Conference materials
UDC 004.852
DOI: https://doi.org/10.18721/JPM.163.259

Development of an algorithm for predicting the strength of laser reconstruction of biological tissue

D.I. Ryabkin[1,2], V.V. Suchkova[1,2]✉, A.Yu. Gerasimenko[1,2]

[1] I.M. Sechenov First Moscow State Medical University, Moscow, Russia;

[2] National Research University of Electronic Technology MIET, Zelenograd, Moscow, Russia

✉ ryabkin@bms.zone

Abstract. Laser reconstruction of biological tissue is a fast and minimally invasive method of wound closure without the risk of stenosis, foreign body reaction or inflammation. However, the strength of the welds is inferior to the traditional suture method in the first few days after surgery. This limitation can be overcome by optimizing the laser beam and solder component composition using machine learning methods. The aim of the study was to develop and experimentally verify an algorithm for predicting the strength of laser reconstruction of biological tissues. The best prediction models were based on an extreme gradient boosting algorithm and random forest. To train the algorithms, a dataset consisting of experiments described in published research papers was used. The dataset contains a total of 394 samples and 39 features on which training was performed. The effectiveness of the model was tested experimentally in two stages. Bovine ex vivo vascular repair was the first step. The second stage was in vivo testing of the algorithm on laboratory animals. The average percentage error of the strength prediction was 19%. This error is due to the large scatter in the strength values obtained experimentally. The strength obtained is sufficient to analyses the laser radiation characteristics and the component composition of the solder prior to laser reconstruction.

Keywords: machine learning, laser soldering, tissue reconstruction

Funding: The study was supported by the Russian Science Foundation grant 22-75-00089, https://rscf.ru/project/22-75-00089/.

Citation: Ryabkin D.I., Suchkova V.V., Gerasimenko A.Yu., Development of an algorithm for predicting the strength of laser reconstruction of biological tissue, St. Petersburg State Polytechnical University Journal. Physics and Mathematics. 16 (3.2) (2023) 338–343. DOI: https://doi.org/10.18721/JPM.163.259

Материалы конференции
УДК 004.852
DOI: https://doi.org/10.18721/JPM.163.259

Разработка алгоритма прогнозирования прочности лазерного восстановления биологических тканей

Д.И. Рябкин[1,2], В.В. Сучкова[1,2]✉, А.Ю. Герасименко[1,2]

[1] Первый Московский государственный медицинский университет имени И.М. Сеченова, Москва, Россия

[2] Национальный исследовательский университет «Московский институт электронной техники», Зеленоград, Москва, Россия

✉ molodykh1999@gmail.com

Аннотация. Цель исследования заключалась в разработке и экспериментальной проверке алгоритма прогнозирования прочности лазерного восстановления биологических тканей. Лучшие модели прогнозирования были основаны на алгоритме экстремального градиентного бустинга и случайном лесе. Для обучения алгоритмов использовался набор данных, состоящий из экспериментальных данных, описанных в опубликованных научных работах по лазерному восстановлению биологических тканей. Средняя процентная ошибка прогноза прочности составила %19. Эта ошибка обусловлена большим разбросом значений прочности, полученных экспериментально. Достигнутая точность значений прочности является достаточной для анализа характеристик лазерного излучения и компонентного состава припоя до операции по лазерному восстановлению биотканей.

Ключевые слова: машинное обучение, лазерная пайка, реконструкция тканей

Финансирование: Работа выполнена в рамках гранта Российского научного фонда № 00089-75-22, https://rscf.ru/project/00089-75-22/.

Ссылка при цитировании: Рябкин Д.И., Сучкова В.В., Герасименко А.Ю. Разработка алгоритма прогнозирования прочности лазерного восстановления биологических тканей // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2023. Т. 16. № 3.2. С. 338–343. DOI: https://doi.org/10.18721/JPM.163.259

Introduction

Laser reconstruction of biological tissues, is a seamless method of wound closure that uses the application of an absorbent bioorganic solder to tissue and the subsequent absorption of laser radiation by the tissue and produces a thermal effect by interaction to achieve the goal of tissue fusion [1–3]. However, despite the speed and painlessness of the procedure, laser reconstruction has not become widespread clinically. As the structure of biological tissue is complex, its thermal and optical characteristics are different, and different laser radiation parameters and solder component compositions will cause different thermal effects, which may result in insufficient tensile strength, especially in the first days after surgery [2].

At the moment, the selection of laser reconstruction parameters is realized by performing numerous tests on laboratory animals, which is a low-efficient and resource-intensive strategy [4, 5]. This paper investigated the possibility of predicting the results of laser reconstruction of biological tissues using machine learning methods.

Machine learning is one of the most important and effective tools for analyzing complex medical data. Machine learning techniques have been shown to be effective in predicting the progression of diseases such as dementia [6], lung cancer, liver disease, etc. [7,8].

The use of machine learning methods for predicting laser reconstruction results will accelerate and automate the process of selecting laser irradiation characteristics and solder component compositions. In this paper, machine learning regression models have been trained, optimized and experimentally tested.

Materials and Methods

The prediction algorithm for all regression models is based on learning from the data set. The accuracy of the prediction depends on the correctness, completeness and homogeneity of the input data. A data set of 400 unique objects based on 45 published scientific sources was generated to train an algorithm for predicting the strength of laser repair of biological tissues. The dataset was generated using the python programming language. Using a library for processing and analysis of structured data pandas. All categorical were converted linear using the one hot encoding method. In one hot coding, each categorical feature value is represented as a separate binary feature column. In this case, the categorical value corresponding to the object received a value of 1, all others 0 [9]. Thus, after conversion, the dataset has 54 features on which the prediction was based. The data provided information on laser restoration of biological tissues including tissue type and type of operation, laser irradiation parameters, component composition of laser solders, and methodology for measuring the strength of welds. The target variable in the data set was the tensile strength of the laser welds expressed in kPa. The minimum strength value in the data set is 2.07 kPa. The maximum strength in the data set is 6000 kPa. The average strength value in the data set is 657 kPa. This scatter of values indicates a heterogeneity in the data. To minimize the role of outliers in prediction, values between the 5th and 95th percentile of the data distribution were used to train the models.

Two strategies were used to fill in the missing values in the data set: filling in the mean value for a feature and filling in the median value for a feature [10]. Normalization (1) and standardization (2) functions were used to bring the values into the same range [11]. For each machine learning regression model, the most efficient combination of data processing was selected.

The normalization of the data is presented in Equation 1:

$$\overline{x} = \frac{x - x_{\min}}{x_{\max} - x_{\min}}, \tag{1}$$

where x is the initial value of the feature, x_{min} is the minimum value of the feature, x_{max} is the maximum value of the feature, $\overline{x}$ is the value of the feature after normalization.

The standardization of the data is presented in Eq. 2:

$$\overline{x} = \frac{x - x_{mean}}{\sigma}, \tag{2}$$

where x_{mean} is the arithmetic average value of the feature, σ is the standard deviation of the feature values.

Seven different models were trained to predict the strength of laser reconstruction of biological tissues: linear regression, k-nearest neighbors, decision tree, random forest, gradient boosting, and variations of gradient boosting (XGBoost, LightGBM). All models based on the sklearn library of the python programming language. A preprocessing strategy was chosen for each model.

The coefficient of determination and the mean absolute percentage error were used to assess performance. The mean absolute percentage error (MAPE) has a dimensionless value and allows the relative error of the prediction to be determined:

$$MAPE = \frac{100\%}{N} \sum_{i=1}^{N} \frac{\left| y_{i(pred)} - y_i \right|}{y_i}, \tag{3}$$

where $y_{i(pred)}$ is the predicted strength value, y_i is the actual strength value, N is the number of data objects.

The coefficient of determination (R^2) is defined as the ratio of the variance of the target variable explained by the predictive model with a given set of inputs:

$$R^2 = 1 - \frac{\frac{1}{N}\sum_{i=1}^{N}\left(y_i - y_{i(pred)}\right)^2}{\frac{1}{N}\sum_{i=1}^{N}\left(y_i - y_{mean}\right)^2}, \tag{4}$$

where y_{mean} is the average strength value. A 5-block cross-validation was used to assess the stability of the performance. The 5-block cross-validation was also used to determine the best combination of preprocessing data. The models that performed the best were also optimized manually by parameter enumeration, using grid-based parameter search, and random search. The training hyper-parameters chosen were: tree depth, the maximum number of features that the model is allowed to sample at each partitioning, the number of decision trees.

To experimentally validate the strength algorithm, a series of *ex vivo* and *in vivo* experiments were conducted on laser reconstruction of biological tissues (skin) using the most commonly used bioorganic solders and a laser complex with the ability to maintain a set temperature in the weld formation zone.

The laser system for biological tissue reconstruction consists of an optical module, a temperature module, and a surgical pencil-tip. To recover the integrity of biological tissues, a diode laser generating continuous infrared radiation with a wavelength of $\lambda = 810 \pm 3$ nm was used.

In order to experimentally test the algorithm for predicting the strength of laser solder, a bioorganic solder consisting of bovine serum albumin (BSA), green indocyanine (ICG) and water was fabricated. The solder was an aqueous solution of 25 wt% BSA and 0.1 wt% ICG.

Ex vivo experiments were performed using bovine aortic tissue and pig skin. 50 µl of solder was applied to the area of the joint. The laser irradiation was performed using the spot method. Each spot was irradiated for 1 minute. A total of approximately 5 dots were placed per weld. Maximum temperature in the laser exposure area was $T = 55$ °C. The laser power was varied from 1.1 to 2.5 W. The diameter of the laser spot was 2 mm.

Chinchilla rabbits were used for the *in vivo* experiment. The skin was dissected with a surgical scalpel, forming a 1 cm linear incision. The reconstructed area was irradiated with laser at 42 °C for 1 minute. Tensile strength was recorded on days 1, 3, 7 and 10 postoperatively.

Results and Discussion

Fig. 1,*a* and 1,*b* shows the cross-validation results of all trained models. Cross-validation results revealed that the worst model for predicting laser reconstruction results is linear regression (average values: MAPE = 75%, $R^2 = 0.7$) this is due to non-linear relationships between laser reconstruction features and strength, as well as non-uniformity of feature distribution. Models based on ensembles of decision trees showed the best results. These models were also optimized manually by parameter oversampling, grid search, and random search for the best result (Fig. 1,*c*). The highest accuracy in strength prediction is achieved by the extreme gradient boosting model XGBoost, optimized using random search with ridge regularization of the decision trees.

A preprocessing strategy was chosen for each model. The best preprocessing strategy in all cases was to fill in the missing values with the median value for the relevant feature and then standardize all values (Table 1).

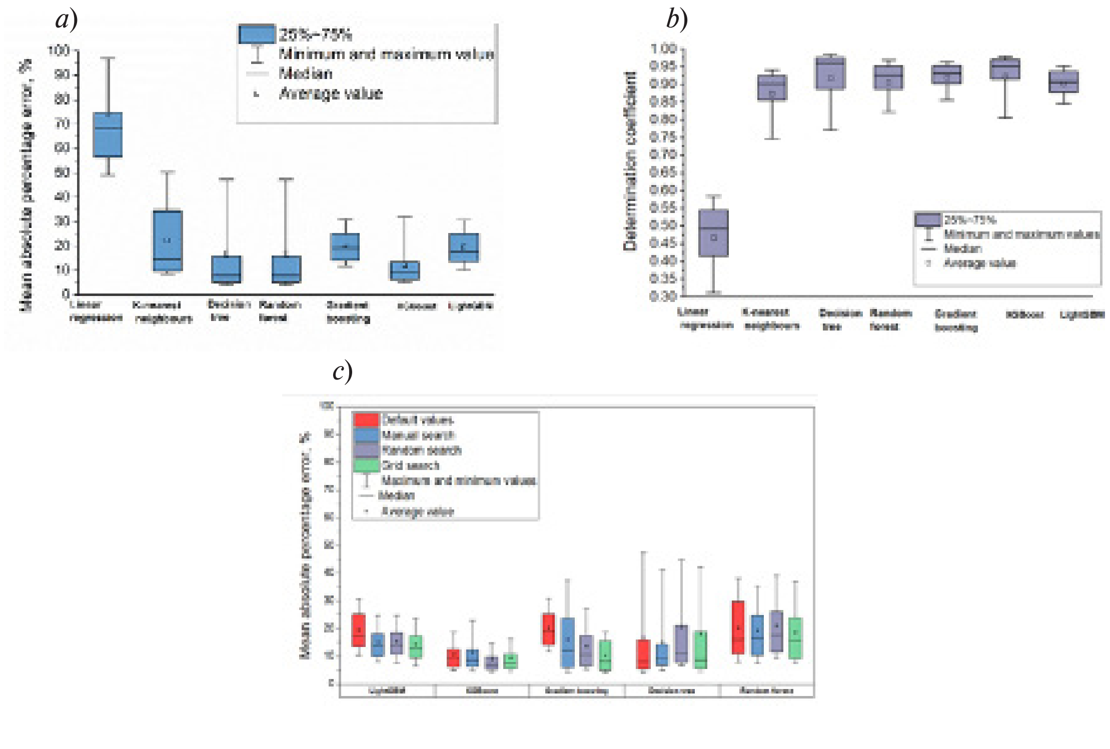

Fig. 1. Results of training and cross-validation of weld strength prediction models after laser reconstruction: MAPE of all models (*a*), determination coefficient of all models (*b*), comparison of MAPE tree-based models after optimization (*c*)

Table 1

Choosing the best data preprocessing strategy

Bringing values into the same range	Filling in the skip patterns	
	Average value	Median
MinMaxScaler	$R^2 = 0.80 \pm 0.12$	$R^2 = 0.92 \pm 0.13$
StandardScaler	$R^2 = 0.90 \pm 0.12$	$R^2 = 0.93 \pm 0.14$

Based on the results of training and cross-validation on the dataset, an extreme gradient boost based XGBoost model was selected for experimental validation of the predictive strength of laser reconstruction of biotissues. For this purpose, laser repair of bovine and porcine skin vessels was performed. A total of 80 specimens were used, which were divided into 4 groups and reconstructed with different types of solder (Fig. 2). The mean value of the coefficient of determination was 0.82, indicating the high accuracy of the predictive model.

After ex vivo validation, an in vivo study of laser tissue repair in rabbits was conducted jointly. The rabbits were divided into 2 groups. The groups differed in the presence or absence of ICG in the component composition of the solders. A comparison of actual strength and predicted strength is shown in Fig. 3. A coefficient of determination of 0.84 and an average absolute percentage error of 19% were obtained.

The error is caused by the large dispersion of strength values obtained experimentally. The achieved strength is sufficient for analyzing the laser radiation characteristics and the component composition of the solder prior to the laser reduction procedure.

a) *b*)

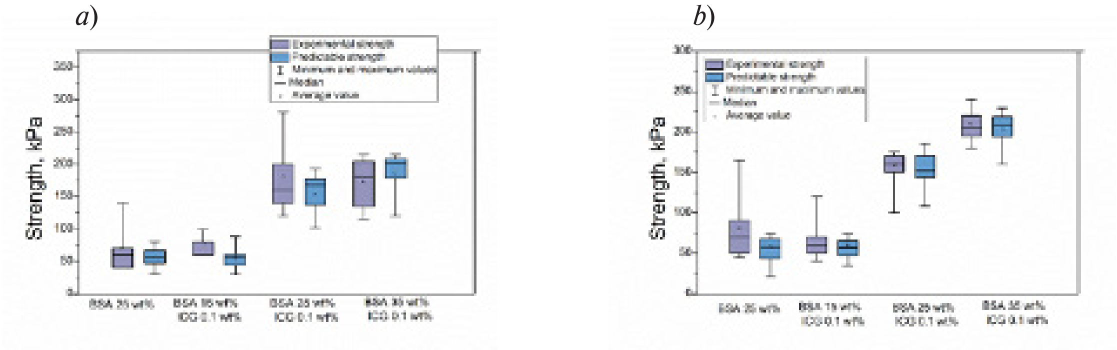

Fig. 2. *Ex vivo* validation of the prediction algorithm: laser vascular reconstruction (*a*), laser skin reconstruction (*b*)

a) *b*)

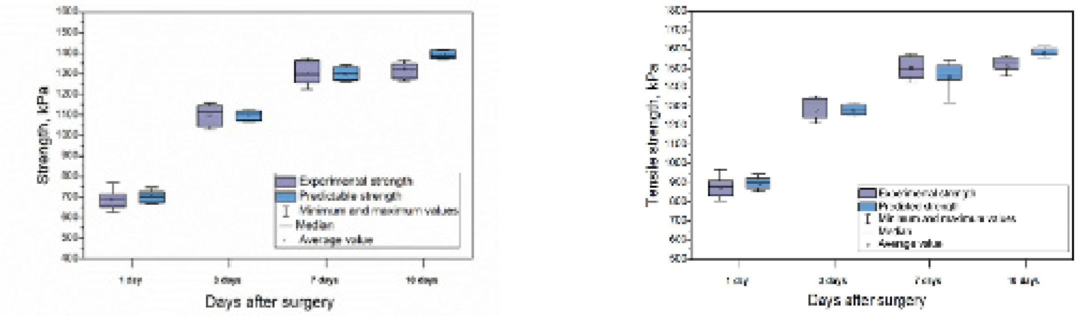

Fig. 3. *In vivo* validation of the prediction algorithm: solder based on BSA 25 wt.% (*a*), solder based on BSA 25 wt% and ICG 0.1 wt% (*b*)

Conclusion

Laser reconstruction of biological tissues is a promising minimally invasive technique, however, in order to achieve maximum recovery strength, each operation requires individual selection of laser irradiation characteristics and solder component compositions. The results of the research confirm the possibility of using trained machine learning models to predict the results of laser reconstruction of biological tissues. The use of predictive models will significantly increase the efficiency of the process of selecting optimal parameters for laser tissue reconstruction, reducing the time and cost of the experimental study.

REFERENCES

1. **Li C., Wang K., Huang J.,** Effect of scanning modes on the tensile strength and stability in laser skin welding in vitro, Optik. 179 (2019) 408–412.

2. **Gomes D.F., Galvana I, Ramos Loja M.A.,** Overview on the Evolution of Laser Welding of Vascular and Nervous Tissues, Applied sciences. 9 (2019) 9102157.

3. **Li C., Wang K.,** Effect of welding temperature and protein denaturation on strength of laser biological tissues welding, Optics and Laser Technology. 138 (2021) 106862.

4. **Fekrazad R., Mortezai O., Pedram M., Kalhori K.A., Joharchi K., Mansoori K., Mashhadiabbas F.,** Transected sciatic nerve repair by diode laser protein soldering, Journal of Photochemistry and Photobiology B: Biology. 173 (2017) 441–447.

5**. Schiavon M.,** Experimental evaluation of a new system for laser tissue welding applied on damaged lungs, Interactive cardiovascular and thoracic surgery. 16 (5) (2017) 441–447.

6. **Kumar S., Oh I., Schindler S., Lai A. M., Payne P.R., Gupta A.** Machine learning for modeling the progression of Alzheimer disease dementia using clinical data: a systematic literature review, Journal of the American Medical Informatics Association: JAMIA open. 4 (3) (2021) ooab052.

7. **Le T.L.,** Fuzzy C-means clustering interval type-2 cerebellar model articulation neural network for medical data classification, IEEE Access. 7 (2019) 20967–20973.

8. **Garg A., Mago V.,** Role of machine learning in medical research: A survey, Computer science review. 40 (2021) 100370.

9. **Potdar K., Pardawala T.S., Pai C.D.,** A comparative study of categorical variable encoding techniques for neural network classifiers, International journal of computer applications. 175 (4) (2017) 7–9.

10**. Wang S., Celebi M.E., Zhang Y.D., Yu X., Lu S., Yao X.,** Advances in data preprocessing for biomedical data fusion: An overview of the methods, challenges, and prospects, Information Fusion. 76 (2021) 376–421.

11. **Ali P. J. M., Faraj R. H., Koya E., Ali P. J. M., Faraj R. H.,** Data normalization and standardization: a technical report, Machine Learning: Science and Technology. 1 (1) (2014) 1–6.

THE AUTHORS

RYABKIN Dmitry I.
ryabkin@bms.zone
ORCID: 0000-0002-1327-5690

SUCHKOVA Victoria V.
molodykh1999@gmail.com
ORCID: 0000-0002-3207-7622

GERASIMENKO Alexander Yu.
gerasimenko@bms.zone
ORCID: 0000-0001-6514-2411

Received 10.07.2023. Approved after reviewing 27.07.2023. Accepted 01.08.2023.