Application of machine learning in joint replacement:A review

Xiaozhuang Man[a], Tianyi Zhang[a], Kai Li[a], Yunzhang Cheng*[,a], Hao Shen*[,b]
[a] School of Health Science and Engineering, University of Shanghai for Science and Technology, 334 Jungong Road, Yangpu District, Shanghai, China 200093; [b] Department of Orthopedics, Shanghai Sixth People's Hospital, Shanghai Jiao Tong University, 600 Yishan Road, Xuhui District, Shanghai, China 200233.

ABSTRACT

Joint replacement is a frequently performed surgical intervention to address end-stage joint ailments. However, successful surgery relies on early diagnosis and appropriate treatment of joint diseases. With the advancement of machine learning technology and its integration with the medical field, the use of machine learning technology in joint replacement has made substantial advancements. This article will explore the implementation and current standing of machine learning in joint replacement from four aspects: preoperative auxiliary diagnosis of arthritis, preoperative auxiliary decision-making, postoperative complication diagnosis and postoperative prediction.

Keywords: Machine learning, joint replacement, assisted diagnosis, clinical decision-making

1. INTRODUCTION

The development of algorithms and models that are capable of learning from data and making decisions or predictions is the focus of machine learning, a branch of artificial intelligence. The analysis flow of machine learning is shown in Figure 1. Machine learning aims to empower computers to enhance their performance in tasks without the need for explicit programming by enabling them to learn and improve through experience. There are various approaches to machine learning, including supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning, each with its own strengths and limitations[1]. The implementation of machine learning in the medical field holds immense promise, and it has been widely used in cardiology, dermatology, ophthalmology and many other medical fields [2-5]. Provide doctors with help in disease diagnosis, health management, image analysis, surgical guidance, treatment plan, adjuvant therapy, and prognosis prediction [6]. As data and computational power become more accessible, machine learning is poised to play a critical role in shaping the future of AI.

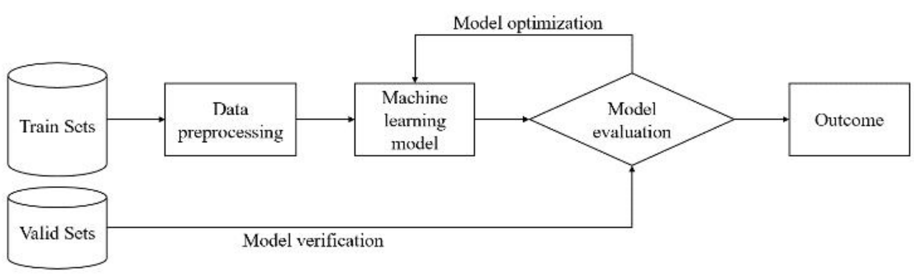

Figure 1. Machine learning process

*Yunzhang Cheng: cyz2008@usst.edu.cn; phone 021-55572161
* Hao Shen: shenhao7212@sina.com; phone 021-24058101
Xiaozhuang Man: 213332718@st.usst.edu.cn; Tianyi Zhang: 8351928@qq.com; Kai Li: 202562459@st.usst.edu.cn

Seventh International Conference on Mechatronics and Intelligent Robotics (ICMIR 2023), edited by Srikanta Patnaik, Tao Shen, Proc. of SPIE Vol. 12779, 127792G · © 2023 SPIE · 0277-786X · Published under a Creative Commons Attribution CC-BY 3.0 License · doi: 10.1117/12.2688988

Joint replacement has become a standard treatment for various advanced joint diseases, providing effective pain relief and restoring joint function. Worldwide, hip and knee arthritis are leading causes of disability [7]. Over the past few years, the application of machine learning in joint replacement has been developed, and many scholars have conducted extensive research and achieved promising results. This paper provides a comprehensive review of current research on the use of machine learning in preoperative arthritis diagnosis, clinical decision-making, postoperative complication prediction, and postoperative evaluation of joint replacement. By utilizing machine learning in these areas, joint replacement procedures can be further improved, benefiting patients and healthcare providers alike.

2. AUXILIARY DIAGNOSIS OF ARTHRITIS

Joint replacement is a common treatment for end-stage joint diseases. However, the progress of various types of osteoarthritis cannot be accurately monitored using current examination methods, and various examinations have limited roles in the early detection of knee osteoarthritis. Machine learning technology is quickly expanding its use in data analysis within the realm of computer vision. In particular, the widespread adoption of Convolutional Neural Networks (CNN) to automatically extract medium and high-level abstract features from images has been observed in various medical image analysis tasks. Machine learning models have demonstrated effectiveness in assisting doctors with the diagnosis and evaluation of the severity of osteoarthritis [8-9].

Xue et al. [10] conducted a study using deep convolutional neural network to analyze 420 hip joint X-ray images and compared the results with traditional manual evaluation. In this study, X-ray images were categorized into "normal" or "arthritic" groups by two experienced physicians. The results showed that the deep learning model accurately diagnosed mid-hip osteoarthritis with a sensitivity of 95% and a specificity of 90.7%, which was comparable to the accuracy of physician diagnosis. These results demonstrate the potential of deep learning in automating the diagnosis of hip osteoarthritis. Using a multimodal machine learning approach, Tiulpin et al. [11] constructed a predictive model for the progression of knee osteoarthritis (OA). The model utilized X-ray images, clinical examination results, and patients' medical history to evaluate the probability of OA progression and the severity of the current disease, and constructed a variety of prediction models. The study used independent test sets for validation, and found that the combination of convolutional neural network and clinical data achieved an Area Under the Receiver Operating Characteristic Curve (AUC) of 0.81 and an Average Precision (AP) of 0.70 in the training set. On the test set, the AUC was 0.80 and the AP was 0.62, demonstrating promising predictive performance of the model.

In a study by Kevin et al. [12], knee radiographs of patients who underwent Total Knee Arthroplasty (TKA) and matched control patients who did not receive TKR were utilized to construct a predictive model for the risk of osteoarthritis (OA) progression. A deep learning model based on ResNet34 was applied to anticipate the likelihood of OA patients undergoing TKA within 9 years. The prediction model achieved an AUC of 0.87, demonstrating good predictive performance.

Machine learning possesses the capability to handle and categorize extensive quantities of data,construct learning models, and improve the accuracy and ease of the diagnosis process. Its application in the early diagnosis and disease progression of arthritis is of great value in evaluating the need for joint replacement or other treatments. In the future, more intelligent image processing systems shall be developed to better assist orthopedic surgeons in diagnosing arthritis at an early stage. These systems will help to improve the accuracy and efficiency of diagnosis and treatment planning for joint diseases.

3. PREOPERATIVE DECISION-MAKING AID

Preoperative analysis of risk factors and clinical decision-making assistance can lead to a more comprehensive surgical plan and implementation of preventive measures for patients. Thus, preoperative risk prediction in patients scheduled for hip or knee replacement can aid physicians in surgical planning and close monitoring to intervene early in the event of complications.

Jo et al. [13] retrospectively analyzed clinical data of 1686 patients undergoing initial total knee arthroplasty and used preoperative variables to predict the risk of postoperative blood transfusion. Data from 43 preoperative variables were collected, and the Recursive Feature Elimination (RFE) algorithm was used to select six key variables, including various factors such as hemoglobin levels, platelet count, surgical type, use of tranexamic acid, age, and weight. A gradient boosting prediction model was established, and an independent data set was used for external validation. The predictive model showed good performance, with an AUC of 0.842 and an AUC of 0.880 on the external validation set. Blood transfusion risk prediction using machine learning can predict the risk of blood transfusion before total knee arthroplasty and guide preventive measures for high-risk patients.

Huang et al. [14] conducted a study using a large dataset of 15,187 patients who underwent lower limb joint replacement. They utilized random forest and logistic regression models to establish an intraoperative blood transfusion prediction model that identifies high-risk factors for intraoperative blood transfusion. The random forest model outperformed the logistic regression model with an AUC of 0.84. According to the study, female sex and American Society of Anesthesiologists II, III, and IV ratings were patient-related factors that were associated with a higher risk. Surgery-related risk factors included operation time, use of a drainage tube, and intraoperative blood loss. Additionally, higher preoperative hemoglobin levels and the use of tranexamic acid were associated with a lower risk of intraoperative blood transfusion. The risk factors that were identified can be used to provide personalized assessments of perioperative transfusion risk for patients who are undergoing lower extremity joint replacement. The random forest algorithm had a significantly higher prediction accuracy than the logistic regression algorithm and can potentially serve as a tool for future personalized perioperative transfusion risk prediction.

In addition, there are also studies aimed at predicting a patient's operation time and postoperative hospital stay to optimize surgical resource allocation and reduce operating room downtime. Abbas et al. [15] used a machine learning algorithm to make predictions about the operation time and hospitalization time of elective unilateral TKA using preoperative factors. The AdaBoost model was trained using linear, tree-based, and Multilayer Perceptron (MLP) models. The MSE of operation time and postoperative hospital stay were 0.003 and 0.001, respectively. Among all the models tested, the AdaBoost model demonstrated the highest classification accuracy and buffering capacity, with all models exhibiting better performance than the mean regressor in terms of accuracy for both outcome measures. On the validation set, the MLP model produced the lowest Mean Square Error (MSE) of 0.914, and all the models demonstrated superior performance over the mean regressor in terms of MSE. On the test set, the MLP performance was also the best. The ability to accurately predict the length of stay and operation time has a significant impact on hospital scheduling and resource allocation while potentially playing a role in postoperative patient planning.

4. DIAGNOSIS OF POSTOPERATIVE COMPLICATIONS

Complications following joint replacement are common, with periprosthetic joint infection (PJI) and prosthesis loosening being the most serious. PJI is particularly devastating and early, timely diagnosis of complications following joint replacement is critical. However, effective diagnosis of PJI and prosthesis loosening remains challenging, with the International Consensus Meeting (ICM) score for PJI and clinicopathological results being the primary diagnostic criteria, while imaging examinations provide only a reference.

Machine learning has demonstrated great potential in the early diagnosis of loosening and infection after joint arthroplasty. Kuo et al. [16]created a machine learning system for PJI and compared its performance to that of the ICM scoring system. To enhance the system's effectiveness, an integrated meta-learner was designed, which synergistically combined five different learning algorithms. A second-level stacked generalization architecture was used for PJI prediction, with four basic classifiers at the bottom: random forest, extreme gradient boosting, logistic regression, and naive Bayesian model. A support vector machine was used at the top level as a meta-classifier to make the final prediction. The machine learning system was compared with the ICM scoring system using cross-validation. The results showed that the machine learning system outperformed the ICM scoring system in several indicators such as accuracy, precision, recall, F1 score, Matthews correlation coefficient, and area under the receiver operating characteristic curve. Furthermore, the machine learning system was able to identify personalized important features that were missed by the ICM scoring system and Offer understandable decision-making assistance for personalized diagnosis.

Machine learning models offer an advantage over ICM by building adaptive diagnostic models using existing patient data, instead of relying on predetermined criteria for diagnosis. Experimental results demonstrate the feasibility of machine learning in diagnosing PJI compared to the widely used ICM scoring system.

Shah et al. [17] collected preoperative anteroposterior and lateral X-ray films and clinical information of 697 patients who underwent total knee arthroplasty and total hip arthroplasty. They then divided the patients into two groups: the loosening group and the fixed group. Transfer learning was applied to a series of CNN models, including ResNet, AlexNet, Inception, and DenseNet. The best performing model, DenseNet, was further enhanced with the addition of the patient's clinical history data to create the final model, which was tested on independent datasets. The results showed that the DenseNet neural network performed the best, with an accuracy of 88.3%, a sensitivity of 70.2%, and a specificity of 95.6% on the independent test dataset. Furthermore, the performance was better in cases of hip replacement, with an accuracy of 90.1%, compared to 85.8% in cases of knee replacement.

The results of the study indicate that machine learning is a viable approach for detecting prosthesis loosening from radiographic images, and that the accuracy of the algorithm can be improved by incorporating clinical data.

5. POSTOPERATIVE PREDICTION

Complications following total hip replacement may lead to a poor prognosis or even death in some patients. Despite surgery, some patients may be unable to function at their preoperative level. Moreover, impairments in lower-limb strength, balance, and gait may continue to persist for a considerable time following the surgical procedure. Predicting patient prognosis after joint replacement is important for identifying high-risk patients and providing timely intervention and treatment to improve their outcomes.

Bozic et al. [18] performed COX regression analysis to assess the impact of 29 different medical conditions on PJI and postoperative mortality. These underlying diseases were found to be associated with increased risk of periprosthetic joint infection and 90-day mortality after total hip arthroplasty. Rheumatic disease was found to be the highest risk factor for PJI, while congestive heart failure had the highest mortality at 90 days after surgery. In another study, Harris et al. [19] used a data set of more than 10,000 primary THA and TKA, collected preoperative demographic and clinical variables, established a model using LASSO regression, and evaluated it using AUC and established a prediction model for 30-day mortality and significant complications. The AUC value was 0.78 in distinguishing renal complications within 30 days after joint replacement. The AUC value for predicting death was 0.73. The AUC value of predicting cardiac complications was 0.73, which had good accuracy. External validation showed high confidence in predicting mortality and cardiac complications, but not in predicting renal complications.

Radiomics is a novel research method developed in recent years that allows for the effective identification of diseases and prediction of patient prognosis by extracting and quantifying lesion features and applying different calculation methods for individualized data analysis [20]. The flow chart of radiomics is shown in Figure 2

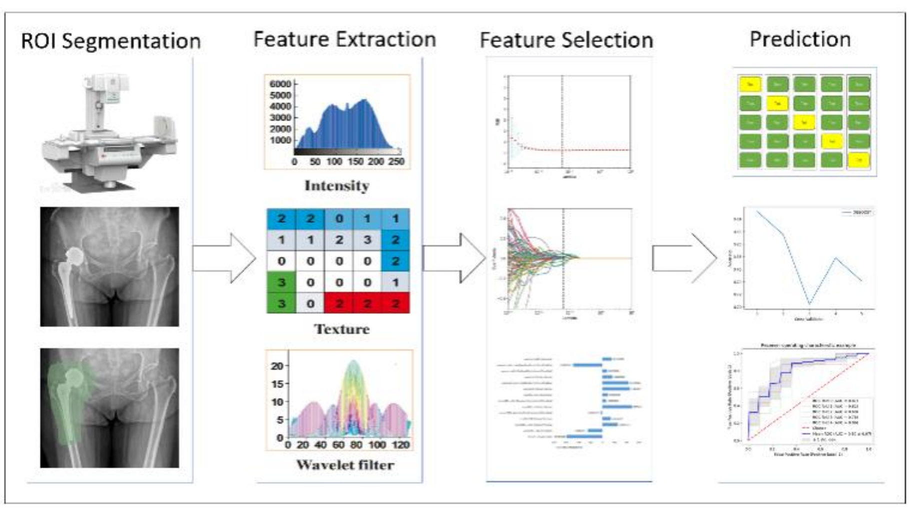

Figure 2. Radiomics analysis process

Zheng et al. [21] collected clinical and laboratory data, as well as preoperative and postoperative CT images of 85 patients who underwent total hip arthroplasty for femoral neck fracture. They preprocessed the images and established a model for the 6-month prognosis of patients using radiomics features. The Pyradiomics Python package was used to extract features from the preprocessed CT images. Lasso regression was used to select the optimal predictors, and the random forest algorithm was used to establish several prediction models. The results showed that the best model was constructed based on three features: clinical data, laboratory indicators, and preoperative radiomics features. The training set yielded an AUC, sensitivity, specificity and accuracy of 0.986, 0.925, 0.983 and 0.953, respectively. The validation set yielded an AUC, sensitivity, specificity and accuracy of 0.949, 0.767, 1.000 and 0.873, respectively. These findings indicate that utilizing a radiomics approach based on CT scans could be useful in identifying significant texture features linked to femoral neck fractures, thereby potentially enabling the prediction of patient outcomes following hip replacement. The prediction model provides clinicians with a useful tool for developing more targeted surgical plans for high-risk patients with a poor prognosis as early as possible.

6. DISCUSS

In conclusion, machine learning has demonstrated its potential to improve decision-making in all stages of joint replacement, leading to better diagnostic and treatment outcomes for patients. By automating repetitive tasks, machine learning systems can help doctors to focus on research and innovation. However, the future development of machine learning in medicine depends on close collaboration between medical professionals and technology experts, so that these tools can provide more intelligent and accurate support to medical workers and ultimately benefit patients. As such, continued research and investment in machine learning technology is essential for advancing the field of joint replacement and improving patient care.

Despite the promise of machine learning in healthcare, there are still many technical challenges that must be addressed. Machine learning approaches rely heavily on large volumes of high-quality data that accurately represent the target patient population. However, data from different hospitals may contain various biases that can prevent a model trained on one hospital's data from being generalizable to another. In the future, machine learning will continue to evolve based on clinical big data and advanced algorithms, as well as integrating fields such as biomarkers, genomics, radiomics, and metabolomics, which can further enhance its clinical predictive and decision-making value. This will lead to the development of precision medicine, enabling personalized medical services for patients, and ultimately improving patient outcomes.

REFERENCES

[1] Shinde, P. P., Shah, S., A Review of Machine Learning and Deep Learning Applications: IEEE, pp. 1-6 (2018).

[2] Rezazade Mehrizi, M. H., van Ooijen, P., Homan, M., Applications of artificial intelligence (AI) in diagnostic radiology: a technography study, European Radiology, 31(4):1805-1811(2021).

[3] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., Thrun, S., Dermatologist-level classification of skin cancer with deep neural networks, Nature, 542(7639):115-118 (2017).

[4] Gargeya, R., Leng, T., Automated Identification of Diabetic Retinopathy Using Deep Learning, Ophthalmology, 124(7):962-969 (2017).

[5] Lopez-Jimenez, F., Attia, Z., Arruda-Olson, A. M., Carter, R., Chareonthaitawee, P., Jouni, H., Kapa, S., et al., Artificial Intelligence in Cardiology: Present and Future, Mayo Clinic Proceedings, 95(5):1015-1039 (2020).

[6] Ramkumar, P. N., Kunze, K. N., Haeberle, H. S., Karnuta, J. M., Luu, B. C., Nwachukwu, B. U., Williams, R. J., Clinical and Research Medical Applications of Artificial Intelligence, Arthroscopy: The Journal of Arthroscopic & Related Surgery, 37(5):1694-1697 (2021).

[7] Cross, M., Smith, E., Hoy, D., Nolte, S., Ackerman, I., Fransen, M., Bridgett, L., et al., The global burden of hip and knee osteoarthritis: estimates from the Global Burden of Disease 2010 study, Annals of the Rheumatic Diseases, 73(7):1323-1330 (2014).

[8] Zeng, K., Hua, Y., Xu, J., Zhang, T., Wang, Z., Jiang, Y., Han, J., et al., Multicentre Study Using Machine Learning Methods in Clinical Diagnosis of Knee Osteoarthritis, Journal of Healthcare Engineering, 2021:1-12 (2021).

[9] Ashinsky, B. G., Bouhrara, M., Coletta, C. E., Lehallier, B., Urish, K. L., Lin, P., Goldberg, I. G., et al., Predicting early symptomatic osteoarthritis in the human knee using machine learning classification of magnetic resonance images from the osteoarthritis initiative, Journal of Orthopaedic Research, 35(10):2243-2250 (2017).

[10] Xue, Y., Zhang, R., Deng, Y., Chen, K., Jiang, T., A preliminary examination of the diagnostic value of deep learning in hip osteoarthritis, PLoS One, 12(6):e178992 (2017).

[11] Tiulpin, A., Klein, S., Bierma-Zeinstra, S. M. A., Thevenot, J., Rahtu, E., Meurs, J. V., Oei, E. H. G., et al., Multimodal Machine Learning-based Knee Osteoarthritis Progression Prediction from Plain Radiographs and Clinical Data, Scientific Reports, 9(1) (2019).

[12] Leung, K., Zhang, B., Tan, J., Shen, Y., Geras, K. J., Babb, J. S. and Cho, K., et al., "Prediction of Total Knee Replacement and Diagnosis of Osteoarthritis by Using Deep Learning on Knee Radiographs: Data from the Osteoarthritis Initiative, Radiology, 296(3):584-593 (2020).

[13] Jo, C., Ko, S., Shin, W. C., Han, H., Lee, M. C., Ko, T., Ro, D. H., Transfusion after total knee arthroplasty can be predicted using the machine learning algorithm, Knee Surgery, Sports Traumatology, Arthroscopy, 28(6):1757-1764 (2020).

[14] Huang, Z., Huang, C., Xie, J., Ma, J., Cao, G., Huang, Q., Shen, B., et al., Analysis of a large data set to identify predictors of blood transfusion in primary total hip and knee arthroplasty, Transfusion, 58(8):1855-1862 (2018).

[15] Abbas, A., Mosseri, J., Lex, J. R., Toor, J., Ravi, B., Khalil, E. B., Whyne, C., Machine learning using preoperative patient factors can predict duration of surgery and length of stay for total knee arthroplasty, International Journal of Medical Informatics, 158:104670 (2022).

[16] Kuo, F., Hu, W., Hu, Y., Periprosthetic Joint Infection Prediction via Machine Learning: Comprehensible Personalized Decision Support for Diagnosis, The Journal of Arthroplasty, 37(1):132-141(2022).

[17] Shah, R. F., Bini, S. A., Martinez, A. M., Pedoia, V., Vail, T. P., Incremental inputs improve the automated detection of implant loosening using machine-learning algorithms, Bone & Joint Journal, 102-B(6_Supple_A):101-106 (2020).

[18] Bozic, K. J., Lau, E., Kurtz, S., Ong, K., Rubash, H., Vail, T. P., Berry, D. J., Patient-Related Risk Factors for Periprosthetic Joint Infection and Postoperative Mortality Following Total Hip Arthroplasty in Medicare Patients, Journal of Bone and Joint Surgery, 94(9):794-800 (2012).

[19] Harris, A. H. S., Kuo, A. C., Weng, Y., Trickey, A. W., Bowe, T., Giori, N. J., Can Machine Learning Methods Produce Accurate and Easy-to-use Prediction Models of 30-day Complications and Mortality After Knee or Hip Arthroplasty?. Clinical Orthopaedics & Related Research, 477(2):452-460 (2019).

[20] Lambin, P., Rios-Velazquez, E., Leijenaar, R., Carvalho, S., van Stiphout, R. G. P. M., Granton, P., Zegers, C. M. L., et al., Radiomics: Extracting more information from medical images using advanced feature analysis, European Journal of Cancer, 48(4):441-446 (2012).

[21] Zheng, X., Xiao, C., Xie, Z., Liu, L., Chen, Y., Prediction Models for Prognosis of Femoral Neck-Fracture Patients 6 Months after Total Hip Arthroplasty, International journal of general medicine, 15:4339-4356 (2022).