Heliyon 10 (2024) e35742

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Review article

Artificial intelligence-based predictive model for guidance on treatment strategy selection in oral and maxillofacial surgery

Fanqiao Dong [a,1], Jingjing Yan [b,1], Xiyue Zhang [a], Yikun Zhang [a], Di Liu [a], Xiyun Pan [a], Lei Xue [a,b,**], Yu Liu [c,*]

[a] School of Stomatology, China Medical University, Shenyang, China
[b] Hospital of Stomatology, China Medical University, Shenyang, China
[c] First Affiliated Hospital of Jinzhou Medical University, Jinzhou, China

A R T I C L E I N F O

Keywords:
Oral and maxillofacial surgery
Artificial intelligence
Disease prediction
Disease diagnosis

A B S T R A C T

Application of deep learning (DL) and machine learning (ML) is rapidly increasing in the medical field. DL is gaining significance for medical image analysis, particularly, in oral and maxillofacial surgeries. Owing to the ability to accurately identify and categorize both diseased and normal soft- and hard-tissue structures, DL has high application potential in the diagnosis and treatment of tumors and in orthognathic surgeries. Moreover, DL and ML can be used to develop prediction models that can aid surgeons to assess prognosis by analyzing the patient's medical history, imaging data, and surgical records, develop more effective treatment strategies, select appropriate surgical modalities, and evaluate the risk of postoperative complications. Such prediction models can play a crucial role in the selection of treatment strategies for oral and maxillofacial surgeries. Their practical application can improve the utilization of medical staff, increase the treatment accuracy and efficiency, reduce surgical risks, and provide an enhanced treatment experience to patients. However, DL and ML face limitations, such as data drift, unstable model results, and vulnerable social trust. With the advancement of social concepts and technologies, the use of these models in oral and maxillofacial surgery is anticipated to become more comprehensive and extensive.

1. Introduction

Oral and maxillofacial surgery (OMS) is a multidisciplinary approach that encompasses oncology, dentistry, and orthopedic surgery [1]. For younger surgeons with limited experience, succeeding in maxillofacial surgery is considerably difficult compared to other surgical fields [2,3]. In some cases, physicians do not have sufficient knowledge to make the correct clinical decision within a limited timeframe. It is worth noting that machine learning (ML) and the broader field of artificial intelligence (AI) can be leveraged to detect anomalies in images that are not readily visible to the naked eye. Prediction models derived from deep learning (DL) serve as noninvasive aids in diagnosing oral and maxillofacial diseases, and these are used not only by surgeons but also by primary care physicians. By collecting a large amount of disease-related imaging data, the model learns the differences between regular imaging

* Corresponding author.
** Corresponding author. School of Stomatology, China Medical University, Shenyang, China.
E-mail addresses: xuelei4970362@sina.com (L. Xue), 18742095369@163.com (Y. Liu).
[1] These authors contributed equally to this work and should be considered co-first authors.

https://doi.org/10.1016/j.heliyon.2024.e35742
Received 5 March 2024; Received in revised form 27 July 2024; Accepted 2 August 2024
Available online 2 August 2024

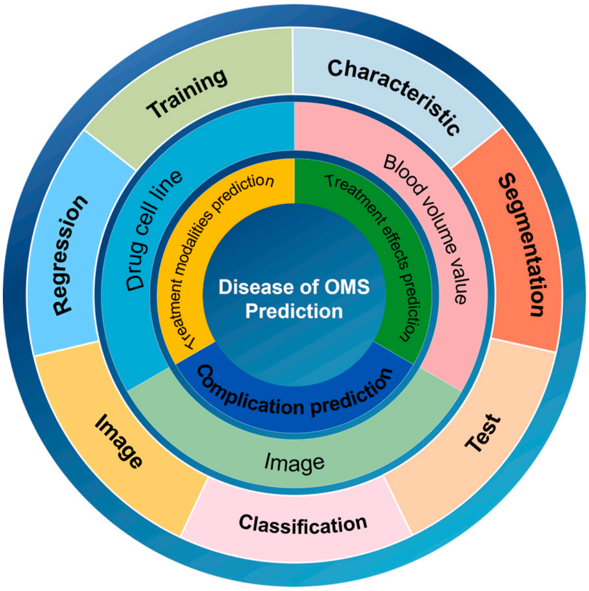

Fig. 1. Disease of OMS prediction.

features and disease-state imaging features after continuous learning and training, and can automatically extract features in unlearned images. Using this classification model, disease and non-disease images can be classified and labelled for diagnosis.

Doctors can significantly reduce the data collation work in the OMS field if the relevant rules are summarized from the existing data. Prediction of the relevant factors can be improved by using AI rather than relying solely on clinical inspiration from doctors or extensive case-screening work. The advantages of AI in disease treatment modes, complication prediction, and prognosis prediction is worth noting (Fig. 1). When predicting disease occurrence, ML models such as the stochastic forest survival model and support vector machine (SVM) model are mainly constructed based on risk factors such as age, sex, family history, and lifestyle. Further, probability analyses are conducted using a statistical algorithm. Even for patients with precancerous lesions, cancer prediction should be combined with assessing of risk factors. Predicting complications is related to several factors, such as surgical technique, doctors' experience, and the patient's medical history. Therefore, prediction of complications requires patient data, such as imaging data, to determine patient-related factors, such as anatomical factors. When predicting the prognosis of certain diseases, it is necessary to collect data on high-risk factors and analyze them together with imaging and pathological data. Therefore, using AI for disease prediction often requires more clinical patient information and is more complex than using AI for disease diagnosis.

In recent years, the popularity of AI-based disease prediction in OMS can be attributed to the increased depth of ML networks. Based on the network depth, ML can be divided into traditional ML and DL. The key to traditional ML is feature extraction; however, the network matrix must be based on the original experience and expert experience in defining the features. DL is a complex model comprising nodes (called neurons), which can be directly learned from the original data to determine the inherent law and representation level of the sample data, providing the possibility of predicting disease diagnosis and treatment [4]. There are three main DL methods: deep neural networks, recurrent neural networks, and convolutional neural networks (CNNs) [5].

This review article outlines the general concepts of AI systems, such as AI, DL, and ML programs, and their use in the diagnosis and treatment planning of various dental diseases. We provide insights into the relevant strengths and weaknesses of these technologies and draw conclusions by discussing future research areas. Learning from data can help solve problems without requiring manual input. The objective of this systematic review was to determine the development of AI applications widely used in OMS and assess their ability to predict treatments, complications, and disease outcomes. The available literature suggests that AI systems have become useful aids for radiologists in analyzing large numbers of diagnostic images, with improved diagnostic speed and accuracy.

2. AI-based oral cancer prediction model

2.1. Predictive model guidance in oral cancer treatment style

Higher levels of head and neck intensity-modulated radiatiotherapy(IMRT) treatment planning require multidisciplinary personnel. Atlas-based imaging techniques, which provide a relationship between imaging and clinical features, are the primary means of dose prediction in patients with head and neck cancers treated with radiotherapy(RT). A summary of previous treatment experiences is an important reference indicator for determining the baseline for RT and historical data from old patients can be used to infer the volumetric dose distribution for new patients [6]. Advanced atlas technologies abstract volumetric information into a subset of structured input data. Such techniques restore a set of images and contour data into a subset of descriptive data points and associate them with the corresponding patient to build a matching algorithm. Using previous patient experiences, McIntosh et al. artificially screened the most similar cases and used a regression forest algorithm to predict the dose of each voxel, ultimately generating a beam geometry for new patients [7]. The computational methods used have also significantly improved, progressing from decision trees (DTs) to random forests (RFs). DTs are mainly used to solve classification and regression problems, but leads to overfitting, resulting in a weak generalization ability. Overfitting is an important challenge in building DT models. RF is an improved approach that overcomes the overfitting issue in DT, but the error increases due to its cumbersome steps, which is a major drawback of atlas prediction.

Fully connected neural networks directly predict doses as an important means of reducing labor force and improving accuracy. Shiraishi et al. trained the target and organ-at-risk distance information and planned bundle intersection information in a simple two-layer densely linked neural network algorithm to obtain the dose information required by the patient [8]. However, this approach is prone to overfitting, unrelated considerations, and memory wastage. CNNs compensate for this defect and can extract spatial features more accurately, while focusing on local features. Dilated CNNs (DCNNs) are widely used for dose prediction [9]. DCNNs can be used to predict doses that vary with anatomy, such as in patients with head and neck cancer. Radiation doses can also be predicted by mimicking the decision of the radiologist. Wang et al. designed a network system that mimics the behavior of radiologists. By selecting a set of initial process-planning parameters from the plan library, the program parses commands directly into the treatment planning system, thereby moving the system towards a clinically acceptable plan [10]. We believe that the application of AI in image processing has evolved from traditional ML algorithms to DL algorithms. DL has significant advantages in image segmentation, classification, and lesion detection for head and neck tumors using a multilayer neural network structure. However, it should be noted that when using a deep neural network model for prediction, even small changes in the input data can lead to different classification results, thereby requiring manual review.

The concept of precision tumor treatment has been explored in studies concerning drug response at the molecular level and large-scale drug screening of novel cancer drugs for precision treatment. The selection of chemotherapy drugs for tumors using AI technology is mainly reflected in the screening of highly sensitive drugs and selection of combination drug strategies to prevent the development of drug resistance [11]. Research on these two aspects has primarily focused on molecular biology. The challenge lies in building a reliable and universal prediction model, which requires not only a large amount of screening data, but also fine-tuned datasets for particular tumor types or classes of drugs. Although this may be achieved, there could be vital issues in the economic and ecological aspects of constructing these models. Therefore, balancing the model performance and clinical applicability is an important future aspect. Several methods for pretreatment of drug reactions and calculation of biomarkers for drug reaction detection have been reported, including ML-based methods such as SVMs [12], Bayesian multitasking, and multicore learning [13]. However, these computational models have considerable room for improvement in their prediction performance and model versatility [14].

By pairing drugs known to have high anticancer potential with cell lines of different species, the prediction model of therapeutic responses and biomarkers of biological responses can be improved. The calculation strategy for pre-experimental drug response data plays an important role in controlling the search space and providing guidance for the detection process. This reduces the required workload in the experiment and has become the most important approach to efficiently leverage the calculation method. Therefore, combining the results of real-world experimental cell preconditioning with AI is a breakthrough in achieving specific cancer drug responses. Using AI to screen for molecular biomarkers of cancer is challenging; however, the most difficult part of cancer chemotherapy drug screening strategies is the mechanism of drug resistance known to accompany the drug therapy, which is the key in the treatment process. To overcome drug resistance, the primary clinical method employed is to use multiple chemotherapy drugs together to obtain a higher combined effect. The biggest advantage of this approach is the improvement achieved in treatment effectiveness without increasing the drug dose. However, predicting the combined effect of a drug is more complex than only predicting its sensitivity to a single drug; it includes evaluating the screening methods and computational power of the model.

By establishing the DeepSynergy system, Preuer et al. proposed for the first time the use of DL models to predict the effects of drug combinations, and trained the model using pharmacological data from Merck [15] composite screening and histological data from GDSC [11,16]. The model achieved high performance. DeepSynergy can be used as a classifier. However, we believe that if the model is applied to an unknown drug or cell line, it remains unclear whether it can maintain a high predictive performance. Xia et al.proposed another DL model for predicting drug-crosslinking effects [17]. Similar to other models used for drug sensitivity, it also inputs data from individual drug experiments. In addition, the most likely drug combinations were correctly identified. However, it was evaluated using only 5x cross-validation, and performance comparisons with other recent methods trained on the same dataset are lacking.

Chen et al. used the deep belief network, a novel DL architecture, to determine the effects of drug combinations as synergistic or nonsynergistic [18]. In contrast to the previously described studies, drug target information plays a crucial role in providing information on each drug target pathway.

Overfitting is a significant obstacle to prediction models for synergistic application of drugs, and a large amount of data are required for model training to achieve the training effect. Therefore, training datasets must be added to a wide range of drugs and cell lines. The most direct method is to integrate the data of different drug screenings in the past, directly unify the data analysis, and use patient data to pretrain the network to help extract clinically appropriate qualities. This method may help understand the correlation between the cell-type environment in the cell line and the patient. However, the main problem with this method is that it is difficult to guarantee consistency of the experimental conditions and methods and to specify the bias coefficient. At the same time, there are differences in the calibration and expression of genes between different test institutions in different regions, and unifying these elements is the key to solving the problem.

2.2. Guidance of prediction model in predicting oral cancer complications from chemoradiotherapy

Because head and neck cancers are in close proximity to normal tissues and organs, treatment-related toxicity is a major issue. It is important to predict the possibility of complications to prevent and treat them [19]. Mucositis is a common complication of radiotherapy for head and neck cancer, often resulting in dysphagia, pain, and decreased quality of life [20]. Dean et al. used a random forest-based classifier to demonstrate that ML can be used to precisely simulate an overview of mucositis episodes based on dose levels, spatial dose metrics, and clinical data of IMRT patients [21]. Sanguineti et al. found that oral mucosal volumes (defined as oral, oropharyngeal, and hypopharyngeal) in patients receiving 10.1 Gy (2.0 Gy per day) per week produced severe mucositis [22]. The authors also observed a positive association between concurrent chemotherapy and severe mucositis. However, the bottlenecks of such studies are reflected in the sample size, which allows for a higher degree of accuracy in broader dose finding.

Xerostomia caused by IMRT in the head and neck is the most common complication affecting survival after radiotherapy [23]. Elnaqa et al. accurately predicted the probability of dry mouth and associated toxicity by entering a logistic regression framework of variables, such as the mean parotid dose [24]. However, the predictive confidence of this side effect is generally only due to the average dose delivered to the parotid gland. However, this parameter lacks sensitivity and specificity for correctly estimating patient-specific treatment outcomes, and additional parameters are required to improve the prediction performance. Lee et al. developed a multi-component predictive model for xerostomia based on clinical data and treatment parameters [25]. The primary endpoint was moderate to severe xerostomia after IMRT. Dosimetric factors for the average dose administered to the contralateral and ipsilateral parotid glands were selected as the most important predictors, incorporating clinical and socioeconomic factors, specifically, age, financial status, T-stage, and education level. The multivariate logistic regression model using the LASSO technique based on bootstrapping significantly improved the accuracy of predicting the incidence of xerostomia in HNSCC and NPCs. Fanizzi et al. included pretreatment CT images, clinical and dose characteristics, and drinking habits as predictors and achieved good results [26].

The spatial relationship between the primary focus of head and neck tumors and angle of radiotherapy is a key factor affecting post-radiotherapy complications. Some studies have investigated the utility of radiological features obtained from images [27,28]and dose features extracted from three-dimensional (3D) dose distributions to develop predictive models for therapeutic complications [29,30]. Manual extraction of features is a common technique for image processing, segmentation of lesion areas using experienced clinicians, and modeling of these features using traditional ML methods. Sheikh et al. collected the CT and MRI images of 266 patients for manual segmentation of the lesion area followed by feature extraction and achieved good results [31]. However, manual extraction of image features is not only time-consuming and laborious, but also easily omits the relevant essential factors. Thus, hierarchical features can be automatically extracted from the data using a step-by-step DL model. Some studies have attempted to use this automated feature extraction from the 3D CNN to predict the radiographic features of the response to cancer therapy and reported the outcomes [32]. Radiation dose distribution plays an essential role in predicting the treatment outcomes. Men et al. designed and tested a model to predict toxic substances based on a three-dimensional residual convolutional neural network(3D rCNN) [33]. CT planning images and radiation therapy dose profiles can be used to obtain spatial features, and 3D filters can provide contours. Comparative experiments showed that the model can predict xerostomia with a good performance of xerostomia.

2.3. Guidance of predictive models in oral cancer prognosis prediction

Currently, the application of intelligent pathology in oral cancer is focused on tumor classification. As key features affecting the risk of malignant transformation, clinicopathological features of the underlying malignant disease of the oral cavity, including epithelial dysplasia, anatomical site, lesion size and occurrence, and various systemic complications, have been noted [34]. However, because the clinical evaluation of leukoplakia (comprising proliferative leukoplakia), erythema, and oral flattening is highly subjective, resulting in significant interclinical differences, it is difficult to predict the prognosis.

The most important aspect of using ML to learn pathological images is feature extraction from digital pathological images, including the structure, grayscale, diameter, shape, texture, and relationship with surrounding tissues of imaged tissue cells, and automatic learning of key parts of features. Exfoliative cytology has been broadly utilized for early diagnosis of oral squamous cell carcinoma (OSCC) [35]. Liu et al. developed an index called OCRI2 to assess the risk of cancer in patients with oral flattened tinea versicolor (OLK) and validated its predictive performance in new patients [36]. However, the pathology remains uncertain, and for large or multiple OLK lesions, incompetent or unqualified sampling of exfoliated cells as well as improper dyeing and imaging procedures may lead to false-negative predictions. Longitudinal detection of exfoliative cytology during follow-up may acquire such high-risk samples, and thus, kinetic data may be beneficial in predicting cancer occurrence.

Adeoye et al. found that DeepSurv and RSF had a strong ability to discern the opportunity to malignantly transform oral leukoplakia and oral flattened moss, and could provide better calibration probability estimates [37]. However, molecular data, which could have promoted the clinical effect of the ML model, were not considered in this study. Future research should consider the potential of these models and include biomarker measurements to provide stronger and more accurate predictive capabilities.

In addition to predicting the deterioration of the precancerous state, predicting the lymph node metastasis using AI is noteworthy. Cervical lymph node metastasis and extranodular expansion (ENE) in head and neck squamous cell carcinoma are key prognostic factors for tumor-bearing survival and distant metastasis and affect therapy decision making [38]. ENE occurs when metastatic tumor cells in the lymph nodes penetrate the nodular sac into the surrounding tissue [39]. Identifying ENE before treatment may help guide subsequent treatments [40].

AI with DL systems have been increasingly used in medicine, particularly in diagnostic imaging [37,41]. A DL system using a multilayer CNN can automatically obtain and analyze quantitative image characteristics and establish a prediction model. Kann et al.

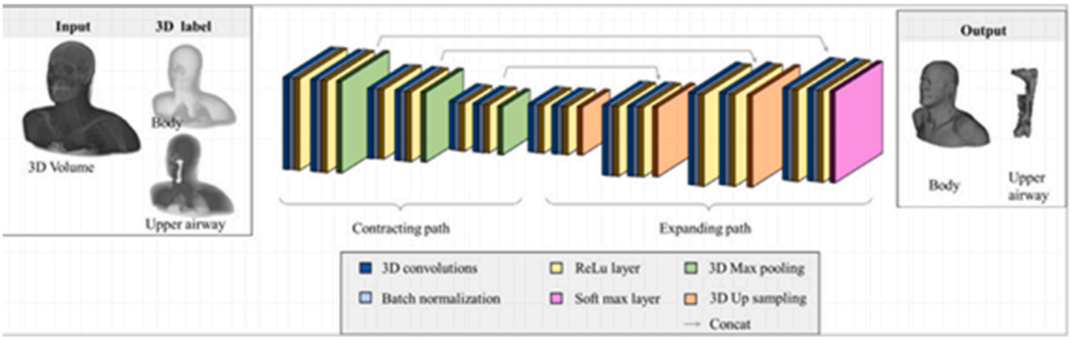

Fig. 2. 3D U-Net architecture, which inputs 3D CT images and outputs the upper airway and body regions. This architecture consists of contracting and expanding paths. Reproduced with permission from Ryu. et al., Comput Methods Programs Biomed; published by Elsevier, 2021.

used the deep learning architecture DualNet to diagnose the AUCs of ENE on CT images up to 0.91 [39]. Ariji et al. used a correlatively low-cost system using AlexNet's neural network and a DL training system on 11-GB GPUs (NVIDIA Inc., Holmdel NJ, USA) to achieve an AUC of 0.80 in the diagnosis of lymph node metastases in patients with oral cancer using enhanced CTs [42]. After studying and analyzing the data, past benign and malignant diagnoses can be re-examined to guide pathologists and clinicians to correct the diagnosis and detect lymph node metastasis of breast cancer using DL to more accurately judge the prognosis. Additional treatment was performed histologically when ENE was confirmed.

Wu et al. proposed a deep multimodal learning network called the MMC-Net [43]. It is used to predict lymph node metastasis in patients with primary thyroid cancer using clinical records and B-mode, as well as CDFI ultrasound images as input data. At the same time, the authors compared the prediction capability of MMC-Net by using only images as input data with the model using only clinical data as input data and found that the MMC-Net with multiple modes integrated had the best prediction capability. This suggests that the predictive power of AI must be improved to integrate multimodal data.

Surgery is the most common early-stage treatment option for OSCC. Nevertheless, additional chemotherapy and/or radiotherapy is used to promote disease restraint, especially in advanced tumors [44]. The prediction of survival time based on pathological diagnosis inevitably shows cohort deviation. Previous studies used AI based on factors such as age, sex, family history, and high-risk habits to predict the development of cancer [45,46]. Alabi et al. pretreated the clinical records of 311 patients to compare the performance of four ML models in predicting the risk of oral and tongue squamous cell carcinoma recurrence [46]. However, this process may not provide sufficient information to make reliable decisions.

ML algorithms depend on the quality and accuracy of input data. When the training set is unbalanced and some data features are inadequate/excessive, artificial neural networks (ANNs) may be prone to overfitting, which means that the model cannot be generalized or may mislead specific populations. Twenty-nine variables were used by Alhazmi et al. to forecast the risk of recurrence of oral cancer [47]. The model was optimized using larger cohorts of patients from multiple sites and more parameters were obtained from electronic medical records. However, it is likely that traditional medical record information currently lacks genetic mapping, biomarker analysis, and histopathological imaging, and is ultimately insufficient for predictive analysis. Tseng et al. conducted a study using ML to predict the prognosis of patients with oral cancer based on their clinical and pathological features [48]. The neural network model yielded better results than those with the traditional statistical model. The prediction of patient outcomes using ML techniques exhibited higher accuracy and superiority than those of Cox regression. Similarly, Chu et al. achieved 70.59 % accuracy, 41.98 % sensitivity, and 84.12 % high specificity in predicting prognosis in patients with oral cancer [49].

Recently, the use of molecular markers for categorizing cancer risk has attracted significant attention. However, prognostic predictions using molecular markers alone do not perform better than accepted clinical risk elements [50]. In an oral cancer study by Saintigny et al., a gene prediction model significantly improved the accuracy of prediction compared to a model utilizing clinicopathological risk factors [51]. However, the expenses and necessity for specific assessing techniques are clear barriers that must be overcome before this model can be used clinically. Moreover, improving the list of consensus genes is a significant challenge.

In recent years, as an effective diagnostic instrument for ranking patients with OSCC, FDG-PET has provided functional information on glucose uptake [52], and maximum standardized uptake values are broadly used as representative PET parameters. FDG-PET tumor function information can be measured by the extent of tumor glucose uptake. In addition, FDG-PET can assess the gross anatomical extent and shape of a tumor [53]. Fujima et al. evaluated the use of FDG-PET DL analysis to predict disease-free survival in patients with OSCC [54]. Although the model ultimately presents better results, future applications may be limited given the small amount of information in the FDG-PET image owing to its low spatial resolution and small amount of background information.

3. Application of AI in orthodontic and surgical treatment

3.1. Predictive model guidance in orthodontic and surgical treatment style

Owing to the high level of individualized need for orthodontic treatment and the higher level of physician experience required, AI technologies face three main issues in orthodontic treatment: when to start, whether to perform, and how to perform.

Table 1
Summary of machine learning and DL models in guiding treatment Style.

Stuady Authors	Yearof Pubiation	Study Theme	Sample	Model	Penformance Metric(s)	Conchusion	Referenas
Mclntosh et al.	2016	Automated dose predi ction in radi ation therapy	659 Patents' treatment planning image from hospital	Regression forests	Tovalidate their pprach on 276 pati ants from 3 clinical treatment plan sites (wholebreast. breast cavit, and prostate), with an overall dose predicin accuracies of 78.686.64.76 %, 86.83 % undr the Gamma metric	Regression forest is expected to be used to predict the radi ation dose of patients	[8]
Shirashi et al.	2016	Predi cti on of treatment planning strategy for radiati on therapy	Previously treated plans	ANN	The average prediction bias for all voxels irspectivef organ ddineation rangedfrom −1%–0 %	The study demonstrates highly accurate knowledge -based 3D dose predictions for radi otherapy pl ans	[9]
Wang et al.	2017	Predi cti on of treatment planning strategy for radiati on therapy	Not mentioned	Not mentioned	Not mentioned	IMRT/VMAT treatment planning can be readily automated	[10]
Preue et al.	2018	Predi cting anti -cancer drug synergy	23 062 samples	DL	The mean Pearson crrelation cfint between the measured and the pedicted values ofDeepSynery was 0.73 andthe AUC of asification of these noved dug combinati onsis 0.90	DeepSynergy could be a valuable tool for sel ecting novel synergistic drug combination	[12]
Xia et al.	2018	Predi cting tumor cell line response to drug pairs	Drug pair response data from NCI-ALMANAC	DL	Their best mod ahieves a mean absolute rror beow 10 %, with cofficient of determnination R2 of 0.94 and Pearson orrelation coficiant of0.97 in 5-fold cross validati on	They pres ent promising results in applying deep learning to predicting combinati onal drug response	[18]
Chan et al.	2018	Predi cting effective drug combinati on	Data from the AstraZeneca-Sanga Drug Combination	Deep belief network	The modd resulted in aprecision of 71.5 %, arecall of 60.2 % and the F score of 65.4 %	Predi cting drug synergy from literature and the omi cs data using advanced artificial intelligence approach is feasible	[19]
Lio et al.	2017	Prediction of OSA	Clinical data collected from two independent sleep centers and TMUH database	SVM	The cross validaion and tsting acauracy for the pedicion were 85.3 % and 76.7 %	It is a applicable prediction model for the severity of obstructive sleep apnea in Asi ans by simply considering their anthropometric features and age	[21]
Ryu et al.	2021	Predi ction of flow characteristics	CT data from 88 patients for automatic segmentation and another 173 patients from	3D UNet dep-learning model, MVGPR modd and SVM	The dssication acuracy, sitvt. seifcit, andF1-score of the dfagnosis agrithm were 81.5 %.89.3 %6, 86.2 %, and 87.6 %.respectively	The convenience and accuracy of sleep apnea diagnosis are improved using deep leaming and machine 1 earning	[23]
Kok et al.	2019	Predi cti on of orthodontic time	Sever ance hospital for d agnosis Cephalometic radiographs from 300 individuals	ANN	The mean acuracy values di ANN in detarmining CVsis about 77.02 %	ANN could be the preferred methodfor determining CVS	[25]
Shin et al.	2021	Predi ction of the needfor	The cephalograms of 840 pati ants	DL	The accum acy, sasitivity., and seifity were0.954,	The DL can determine the need for orthognathic surgery	[26]

(continued on next page)

Table 1 (*continued*)

Stuady Authors	Yearof Pubiation	Study Theme	Sample	Model	Penformance Metric(s)	Conchusion	Referenas
		orthodontic surgery			0.844, and0.993, respectively	with relative accuracy when using cephalogram	
Xie et al.	2010	Predi cti on of extractions	dinical data of 200 subjects	ANN	The umtr ained data from 20 ptiets in the esting set wre 80 % correct	The constructed atifi cial neural network in this study was effective	[29]
Jung et al.	2016	Predi cti on of extractions	Lateral cephalograms of 156 patients	ML	The scss rates ofthe models were 93 % for the diagnosis. of extraction vs nonextraction and 84 % for the detaild diagnosis of the extraction patterns	Artificial intelligence expert systems with neur al network machine learni ng could be useful in orthodontics	[30]

* ANN, artificial neural network; DL, deep learning; SVM, support vector machine; ML, machine learning.

The maturity of growth and development is an essential reference for judging the timing of orthodontic intervention, and changes in the cervical spine development in terms of shaping and ranking are considered the most appropriate approaches to assess the biological maturity of individuals [55]. The advantages of AI in image processing are important in determining the developmental stages of the cervical spine. One study reported that models based on ANNs showed an average accuracy of 77.02 % in determining the growth and development stages of the cervical spine [56]. These consequences are analogous to those of another study in which an AI-based model was better represented. Based on the results of these studies, AI-based automated pairing systems should be viewed as safe systems for determining the developmental stages of the cervical spine, and are of higher value to correctors.

In patients with occlusal deformities, it is necessary to determine whether patients require orthognathic surgery. Expert treatment plans are personalized, and it is difficult for clinicians with limited experience to make such judgments. The most important aspect of these decisions is an accurate and detailed diagnosis. Patient characteristics were identified using cephalometry, and expert experience was used to determine whether the difference between the maxilla and mandible could be corrected.

Shin et al. used an automatic DL structure to predict whether orthognathic surgery was required based on a patient's skull map [57]. A dataset containing 840 cases was constructed and used to estimate the proposed network. The results showed an accuracy of 0.954, sensitivity of 0.844, and specificity of 0.993. This model aids regular dentists in screening patients with malocclusion and/or malocclusion and devise a general therapy. Thus, it is beneficial for oral and maxillofacial surgeons, orthodontists, and general dentists. DL procedures make it possible to standardize decision-making processes and help in understanding the need for orthognathic surgery.

The decision to perform orthodontic tooth extraction has a significant influence on treatment prognosis. This is considered essential because of the irreversibility of the extraction process. In contrast to the difficulty in determining whether to perform orthodontic treatment, the decision to perform dental extraction requires additional factor analyses. Several multi-factor analysis methods are currently available. The most commonly used method is a statistical process known as fuzzy grouping analysis. Fuzzy grouping analysis regroups multiple factors according to their degree of compactness, which affects the extraction decision. Classification using this algorithm is applicable to several patients. Xie et al. used an AI model based on ANNs to determine whether extraction was required before correction, which showed significant results with an accuracy of 80 %, and was confirmed to be an effective method for decision-making [58]. This is similar to another study based on AI, which reported an accuracy of 92 % and can be considered an impactful model [59]. The results of these studies show that clinical decision-making for orthodontics profits from AI-based automated systems can be utilized to support professionals with less clinical experience. However, these studies require extensive data pre-processing, and the work of human doctors is not reduced.

For patients requiring orthognathic surgery, digital model surgical planning is a key step. Identifying the anatomical characteristics of the patient and developing a detailed surgical plan are of paramount importance [60]. Automatic identification of anatomical marker points can significantly assist in predicting postoperative outcomes, owing to the rapid increase in the number of patients undergoing orthognathic surgery. Therefore, although there is still a large workload in collecting clinical data and constructing neural networks in the early stages, the model can be popularized and reused on a large scale after training. Ma et al. used cascaded network structures to achieve automatic prediction of the optimal spatial position of anatomical markers using preoperative and postoperative CT images to summarize the surgeons' experience in surgical planning to help surgeons improve decision-making [61].

The 3D virtual soft tissue simulations previously used were used to predict the position of soft tissue based on the correlation between soft tissue and hard tissue and to assume a linear correspondence between soft tissue movement and bone tissue dis-placement [62]. It is affected by multiple factors such as age, sex, and preoperative facial features [63]. To promote precision, patient- and procedure-related elements should be included in soft tissue simulation models. However, to the best of our knowledge, no such model has been developed thus far.

Precision determines the success rate of orthognathic surgery and patient satisfaction [64]. Because current ML algorithms do not

Table 2
Summary of machine learning and DL models in prediction of complications.

Stuady Authors	Yearof Pubiation	Study Theme	Sample	Model	Penformance Metric(s)	Conchusion	Referenas
Dean et al.	2016	Prediction of mucositis	Data from 351 head and neck RT patients	RFC standard model	The mean AUC and calibration slope for this model were 71 %	The RFCstandard model performance is modest-to-good, but should be improved, and requires extemal validation	[23]
Le et al.	2014	Prediction of xerostomia	236 quality of life questionnaire datasets of patients	LASSO	The AUC for the HNSCC model was 0.88 and 098 for the time points of3 and 12 months, respectively, for the NPC model, the AUC was 0.87 and 0.96 for the time points of3 and 12 months, respectively	The predictedincidence of xerostomia for HNSCC and NPC patients can be improvedby using multivariable logistic regression models with LASSO technique	[25]
Fanizzi et al.	2022	Prediction of xerostomia	dinical data of 61 patients	Transfer learning	The model reached median AUC, accuacy, sensi tivity, and specilaty values of81.17 %, 83 33 %, 71.43 %, and 90.91 %, respectively	Radiomic analysis could help to develop a valid support tool for dinicians in pl anning radi otherapy treatment, by providing a risk score of the toxi city development for each individual patient	[26]
Sheikh et al.	2019	Prediction of xerostomia	CT and MR images fom 266 patients	LASSO	ROC-AUC values forClinical + DVH + CT + MR features were 0.68	The integration of baseline image features into prediction models has the potential to improve xerostomia risk strtfcation	[31]
Men et al.	2019	Prediction of xerostomia	CT, 3D dose distributionsof and contours of the parotid and submandibular glands of 784 patients	3D rCNN	The model achieves accuracy, sensitiviy,pecifity. F-score and AUC are 0.76.0.76.0.76, 0.70. and 0.84. respectively	This is a potentally ffecive model for predicting objective toxicity for supporting dinical decision-making	[33]
Yoo et al.	2021	Prediction of extraction difioulty for mandibular third molars	1053 mandibular third molars from 600 preoperative panorami cradiographic images	CNN	The prediction accuracies for C1 (depth). C2 (ramal relationship), and C3 (angulation) were 78.91 %, 82.03 %, and 90 23 %。respectivelty	The proposed CNN-based deep learning model could be used to predict the dificulty of extacting a mandibular third molar using a panoramic radi ographic image	[75]
Kim et al.	2021	To predic paresthesia after third molar extracion	panorami cradiographic images of 300 patients	CNN	The average accuracy. sensitivity, specificity, and area under the aurve were 0.827, 0.84, 0.82. and 0917. respectively	CNNs can assist in the prediction of paresthesia of the inferior alveolar nerve after third molar extraction using panoramic radiographic images	[76]

(continued on next page)

Table 2 (*continued*)

Stuady Authors	Yearof Pubiation	Study Theme	Sample	Model	Penformance Metric(s)	Conchusion	Referenas
Stehrer et al.	2019	To predict perioperative blood loss of orthognathic surgery	1472 patients	random forest algorithm	The predicted perioperative blood 1oss deviated on average only 7.4 ml	The application of random forest al gorithm allows a prediction of perioperative blood 1oss prior to orthognathic surgery	[70]

* LASSO, the least absolute shrinkage and selection operator; CNN, convolutional neural network.

yet combine patient-specific and surgically correlated elements, DL-based algorithms must be used to promote the precision of virtual soft tissue simulations. Ideally, stability over time after surgery is included in the algorithm [65]. The error range of the DL-based lower lip and chin area simulations was within 2 mm, and is clinically acceptable [66,67]. Errors within this scope do not materially affect the therapeutic development or patient communication. Ter et al. developed a DL-based algorithm to predict virtual soft tissue contours after mandibular retraction surgery and compared its accuracy with that of a mass tensor model [68]. The DL model was drilled using 3D photographs and CBCT data. The mean absolute error of the lower face region DL simulation was 1.0 ± 0.6 mm, which was significantly lower than that of the MTM-based simulation (1.5 ± 0.5 mm). Tanikawa et al. used landmark-based geometry measurements and DLs to predict facial morphology in orthognathic surgery based on the results of previous treatments, with yielded systematic errors of 0.94 mm and 0.69 mm [69]. These results are applicable to clinical scenarios.

3.2. AI for predictive model guidance for orthognathic surgery complications

Although orthognathic surgery is considered relatively safe, blood loss is a serious surgical complication [70]. Therefore, Stehrer et al. used the RF algorithm to predict perioperative blood loss, which integrates multiple decisions through ensemble learning and significantly improves the prediction accuracy without increasing the number of calculations [71]. This study included 950 patients who underwent the hemoglobin balance method to calculate perioperative blood loss based on blood volume and preoperative and postoperative hemoglobin values, which were used to predict perioperative blood loss in the test set after the training set was used to generate a predictive model. The average value of perioperative blood loss predicted by this algorithm exhibited a difference of only 7.4 ml from the actual blood loss, and bimaxillary surgery was found to have the highest characteristic importance, and may be the major cause of perioperative bleeding. The results showed that the algorithm could effectively predict perioperative bleeding before orthognathic surgery.

4. Predicting mode of diagnosis of obstructive sleep apnea (OSA)

The biggest obstacle in the treatment of OSA is diagnosis. Although polysomnography is currently used as the gold standard for diagnosis, heart rate variability and breathing sounds can only be recorded when at least 6 h of sleep data are available, and time and space costs are not conducive to the promotion and application of this technique [72].

Questionnaires and human test features can be used as rapid methods to diagnose OSA. Liu et al. designed a prediction model based on the correlation between BMI, neck circumference, waist circumference, and OSA severity, and applied the SVM algorithm. The accuracy of the results was 72.3 % [73]. We believe that although ML considers an algorithm that draws a conclusion by solving the multi-dimensional interaction between variables simultaneously to reduce the sensitivity of questionnaire detection owing to multiple factors, the character is very low, and the parameter setting does not have enough personalization.

The use of images to predict the occurrence of OSA is a more reliable method, as the shape and aerodynamic characteristics of the airway influence the tendency to obstruct the upper airway.

Computational fluid dynamics (CFD) -derived variables are closely related to the severity of OSA, where flow resistance in the narrowest airway is 3.3 times higher than that in normal subjects [74]. SVM is a representative classifier that produces the best prediction results for a binary dataset. Ryu et al. employed a 3D U-Net DL model for medical image segmentation [75]. The 3D U-Net prevented possible feature extraction losses in the junction layer and extracted the precoordination and width of the airway morphology. Changes in flow characteristics according to the upper airway morphology were analyzed using CFD. As a diagnostic step, this study used an SVM to predict aerodynamics and biometrics, classifying patients as healthy or with moderate OSA (Fig. 2). The classification accuracy, sensitivity, specificity, and F1 score of the diagnostic algorithm were 81.5 %, 89.3 %, 86.2 %, and 87.6 %, respectively.

5. Discussion

AI is no longer a distant concept and can be applied in practical clinical applications. To date, image information has been assessed in a powerful manner, particularly in computer vision. For example, because of the high level of impaction of the third mandibular molar (M3), a large number of patients experience IAN injury due to M3 extraction. Yoo et al. proposed a DL model based on CNN, and

Table 3
Summary of ML and DL models in other guided treatment effects.

Stuady Authors	Yearof Pubiation	Study Theme	Sample	Model	Penformance Metric(s)	Conchusion	Referenas
Ter et al.	2021	To predict thevirtual soft tissue profile after mandibular advancement surgery	3D photographs and CBCT of 133 patients	DL	The perentage ratis of simulations wit a high degre of accuracy (erors 1 mm) and with a medium degree of accunacy (erors2 mm) were 64.3 % and 929 %, respectively	The DL-based algrithm can predict 3D soft tssue profiles following mandiboular advancement surgey	[67]
Tanikawa et al.	2021	To predict facial morphology after orthognathic surgery	3-D facial image of 137 patients	DL and GMM	Eleven- fold CrOs9 va lidation showed that the avenge system erors were 0 94 mm for the modd ad the total success rate of<1 mm was 74 %	AI systems to predict facial mophology afer treatment were therefome confirued to be cinically acepable	[68]
Liu et al.	2018	To quantiativly predict of cancer risk in patients with oral leukoplakia	Exfoliative cytology, histopathology, and cinical follow-up data of 364 patients	Peaks-: Random Forest model	Random forest has high sensitrity (100 %) and speificity (99.2 %)	The efliative cytology-based method for quantitative prediction of cancer risk in patients wit oral leukoplakia is acceptable	[36]
Adeoye et al.	2021	To predict the probability of malignant transformation of oral potentially malignant disorders	26 features available from electronic health of 1098 patients	DL	The concordance indes and integrated Brier score are about 093 and 0035	This study scssolly utilized time to event algrithms t如mnod the malignant taosformation risk for oral lukoplakdia and oral lidumoid lesious	[37]
Wu et al.	2022	To predic lymph node metastasis in primary thyroid cancer patients	ultrasound images and cilnical data from 1131 patients	MCC-Net	The MMC-Net achieved amn avengeF1 soore of0 888 and an AUC value of0. 973 in two independent validation sets	Their work is beneficial t如o prospective cinic tials of radiologists on the diagnosis of lymph node metastasis in primary thyroid cancer	[43]
Fujima et al.	2020	To predict disease fee survival in patients with OCSCC	FDG-PET/CT images of 113 patients	DL	The diagnostic accuraay of 08 was otained using deep earning dassifiation, with a sensitvity of0.8, spedfiaity of 08, positive predictivevalue of 0.89, and negative predictive value of 0.67	Deep learning based diagnosis with FDG-PET images my predict teatmant outcome in patients with OCSCC	[54]
Alabi et al.	2020	To Predict risk of recurance of oral tonguesquamous cell carcinoma(OTSCC)	311 patient datasets	ML	Accuacy of 68 % for SVM 70 % Naive Bayes (NB), 81 % Boosted Decisio Tree (BDT) and 78 % Decision Forest (DF)	Machine algoritms should be considened in medical applications	[49]
Tseng et al.	2019	Topredict survival in patients with OCSCC	255 patient datasets	DeepSurv	c-indesx of testing sets reaching 0.781	This model can be fetire in predicting with higher acuncy and can guide cinicians both in	[48]

(continued on next page)

Table 3 (*continued*)

Stuady Authors	Yearof Pubiation	Study Theme	Sample	Model	Penformance Metric(s)	Conchusion	Referenas
						choosing treatment options and avoiding urnecessary treatments	
Saintigny et al.	2015	To predict oral cancer prognosis	Datasets	ML	Not Mentioned	Both decision tree and artifcial neural network models showed superiorityto the tnaditioual statistical model	[51]
Chu et al.	2020	To predict outome of Oral Cancer	Clinicopathological data of 467 patients	ML	70.59 % acouracy (AUC 0.67), 41 989 % sensitvity, and a high specificity of 84.129	AI models in this stady have shown promise in predicing progressive OSCC disease outomes	[49]

* GMM, geometric morphometric methods; DL, deep learning; MMC-Net, multimodal classifica-tion network; ML, machine.

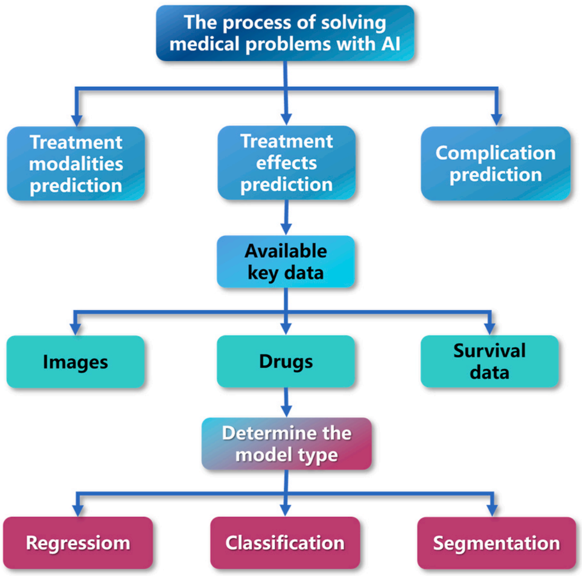

Fig. 3. The process of solving medical problems with AI.

found that it could predict the challenges in extracting M3 using panoramic radiographs [76]. Kim et al. used SSDs 300 and ResNet-18 to perform DL on panoramic films to determine whether paresthesia occurred after wisdom tooth extraction and showed that CNNs could assist in predicting paresthesia in the inferior alveolar nerve after third molar extraction [77]. Several recent studies have used AI in diagnostic imaging to overcome individual differences and ensure the repeatability and comparability of diagnostic results. The potential of this technique lies in OMS research, particularly for studying the quality of bone healing after condylar fractures [78]. The bone healing quality of the absorbable magnesium screw group was reflected through a complex calculation method and image texture feature index. Kozakiewicz discussed the limitations of one of the indices used in the results. If the deep learning method can be used to observe the texture features of CBCT images and predict the best calculation formula, it can not only save considerable human resources, but also further overcome the limitations of calculation indicators from different perspectives.

In addition to image learning and diagnostic imaging, AI is effective in integrating heterogeneous data from different fields. The majority of current research on complications and related risk factors, such as studying the complications related to the treatment of the mandibular condyles in adolescents and children, involved retrospective studies that manually collected patient data; therefore, different research methods resulted in different conclusions on surgical treatment versus conservative treatment.

Contradictions often lead to advances in medicine; however, the use of AI could lead to breakthroughs. The workload can be doubled and a geometric increase in computational engineering can be realized when multiple variables are added to an experiment. The documentation work of medical staff can be reduced by integrating these variables with AI [79]. The results are more credible and predictable because of the computing power. Moreover, automation of dental work and simple routine tasks can be assisted by AI [80, 81]. Freeing time for the medical staff often enables them to provide more care to patients during medical treatment.

AI has room for development in the prediction of complications after cleft lip and palate reconstruction [82], reduction of obstructive breathing disorder after palatal surgery [83], and prediction of complications such as hypotension and hypoxemia during surgery [84,85]. The use of AI in the OMS, especially in the prediction of hazard indicators, should be expanded. To more clearly show the status quo of OMS research on AI, we have listed the examples mentioned in this paper in Table 1, Tables 2, and Table 3. To make a long story short, AI has significant potential for predicting treatment options, complications, and prognostic factors. As a result, the costs of diagnosis and treatment and load on the healthcare system can be reduced.

In summary, the prospects for the application of AI in the medical field have been widely reported, and its potential has been demonstrated in various fields of medicine. However, this does not imply that it can be used in clinical practice. The application of AI in predicting maxillofacial surgical diseases is mainly focused on three aspects: one is the credibility of the results, the other is incorporation of AI into the daily diagnosis and treatment system, and the third is allowing doctors and patients to accept the results from AI (Fig. 3).

The first aspect has been extensively studied. From a practical perspective, the inapplicability of algorithms and low performance of models are mainly caused by the data drift problem. Current AI models use only one technical index and seldom consider multimodal data in the diagnosis process. This requires a large amount of data to train the model to reach the acceptable performance level, and can significantly improve the accuracy and performance of judgment combined with other relevant checks or clinical indexes. Therefore, image genomics based on the manual extraction of high-throughput features is often used in relatively small data volumes, and some descriptions of shape, gray scale, and texture characteristics are used to obtain more effective data.

The second aspect is an urgent problem that needs to be solved; in the process of testing the model, the results are often more ideal, but different equipment and parameters in different institutions will lead to unstable results in application. Moreover, the data input in the model training process has a certain specificity, which inevitably leads to inapplicability after different datasets are replaced and the data distribution shifts. Because most AI disease predictions are based on past experience and clinical practice, they may not be available in previous models as medical knowledge is frequently updated. The solution to this problem is not to build an extremely accurate model, but to build a more generic model. Using first principles, the essential characteristics of disease diagnosis and treatment are interpreted, so that the model can be self-updated and actively updated to ensure its applicability. In addition to the versatility of the computing model, the practicality for applications of AI is more important in economically and medically underdeveloped regions. Hardware levels in these regions are not widespread; for instance, histopathological image files are large and storage and processing require high-specification hardware. Overcoming these problems with cost considerations to establish efficient, scalable storage and computing systems to analyze images is important. The use of cloud platforms may be a good approach to transfer gigabit-size WSI images to data clouds using 5G communications.

The third and most difficult aspect is how doctors and patients can receive and trust the conclusions and data from AI. This issue involves two aspects: the interpretability of forecasts and protection of the rights of doctors and patients by third parties.

Because of the existence of the black box theory, the process of producing DL AI cannot be intuitively seen, and the working principle is generally doubtful, which makes it less interpretable in biological logic. The basic mathematical principles of such a model are well-known. However, it is difficult to understand how and why the system makes decisions, which is an important question for medical applications. Medical applications require not only high performance, but also trustworthy, transparent, and well-interpretable features. An increasing number of policies and regulations require traceable decision results from AI algorithms. Other important questions are how to get doctors to accept new workflows, including pathologists and clinicians; how to relay the results of AI to patients; and how to avoid doctor-patient disputes.

This problem cannot be solved by adding review steps or re-enhancing the performance of the algorithm, which requires policy and legal support for doctors. In China, it is difficult for social health insurance and insurance companies to cover AI-based tools. In the European Union, no AI solution with prognostic or predictive intent has qualified European markers. In conclusion, the above three problems suggest that AI is not only hindered by algorithms and computer technology, but also by the development of ethics, laws, and regulations.

To increase trust in AI, scientists must make it more prominent in people's field of vision. It should be popularized that AI can and has been involved in all aspects of human life, and brings convenience to humans. The gradual disappearance of the fear of unfamiliar objects will allow AI to develop further. Vekaria et al. [86] proposed a new method for predicting the spread of COVID-19 by combining AI and data analysis techniques, and they provided guidance for economic revitalization [87]. People's psychological receptiveness to AI will increase if it is used to solve social problems in public health crises and economic revitalization. AI technology has the potential to become a powerful tool for clinical diagnosis and treatment, but only by continuously solving social and technological problems. New ideas for updating AI technology are continuously being provided by scientists in related fields. The computational complexity of traditional ML and DL methods is high; moreover, they lack real-time and personalized data, have lengthy response times, and demonstrate low efficiency. Bhatt et al. [88]proposed a lightweight 5G-assisted federated learning architecture to improve the detection rate of strokes in cardiovascular and cerebrovascular diseases, which improved the performance of the model while protecting user privacy. Lightweight 5G-assisted federated learning methods can be utilized on smart devices to monitor individual physiological indicators in real time, enhancing diagnosis efficiency and accuracy. As the technical problems associated with

AI continue to be solved, new value will be created.

Funding

This research was funded by Shenyang Bureau Of Science and Technology, China grant number “22-321-33-76” and “2022-BS-352”.

Data availability statement

No data was used for the research described in the article.

Ethical statement

Not applicable.

CRediT authorship contribution statement

Fanqiao Dong: Writing – review & editing, Writing – original draft, Methodology. **Jingjing Yan:** Writing – review & editing, Writing – original draft. **Xiyue Zhang:** Writing – review & editing. **Yikun Zhang:** Visualization, Validation. **Di Liu:** Data curation. **Xiyun Pan:** Validation, Data curation. **Lei Xue:** Writing – review & editing, Funding acquisition. **Yu Liu:** Writing – review & editing, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank all the doctors, authors and team members who have contributed to the relevant fields, and thank the mentors for their guidance. This research was funded by, Shenyang Bureau Of Science and Technology grant number ”22-321-33-76″ and “2022-BS-352”.

References

[1] J.S. Lee, S.M. Roser, S.R. Aziz, Oral and maxillofacial surgery in low-income and middle-income countries, Oral Maxillofac. Surg. Clin. 32 (3) (2020) 355–365, https://doi.org/10.1016/j.coms.2020.04.001.
[2] J.E. Bertz, G.E. Ghali, T.P. Williams, Creating the successful global maxillofacial surgeon: a 35-year perspective, Oral Maxillofac. Surg. Clin. 32 (3) (2020) 489–493, https://doi.org/10.1016/j.coms.2020.04.011.
[3] A.L. Felsenfeld, A. Casagrande, Oral and maxillofacial surgery residency education, J. Calif. Dent. Assoc. 32 (10) (2004) 817–822, https://doi.org/10.1080/19424396.2004.12224029.
[4] S.B. Khanagar, A. Al-Ehaideb, P.C. Maganur, et al., Developments, application, and performance of artificial intelligence in dentistry - a systematic review, J. Dent. Sci. 16 (1) (2021) 508–522, https://doi.org/10.1016/j.jds.2020.06.019.
[5] Y. Lecun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (7553) (2015) 436–444, https://doi.org/10.1038/nature14539.
[6] G. Valdes, C.B. Simone, J. Chen, et al., Clinical decision support of radiotherapy treatment planning: a data-driven machine learning strategy for patient-specific dosimetric decision making, Radiother. Oncol. 125 (3) (2017) 392–397, https://doi.org/10.1016/j.radonc.2017.10.014.
[7] C. Mcintosh, T.G. Purdie, Contextual atlas regression forests: multiple-atlas-based automated dose prediction in radiation therapy, IEEE Trans. Med. Imag. 35 (4) (2016) 1000–1012, https://doi.org/10.1109/tmi.2015.2505188.
[8] S. Shiraishi, K.L. Moore, Knowledge-based prediction of three-dimensional dose distributions for external beam radiotherapy, Med. Phys. 43 (1) (2016) 378, https://doi.org/10.1118/1.4938583.
[9] V. Kearney, J.W. Chan, G. Valdes, et al., The application of artificial intelligence in the IMRT planning process for head and neck cancer, Oral Oncol. 87 (2018) 111–116, https://doi.org/10.1016/j.oraloncology.2018.10.026.
[10] H. Wang, P. Dong, H. Liu, et al., Development of an autonomous treatment planning strategy for radiation therapy with effective use of population-based prior data, Med. Phys. 44 (2) (2017) 389–396, https://doi.org/10.1002/mp.12058.
[11] K. Preuer, R.P.I. Lewis, S. Hochreiter, et al., DeepSynergy: predicting anti-cancer drug synergy with deep learning, Bioinformatics 34 (9) (2018) 1538–1546, https://doi.org/10.1093/bioinformatics/btx806.
[12] C. Huang, R. Mezencev, J.F. Mcdonald, et al., Open source machine-learning algorithms for the prediction of optimal cancer drug therapies, PLoS One 12 (10) (2017) e0186906, https://doi.org/10.1371/journal.pone.0186906.
[13] J.C. Costello, L.M. Heiser, E. Georgii, et al., A community effort to assess and improve drug sensitivity prediction algorithms, Nat. Biotechnol. 32 (12) (2014) 1202–1212, https://doi.org/10.1038/nbt.2877.
[14] D. Baptista, P.G. Ferreira, M. Rocha, Deep learning for drug response prediction in cancer, Briefings Bioinf. 22 (1) (2021) 360–379, https://doi.org/10.1093/bib/bbz171.
[15] J. O'neil, Y. Benita, I. Feldman, et al., An unbiased oncology compound screen to identify novel combination strategies, Mol. Cancer Therapeut. 15 (6) (2016) 1155–1162, https://doi.org/10.1158/1535-7163.Mct-15-0843.
[16] F. Iorio, T.A. Knijnenburg, D.J. Vis, et al., A landscape of pharmacogenomic interactions in cancer, Cell 166 (3) (2016) 740–754, https://doi.org/10.1016/j.cell.2016.06.017.
[17] F. Xia, M. Shukla, T. Brettin, et al., Predicting tumor cell line response to drug pairs with deep learning, BMC Bioinf. 19 (Suppl 18) (2018) 486, https://doi.org/10.1186/s12859-018-2509-3.

[18] G. Chen, A. Tsoi, H. Xu, et al., Predict effective drug combination by deep belief network and ontology fingerprints, J. Biomed. Inf. 85 (2018) 149–154, https://doi.org/10.1016/j.jbi.2018.07.024.
[19] P.G. Hawkins, J.Y. Lee, Y. Mao, et al., Sparing all salivary glands with IMRT for head and neck cancer: longitudinal study of patient-reported xerostomia and head-and-neck quality of life, Radiother. Oncol. 126 (1) (2018) 68–74, https://doi.org/10.1016/j.radonc.2017.08.002.
[20] G. Sanguineti, N. Rao, B. Gunn, et al., Predictors of PEG dependence after IMRT±chemotherapy for oropharyngeal cancer, Radiother. Oncol. 107 (3) (2013) 300–304, https://doi.org/10.1016/j.radonc.2013.05.021.
[21] J.A. Dean, K.H. Wong, L.C. Welsh, et al., Normal tissue complication probability (NTCP) modelling using spatial dose metrics and machine learning methods for severe acute oral mucositis resulting from head and neck radiotherapy, Radiother. Oncol. 120 (1) (2016) 21–27, https://doi.org/10.1016/j.radonc.2016.05.015.
[22] G. Sanguineti, M.P. Sormani, S. Marur, et al., Effect of radiotherapy and chemotherapy on the risk of mucositis during intensity-modulated radiation therapy for oropharyngeal cancer, Int. J. Radiat. Oncol. Biol. Phys. 83 (1) (2012) 235–242, https://doi.org/10.1016/j.ijrobp.2011.06.2000.
[23] A. Trotti, Toxicity in head and neck cancer: a review of trends and issues, Int. J. Radiat. Oncol. Biol. Phys. 47 (1) (2000) 1–12, https://doi.org/10.1016/s0360-3016(99)00558-1.
[24] I. El Naqa, J. Bradley, A.I. Blanco, et al., Multivariable modeling of radiotherapy outcomes, including dose-volume and clinical factors, Int. J. Radiat. Oncol. Biol. Phys. 64 (4) (2006) 1275–1286, https://doi.org/10.1016/j.ijrobp.2005.11.022.
[25] T.F. Lee, M.H. Liou, Y.J. Huang, et al., LASSO NTCP predictors for the incidence of xerostomia in patients with head and neck squamous cell carcinoma and nasopharyngeal carcinoma, Sci. Rep. 4 (2014) 6217, https://doi.org/10.1038/srep06217.
[26] A. Fanizzi, G. Scognamillo, A. Nestola, et al., Transfer learning approach based on computed tomography images for predicting late xerostomia after radiotherapy in patients with oropharyngeal cancer, Front. Med. 9 (2022) 993395, https://doi.org/10.3389/fmed.2022.993395.
[27] H.S. Gabryś, F. Buettner, F. Sterzing, et al., Design and selection of machine learning methods using radiomics and dosiomics for normal tissue complication probability modeling of xerostomia, Front. Oncol. 8 (2018) 35, https://doi.org/10.3389/fonc.2018.00035.
[28] B.S. Rosen, P.G. Hawkins, D.F. Polan, et al., Early changes in serial CBCT-measured parotid gland biomarkers predict chronic xerostomia after head and neck radiation therapy, Int. J. Radiat. Oncol. Biol. Phys. 102 (4) (2018) 1319–1329, https://doi.org/10.1016/j.ijrobp.2018.06.048.
[29] I. Improta, F. Palorini, C. Cozzarini, et al., Bladder spatial-dose descriptors correlate with acute urinary toxicity after radiation therapy for prostate cancer, Phys. Med. 32 (12) (2016) 1681–1689, https://doi.org/10.1016/j.ejmp.2016.08.013.
[30] L. Rossi, R. Bijman, W. Schillemans, et al., Texture analysis of 3D dose distributions for predictive modelling of toxicity rates in radiotherapy, Radiother. Oncol. 129 (3) (2018) 548–553, https://doi.org/10.1016/j.radonc.2018.07.027.
[31] K. Sheikh, S.H. Lee, Z. Cheng, et al., Predicting acute radiation induced xerostomia in head and neck cancer using MR and CT radiomics of parotid and submandibular glands, Radiat. Oncol. 14 (1) (2019) 131, https://doi.org/10.1186/s13014-019-1339-4.
[32] Z. Zhou, L. Chen, D. Sher, et al., Predicting lymph node metastasis in head and neck cancer by combining many-objective radiomics and 3-dimensioal convolutional neural network through evidential reasoning, Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. 2018 (2018) 1–4, https://doi.org/10.1109/embc.2018.8513070.
[33] K. Men, H. Geng, H. Zhong, et al., A deep learning model for predicting xerostomia due to radiation therapy for head and neck squamous cell carcinoma in the RTOG 0522 clinical trial, Int. J. Radiat. Oncol. Biol. Phys. 105 (2) (2019) 440–447, https://doi.org/10.1016/j.ijrobp.2019.06.009.
[34] M. González-Moles, I. Ruiz-Ávila, L. González-Ruiz, et al., Malignant transformation risk of oral lichen planus: a systematic review and comprehensive meta-analysis, Oral Oncol. 96 (2019) 121–130, https://doi.org/10.1016/j.oraloncology.2019.07.012.
[35] A. Sahu, P. Gera, V. Pai, et al., Raman exfoliative cytology for oral precancer diagnosis, J. Biomed. Opt. 22 (11) (2017) 1–12, https://doi.org/10.1117/1.Jbo.22.11.115003.
[36] Y. Liu, Y. Li, Y. Fu, et al., Quantitative prediction of oral cancer risk in patients with oral leukoplakia, Oncotarget 8 (28) (2017) 46057–46064, https://doi.org/10.18632/oncotarget.17550.
[37] J. Adeoye, M. Koohi-Moghadam, A.W.I. Lo, et al., Deep learning predicts the malignant-transformation-free survival of oral potentially malignant disorders, Cancers 13 (23) (2021) 6054, https://doi.org/10.3390/cancers13236054.
[38] F. Yan, H. Li, J.R. De Almeida, et al., Microscopic extranodal extension in HPV-negative head and neck cancer and the role of adjuvant chemoradiation, Otolaryngol. Head Neck Surg. 165 (4) (2021) 536–549, https://doi.org/10.1177/0194599821989637.
[39] B.H. Kann, S. Aneja, G.V. Loganadane, et al., Pretreatment identification of head and neck cancer nodal metastasis and extranodal extension using deep learning neural networks, Sci. Rep. 8 (1) (2018) 14036, https://doi.org/10.1038/s41598-018-32441-y.
[40] E.M. Anderson, M. Luu, B.L. Balzer, et al., Variations in the association of grade with survival across the head and neck cancer landscape, Head Neck 43 (4) (2021) 1105–1115, https://doi.org/10.1002/hed.26566.
[41] M.Q. Liu, Z.N. Xu, W.Y. Mao, et al., Deep learning-based evaluation of the relationship between mandibular third molar and mandibular canal on CBCT, Clin. Oral Invest. 26 (1) (2022) 981–991, https://doi.org/10.1007/s00784-021-04082-5.
[42] Y. Ariji, M. Fukuda, Y. Kise, et al., 42l intelligence, Oral Surg. Oral Med. Oral Pathol. Oral Radiol. 127 (5) (2019) 458–463, https://doi.org/10.1016/j.oooo.2018.10.002.
[43] X. Wu, M. Li, X.W. Cui, et al., Deep multimodal learning for lymph node metastasis prediction of primary thyroid cancer, Phys. Med. Biol. 67 (3) (2022) 035008, https://doi.org/10.1088/1361-6560/ac4c47.
[44] Y.J. Cheng, M.H. Tsai, C.J. Chiang, et al., Adjuvant radiotherapy after curative surgery for oral cavity squamous cell carcinoma and treatment effect of timing and duration on outcome-A Taiwan Cancer Registry national database analysis, Cancer Med. 7 (7) (2018) 3073–3083, https://doi.org/10.1002/cam4.1611.
[45] D.W. Kim, S. Lee, S. Kwon, et al., Deep learning-based survival prediction of oral cancer patients, Sci. Rep. 9 (1) (2019) 6994, https://doi.org/10.1038/s41598-019-43372-7.
[46] R.O. Alabi, M. Elmusrati, I. Sawazaki-Calone, et al., Comparison of supervised machine learning classification techniques in prediction of locoregional recurrences in early oral tongue cancer, Int. J. Med. Inf. 136 (2020) 104068, https://doi.org/10.1016/j.ijmedinf.2019.104068.
[47] A. Alhazmi, Y. Alhazmi, A. Makrami, et al., Application of artificial intelligence and machine learning for prediction of oral cancer risk, J. Oral Pathol, Méd. 50 (5) (2021) 444–450, https://doi.org/10.1111/jop.13157.
[48] W.T. Tseng, W.F. Chiang, S.Y. Liu, et al., The application of data mining techniques to oral cancer prognosis, J. Med. Syst. 39 (5) (2015) 59, https://doi.org/10.1007/s10916-015-0241-3.
[49] C.S. Chu, N.P. Lee, J. Adeoye, et al., Machine learning and treatment outcome prediction for oral cancer, J. Oral Pathol. Med. 49 (10) (2020) 977–985, https://doi.org/10.1111/jop.13089.
[50] J. Carmicheal, A. Patel, V. Dalal, et al., Elevating pancreatic cystic lesion stratification: current and future pancreatic cancer biomarker(s), Biochim. Biophys. Acta Rev. Canc 1873 (1) (2020) 188318, https://doi.org/10.1016/j.bbcan.2019.188318.
[51] P. Saintigny, L. Zhang, Y.H. Fan, et al., Gene expression profiling predicts the development of oral cancer, Cancer Prev. Res. 4 (2) (2011) 218–229, https://doi.org/10.1158/1940-6207.Capr-10-0155.
[52] M.A. Pasha, C. Marcus, C. Fakhry, et al., FDG PET/CT for management and assessing outcomes of squamous cell cancer of the oral cavity, AJR Am. J. Roentgenol. 205 (2) (2015) W150–W161, https://doi.org/10.2214/ajr.14.13830.
[53] M. Kimura, I. Kato, K. Ishibashi, et al., The prognostic significance of intratumoral heterogeneity of 18F-FDG uptake in patients with oral cavity squamous cell carcinoma, Eur. J. Radiol. 114 (2019) 99–104, https://doi.org/10.1016/j.ejrad.2019.03.004.
[54] N. Fujima, V.C. Andreu-Arasa, S.K. Meibom, et al., Deep learning analysis using FDG-PET to predict treatment outcome in patients with oral cavity squamous cell carcinoma, Eur. Radiol. 30 (11) (2020) 6322–6330, https://doi.org/10.1007/s00330-020-06982-8.
[55] H.A. Alkhal, R.W. Wong, A.B. Rabie, Correlation between chronological age, cervical vertebral maturation and Fishman's skeletal maturity indicators in southern Chinese, Angle Orthod. 78 (4) (2008) 591–596, https://doi.org/10.2319/0003-3219(2008)078[0591:CBCACV]2.0.CO;2.

[56] H. Kök, A.M. Acilar, M.S. İzgi, Usage and comparison of artificial intelligence algorithms for determination of growth and development by cervical vertebrae stages in orthodontics, Prog. Orthod. 20 (1) (2019) 41, https://doi.org/10.1186/s40510-019-0295-8.
[57] W. Shin, H.G. Yeom, G.H. Lee, et al., Deep learning based prediction of necessity for orthognathic surgery of skeletal malocclusion using cephalogram in Korean individuals, BMC Oral Health 21 (1) (2021) 130, https://doi.org/10.1186/s12903-021-01513-3.
[58] X. Xie, L. Wang, A. Wang, Artificial neural network modeling for deciding if extractions are necessary prior to orthodontic treatment, Angle Orthod. 80 (2) (2010) 262–266, https://doi.org/10.2319/111608-588.1.
[59] S.K. Jung, T.W. Kim, New approach for the diagnosis of extractions with neural network machine learning, Am. J. Orthod. Dentofacial Orthop. 149 (1) (2016) 127–133, https://doi.org/10.1016/j.ajodo.2015.07.030.
[60] N. Zamora, J.M. Llamas, R. Cibrián, et al., A study on the reproducibility of cephalometric landmarks when undertaking a three-dimensional (3D) cephalometric analysis, Med. Oral Patol. Oral Cir, Bucal 17 (4) (2012) e678–e688, https://doi.org/10.4317/medoral.17721.
[61] Q. Ma, E. Kobayashi, B. Fan, et al., Machine-learning-based approach for predicting postoperative skeletal changes for orthognathic surgical planning, Int. J. Med. Robot 18 (3) (2022) e2379, https://doi.org/10.1002/rcs.2379.
[62] J. San Miguel Moragas, W. Van Cauteren, M.Y. Mommaerts, A systematic review on soft-to-hard tissue ratios in orthognathic surgery part I: maxillary repositioning osteotomy, J. Cranio-Maxillo-Fac. Surg. 42 (7) (2014) 1341–1351, https://doi.org/10.1016/j.jcms.2014.03.024.
[63] A. Bral, S. Olate, C. Zaror, et al., A prospective study of soft- and hard-tissue changes after mandibular advancement surgery: midline changes in the chin area, Am. J. Orthod. Dentofacial Orthop. 157 (5) (2020) 662–667, https://doi.org/10.1016/j.ajodo.2019.05.022.
[64] F. Ferraz, L. Iwaki-Filho, G.N. Souza-Pinto, et al., A comparative study of the accuracy between two computer-aided surgical simulation methods in virtual surgical planning, J. Cranio-Maxillo-Fac. Surg. 49 (2) (2021) 84–92, https://doi.org/10.1016/j.jcms.2020.12.002.
[65] K. Stokbro, S.W. Borg, M. Andersen, et al., Patient-specific 3D printed plates improve stability of Le Fort 1 osteotomies in vitro, J. Cranio-Maxillo-Fac. Surg. 47 (3) (2019) 394–399, https://doi.org/10.1016/j.jcms.2018.12.015.
[66] P.G.M. Knoops, A. Papaioannou, A. Borghi, et al., A machine learning framework for automated diagnosis and computer-assisted planning in plastic and reconstructive surgery, Sci. Rep. 9 (1) (2019) 13597, https://doi.org/10.1038/s41598-019-49506-1.
[67] H.S. Cunha, C.A. Da Costa Moraes, R. De Faria Valle Dornelles, et al., Accuracy of three-dimensional virtual simulation of the soft tissues of the face in OrtogOnBlender for correction of class II dentofacial deformities: an uncontrolled experimental case-series study, Oral Maxillofac, Surgery (St Louis) 25 (3) (2021) 319–335, https://doi.org/10.1007/s10006-020-00920-0.
[68] R. Ter Horst, H. Van Weert, T. Loonen, et al., Three-dimensional virtual planning in mandibular advancement surgery: soft tissue prediction based on deep learning, J. Cranio-Maxillo-Fac. Surg. 49 (9) (2021) 775–782, https://doi.org/10.1016/j.jcms.2021.04.001.
[69] C. Tanikawa, T. Yamashiro, Development of novel artificial intelligence systems to predict facial morphology after orthognathic surgery and orthodontic treatment in Japanese patients, Sci. Rep. 11 (1) (2021) 15853, https://doi.org/10.1038/s41598-021-95002-w.
[70] A. Piñeiro-Aguilar, M. Somoza-Martín, J.M. Gandara-Rey, et al., Blood loss in orthognathic surgery: a systematic review, J. Oral Maxillofac. Surg. 69 (3) (2011) 885–892, https://doi.org/10.1016/j.joms.2010.07.019.
[71] R. Stehrer, L. Hingsammer, C. Staudigl, et al., Machine learning based prediction of perioperative blood loss in orthognathic surgery, J. Cranio-Maxillo-Fac. Surg. 47 (11) (2019) 1676–1681, https://doi.org/10.1016/j.jcms.2019.08.005.
[72] S. Wu, D. Liang, Q. Yang, et al., Regularity of heart rate fluctuations analysis in obstructive sleep apnea patients using information-based similarity, Biomed. Signal Process Control 65 (2021) 102370, https://doi.org/10.1016/j.bspc.2020.102370.
[73] W.T. Liu, H.T. Wu, J.N. Juang, et al., Prediction of the severity of obstructive sleep apnea by anthropometric features via support vector machine, PLoS One 12 (5) (2017) e0176991, https://doi.org/10.1371/journal.pone.0176991.
[74] D.M. Wootton, S. Sin, H. Luo, et al., Computational fluid dynamics upper airway effective compliance, critical closing pressure, and obstructive sleep apnea severity in obese adolescent girls, J. Appl. Physiol. 121 (4) (1985) 925–931, https://doi.org/10.1152/japplphysiol.00190.2016, 2016.
[75] S. Ryu, J.H. Kim, H. Yu, et al., Diagnosis of obstructive sleep apnea with prediction of flow characteristics according to airway morphology automatically extracted from medical images: computational fluid dynamics and artificial intelligence approach, Comput. Methods Progr. Biomed. 208 (2021) 106243, https://doi.org/10.1016/j.cmpb.2021.106243.
[76] J.H. Yoo, H.G. Yeom, W. Shin, et al., Deep learning based prediction of extraction difficulty for mandibular third molars, Sci. Rep. 11 (1) (2021) 1954, https://doi.org/10.1038/s41598-021-81449-4.
[77] B.S. Kim, H.G. Yeom, J.H. Lee, et al., Deep learning-based prediction of paresthesia after third molar extraction: a preliminary study, Diagnostics 11 (9) (2021) 1572, https://doi.org/10.3390/diagnostics11091572.
[78] M. Kozakiewicz, I. Gabryelczak, Bone union quality after fracture fixation of mandibular head with compression magnesium screws, Materials (6) (2022) 2230, https://doi.org/10.3390/ma15062230.
[79] S.T. Israni, A. Verghese, Humanizing artificial intelligence, JAMA 321 (1) (2019) 29–30, https://doi.org/10.1001/jama.2018.19398.
[80] J.H. Lee, D.H. Kim, S.N. Jeong, et al., Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm, J. Dent. 77 (2018) 106–111, https://doi.org/10.1016/j.jdent.2018.07.015.
[81] J. Grischke, L. Johannsmeier, L. Eich, et al., Dentronics: review, first concepts and pilot study of a new application domain for collaborative robots in dental assistance, in: 2019 International Conference on Robotics and Automation (ICRA), IEEE, Piscataway, 2019, pp. 6525–6532, https://doi.org/10.1109/ICRA.2019.8794139.
[82] Y. Li, J. Cheng, H. Mei, et al., CLPNet: cleft lip and palate surgery support with deep learning, Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. 2019 (2019) 3666–3672, https://doi.org/10.1109/embc.2019.8857799.
[83] S.J. Yang, J.S. Kim, S.K. Chung, et al., Machine learning-based model for prediction of outcomes in palatal surgery for obstructive sleep apnoea, Clin. Otolaryngol. 46 (6) (2021) 1242–1246, https://doi.org/10.1111/coa.13823.
[84] M. Wijnberge, B.F. Geerts, L. Hol, et al., Effect of a machine learning-derived early warning system for intraoperative hypotension vs standard care on depth and duration of intraoperative hypotension during elective noncardiac surgery: the HYPE randomized clinical trial, JAMA 323 (11) (2020) 1052–1060, https://doi.org/10.1001/jama.2020.0592.
[85] S.M. Lundberg, B. Nair, M.S. Vavilala, et al., Explainable machine-learning predictions for the prevention of hypoxaemia during surgery, Nat. Biomed. Eng. 2 (10) (2018) 749–760, https://doi.org/10.1038/s41551-018-0304-0.
[86] H. Bhatt, N.K. Jadav, A. Kumari, et al., Artificial neural network-driven federated learning for heart stroke prediction in healthcare 4.0 underlying 5G, Concurr. Comput. 36 (3) (2024) e7911, https://doi.org/10.1002/cpe.7911.
[87] I. Shah, C. Doshi, M. Patel, et al., A comprehensive review of the technological solutions to analyse the effects of pandemic outbreak on human lives, Medicina 58 (2) (2022) 311, https://doi.org/10.3390/medicina58020311.
[88] D. Vekaria, A. Kumari, S. Tanwar, et al., ξboost: an AI-based data analytics scheme for COVID-19 prediction and economy boosting, IEEE Internet Things J. 8 (21) (2021) 15977–15989, https://doi.org/10.1109/jiot.2020.3047539.