

# Applications and challenges of neural networks in otolaryngology (Review)

IULIAN-ALEXANDRU TACIUC<sup>1</sup>, MIHAI DUMITRU<sup>2</sup>, DANIELA VRINCEANU<sup>2</sup>,  
MIRELA GHERGHE<sup>3</sup>, FELICIA MANOLE<sup>4</sup>, ANDREEA MARINESCU<sup>5</sup>,  
CRENGUTA SERBOIU<sup>6</sup>, ADRIANA NEAGOS<sup>7</sup> and ADRIAN COSTACHE<sup>1</sup>

<sup>1</sup>Department of Pathology, 'Carol Davila' University of Medicine and Pharmacy, 020021 Bucharest, Romania;

<sup>2</sup>Department of ENT, 'Carol Davila' University of Medicine and Pharmacy, 050751 Bucharest, Romania;

<sup>3</sup>Department of Nuclear Medicine, 'Carol Davila' University of Medicine and Pharmacy, 022328 Bucharest, Romania;

<sup>4</sup>Department of ENT, Faculty of Medicine University of Oradea, 410073 Oradea, Romania; <sup>5</sup>Department of Radiology and Medical Imaging 'Carol Davila' University of Medicine and Pharmacy, 050096 Bucharest, Romania;

<sup>6</sup>Department of Cell Biology, Molecular and Histology, 'Carol Davila' University of Medicine and Pharmacy, 050096 Bucharest, Romania; <sup>7</sup>Department of ENT, 'George Emil Palade' University of Medicine, Pharmacy, Science, and Technology of Targu Mures, 540142 Mures, Romania

Received January 28, 2024; Accepted April 5, 2024

DOI: 10.3892/br.2024.1781

**Abstract.** Artificial Intelligence (AI) has become a topic of interest that is frequently debated in all research fields. The medical field is no exception, where several unanswered questions remain. When and how this field can benefit from AI support in daily routines are the most frequently asked questions. The present review aims to present the types of neural networks (NNs) available for development, discussing their advantages, disadvantages and how they can be applied practically. In addition, the present review summarizes how NNs (combined with various other features) have already been applied in studies in the ear nose throat research field, from assisting diagnosis to treatment management. Although the answer to this question regarding AI remains elusive, understanding the basics and types of applicable NNs can lead to future studies possibly using more than one type of NN. This approach may bypass the actual limitations in accuracy and relevance of information generated by AI. The proposed studies, the majority of which used convolutional NNs, obtained accuracies varying 70-98%, with a number of studies having the AI trained on a limited number of cases (<100 patients). The lack of standardization in AI protocols for research negatively affects data homogeneity and transparency of databases.

## Contents

1. Introduction
2. Methods
3. Concepts
4. Types of neural network models
5. AI and ENT
6. Discussions
7. Limitations
8. Conclusions

## 1. Introduction

Over the past number of decades, various attempts have been made to define 'Artificial Intelligence' (AI) to capture its essence. At present, the most accepted definition appears to have formulated by Elaine Rich, as follows: 'Artificial Intelligence is the study of how to make computers do things at which, at the moment people are better' (1). For AI, the idea behind these applications resides in an intelligent agent that can receive percepts and perform actions. For a program to be able to 'understand' and 'react rationally', it must learn. Therefore, 'machine learning' (ML) is not a synonym for AI but is an important sub-field of AI (1,2). After the ML process, neural learning has been developed, which allows for the generation of multiple non-linear combinations with the data given, by creating layers of artificial neurons. This enables the software to mimic human reasoning, by passing the data inputted through multiple layers of artificial neurons to obtain the optimal output. A combination of ML with neural learning defines the 'deep learning' process (Fig. 1) (3).

In healthcare systems, AI has already been applied in a number of subdomains as a tool to increase the insights for classification, prediction, disease diagnosis and patient management.

*Correspondence to:* Dr Mihai Dumitru, Department of ENT, 'Carol Davila' University of Medicine and Pharmacy, 1 Sarandy Frosa, 050751 Bucharest, Romania  
E-mail: orldumitrumihai@yahoo.com

**Key words:** neural networks, otorhinolaryngology, artificial intelligence, head and neck cancer

Examples of procedures AI programs are currently being used for in daily clinical practice include skin lesion identification, fundus retinography for the detection of diabetic retinopathy, radiologic diagnosis based on Rx scans and identification of malignant cell populations through digital microscopy (4,5).

Neural networks (NNs) and deep learning are currently revolutionizing medical imaging and diagnosis due to the complementary manner in which AI is being used alongside manual procedures. From image acquisition and reconstruction to enhancement and quantifications (using radiomics), all outcomes are possible under the supervision of specialists. The continuing evolution of digital hardware and software have made the evolution of novel algorithms with higher processing power possible, which if used correctly, can increase both the efficiency and efficacy of medical actions (6).

Otolaryngology [ear, nose and throat (ENT)] is a vast specialist field that focuses on infectious, inflammatory, immunological and neoplastic disorders of the head and neck, along with their medical and surgical therapy (7). According to data from a previous study, non-communicable diseases (including ENT disorders) accounted for 60% of the global mortality cases in 2008. In addition, the most common disability internationally was disabling hearing impairment (8). ENT diseases treated at health facilities constitute 20-50% of the cases, with an incidence of diagnostic error varying 0.7-15% in developed countries (9) ENT surgical sub-specialties are also evolving, where risk awareness is increasing for specialists becoming deskilled in the case of certain emergencies (10). Apart from surgical interventions, other diagnostic aspects are also of importance, such as rhinology, phoniatrics, laryngology and audiology (7). Due to the vastness of the field and the incidence of diseases, there is now an opportunity to develop AI tools to evaluate the performance of newly developed AI programs (11,12). The key challenges in ENT that AI can address (or improve) are as follows: i) Assistance in the early detection of certain diseases (such as throat cancer); ii) assistance in reducing diagnostic errors; iii) fast screening to optimize treatment (such as in genetic and molecular profiling for tumor-specific targeted therapies); and iv) image enhancements for optimizing diagnostic accuracy.

The objective of the present review is to familiarize the reader with the various NNs to understand their advantages and disadvantages, in addition to understanding why certain networks are preferred in certain cases [such as why convolutional NNs (CNNs) are preferred in image processing]. Furthermore, the present review would like to emphasize that the current state of AI technology advancements actually permit the development of assistance AI programs that will likely enhance the performance of medical actions in the future.

The present article will address the following two main topics: i) The types of NNs (preceded by an explanation of key concepts intended to facilitate understanding); and ii) the application of NNs in the ENT field (covering NNs applied in head and neck cancer, treatment management, image enhancement, voice analysis, hearing prediction and head and neck surgery).

## 2. Methods

The present review used the PubMed database. The article searches included key words ‘artificial intelligence’, ‘neural

networks’, ‘deep learning’, ‘otolaryngology’ or ‘ENT’. The Boolean operators used were Neural Networks AND ENT; Artificial Intelligence AND ENT; (Artificial Intelligence OR Neural Networks OR Deep Learning OR Machine Learning) AND Otolaryngology. A number of references were obtained using words from ENT subfields (such as osteomeatal complex and laryngoscopy) to identify more accurate studies after outlining the subsections of this article: Artificial Intelligence AND laryngoscopy. Due to the novelty of the topic, the number of citations was not included in any of the inclusion or exclusion criteria. The search was conducted in 2023, throughout the year. From the 1,803 articles returned by PubMed, 59 articles were selected using the following inclusion criteria: i) Must have faithful information (complete articles, with information in accordance with reality and achievable results); ii) must be published later than 2011; iii) can be reviewed or an original study; and iv) must include clinical information. Articles with misleading titles or key words were excluded. In total, seven articles were excluded after full text review due to the lack of sufficient clinical information. A few articles that were used in the present review were published before 2015 and were therefore not counted in the aforementioned 59 papers.

## 3. Concepts

ML is a subfield of AI that has changed its goal from achieving autonomous AI to using it to solve practical problems in the 1990s. There are four types of ML algorithms: Supervised, semi-supervised, unsupervised and reinforcement. Functions of a ML system can vary from descriptive (using data to explain), predictive (using data to predict) or prescriptive (using data to suggest) (13).

Deep learning is an advanced type of ML that is capable of establishing nested hierarchical models to process and analyze data through multilayer NNs (14). Datasets are one of the most important pieces when building an AI program. They must be clear, sorted and annotated by specialists in the field to facilitate the ML process. When building a dataset, it must be noted that a model cannot handle multiple tasks. Under most circumstances, an NN that can only perform one task with optimal performance is present, but it is suboptimal at performing other tasks. Therefore, the dataset should only contain relevant information that can facilitate the training process. Although the entire process of annotation appears to be a burden, the currently available learning algorithms require vast amounts of labeled information for training. The annotation should be performed by a team to boost the precision of the results. (15,16). If however, some data are missing, options to ‘recover’ them do exist, such as by using clustering algorithms to sort the data by creating clusters of information with similar attributes before filling in the blanks with similar data from the cluster (17). In ML, node, hidden layer and NNs are commonly used terminologies. A node is a structure made up from data and is capable of storing any type of data. It typically has a link to another node, where each node can have  $\geq 0$  derivative nodes (14). A hidden layer is a layer of artificial neurons between the input and output layers, where conditions and functions are activated by certain input data. Finally, NNs represent a subset of ML that forms the core of deep learning. They are structured to mimic the

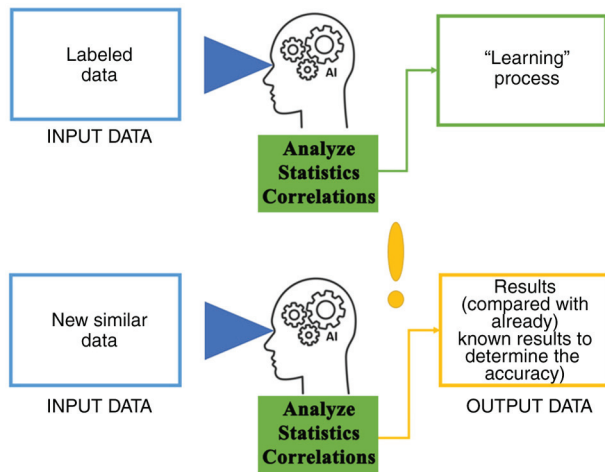


Figure 1. Learning process. First, the labeled data is processed by the AI program, which searches for patterns of recognition. When new similar data is given to the AI, it will provide a result based on the recognition patterns developed during the training process. AI, artificial intelligence.

human brain, specifically interactions among organic neurons. In addition, this algorithm can be used to create an adaptive network that is capable of continuous auto-optimization by learning from their 'failures'. The NNs learn by analyzing the various labeled examples provided during the training phase before using key parameters to determine the quality of the input information and assess what are required to produce the 'correct' output (18).

In medical imaging, segmentation represents the process of dividing a digital image into subgroups (image segments) based on similar characteristics (texture, color, contrast and brightness) (19). There are several techniques available for image segmentation, namely edge-based, threshold-based, region-based, cluster-based and watershed. Each of the aforementioned holds an advantage depending on the common feature of interest and can all be applied in medical imaging (20). The dice similarity coefficient is a statistical instrument used to measure the degree of overlap by two areas. It can be used to verify the accuracy of AI programs that are trained to interpret images (21).

#### 4. Types of NN models

**General characteristics.** There are nine types of NN models, which are defined by numerous parameters, with the most important including the input/output capabilities of the used nodes, the depth of the network, the number of hidden layers and the number of connections between the artificial neurons. It is paramount to have knowledge of the different types of NNs, since each of them has their advantages and disadvantages that must be tailored to each situation (14).

**Perceptron.** Perceptron is a network that is constructed by connecting the nodes. The signal obtained is a simple linear regression. This means that the relationship between two quantitative variables is estimated (22). An example of the application of this NN is the estimation of how strong the relationship is between (the variables) transaminases and the degree of liver fibrosis damage (23). Advantages of using

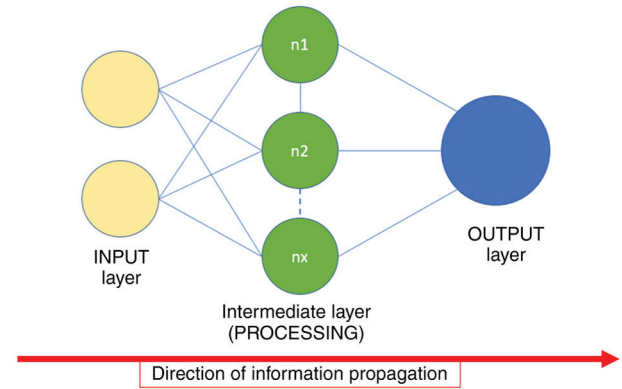


Figure 2. Schematic of a simple neural network. The most basic neural network structure is shown, with three layers, always having the same direction of information propagation.

perceptron is that one can implement logic operators (such as 'AND', 'OR' and 'NAND'), which improves the ability to anticipate the various outcomes. Disadvantages of perceptron is that it can only take two values (namely 0 and 1) and can categorize vectors that can be separated linearly.

**Multilayer perceptron.** This NN is a fully convolutional network that takes in a set of inputs, using it to create a collection of outputs. It is a network that is ideal for machine translation and 'complex classification'. Advantages of multilayer perceptron are that it can be used for deep learning and learning non-linear models. It can also handle a large quantity of input data. Disadvantages of this NN are that due to the vast networks among the nodes; this NN is difficult to design and manage, which slows down the learning process. In addition, even though it can handle large quantities of input data, if the variability is too high the output results will be poor. This disadvantage has been validated using a program that is incapable of recognizing rotated versions of the handwritten digits but was trained on canonical digits (24).

**Feed-forward NN.** This is one of the most basic types of NNs. The input information only goes in one direction until it arrives at the output node (Fig. 2). Feed-forward NNs can have hidden layers of artificial neurons, but it is not mandatory. An advantage of this NN is the fast processing time due to the one-way propagation of information using a pipeline that is easy to design. In addition, it can be used for simple forms of classification, such as face and speech recognition. A disadvantage of a feed-forward NN is that it is not suitable for deep learning. Although possible, the necessity of additional variables to optimize the process and the data propagation process form major obstacles for deep learning using this NN (25).

**CNNs.** In the majority of circumstances, NNs are networks that form a two-dimensional array. By contrast, CNNs are typically built in three dimension. This structure is due to the requirement of the input information being vectorized. In imaging, using a two-dimensional array would destroy the structural information of the medical image. Therefore, each neuron will be requested to process a small portion of the image, such that as the deeper layers of the network are approached, the pixels will be recognized in shapes. A CNN typically has three types of layers, namely a convolutional layer, a pooling layer (with which you can reduce the number

of parameters needed to process the image by reducing the size of the image before extracting the features) and fully connected layers (which process the output). If the image is resized (such as by maintaining the proportions), it would be almost impossible to spot the imaging details using the naked human eye. However, using CNN, the same regions of interest can be detected. In addition, CCNs can be trained with high accuracy to use lower numbers of parameters (20,26). This is the reason for CNNs being so commonly used in training AI for interpreting medical images, with particular interest in oncologic radiology and pathology, for macroscopic and microscopic tumor detection (4). An advantage of CNNs is their powerful deep learning network that does not require high numbers of learning parameters. A disadvantage of CNNs is that they are difficult to design and although they require fewer numbers of operational parameters, they are typically slow (26).

**Radial basis functional (RBF) NN.** This NN is preferable for approximation and classification. In addition to the input cells, each of which must contain a vector variable, the network also consists of a layer of hidden RBF neurons and a single output neuron for each category. The target is to train the network to store prototype examples. When a new input is inserted in the program, the data will be compared with each prototype in every neuron, before they are classified based on the highest degree of similarity with one of the prototypes. An advantage of this NN is that it can be used to create various adaptable control systems. However, due to the data variation and gradient issue, the training process can be challenging (27). This model has been used in the medical field to create an NN that can classify six different types of heartbeats in electrocardiograms (ECG), with an average sensitivity of 96.25% and sensibility of 99.10% (28).

**Recurrent NN (RNN).** An RNN is used to create a cyclic connection between the artificial nodes. This allows the output nodes to influence the subsequent input nodes. RNNs can improve their results (predictions) by reusing the early or later node activations in the sequence (29). The first layer will always be a feed forward NN. The next layers would then consist of recurrent layers. If the results are 'wrong', small changes would be made in those neurons to gradually enhance the accuracy of prediction during the backpropagation. This type of NN is best used for translations, text processing applications (such as grammar checks, auto-suggest), image-tagging tasks, image-captioning and even for pixel-wise classification (which can increase the CNN accuracy). An advantage of an RNN is that each sample from the data can be assumed to be dependent on its predecessor. This way, the NN will continue to improve the predictions continuously. In addition, combining a CNN with an RNN can extend the pixel power effectiveness (30). However, a disadvantage of an RNN is that training the nets of the RNNs is a difficult task. Additionally, fading gradients (where information from some neurons 'fades', obstructing improvements in those nets) and the exploding gradients (where the gradient expands, leading to an unacceptably large weight update in the whole NN) remain issues that must be addressed (31).

**Long short-term memory (LSTM).** LSTM is an enhanced version of an RNN. It is one of the most advanced NNs used for learning tasks (such as predictions and sequence learning

tasks) (32). This neural network includes a 'memory cell' that stores information in its memory for long periods of time. To be able to control the information entering the memory and information that is forgotten, a set of three 'gates' must be coded. The input gates will then decide on which information from the last entry will be kept as memory, whereas the output gates will control which data will pass onto the next layer of the NN. The forget gates will also select the rate at which data is deleted from the memory cell (33). An advantage of LSTM is the wide range of parameters to choose from. A disadvantage of LSTM is the necessity for significant resources and time for training this NN. This type of network has been used for detecting anomalies in ECG time signals, by developing a predictive model for healthy ECG signals (34). It yielded promising results to support the application of using LSTM for ECG interpretation (34).

**Sequence to sequence models.** These NNs are two joined RNNs. This type of NN has an encoder (which processes the input data) and a decoder (which processes the output), both of which operate simultaneously. In addition, they can either use the same parameter or different ones. One of the most important things to note is that this NN is particularly applicable if the length of the input and output data are of the same length. An advantage of this model is that its architecture allows for the processing of several inputs and outputs at the same time. A disadvantage of this model is that little memory is present. This neural network particularly excels in speech recognition (35).

**Modular NN.** This NN is built with several distinct networks, each having a specific task. Although all the NNs are within the same program, little interaction or communication exists among them. Their contribution to the outcome is processed individually. The advantage of this NN is its efficient and independent training. By contrast, a disadvantage of a modular NN is that there are issues with moving targets. This type of network may prove to be the future if the accuracy can be optimized and sufficient data are input. A recent study devised a program for lung disease classification based on multiple objective feature optimization for chest X-ray images (36).

The potential for applications of NN are vast, where the list of advantages and disadvantages are shown to facilitate selection depending on the situation and the desired outcomes. However, a direct comparison among them would prove futile, since each of the nine aforementioned types of NNs excels under different circumstances.

## 5. AI and ENT

At present, AI applications have been established for three main functions clinically (Fig. 3). The most important of these is in medical imaging, where it is being applied for aiding the workflow in image interpretation, classification and analysis, especially for radiologists (37). AI can potentially be used to optimize image quality through noise/artifact reduction whilst reducing the investigation time through image generation. The development of ML and deep learning algorithms, combined with the processing power tools that are currently available, is a field that is garnering interest in the medical setting (38,39).



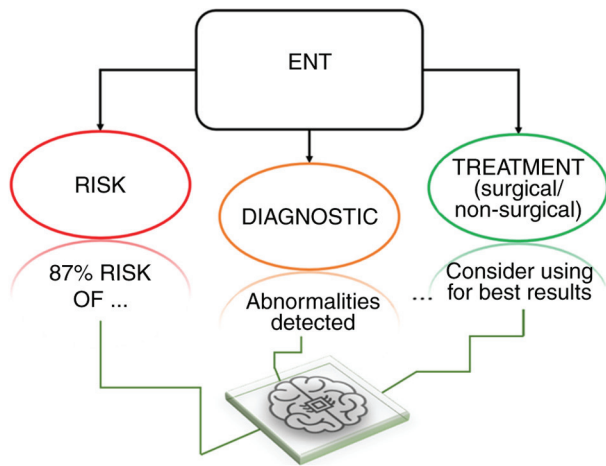


Figure 3. Ways in which artificial intelligence can help the ENT physician. Summaries of the three main ways in which artificial intelligence can be applied to assist the ENT physician. ENT, ear, nose and throat.

*AI in head and neck cancer (HNC).* Over the past decade, the high demand for oncological specialties and advances in the field in turn increased the demand for developing ML tools. Specifically, they can be trained to detect certain oncological diseases (such as early stage lesions) and stratify patients for treatment (40).

Previously, two studies have reported that HNC is the sixth most common neoplasm, with ~520,000 new cases and 350,00 cases of mortality annually (41,42). The American Joint Committee on Cancer defined HNC as malignancies originating from the salivary glands, oral cavity, paranasal sinuses, pharynx and larynx. HNC is frequently first detected already at advanced stages, whereas diagnostic errors can markedly worsen the outcome (43,44). The most important function of AI for the radiologist is tumor segmentation and risk stratification based on the images. A previous systematic review found that the majority of studies included used highly heterogeneous classification and segmentation techniques, with unicentric datasets (consisting of 40-270 images) and accuracies varying between 79 and 100%. However, the quality of evidence is low, due to the small unicentric datasets and overestimated accuracy rates (45). Another study has previously explored the use of deep learning in head and neck tumor multi-omics diagnosis, where multiple sources of information, such as radiomics, pathomics and proteomics, were incorporated into AI programs to analyze the tumors. The purpose was to integrate all the collected data and generate protocols for personalized diagnosis and treatment (46).

In particular, a study even attempted to create an NN for multi-organ segmentation on head and neck CT images, also using a self-channel and spatial attention NN. This NN functions by incorporating features extracted from the CT images. In turn, the 'useful' features adaptively force the network to emphasize them, whilst the 'irrelevant' features are ignored. Although the results from this study appear promising, the dataset used consisted of only 48 patients, where the accuracy rates may have been overestimated (45,47).

The use of deep learning algorithms in medical imaging has also been applied in MRI imaging for head a neck tumor segmentation. A previous study assessed this possibility in patients with HNC who underwent multi-parametric MRI. To train the CNN, manually defined tumor volumes from seven different MRI contrasts were used for segmentation. This program obtained a dice similarity coefficient of 65% when detecting the gross tumor volume and 58% when detecting the malignant lymph nodules (48). In addition to CT and MRI, several studies have also assessed AI use for analyzing PET/CT images for HNC prognosis. By automatically generating the gross tumor volume from both the PET and CT images, the predictive performance can be increased (49). If the patient has only one of either PET or CT available, then a channel dropout technique can be used, where the network would be prevented from developing co-adaptations whilst also enhancing the performance of single modality images (CT or PET) and/or images from both modalities (PET/CT) (50).

Another application of AI that is currently under investigation is the real-time detection of lesions. A previous study developed an NN that can potentially identify and classify nasal polyps during endoscopy using video-frames. This program uses a light binary pattern to identify the lesions based on color change, before feeding them through a CNN architecture for classification workflow. According to the results, this program obtained a maximal accuracy of 98.3% (51).

*AI in treatment management.* In a previous cohort study of 33,526 patients, three ML models were used for recommending chemoradiation treatment for patients with intermediate risk of head and neck squamous cell carcinoma. The AI programs were trained and validated using the National Cancer Database. Although the results are not final, the available findings suggested that these ML models have promising potential for accurately selecting patients most in need of chemoradiotherapy (52).

*AI for image enhancement.* Cone-beam CT (CBCT) was introduced to overcome the limitations of two-dimensional imaging techniques. Although it is superior in terms of diagnostic accuracy, CBCT increases radiation exposure and is less cost-effective. However, low-dose (LD) CBCT is possible to mitigate this in exchange for sacrificing image quality (53). A previous study used CNNs to train an algorithm for enhancing the quality of LD-CBCT used for head and neck radiotherapy. The enhanced images generated by the AI exhibited superior soft tissue contrast and less artifact distortion (54). In another previous study, which aimed to increase the accuracy of radiation therapy delivery, a deep NN was trained using ultra LD-CBCT images. This new proposed NN (modified U-Net++) was found to outperform the previous networks used (U-Net and wide U-Net) in terms of image contrast, accuracy of soft tissue regions and mean absolute error (55).

*Voice analysis.* Any alterations in the geometry of the vocal cords will lead to dysphonia or changes of the voice. These types of diseases are mainly diagnosed auditorily and/or through laryngoscopy, both of which are unfortunately based on a subjective judgement by the clinicians (56,57). A previous study proposed the use of a CNN-transfer learning technique

for diagnosing vocal cord lesions in patients who underwent laryngoscopy. Based on the morphological features of normal and different types of pathological vocal cords, the AI was able to diagnose normal vocal cords, polyps and carcinoma with higher sensitivity compared with clinician judgement. By contrast, the keratinization lesions were diagnosed by clinicians with ~2X higher sensitivity compared with the AI (58). Therefore, although availability of an image does indeed provide important visual information (57), laryngoscopy alone is not sufficient as a screening method. Voice analysis, on the other hand, requires the patient to only speak a vowel. Using speech analysis in phonetics and Mel-frequency cepstral coefficients from the /a:/vowel, a trained AI can be used to potentially detect the voice changes due to laryngeal cancer. A study previously assessed this hypothesis, by using and comparing six different ML algorithms. The results revealed that the one-dimensional CNN yielded the highest accuracy (85%), sensitivity (78%) and specificity (93%) for detecting laryngeal cancer, compared with the accuracy of 69.9% and a sensitivity of only 44% manually yielded by two laryngologists (59).

*Hearing prediction.* Hearing is an important sense, the loss of which is associated with high rates of morbidity (60). Since there are various causes of hearing loss, such as aging, surgery, accidents and sound trauma, a number of studies have been aiming to build an AI model capable of predicting hearing prognosis. In one of the studies, the possibility of building an NN for differentiating chronic otitis media (COM) with and without cholesteatoma was investigated. The dataset was formed using CT images from 200 patients (100 patients with COM with cholesteatoma). This proposed model correctly predicted 2,952 out of 3,093 images, thus achieving an accuracy of 95.4% (61). Another previous study compared the performances of three types of multilayer perceptron NN models and a k-nearest neighbor model by applying them for determining the prognosis of patients who underwent tympanoplasty after chronic suppurative otitis media. The k-nearest neighbor is a non-parametric supervised learning classifier. The optimal configuration of the three multilayer perceptron was found to achieve an accuracy of 84%, compared with the 75.8% achieved by the k-NN model. This obtained accuracy suggests a complex and non-linear relationship between clinical and surgical variables, meaning that results generated by predictive models should be assessed with caution to avoid pitfalls (62).

*Head and neck surgery assistant.* ENT surgery can be supported through image analysis. Pre-procedural imaging, mostly CT, is indispensable for deciding the optimal approach of intervention. Studies have been performed using CNN to train the AI to detect regions of interest in the CT images (63). Prior to sinonasal surgery, CT images can be evaluated to identify any anatomic variations, determine the extent of the disease and then devise the most appropriate surgical plan (64). In addition, three studies each presented three different programs consisting of CNNs. One was used to detect osteomeatal complex occlusion, one was used to predict the location of the anterior ethmoid artery and the other was used to identify a concha bullosa at the level of the

osteomeatal complex. However, all three of the models had the same underlying purpose, which was to reduce the incidence of intraoperative complications (65-68). The CNN developed to detect the osteomeatal complex occlusion specifically incorporated individual two-dimensional coronal CT images. This model was trained using 1.28 million images, which obtained a classification accuracy of 85% (66). To predict the location of the anterior ethmoid artery, the CNN algorithm was trained using 675 CT images, which acquired an accuracy of 82.7%. With this result, this previous study concluded that this form of identification of clinically important structures before functional endoscopic sinus surgery can reduce the risk of cutting the ethmoid artery (67). The identification of middle turbinate pneumatization (concha bullosa) was performed using a CNN trained by incorporating high-resolution CT images, obtaining an accuracy of 81% (68).

The primary objective of surgical treatment for chronic otitis media is the removal of any pathological changes and the restoration of middle ear anatomy. However, this form of surgical intervention has its risks, meaning that it should only be performed when there is a postoperative hearing improvement procedure (61). Using AI, associating image features and parameters with hearing improvement prognosis is becoming possible.

Robotic surgery in ENT is a field that is currently undergoing rapid evolution after it was first introduced. Research into this technique is currently ongoing, with the most important aspects including image guidance and augmented reality, microscopic imaging, possibility of telesurgery and the heavily in-demand haptic feedback. However, the most promising prospect is the development of autonomous surgical robots powered by AI instead of surgeon-controlled robots (69).

Periprocedural anaphylaxis presents a serious threat for the patient. A previous study suggested the theoretical possibility of creating an AI program to warn the medical staff about the risk of anaphylaxis in certain patients. However, due to the small number of cases, a multicentric study must be conducted to verify such possibilities (70).

## 6. Discussions

Numerous studies have documented various AI programs, each designed to perform a specific task. The present review only focuses on a selective few, which use novel features to enhance the results (Fig. 4).

According to 'Designing Machine Learning Systems' by Huyen (71), creating an AI program requires several steps, each of which is described in detail in this book (71). Although there are various additional sub-steps, the following seven main steps can be recognized in all studies addressed: i) Gathering data; ii) preparing the data (generally referring to the randomization and labeling process); iii) choosing the right NN model (or generating more for comparison); iv) training (which creates patterns using the data so that it can make decisions); v) evaluation (to check the performance unseen data is used, where results are measured in accuracy, specificity, sensibility or even dice similarity coefficient); vi) hyperparameter tuning (refers to finding the right training parameter values to generate the maximum accuracy); and vii) making predictions.

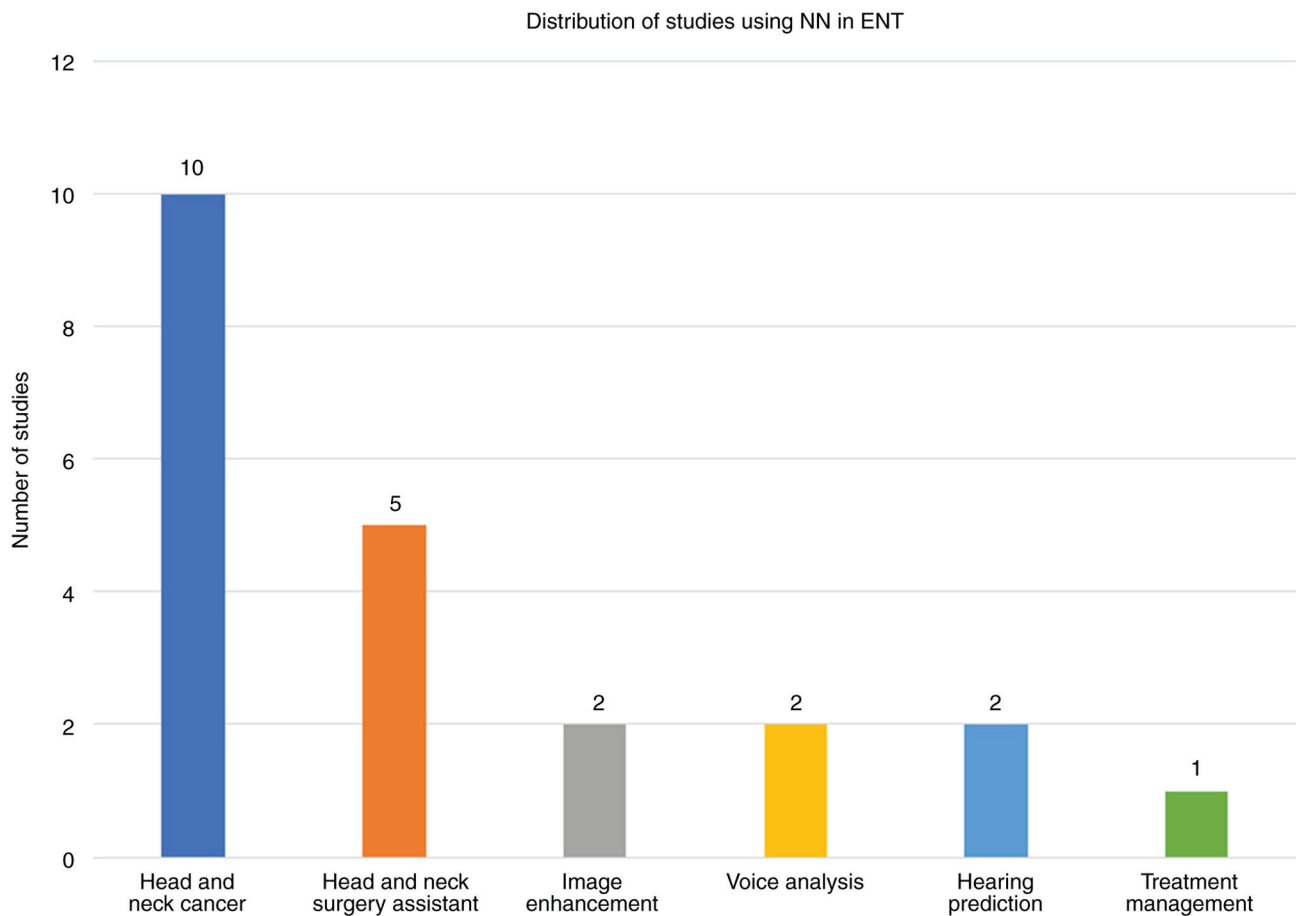


Figure 4. Relevant studies used in the present article. Examples of artificial intelligence used in ear, nose and throat application are shown. NN, neural network; ENT, ear, nose and throat.

The implementation of AI during the daily clinical routine is met with a number of unresolved issues regarding process and data standardization, transparency and concerns regarding sharing and privacy of the used data (72). In addition, some differences remain among ML models used for research and those used for daily use in any domain (Table I) (71).

Data management is one of the most important steps in AI creation. The majority of the NNs used are supervised, meaning that AI is taught using given examples (14). The operator must provide the ML algorithm with a known dataset, including the inputs and the desired outputs. This means that the operator must know the correct answers and the algorithm is trained to find a pattern in the data using what it learns to make various predictions (73). This form of training requires specific parameter adjustments. The dataset is generally divided into two sets, namely the training set and the testing set, with the most common division being 80% for training and 20% for testing (65). The training process uses the input data and the output data (labeled 'da-ta') to learn patterns. The testing set then gives the program only the input data and compares the result yielded by the NN with the known output data to determine its performance (1). Therefore, to incorporate AI into clinical practice, the complexities surrounding data management, refining the training process and the correlation between research models and real-world applications must first be unraveled.

Depending on the dataset and objective, the labeling process is different. When the AI is used to make predictions based on numerical/valued data, a 'csv.' or an 'xlsx' file would be sufficient. The labeling will then be the known outcome value (23,52). By contrast, a more difficult labeling process exists for CNN algorithms, because it learns based on image characteristics, such that each image must be labeled independently. Although CNNs can learn efficiently even using small datasets (26), labeling must be executed carefully in specific software, especially if the dice coefficient is used to determine the accuracy (48,68). Due to the diversity of the imagistic machines (CT and MRI), in multicentre studies the images may not have the same dimensions, which poses a problem in the learning process. To avoid this, an extra step is typically required, which is the interpolation/extrapolation of the base resolution based on the highest/lowest resolution available (48). In one previous study, an NN was created to analyze the voice of the patient. In this case, the vowel 'a' was used as the input, but the output was the anatomopathological result (presence or absence of malignancy). In addition, these results were then compared with those from two physicians, who also heard the vowel from the patients (59).

Various ML models also have unsupervised learning algorithms that analyze unlabeled data to identify patterns. A disadvantage of these algorithms is that the program is

Table I. Differences between ML characteristics used in research vs. daily-use models.

Category	Research	Production
Main focus	The best performance (the use of latest discoveries)	Based the requirements of the task, using the most suited algorithms
Computational priority	Fastest training with greatest throughput	Best results with lowest latency
Data set	Static	Dynamic (constantly changing)
Fairness	May not be a focus	Must be considered
Interpretability	May not be a focus	Must be considered

Adapted from Ref (71).

left to interpret large datasets without instructions from a human operator (13). A previous study used a form of an unsupervised ML model to virtually reconstruct bilateral maxillary defects to optimize the virtual surgical planning workflows. To determine the accuracy, the 95th percentile Hausdorff distance, volumetric dice coefficient, anterior nasal spine deviation and midfacial projection deviation were used. Although the results are promising in terms of virtually reconstructing the missing maxilla, the overall dice coefficient only reached a maximum of 68% (74). In another previous study, a deep-learning-based automatic facial bone segmentation model using a CNN (named U-Net) was applied, and its accuracy was also measured using the dice coefficient. Automatic segmentation can increase the workflow and provide more information regarding where the lesions are precisely. The dice coefficient of the maxilla in this case was found to be ~90% (75). A previous online survey regarding the awareness and utility of AI found that the majority of participants believe in the future of this type of program. However, concerns were raised when the topic of unsupervised ML was raised, since it is not yet possible to use this to obtain faithful results (72). In addition, the labeling process of AI model building, especially using CT images, is one of the most difficult and tedious stages. Since it requires particular focus and attention, labeling is typically time-consuming, because of the huge quantity of data requiring processing in the form of images. Unsupervised learning can potentially be honed to bypass this step, but at the cost of lower control over the entire ML process and accuracy (76). Furthermore, the necessity of more than one individual to process or label the data can be considered as another limitation due to the risk of inter-rater variability, since different individuals can interpret and label the data differently. By contrast, delegating a single individual for the entire labeling process runs the risk of bias, since the labeling will become dependent on the perspective of the specific individual.

A type of ML algorithm that is noteworthy is decision trees. Decision trees are supervised learning algorithms that have hierarchical ‘tree structures’, making them ideal for classification and regression tasks. The main three parts of such structures are the root node, branches and leaf nodes. Starting from the root information, the data are then diverted into certain branches of the tree to classify each label (77).

Such decision trees have been used for classifying lipomas from liposarcoma using MRI images. To build the classification tree, six attributes from MRI images were used, namely size, location, depth, shape pattern of the tumor, the intermingled muscle fibers, sex and age of the patients. This tree model was able to distinguish lipoma from well-differentiated lipo-sarcoma with a sensitivity of 80%, specificity of 94.6% and an overall accuracy of 91.7% (78). However, a small change in the training data can drastically alter the predictions (79). In addition, overwhelming data trees that cannot generalize the training data efficiently can be created. A solution to this problem is pruning, specifically by reducing the aberrant branches and leaf nodes (77,80). A retrospective cross-sectional study previously evaluated the performance of an ML model in analyzing major prognostic factors for hearing outcomes after mastoidectomy with tympanoplasty. A multilayer perceptron and a light gradient boosting machine (LightGBM; which is based on decision trees) were trained using the same data. Although the results were found to be similar, the LightGBM yielded slightly superior results (81 vs. 79%) (81). Another previous study used MRI images to train a regression tree for differentiating lacrimal gland tumors into four different types. The predictive accuracy was 86.1%. However, this previous study did not correspondingly train NNs for comparison (82). Although the strengths of decision trees lie in their flexibility and transparency, sensitivity to small changes in the training data can pose serious problems in clinical application, especially in cases where the variability found among patients with the same disease can be high.

Although discoveries are continuously made in this field, skepticism remains with regard to obtaining a fully capable AI for facilitating patient decision support in conjunction with clinicians (83).

Different NN architectures can provide different results and accuracies using the same dataset. A previous study proposed a support vector machine to increase the diagnostic accuracy of obstructive sleep apnea compared with conventional logistic regression models (84).

The performance of AI programs in relation to the performance of expert clinicians has been previously analyzed in a systematic review. This previous study discovered various key findings that are relevant to the studies included in the present review. Limited studies were actually tested in a real-world



clinical environment, where the testing was only performed by one-four specialists. In addition, the availability of datasets and code is limited, meaning that they cannot be reproduced or tested in other centers (85). However, improving the performance of each of the AI programs should be a top priority. In particular, the optimization strategies must address the limitations that persist, specifically data sizes (which should contain thousands of examples), populations from different areas that can cause bias in the dataset and the most importantly, lack of multicenter validation.

AI programs should also be developed and approached under the jurisdiction of ethical laws of research, centering around benevolence and non-maleficence. Furthermore, it is important to have a multidisciplinary team, since diverse perspectives can address different ethical challenges more efficiently. Regarding transparency and privacy, there are several privacy-enhancing techniques that can be used to anonymize data. Regulatory processes are required for real-world implementation and should start from the standardization of imaging, annotations and metrics used to quantify the performance (86).

## 7. Limitations

The present review is not a meta-analysis or a systematic review, which is a limitation. Due to the lack of standardization regarding the currently available AI studies, reviewing data homogeneity is difficult. Differences appear even in studies sharing a common objective regarding ENT disease. The most common data found to be missing tend to be performance metrics. Not all studies have all area under the curve, sensitivity, specificity and accuracy parameters. When working with CNNs and images, although dice similarity coefficients are more faithful parameters, they cannot be directly compared with accuracy, which is commonly used in similar studies that do not report dice similarity coefficients.

Selecting studies with stricter criteria would create an insufficient database and would also eliminate potentially important studies from the meta-analysis. However, having articles that allow future studies to improve and include all relevant metrics data may prove helpful but will not replace the need for standardization. The methods used for data collection (which are usually provided within the same study center and/or images from a single type of imaging device) and the sample size represent a source of bias when trying to address the issue of generalizability. Having standard measurement tools and instruments used in all the studies that included AI programs would overcome such limits to address the issue of generalization.

## 8. Conclusions

To conclude, physicians can benefit from AI-assisted tools for daily practice to increase the workflow and treatment success rates, especially ENT clinicians and surgeons. Knowing the different types of NNs, their advantages and disadvantages can open a path of exploring methods for exploiting such modular networks. The dice similarity coefficient should be used as the default standard in image-processing AI programs, since it most accurately reflects the actual performance of the

programs. The high interest in this field and the promising data acquired thus far are indicators of the acceptance of AI in daily clinical practice. The present review documented the potential difficulties that may arise during the development of an AI program, the multitude of methods that can be supported and improved in the ENT field and the subtle call for multicenter collaboration so that optimal results can be obtained in the future.

Regarding the future research directions, a direct comparison between the same subtypes of NNs, such as those between the existing CNN architectures, applied on a multicenter database, may be the next step. Taken together, at present the answer to the question ‘when can we use AI in daily practice?’ remains elusive, but understanding the basics and the types of NNs currently available can lead to future studies that will possibly use more than one type of NN. This approach may bypass the actual limitations in terms of accuracy and relevance of information generated by the AI.

## Acknowledgements

Publication of this article was supported by the ‘Carol Davila’ University of Medicine and Pharmacy (Bucharest, Romania), through the institutional program ‘Publish not Perish’.

## Funding

No funding was received.

## Availability of data and materials

Not applicable.

## Authors' contributions

IAT and MD contributed substantially to the conception and design of the study, the acquisition, analysis and interpretation of the data from the literature for inclusion in the review, and were involved in the drafting of the manuscript. DV and AC contributed substantially to the acquisition, analysis and interpretation of the data from the literature for inclusion in the review, and were involved in the drafting of the manuscript. MG, FM, AM, CS and AN contributed substantially to the acquisition of the data from the literature for inclusion in the review, and were involved in the critical revisions of the manuscript for important intellectual content. Data authentication is not applicable. All authors have read and agreed to the published version of the manuscript.

## Ethics approval and consent to participate

Not applicable.

## Patient consent for publication

Not applicable.

## Competing interests

The authors declare that they have no competing interests.

## References

- Ertel W: Introduction to artificial intelligence. 2nd edition. Springer International Publishing, pp2-3, 2017.
- Russell S and Norvig P: Artificial intelligence. A modern approach. 4th edition. Pearson Education Inc., pp2-5, 2021.
- Wang F, Casalino LP and Khullar D: Deep learning in medicine: promise, progress, and challenges. *JAMA Intern Med* 179: 293-294, 2019.
- Martinez-Millana A, Saez-Saez A, Tornero-Costa R, Azzopardi-Muscat N, Traver V and Novillo-Ortiz D: Artificial intelligence and its impact on the domains of universal health coverage, health emergencies and health promotion: An overview of systematic reviews. *Int J Med Inform* 166: 104855, 2022.
- Ahmad Z, Rahim S, Zubair M and Abdul-Ghafar J: Artificial intelligence (AI) in medicine, current applications and future role with special emphasis on its potential and promise in pathology: Present and future impact, obstacles including costs and acceptance among pathologists, practical and philosophical considerations. A comprehensive review. *Diagn Pathol* 16: 24, 2021.
- Streiner DL, Saboury B and Zukotynski KA: Evidence-based artificial intelligence in medical imaging. *PET Clin* 17: 51-55, 2022.
- Noto A, Piras C, Atzori L, Mussap M, Albera A, Albera R, Casani AP, Capobianco S and Fanos V: Metabolomics in otorhinolaryngology. *Front Mol Biosci* 9: 934311, 2022.
- Ta NH: ENT in the context of global health. *Ann R Coll Surg Engl* 101: 93-96, 2019.
- Lukama L, Aldous C, Michelo C and Kalinda C: Ear, nose and throat (ENT) disease diagnostic error in low-resource health care: Observations from a hospital-based cross-sectional study. *PLoS One* 18: e0281686, 2023.
- Rouhani MJ: In the face of increasing subspecialisation, how does the specialty ensure that the management of ENT emergencies is timely, appropriate and safe? *J Laryngol Otol* 130: 516-520, 2016.
- Wilson BS, Tucci DL, Moses DA, Chang EF, Young NM, Zeng FG, Lesica NA, Bur AM, Kavookjian H, Mussatto C, *et al*: Harnessing the power of artificial intelligence in otolaryngology and the communication sciences. *J Assoc Res Otolaryngol* 23: 319-349, 2022.
- Lechien JR, Maniaci A, Gengler I, Hans S, Chiesa-Estomba CM and Vaira LA: Validity and reliability of an instrument evaluating the performance of intelligent chatbot: The artificial intelligence performance instrument (AIPI). *Eur Arch Otorhinolaryngol* 281: 2063-2079, 2024.
- Rebala G, Ravi A and Churiwala S: Machine learning definition and basics. In: *An Introduction to Machine Learning*. Springer, Cham, pp1-17, 2019.
- Alzubaidi L, Zhang J, Humaidi AJ, Al-Dujaili A, Duan Y, Al-Shamma O, Santamaría J, Fadhel MA, Al-Amidie M and Farhan L: Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *J Big Data* 8: 53, 2021.
- Wong STC: Is pathology prepared for the adoption of artificial intelligence? *Cancer Cytopathol* 126: 373-375, 2018.
- Benjamin RM, Marion CC and Stanley C: Dealing with multi-dimensional data and the burden of annotation: Easing the burden of annotation. *Am J Pathol* 191: 1709-1716, 2021.
- Charu CA and Chandan KR: Data clustering: Algorithms and applications. CRC Press: Boca Raton, FL, USA, pp2-22, 2018.
- Zhou SZ, Hayit G and Dinggang S: Deep Learning for Medical Image Analysis. The Elsevier and MICCAI Society Book Series, 2017.
- Gonzalez RC and Woods RE: Digital image processing. 2nd edition. Pearson Education, 2004.
- Sharma N and Aggarwal LM: Automated medical image segmentation techniques. *J Med Phys* 35: 3-14, 2010.
- Moore C and Bell D: Dice similarity coefficient. Reference article, Radiopaedia.org (Accessed on 25 Aug 2023) <https://doi.org/10.5334/rID-75056>.
- Stéphane GS, Jérémie D and Carlos Andrés AR: Improving neural network interpretability via rule extraction. In: *Artificial Neural Networks and Machine Learning Part I*. Springer, pp811-813, 2018.
- Thong VD and Quynh BTH: Correlation of serum transaminase levels with liver fibrosis assessed by transient elastography in vietnamese patients with nonalcoholic fatty liver disease. *Int J Gen Med* 14: 1349-1355, 2021.
- Yoon Y, Lee LK and Oh SY: Semi-rotation invariant feature descriptors using Zernike moments for MLP classifier. In: *IJCNN'16 Proceedings of 2016 International joint conference on neural networks*. IEEE, Vancouver, British Columbia, Canada, pp3990-3994, 2016.
- Sazlı MH: A brief review of feed-forward neural networks. *Commun Fac Sci Univ Ank Series A2-A3 Phys Sci Eng* 50: 11-17, 2006.
- Yamashita R, Nishio M, Do RKG and Togashi K: Convolutional neural networks: An overview and application in radiology. *Insights Imaging* 9: 611-629, 2018.
- Young-Sup H and Sung-Yang B: An efficient method to construct a radial basis function neural network classifier. *Neural Netw* 10: 1495-1503, 1997.
- Korürek M and Doğan B: ECG beat classification using particle swarm optimization and radial basis function neural network. *Expert Syst Appl* 37: 7563-7569, 2010.
- DiPietro R and Hager GD: Chapter 21-deep learning: RNNs and LSTM. In: *Handbook of medical image computing and computer assisted intervention*. The Elsevier and MICCAI Society Book Series. Zhou SK, Rueckert D and Fichtinger G (eds). Academic Press, pp503-519, 2020.
- Tyagi AC and Abraham A: Recurrent neural networks concepts and applications. CRC Press Taylor Francis Group, 2023.
- Pascanu R, Mikolov T and Bengio Y: On the difficulty of training recurrent neural networks. *Proceedings of the 30th International Conference on Machine Learning. Proc Mach Learn Res* 28: 1310-1318, 2013.
- Fischer T and Krauss C: Deep learning with long short-term memory networks for financial market predictions. *Eur J Oper Res* 270: 654-669, 2018.
- Xu Y, Mou L, Li G, Chen Y, Peng H and Jin Z: Classifying relations via long short-term memory networks along shortest dependency paths. In: *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*. Lisbon, Portugal, pp1785-1794, 2015.
- Chauhan S and Vig L: Anomaly detection in ECG time signals via deep long short-term memory networks. In: *Proceedings of the 2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*. IEEE, Paris, pp1-7, 2015.
- Enarvi S, Amoia M, Teba MDA, Delaney B, Diehl F, Hahn S, Harris K, McGrath L, Pan Y, Pinto J, *et al*: Generating medical reports from patient-doctor conversations using sequence-to-sequence models. In: *Proceedings of the First Workshop on Natural Language Processing for Medical Conversations*. Association for Computational Linguistics, pp22-30, 2020.
- Sergio VS and Patricia M: A new modular neural network approach with fuzzy response integration for lung disease classification based on multiple objective feature optimization in chest X-ray images. *Expert Syst Appl* 168: 114361, 2021.
- Letourneau-Guillon L, Camirand D, Guilbert F and Forghani R: Artificial intelligence applications for workflow, process optimization and predictive analytics. *Neuroimaging Clin N Am* 30: e1-e15, 2020.
- Erickson BJ, Korfiatis P, Akkus Z and Kline TL: Machine learning for medical imaging. *Radiographics* 37: 505-515, 2017.
- Choy G, Khalilzadeh O, Michalski M, Do S, Samir AE, Pianykh OS, Geis JR, Pandharipande PV, Brink JA and Dreyer KJ: Current applications and future impact of machine learning in radiology. *Radiology* 288: 318-328, 2018.
- Bera K, Schalper KA, Rimm DL, Velcheti V and Madabhushi A: Artificial intelligence in digital pathology-new tools for diagnosis and precision oncology. *Nat Rev Clin Oncol* 16: 703-715, 2019.
- Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D and Bray F: Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *Int J Cancer* 136: E359-E386, 2015.
- Fidler MM, Bray F, Vaccarella S and Soerjomataram I: Assessing global transitions in human development and colorectal cancer incidence. *Int J Cancer* 140: 2709-2715, 2017.
- Cohen N, Fedewa S and Chen AY: Epidemiology and demographics of the head and neck cancer population. *Oral Maxillofac Surg Clin North Am* 30: 381-395, 2018.
- Bassani S, Santonicco N, Eccher A, Scarpa A, Vianini M, Brunelli M, Bisi N, Nocini R, Sacchetto L, Munari E, *et al*: Artificial intelligence in head and neck cancer diagnosis. *J Pathol Inform* 13: 100153, 2022.
- Mahmood H, Shaban M, Indave BI, Santos-Silva AR, Rajpoot N and Khurram SA: Use of artificial intelligence in diagnosis of head and neck precancerous and cancerous lesions: A systematic review. *Oral Oncol* 110: 104885, 2020.

46. Wang X and Li BB: Deep learning in head and neck tumor multimomics diagnosis and analysis: Review of the literature. *Front Genet* 12: 624820, 2021.
47. Gou S, Tong N, Qi S, Yang S, Chin R and Sheng K: Self-channel-and-spatial-attention neural network for automated multi-organ segmentation on head and neck CT images. *Phys Med Biol* 65: 245034, 2020.
48. Bielak L, Wiedenmann N, Berlin A, Nicolay NH, Gunashekar DD, Hägele L, Lottner T, Grosu AL and Böck M: Convolutional neural networks for head and neck tumor segmentation on 7-channel multiparametric MRI: A leave-one-out analysis. *Radiat Oncol* 15: 181, 2020.
49. Wang Y, Lombardo E, Avanzo M, Zschaek S, Weingärtner J, Holzgreve A, Albert NL, Marschner S, Fanetti G, Franchin G, *et al*: Deep learning based time-to-event analysis with PET, CT and joint PET/CT for head and neck cancer prognosis. *Comput Methods Programs Biomed* 222: 106948, 2022.
50. Zhao LM, Zhang H, Kim DD, Ghimire K, Hu R, Kargilis DC, Tang L, Meng S, Chen Q, Liao WH, *et al*: Head and neck tumor segmentation convolutional neural network robust to missing PET/CT modalities using channel dropout. *Phys Med Biol* 68: 095011, 2023.
51. Ay B, Turker C, Emre E, Ay K and Aydin G: Automated classification of nasal polyps in endoscopy video-frames using handcrafted and CNN features. *Comput Biol Med* 147: 105725, 2022.
52. Howard FM, Kochanny S, Koshy M, Spiotto M and Pearson AT: Machine learning-guided adjuvant treatment of head and neck cancer. *JAMA Netw Open* 3: e2025881, 2020.
53. Nomier AS, Gaweesh YSE, Taalab MR and El Sadat SA: Efficacy of low-dose cone beam computed tomography and metal artifact reduction tool for assessment of peri-implant bone defects: An in vitro study. *BMC Oral Health* 22: 615, 2022.
54. Yuan N, Dyer B, Rao S, Chen Q, Benedict S, Shang L, Kang Y, Qi J and Rong Y: Convolutional neural network enhancement of fast-scan low-dose cone-beam CT images for head and neck radiotherapy. *Phys Med Biol* 65: 035003, 2020.
55. Yuan N, Rao S, Chen Q, Sensoy L, Qi J and Rong Y: Head and neck synthetic CT generated from ultra-low-dose cone-beam CT following image gently Protocol using deep neural network. *Med Phys* 49: 3263-3277, 2022.
56. Chediak Coelho Mdo N, Guimarães Vde C, Rodrigues SO, Costa CC and Ramos HVL: Correlation between clinical diagnosis and pathological diagnoses in laryngeal lesions. *J Voice* 30: 595-599, 2016.
57. Grant NN, Holliday MA and Lima R: Use of the video-laryngoscope (GlideScope) in vocal fold injection medialization. *Laryngoscope* 124: 2136-2138, 2014.
58. Zhao Q, He Y, Wu Y, Huang D, Wang Y, Sun C, Ju J, Wang J and Mahr JJJ: Vocal cord lesions classification based on deep convolutional neural network and transfer learning. *Med Phys* 49: 432-442, 2022.
59. Kim H, Jeon J, Han YJ, Joo Y, Lee J, Lee S and Im S: Convolutional neural network classifies pathological voice change in laryngeal cancer with high accuracy. *J Clin Med* 9: 3415, 2020.
60. Contrera KJ, Betz J, Genter DJ and Lin FR: Association of hearing impairment and mortality in the national health and nutrition examination survey. *JAMA Otolaryngol Head Neck Surg* 141: 944-946, 2015.
61. Eroğlu O, Eroğlu Y, Yıldırım M, Karlıdağ T, Çınar A, Akyiğit A, Kaygusuz İ, Yıldırım H, Keleş E and Yalçın Ş: Is it useful to use computerized tomography image-based artificial intelligence modelling in the differential diagnosis of chronic otitis media with and without cholesteatoma? *Am J Otolaryngol* 43: 103395, 2022.
62. Szaleniec J, Wiatr M, Szaleniec M, Składzień J, Tomik J, Oleś K and Tadeusiewicz R: Artificial neural network modelling of the results of tympanoplasty in chronic suppurative otitis media patients. *Comput Biol Med* 43: 16-22, 2013.
63. Tama BA, Kim DH, Kim G, Kim SW and Lee S: Recent advances in the application of artificial intelligence in otorhinolaryngology-head and neck surgery. *Clin Exp Otorhinolaryngol* 13: 326-339, 2020.
64. O'Brien WT Sr, Hamelin S and Weitzel EK: The preoperative sinus CT: Avoiding a 'CLOSE' call with surgical complications. *Radiology* 281: 10-21, 2016.
65. Amanian A, Heffernan A, Ishii M, Creighton FX and Thamboo A: The evolution and application of artificial intelligence in rhinology: A state of the art review. *Otolaryngol Head Neck Surg* 169: 21-30, 2023.
66. Chowdhury NI, Smith TL, Chandra RK and Turner JH: Automated classification of osteomeatal complex inflammation on computed tomography using convolutional neural networks. *Int Forum Allergy Rhinol* 9: 46-52, 2019.
67. Huang J, Habib AR, Mendis D, Chong J, Smith M, Duvnjak M, Chiu C, Singh N and Wong E: An artificial intelligence algorithm that differentiates anterior ethmoidal artery location on sinus computed tomography scans. *J Laryngol Otol* 134: 52-55, 2020.
68. Parmar P, Habib AR, Mendis D, Daniel A, Duvnjak M, Ho J, Smith M, Roshan D, Wong E and Singh N: An artificial intelligence algorithm that identifies middle turbinate pneumatization (concha bullosa) on sinus computed tomography scans. *J Laryngol Otol* 134: 328-331, 2020.
69. Tamaki A, Rocco JW and Ozer E: The future of robotic surgery in otolaryngology-head and neck surgery. *Oral Oncol* 101: 104510, 2020.
70. Dumitru M, Berghi ON, Taciuc IA, Vrinceanu D, Manole F and Costache A: Could artificial intelligence prevent intraoperative anaphylaxis? Reference review and proof of concept. *Medicina (Kaunas)* 58: 1530, 2022.
71. Huyen C: Designing Machine Learning Systems: An Iterative Process for Production-ready Applications. 1st edition. O'Reilly Media, Inc., 2022.
72. Liyanage H, Liaw ST, Jonnagaddala J, Schreiber R, Kuziemy C, Terry AL and de Lusignan S: Artificial intelligence in primary health care: Perceptions, issues, and challenges. *Yearb Med Inform* 28: 41-46, 2019.
73. Maleki F, Ovens K, Najafian K, Forghani B, Reinhold C and Forghani R: Overview of machine learning part 1: Fundamentals and classic approaches. *Neuroimaging Clin N Am* 30: e17-e32, 2020.
74. Zhou KX, Patel M, Shimizu M, Wang E, Prisman E and Thang T: Development and validation of a novel craniofacial statistical shape model for the virtual reconstruction of bilateral maxillary defects. *Int J Oral Maxillofac Surg* 53: 146-155, 2024.
75. Morita D, Mazen S, Tsujiko S, Otake Y, Sato Y and Numajiri T: Deep-learning-based automatic facial bone segmentation using a two-dimensional U-Net. *Int J Oral Maxillofac Surg* 52: 787-792, 2023.
76. Nasteski V: An overview of the supervised machine learning methods. *Horizons B* 4: 51-62, 2017.
77. Oliver P: Random forests for medical applications, PhD Thesis, Technische Universität München, pp23-32, 2012.
78. Shim EJ, Yoon MA, Yoo HJ, Chee CG, Lee MH, Lee SH, Chung HW and Shin MJ: An MRI-based decision tree to distinguish lipomas and lipoma variants from well-differentiated liposarcoma of the extremity and superficial trunk: Classification and regression tree (CART) analysis. *Eur J Radiol* 127: 109012, 2020.
79. James G, Witten D, Hastie T, and Tibshirani R: An introduction to statistical learning. New York: Springer, pp315, 2015.
80. Ahmed AM, Rizaner A and Ulusoy AH: A novel decision tree classification based on post-pruning with Bayes minimum risk. *PLoS One* 13: e0194168, 2018.
81. Lim SJ, Jeon ET, Baek N, Chung YH, Kim SY, Song I, Rah YC, Oh KH and Choi J: Prediction of hearing prognosis after intact canal wall mastoidectomy with tympanoplasty using artificial intelligence. *Otolaryngol Head Neck Surg* 169: 1597-1605, 2023.
82. Li X, Wu X, Qian J, Yuan Y, Wang S, Ye X, Sha Y, Zhang R and Ren H: Differentiation of lacrimal gland tumors using the multi-model MRI: Classification and regression tree (CART)-based analysis. *Acta Radiol* 63: 923-932, 2022.
83. Ghassemi M, Oakden-Rayner L and Beam AL: The false hope of current approaches to explainable artificial intelligence in health care. *Lancet Digit Health* 3: e745-e750, 2021.
84. Maniaci A, Riela PM, Iannella G, Lechien JR, La Mantia I, De Vincentiis M, Cammaroto G, Calvo-Henriquez C, Di Luca M, Chiesa Estomba C, *et al*: Machine learning identification of obstructive sleep apnea severity through the patient clinical features: A retrospective study. *Life (Basel)* 13: 702, 2023.
85. Nagendran M, Chen Y, Lovejoy CA, Gordon AC, Komorowski M, Harvey H, Topol EJ, Ioannidis JPA, Collins GS and Maruthappu M: Artificial intelligence versus clinicians: Systematic review of design, reporting standards, and claims of deep learning studies. *BMJ* 368: m689, 2020.
86. Le ÉPV, Wang Y, Huang Y, Hickman S and Gilbert FJ: Artificial intelligence in breast imaging. *Clin Radiol* 74: 357-366, 2019.

