Spring 4-14-2021

Artificial Intelligence-Driven Remaining Useful Life Prediction of a Machinery-A Review

Dvij Barot
Symbiosis International University, barot.dvij@sitpune.edu.in

Honey Sharma
Symbiosis International University, honey.sharma@sitpune.edu.in

Mahika Yadav
Symbiosis International University, mahika.yadav@sitpune.edu.in

Pooja Kamat
Symbiosis International University, pooja.kamat@sitpune.edu.in

Artificial Intelligence-Driven Remaining Useful Life Prediction of a Machinery-A Review

Dvij Barot[1], Honey Sharma[2], Mahika Yadav[3], Pooja Kamat[4]
[1]Symbiosis Institute of Technology (SIT), Symbiosis International (Deemed University), Pune, India.
Email: barot.dvij@sitpune.edu.in
[2]Symbiosis Institute of Technology (SIT), Symbiosis International (Deemed University), Pune, India.
Email: honey.sharma@sitpune.edu.in
[3]Symbiosis Institute of Technology (SIT), Symbiosis International (Deemed University), Pune, India.
Email: mahika.yadav@sitpune.edu.in
[4]Assistant Professor (Computer Science Engineering), Symbiosis Institute of Technology (SIT), Symbiosis International (Deemed University), Pune, India.
Email: pooja.kamat@sitpune.edu.in

ABSTRACT

The Remaining Useful Life of a machine is very useful statistical information for the operator and manufacturer. It provides a very clear perspective to the user how long the machine can be operated and if any faults are detected how can they be prevented and ultimately increase the Remaining Useful Life. If the operators are aware of the forthcoming issues of the machine the downtime caused in the inspection, part delivery and eventually replacing parts is significantly reduced.

The paper presents a study on the remaining useful life of machinery as it is an emerging technique, starting from the year 2010 to the current year of 2021. The bibliometric research is carried out using the Scopus database and software such as the VOS Viewer and GPS Visualizer. The study tells that the maximum publications are from articles and conference papers affiliated to remaining useful life, Chinese lead the publication followed by the United States and then France. Remaining useful lives is the most used keyword followed by forecasting and the least contributing keywords are also found.

Keywords: Bibliometric Analysis; Deep Learning; Remaining Useful Life; Predictive Maintenance; Anomaly Detection; Machinery; LSTM; Autoencoders

1. INTRODUCTION

The modern era is filled with modern electronic devices. Each device is built with certain complex designs and parts and with so many possible points of failure, and after extensive use, it is likely to cause damage to the device. Devices need to be properly analyzed to maximize this flow of control by ensuring the reliability, efficiency, and durability of the machine. Irrespective of the scale of work machines are used for increasing production and efficiency. Nowadays, a variety of industries use electric machinery for different purposes. For certain applications, these machines may not operate as required in certain conditions like high temperature or high moisture which leads to motor malfunction and that may lead to high maintenance cost and fewer financial loss. So it becomes essential to ensure the smooth working of these machines and prevent costly unplanned maintenance. The project aims to propose an approach to monitor the remaining useful life of machinery and build predictive fault detection models. Deep learning is used in our project to detect the anomalies in the machine and the autoencoder to extract the performance features. In many industries, the maintenance cost is quite high and the present systems do not provide a systematic way of modeling early failures.

Machinery is widely used in the industry which involves using gearboxes, motors, heavy-duty lathers, and many more moving parts. Such machines reduce the operating cost and increase efficiency. But to maintain this efficiency, timely maintenance is required, to avoid downtimes and consequent losses. Some important parts of a machine that requires frequent maintenance are bearings, hydraulics, gears, motors, and timing belts. Figure 1 shows a new bearing and a damaged bearing.

Figure 1: Bearing

Machinery is prone to failure of parts over time and many variables are contributing to it. Some of the common causes of failure are having extreme weather conditions, improper working conditions of the machinery like the wrong elevation, and misalignment of the machines resulting in abnormal vibrations and frequent failures. Some factors are caused by wrong torquing of bolts and overrunning machines, which drastically affects the remaining life of machines. Minor mistakes of improper maintenance and untrained operators, ignoring warning signals, or not reading the operator's instructions/manual can cause critical damage resulting in downtime of machinery. Hence, we need to ensure such mistakes and faults are prevented at any point in time to significantly reduce terminal failure and increase the remaining useful life of machinery.

In case of diagnosis of gear pitting faults, the model uses Augmented Deep Sparse Autoencoder (ADSAE) with data enhancement techniques. ADSAE uses Sparse Autoencoders (SAE) to extract information from high dimensional signals into lower dimension signals. The ADSAE algorithm is used to extract features using unsupervised learning methods and process the raw vibration signal, and the network training portion primarily uses the back-propagation algorithm. [1].

To classify a good or bad gearbox, LSTM and BLSTM gear-box diagnosis models are proposed. GMM and KNN are used separately to identify the classification. The input of the model is in the time or frequency domain. BLSTM model outperformed the LSTM model in both the time domain and frequency domain [2].

The model uses Deep Bi-directional Long-Short-term Memory (DB-LSTM) for wind turbine gearbox fault diagnosis, which processes the sequence data in a forward path as well as a backward path with two different hidden layers. Raw time series & frequency data were input of DBLSTM. It gave an accuracy of 99.08% [3].

To diagnose a fault in Oil-Immersed Power Transformers, Dynamic Adam dropout based deep neural network (DADDNN) was used. For the large and high dimensional dataset, gradient-based optimization of objective stochastic functions was used. The dropout technique was used as an efficient way to combine exponentially different neural network architecture which decreases the dynamic co-adaptability of neurons and increases intensity. DNN was used directly to extract high-level features from original data, The DADDNN with its characteristics and Sigmoid and Softplus training functions is triggered with the Relu function.[4].

Fault diagnosis of rotary machine bearings is done by envelope spectra and stacked sparse autoencoder-based deep neural networks (SSAE-based DNNs). The autoencoders were used to extract useful features and then the model was fine-tuned using backpropagation. Discriminant characteristics increased DNN efficiency by doing successful classification of faults when the fluctuations occurred and gave an accuracy of 90% [5].

The DCGAN-based AE model was learned in an unsupervised fashion in order to convert high dimensional readings to low dimensional readings. A double-nested model was proposed to enhance the quality of extracted representation. AE was embedded in DCGAN as a generator. For high-level extraction, normalized data are sent to the model. LSTM was used to capture the temporal info of features in supervised learning. To create images from random noise, Generator G was used. To use multivariate temporal information, it applied the sliding window strategy. In the fine-tuning process, LSTM and FNN were combined to produce the goal RUL.[6].

To estimate the RUL of bearings a two-stage automated approach is used using DNN. To group the obtained signals of bearings into various degradation levels, a denoising autoencoder was used. Each layer of DNN is trained sequentially. The DNN model consists of 3 hidden layers and is created using the raw signal to identify the level of the tracked bearing. A stacked denoising autoencoder (SDA) is used to capture more general features. For the formulation of the RUL predictors, the shallow ANN model with two hidden layers was used. Layer by layer, SDA was used to configure each layer's parameters. In order to calculate the prediction efficiency, the average prediction error was used [7].

Figure 2 shows the advantages and the disadvantages of previously tested algorithms, namely, Recurrent Neural Network (RNN), Artificial Neural Network (ANN), Convolutional Neural Network (CNN), and Autoencoder (AE), along with the algorithm we have implemented, i.e. Long Short Term Memory-Autoencoder.

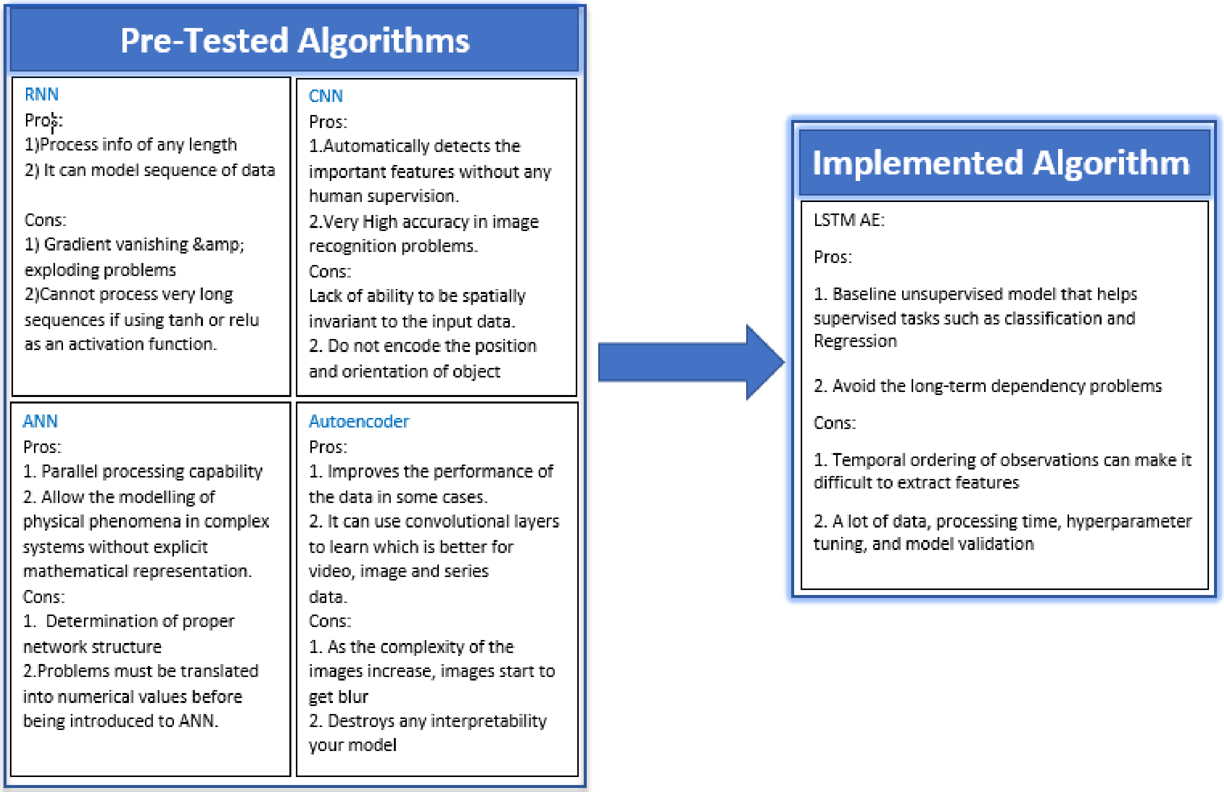

Figure 2: Comparison of previously tested algorithms and the algorithm being implemented in this paper

The basic structure of RNN is shown in figure 3, Autoencoder in figure 4, and LSTM-AE in figure 5.

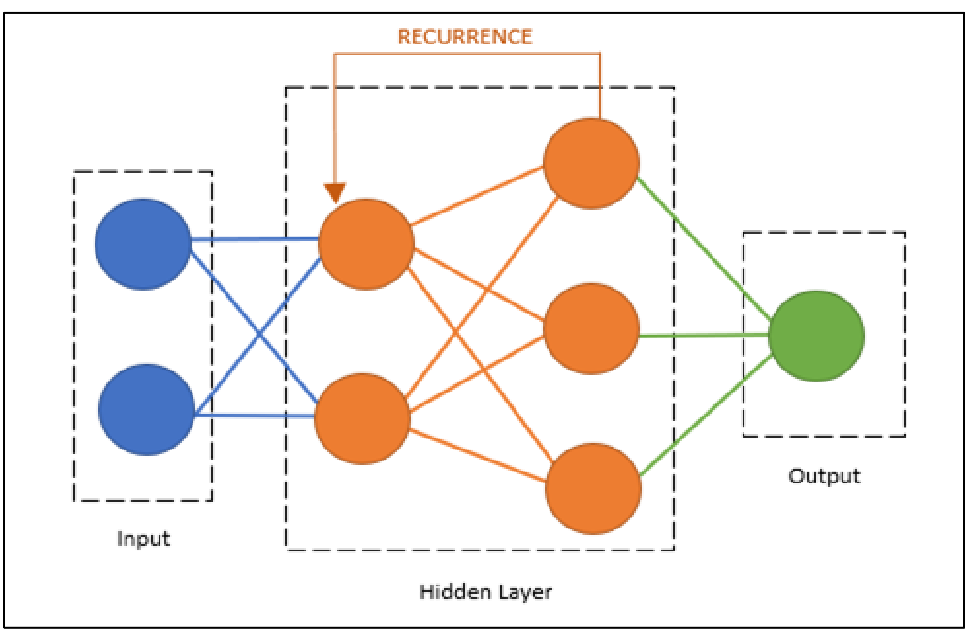

Figure 3: RNN (Recurrent Neural Network)

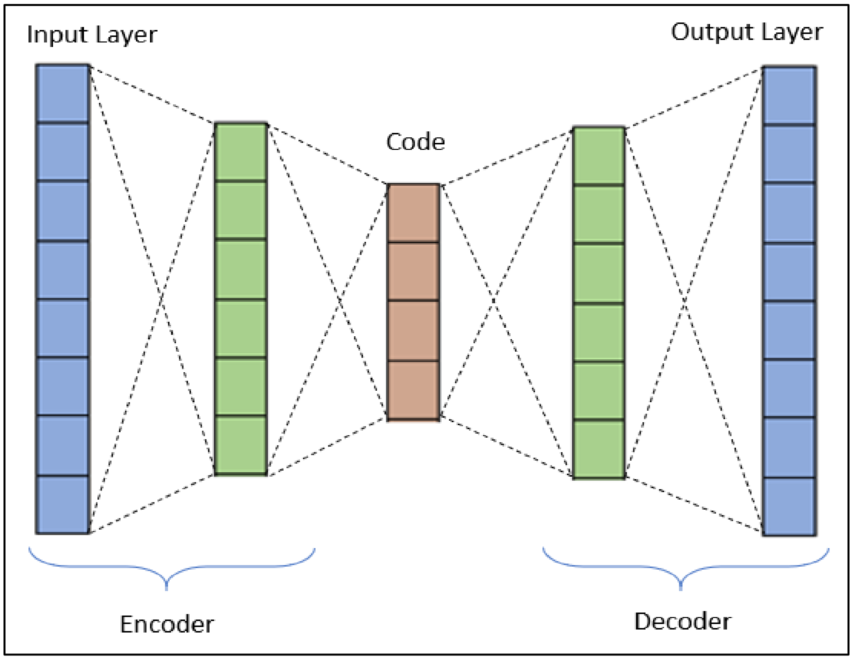

Figure 4: Autoencoder

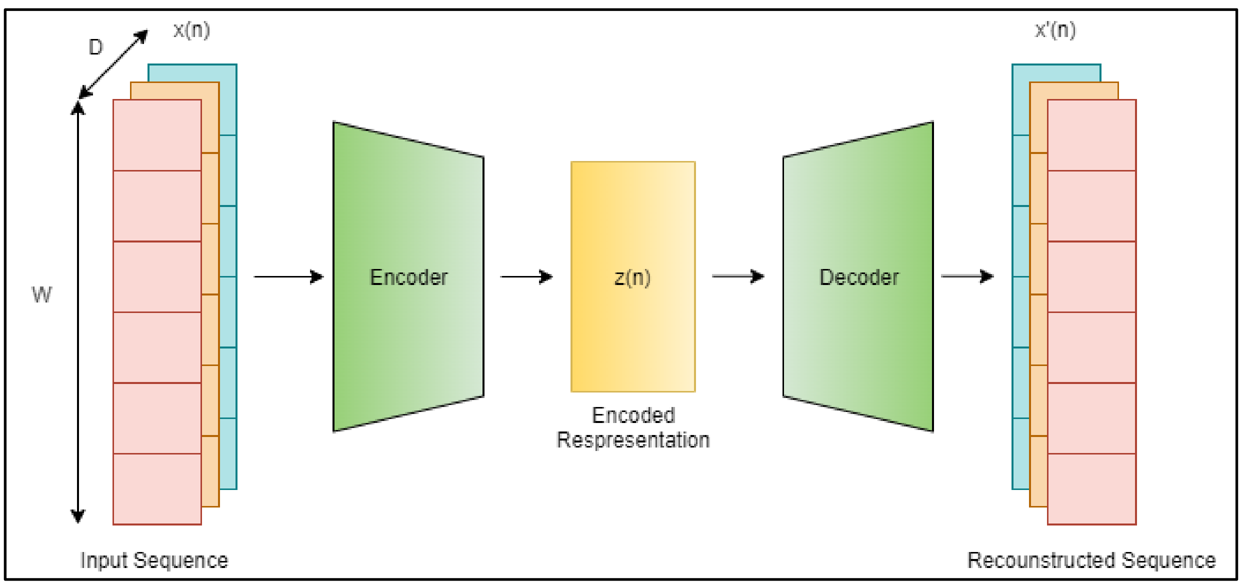

Figure 5: LSTM-AE (Long Short Term Memory-Autoencoder)

2. PRELIMINARY DATA COLLECTION

Two types of publication databases can be gathered through open and paid access. These publication details can be downloaded either from their portals in the university library or by logging independently on various web sources. There are also various common methods for obtaining the data from the appropriate databases. The popular publication databases are Scopus, ScienceDirect, ResearchGate, Web of Science, Google Scholar, SCImago, Clarivate, Mendeley and many more. Scopus is one of the largest peer-reviewed citation databases of published research in the field of technology, engineering, science, and medicine along with social sciences, arts, and humanities. A significant region-based keyword analysis is identified in section 2.1, using the Scopus database.

2.1 Significant keywords

All the important keywords related to the remaining useful life were divided into two categories namely, primary and secondary type. For this analysis, table 1 provides the list of keywords used in the search mechanism.

Table 1: Planned structure of keywords assigned

Primary Keyword	"remaining useful life"
Secondary Keyword (AND)	"machines" OR "artificial intelligence" OR "deep learning" OR "autoencoders" OR "fault detection"

2.2 Primary search outcomes

All the analysis is done using the Scopus database. Initially, investigation resulted in a total of 1223 publications from the intended keyword search. This is then reduced to 1173 English publications only (Table 2). Figure 6 provides the graphical representation of publication count in different languages.

Table 2: remaining useful life publications in different languages

Sr. No.	Publishing Language	Publications Count
1	English	1173
2	Chinese	45
3	German	3
4	Portuguese	1
5	Spanish	1
Total		1223

Source: https://www.scopus.com (accessed on 11th December 2020)

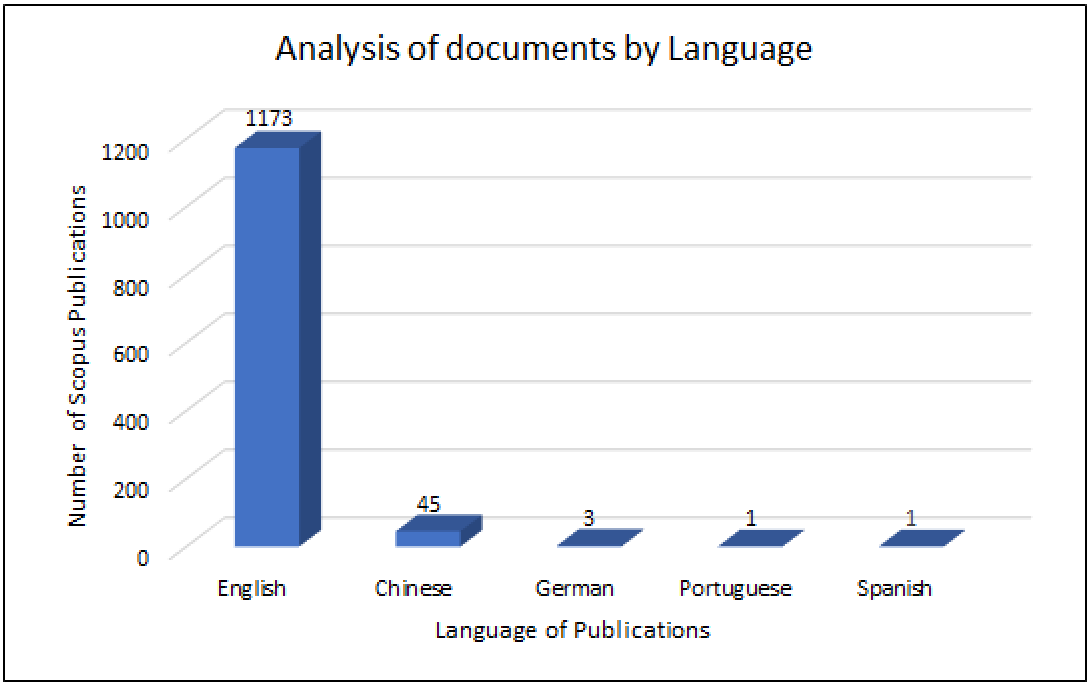

Figure 6: Graphical representation of document analysis by language in remaining useful life
Source: https://www.scopus.com (accessed on 11th December 2020)

All sorts of published and unpublished publications were considered for review purpose. The researchers in the area of remaining useful life have more focus on publications in articles and conference papers. 47.26% of articles were found and 47.09% of conference papers were found (Table 3).

Table 3: Publication types in the remaining useful life

Publication Type	**Number of publications**	**Percentage of 1223**
Article	578	47.26%
Conference Paper	576	47.10%
Conference Review	34	2.78%
Book Chapter	16	1.30%
Review	12	0.98%
Book	6	0.49%
Note	1	0.08%
Total	1223	100%

Source: https://www.scopus.com (accessed on 11th December 2020)

2.3 Initial highlights of data

The documents were collected from articles, conference papers, conference reviews, etc. for the span of twelve years from 2010 to 2021. Analysis for yearly publications is shown in table 4 and figure 7 shows its graphical representation for remaining useful life research areas.

Table 4: Analysis for Yearly publications in remaining useful life

Year	Publications Count	Year	Publications Count
2021	12	2015	87
2020	269	2014	47
2019	260	2013	57
2018	144	2012	46
2017	132	2011	37
2016	99	2010	33
Total			1223

Source: https://www.scopus.com (accessed on 11th December 2020)

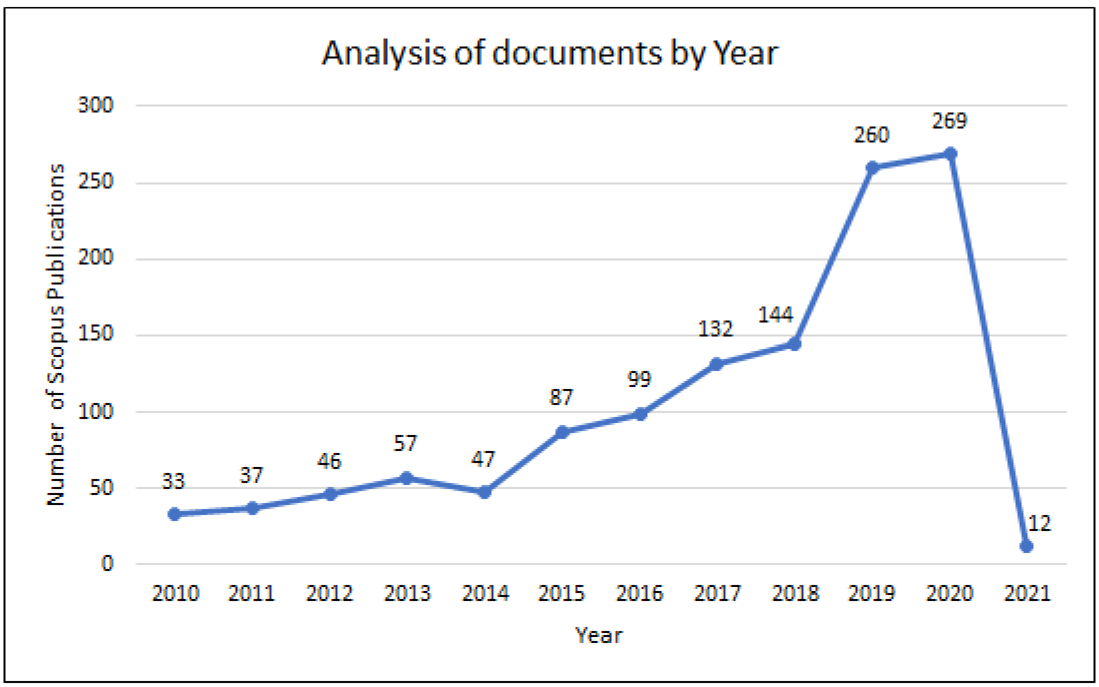

Figure 7: Graphical representation for Yearly publications in remaining useful life
Source: https://www.scopus.com (accessed on 11th December 2020)

2.4 Data validation through databases

In section 3, the detailed bibliometric study was performed to know the uniqueness of literature and leading researchers in the area of remaining useful life through geographical locations, author contributions, affiliation statistics, leading journals and their statistics, citation analysis, and the research study.

3. BIBLIOMETRIC ANALYSIS

To perform a good bibliometric analysis on the remaining useful life, the following two ways are applied which are:

- Analysis of geometric region, their network, and citation, etc.
- Statistical analysis of the keyword, university, affiliation, author, and journal to study data at popularity level.

3.1 Region based geographical analysis

To display the most preferred published papers' geographical position, figure 8 is drawn using www.gpsvisualizer.com. The map clearly shows that Europe is the most densely populated in terms of the number of researchers.

Figure 8: Geographic locations of remaining useful life

Figure 9 depicts the first ten countries having the maximum number of publications in the domain of remaining useful life. Chinese researchers lead by 36%, which is followed by the United States (US) with 22% and France with 7% in publications.

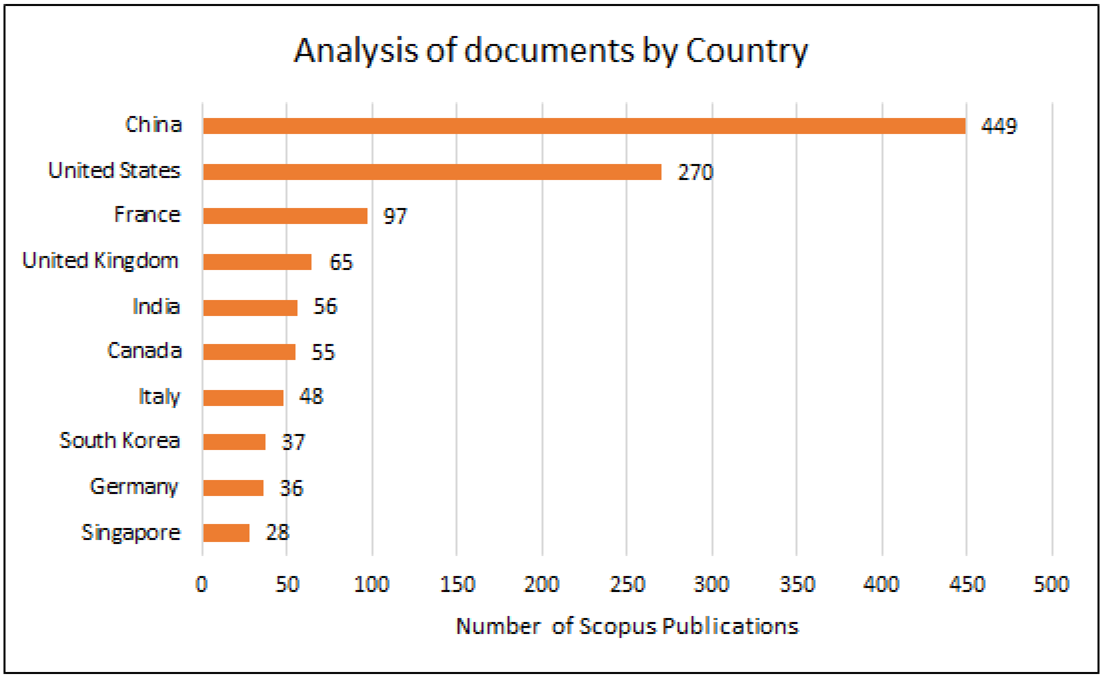

Figure 9: Top ten countries published papers on the remaining useful life
Source: https://www.scopus.com (accessed on 11th December 2020)

3.2 Keywords analysis

Some of the popular keywords are enlisted below. The correct combination of keywords gives the targeted count in the significant research area. Table 5 enlists the first ten keywords for publications on the remaining useful life. Figure 10 shows a graphical analysis of the top ten keywords used in the area of remaining useful life.

Table 5: Top ten keywords for remaining useful life

Keywords	**Number of Publications**
Remaining Useful Lives	783
Forecasting	430
Remaining Useful Life Predictions	323
Remaining Useful Life	316
Systems Engineering	266
Fault Detection	183
Deep Learning	176
Condition Monitoring	174
Learning Systems	167

Health	165

Source: https://www.scopus.com (accessed on 11th December 2020)

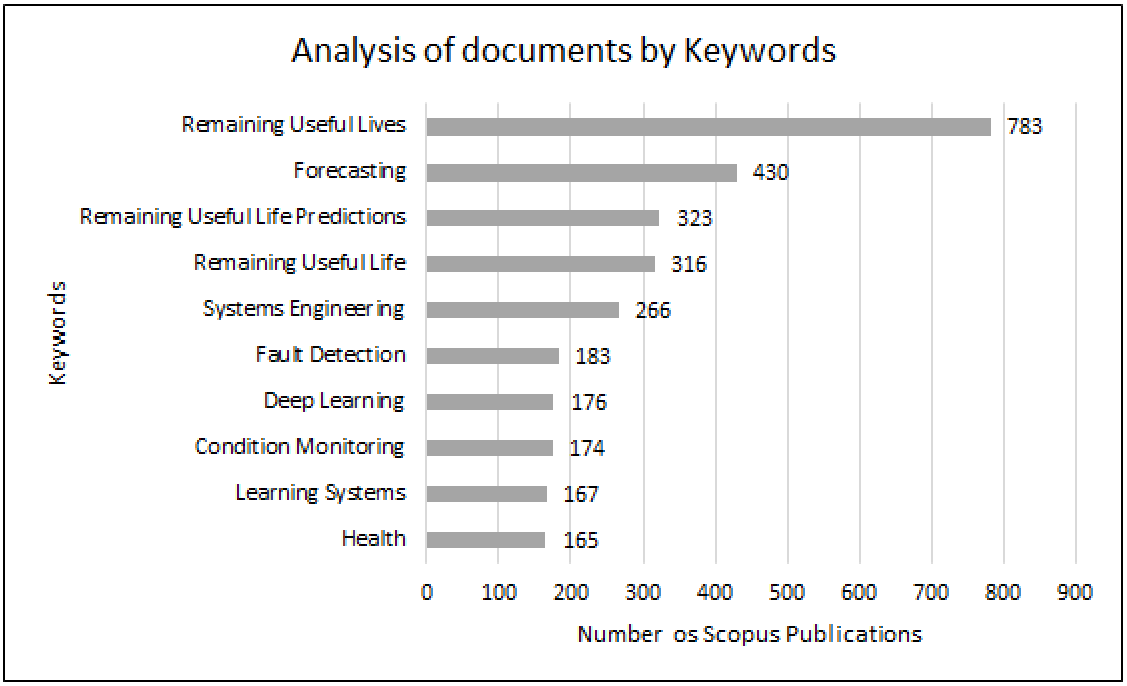

Figure 10: Analysis by Keywords for remaining useful life
Source: https://www.scopus.com (accessed on 11th December 2020)

3.3 Analysis based on Subject Areas

The subject area wise comparison for remaining useful life publications is shown in Figure 11. From the figure, it can be seen that most of the research is carried out in the domain of engineering, computer science, and mathematics. It is also observed that economics and pharmacology have the least amount of research.

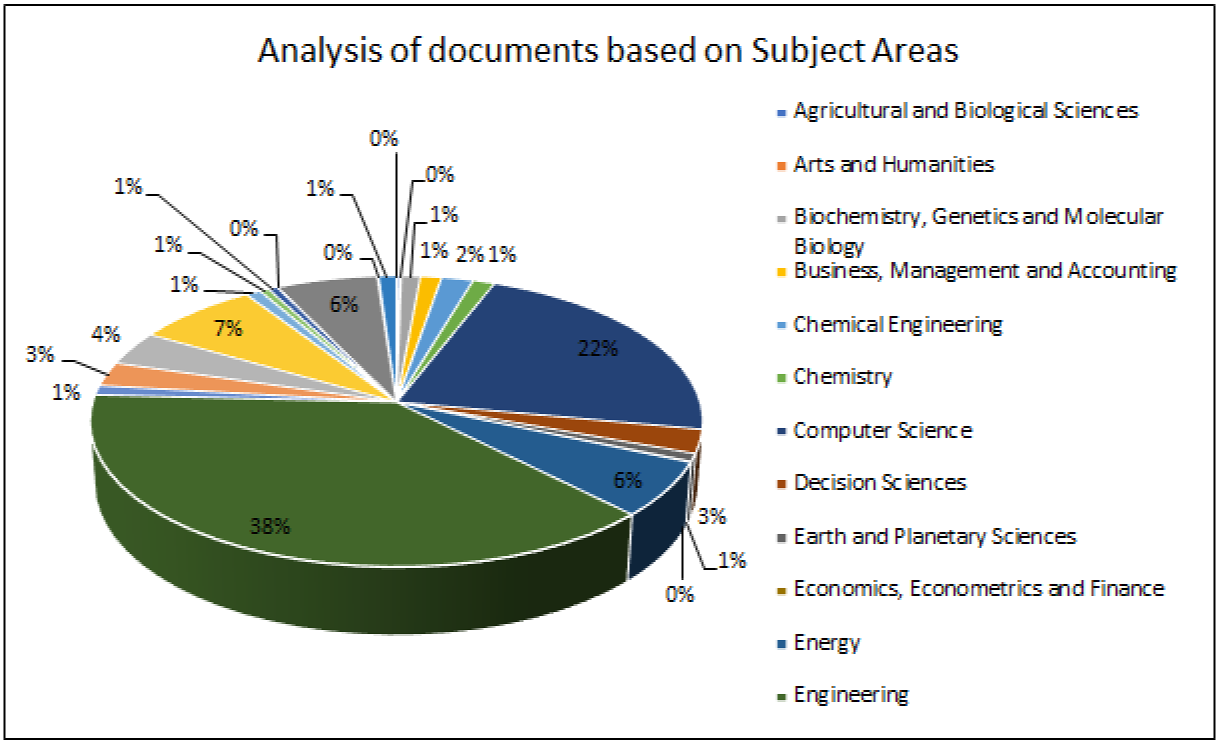

Figure 11: Analysis by Subject area for remaining useful life
Source: https://www.scopus.com (accessed on 11th December 2020)

3.4 Network-based Analysis

By using network analysis tools like VOS Viewer, the relationship between various parameters can be displayed. It is software for analysis of parameters like different authors, keywords, its year, etc. It is capable of generating maps for navigating and clustering network data. Figure 12 shows the cluster analysis of the keywords used by the authors in the prediction of remaining useful life. The larger circle shows the keywords which are widely used.

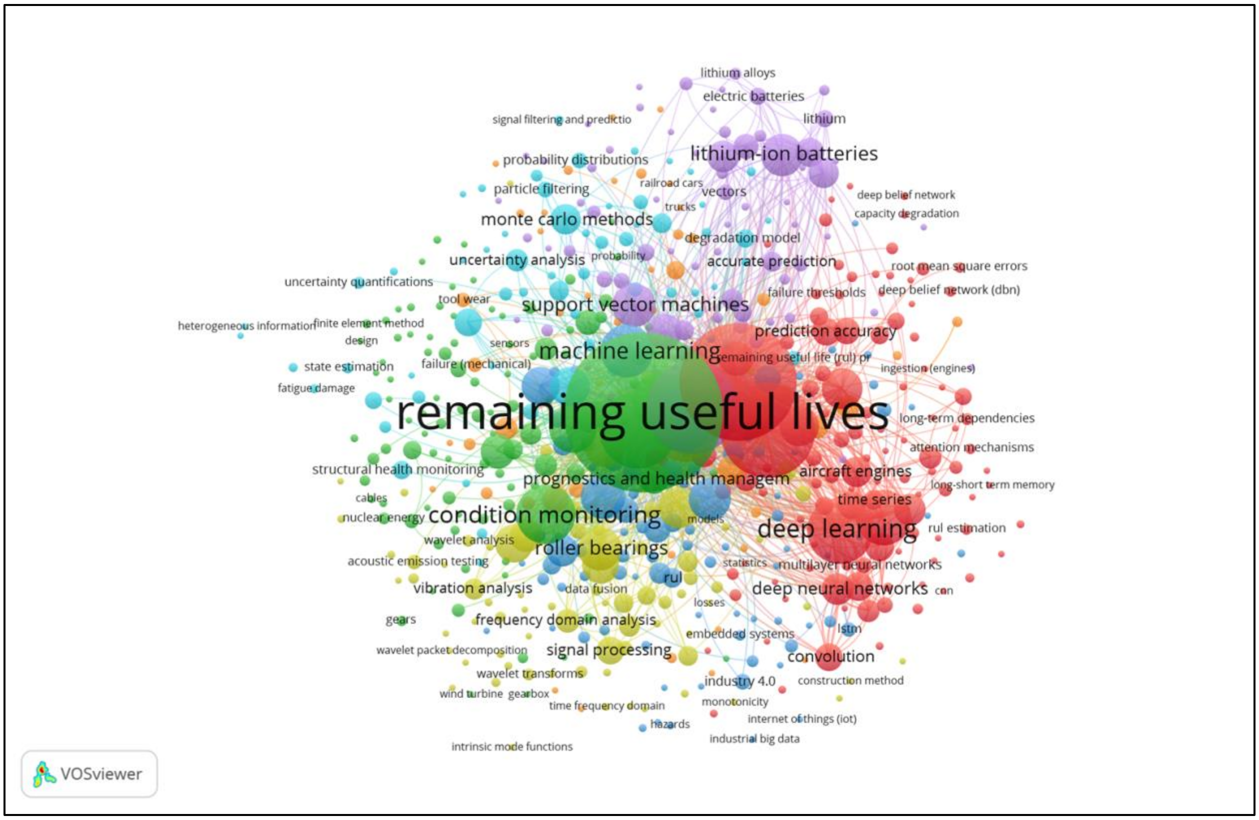

Figure 12: Cluster Analysis of Keywords and Source Title(Data access till 11th December 2020)

Figure 13 shows the cluster analysis of authors and co-authors in the papers. The larger circle shows the authors who wrote maximum papers in the field of remaining useful life.

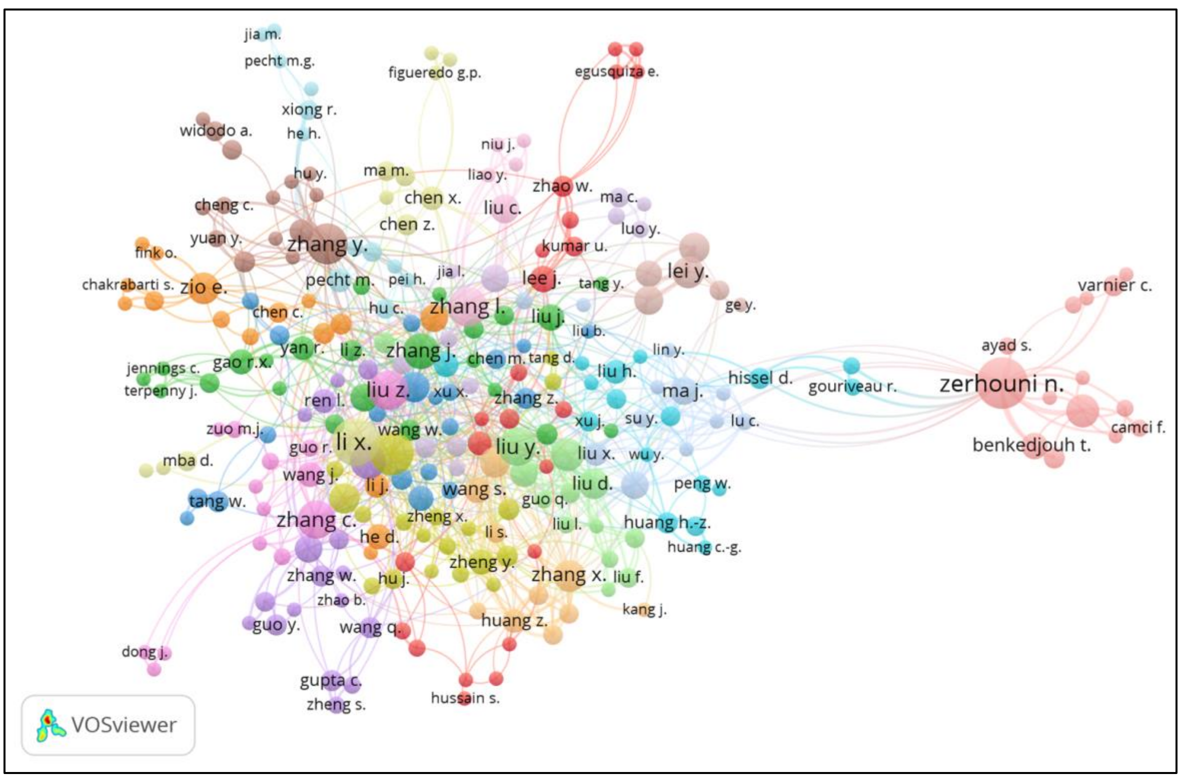

Figure 13: Cluster Analysis of Authors and Co-authors(Data access till 11th December 2020)

3.5 Affiliation based statistics

The top ten universities/organizational affiliations are provided in Figure 14. The remaining useful life for a while has been a domain of research in FEMTO-ST Sciences and Technologies and Beihang University. Also, out of the top ten universities or organizations contributing, six of them belong to China.

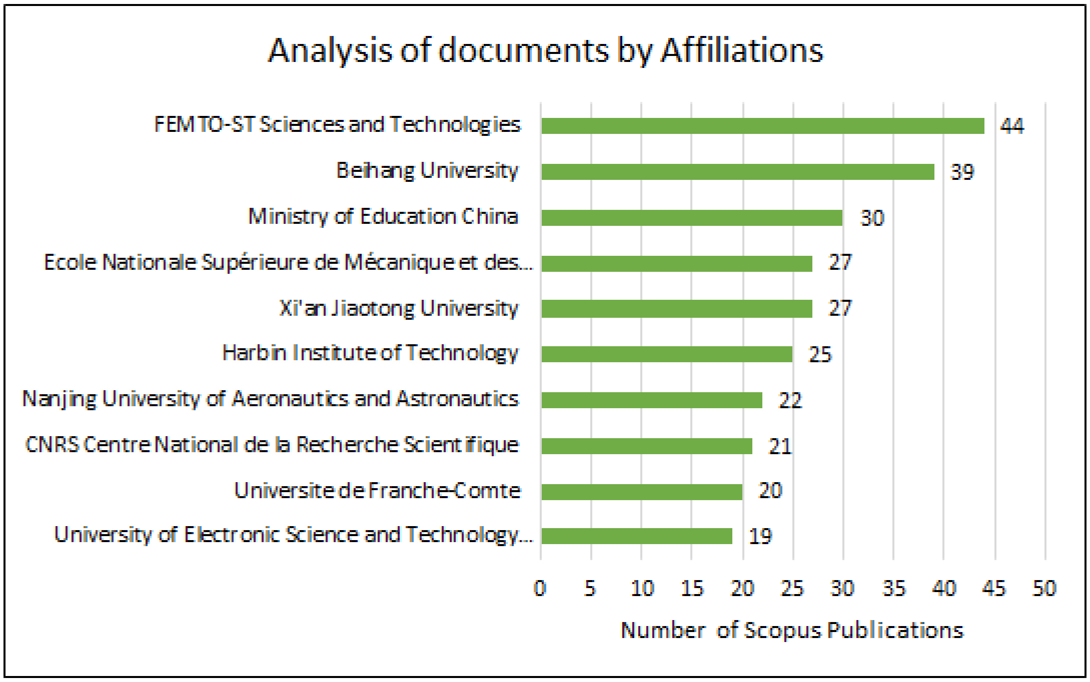

Figure 14: Analysis by Subject area for remaining useful life
Source: https://www.scopus.com (accessed on 11th December 2020)

3.6 Author based statistics

Figure 15 illustrates the top ten authors which have contributed to the domain of remaining useful life to notice the impact of a particular author.

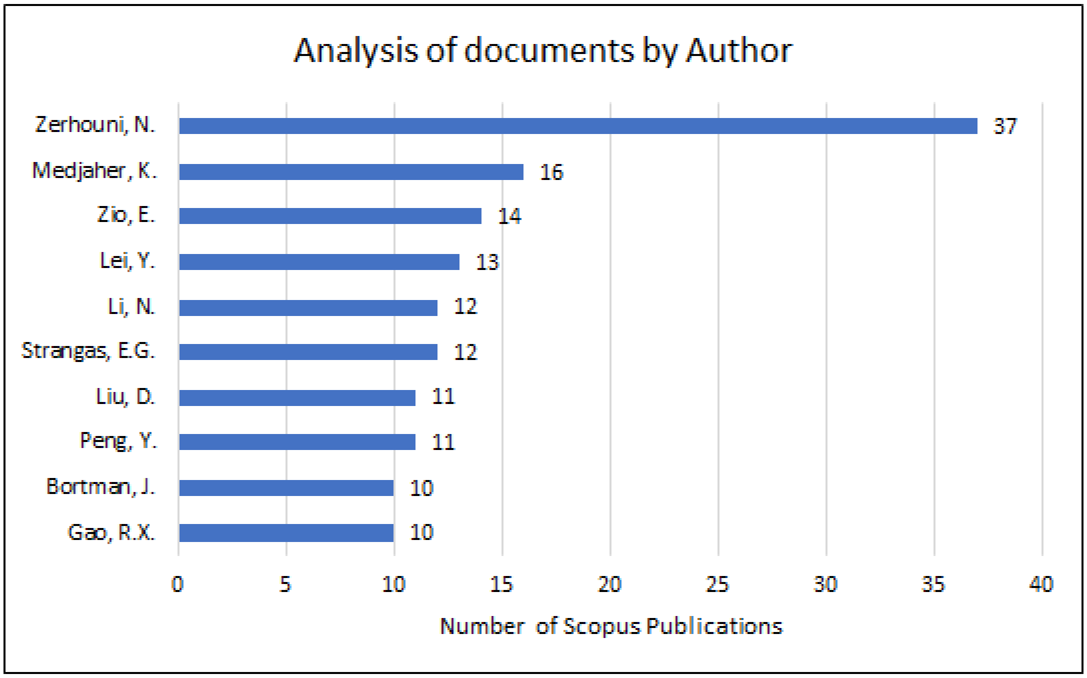

Figure 15: Key contributing authors in the field of remaining useful life
Source: https://www.scopus.com (accessed on 11th December 2020)

3.7 Journal statistics

Figure 16 provides the analysis of publication types for remaining useful life. The collected data indicates that 94% of publications are either from articles or conference papers. Also, it can be seen that only 0.98% of contributors follow review papers. Therefore, this review is written on the remaining useful life.

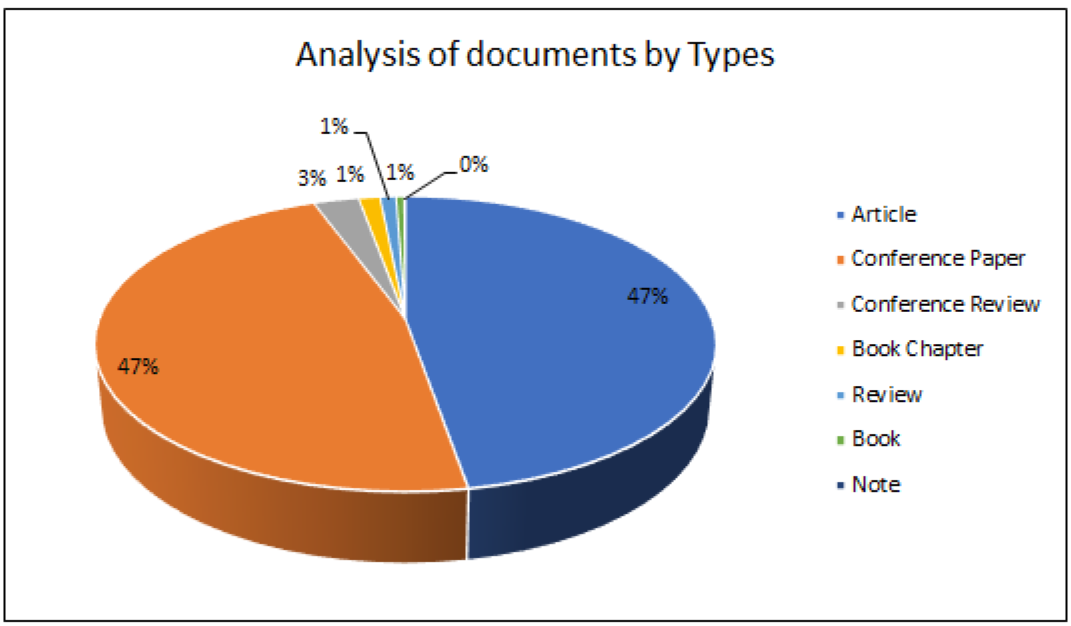

Figure 16: Various types of sources for publications in the remaining useful life
Source: https://www.scopus.com (accessed on 11th December 2020)

3.8 Review based on Citation

Yearwise citations collected in the domain of remaining useful life is shown in Table 6. The overall count of citation was 14466 up to 11th December 2020 out of 1233 publications. Table 7 shows the first top ten papers along with their year-wise citations received to them.

Table 6: Analysis for Citation for publications in remaining useful life

Year	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	>2020	Total
No. of Citations	3	47	107	254	360	587	746	1256	1989	3703	5086	328	14466

Source: https://www.scopus.com (accessed on 11th December 2020)

Table 7: A citation review of the top ten papers in remaining useful life

Title of Publication	Year-wise Citations collected via the publication											
	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	>2020	Total
A review on prognostics and health monitoring of Li-ion battery [8]	2	14	47	44	49	32	66	51	78	47	2	432
A recurrent neural network-based health indicator for remaining useful life prediction of bearings [9]	-	-	-	-	-	-	10	51	102	149	10	322
Bearing health monitoring based on Hilbert-Huang transform, support vector machine, and regression [10]	-	-	-	-	6	27	33	48	70	70	8	262
Remaining useful life estimation in prognostics using deep convolution neural networks [11]	-	-	-	-	-	-	-	13	100	135	6	254
Health diagnosis and remaining useful life prognostics of	-	-	1	13	24	22	27	35	61	48	2	233

lithium-ion batteries using data-driven methods [12]												
Accurate bearing remaining useful life prediction based on Weibull distribution and artificial neural network [13]	-	-	-	-	4	19	32	36	62	57	5	215
Machine health monitoring using local feature-based gated recurrent unit networks [14]	-	-	-	-	-	-	-	17	71	97	8	193
Deep convolutional neural network-based regression approach for estimation of remaining useful life [15]	-	-	-	-	-	2	11	27	60	87	6	193
Battery health prognosis for electric vehicles using sample entropy and sparse Bayesian predictive modeling [16]	-	-	-	-	-	10	33	42	44	54	3	186
Prognostics of lithium-ion batteries based on relevance vectors and a conditional three-parameter capacity degradation model [17]	-	-	3	14	27	21	38	28	30	24	1	186

Source: https://www.scopus.com (accessed on 11th December 2020)

3.9 Source Statistics

Figure 17 shows the source statistics for publications in remaining useful life, it is clear that the maximum number of papers are from Proceedings Of The Annual Conference Of The Prognostics And Health Management Society Phm.

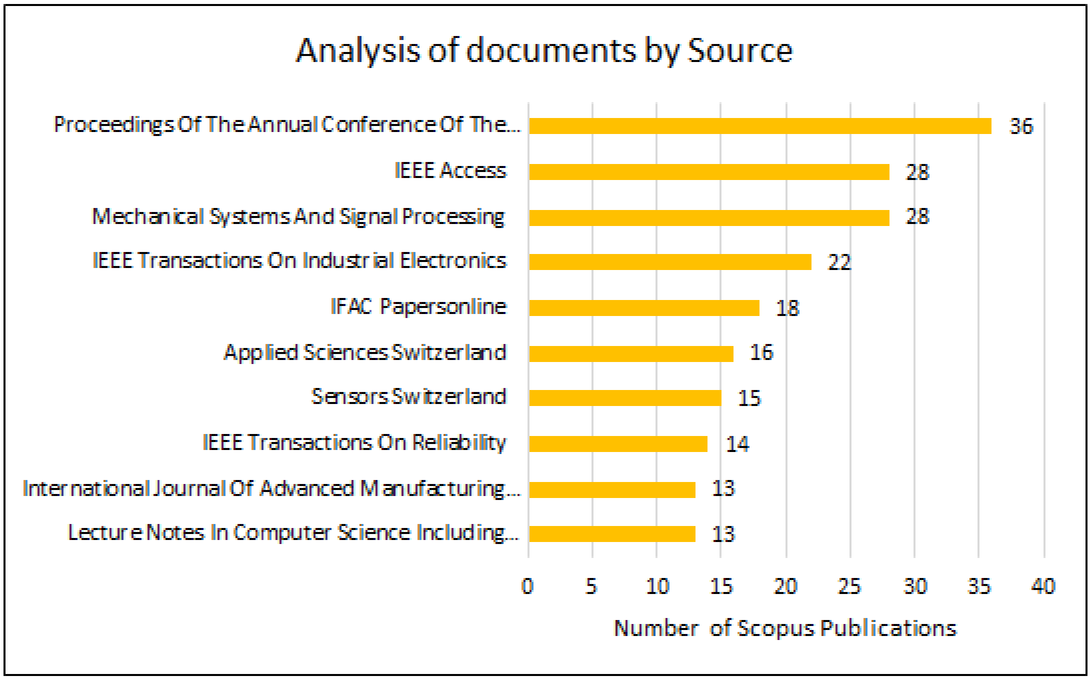

Figure 17: Source statistics for publications in the remaining useful life
Source: https://www.scopus.com (accessed on 11th December 2020)

3.10 Analysis by Funding Sponsor

Figure 18 gives the top ten funding sponsors in the area of remaining useful life. The topmost funding agency is the National Natural Science Foundation (NNSF) of China. More than 190 publications are funded by NNSF to the researchers working on remaining useful life.

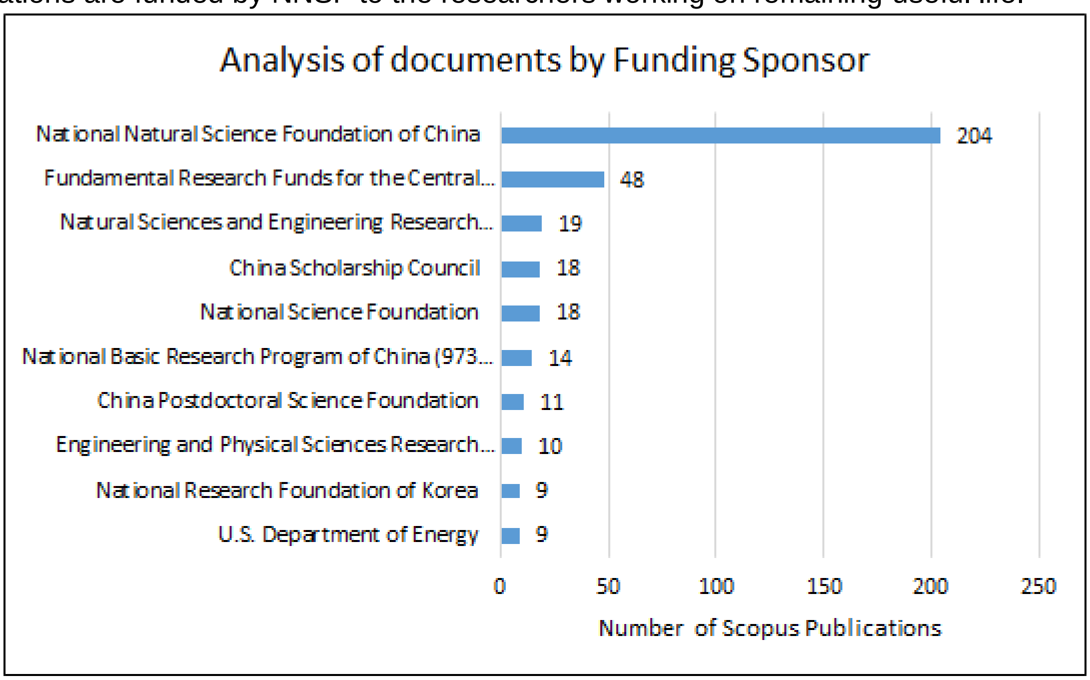

Figure 18: Top funding sponsors for publications in the remaining useful life
Source: https://www.scopus.com (accessed on 11th December 2020)

4. RESEARCH INFERENCE OF THE STUDY

Study in the area of the remaining useful life is continuously increasing. This research laid the preliminary work that will induce innovation and will inform the significance of remaining useful life to bring change by improving their processes for further research.

In table 6 of this bibliometric review, the bottom level keywords on the remaining useful life are parameter estimation, network architecture, machine learning models, maintainability, etc. This indicates a major research gap in these domains and hence critical state should be provided to analysts to concentrate on these domains. They have more scope and hence need to be explored. China and the US are leading in the research field. There exist only seven review papers out of 1073, in the area of remaining useful life.

5. LIMITATIONS OF THE STUDY

For analysis purposes, this review analyzes the Scopus database by using a mixture of multiple keywords. Many important journals and articles were not included in the Scopus collection during the preparation of this report and could therefore not be integrated into the study. This review is limited to English language research papers only.

6. CONCLUSION

This paper presents a bibliometric analysis of the remaining useful life performed using the Scopus database. The study shows that the maximum publications are from journals and conferences. And the past research is mainly in the area of engineering and computer science. China has the maximum number of publications followed by the US and France. The number of publications is showing an increase from 2014, indicating that the research area is growing.

The reviewed literature proposed that more research is needed in the field of predictive maintenance and RUL prediction of machinery to improve the accuracy of the models and reduce the losses. To achieve this many algorithms such as RNN, DNN, Autoencoders, LSTM, etc. were required. The RUL of machinery is an untapped domain of machine learning and needs more research and bibliometric research to develop the technologies and push innovation further.

As per the review literature, we find that the fault tolerances in machinery are undetected most of the time, rendering longer downtime and unnecessary breakdowns and failures. More practical research is needed to improve the current scenario and effectively create a more smooth operation and maintenance procedures for various machinery. There’s a lack of research on acoustic and thermal-based fault detection and remaining useful life detection on Scopus, providing a clear scope of future research targets.

REFERENCES

1. LI, X., LI, J., QU, Y., & HE, D. (2020). Semi-supervised gear fault diagnosis using raw vibration signal based on deep learning. Chinese Journal of Aeronautics, 33(2), 418-426. doi:10.1016/j.cja.2019.04.018

2. Mallikarjuna, P. B., Sreenatha, M., Manjunath, S., & Kundur, N. C. (2020). Aircraft gearbox fault diagnosis system: An approach based on deep learning techniques. Journal of Intelligent Systems, 30(1), 258-272. doi:10.1515/jisys-2019-0237

3. Cao, L., Qian, Z., Zareipour, H., Huang, Z., & Zhang, F. (2019). Fault diagnosis of wind turbine gearbox based on deep bi-directional long short-term memory under time-varying non-stationary operating conditions. IEEE Access, 7, 155219-155228. doi:10.1109/ACCESS.2019.2947501

4. Ou, M., Wei, H., Zhang, Y., & Tan, J. (2019). A dynamic adam based deep neural network for fault diagnosis of oil-immersed power transformers. Energies, 12(6) doi:10.3390/en12060995

5. Sohaib, M., & Kim, J. -. (2018). Reliable fault diagnosis of rotary machine bearings using a stacked sparse autoencoder-based deep neural network. Shock and Vibration, 2018 doi:10.1155/2018/2919637

6. Hou, G., Xu, S., Zhou, N., Yang, L., & Fu, Q. (2020). Remaining useful life estimation using deep convolutional generative adversarial networks based on an autoencoder scheme. Computational Intelligence and Neuroscience, 2020 doi:10.1155/2020/9601389

7. Xia, M., Li, T., Shu, T., Wan, J., De Silva, C. W., & Wang, Z. (2019). A two-stage approach for the remaining useful life prediction of bearings using deep neural networks. IEEE Transactions on Industrial Informatics, 15(6), 3703-3711. doi:10.1109/TII.2018.2868687

8. Zhang, J., & Lee, J. (2011). A review on prognostics and health monitoring of Li-ion battery. Journal of Power Sources, 196(15), 6007-6014. doi:10.1016/j.jpowsour.2011.03.101

9. Guo, L., Li, N., Jia, F., Lei, Y., & Lin, J. (2017). A recurrent neural network-based health indicator for remaining useful life prediction of bearings. Neurocomputing, 240, 98-109. doi:10.1016/j.neucom.2017.02.045

10. Soualhi, A., Medjaher, K., & Zerhouni, N. (2015). Bearing health monitoring based on hilbert-huang transform, support vector machine, and regression. IEEE Transactions on Instrumentation and Measurement, 64(1), 52-62. doi:10.1109/TIM.2014.2330494

11. Li, X., Ding, Q., & Sun, J. -. (2018). Remaining useful life estimation in prognostics using deep convolution neural networks. Reliability Engineering and System Safety, 172, 1-11. doi:10.1016/j.ress.2017.11.021

12. Nuhic, A., Terzimehic, T., Soczka-Guth, T., Buchholz, M., & Dietmayer, K. (2013). Health diagnosis and remaining useful life prognostics of lithium-ion batteries using data-driven methods. Journal of Power Sources, 239, 680-688. doi:10.1016/j.jpowsour.2012.11.146

13. Ben Ali, J., Chebel-Morello, B., Saidi, L., Malinowski, S., & Fnaiech, F. (2015). Accurate bearing remaining useful life prediction based on weibull distribution and artificial neural network. Mechanical Systems and Signal Processing, 56, 150-172. doi:10.1016/j.ymssp.2014.10.014

14. Wang, J., Zhao, R., Wang, D., Yan, R., Mao, K., & Shen, F. (2017). Machine health monitoring using local feature-based gated recurrent unit networks. IEEE Transactions on Industrial Electronics, 65(2), 1539-1548. doi:10.1109/TIE.2017.2733438

15. Babu, G. S., Zhao, P., & Li, X. -. (2016). Deep convolutional neural network-based regression approach for estimation of remaining useful life doi:10.1007/978-3-319-32025-0_14

16. Hu, X., Jiang, J., Cao, D., & Egardt, B. (2016). Battery health prognosis for electric vehicles using sample entropy and sparse bayesian predictive modeling. IEEE Transactions on Industrial Electronics, 63(4), 2645-2656. doi:10.1109/TIE.2015.2461523

17. Wang, D., Miao, Q., & Pecht, M. (2013). Prognostics of lithium-ion batteries based on relevance vectors and a conditional three-parameter capacity degradation model. Journal of Power Sources, 239, 253-264. doi:10.1016/j.jpowsour.2013.03.129