

The experience-based models, which represent the traditional methods for predictive maintenance, are widely used. Moreover, the experience-based models can be well explained and easy debugged using IF-THEN rules and fuzzy logic. For instance, Jindal and Aggarwal [7] proposed a user-friendly expert system that is capable to assist the drivers to cope with their car problems by providing a logical solution. However, the experience-based models cannot deal with a large number of queries in expert systems, and they mostly rely on expert knowledge and engineering experience. The similar deficiency exits in physics-based models. The physics-based models require insight in system failure mechanisms, which are supposed to be converted into mathematical expressions. Zhao *et al.* [8] provided a method for prediction of remaining useful life (RUL) for gears based on Bayesian framework, which updates the parameters. Daigle and Goebel [9] used the particle filters to predict RUL of solenoid valves. However, in certain cases, the physics-based models are not appropriate for complicated systems because the humans cannot understand all failure modes and behaviors.

Due to increasing data availability, the data-driven models provide a new approach in predictive maintenance.

The historical data of equipment, statistical models, reliable functions, and artificial intelligence methods, especially machine learning, are widely used for both RUL estimation and other applications. Namely, the hidden Markov models (HMMs) [10], grey models [11] and other methods were successfully applied to predictive maintenance. In 2004, Gebraeel *et al.* [12] proposed the use of neural networks in machine learning to predict the bearing RUL. He *et al.* [13] provided a method for prediction of axial piston pump based on the support vector machines (SVMs) [14], empirical mode decomposition (EMD) of time series [15], and particle swarm optimization (PSO) [16]. With the advent of data explosion era, the deep learning based on layers of neurons [17], which is evolved from the artificial neural networks, have a strong learning ability. Namely, the deep learning has improved the state-of-the-art in many aspects dramatically, especially in images processing [18] and natural language processing [19]. Due to recent improvements in cyber-physical systems (CPS) [20], big data [21]–[24], cloud-assisted equipment [25], software-defined networks [26]–[28], artificial intelligence [29]–[31], and industrial wireless networks [32], [33], there is a variety of applications wherein a deep learning is employed for environment manufacturing.

In this paper, the device electrocardiogram (DECG) principle is introduced, and a new methodology based on deep learning and DECG is proposed for prediction of RUL of equipment as well as production line. DECG, which is similar to monitor the health of the human body, records devices' cycle time with all its sub-processes. Due to much more data collected from DECG, it's possible to introduce deep learning and fully enhance the performance of deep learning for specific application. Based on deep learning and a large number of run-to-failure samples, the proposed algorithm can provide an accurate prediction of device RUL.

In summary, there are three main contributions of this paper:

- A concept of DECG in manufacturing environment, which provides a fine-grained status observation and reduces dependency on experts' knowledge greatly, is proposed.
- A RUL predicting methodology based on regression and deep denoising auto-encoders (DDA) is proposed to achieve an automatic feature engineering and a high-level features extraction.
- The proposed algorithm and traditional factory information system are compared, and the experiment is performed in order to validate the feasibility and effectiveness of proposed algorithm.

The paper is organized as follows. In Section II, the main idea and details of DECG are presented. In Section III, the problems are defined and described in detail. The algorithm based on deep learning and industrial big data analysis is proposed in Section IV. In Section V, the experimental results are provided. Lastly, a brief conclusion is given in Section VI.

II. DEVICE ELECTROCARDIOGRAM

The existing technologies for industrial maintenance, such as factory information system (FIS) [34] and prognostic health management (PHM) [35], are widely used in manufacturing. However, FIS can only detect whether an over cycle appears, while the specific motion that causes delay cannot be detected. Additionally, this coarse-grained monitoring method is able only to recognize the critical issues after they occurred, which means that no sign is detected before the issue happens. In contrary, PHM can detect early signs of potential failure of machines, and currently it is one of the most popular methods. However, many PHM requires additional sensors, which affects the detection accuracy because of installation, calibration, and environmental issue. On the other hand, PHM cannot detect the cycle end. Namely, in production scenarios, especially in vehicle production, the production cycle which is consisted of thousands of motions is often accomplished within one minute. Therefore, a lack of detection of cycle end might cause huge losses.

The DECG for industrial maintenance is similar to physical examination of human health condition. With fine-grained monitoring of sub-processes within a production cycle, DECG can visually represent current or historical processing time of the sub-processes for further scheduling and maintenance.

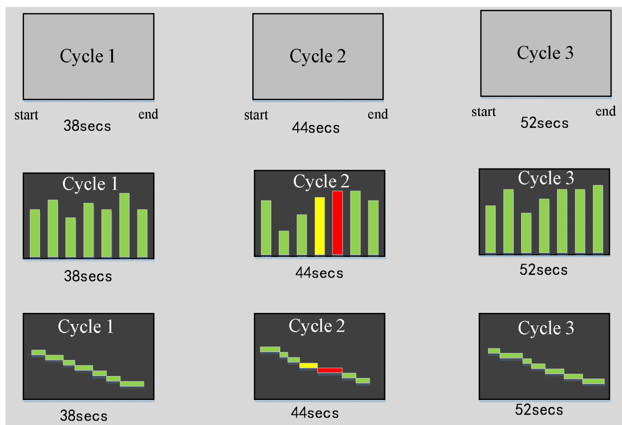


FIGURE 1. **DECG principle.**

The DECG principle is presented in Fig. 1. In many manufacturing scenarios, a final product is made by a variety of processing steps or production cycles. For instance, in vehicle production, there are often four typical production cycles, namely stamping, coating, welding, and general assembly. Each of the listed production cycles is made up with multiple operations. For each specific operation, a baseline and appropriate tolerances of working time are set. In DECG, the working time of each operation in the cycle is recorded in form of bar chart and colored in green, yellow, orange or red that corresponds to good, watch, warning or fault respectively. These bar charts are presented in a sequence view and displayed as a set of cycle operations.

Unlike many PHM, DECG can avoid sensors installation and retrieve time information for each operation to the device or operator level from the programmable logic controller (PLC). Another important DECG advantage is the great reduction of dependency on experienced experts in complicated production issues, which significantly reduces maintenance cost for maintenance stuffs. In addition, unlike FIS, DECG monitors all operations within a production cycle and provides predictive insights in production lines, which improves operator efficiency and unplanned downtime.

For the aspect of implementation, considering the PLCs' important role in manufacturing control, additional devices are required to extract the information (working time of process and its sub-operations) from PLC. With the OPC UA Server deployed and directly connected to PLC, data changes can be attained and the production logs can be generated. Therefore, OPC UA Sever can be treated as a data collector for specific application. In summary, the spatially distributed devices and corresponding OPC UA Servers are linked together to collect the records and take it as input for DECG.

III. PROBLEM STATEMENT

In this section, a typical industrial scenario and problem statement are introduced. In manufacturing process, a very common scene is that several identical devices do the same job in parallel. This job usually contains more than one subprocess. Since multiple similar devices and processes provide sufficient information support, it is assumed that if the state data from adequate run-to-failure devices are recorded, the unplanned downtime of identical or similar devices can be predicted before failure happened.

Thus, in this study, the problem statement is defined as follows. There are N identical run-to-failure devices, $D = \{d_1, d_2, \ldots, d_N\}$, performing a specific assignment that contains M sub-processes, $P = \{p_1, p_2, \ldots, p_M\}$. In production cycle, the working time of each sub-process is recorded and these records tend to change gradually according to certain rule. Whenever identical devices or devices with similar processes require the RUL prediction, the corresponding working time within the production cycle is compared with the recorded working time. Since the recorded working time is collected from run-to-failure devices, when numerical value and numerical distribution characteristics of device are similar to recorded ones, the device RUL can be predicted based on matched records. Hence, our goal is to predict the unplanned downtime or RUL of working devices using the run-to-failure devices.

IV. INDUSTRIAL BIG DATA ANALYTICS

In this section, the algorithm for industrial big data analytics, data acquisition and preprocessing, is introduced. As it was mentioned in previous section, in this study, plenty of similar devices or a group of devices generates a large amount of data. Data acquisition and preprocessing should be accomplished before data analysis. In order to achieve higher prediction accuracy and the best use of cleaned data, the ability of deep learning to attain insight and knowledge from big data is used. The algorithm details are presented in the following.

A. DATA ACQUISITION AND PREPROCESSING

The data analytics proposed in this paper is based on a two-phase prediction of unplanned downtime, which consists of training phase and predicting phase.

As mentioned in the preceding section, it is supposed that there are N run-to-failure identical devices $D = \{d_1, d_2, \ldots, d_N\})$ performing a specific assignment of production cycle, which contains M sub-processes $P = \{p_1, p_2, \ldots, p_M\}$. The training of deep neural network is performed with state records of these N run-to-failure devices. Namely, for each device, the state data at K time moments before failure $T = \{t_1, t_2, \ldots, t_K\}$ are recorded. Here, we use R_i, where $i = 1, 2, \ldots, N$, in $S = \{R_1, R_2, \ldots, R_N\}$ to denote the i^{th} device records set for K time moments. More specifically, R_i is a set $\{R_{i1}, R_{i2}, \ldots, R_{iK}\}$, which contains K records for K time moments. Furthermore, since each process (record set) consists of M sub processes, the j^{th} $(j = 1, 2, \ldots, K)$ record set of i^{th} $(i = 1, 2, \ldots, N)$ device R_{ij} will embrace M sub-records for these sub-processes $R_{ij} = \{r_1^{ij}, r_2^{ij}, \ldots, r_M^{ij}\}$. When the information is collected, data cleaning such as feature discretization, missing data filling are performed for further operation.

Deep learning is performed with mentioned records, which are used as input and output data. In prediction phase, if one or more similar devices are in a sharp wear stage, the state data of these devices are treated as input of deep neural network, while network output is the predicted RUL.

B. RUL PREDICTION BASED ON DEEP LEARNING

In recent years, the data-driven models have a very important role in active maintenance and RUL prediction. However, one of the most significant model tasks is feature engineering, which involves several professional data operations, such as dimensionality reduction, and relies greatly on specific production scenario. In manufacturing environment, different types of processing units usually have their own working mechanisms and maintenance logics. Therefore, it might be impossible to achieve a large number of corresponding advanced data-driven models for maintenance, which is unnecessary and causes a resource wasting.

In this paper, we propose deep learning for prediction of the remaining useful life. One of the most obvious advantages of deep learning is its ability to extract the features automatically such as convolutional neural network (CNN) and recurrent neural network (RNN) [36]. Additionally, as the layer gets deeper, the number of features usually reduces and obtained features become more abstract. In our application scenario, the historical data of each process working time are used for training. Based on the production background interpreted before, and in order to achieve the best feature extraction using deep learning, the integrated deep denoising auto-encoder (IDDA) is introduced into manufacturing environment.

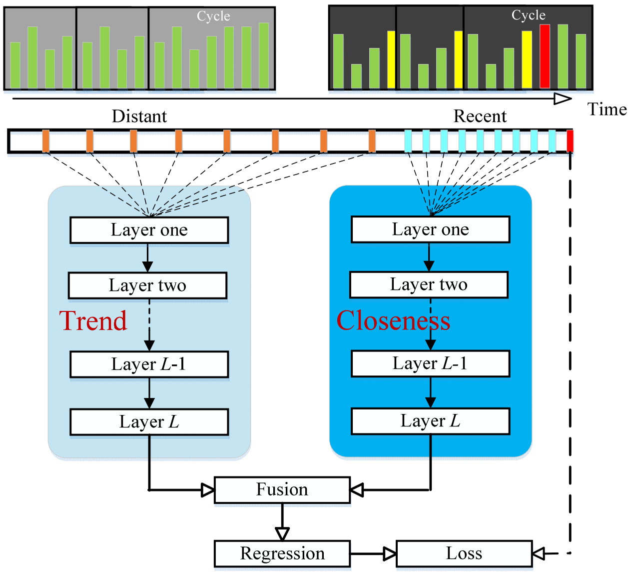


FIGURE 2. IDDA architecture.

The proposed IDDA architecture shown in Fig. 2 consists of two DDA and a linear regression analysis. As mentioned before, our goal is to predict the equipment RUL based on current state, which is in our case, the working time of equipment processes. Considering that collected data consist of time series, in order to achieve a better prediction, at each time moment during the training phase the record is split into distant records and recent records, and then, they are used as inputs of two different DDA. The distant records denote the records that are far away from current time moment, while the recent records denote records that are close to current time moment. From our perspective, an accurate prediction requires reasonable fusion of damage tendency and current states. Thus, the distant records are used to simulate the damage trend, while the recent records are used to simulate the smoothing process of recent change. With the aim to avoid the overfitting, a famous trick called dropout is applied to data flow in two deep models. Afterwards, two outputs are fused and eventual linear regression is performed to transform the discrete records to equipment RUL. The details of proposed algorithm and workflow (see Algorithm 1), including both training phase and predicting phase, are presented in the following.

When the historical data from K time moments are gathered and cleaned, the first step is to split the data into distant records and recent records. In this paper, records collected at K time moments do not have a uniform distribution all the time. As shown in line 2 of proposed algorithm, time moments from $NO.1$ to $NO.S$ have uniform distribution, which provides a large scale for damage tendency simulation. On the other hand, the rest of time moments, which form the time interval from $NO.S$ to $NO.N$, shrink to embody the recent characteristic fully. Additionally, since the records have the same data form, DDAs have the same structure. In the training phase, after data fusion and regression, the difference of prediction and recorded values is defined as Euclidean loss. Moreover, if Adam stochastic optimization is adopted, faster iterators can be achieved compared with traditional SGD or batch processing.

Algorithm 1 Algorithm of Integrated Deep Denoising Auto-Encoder

Training Input: $D = \{(R_i, Y_i)\}_{i=1}^{N}$ where R_i and Y_i denote the i^{th} equipment's records and the remaining useful life, respectively
Prediction input: distant records $I_d^p = \{P_i\}_{i=1}^{S}$ and recent records $I_r^p = \{P_i\}_{i=S}^{K}$

1. //training phase
2. Split D into $I_d \leftarrow \{R_{ij}\}_{i=1}^{N}{}_{j=1}^{S}$, $I_r \leftarrow \{R_{ij}\}_{i=1}^{N}{}_{j=S}^{K}$ and $Y \leftarrow \{Y_i\}_{i=1}^{N}$
3. **for** $i \leftarrow 1$ to N //for each equipment
4. $I_d^i \leftarrow \{R_{ij}\}_{j=1}^{S}$
5. $I_r^i \leftarrow \{R_{ij}\}_{j=S}^{K}$
6. *Output1* $\leftarrow$ **DeepDenoisingAuto-encoder1**(input $\leftarrow I_d^i$)
7. *Output2* $\leftarrow$ **DeepDenoisingAuto-encoder2**(input $\leftarrow I_r^i$)
8. //integrate the outputs of two deep neural networks into one output
9. *Output* $\leftarrow$ **Fusion**(*Output1*, *Output2*)
10. //perform regression to transform the discrete output into prediction
11. $Y_i^{\Delta} \leftarrow$ **Regression**(*Output*)
12. // calculate Euclidean loss between predictions and records
13. $Loss \leftarrow \frac{1}{2N}\sum_{i=1}^{N} ||Y_i - Y_i^{\Delta}||_2^2$
14. //use stochastic optimization to update the parameters of the IDDA
15. // α is the learning rate
16. **StochasticOptimization**(*Loss*, α)
17. **End for**
18. // prediction phase
19. *Poutput1* $\leftarrow$ **DeepDenoisingAuto-encoder1** (input $\leftarrow I_d^p$)
20. *Poutput2* $\leftarrow$ **DeepDenoisingAuto-encoder2** (input $\leftarrow I_r^p$)
21. *Poutput* $\leftarrow$ **Fusion**(*POutput1*,*POutput2*)
22. *Prediction* $\leftarrow$ **Regression**(*POutput*)

V. EXPERIMENTS

A. EXPERIMENTS DESCRIPTION

In this section, the proposed algorithm is validated and compared with FIS. Specifically, these methods were implemented in CNC machining center, which has been working for a long time and the status information was recorded until the machining center was exhausted. The machining center was used to remove redundant material from the mechanical connecting rod and it had 32 sub-operations, such as axis rotating, cutting, tool changing, etc. In order to guarantee the acquisition and precision of working time of operations, the OPC UA server was launched over the embedded device and it was connected directly to PLC of machining center. Using the Server/Client mode of OPC UA, we managed to record the working time of all sub-operations within a production cycle. Since the production cycle time was around 52 seconds, it was possible to obtain a large amount of status information, thus the generalization ability of complex deep model was greatly improved.

As it was previously presented, the proposed algorithm embraces IDDA and numeric regression. Namely, the algorithm first had two DDA with the same structure, which were supposed to receive the input of distant and recent records respectively, and to capsule the fused output as an input for the next regression operation. Then, a regression was added to predict the RUL. One of the most significant feature of DDA is that at the training part, there is only one weight matrix to be trained between two layers ordered by data flow. Once the farmer weight matrix trained, the next training begins. Additionally, after training, DDA can represent the features with less dimension but with little information loss. That's because the input and output of weight matrix are artificially the same, which lead DDA to make the feature another type of representation (usually the dimension of the output is less than the input) but with little information loss. In our experiment, to make the feature transformation smoothly, we decided to have three weight matrixes where the first two of them are trained as DDA unit does and a fully connected layer. Details of DDA are shown in Table 1.

TABLE 1. Parameters of integrated deep denoising auto-encoders.

Layer	Number of Neurons	Connection Type
Input Layer	32	Drop out
First Hidden Layer	24	Drop out
Second Hidden Layer	18	Drop out
Third Hidden Layer	10	Fully Connected
Output Layer	4	Null

As it can be seen in Table 1, DDAs have three hidden layers, which are used to extract high-level features. Although huge amount of data was obtained, in order to avoid the overfitting, the drop out operation was implemented, which represents a popular trick in deep learning that makes some neurons artificially dead when data flows forward. Additionally, since we trained the model in Caffe [37], Xavier initialization [38] was applied to adjust the weights automatically based on numbers of input layer and output layer. The training of DDAs was performed with *Softmax* activation function and Gaussian noise was added to improve the model robustness. At the last hidden layer, instead of drop out and *Softmax*, a fully connection was employed and Sigmoid activation function was used. In the training, after the linear regression, Euclidean loss layer was introduced to build-in module in Caffe to calculate the loss. Moreover, in order to make a faster convergence, Adam Optimization was implemented.

The number of neurons in input layer was chosen according to number of sub-processes in machining center. Since there were 30 sub-processes, the input layer consisted of 32 neurons, whose inputs were the working times. In neural network output layer was a set of four high-level features, which were used as an input for the following regression operation. With the aim to simulate the difference between distant records and recent records, the records at every 15 records and 2 records were used as distant records and recent records, respectively. Considering the experiment cost, two similar CNC machining centers were used as a provider of training samples and it was assumed that if the production cycle exceeds the expectation (52 seconds in this case) by 30%, the machining center is treated as failed.

B. RESULTS AND ANALYSIS

One of FIS characteristic is that it is able to recognize critical issues after they occurred, but it cannot determine which part of production cycle caused the delay. Compared with FIS, DECG provides a fine-grained methodology without any additional sensors. Moreover, the fusion of DDA and proposed algorithm has less dependence on artificial decisions in RUL prediction. The snapshot of FIS and DECG taken during the experiment is presented in Fig. 3.

TABLE 2. Experiment results for FIS and DECG from the first warning to the failure.

	Time Interval (min)	Number of Warnings
FIS	62	52
DECG	210	Continuous warning since the first warning was sent

As it is shown in Fig. 3, FIS reflects the coarse-grained state of specific process, which is composed of time stamp, working time and evaluation of past process. Namely, FIS is able to estimate weather the total working time is normal, but it cannot determine the specific abnormal sub. In contrary, DECG (lower subs in Fig. 3) is capable to record a fine-grained information, which discovers the abnormal sub operation and sends the warning signal. This significant difference between FIS and DECG has a great impact on RUL prediction. The time intervals of FIS and DECG from the first warning to the failure obtained in experiment and the response speed in equipment abnormal performance are presented in Table 2. As it is shown in Table 2, the main difference between FIS and DECG is related to number of warnings before the failure. In the experiment, DECG recognized the abnormal operation in one-axis motion 210mins before failure, while

Runtime Record

Machine: NO.2
From: 05/26/2016 11:00:00 am
To: 05/26/2016 11:15:32 am

05/26/2016 11:15:32 am	0:00:52	M			Normal
05/26/2016 11:14:39 am	0:00:53	M			Normal
05/26/2016 11:13:47 am	0:00:52	M			Normal
05/26/2016 11:12:52 am	0:00:55	M			Normal
05/26/2016 11:11:55 am	0:00:57	M			Attention
05/26/2016 11:10:55 am	0:01:00	M			Attention
05/26/2016 11:09:01 am	0:00:54	M			Normal
05/26/2016 11:08:09 am	0:00:52	M			Normal
05/26/2016 11:07:17 am	0:00:52	M			Normal
05/26/2016 11:06:24 am	0:00:53	M			Normal
05/26/2016 11:05:33 am	0:00:51	M			Normal
05/26/2016 11:04:35 am	0:00:58	M			Attention
05/26/2016 11:03:40 am	0:00:55	M			Normal
05/26/2016 11:02:47 am	0:00:53	M			Normal
05/26/2016 11:02:55 am	0:00:52	M			Normal
05/26/2016 11:01:55 am	0:01:00	M			Attention
05/26/2016 11:01:01 am	0:00:54	M			Normal
05/26/2016 11:00:52 am	0:01:01	M			Attention
05/26/2016 11:00:00 am	0:00:52	M			Normal

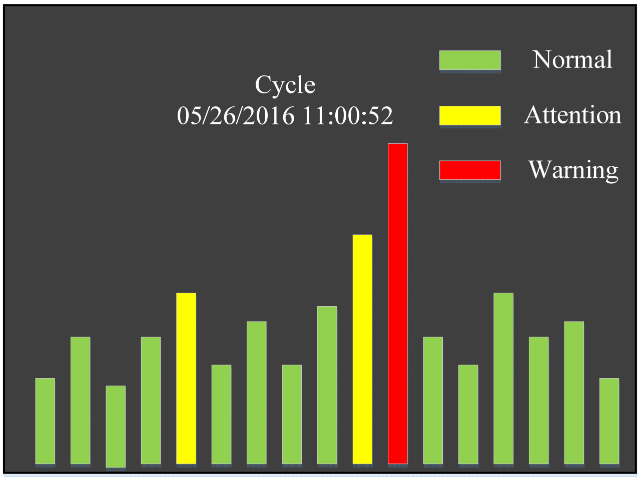


FIGURE 3. The snapshot of FIS and DECG.

FIS detected the delay just 52 minutes before the failure. Consequently, FIS had 52 warning times in total, including a few unexplained warnings around the 62th minute and frequent warnings that occurred 48 minutes before the failure. In contrary, DECG was warning continuously since the first warning was sent, which greatly improved the credibility of failure detection.

The additional experiment was performed in order to validate the prediction accuracy. In order to present the prediction precision properly, the predictions were recorded at every 5% of RUL exceeding, and obtained results are shown in Fig. 4. Usually, the estimated RUL is shorter than true RUL. Therefore in Fig. 4, the X axis represents the exceeding percentage of cumulative working time against the estimated RUL.

Similar to previous experiment, this experiment was conducted on two similar CNC machining centers. Regardless the short production cycle, a large amount of data was recorded, but the statistical data from BEET Company [39] were also considered in order to achieve better validation of proposed algorithm and to avoid abnormal samples from limited experimental object. The statistical data were focused on similar application of a large truck engine. Based on records in experimental data and statistical data from BEET Company, the actual RUL, which is labeled as a true RUL in this paper, of experiment objects had better credibility.

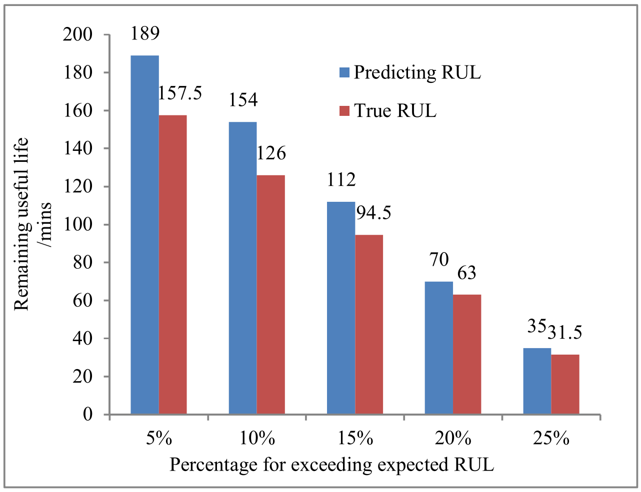


FIGURE 4. Comparison of RUL predicted by proposed model and true RUL.

The experimental results are presented in Fig. 4, the error between predicted and true value is about 20%. In addition, the contrast demonstrated in Fig. 3 reflects a high accuracy when DECG is used to refine production cycle and to predict RUL. From our perspective, DECG both interprets the trend of operation working time and provides much more fine-grained training samples, which reduces deep learning overfitting and improves prediction accuracy significantly. Additionally, IDDA reduces the dependency on expert experience and human decision, and the integration with deep learning provides an automatic feature engineering, which fully embodies the concept of smart manufacturing.

VI. CONCLUSIONS

In this paper, a new algorithm for RUL prediction based on DECG and deep learning is presented. Firstly, the concept of DECG was introduced. Then, the problem statement in manufacturing environment was explained. In addition, in order to reduce the impact of experts' experience and human decision on prediction, a deep learning methodology, which embraces IDDA and regression operation, was used. The proposed algorithm was verified by experiments, wherein DECG was compared with FIS. The experimental result have proven DECG superiority over FIS in terms of response and reliability. Furthermore, the prediction accuracy of IDDA was validated by comparison with true RUL. The obtained results have shown a high effectiveness of proposed algorithm. Nevertheless, the comparison results have indicated superiority of proposed algorithm and its feasibility to accelerate the implementation of Industry 4.0.

REFERENCES

[1] J. Wan, S. Tang, Q. Hua, D. Li, C. Liu, and J. Lloret, "Context-aware cloud robotics for material handling in cognitive industrial Internet of Things," *IEEE Internet Things J.*, to be published, doi: 10.1109/JIOT.2017.2728722.

[2] E. Ahmed *et al.*, "The role of big data analytics in Internet of Things," *Comput. Netw.*, vol. 129, pp. 459–471, Dec. 2017.

[3] J. Wan, D. Zhang, Y. Sun, K. Lin, C. Zou, and H. Cai, "VCMIA: A novel architecture for integrating vehicular cyber-physical systems and mobile cloud computing," *Mobile Netw. Appl.*, vol. 19, no. 2, pp. 153–160, 2014.

[4] E. Ahmed *et al.*, "Bringing computation closer towards user network: Is edge computing the solution?" *IEEE Commun. Mag.*, vol. 55, no. 11, pp. 138–144, Nov. 2017.

[5] Y. Xu, Y. Sun, J. Wan, X. Liu, and Z. Song, "Industrial big data for fault diagnosis: Taxonomy, review, and applications," *IEEE Access*, to be published, doi: 10.1109/ACCESS.2017.2731945.

[6] L. Linxia and F. Köttig, "Review of hybrid prognostics approaches for remaining useful life prediction of engineered systems, and an application to battery life prediction," *IEEE Trans. Rel.*, vol. 63, no. 1, pp. 191–207, Mar. 2014.

[7] V. Jindal and J. Aggarwal, "An approach towards designing of car troubleshooting expert system," *Int. J. Comput. Appl.*, vol. 1, no. 23, pp. 107–114, 2010.

[8] F. Zhao, Z. Tian, and Y. Zeng, "Uncertainty quantification in gear remaining useful life prediction through an integrated prognostics method," *IEEE Trans. Rel.*, vol. 62, no. 1, pp. 146–159, Mar. 2013.

[9] M. Daigle and K. Goebel, "Improving computational efficiency of prediction in model-based prognostics using the unscented transform," DTIC, Fort Belvoir, VA, USA, Tech. Rep. ARC-E-DAA-TN2287, 2010.

[10] K. Medjaher, D. A. Tobon-Mejia, and N. Zerhouni, "Remaining useful life estimation of critical components with application to bearings," *IEEE Trans. Rel.*, vol. 61, no. 2, pp. 292–302, Jun. 2012.

[11] Y. Peng and M. Dong, "A hybrid approach of HMM and grey modelfor age-dependent health prediction of engineering assets," *Expert Syst. Appl.*, vol. 38, no. 10, pp. 12946–12953, 2011.

[12] N. Gebraeel, M. Lawley, R. Liu, and V. Parmeshwaran, "Residual life predictions from vibration-based degradation signals: A neural network approach," *IEEE Trans. Ind. Electron.*, vol. 51, no. 3, pp. 694–700, Jun. 2004.

[13] Z. He, S. Wang, K. Wang, and K. Li, "Prognostic analysis based on hybrid prediction method for axial piston pump," in *Proc. 10th IEEE Int. Conf. Ind. Inform. (INDIN)*, Jul. 2012, pp. 688–692.

[14] V. Cherkassky and Y. Ma, "Practical selection of SVM parameters and noise estimation for SVM regression," *Neural Netw.*, vol. 17, no. 1, pp. 113–126, Jan. 2004.

[15] P. Rajesh, M. PandurangeRao, S. Sengottuvel, K. Gireesan, and T. S. Radhakrishnan, "Effective extraction of visual event-related pattern by combining template matching with ensemble empirical mode decomposition," *IEEE Sensors J.*, vol. 17, no. 7, pp. 2146–2153, Jul. 2017.

[16] M. Salucci, L. Poli, N. Anselmi, and A. Massa, "Multifrequency particle swarm optimization for enhanced multiresolution GPR microwave imaging," *IEEE Trans. Geosci. Remote Sens.*, vol. 15, no. 2, pp. 283–289, Feb. 2016.

[17] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.

[18] A. Krizhevsky, I. Sutskever, and G. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.

[19] R. Collobert and J. Weston, "A unified architecture for natural language processing: Deep neural networks with multitask learning," in *Proc. 25th Int. Conf. Mach. Learn.*, 2008, pp. 160–167.

[20] S. Wang, J. Wan, D. Zhang, D. Li, and C. Zhang, "TowTowards smart factory for industry 4.0: A self-organized multi-agent system with big data based feedback and coordination," *Comput. Netw.*, vol. 101, pp. 158–168, Oct. 2016.

[21] W. Yuan, P. Deng, T. Taleb, J. Wan, and C. Bi, "An unlicensed taxi identification model based on big data analysis," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 6, pp. 1703–1713, Jun. 2016.

[22] J. Wan *et al.*, "A manufacturing big data solution for active preventive maintenance," *IEEE Trans. Ind. Informat.*, vol. 13, no. 4, pp. 2039–2047, Aug. 2017.

[23] X. Wang, Y. Zhang, V. Leung, N. Guizani, and T. Jiang, "D2D big data: Content deliveries over wireless device-to-device sharing in realistic large scale mobile networks," *IEEE Wireless Commun.*, vol. 25, no. 1, pp. 1–10, Jan. 2018.

[24] A. Karim *et al.*, "Big data management in participatory sensing: Issues, trends and future directions," *Future Generat. Comput. Syst.*, to be published, doi: 10.1016/j.future.2017.10.007.

[25] J. Wan, S. Tang, H. Yan, D. Li, S. Wang, and A. Vasilakos, "Cloud robotics: Current status and open issues," *IEEE Access*, vol. 4, pp. 2797–2807, Jun. 2016.

[26] J. Liu, J. Wan, B. Zeng, Q. Wang, H. Song, and M. Qiu, "A scalable and quick-response software defined vehicular network assisted by mobile-edge computing," *IEEE Commun. Mag.*, vol. 55, no. 7, pp. 94–100, Jul. 2017.

[27] Y. Zhang, M. Chen, N. Guizani, D. Wu, and V. Leung, "SOVCAN: Safety-oriented vehicular controller area network," *IEEE Commun. Mag.*, vol. 55, no. 8, pp. 94–99, Aug. 2017.

[28] J. Wan, S. Tang, Z. Shu, D. Li, S. Wang, M. Imran, and A. Vasilakos, "Software-defined industrial Internet of Things in the context of industry 4.0," *IEEE Sensors J.*, vol. 16, no. 20, pp. 7373–7380, Oct. 2016.

[29] M. Chen, X. Shi, Y. Zhang, D. Wu, and M. Guizani, "Deep features learning for medical image analysis with convolutional autoencoder neural network," *IEEE Trans. Big Data*, to be published, doi: 10.1109/TBDATA.2017.2717439.

[30] J. Xiong, Q. Zhang, J. Wan, L. Liang, P. Cheng, and Q. Liang, "Data fusion method based on mutual dimensionless," *IEEE/ASME Trans. Mechatronics*, to be published, doi: 10.1109/TMECH.2017.2759791.

[31] J. Wan *et al.*, "Toward dynamic resources management for IoT-based manufacturing," *IEEE Commun. Mag.*, vol. 56, no. 2, pp. 52–59, Feb. 2018.

[32] J. Liu, Q. Wang, J. Wan, J. Xiong, and B. Zeng, "Towards key issues of disaster aid based on wireless body area networks," *KSII Trans. Internet Inf. Syst.*, vol. 7, no. 5, pp. 1014–1035, 2013.

[33] M. Chen, J. Wan, S. Gonzalez, X. Liao, and V. C. M. Leung, "A survey of recent developments in home M2M networks," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 1, pp. 98–114, 1st Quart., 2014.

[34] Y. Hao, A. Shamsuzzoha, and P. Helo, "Strategy for virtual factory information system design," in *Proc. IEEE Int. Conf. Ind. Eng. Manage. (IEEM)*, Dec. 2012, pp. 2023–2027.

[35] T. D. Batzel and D. C. Swanson, "Prognostic health management of aircraft power generators," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 45, no. 2, pp. 473–482, Apr. 2009.

[36] T. Mikolov, M. T. Karafi, L. Burget, J. Cernock, and S. Khudanpur, "Recurrent neural network based language model," *Interspeech*, vol. 2, p. 3, Sep. 2010.

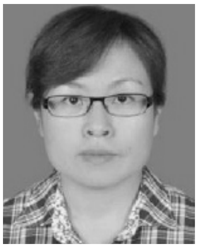
[37] Y. Jia *et al.*, "Caffe: Convolutional architecture for fast feature embedding," in *Proc. 22nd ACM Int. Conf. Multimedia*, 2014, pp. 675–678.

[38] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," *J. Mach. Learn. Res.*, vol. 9, pp. 249–256, May 2010.

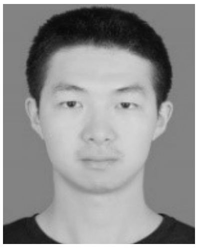
[39] Accessed: Jan. 10, 2016. [Online]. Available: https://docs.beet.com/

HEHUA YAN is currently an Associate Professor with the School of Electrical Engineering, Guangdong Mechanical and Electrical College, China. She directed three research projects, including the Natural Science Foundation of Guangdong Province. She has authored or co-authored 20+ scientific papers. Her research interests include embedded systems, Internet of Things, and cyber-physical systems.

JIAFU WAN is currently a Professor with the School of Mechanical and Automotive Engineering, South China University of Technology, China. He directed 16 research projects, including the National Key Research and Development Project and the National Natural Science Foundation of China. He has authored over 140 scientific papers, including 80+ SCI-indexed papers, 30+ IEEE Transactions/Journal papers, 13 ESI Highly Cited Papers, and four ESI Hot Papers. According to Google Scholar, his published work has been cited more than 4200 times. His SCI other citations (sum of times cited without self-citations) reached 1269 times according to Web of Science Core Collection. His research interests include industry 4.0, industrial wireless networks, cyber physical systems, Internet of Things, clouding computing, embedded systems, and industrial robotics.



CHUNHUA ZHANG is currently a Lecturer with the School of Mechanical and Automotive Engineering, South China University of Technology, China. She directed three research projects, including the Applied Science and Technology Research and Development Program of Guangdong Province. She has authored or co-authored 30+ scientific papers. Her research interests include networked control systems, industrial big data, and cyber-physical systems. She was a recipient of the First Prize for Science and Technology Development of Guangdong Province in 2009.



SHENGLONG TANG received the B.A. degree in mechanical engineering from Yangtze University, China, in 2015. He is currently pursuing the M.S. degree with the School of Mechanical and Automotive Engineering, South China University of Technology, China. His research interests include cloud robotics, cyber-physical systems, and embedded systems.

QINGSONG HUA received the Ph.D. degree in mechanical engineering from the Clausthal University of Technology, Germany, in 2006. He is currently a Professor with the School of Mechanical and Electrical Engineering, Qingdao University, Qingdao, China. His research interests are renewable energy systems, including multienergy system, microgrid, and distributional SOFC power generation.

ZHONGREN WANG (M'10) was born in 1974. He received the B.E. degree from the Beijing University of Chemical Technology, China, in 1997, and the M.E. degree and the Ph.D. degree in mechanical engineering from the South China University of Technology, Guangzhou, in 2003 and 2009, respectively. Since 2003, he has been with the School of Mechanical Engineering, Hubei University of Arts and Science. He is currently a Professor and the Leader of the Laboratory of Intelligent Manufacturing and Machine Vision. Over the last ten years, he has authored over 50 journal and conference papers. His current research interests cover smart factory, machine vision, and image processing. He is a Senior Member of the Chinese Mechanical Engineering Society.