Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 207 (2022) 1057–1066

Procedia Computer Science

www.elsevier.com/locate/procedia

26th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2022)

Data modelling and Remaining Useful Life estimation of rolls in a steel making cold rolling process

Kayal Lakshmanan[a], Eugenio Borghini[a], Arnold Beckmann[a], Cameron Pleydell-Pearce[a], Cinzia Giannetti[a]

[a] *Faculty of Science and Engineering, Swansea University, Swansea SA1 8EN, United Kingdom*

Abstract

The economic cost of roll refurbishment in the steel-making industry is considerable. In a cold rolling mill, wear and damage of rolls disrupt the industrial environment, so it is critical to predict the remaining useful life early and change the roll without causing disruption to the manufacturing process. However, since cold rolling is a complex process affected by multiple variables which are operated in adverse conditions, it is very challenging to mathematically analyse the roll wear and failure. For this reason, in the present paper, a data-driven solution is proposed to predict the correct time for changing individual rolls. To develop an accurate predictive model, several datasets containing high-resolution production data and roll refurbishment data collected from a UK based steel plant have been acquired and processed in a way that the roll wear is modelled as a Remaining Useful Life (RUL) problem, where the number of coils that a roll is able to process is viewed as the remaining cycles. Then hybrid deep learning models are used to predict the Remaining Useful Life of rolls used in steel making. This novel data-driven approach achieves high prediction accuracy and has been validated on a real-world dataset. The proposed approach not only helps avoiding early failure but also can serve as a critical step towards the design of an optimal, automated maintenance schedule for the roll management.

Peer-review under responsibility of the scientific committee of the 26th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2022)

Keywords: Cold mill rolls; Remaining useful life; Convolution Neural Network; LSTM; Bidirectional LSTM

1. Introduction

In steel manufacturing, the cold-rolled strip is produced in a cold strip mill, where the work rolls flatten the strip to a deformed flat shape. The work rolls play the dominant role in a cold rolling mill, making the strip deformed to achieve the required shape. Due to constant use, the material on the roll wears with use and needs refurbishing, which affects the quality of the completed product. Knowing the optimal time to take out the roll for refurbishment, at present, is done based on heuristics or intuition from operators. However, due to the immense cost of refurbishment (up to £500/mm) and potential defects on rolled material, developing a better understanding of the causes of roll wear

E-mail address: k.lakshmanan@swansea.ac.uk

1877-0509 © 2022 The Authors. Published by Elsevier B.V.

Peer-review under responsibility of the scientific committee of the 26th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2022)
10.1016/j.procs.2022.09.161

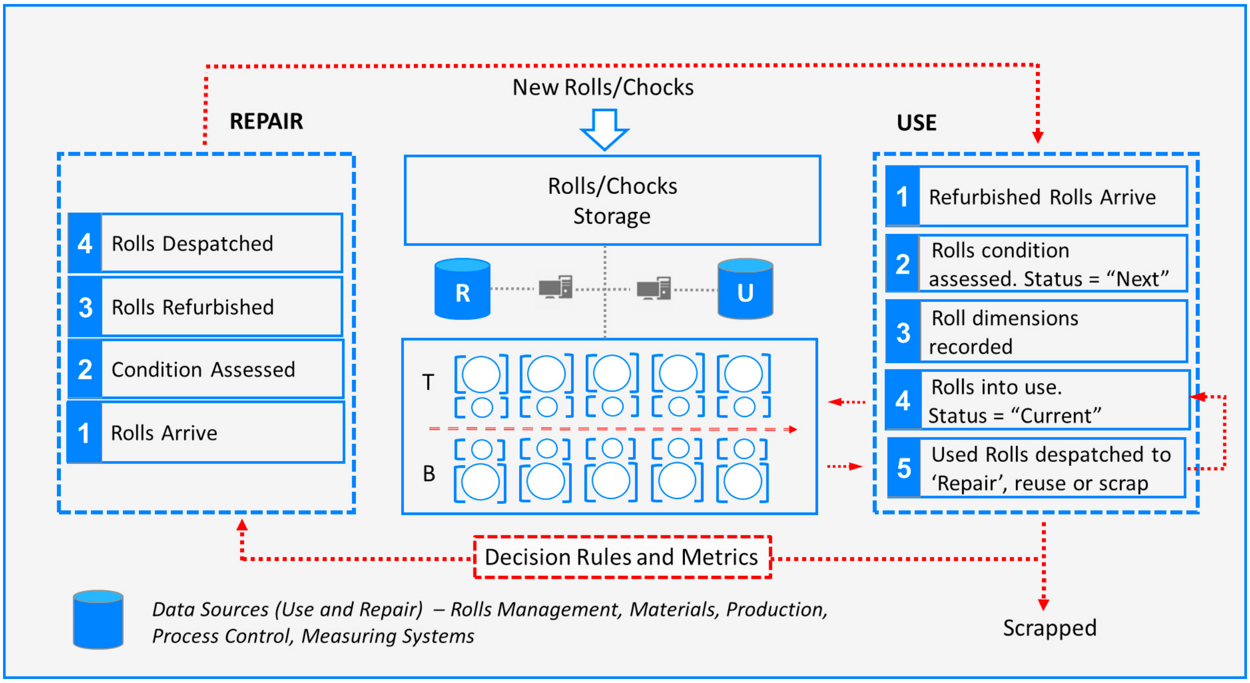


Fig. 1. High-level visualisation of roll management system

will result in financial and experiential benefits to the company. On the other hand, a roll that has passed its time for refurbishment can produce defects in the rolled coils, which also translates to increased spoilage. In addition, the machine sensors in the cold rolling mill automatically collect a vast amount of data in the cold rolling mill that can be exploited to uncover patterns in the dataset, which can be used to explain the roll wear and when to change them.

Currently, the UK based steel industry adopts a heuristic approach to estimating this optimal roll extraction time. However, this is complex, as the RUL of a roll depends on many factors, which need to be understood before developing an advanced predictive maintenance model. This context forms the motivation for this work.

The hypothesis is that a data-driven method of analysing the available data from various sources (i.e., process data, the chemical composition of coils, historical roll wear data, operating conditions, etc.) can lead to an understanding of the causes of roll wear in the cold rolling mill.

1.1. Related Work

The cold-rolled strip is produced on a cold strip mill in steel manufacturing. The work rolls flatten the strip to a deformed flat shape. The work rolls play the dominant role in a cold rolling mill, making the strip deformed to achieve the required shape. One of the main segments of operating costs in a cold rolling mill relates to the work rolls [6]. Many studies, including Lundberg et al. [15] and Robinson et al. [20], have investigated the impact of material, deformation, thermal crown, oxide, and roughness of the roll wear. Roll wear affects the strip quality and works roll service life significantly. The deformation resistance of rolled materials is much higher in cold rolling mills than in hot rolling. Sometimes, the roll bite is subjected to pressures higher than 10,000 MPa. Premature failure of work roll increases rolling cost and causes downtime, affecting productivity significantly. The causes of premature failure can be due to the effects of operating techniques (i.e., rolling load, lubrication, rolling speed, operator experience, etc.) and roll factors. The work roll quality includes non-metallic inclusions, defects, and phase transformations.

Figure 1 shows a graphical conceptualisation of the roll management procedure in a cold rolling mill. The cold rolling mill has five stands, with each stand having a work and backup roll. There is also a top and bottom roll pair (i.e., work and backup). After a fixed amount of tonnage (determined by steel industry), the rolls are extracted and sent for refurbishment.

Several studies have applied data-driven techniques in the literature towards modelling, simulating, optimising, visualizing, and understanding roll wear analysis. For instance, mathematical sets of equations are used to model and simulate tandem cold metal rolling, where Alves et al. [3] developed a state-space model used to conduct analyses in an open loop in the study. Furthermore, a machine learning algorithm is used to quantify the probability of strip snap occurrence by chen et al. [4]. Li et al. [14] investigated the premature failures of several work rolls on a cold strip mill where the study evaluated the chemical compositions, microstructures, and the hardness of roll materials. Findings from the study revealed that the operating factors and metallurgical defects affected the roll service life in cold strip rolling. Similarly, Lanzutti et al. [13] performed the failure analysis of a roller used for cold rolling starting from a comprehensive material investigation to a fracture surface characterisation utilising scanning electron microscopy with energy-dispersive X-ray spectroscopy. Interactive data visualisation of chatter conditions in a cold rolling mill is presented by Perez et al. [18]; this study suggests a visual analytic approach, which is considered exploratory analysis to understand and discover the conditions under which chatter appears. An interactive web-based interface was also developed, which allows a user to explore a map of dynamical conditions, and visualize details of each chatter onset. Mao et al. [16] presented a data-driven remaining useful life (RUL) prediction of roll bearing where the authors present a deep feature representation of learning and transfer learning. The proposed method included two stages – offline and online stages, respectively. In the offline phase, the Hilbert–Huang transform marginal spectra of the raw vibration signal of auxiliary bearings are first calculated as the input, and then contractive denoising autoencoder is introduced to extract deep features with good and stable fault representation. Second, using the obtained deep features and Pearson's correlation coefficient, a new health condition assessment method was implemented, which divided the entire life of each bearing into normal and fast-degradation states. Finally, using the extracted deep features and their RUL values, an RUL prediction model for the fast-degradation state is trained through a least-square support vector machine. In the online stage, the study adopted a transfer learning algorithm using transfer component analysis, where the target features are sequentially adjusted from auxiliary bearings, and then the corresponding RUL is predicted using the corrected features. Pan et al. [17] proposed a two-stage prediction model of the residual life of bearings, wherein rolling-element bearings operation state and the stage before failure of rolling-element bearings failure is divided into normal operation stage and degradation stage. During the normal operation stage, the bearings degradation trend is based on the univariate prediction principle and the 'feedback extreme learning machine model,' and in the degradation stage of rolling-element bearings, the RUL prediction is performed by the Multivariate feedback extreme learning machine model.

Machine learning and deep learning algorithms, namely Multilayer perceptron, recurrent neural network, deep neural network, and Convolutional Neural Network (CNN), etc., have been widely applied to improvement of manufacturing processes. Essian et al. [9] proposed a end-to-end model for multistep machine speed prediction for aluminium can manufacturing based on deep convolutional LSTM encoder-decoder architecture. Torquato et al. [21] performed several machine learning algorithms for improving manufacturing process of electric powertrain. In addition, machine learning and deep learning algorithms have been widely used for fault detection [12], classification of changeover events [8] and RUL prediction in various applications [5].

Zhang et al. [22] designed a hybrid CNN-LSTM algorithm to combine feature extraction, fusion, and regression efficiently to facilitate the prediction based on segmented signals for multiple cutting tools. Remanda et al. [19] investigated the RUL prediction of an aircraft engine and validated with an open-source C-MAPSS dataset using hybrid architecture. A novel multi-sensor data-driven RUL prognostic approach using Auto-Encoder Quasi-Recurrent Neural Networks for RUL Prediction of Engineering Systems is proposed by Cheng et al. [5]; this study adopted three prognostic benchmark datasets to evaluate the proposed model, including a turbofan engine dataset, a rolling bearing dataset, and a machining tool dataset. Beside industrial applications hybrid CNN and LSTM models have been used in univariate timeseries forecasting problems such as traffic flow predictions and temperature predictions [10].

1.2. Main Contributions

The operation of industrial metallurgical plants is associated with compelling wear in spindles, gearboxes, and bearings where it is difficult to implement digital diagnostic tools due to harsh operating conditions [11]. A reliable vibration monitoring is very complicated due to inherent changes of technological regimes. Traditional computational approaches to estimate rolls wear that use simulations or mathematical models may require specialised resources and may fail to capture specific conditions due to the complexity of rolls' operating conditions. Motivated by recent advances in the field of Deep Neural Networks this paper aims to use various sources of data (including chemical and material properties of the coils, sensor collected data, historical refurbishment data) to understand roll wear in the cold rolling mill through the application of data analytics techniques. These data are aggregated in a unique way that allows modelling the life cycle of a given roll within the mill as a sequence of steps, each representing a coil flattened by the mill. The combination of the employed predictive analytics with the data aggregation scheme leads to the development of a novel predictive analytical model, which can be used to predict the remaining useful life (RUL) of a given roll.

For the prediction of RUL, Deep learning methods are utilised. CNN, LSTM, and bi-directional LSTM are commonly used deep learning methods. Since our dataset is modelled as time instances, the aim is to design a hybrid model of CNN, bi-directional LSTM, and LSTM layers to exploit the characteristics for developing an efficient RUL prediction problem. For example, CNN is used for feature extraction purposes, and bi-directional LSTM and LSTM are used for sequence learning purposes. Even though this type of hybrid methodology was used for different applications previously, it has never been used to predict the remaining life of rolls in the cold rolling mill, as well as the real-world experimental chemical and material properties dataset. Also, a comparison is made with different hybrid deep learning methods to identify the best possible model in terms of prediction. The proposed model can be used as a step towards achieving improved predictive maintenance schedules, which will be beneficial to avoid the sudden disruption.

2. Steel manufacturing process dataset

The roll management was based on a comprehensive database made available by the UK-based Steel industry. With the aim of modelling the remaining useful life of rolls, the tables containing sensor information recorded with high resolution, production data (including the chemical and material composition of the coils), and historical refurbishment records (containing chamber and diameter targets, stress tests required and reason for raising refurbishment orders among other columns) together with a tracking of the life-cycle of each roll have been consulted and aggregated.

The life-cycle of rolls is organized abstractly in *roll unit trips*. A trip is a period spent by a pair of rolls in the mill. Usually, one of the rolls serves as a backup, and they are paired in advance according to their size and condition to be inserted into specific mill stands at the following change. The main features of trips include a unique identifier, identifiers for the pair of rolls, timestamps that indicate their start and end, and the reason why they were changed, which could be tonnage rolled, damage, or schedule reasons, among others.

As it is hypothesized that the health status of the rolls is affected not only by the rolled tonnage but also by the composition of the material, the roll unit trips information is enriched with the production dataset. In this table, each row represents a rolled coil and contains details such as a unique identifier, a timestamp, the length and weight, the desired final width and chemical composition (percentage of carbon, silicon, sulphur, copper, phosphorus, arsenic, etc.).

Building on these data, the precise way the prediction of the wear of rolls is framed as an RUL estimation problem is one of the novel contributions of this work. In contrast to the standard RUL estimation setting, there is no lifetime or run-to-failure time series out of the box in the available data. Instead, a sequence was built for each roll unit trip where each step corresponds to a rolled coil. For the present paper, the *life span* of a roll unit trip is defined as the length of the corresponding sequence. Thus, for a given step in the sequence, the RUL of the roll unit trip is the number of steps that separates it from the end of the sequence (see Figure 2 for a schematic explanation of the data aggregation process). It is worth noticing that most roll unit trips are terminated due to scheduled maintenance, while only about 20% of them ended up in some form of damage to the roll.

The resulting dataset consists of 726 sequences with an average length of 81 steps (i.e., the average life of the roll). Besides tracking information, such as trip sequence, roll, and coil identifiers, each step contains details on the cumulative work performed by the roll so far (such as the total tonnage and meterage rolled) and the chemical composition of each coil, totalling 172 variables.

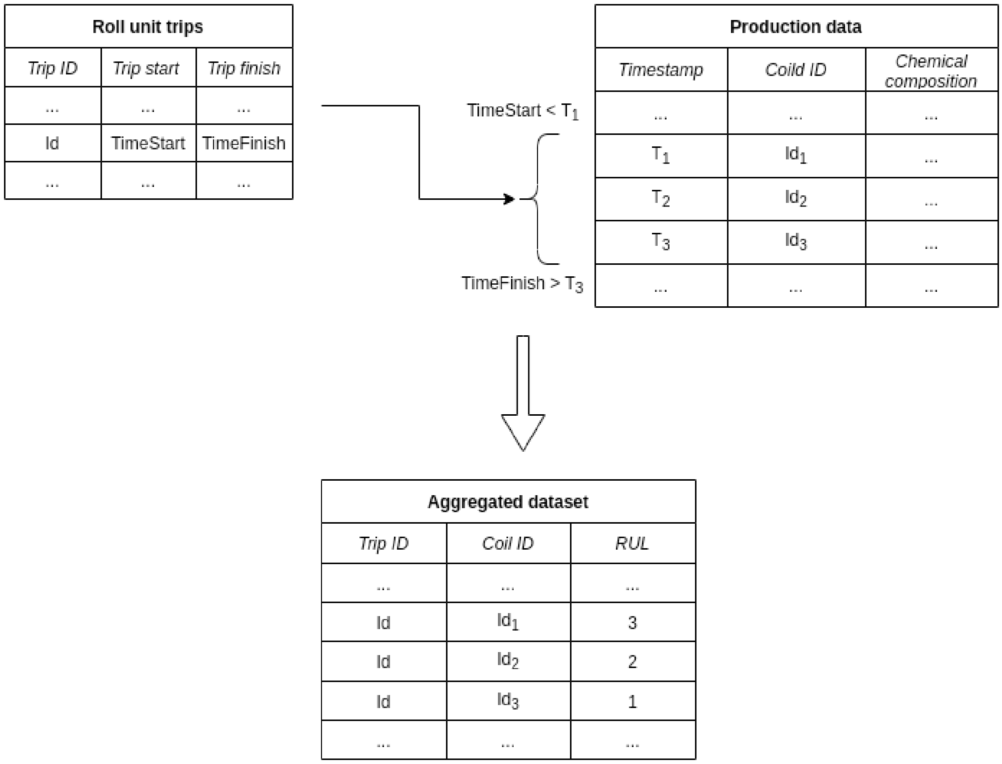


Fig. 2. Illustration of the data aggregation procedure.

3. CNN and LSTM Framework for RUL prediction

A Convolutional Neural network (CNN) is a multi-layered feed-forward neural network model. A CNN consists of a convolutional layer, max-pooling layer, and a fully connected layer. Unlike traditional artificial neural networks, CNN extracts features by itself according to the spatial principle, and a max-pooling layer is utilised for feature reduction. CNN can extract the spatial features from the raw input data and realise a high-dimensional feature representation of original data. The convolutional layer uses convolutional kernels to convolve local regions of the data to develop corresponding features. The parameter sharing of a CNN decreases the parameters in the neural network's calculation process, which effectively avoids having too many parameters, which leads to overfitting the neural network. In this way, the overfitting of the neural network is resolved [1].

The Long Short-Term Memory (LSTM) neural network as a specific RNN has a complex LSTM cell structure in its hidden layer. The LSTM has three gates: input gate, forget gate, and output gate, which control the information flow through the cell and the neural network. LSTM affects the state of RNN at each time step through the gate structure. The gates only limit directions of the information flow. The output value at any neuron $(t + 1)$ depends on its input at the moment t [7].

A Bidirectional LSTM (BiLSTM) was used to extract the long-term dependency characteristics of the sample data. The BiLSTM was comprised of two independent LSTM neural networks, and it has two directions, namely forward and backward propagation.

For predicting the RUL of the individual roll in cold rolling mill, a hybrid deep learning method is employed. A combination of CNN, LSTM, and BiLSTM LSTM layers are considered for the prediction. Specifically, CNN is used for the feature extraction method, and BiLSTM LSTM and LSTM are used for temporal sequence learning. The considered deep learning methodology is depicted in Figure 3.

The raw input data acquired from the cold rolling mill, which are modelled as time sequences to compute RUL, is discussed in section 2. The time sequences are given as input to the convolutional layer for the feature extraction to

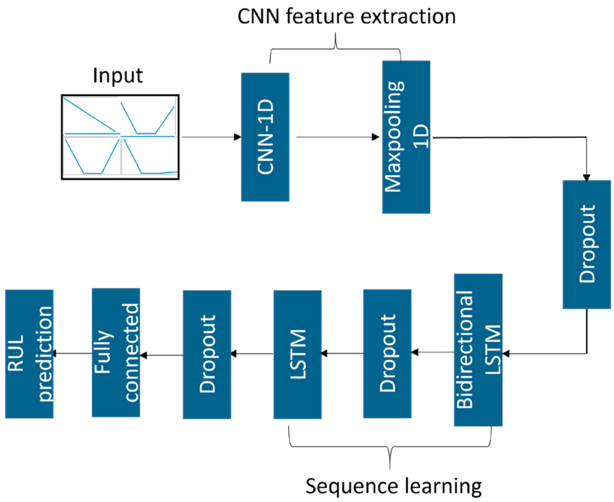


Fig. 3. Methodology of the utilised hybrid deep learning model

obtain the most critical attributes from the dataset. Since all the variables are time instances, the input dataset was first applied to the min-max normalisation technique. The CNN feature extraction block consists of one 1D convolutional layer, a max-polling layer, and a Rectified Linear Unit (ReLU) layer. The convolution layer's output is called feature map, which has a limitation that it keeps track of the precise location of the features in the input, which is to say if there is a small improvement in the location of the feature in the input will lead to a different feature map. To mitigate the limitation of the invariance of the produced feature map, the pooling layer is widely added after the convolutional layer. The ReLU activation function helps to enhance the capability of the model for learning complex structures and is resilient against the gradient vanishing problem [2]. The dropout layer includes the random selection of neurons and deactivating some of them in the training process, which helps prevent overfitting. In this work, the dropout layer is added between CNN feature extraction block and LSTM layers, between BiLSTM and LSTM layers, and between the LSTM sequence learning block and fully connected layer.

In the sequence learning block, one BiLSTM layer and one unidirectional LSTM are employed. In the BiLSTM layer, the return sequence is set to be true so that the network will output the complete sequence of hidden states whereas, in the last unidirectional LSTM layer, the return sequence is said to be false, so that the LSTM network will output the hidden state at the final step. Finally, the fully connected layer has one neuron evaluating each RUL sequence.

Table 1. Configuration of the deep learning layers of the developed hybrid model

Layers (type)	Parameters	
Convolution-1D	Kernels	64
MaxPooling	-	-
ReLU	-	-
Dropout1	-	0.2
BiLSTM	Hidden node	32
	Return sequence	True
Dropout2	-	0.2
LSTM	Hidden node	16
	Return sequence	False
Dropout3	-	0.2
Time distributed Dense	-	1 (sequence output)

Since each dataset has its own unique representation, the number of CNN layers and the BiLSTM and LSTM layers are adopted sequentially by following the output and assessing Mean Squared Error (MSE) and R^2 score values. The CNN's kernel, BiLSTm neurons, and LSTM layers neurons batch size and learning rate are selected by the random search method. The well-known 'Adam' optimiser is utilised in this work, and MSE is used as a loss function. Table 2 shows the parameter settings of the developed hybrid model.

Table 2. Parameter settings of the developed hybrid model

Model parameters	values
Optimiser	Adam
Loss function	MSE
Learning rate	0.001
Epoch	300
Early stopping	Patience=10
Batch size	32

In this work, the dataset is split into 70% training, 20% testing, and 10% validation. The training and validation data is first loaded then the training process is initialised. Meanwhile, the epoch is assigned to 300 with early stopping criteria. This criterion helps stop the training when the training process does not progress anymore, which is assessed through the MSE, used for monitoring validation loss for each epoch. Once the training process is completed, the test dataset, which is not involved in the training process, has been tested, and evaluation metrics, namely R^2-score, Mean Squared Error (MSE), Mean Absolute error (MAE), have been computed.

The presented model is compared with different hybrid deep learning models, namely CNN feature extraction block with other sequence learning networks, namely BiLSTM layer, LSTM layer, or two LSTM layers.

Another assessment has been made considering the test dataset, which contains only the damaged rolls that ran below the roll's average lifetime. By considering this, the credibility of the trained model can be proved since the model is trained on combined normal working conditions and damaged rolls.

4. Comparative results of hybrid deep learning models

In this work, the database of the roll management system from the steel industry has been collected. Without a lifetime or run-to-failure dataset, the roll management system's dataset is modelled as an RUL problem to be able to schedule the change the roll before it wears off successfully. Also, among the time sequences, the damaged rolls are identified, i.e., those with life below the average lifetime. Data modelling has been discussed in Section 2. Finally, the modelled time sequences are partitioned into training 70%, testing 20% and validation 10%.

The developed hybrid model (Model A) discussed in Section 3 are trained using the training dataset and validated using the validation dataset during the training process. Then the test dataset is tested against each trained model at the end of the training process. Finally, the evaluation metrics such as R^2 score and MSE are computed. Solely for the comparison reasons, different but similar existing hybrid models such as CNN-BiLSTM (Model B), CNN-LSTM with LSTM 1 layer (Model C), and CNN-LSTM with LSTM 2 layers (Model D) are trained and tested on the test dataset. Table 3 shows the evaluation metrics calculated for all hybrid models for the test dataset, where model A has a higher R^2 score and lower MAE compared to other models. In addition, for the hybrid models listed in table 3, the train and validation losses are shown in Figure 4, where except model C, other models train and validation losses are converged smoothly.

In the cold rolling mill, the roll is removed after it has rolled for a specific amount of tonnage. Currently, this is evaluated using preventive maintenance, i.e., the maintenance is regularly scheduled according to the usage. In this context, this is assessed after the cumulative amount of tonnage has been rolled. The roll that wears early in the process causes disruption in the industrial environment; that is why it is essential to predict the damaged rolls early and change the roll without causing disruption to the manufacturing process. The damaged rolls are the ones that ran below the roll's average lifetime. Utilizing the same trained models, the evaluation metrics calculated for all hybrid models for

Table 3. Evaluation metrics for the test dataset tested against trained models.

Hybrid models	R^2 score	MSE	MAE	Computational time
CNN-BiLSTM-LSTM (Model A)	**0.9819**	**3.659269e-09**	**3.4542e-05**	518.83 s
CNN-BiLSTM (Model B)	0.9382	1.248579e-08	8.0111e-05	**476.88 s**
CNN-LSTM (1 layer LSTM) (Model C)	0.6396	7.287792e-08	2.2212e-04	699.38 s
CNN-LSTM (2 layers LSTM) (Model D)	0.6404	7.272861e-08	1.7002e-04	769.19 s

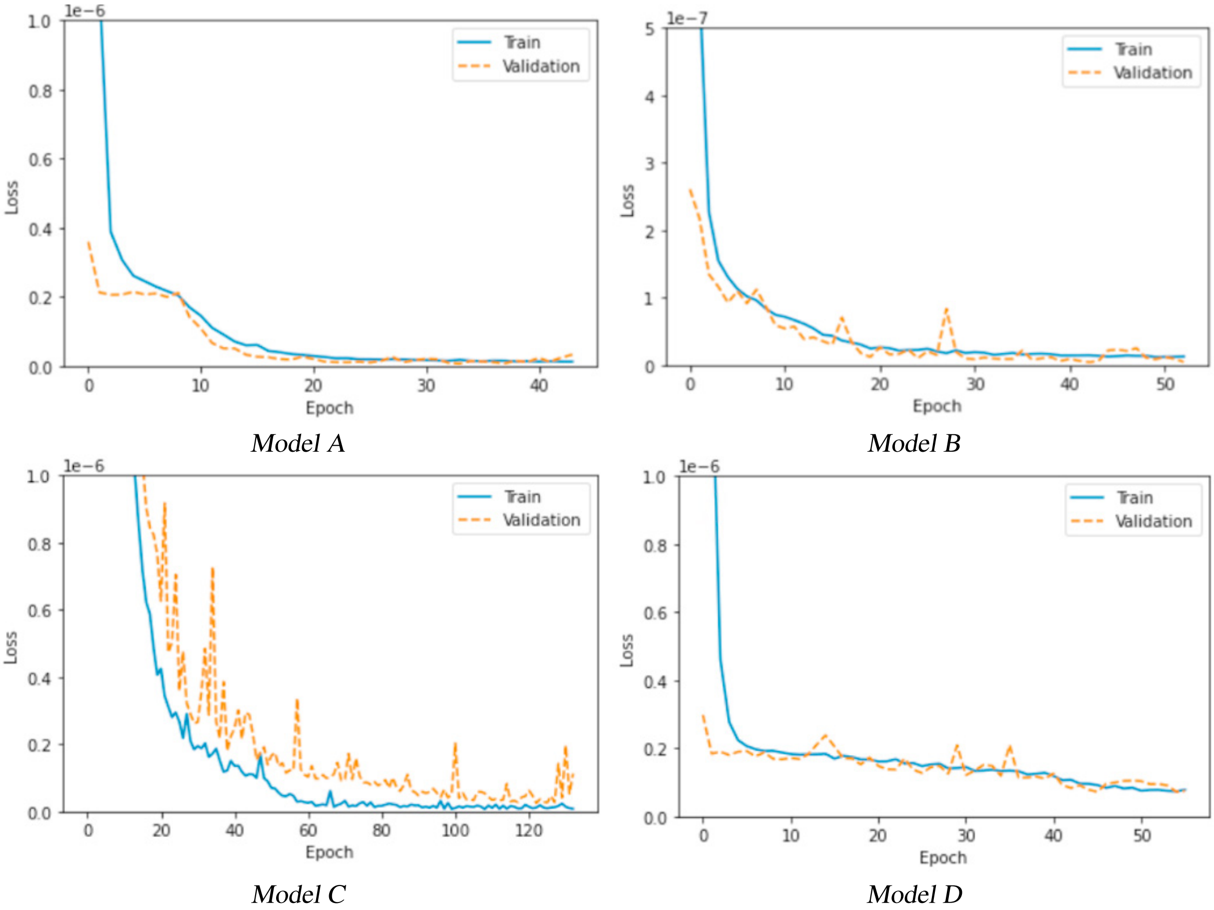


Fig. 4. Train and validation Loss curves

the test dataset only with damaged rolls are demonstrated in Table 4, where model A has a higher R^2 score and lower MAE compared to other models.

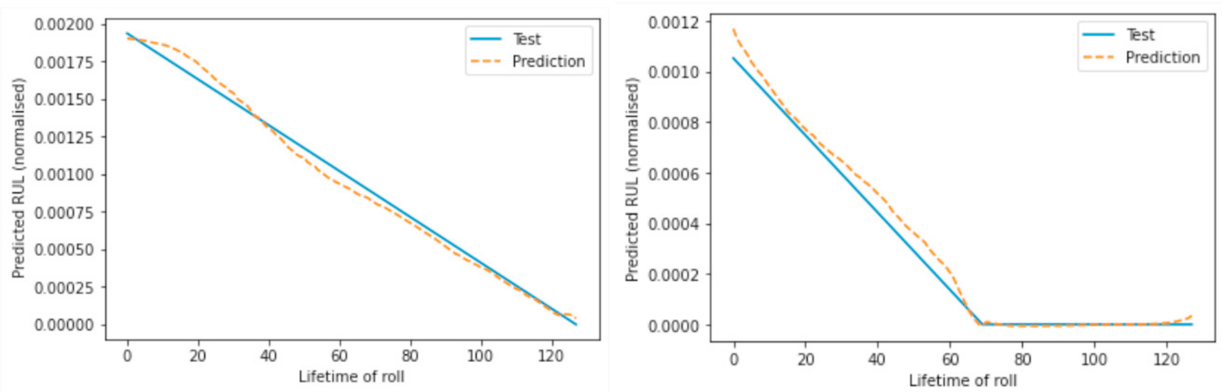
Figure 5 shows the test RUL vs predicted RUL for roll ran in normal condition (left), and roll ran in a damaged condition (right). After the 64th cycle, the data is padded with zero for training purposes in the damaged condition roll. The cold mill roll ran in normal condition has a better prediction compared to the damaged roll; however, the model has predicted the damaged roll efficiently.

5. Conclusions and remarks

This paper has presented a novel data-driven methodology to estimate remaining useful life of rolls in a cold rolling mill which aggregates heterogeneous data sources from various databases, including historical refurbishment

Table 4. Evaluation metrics for the test dataset containing damaged rolls tested against trained models.

Hybrid models	R^2 score	MSE	MAE	Computational time
CNN-BiLSTM-LSTM (Model A)	**0.9579**	**2.7274425e-09**	**2.0857e-05**	518.83 s
CNN-BiLSTM (Model B)	0.8646	3.8777575e-09	5.3465e-05	**476.88 s**
CNN-LSTM (1 layer LSTM)(Model C)	0.3093	4.4804505e-08	1.1122e-04	699.38 s
CNN-LSTM-LSTM (2 layers LSTM)(Model D)	0.4064	3.8510021e-08	8.9739e-05	769.19 s



RUL prediction of the roll in normal condition *RUL prediction of the damaged roll*

Fig. 5. RUL prediction of the roll in normal condition (left) and in damaged condition (right)

data, chemical and material properties of the coils and sensor collected data. The proposed approach introduces a novel formulation of the RUL framework by organising the data sets in sequences, each one corresponding to a specific roll unit trip. The RUL was calculated depending on how many coils the roll has processed in each trip. These sequences were used to train and compare several hybrid deep learning networks combining CNN and LSTM layers. It is shown that the best accuracy is achieved by the hybrid model CNN-BiLSTM-LSTM.

The roll that wears early in the process disrupts the industrial environment, so it is critical to predict the remaining useful life of rolls early and change the roll without causing disruption to the manufacturing process; the developed CNN-BiLSTM-LSTM model effectively predicted the remaining useful life at the early stage with reasonable accuracy. The model could be used as part for the roll management systems to plan and optimise the refurbishment schedule. Future work will explore the use of model agnostic interpretation methods to extract useful information on most impost factors that affect the life of the rolls and further integration of the model with optimisation algorithms to devise an optimal refurbishment schedule.

Acknowledgements

This work was supported by the UK Engineering and Physical Sciences Research Council (EPSRC) projects EP/V061798/1 and EP/S018107/1. Cinzia Giannetti would like to acknowledge the support of the IMPACT, Supercomputing Wales and Accelerate AI projects, which are part-funded by the European Regional Development Fund (ERDF) via Welsh Government. The authors would like to thank Tata Steel UK for data access and Steve Thornton, Scientific Fellow at Tata Steel UK for discussion and feedback on the manuscript.

References

[1] Albawi, S., Mohammed, T.A., Al-Zawi, S., 2017. Understanding of a convolutional neural network, in: 2017 international conference on engineering and technology (ICET), Ieee. pp. 1–6.
[2] Alhussein, M., Aurangzeb, K., Haider, S.I., 2020. Hybrid cnn-lstm model for short-term individual household load forecasting. IEEE Access 8, 180544–180557.

[3] Alves, P.G., Castro, J.A.d., Moreira, L.P., Hemerly, E.M., 2012. Modeling, simulation and identification for control of tandem cold metal rolling. Materials Research 15, 928–936.

[4] Chen, Z., Liu, Y., Valera-Medina, A., Robinson, F., 2019. Strip snap analytics in cold rolling process using machine learning, in: 2019 IEEE 15th International Conference on Automation Science and Engineering (CASE), IEEE. pp. 368–373.

[5] Cheng, Y., Hu, K., Wu, J., Zhu, H., Shao, X., 2021. Auto-encoder quasi-recurrent neural networks for remaining useful life prediction of engineering systems. IEEE/ASME Transactions on Mechatronics .

[6] Colas, R., Ramırez, J., Sandoval, I., Morales, J.C., Leduc, L.A., 1999. Damage in hot rolling work rolls. Wear 230, 56–60.

[7] Duan, Y., Yisheng, L., Wang, F.Y., 2016. Travel time prediction with lstm neural network, in: 2016 IEEE 19th international conference on intelligent transportation systems (ITSC), IEEE. pp. 1053–1058.

[8] Essien, A., Giannetti, C., 2019. A deep learning framework for univariate time series prediction using convolutional lstm stacked autoencoders. doi:10.1109/INISTA.2019.8778417.

[9] Essien, A., Giannetti, C., 2020. A deep learning model for smart manufacturing using convolutional lstm neural network autoencoders. IEEE Transactions on Industrial Informatics 16, 6069–6078.

[10] Giannetti, C., Essien, A., Pang, Y.O., 2019. A novel deep learning approach for event detection in smart manufacturing. CIE49 proceedings, Beijing .

[11] Krot, P., Prykhodko, I., Raznosilin, V., Zimroz, R., 2020. Model based monitoring of dynamic loads and remaining useful life prediction in rolling mills and heavy machinery, in: Advances in asset management and condition monitoring. Springer, pp. 399–416.

[12] Lakshmanan, K., Gil, A.J., Auricchio, F., Tessicini, F., 2020. A fault diagnosis methodology for an external gear pump with the use of machine learning classification algorithms: Support vector machine and multilayer perceptron, Loughborough University. doi:https://doi.org/10.17028/rd.lboro.12097668.v1.

[13] Lanzutti, A., Novak, J.S., De Bona, F., Bearzi, D., Magnan, M., Fedrizzi, L., 2020. Failure analysis of cemented carbide roller for cold rolling: Material characterisation, numerical analysis, and material modelling. Engineering Failure Analysis 116, 104755.

[14] Li, H., Jiang, Z., Tieu, K.A., Sun, W., 2007. Analysis of premature failure of work rolls in a cold strip plant. Wear 263, 1442–1446.

[15] Lundberg, S.E., 1993. Evaluation of deterioration mechanisms and roll life of different roll materials. Steel research 64, 597–603.

[16] Mao, W., He, J., Zuo, M.J., 2019. Predicting remaining useful life of rolling bearings based on deep feature representation and transfer learning. IEEE Transactions on Instrumentation and Measurement 69, 1594–1608.

[17] Pan, Z., Meng, Z., Chen, Z., Gao, W., Shi, Y., 2020. A two-stage method based on extreme learning machine for predicting the remaining useful life of rolling-element bearings. Mechanical Systems and Signal Processing 144, 106899.

[18] Pérez, D., Díaz, I., Cuadrado, A.A., Rendueles, J.L., García, D., 2018. Interactive data visualization of chatter conditions in a cold rolling mill. Computers in Industry 103, 86–96.

[19] Remadna, I., Terrissa, S.L., Zemouri, R., Ayad, S., Zerhouni, N., 2020. Leveraging the power of the combination of cnn and bi-directional lstm networks for aircraft engine rul estimation, in: 2020 Prognostics and Health Management Conference (PHM-Besançon), IEEE. pp. 116–121.

[20] Robinson, J.J., Van Steden, G., Ter Lingen, F., 1996. Effect of back-up roll wear on operation and strip shape of a cvc cold mill. Iron and Steel Engineer(USA) 73, 15–19.

[21] Torquato, M.F., Lakshmanan, K., Narożańska, N., Potter, R., Williams, A., Belblidia, F., Fahmy, A.A., Sienz, J., 2021. Cascade optimisation of battery electric vehicle powertrains. Procedia Computer Science 192, 592–601.

[22] Zhang, X., Lu, X., Li, W., Wang, S., 2021. Cnn-lstm enabled prediction of remaining useful life of cutting tool, in: Data Driven Smart Manufacturing Technologies and Applications. Springer, pp. 91–123.