Advances in Reliability, Safety and Security, Part 2 – Kołowrocki, Magryta-Mut (eds.)

Advances in Reliability, Safety and Security

ESREL 2024
Monograph Book Series

Remaining Useful Life Prediction From Degradation Zone Detected Of Microgripper For Intracorporeal Surgeries

Liseth Pasaguayo[a,b], Zeina Al Masry[a], Noureddine Zerhouni[a], Sergio Lescano[b]

[a]*SUPMICROTECH, CNRS, institut FEMTO-ST, 24 rue Alain Savary, 25000 Besançon, France,*
[b]*Amarob technologies, Besançon, France*

Abstract

The new generation of surgical microgrippers in the field of medical robotics aims to be less and less invasive for the patient. However, these medical devices must undergo several tests to evaluate and ensure their performance before their use in patients. This work is focused on predicting the remaining number of operating cycles of a microgripper for intracorporeal surgeries. Considering that the microgripper is under development, the study starts with data acquisition. Once the data are collected, their analysis takes place. Data analysis allowed the identification of statistical features that provides information on microgripper degradation. These features are used to divide the behavior of the microgripper into three zones: safety, degradation, and critical. Then, a multi-class classification Gradient Boosting model was implemented to predict the defined zones. This model exhibited challenges in accurately predicting the degradation and critical zone. Consequently, relying on zone predictions for estimating the RUL of the microgripper system was not an option. To address this challenge, a Long Short-Term Memory approach is considered to predict the exact number of remaining cycles for the microgripper when starting the degradation zone. The results were assessed through several metrics. In conclusion, RUL estimation provided us with a more precise means of determining when to cease the usage of the microgripper and initiate timely repairs for components nearing failure.

Keywords: surgical microgripper, remaining useful life (RUL), prognostics, PHM for medical devices, data-driven approach

1. Introduction

Prognostics and health management approaches (PHM) have emerged as one key to analyzing, evaluating, and monitoring the real-time health state of the systems. PHM uses a combination of many techniques to sense, record, and interpret some indicative parameters of the systems' health. For example, performance degradation, physical or electronic degradation, and change in the life cycle environment. (Pecht et al., 2017). PHM also incorporates disciplines based on failure physics, machine learning, modern statistics, and reliability engineering to predict the future state of the systems (Kim et al., 2017).

Although PHM approaches continue to expand in areas such as microelectronics, micromechanics, and micro mechatronics, to our knowledge, PHM implementation for medical microdevices has not been much studied in the literature. This may be due to PHM development for medical microdevice depends mainly on physics, data availability (Zio et al., 2022), and data collection. (1) Physics is related to the complexity and the number of components. When devices have many components, it is difficult to identify the critical components and therefore their failure mechanism. (2) The data availability is a challenging part since there are no existing databases that provide information on the behavior and performance of similar medical microdevices. (3) The data collection can pose challenges contingent upon factors such as the system's nature, complexity, the number of components, size, and the availability of measurement tools. Data collection plays an important role since once data is available this enables researchers to test hypotheses and address their research questions effectively.

A key pillar of PHM is prognostics, which enables the remaining useful life (RUL) estimation of the system by considering its current health state and its future operations. Within prognostics, PHM is classified into three types: data-driven, physics-based, and hybrid (Ferreira and Gonçalves, 2022).

This work is based on a data-driven approach, to predict the RUL on a microgripper to be used in intracorporeal surgeries. In this study RUL focuses on estimating how many operating cycles or how many opening and closing movements can perform the microgripper before failure occurs.

RUL of a system or one of its components is defined as the length of time left until the system reaches its end of useful life (Wang et al., 2019). For this work RUL is given by:

$$RUL(c) = C_{EOL} - c \tag{1}$$

where C_{EOL} is the total expected lifetime of the system in cycles, and c is the current number of cycles.

The RUL can have different measurement parameters depending on the specific context of the application. For example, in the case of batteries can be measured the remaining number of charges and discharges, in the case of motors the remaining rotation number, in mechanical, electronics, or electric systems the remaining cycle number, and in our microgripper system the remaining number of operating cycles before failure occurs. In addition, RUL can also be expressed in terms of hours, days, months, and years.

In this work, to implement the PHM, the first step is data collection of the microgripper. To collect data, the microgripper undergoes several experimental tests under five scenarios. Once data is available, data visualization and interpretation, data analysis, and feature processing are performed to implement a model for RUL predictions.

This paper is organized as follows. Section 2 explains the microgripper system and problem statement. Section 3 is focused on analyzing and interpreting the collected data for the microgripper. Section 4 introduces the approach developed to predict the RUL of the microgripper system. In this section, the results of the GB model and LSTM are explained. Section 5 presents some metrics to evaluate the proposed model and provide an interpretation of the results. Finally, Section 6 provides the conclusions of this work.

2. Microgripper system description

Amarob company offers a robotic system to perform intracorporeal laser surgeries. This system has an endoscope equipped with a micro-robot directing a laser beam during a surgical intervention. Moreover, Amarob aims to develop new complex medical devices for intracorporeal surgeries, and in the short term, to sell the robotic system to perform laryngoscopies at the back of the throat. Currently, to achieve its goal, Amarob is integrating a microgripper into its robotic system. The main tasks of the microgripper are to hold, pull and expose the vocal cord tissue to the laser beam provided by the microrobot during a surgical procedure. Figure 1. (a) illustrates how the robotic system composed of the microrobot and the microgripper would be implemented. The camera, light, and microrobot are introduced through an endoscope and the microgripper through a tube called catheter. The microrobot was already designed and tested under laboratory conditions by (Lescano, 2015). On the other hand, the microgripper is under development, therefore, this must be tested before being used inside the human body. The microgripper is composed of two jaws, a shaft of rotation, one support, and four 316L stainless steel wires of 0.08 mm in diameter as shown in Figure 1. (b). The 316L stainless steel wires are used to perform the opening and closing movements of the microgripper jaws by using four motors.

In a previous study, the 316L stainless steel wires and the shaft of rotation were identified as critical components by implementing a failure mode and effects analysis at the design level (D-FMEA) (Pasaguayo et al., 2022).

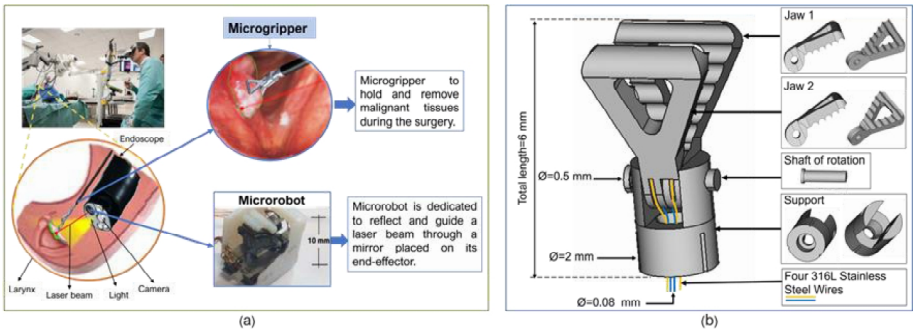

Fig. 1. Robotic microsystem description, (a) Microrobot and microgripper during a surgical procedure; (b) Microgripper components.

Once the critical components are selected, the microgripper undergoes several tests to collect data to assess its performance. Its performance can be evaluated by implementing PHM study, mainly to compute its RUL.

The next section explains how the data from the microgripper was collected.

3. Data for the microgripper

To collect data on the microgripper we responded to the questions: Why do we need to collect data?, What kind of data do we need?, and What do we hope to achieve from the collected data?. For the first question, the main reason to collect data is to understand how the microgripper system works and how it behaves over time. Additionally, the microgripper system is new and has never been tested. Therefore, there is no historical data that represents its behavior. There are also no large data sets from similar microgrippers. Although small sets of data were identified from the literature (Podolsky et al.; 2018; Haraguchi et al., 2015) they cannot be used in this study, since the data obtained from each researcher has a unique reason, and they do not align with our objective.

Regarding the second question, the data required for this work is primary experimental data since it is collected firsthand to address our research objective. Regarding the third question, what is sought in the data collected is to identify trends, patterns of degradation, or changes over time, which lead to an eventual failure of the microgripper. The data must help to understand how variables or factors relate to each other and how they evolve. In addition, data patterns and trends, identified in the collected data, must serve as the foundation for predictive modeling by using machine learning and/or deep learning algorithms. Consequently, data collection, data analysis, and data modeling will allow for the optimization and improvement of the microgripper.

3.1. Data collection

Data for the microgripper was collected by an experimental setup that allowed subjecting the microgripper to several tests under five scenarios. The scenarios focus on performing consecutive opening and closing movements of only one jaw, mainly varying the closing range. The scenarios to which one jaw of the microgripper was subjected are detailed below.

Scenario 1. Only one jaw of the microgripper will be subjected to cyclical opening and closing movements, with a maximum opening range of 2 mm. ***Scenario 2.*** Only one jaw of the microgripper will be subjected to cyclical opening and closing movements, changing the closing range by holding a small object of 0.39 mm. ***Scenario 3.*** Only one jaw of the microgripper will be subjected to cyclical opening and closing movements, changing the closing range by holding a small object of 0.65 mm. ***Scenario 4.*** Only one jaw of the microgripper will be subjected to cyclical opening and closing movements, changing the closing range by holding a small object of 0.91 mm. ***Scenario 5.*** Only one jaw of the microgripper will be subjected to cyclical opening and closing movements, holding small pieces of raw chicken. The five scenarios were considered since during a surgical procedure the microgripper can hold tissues or abnormalities of different sizes. Thus, the objective is to vary the closing range by testing objects of different sizes.

To test the opening and closing movements of only one jaw of the microgripper, two 316L stainless steel wires and two motors were used. One way to analyze the microgripper behavior is through the parameters that describe the movement of the jaw in the context of kinematics such as position, velocity acceleration, and jerk. Position data was collected directly by encoders placed in the motors. On the other hand, velocity, acceleration, and jerk data were generated by the successive derivation of position to calculate velocity, velocity to calculate acceleration, and acceleration to determine jerk over time. During each scenario, 25 tests were performed, resulting in 125 tests for position, 125 tests for velocity, 125 tests for acceleration, and 125 tests for jerk.

3.2. Data visualization and interpretation

The data shown in Figures 2. (a), (b), (c) and (d) are the position, velocity, acceleration, and jerk of the test 4 (T1004) belonging to scenario 1 after performing filtering data. Data filtering and data quality are not addressed in this article. However, it is worth mentioning that data quality was ensured by meticulous compliance with established data collection methods and protocols.

Figure 2. (a) shows how the position begins to increase as the number of cycles increases. The position data provide information on how to evolve the degradation of the microgripper according to the number of cycles executed. The position data in each scenario showed to have the same form, however, position values grew differently in each scenario. Figure 2. (b) represents the behavior of the velocity data. In these variables, it was difficult to precisely identify any common information, because each one changed in each scenario, possibly due to the scenario's conditions or external factors that cannot be controlled. Concerning acceleration and jerk, shown in Figures 2. (c) and (d) respectively, a similar effect was observed. In these variables, no common information could be seen either among the scenarios.

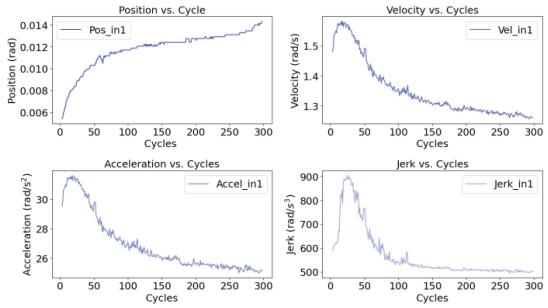

Fig. 2. Collected data. (a) Position vs cycles data; (b) Velocity vs cycles data; (c) Acceleration vs cycles data and (d) Jerk vs cycles data.

Henceforth, our focus will primarily center on presenting a more comprehensive analysis of the position data obtained vs the number of cycles. This emphasis is attributed to the presence of a discernible degradation pattern within the position profiles.

This analysis is performed for the 25 position profiles of scenarios 1, 2, 3, 4 and 5 illustrated in Figures 3. (a), (b), (c), (d), and (e).

To understand degradation behavior, two analyses are performed. The first one consists of an analysis of the number of cycles for each position profile. This analysis aims to identify how the number of cycles varies depending on the conditions of each scenario. Second, an analysis to understand the relationship between the position data vs the number of cycles is proposed. For both analyses programming scripts in the Python® language were developed.

For the analysis of the number of cycles, a box plot is utilized, where the results are displayed in Table 1.

According to the results of the median values of Table 1, there is a clear change in the number of cycles from scenarios 1 to 4. The median represents the midpoint of the number of cycles before failure, ranging from 269 in scenario 1 to 927.5 in scenario 4, indicating that the conditions of scenario 4 allow for a greater duration of the number of cycles before failure. The quartile 1 (Q1) and quartile 3 (Q3) results of each scenario represent that about 25% of data are lower than Q1 and about 75% are above Q3. The interquartile (IQR) results indicate the variability in the duration of the number of cycles. These results showed greater variability in the duration of the number of cycles under the conditions of scenario 4. In scenario 4, two outlier values appear. These outliers represent the two tests that have more than 1500 cycles, indicating that some tests had exceptionally high performance before failing. In each scenario, the closing range increases or decreases, depending on the size of the object that the jaws hold to. According to the analysis of the number of cycles, it is concluded that the smaller the object to be held, the greater the closing range. Therefore, the greater the closing range, the lower the number of operating cycles of the microgripper system. This demonstrates that the performance of the microgripper system is limited by the closing range of the jaw since for all tests the microgripper starts from the maximum opening range of 2 mm. Scenario 5 is not compared with the other scenarios because the raw chicken piece could not be measured.

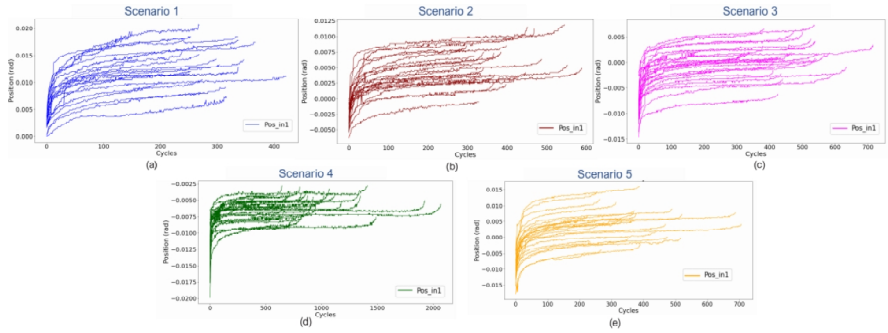

Fig. 2. Filtered position data sets, (a) filtered position data set under scenario 1; (b) filtered position data set under scenario 2; (c) filtered position data set under scenario 3, (d), filtered position data set under scenario 4 and (e) filtered position data set under scenario 5.

Table 1. Box plot statistics for the number of cycles for all the tests under each scenario.

Scenario	Median	Q1	Q3	IQR	Min	Max	Outliers
1	269	254	337	83	151	422	0
2	369	334	447	113	293	589	0
3	468	426	541	115	324	650	0
4	927	819	1183	364	645	2069	2
5	373	275	470	195	225	707	0

After finishing the analysis of the number of cycles, we proceed to the second analysis to interpret the relation between the variables of position and the number of cycles. This analysis aims to provide an understanding of the degradation of the microgripper system, as well as to identify relevant degradation patterns to implement models for RUL predictions. When performing an analysis of all the position tests vs. the number of cycles, it is observed that in most of the tests under the five scenarios, the position has a similar behavior. First, a pronounced growth in position is observed during the first operating cycles. Then a less pronounced and slower growth, which continues to grow as the cycles are executed. And finally, an abrupt growth occurs before failure during the last operating cycles. These behaviors allow the recognition of three zones during the evolution of microgripper degradation, as illustrated in Figure 4. (a) where a random test of scenario 1 is taken. The first zone is represented in green and belongs to the pronounced growth of the position. The second zone is represented in blue and indicates the slow growth of the position. The third zone is represented in red and belongs to the abrupt growth of the position before reaching failure. However, to prove whether what is observed with the naked eye occurs for all tests, a statistical analysis is carried out to verify this behavior in all tests. This analysis consists of analyzing how the data of each position profile is distributed. To do that, probability distributions such as the Normal, Lognormal, and Weibull distributions are considered (Lai et al., 2006), which are defined by their probability functions as outlined in equations (2), (3), and (4) respectively.

$$f(c) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{c-\mu}{\sigma}\right)^2}, \quad -\infty < c < \infty, \tag{2}$$

where μ is the mean value of the distribution, and σ is the standard deviation of the distribution.

$$f(c) = \frac{1}{c\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln c-\mu}{\sigma}\right)^2}, \quad c \geq 0, \tag{3}$$

where μ and σ are parameters such that $-\infty < \mu < \infty$, and $\sigma > 0$. Note that μ, and σ are not the mean and standard deviations of the distribution.

$$f(c) = \frac{\beta(c-\gamma)^{\beta-1}}{\theta^{\beta}} e^{-\frac{(c-\gamma)^2}{\theta}}, \quad c \geq \gamma \geq 0, \tag{4}$$

where β is a shape parameter, γ is a location parameter and θ is a scale parameter.

Then, to determine the best fit among the three distributions, evaluation metrics such as Kolmogorov-Smirnov Goodness-of-Fit Test (KS) (Lai et al., 2006), Akaike information criterion (AIC), Bayesian information criteria (BIC) (Mohammed et al., 2015), and root mean square error (RMSE) (Zraibi et al., 2021) are computed by using the equation (5), (6), (7) and (9) respectively.

$$D_n = \sup_{-\infty < c < \infty} |F_n(c) - F_0(c)| \tag{5}$$

where $F_n(c)$ is unknown distribution, and $F_0(c)$ is a given continuos distribution.

$$AIC = -2 * \ln(L) + 2 * k \tag{6}$$

$$BIC = -2 * \ln(L) + 2 * \ln(N) * k \tag{7}$$

where L is the values of the likelihood, N is the number of samples, and k is the number of estimated parameters.

Table 2 shows the results of the metrics mentioned above. According to these results, the normal distribution fits best for all the tests. This is supported by RMSE, AIC, and BIC metrics results, which have the lowest value compared to the other distributions. Although the Weibull distribution has the highest value in the Kolmogorov-Smirnov (KS) test, the normal distribution outperforms the others in the most important global metrics. Therefore, the normal distribution is considered for position data. All the calculations were performed in Python® language using the equations provided above.

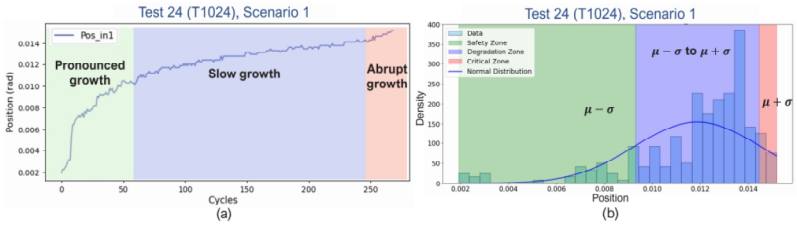

Fig. 3. Test 24 (T1024) under scenario 1 (a) Position data vs number of cycles interpretation; (b) Normal distribution of position data.

Table 2. Metrics to evaluate the distribution that fits best to data.

Distribution	RMSE	KS	AIC	BIC
Weibull	232.22	0.9981	5782.27	5790.94
Normal	160.57	0.9763	5311.76	5319.98
Lognormal	170	0.977	5357.85	5370.18

Figure 4. (b) provides a visual representation of the distribution, focusing on a single test from scenario 1. Here, most of the values lie in a normal distribution, where around 68% of values are within plus or minus one standard deviation (σ) from the mean (μ). Moreover, this figure illustrates that data are distributed with less frequency at the beginning of the first operating cycles, then the largest quantity of data is represented by the bell zone, and finally, the frequency of the data decreases rapidly until failure occurs. Therefore, according to these results, it is possible to identify three zones in position data. These zones are separated based on the μ and σ of a normal distribution. Then, the zones are defined as follows:

Safety zone: This zone contains values that are below (μ-σ).

Degradation zone: This zone contains values that are between (μ-σ to μ+σ).

Critical zone: This zone contains values that are above (μ+σ).

After data analysis, an approach for RUL predictions is proposed, which is explained in the next section.

4. RUL prediction of the microgripper system

The methodology proposed to predict the RUL of the microgripper system by considering a data-driven approach is illustrated in Figure 5. For RUL prediction a GB model and an LSTM model are considered. The GB model is used to predict the three identified zones that are: safety zone, degradation zone, and critical zone. The LSTM model is used to predict the RUL of the microgripper from the degradation zone that predicts the GB model. This approach is proposed since when predicting the zones in this work, it is not possible to know exactly when the maintenance action should be performed. On the contrary, the LSTM model allows knowing exactly how many operating cycles remain before the microgripper fails. The proposed approach was developed in Python® language.

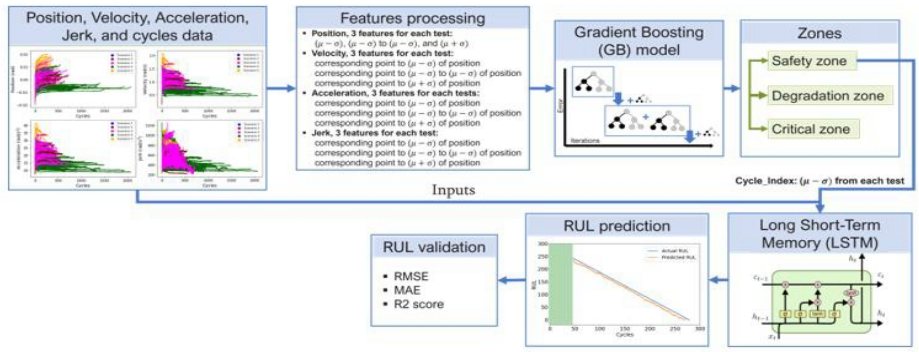

Fig. 4. Bloc diagram to predict the remaining useful life (RUL) of the microgripper for intracorporeal surgeries.

4.1. Gradient boosting model to predict the zones.

The GB is a machine-learning technique that builds predictive models in the form of an ensemble of weak learners, typically decision trees, in a sequential manner. Each new tree in the ensemble corrects the errors made by the previous ones, hence boosting the overall performance. (Asongo et al., 2021).

Before implementing the GB model, some features are processed from the data.

Features processing: From the position data analysis presented above, it is identified that the degradation behavior of the microgripper is divided into three zones, which are bounded by (μ−σ) for the safety zone, (μ−σ) to (μ+σ) for the degradation zone, and (μ+σ) for the critical zone.

As for the velocity, acceleration, and jerk data, they may not explicitly exhibit degradation behavior, they are important since they contain complementary information on the kinematics of the microgripper system. This is because these variables are derived from the rate of change of the position over time. Therefore, the features of the velocity profiles of each test were the corresponding point to (μ−σ) of position profile, the corresponding point to (μ − σ) until (μ + σ) of position profile, and the corresponding point to (μ + σ) of position profile. The same process was done for the acceleration and jerk profiles of each test. From this process, three features from each profile of position, velocity, acceleration, and jerk are obtained, resulting in twelve features for each test. These features are plotted in Figures 6. (a), (b), (c), and (d). The points that limit the safety zones of position, velocity, and acceleration profiles, are in green, the points that limit the degradation zones of position, velocity, acceleration, and jerk profiles, are in blue, and the points that limit the critical zones of position, velocity, and acceleration profiles, are in red. In the four Figures it is possible to easily observe the safety zone, which is separated from the degradation zone. However, the end of the degradation zone and the end of the critical zone cannot be easily identified. This is because the last point that limits the degradation zone is close to the last point of the critical zone. These three zones serve as crucial indicators for assessing the health state of the microgripper. The ability to identify the specific operational zone of the microgripper facilitates informed decision-making regarding maintenance interventions or the decision to cease its operation. Consequently, to predict these zones, a model for classification tasks will be implemented below.

Implementation: Once the features are extracted, a multi-class classification GB model is implemented to learn the points that separate the three zones previously mentioned. The inputs of the GB model are the 12 features that separate the safety, degradation, and critical zone of position, velocity, acceleration, and jerk profiles of each test, and the number of cycles of each test. The outputs of the model are the safety, degradation, and critical zones are illustrated in Figure 5. The GB model was configured with 100 estimators, a learning rate of 0.2, a maximum depth of 3, and a random state of 42. The training and testing sets are 80% and 20% respectively. The metrics used to evaluate the proposed model performance on the test dataset were accuracy, precision, sensitivity, specificity, F-score, and G-mean (Grandini et al., 2020), which are displayed in Table 3. These metrics could be assessed with |TP| the number of the true positive, |TN| the number of the true negative, |FP| the number of the false positive, and |FN| the number of the false negative (Remadna et al., 2022).

The results of the training vs the predicted points are illustrated in Figure 6. (e). Notably, in the safety zone, the training and predicted points align closely, indicating a high degree of accuracy in the model's predictions. However, when it comes to distinguishing between the degradation zone and the critical zone, observable discrepancies emerge. Therefore, the model exhibits challenges in accurately predicting points that fall within the

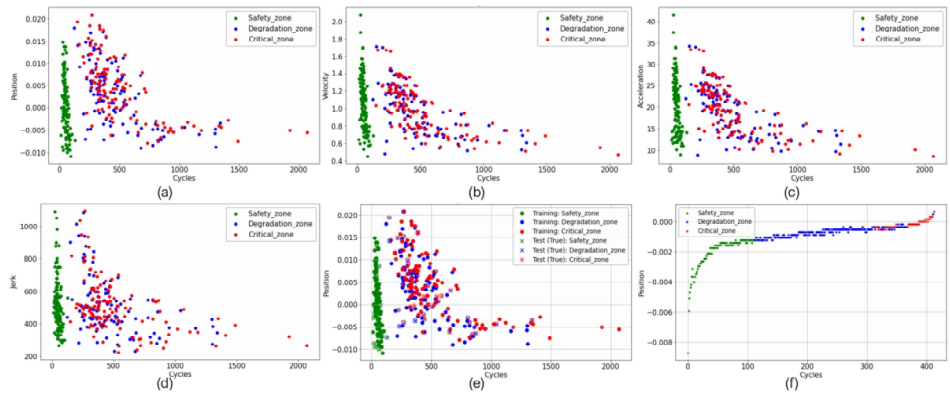

Fig. 6. Features of position, velocity, acceleration, and jerk, (a) Features of 125 position tests; (b) Features of 125 velocity tests; (c) Features of 125 acceleration tests; (d) Features of 125 jerk tests; (e) Results of the training and testing of the multi-class classification GB model; and (e) Results predictions of safety, degradation, and safety zone of the microgripper.

boundary separating these two zones. The degradation zone marks the onset of deterioration in the microgripper, signaling a critical phase that necessitates decision-making. However, determining the optimal timing for intervention, whether at the beginning or the end of the degradation zone, poses a challenge. As for the critical zone, it signifies the final cycles of operation for the microgripper. Due to the rapid nature of failures within this zone, it is impractical for maintenance actions.

This interpretation is supported by the metrics results shown in Table 3, the model is highly reliable when predicting the safety zone. However, regarding the degradation zone and critical zone, the model performance is less satisfactory, with low precision and recall values. The GB model was assessed alongside a Random Forest (RF) and Neural Network (NN) model. The evaluation revealed that the GB model outperformed the others in terms of both accuracy and precision, establishing it as the superior choice.

Consequently, considering that the microgripper is intended for use inside the human body, there arises a need to precisely estimate the remaining operating cycles. To address this challenge, an LSTM model is proposed to predict the exact number of cycles remaining for the microgripper when starting the degradation zone.

Table 3. Metrics to evaluate the classification model's performance.

Metric	Safety zone	Degradation zone	Critical zone
$Accuracy = \frac{\|TP\| + \|TN\|}{\|TP\| + \|FP\| + \|FN\| + \|TN\|}$	0.61		
$G - Mean = \sqrt{Sensitivity \times Specificity}$	0.98	0.41	0.25
$Precision = \frac{\|TP\|}{\|TP\| + \|FP\|}$	1.00	0.43	0.23
$Recall = \frac{\|TP\|}{\|TP\| + \|FN\|}$	0.96	0.40	0.27
$F1 - score = \frac{Precision \times Sensitivity}{Precision + Sensitivity}$	0.98	0.41	0.25
$Specificity = \frac{\|TN\|}{\|TN\| + \|FP\|}$	1.00	0.43	0.23

4.2. LSTM model to predict RUL

LSTM model is a type of recurrent neural network (RNN) architecture belonging to the deep learning models. This model is one of the most used for times series data since it captures dependencies within sequential data (Abdelli et al., 2021). To implement the GB and LSTM approach, the inputs of the LSTM model are the 125 profiles of each variable, which are position, velocity, acceleration, and jerk, as well as the points that bound the safety zone predicted by the GB (cycle_index), as illustrated in Figure 5. The target of the model is the actual RUL computed for each test by using equation (1). Therefore, LSTM has a total number of six features. After some

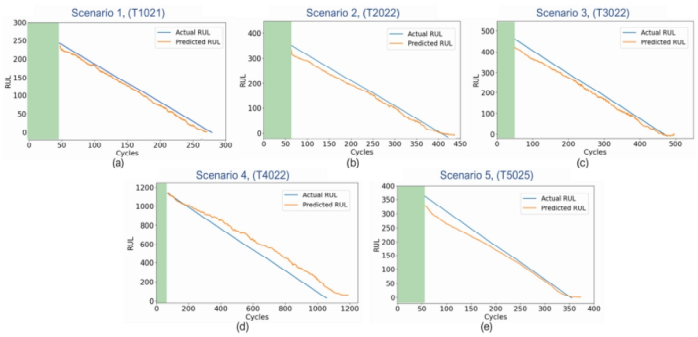

Fig. 7. RUL predictions, (a) RUL prediction of test 21 (T1021) under scenario 1; (b) RUL prediction of test 22 (T2022) under scenario 2; (c) RUL prediction of test 22 (T3022) under scenario 3; (d) RUL prediction of test 22 (T4022) under scenario 4; and (e) RUL prediction of test 25 (T5025) under scenario 5.

attempts, the LSTM model is configured with the following hyperparameters. A window size of 10, batch size of 1, a learning rate of 1, epochs set to 200, an LSTM layer of 128 units, a dense layer with one unit, a Rectified Linear Unit (ReLu) as activation function used in the output layer, a loss function (mean squared error) used for optimization during training with the Adam optimizer algorithm. For data training and validation, 100 profiles of each variable are considered divided into 80% and 20% respectively. For the testing of the model, the remaining 25 profiles of each of the variables are taken, these are the unseen data.

Figures 7. (a), (b), (c), (d) and (e) illustrate the actual RUL and predicted RUL for one test conducted under each scenario. The actual RUL is represented by the blue line, and the predicted RUL by the orange line. To estimate the RUL, an unseen test is entered for each scenario. Once the GB model detects the degradation zone, the LSTM model starts to predict the RUL. Figure 7. (a) illustrates that the actual RUL when finishing the safety zone and starting the degradation zone for test 21 is 250 cycles. As the number of cycles increases, the RUL decreases. In the same way occurs for the scenarios 2, 3, 4, and 5. After finishing the safety zone, the actual RUL for each case is predicted. Then, as the number of cycles increases, the RUL decreases. In this study, an RUL analysis for each scenario is presented since allows us to cover the greatest information about the opening and closing behavior of the microgripper jaws during a surgical procedure. The performance of the proposed approach was evaluated considering some metrics. The metrics used and the RUL evaluation are presented in the next section.

5. RUL evaluation

In this work, the mean absolute error (MAE), root mean squared error (RMSE), and R2 score (Zraibi et al., 2021; Abdelli et al., 2021) are used as indicators to evaluate the performance of proposed method. MAE, RMSE and R2 score metrics are given by equations (8), (9) and (10) respectively.

$$MAE = \frac{1}{n}\sum_{i=1}^{m}|y(i) - \hat{y}(i)| \tag{8}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{m}(y(i) - \hat{y}(i))^2} \tag{9}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{m}\left(y(i)-\hat{y}(i)\right)^2}{\sum_{i=1}^{m}\left(y(i)-\bar{y}(i)\right)^2} \tag{10}$$

Where n denotes the total number of test samples, $y(i)$ is the actual value of RUL, $\hat{y}(i)$ is the predicted value of RUL, and $\bar{y}(i)$ is the average of the true values.

The RUL for the microgripper was predicted by considering one test by scenario. Therefore, the RMSE, MAE, and R2 score metrics for each scenario are computed. The results of these metrics are shown in Table 4.

For scenario 1, the RMSE is 10.93, which means, on average, the predictions under this scenario deviate by approximately 10.93 cycles from the actual values. The MAE of 8.26 suggests that, on average, the absolute error between predicted RUL and actual RUL values is 8.26 cycles. The R2 score is 0.992, indicating a significant portion of the variance in the actual RUL. Some similarities occur in scenarios 2, 3, 4, and 5. The RMSE varies between 9.53 to 17.58, the MAE varies between 7.20 and 13.29, and the R2 score varies between 0.97 to 0.987. The R2 score values in the five scenarios are very high, ranging from 0.97 to 0.992. This suggests that the model, in each scenario, explains a substantial portion of the variance in the RUL values. The scenario 5 shows the bests of accuracy in terms of error with lower values of MAE and RMSE compared to the results of the scenarios 1, 2, 3, and 4. According to R2 results, these values suggest that the proposed approach fits well with the data.

Table 4. LSTM prediction results.

Scenario	1	2	3	4	5
RMSE	10.93	12.21	11.24	17.58	9.53
MAE	8.26	9.23	8.49	13.29	7.20
R2_score	0.992	0.987	0.982	0.97	0.971

Unlike the GB approach which only allowed us to precisely know the safety zone established for the microgripper system, the LSTM approach offers the capability to anticipate the RUL of the system. This approach indicates how many more operating cycles the microgripper system can undergo before potential failure. This approach enables both corrective and potentially preventive maintenance strategies, depending on the insights gained from the analysis.

Conclusion

This work presented the first RUL results of a new microgripper to be used in intracorporeal surgeries. Data collection was the first step for RUL implementation. The amount of performed tests provided a robust dataset, which allowed us to successfully calculate the RUL by using machine learning and deep learning techniques. The data used was sufficient to support the analysis and obtain accurate results.

The RUL predictions allowed us to consider the implementation of proactive and preventive maintenance techniques, to ensure the safety of the microgripper. In this case study, maintenance techniques could be focused on replacing the 316L stainless steel wire, which was the component that failed in all the tests.

This first study confirmed that the 316L stainless steel wire, identified as critical in a previous work (Pasaguayo et al., 2022), is indeed the critical component. Therefore, it is concluded that the failure mode is wire breakage, the failure effect is the unavailability of the microgripper, and the failure cause is the cyclical movements of the opening and closing movements to what is subjected. This finding is valuable as it allows for potential modifications to the microgripper design before its final implementation. One option could involve exploring alternative actuation principles to eliminate the need for wires and incorporating additional mechanisms to enhance the safety, reliability, and RUL of the microgripper system.

The historical data collected in this work can be used to implement a quantitative failure mode and effects analysis (FMEA) to enhance the safety and reliability of the microgripper by prioritizing focus areas based on the severity, occurrence, and detection of possible failures in the microgripper.

Although this approach allowed us to know exactly the number of cycles remaining before failure, it is recommended to test the microgripper considering other scenarios. This work opens a window for future studies to evaluate the microgripper in real-time by implementing adaptive approaches to provide dynamic and updated estimates of the RUL as the system evolves under a new scenario.

Acknowledgements

This work has been supported by the ANRT-AMAROB CIFRE Thesis and the EIPHI Graduate school (contract "ANR-17-EURE-0002").

References

Abdelli, K., Grießer, H. and Pachnicke, S., 2021, July. A hybrid CNN-LSTM approach for laser remaining useful life prediction. In Optoelectronics and Communications Conference (pp. S3D-3). Optica Publishing Group.

Asongo, A.I., Barma, M., and Muazu, H.g. 2021. Machine Learning Techniques, methods and Algorithms: Conceptual and Practical Insights. International Journal of Engineering Research and Applications Vol. 11, Issue 8, pp. 55-64.

Ferreira, C. and Gonçalves, G., 2022. Remaining Useful Life prediction and challenges: A literature review on the use of Machine Learning Methods. Journal of Manufacturing Systems, 63, pp.550-562.

Grandini, M., Bagli, E. and Visani, G., 2020. Metrics for multi-class classification: an overview. arXiv preprint arXiv:2008.05756.

Haraguchi D., Kanno, T., Tadano, K. and Kawashima, K., 2015. A pneumatically driven surgical manipulator with a flexible distal joint capable of force sensing. IEEE/ASME Transactions on Mechatronics, 20(6), pp.2950-2961.

Kim, N.H., An, D. and Choi, J.H., 2017. Prognostics and health management of engineering systems. Switzerland: Springer International Publishing.

Lai, D., Murthy, D., and Xie M. 2006. Basic Statistical Concepts. Springer Handbook of Engineering Statistics, vol. Chapter 1, page 03–26.

Lescano, S., 2015. Design, Fabrication and Control of a Microrobot for Laser Phonomicrosurgery. Université de Franche-Comté.

Mohammed, E. A., Naugler, C., & Far, B. H. 2015. Emerging Business Intelligence Framework for a Clinical Laboratory Through Big Data Analytics. Emerging Trends in Computational Biology, Bioinformatics, and Systems Biology, pp. 577–602.

Pasaguayo, L., Al Masry, Z., Lescano, S., & Zerhouni, N. 2022. Critical Components' Selection Methodology of a Microgripper for Intracorporeal Surgery. Annual Conference of the PHM Society, 14(1)

Pecht, M., Tuchband, B., Vichare, N., & Ying. Q. J., 2007. Prognostics and Health Monitoring of Electronics. International. Conference Thermal, Mechanical and Multi-Physics Simulation Experiments in Micro-electronics and Micro-Systems. EuroSime, 1–8.

Podolsky, D.J., Diller, E.D., Fisher, D.M., Riff, K.W., Looi, T., Drake, J.M., & Forrest, C.R. 2018. Utilization of Cable Guide Channels for Compact Articulation Within a Dexterous Three Degrees-of-Freedom Surgical Wrist Design. Journal of Medical Devices.

Remadna, I., Terrissa, L.S., Al Masry, Z. and Zerhouni, N., 2022. RUL Prediction Using a Fusion of Attention-Based Convolutional Variational AutoEncoder and Ensemble Learning Classifier. IEEE Transactions on Reliability, 72(1), pp.106-124.

Wang, Q., Zheng, S., Farahat, A., Serita, S. and Gupta, C., 2019, June. Remaining useful life estimation using functional data analysis. In 2019 ieee international conference on prognostics and health management (icphm) (pp. 1-8).

Zio, E., 2022. Prognostics and Health Management (PHM): Where are we and where do we (need to) go in theory and practice. Reliability Engineering & System Safety, vol 218, 108119.

Zraibi, B., Okar, C., Chaoui, H. and Mansouri, M., 2021. Remaining useful life assessment for lithium-ion batteries using CNN-LSTM DNN hybrid method. IEEE Transactions on Vehicular Technology, 70(5), pp.4252-4261.