Fatigue life Prediction based on Machine Learning: A Literature Review.

Bruna Balzer Maciel [a]

[a] Department of Mechanical Engineering, University of Federal Technological University of Paraná, Brazil, Ponta Grossa, brunabalzer@alunos.utfpr.edu.br

Abstract. When dealing with mechanical fatigue, making predictions about the life of materials subject to this type of failure presents itself as a persistent challenge. Throughout this continuous journey of studies and challenges, different methods have emerged to approach this complex issue. Among this process, the use of machine learning stands out as a promising way to predict life more accurately and efficiently. This article, based on a literature review, analyses and summarizes various current articles on fatigue analysis development using machine learning methods and presents relevant research findings. It is verified that machine learning models, in general, demonstrates remarkable capabilities in fatigue life prediction, outperforming traditional models across a wide compliable range of applications. In conclusion, is possible to affirm that the adaptability and precision of machine learning techniques in handling the complexities of fatigue analysis represent a significant advancement in predictions life, marking a notable shift in the field.

Keywords. Fatigue life prediction, machine learning, fatigue failure.

1. Introduction

The process of failure due to cyclic loading is called fatigue. The consequences of this phenomenon represent a major concern in engineering problems. It is expected that at least half of all mechanical failures are attributed to fatigue. As a result, the annual cost of materials fatigue to the United States economy is about 3% of gross domestic product (1).

The prediction of fatigue life, defined as the number of cycles until fatigue failure, is a fact of difficult precision in real-world scenarios, requiring the consideration of various variables to prevent unwanted and dangerous failures. Initially, analytical approaches to study fatigue issues relied mainly on the deduction of classical theory to understand this phenomenon [9]. Subsequently, experimental data were incorporated into abstract theoretical models and resorted to simulation on finite element-based simulation. In this context, several methods for predicting fatigue life have been proposed, such as the local stress-strain method, stress field intensity approach, critical plane method, energy-based method, and damage tolerance, etc. [21].

However, these conventional fatigue life models face challenges when aiming to achieve high precision and reliability in their predictions, especially when dealing with complex components or loads, to which sampling-related difficulties are added, as well as to being time-consuming and expensive processes [6]. In recent years, with the development of artificial intelligence and interdisciplinarity, machine learning (ML) is increasingly being used to uncover hidden relationships in data and make property predictions [5]. As ML algorithms possess greater capabilities for tackling the data relationships, it may be a promising tool for addressing the complex issues of fatigue life prediction [18].

Therefore, an increasing number of studies have applied ML methods to fatigue life prediction [18]. When developing models based on ML, the first crucial step is to properly define the input features to accurately map the target variable. Next, the process includes data collection and preparation, choosing the appropriate algorithm, training the model using a training set, evaluating performance with a test set, and ultimately implement the model [12]. Different algorithms have been adopted for different materials, e.g., neural networks (NN), decision trees (DT), regression analysis (RA), support vector machines (SVM), random forest (RF), boosting algorithms (BA), and Physics-Informed Neural Networks (PINN).

However, predictive capacity is determined by the quality and quantity of available data. For fatigue life prediction, when experimental data is sufficient, machine learning can be a powerful tool [8].

In this article, a review of fatigue prediction methods using machine learning was conducted. The analysis aimed to catalog the various ML techniques applied to different materials and evaluate the effectiveness of the predictions generated. The reading and analysis of these articles were performed by considering the following questions:

- Which ML technique is being utilized?
- What materials/manufacturing process are being applied?
- Is it a good prediction?
- What are the pros/cons of the method?

2. Methodology

In the course of conducting a comprehensive review for this article, we employed a narrative literature review methodology. The selection of the studies was made through searches in the Research Databases of Google Scholar and Science Direct. Advanced searches were conducted using terms such as "fatigue life prediction" + "fatigue failure" AND "machine learning." The focus was on studies published after 2018, a year in which there was a rising trend in the number of publications in this field.

A total of 852 documents were found from the databases. In the filtering process, initially, 800 articles were excluded after a preliminary analysis of titles and abstracts that did not align with the desired search terms. In the second step, 14 duplicates were removed. In the final step, only those articles that addressed the research questions and matched the specified publication date criteria were selected.

As a result, 19 articles remained for reading and analysis in this review study. The following sections are structured as follows: the results section offers a brief summary and exploration of each selected article, while the discussion section synthesizes the information acquired from these articles. Finally, the conclusion section provides a concise recapitulation of the entire study.

3. Results

In this paper [2], a unified SVR fatigue life model with 17 inputs, categorized into Mechanical properties (Elastic modulus, yield tensile stress, Ultimate tensile stress, elongation, surface hardness, etc.) and Loading parameters (maximum stress, stress rate, loading frequency), has been developed for the Ni-based superalloy family. This model addresses predictions for both Low Cycle Fatigue (LCF) which happens under with high stress amplitude and fewer loading cycles and High Cycle Fatigue (HCF) regimes with a high number of loading cycles with lower stress amplitude. It also investigates the impact of dataset size and highly correlated variables using the pairwise Pearson correlation. In conclusion, the SVR model achieve significantly higher predictive accuracy for fatigue life compared to classical models. Notably, the SVR model's prediction accuracy, when trained with only 526 data points, reached an acceptable level. However, the model exhibits increased prediction stability as the training dataset size increases. And, it's recognized that both the total strain range and test frequency are highly correlated variables with the investigated from a dataset perspective.

In this work [3], the utilization of a neuro-fuzzy-based machine learning method for predicting the HCF of a stainless steel 316L. Notably, this material was manufactured using an additive manufacturing process, specifically the Laser Powder Bed Fusion (LPBF) technique. Additive manufacturing, or 3D printing, in this process, a high-power laser beam is directed to melt thin layers of metal powder, one layer at a time, to build three-dimensional parts layer by layer. LPBF, a specific additive manufacturing technique, employs a high-power laser to selectively melt thin layers of metal powder, allowing for precise fabrication of intricate metal parts. A dataset comprising 139 experimental fatigue life data points was utilized to develop two models using different sets of input variables: one based on processing parameters and post-processing conditions and the other based on tensile properties. In summary, the results demonstrate that the neuro-fuzzy approach successfully predicted the fatigue life, indicating that both sets of parameters can be utilized for constructing goods predictive models.

In this study [4], the SVR model was utilized to predict fatigue life while considering the influence of the location, size, and morphology of defects in a LBBF Ti-6Al-4V selective alloy. It was observed that the coefficient of determination (R^2)- which provides a measure of how well a statistic model predicts an outcome, can reach up to 0.99.

In this paper [5], different machine learning models (RF, ERT, GBDT and XGBoost) were used to analyze and predict the fatigue life of high-strength bolts, and the relationship between fatigue life of the bolts and the influencing factors was analyzed by SHAP method, that is a theoretic approach to explain the output of any machine learning model. Geometric dimensions and stress states of bolts were as input features, the errors of bolt fatigue life prediction of

Reference	Material and manufacturing processes	Machine Learning algorithms	Input variables
[2]	Ni-based superalloys	SVR	Mechanical properties, test parameters
[3]	Stainless steel-LBPF	Neuro-fuzzy	Processing and post-processing parameters, and tensile properties
[4]	Ti-6Al-4 V-SLM	SVM	Location, size, morphology of the defect
[5]	High-strength bolts, IRON 40Cr	SVM, KNN, RF, ERT, GBDT, XGboost	Geometric dimensions, stress states, stress amplitude, test
[6]	Sucker rod, 30CrMo	BP, Elman e SVR	F rod diameter, defect diameter, defect.
[7]	FGH96 superalloy AM	BP, SVR, RF	Geometry data, environmental data and fatigue process data.
[8]	AZ61A magnesium alloy and titanium alloy TC4,	PINN	Stress-strain data and fatigue parameters
[9]	Liga alsi10mg-LPBF	SVM, RF	Experimental conditions, mechanical properties, porosity analyses, surface morphologies
[10]	Stainless steel 316L-AM	ANN, RF, SVM	Process parameters and fatigue loads
[11]	Lead-free tin-based solders	SVM, RBF, Boosting, CNN, ANN	Composition, loading and geometry factors.
[12]	TA2-TA15 and TC4-TC11 (LMD)	RF	Mechanical properties
[13]	Ti6Al4V (AM), with additives: TA2-TA15 and TC4-TC11	GABP-ANN	Type of structural parts, nominal and concentration stress, temperature
[14]	Ti-6.5Al-2Zr-Mo-V (LDE)	SVR	Variables stress intensity factor range and pore types
[15]	Hollow adhesive joints with sealant DB527	ANN	Path coefficient, fatigue loads
[16]	Ti-6Al-4V (SLM)	ANN-BP	Material defect properties, fatigue load, and build orientation
[17]	Small-scale butt-welded joints	GBT	Load level, local weld geometry, parent material strength
[18]	PA38-T6 and E355	IRT, RF	Strain amplitude, and loading paths
[19]	AlSi10Mg (AM)	PINN	Volume of the defects, the external surface of the defects, direction of the applied load.

Fig. 1 - Resume Table.

six machine learning models were compared. Among them, XGBoost has the best prediction level, with R^2 values of 0.883 in training set and 0.774 in test set. In addition, according to the analysis of SHAP value, the stress amplitude applied on the bolt has the greatest impact on the fatigue life. Nevertheless, there are still some limitations in this study, the high-strength bolt dataset is not broad enough.

In this section [6], three typical ML models (BP, Elman, and SVR) are selected for predicting the fatigue life of pumping rods- 30CrM alloy structural steel. A large sample dataset comprising rod diameter, defect diameter, defect depth, axial load, and fatigue life is used to create the training and test sets for the ML models. The results indicate that BPNN demonstrates strong generalization ability when compared to the prediction results of the others ML models, it can be observed that the maximum and minimum errors for BPNN are 18.65% and 2.66%, respectively. For Elman, the maximum and minimum errors are 111.33% and 6.78%, respectively, while for SVR, they are 43.93% and 2.97%, respectively.

In [7], (BP, SVR, RF) were employed to predict the fatigue life of FGH96 powder metallurgy superalloy based on a dataset encompassing geometry data (grain size, inclusion position and size, notch size, surface roughness, microstructure, and other factors), environmental data (ambient temperatures, and others conditions), and loading data. The results demonstrates that the trained ML models are effective and can be utilized for fatigue life predictions and in conclusion suggest incorporating some hidden physical fatigue information into the ML models.

This article proposes a physics-informed neural network (PINN) based on the multiaxial fatigue life prediction equation and introduces the life model into the loss function of the artificial neural network [8]. PINN can learn the knowledge of physical equations and make the generated data have definite physical meaning. The inputs are variables calculated from stress-strain data and fatigue parameters that guide the physical equations. The study focuses on two materials: magnesium alloy AZ61A and titanium alloy TC4. Three models, namely SWT, FS, and Shang-Wang, are compared for predicting the fatigue life of the experimental data for these two materials. The FS and Shang-Wang models demonstrate good prediction results, while the SWT model provides non-conservative results for torsional and out-of-phase loading paths. The results of feature importance analysis using SHAP analysis reveal that the most critical variables affecting fatigue life are the normal stress amplitude and the shear stress amplitude.

In this study [9], the LPBF-built alsi10mg samples are used for the training of SVM and RF using different strategies and physical information. The dataset is divided into four groups of input features: experimental conditions, mechanical properties, porosity analyses, and surface morphologies. After comparing the predicted results with experimental outcomes, SVM exhibits better prediction performance, and the framework using physical information achieves higher prediction accuracy and generalization performance.

A data-driven fatigue life prediction analysis is developed for AM stainless steel 316L based on continuous damage mechanics (CDM)- physical information (10). Various machine learning models, including (ANN, RF, SVM), are used to study fatigue lives considering AM process parameters (laser power, scanning speed, hatch space and powder layer thickness) and fatigue loads. Among them, RF exhibits the best performance, with results closely matching experimental data and predictions are sensitive to maximum stress and powder layer thickness.

In [11] five machine learning models were employed to predict the LCF of four distinct series of tin-based solders by considering the composition, loading and geometry factors. Among these models, Boosting demonstrated the highest performance, indicating its capacity to associate different features and favorable non-linear fitting for predicting. Additionally, this method proves to be both precise and cost-effective. The performance order is Boosting > ANN > CNN > SVM > RBF.

Fatigue models based on damage mechanics and the RF are conducted for smooth and notched samples of AM titanium alloy using the type of structural parts, nominal stress, temperature and stress concentration as parameters. The R^2 is 0.875 for the TC4-TC11, outperforming the TA2-TA15 additive with an R^2 of 0.682. Overall, the prediction performance of the CDM-RF method is superior to that of the CDM (12)

This article [13], introduces a new method for estimating LCF in titanium alloy (AM-Ti6Al4V) with additives AM-TA2-TA15 and AM-TC4-TC1. It employs the Continuous Damage Mechanics (CDM) theory, along with the Genetic Algorithm-optimized Backpropagation Artificial Neural Network (GABP-ANN) model. The input features were structural component type, nominal stress, temperature, and stress concentration factor. The results indicate that the BP-ANN model is initially established. A genetic algorithm is used to optimize the initial parameters of the model, and the established GABP-ANN model further improved the accuracy and stability of the fatigue life prediction reached $R^2 = 0.9801$.

This study [14], employed a SVR algorithm to develop a fatigue life prediction model for the Laser-directed energy deposition (LDED) Ti-6.5Al-2Zr-Mo-V titanium alloy based on post-mortem fractography analysis. The model presented that the variables stress intensity factor range and pore types achieved a significant increase of at least 18.9% in correlation compared to other models. Additionally, other ML algorithms such as Multilayer Perceptron, ANN, RF, Gaussian Process Regression (GPR) were compared for validation. The RF model exhibited larger prediction errors, while the ANN model tended to be non-conservative. SVR emerged as an effective and accurate approach for prediction.

This study investigates LCF in hollow adhesive joints with sealant DB527, emphasizing key factors affecting fatigue life [15]. The stress-based and energy-based fatigue life prediction models were developed and, in addition, one neural network-based method was adopted to predict the multiaxial fatigue life. The neural network method demonstrated superior predictive accuracy, supported by ample experimental data.

This study aims to overcome data dispersion and propose an easy-to-use, non-redundant ML model for very high cycle fatigue (VHCF) analysis of titanium alloys, particularly SLM Ti-6Al-4V [16]. Monte Carlo simulation of the Beta-PERT type generates ample data scarcity. Backpropagation Neural is applied to model the nonlinear relationship between variables. The dataset comprises five inputs, including material defect properties, loading parameters and build orientation. The model achieved an impressive coefficient of determination (R^2) and low Mean Squared Error (MSE) and effectively predicts VHCF behavior, saving researchers from labor-intensive experiments.

In [17], the fatigue behavior of welded joints depends on factors like loading parameters, local weld geometry, and material strength. Small-scale welded joint samples were analyzed using Gradient Boosted Trees, a black-box model. The SHAP method was used to explain predictions, revealing loading data and angular misalignment as highly influential. The model's robustness and generalization capability were confirmed through cross-validation, even with a relatively small database. As more data becomes available, model performance is expected to improve.

This article [18] introduces a method to predict multiaxial fatigue life for different loading paths using experimental data from aluminum alloys PA38-T6 and non-alloy steel E355. The proposed approach combines Image Recognition Technology (IRT) and RF. Input features include strain amplitude and loading path, converted into vector data by IRT. The method demonstrates high accuracy, with almost all predictions falling within a R^2 of approximately 0.96. This method outperforms empirical models in prediction accuracy and fitting precision.

This paper [19] presents a novel approach that combines ML with Physics-Informed Neural Networks (PINN) to predict the finite fatigue life of AM-metallic AlSi10Mg. The approach incorporates fracture mechanics constraints, overcoming data limitations. Experimental validation shows significantly improved predictions compared to pure ML tools, with an 83% increase in R^2 accuracy. The article also discusses the limitations of traditional fracture mechanics and highlights the emerging trend of Physics-Informed Machine Learning (PINN) to enhance predictions, particularly in data-scarce scenarios.

4. Discussion

The progress of the literature review reveals a successful variety of approaches across different materials, manufacturing processes, loading conditions, input variable, data quantities, and machine learning algorithms. This fact demonstrates the significant potential of the machine learning approach to enhance the accuracy of fatigue predictions compared to traditional models. In this section, the discussion focuses on a comprehensive discussion of the research results and findings.

One of the remarkable characteristics is the wide variety of materials approach, such as stainless steels, titanium, nickel-based superalloys. Several studies have specifically focused on predicting fatigue life of components produced through additive manufacturing processes and powder metallurgy. AM techniques possess the capacity to economically and efficiently produce complex parts. However, these processes are associated with a multitude of parameters, including laser power, scanning speed, hatch spacing, and powder layer thickness, all of which can significantly impact fatigue life predictions; nonetheless, ML has demonstrated to be adept at handling and drawing analogies among these various parameters, making it a highly viable solution.

Another noteworthy aspect is the variety of machine learning algorithms employed, such as Support Vector Regression (SVR), Random Forest (RF), Backpropagation Neural Networks (BPNN), and others. Each algorithm comes with its own set of advantages and limitations. Comparing these methods enables us to clarify their individual characteristics and pinpoint which prove most effective. Physics-based methods can incorporate prior physical knowledge, potentially enhancing prediction accuracy [8], [13], [19].

Another crucial aspect, which was explored in several articles, is the impact of dataset size and quality. In general, expanding the training dataset size tends to enhance prediction stability and accuracy. The data quality is equally essential, as eliminating outliers and errors is important to achieve precise and meaningful results, ensuring the reliability of the predictions. Some studies have proposed efficient strategies to supply that lack i.e., Monte Carlo simulation of the Beta-PERT [17], and using Physics-Informed like fracture mechanics constraints that have shown notable resilience in handling smaller or data-poor [19].

The appropriate selection of input variables is another critical step in fatigue prediction, and several articles have emphasized the importance of these factors in achieving accurate predictions. Some articles have proposed the application of methods such as SHAP [5], [8], [17], providing insights into the inputs and Pearson Correlation, which is a common way of measuring a linear correlation between two datasets [2]. These methods have been of great help in interpreting and understanding the factors that influence fatigue life predictions.

In addition to these aspects, several articles addressed the challenge of predicting fatigue life under multiaxial loading conditions [7], [8], which are common in real-world applications. And, many reported high accuracy in fatigue life predictions, even in lower, high and very-high fatigue regimes.

5. Conclusion

In this paper, a literature review was conducted to analyze the techniques of machine learning used in fatigue prediction. First, after a thorough analysis by following a series of steps and analyzing the quality of the studies was identified 19 primary studies (2018–2023). Second, the main findings of each proposed article were summarized. Third, a discussion section was presented, in which, after the insights from all the articles, a general overview of the approaches was write. The main conclusions are as follows:

- Machine learning methods demonstrates the versatility and reliability in predicting fatigue life, outperform traditional models across a wide range of applications.
- There is no universal algorithm that can be applied to all variables, highlighting the

necessity for tailored approaches, taking into account factors such as data complexity, dimensionality, and the nature of the fatigue phenomenon.

- Techniques such as SHAP offer a great solution for analyzing the correlation between input data and fatigue prediction.
- Both size and quality are extremely crucial factors.

Finally, it is essential to acknowledge that the approach of machine learning in fatigue prediction is in an increasingly growing field, and there is a trend toward the development accurate and precise of prediction methods.

6. Acknowledgement

I am grateful for the opportunity to participate in the scientific training of UNIGOU Program. It was certainly a job in which the learning journey was very useful and enriching.

7. References

1. Norton RL. Machine design: an integrated approach. Prentice Hall; 2011. 1028 p.
2. Tan L, Yang XG, Shi DQ, Hao WQ, Fan YS. Unified fatigue life modelling and uncertainty estimation of Ni-based superalloy family with a supervised machine learning approach. 2022 Nov.
3. Zhang M, Sun CN, Zhang X, Goh PC, Wei J, Hardacre D, et al. High cycle fatigue life prediction of laser additive manufactured stainless steel: A machine learning approach. Int J Fatigue. 2019 Nov.
4. Bao H, Wu S, Wu Z, Kang G, Peng X, Withers PJ. A machine-learning fatigue life prediction approach of additively manufactured metals. 2021 Feb.
5. Zhang S, Lei H, Zhou Z, Wang G, Qiu B. Fatigue life analysis of high-strength bolts based on machine learning method and SHapley Additive exPlanations [SHAP] approach. Structures. 2023 May 1;51:275–87.
6. Yuandeng W, Shaohu L, Weiguo M, Xiaofeng R, Baolong Q. Machine learning method for predicting the fatigue life of sucker rods. Eng Fract Mech. 2023 Apr.
7. Xie RK, Zhong XC, Qin SH, Zhang KS, Wang YR, Wei DS. Predicting multiaxial fatigue life of FGH96 superalloy based on machine learning models by considering failure process and loading paths. 2023 Oct.
8. He GY, Zhao YX, Yan CL. Multiaxial fatigue life prediction using physics-informed neural networks with sensitive features. Eng Fract Mech. 2023 Sep.
9. Wang H, Li B, Xuan FZ. Fatigue-life prediction of additively manufactured metals by continuous damage mechanics [CDM]-informed machine learning with sensitive features. Int J Fatigue. 2022 Nov.
10. Zhan Z, Li H. Machine learning based fatigue life prediction with effects of additive manufacturing process parameters for printed SS 316L. 2021 Jan.
11. Long X, Lu C, Su Y, Dai Y. Machine learning framework for predicting the low cycle fatigue life of lead-free solders. Eng Fail Anal. 2023 Jun.
12. Zhan Z, Hu W, Meng Q. Data-driven fatigue life prediction in additive manufactured titanium alloy: A damage mechanics-based machine learning framework. Eng Fract Mech. 2021 Jul.
13. Wang Y, Zhu Z, Sha A, Hao W. Low cycle fatigue life prediction of titanium alloy using genetic algorithm-optimized BP artificial neural network. 2023 Jul.
14. Dang L, He X, Tang D, Li Y, Wang T. A fatigue life prediction approach for laser-directed energy deposition titanium alloys by using support vector regression based on pore-induced failures. 2022 Jun.
15. Zhang J, Li H, Yu Li H. Evaluation of multiaxial fatigue life prediction approach for adhesively bonded hollow cylinder butt-joints. Int J Fatigue. 2022 Mar.
16. Li J, Yang Z, Qian G, Berto F. Machine learning based very-high-cycle fatigue life prediction of Ti-6Al-4V alloy fabricated by selective laser melting. Int J Fatigue. 2022 May 1;158:106764.
17. Braun M, Kellner L, Schreiber S, Ehlers S. Prediction of fatigue failure in small-scale butt-welded joints with explainable machine learning. 2022 Jan.
18. Zheng Z, Li X, Sun T, Huang Z, Xie C. Multiaxial fatigue life prediction of metals considering loading paths by image recognition and machine learning. 2023.
19. Salvati E, Tognan A, Laurenti L, Pelegatti M, De Bona F. A defect-based physics-informed machine learning framework for fatigue finite life prediction in additive manufacturing. Mater Des. 2022 Oct.
20. Jia Y, Fu R, Ling C, Shen Z, Zheng L, Zhong Z, et al. Fatigue life prediction based on a deep learning method for Ti-6Al-4V fabricated by laser powder bed fusion up to very-high-cycle fatigue regime. 2023.