Universiteit Leiden
The Netherlands

A review on the application of deep learning for CT reconstruction, bone segmentation and surgical planning in oral and maxillofacial surgery

Minnema, J.; Ernst, A.; Eijnatten, M. van; Pauwels, R.; Forouzanfar, T.; Batenburg, K.J.; Wolff, J.

REVIEW ARTICLE

A review on the application of deep learning for CT reconstruction, bone segmentation and surgical planning in oral and maxillofacial surgery

[1]Jordi Minnema, [2]Anne Ernst, [3,4]Maureen van Eijnatten, [5]Ruben Pauwels, [1]Tymour Forouzanfar, [3,6]Kees Joost Batenburg and [7]Jan Wolff

[1]Department of Oral and Maxillofacial Surgery/Pathology, Amsterdam UMC and Academic Centre for Dentistry Amsterdam (ACTA), Vrije Universiteit Amsterdam, 3D Innovationlab, Amsterdam Movement Sciences, Amsterdam, The Netherlands; [2]Institute for Medical Systems Biology, University Hospital Hamburg-Eppendorf, Hamburg, Germany; [3]Computational imaging group, Centrum Wiskunde & Informatica (CWI), Amsterdam, The Netherlands; [4]Department of Biomedical Engineering, Eindhoven University of Technology, Medical Image Analysis Group, Eindhoven, The Netherlands; [5]Aarhus Institute of Advanced Studies, Aarhus University, Aarhus, Denmark; [6]Leiden Institute of Advanced Computer Science (LIACS), Leiden University, Leiden, The Netherlands; [7]Department of Dentistry and Oral Health, Aarhus University, Vennelyst Boulevard, Aarhus, Denmark

Computer-assisted surgery (CAS) allows clinicians to personalize treatments and surgical interventions and has therefore become an increasingly popular treatment modality in maxillofacial surgery. The current maxillofacial CAS consists of three main steps: (1) CT image reconstruction, (2) bone segmentation, and (3) surgical planning. However, each of these three steps can introduce errors that can heavily affect the treatment outcome. As a consequence, tedious and time-consuming manual post-processing is often necessary to ensure that each step is performed adequately. One way to overcome this issue is by developing and implementing neural networks (NNs) within the maxillofacial CAS workflow. These learning algorithms can be trained to perform specific tasks without the need for explicitly defined rules. In recent years, an extremely large number of novel NN approaches have been proposed for a wide variety of applications, which makes it a difficult task to keep up with all relevant developments. This study therefore aimed to summarize and review all relevant NN approaches applied for CT image reconstruction, bone segmentation, and surgical planning. After full text screening, 76 publications were identified: 32 focusing on CT image reconstruction, 33 focusing on bone segmentation and 11 focusing on surgical planning. Generally, convolutional NNs were most widely used in the identified studies, although the multilayer perceptron was most commonly applied in surgical planning tasks. Moreover, the drawbacks of current approaches and promising research avenues are discussed.

Correspondence to: Mr Jordi Minnema, E-mail: j.minnema@amsterdamumc.nl

Introduction

Spatial information embedded in medical three-dimensional (3D) images is being increasingly used to personalize treatments by means of computer-assisted surgery (CAS). This novel image-based treatment modality enables clinicians to perform patient-specific virtual operations, 3D-print personalized medical constructs and perform robot-guided surgery.[1] Moreover, CAS offers the unique possibility to conduct a limitless amount of different surgical simulations (osteotomies, grafts, implants, etc.) prior to surgery in a stress-free environment,[1] and to predict surgical outcomes with minimal risk to the patient. Furthermore, such virtual simulations have proven to be useful for patient communication and medical education.[2,3] As a result, CAS is currently being employed in multiple surgical branches involving the musculoskeletal system. In particular, CAS has advanced in the area of maxillofacial surgery.[4]

The current maxillofacial CAS workflow consists of different steps that are illustrated in Figure 1. The first step in the workflow is image acquisition. To date, numerous imaging modalities have become available on the market, including CT and MRI. CT imaging modalities are most commonly used to visualize bony structures due to their superior hard tissue contrast. CT scanners acquire X-ray projections of the patients' anatomy from multiple angles. These projection data can be subsequently reconstructed into a 3D image using a wide variety of reconstruction methods. After the CT image acquisition step, image processing is necessary to convert the CT scans into a virtual 3D model in the standard tessellation language (STL) file format. This file format is supported by all FDA-approved medical software packages that are currently used for computer-aided design (CAD) and computer-aided manufacturing. The most important step in this CT-to-STL conversion is image segmentation (Figure 1), in which clinicians define and delineate anatomies of interest such as bone. In the final step of the CAS workflow, the acquired STL models are exported to dedicated medical CAD software packages and used for surgical planning by virtually designing patient-specific implants, surgical guides and radiotherapy boluses.[5]

Each of the aforementioned steps (*i.e.* CT image reconstruction, bone segmentation and surgical planning) is a potential source of errors which can lead to inaccuracies in the final STL models and impair the treatment outcome.[6] For example, imaging noise, metallic structures and patient movements can heavily affect the CT image quality after reconstruction. For image segmentation, the segmentation technique can have a considerable effect on the accuracy of the resulting model.[6] Furthermore, the surgical planning step currently relies on extensive domain expertise and manual software input, which often hampers its reproducibility.

One way to overcome these limitations is to employ neural networks (NNs) during the different steps of this maxillofacial CAS workflow. These learning algorithms are different from traditional computer methods in that they can be trained to find characteristic features and patterns in data, without the need for explicit rules specified by domain experts. The most common NN is the multilayer perceptron (MLP), which consists of an input layer, several hidden layers and an output layer. Each of these layers comprises several computational building blocks called neurons. Within an MLP, neurons are connected to neurons in subsequent layers. The output of each neuron is the product of its input with a learned set of weights plus a learned bias. Finally, a non-linear activation function is applied.[7] It can be proven that MLPs can approximate any continuous function (universal approximation theorem,[8] which gives them the ability to infer descriptive functions from data).

In the training phase, the weights and biases of a NN are learned from training data. During this training process, a large amount of input data is propagated through the NN to predict the values of the output layer. The goal of training is to minimize the difference between the NN prediction and the desired output by iteratively updating the weights and biases of the network. After optimizing these trainable parameters, NNs can be used to automatically perform specific tasks of the maxillofacial CAS workflow.

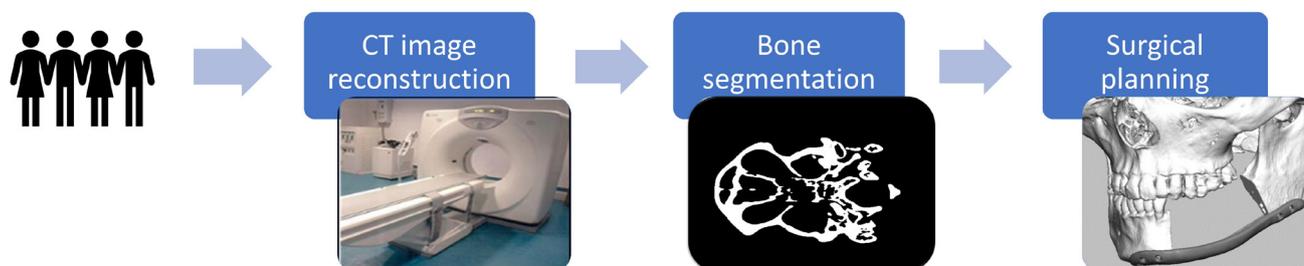Recent advances in computational power and the development of novel NN algorithms have brought

Figure 1 Schematic overview of the maxillofacial CAS workflow. CAS, computer-assisted surgery.

about a paradigm shift in the CAS workflow.[9] A wide variety of advanced NN architectures have been successfully employed for various tasks such as image reconstruction and segmentation. For example, convolutional neural networks (CNNs) are especially useful for processing image data. These networks apply convolutions instead of a set of multiplications to compute the output of the individual layers. Another important type of neural networks is the recurrent neural network (RNN), which can handle temporal dynamic data. However, due to the rapidly increasing number of studies published in the field, maxillofacial surgeons and medical engineers have been facing the difficult task of keeping up with all developments. Therefore, this scoping review aims to provide an overview of the different types of NN approaches that have been used during the three main steps required in the CAS workflow, *i.e.* CT image reconstruction, bone segmentation and surgical planning. Furthermore, the secondary goal of this review paper is to identify the current bottlenecks and possible next research steps regarding the application of NNs in the maxillofacial CAS workflow.

Methods and materials

Existing literature on the application of NNs in the maxillofacial CAS workflow was obtained using Pubmed, Embase, Scopus, Web of Science, and Google Scholar. An initial database was generated with the following search terms:

(1) (CT OR CBCT OR computed tomography OR cone-beam computed tomography) AND (image reconstruction OR image processing OR image analysis OR image segmentation) AND (artificial intelligence OR deep learning OR neural network)
(2) (bone OR bones OR bony) AND ((implant OR prosthesis OR virtual model) AND (design OR planning OR construct OR model)) AND (artificial intelligence OR deep learning OR neural network)

It must be noted that there was no specific focus on maxillofacial surgery when defining the search terms. The reason for this is that we believe that many of the techniques and methods used in other fields can also be of relevance to maxillofacial CAS. Choosing more generic search terms thus allowed us to identify a wider variety of literature relevant that was potentially relevant to maxillofacial CAS.

Publications were only included in the initial database if the search terms one or two were found in their title, abstract or keywords. After removing duplicates and adding literature found from references, a database of 6994 papers was acquired. The title and abstract of these publications were screened, resulting in 248 publications that were eligible for full-text reading. In order to assess whether papers were eligible for inclusion, the following inclusion and exclusion criteria were used:

Inclusion criteria:

- Described the development or implementation of a NN.
- Performed at least one of the three main steps required in the maxillofacial CAS workflow, *i.e.* CT image reconstruction, bone segmentation and surgical planning. Although surgical planning is an extremely broad field, this review specifically focused on the designing and optimization of implants and virtual models.
- Evaluated on medical data sets or artificial medical phantoms.

Exclusion criteria:

- Used non-CT based imaging modalities.

The study selection process of the present study is shown in Figure 2.

Results and discussion

This review aimed to identify NN architectures, training strategies and workflows that can potentially benefit CT image reconstruction, bone segmentation or surgical planning, since these steps are pivotal in the CAS workflow. In total, 76 studies were included in this review: 32 focusing on CT image reconstruction, 33 focusing on bone segmentation and 11 focusing on surgical planning. All studies are summarized in Table 1. In addition, Figure 3 shows the most popular NN approaches used in the reviewed studies.

Of the 32 reviewed studies for CT image reconstruction, 19 studies used simulated CT data to train and test their NN approach, 15 studies used clinical CT data, and 2 studies used CT data of physical phantoms. In contrast, all 33 bone segmentation studies used clinical data sets to train and test the NN approaches. two surgical planning studies were performed based on clinical data sets in two studies, and the remaining nine studies were performed based on simulated data. Details of the data sets used to train the NN approaches are provided in Figure 4.

Training and testing of the NNs with clinical data was performed with a mean of 33 ± 42 CT volumes and 12 ± 16 CT volumes, respectively. The mean ratio between the amount of training and testing data was approximately 9:2. Furthermore, 13 of the 76 reviewed studies employed a leave-k-out testing strategy to improve validation of the NNs. In this cross-validation strategy, the available data are split into k folds, where one fold is alternately used as testing data, and the remaining folds serve as training data for the NN. This process is repeated k times such that all folds have been used for testing.

Quantitative evaluation of the NNs' performances for CT image reconstruction tasks was most commonly performed using the peak signal-to-noise ratio (PSNR)
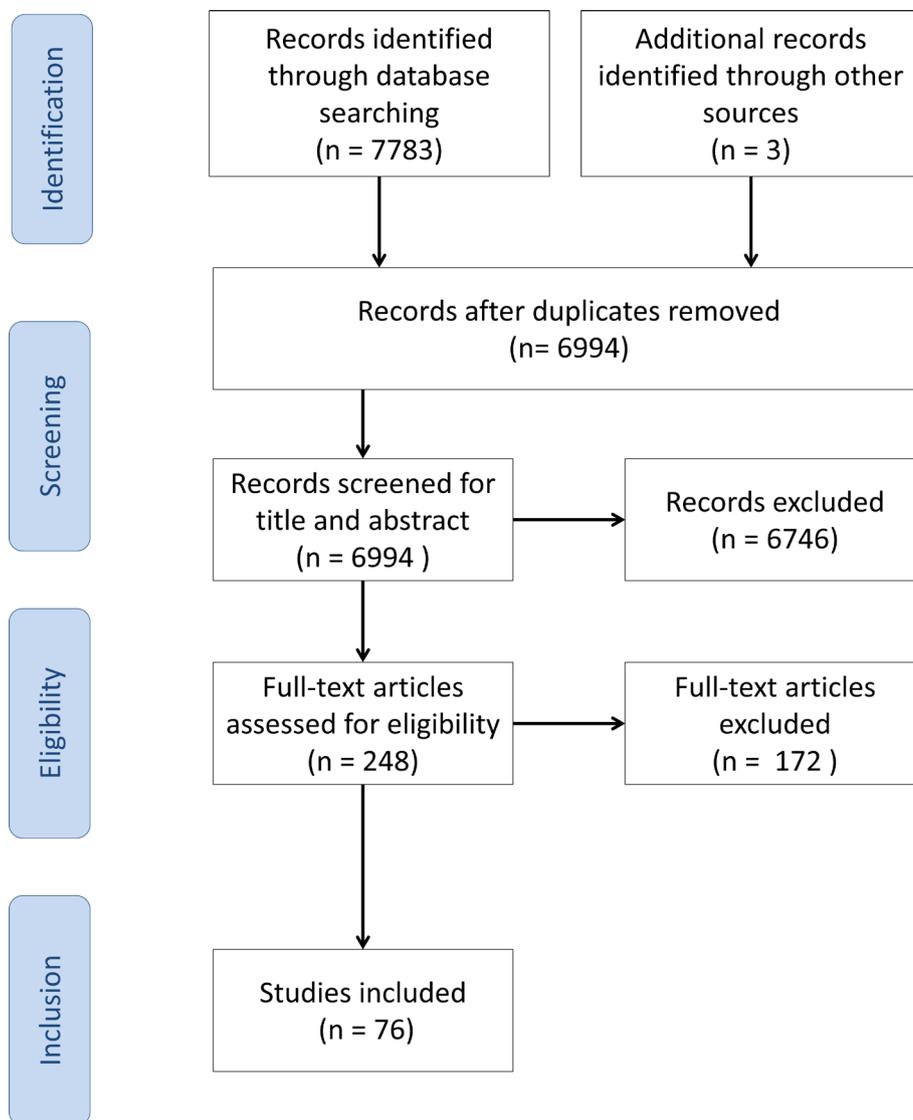
Figure 2 Overview of the study selection process of the present study.

and the structural similarity index measure (SSIM) (Figure 5a). The bone segmentation NNs were most commonly evaluated using the dice similarity coefficient (DSC) (Figure 5b). No consistency was observed in the performance metrics used to evaluate the surgical planning step (Figure 5c). Moreover, 3 of the 11 surgical planning studies did not include quantitative performance evaluations.

In the following subsections, we elaborate on the NN approaches that have been employed in each of the three steps of maxillofacial CAS.

Image reconstruction
Historically, CT image reconstruction (Figure 6) has been a notoriously difficult task, which aims to compute the density of objects or anatomical structures based on the attenuation of X-rays. To date, two different reconstruction methods have been predominantly employed in clinical settings: filtered backprojection (FBP) and iterative reconstruction (IR). FBP is an analytical method in which measured projection data are uniformly distributed across the CT scan with an angle that corresponds to the acquisition of the projection data. A filter is subsequently applied to reduce blurring in the CT scan. By using projection data acquired at multiple angles with respect to the patient, a 3D CT scan can be reconstructed. IR approaches start similar to the FBP in that they use the measured projection data to reconstruct an initial CT scan. Based on this initial scan, a forward operation is performed to create artificial projection data. The artificial projection data are then compared to the measured projection data, which are used to update the initial CT scan. The forward operation and the scan update are repeated until the quality

Table 1 Overview of the studies included in this review

CAS step	Year	CT imaging modality	Anatomy	Neural network architecture	Authors
CT image reconstruction	2005	Fan-beam	Abdomen	Radial basis-function NN	Hu[10]
	2006	Parallel-beam	Shepp-Logan phantom	Radial basis-function NN	Guo[11]
	2008	Parallel-beam	Shepp-Logan phantom	RNN	Cierniak[12]
	2008	Parallel-beam	Shepp-Logan phantom	RNN	Cierniak[13]
	2009	Fan-beam	Shepp-Logan phantom	RNN	Cierniak[14]
	2010	Parallel-beam	Shepp-Logan phantom	RNN	Cierniak[15]
	2010	Parallel-beam	Shepp-Logan phantom	RNN	Cierniak[16]
	2011	Parallel-beam; Fan-beam	Shepp-Logan phantom	RNN	Cierniak[17]
	2012	Parallel-beam	Shepp-Logan phantom	RNN	Cierniak and Lorent[18]
	2016	Parallel-beam, Fan-beam	Chest	CNN	Würfl et al.[19]
	2017	Fan-beam	Shepp-Logan phantom, anthropomorphic phantom head	CNN	Adler and Öktem[20]
	2018	Parallel-beam, Fan-beam	Ellipse phantom, Shepp-Logan phantom, Human phantoms	CNN	Adler and Öktem[21]
	2018	Multidetector row	Abdomen	CNN	Chen et al.[22]
	2018	Multidetector row	Abdomen, Chest, Rat brain	CNN	Gupta et al.[23]
	2018	Multidetector row	Abdomen	CNN	Han et al.[24]
	2018	Fan-beam	Chest	CNN	Liang et al.[25]
	2018	Cone-beam	Abdomen	CNN	Würfl et al.[26]
	2019	Fan-beam	Skull	CNN	Dong et al.[27]
	2019	Parallel-beam	Chest	CNN	Fu and de Man[28]
	2019	Multidetector row	Torso	CNN	He et al.[29]
	2019	Multidetector row	Chest	CNN	Lee et al.[30]
	2019	Fan-beam	Skull	GAN	Li et al.[31]
	2019	Multidetector row	Chest	CNN	Shen et al.[32]
	2019	Fan-beam	Abdomen	CNN	Wu et al.[33]
	2019	–	Shepp-Logan phantom	CNN	Zhang and Zuo[34]
	2019	Cone-beam	Chest	CNN	Zhang et al.[35]
	2020	Fan-beam	Prostate	CNN	Chen et al.[36]
	2020	Translational CT	Chest	CNN	Wang et al.[37]
	2020	Parallel-beam	Chest	CNN	Baguer et al.[38]
	2020	Parallel-beam	Chest	CNN	Ma et al.[39]
	2020	Fan-beam; Cone-beam	Chest	CNN	Wang et al.[40]
	2020	Cone-beam	Breast	GAN	Xie et al.[41]
Bone segmentation	2002	Multidetector row	Chest, Skull	MLP	Zhang and Valentino[42]
	2008	Multidetector row	Phalanx	MLP	Gassman et al.[43]
	2013	-	Jaw, mouth, nose, eye, brain	MLP	Kuo et al.[44]
	2017	Multidetector row	Femur	CNN	Chen et al.[45]
	2017	–	Mandible, Spinal cord	CNN	Ibragimov and Xing[46]
	2017	Multidetector row	Femoral head, bladder, intestine, colon	CNN	Males et al.[47]
	2018	Multidetector row	Whole body	CNN	Klein et al.[48]
	2018	Multidetector row	Vertebrae	CNN	Lessman et al.[49]
	2018	Multidetector row	Skull	CNN	Minnema et al.[50]
	2018	–	Vertebrae	Deep-belief network	Qadri et al.[51]

(Continued)

Table 1 (Continued)

CAS step	Year	CT imaging modality	Anatomy	Neural network architecture	Authors
	2018	Multidetector row	Mandible	CNN	Yan et al.[52]
	2018	Multidetector row	Vertebrae	CNN	Zhou et al.[53]
	2019	–	Vertebrae	CNN	Dutta et al.[54]
	2019	–	Teeth	CNN	Gou et al.[55]
	2019	Multidetector row	Whole-body	CNN	Klein et al.[56]
	2019	Multidetector row	Orbital bones	CNN	Lee et al.[57]
	2019	Multidetector row	Vertebrae	CNN	Lessmann et al.[58]
	2019	Cone-beam	Mandible, teeth	CNN	Minnema et al.[4]
	2019	Multidetector row	Vertebrae	CNN	Rehman et al.[59]
	2019	Cone-beam	Skull	CNN	Torosdagli et al.[60]
	2019	–	Vertebrae	CNN	Vania et al.[61]
	2019	Multidetector row	Pelvic bones	CNN	Wang et al.[62]
	2019	Micro-CT	Teeth	CNN	Yazdani et al.[63]
	2020	Cone-beam	Teeth	CNN	Lee et al.[64]
	2020	–	Temporal bone	CNN	Li et al.[65]
	2020	–	Vertebrae	CNN	Yin et al.[66]
	2020	Multidetector row	Whole-body	CNN	Noguchi et al.[67]
	2020	Multidetector row	Vertebrae	CNN	Bae et al.[68]
	2020	Multidetector row	Cochleae	CNN	Heutink et al.[69]
	2020	Cone-beam	Skull	CNN	Zhang et al.[70]
	2020	–	Vertebrae	Cascaded CNN	Xia et al.[71]
	2020	Cone-beam	Teeth	CNN	Chen et al.[72]
	2020	Cone-beam	Teeth	CNN	Rao et al.[73]
Surgical planning	2000	–	Skull	MLP based on legendre polynomials	Hsu and Tseng[74]
	2001	–	Skull	MLP based on legendre polynomials	Hsu and Tseng[75]
	2001	Multidetector row	Radius	Bernstein Basis function network	Knopf and Al-Naji[76]
	2010	n.a.	Femur	MLP	Hambli[77]
	2012	n.a.	Femur	MLP	Campoli et al.[78]
	2012	n.a.	Bone microstructure	Meshing growing neural gas (MGNG)	Fischer and Holdstein[79]
	2016	n.a.	Femur	MLP	Chanda et al.[80]
	2018	n.a.	Dental implant	MLP	Roy et al.[81]
	2019	n.a.	Spinal implant	MLP	Biswas et al.[82]
	2020	–	Skull	GAN	Kodym et al.[83]
	2020	–	Mandible	GAN	Liang et al.[84]

- : not specified; CNN: convolutional neural network; GAN: generative adversarial network; MLP: multilayer perceptron;NN: neural network; RNN: recurrent neural network; n.a.: not applicable.

of the CT scan is satisfactory or for a fixed number of iterations. To date, a wide variety of different IR algorithms have been developed, including the algebraic reconstruction technique (ART), the simultaneous iterative reconstruction technique (SIRT) and model-based iterative reconstruction (MBIR).

Over the last decade, NNs have opened up a wealth of opportunities in the field of CT image reconstruction, offering CT images with higher quality than with FBP, while requiring shorter reconstruction times than current IR reconstruction approaches. One of the first efforts in developing NNs for medical CT image reconstruction can be traced back to 2005, when Hu et al proposed two different NN-based approaches.[10] The first approach aimed to reconstruct 2D CT images using a Radial Basis Function NN (RBF-NN), which is a similar to the classical MLP, but uses a radial basis function as non-linear activation function. The input of this RBF-NN consisted of CT projection data, and the desired target consisted of previously reconstructed
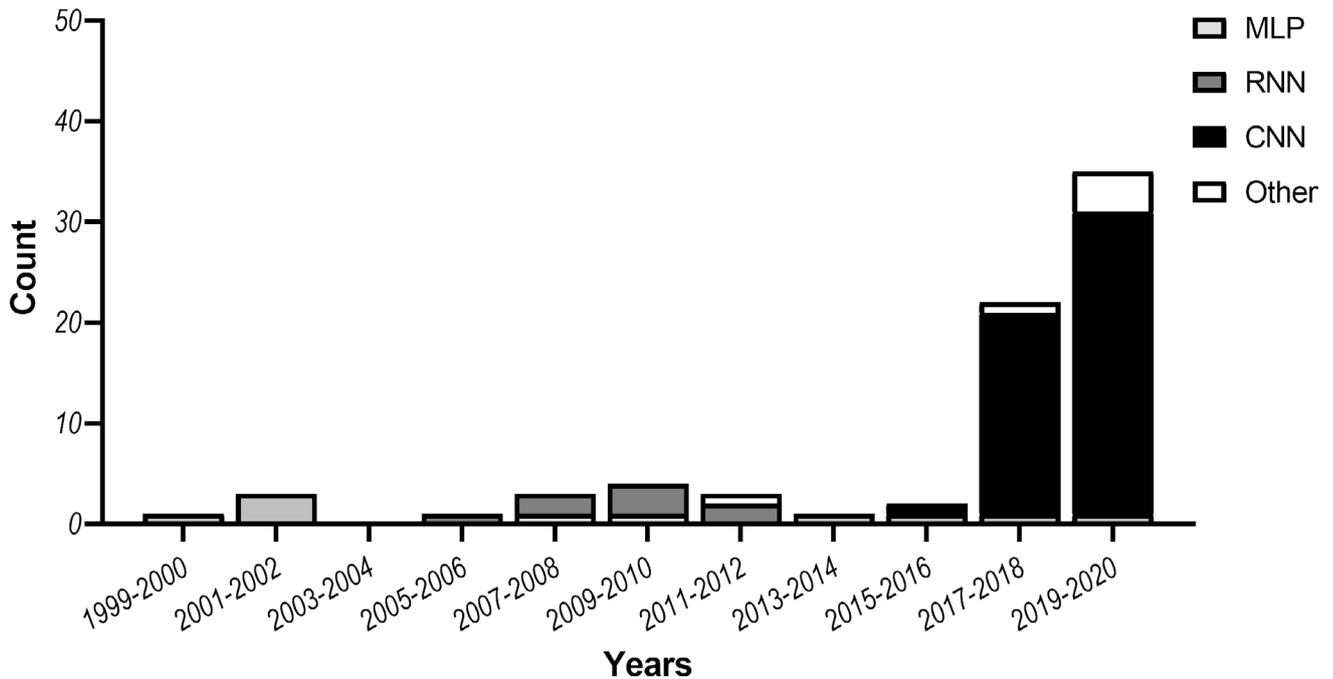
Figure 3 Most popular NN approaches in the maxillofacial CAS workflow over the past two decades. CAS, computer-assisted surgery; CNN, convolutional neural network; MLP, multilayer perceptron; NN, neural network; RNN, recurrent neural network.

CT scans. In the second approach, the RBF-NN was employed to iteratively estimate the intensities of the voxels in the CT scan. More specifically, an IR scheme was employed in which the RBF-NN was trained to update the voxel values in the CT scans. Although both RBF-NN-based approaches were initially used to reconstruct small 32×32 images with only 8–16 projection angles, image sizes were increased to 128×128 in a later study by Guo et al.[11]

In a series of publications between 2008 and 2012,[12–18] Cierniak et al developed multiple RNNs to iteratively reconstruct CT scans after using traditional backprojection to create an initial CT scan. The proposed RNNs were essentially used as learnable filters for the FBP reconstruction method, replacing the fixed filters in

Figure 4 The data sets that were used to train and test the NN approaches in the reviewed studies. NN, neural network.

FBP.[17] Cierniak demonstrated that the proposed RNNs improved CT image quality compared to FBP. Moreover, it was shown that the proposed method was able to reconstruct projection data acquired from various CT scanning geometries such as fan-beam and parallel-beam (Figure 6).

Although the aforementioned RBF-NN and RNNs initially demonstrated promising results, they have been rapidly surpassed by CNNs. CNNs have the ability of capturing spatially oriented patterns in imaging data, which makes them particularly suited to reconstruct CT scans. An example of such a CNN approach was proposed by Würfl et al,[19] who demonstrated that the traditional FBP method can be expressed in terms of CNNs. Their CNN consisted of a single convolutional layer to mimic the filtering of FBP, and a fully connected layer to learn the backprojection step. They found that the CNN achieved comparable results to traditional FBP while markedly reducing the computational complexity required to perform the reconstruction. In addition, they showed that their framework can be extended to mimic the Feldkamp David and Kress (FDK) algorithm that is commonly used to reconstruct CT scans acquired with the cone-beam geometry (Figure 6).[26]

Another way of using CNNs for CT image reconstruction is to incorporate them within IR algorithms. For example, Adler and Öktem replaced forward operations of the iterative gradient descent reconstruction algorithm[20] and the primal-dual hybrid gradient algorithm[21] by partially trainable CNNs. In addition, various CNN approaches have been developed to
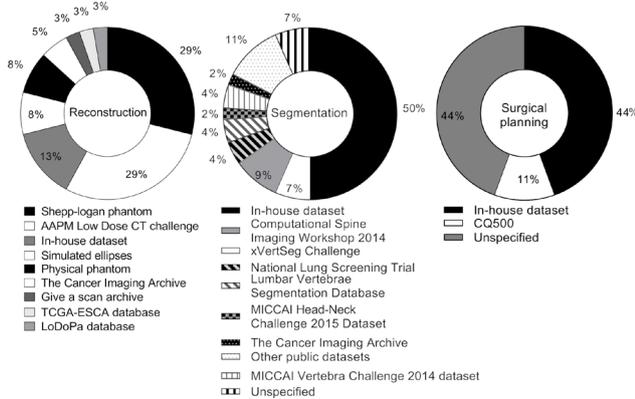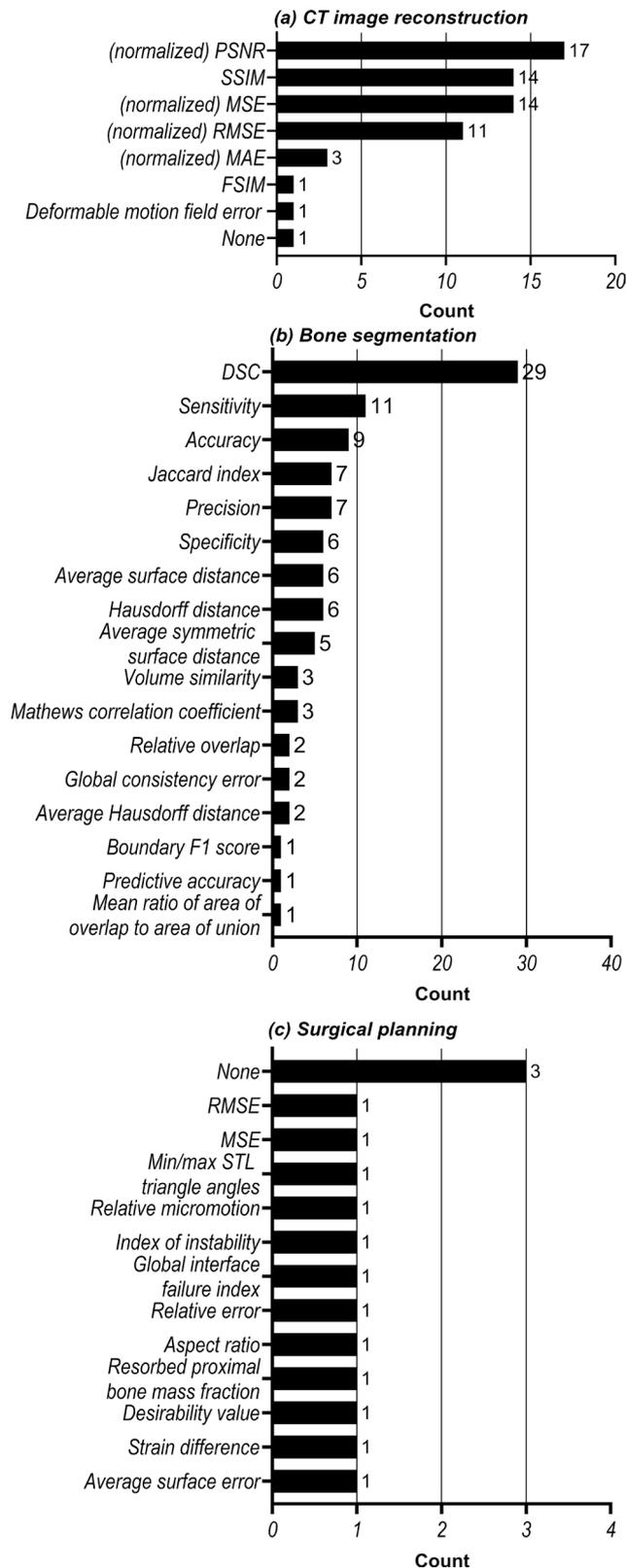
Figure 5 Analysis of the evaluation metrics used to quantify the performance of NN approaches in (a) CT image reconstruction, (b) bone segmentation and (c) surgical planning. DSC, dice similarity coefficient; RMSE, root meat-squared error; MSE, mean-squared error; PSNR, peak signal-to-noise ratio; SSIM, structural similarity index measure; STL, standard tessellation language.
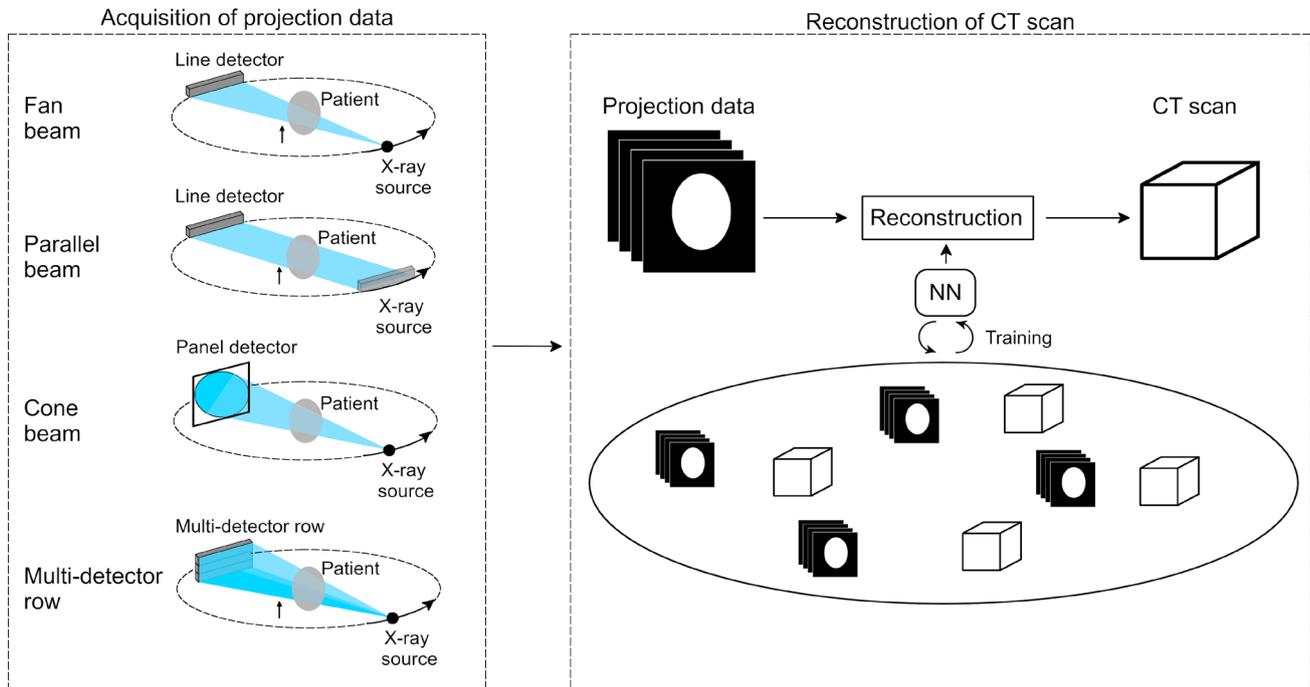
Figure 6   Schematic overview of CT image reconstruction. First, CT projection data are acquired with one of the four commonly used CT geometries (*i.e.* fan-beam, parallel-beam, cone-beam and multidetector row). The acquired projection data are used to reconstruct a CT scan. This reconstruction step can be improved by training a NN. NN, NN, neural network.

improve reconstruction quality[34,36,37] and computational efficiency[33] of IR algorithms.

A different strategy was taken by Chen et al, who developed a CNN-based framework to find a direct mapping between CT projection data and reconstructed CT scans.[22] Their framework used 50 iteration-inspired layers that each consisted of three learned convolutional operations. This framework significantly outperformed state-of-the-art reconstruction approaches. Moreover, they showed that this approach can be effectively used on incomplete projection data, which is a common problem in clinical practice, since radiation dose often needs to be reduced in order to comply with the 'as low as reasonably achievable' (ALARA) principle. The CNN-based IR framework was further improved by Xie et al,[41] who implemented a learnable back-projection step that was previously fixed.

A different way of finding a mapping between projection data and reconstructed CT scans was developed by Fu and De Man.[28] However, instead of finding a *direct* mapping, they proposed a hierarchical CNN in which the difficult reconstruction problem was split up into multiple intermediate steps that can be easily learned. This approach does not only improve the quality of reconstructed CT scans, but also gives clinicians an insight into the intermediate steps learned by the CNN. Similarly, Wang et al,[40] developed two coupled CNNs to convert sinograms into CT images. The first CNN takes sinograms as input and converts them to data that are better suitable for the FBP or FDK algorithms. The output of the FBP or FDK algorithms (in the image

domain) is subsequently fed to the second CNN, which further improves the reconstructed image quality. Ma et al[39] also aimed to reconstruct CT images directly from the projection data. However, instead of breaking down the reconstruction challenge into multiple steps, they applied a combination of fully connected layers and convolutional layers to reduce the memory space requirement. They showed that their proposed approach results in substantially better image quality than standard FBP.

Ever since it was introduced in 2015, the U-Net has become a popular CNN architecture for medical image analysis.[85] The U-Net consists of two convolutional paths. The downsampling path (*i.e.* encoder) creates a low-dimensional representation of the input data in order to capture local patterns. The upsampling path (*i.e.* decoder) subsequently captures global patterns through a series of upsampling steps. Both paths of the U-net are interconnected with skip-connections that allow the network to combine learned patterns at various scales. Applications of U-Nets for medical CT image reconstruction include the estimation of incomplete CT projection data (*e.g.* sparse-view and limited-angle CT)[27,37,38] and improving IR, specifically the projected gradient descent reconstruction algorithm, by replacing the forward projector with a U-Net.[23] Furthermore, variants of U-Net such as dual-frame and tight-frame U-Net have been proposed to reduce streaking and blurring artifacts in sparse-view CT scans during post-processing.[24] Finally, U-Net has also been employed to predict deformation vector fields used to reconstruct 4DCT images.[35]
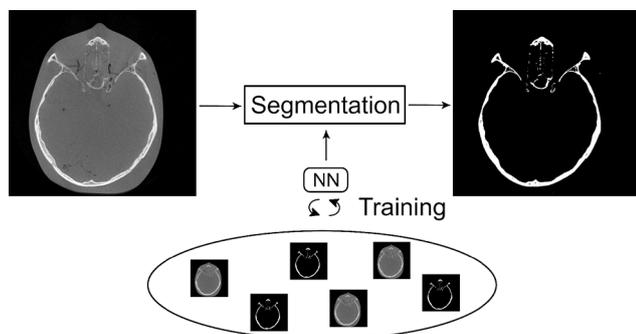
Figure 7 Schematic representation of the bone segmentation task required for maxillofacial CAS. A NN can be trained to automatically perform this segmentation task. CAS, computer-assisted surgery; NN, neural network.

In summary, a wide variety of NN approaches have been proposed for CT image reconstruction (Table 1). In particular, CNNs have shown to be an incredibly interesting research topic with many new recent publications. Three CNN approaches were identified that are particularly interesting for the current maxillofacial CAS workflow. The first CNN approach aims to replace the computationally demanding forward operations within current IR methods.[20,21] This markedly reduces the time constraint of applying IR methods in the maxillofacial CAS workflow. The second interesting CNN approach identified in this study is capable of reconstructing CT scans from incomplete projection data.[27,31,37] Such approaches would enable clinicians to acquire high-quality CT scans of patients using low dose protocols. The third CNN approach replaces the total reconstruction process,[22,28] which means that the CNN is trained to directly reconstruct CT scans from raw projection data. However, it must be noted that such fully learned reconstruction approaches are computationally expensive and require large amounts of annotated training data, which are not always available.

Image segmentation
Image segmentation refers to the task of labelling voxels of an image as a particular class (Figure 7). In the context of maxillofacial surgery, this image segmentation step is typically used to distinguish bony structures from soft tissues or air. Although a large variety of statistical methods have been developed for bone segmentation, they usually require the intervention of a medical professional in order to produce an accurate output, mainly due to the lack of reliable Hounsfield units and the limited signal-to-noise ratio of CBCT scans. It is therefore desirable to automate this task as far as possible, thereby relieving the medical professional from this labor-intensive and time-consuming task, while also increasing the accuracy and consistency of the segmentation results. We therefore reviewed the NNs used to automate CT bone segmentation tasks.

The first study describing the segmentation of bone in CT scans using NNs was published in 2002.[42] In this

study, a hierarchy of MLPs was trained on small patches of head and chest CT scans. The trained MLPs subsequently classified the center pixels of the patches and combined all separate pixel classifications to create a segmented image. Although similar MLPs were adopted by Gassman et al[43] (2008) and Kuo et al[44] (2013) to segment the phalanges and the nasal septum, respectively, different input data were used to train the MLPs. Namely, Gassman et al provided spherical co-ordinates, probabilities and intensities of individual pixels as input for their MLP, whilst Kuo et al used single rows of CT scans.

Segmentation using NNs took a significant leap after the ground-breaking performance of a novel CNN architecture (AlexNet) developed by Krizhevsky et al. in 2012.[86] Similar to the aforementioned MLPs, this CNN architecture was trained using a patch-based approach in which the CNN aimed to classify the center voxels of small image patches. Inspired by the performance of this patch-based CNN, researchers have shown that such CNNs can achieve similar and in some cases superior performances compared to state-of-the art statistical methods when segmenting the mandible,[46,52] the spine[51,61] and the skull[50] in CT scans. Nevertheless, the clinical application of patch-based CNNs for segmentation has been limited since many redundant convolution operations are necessary to classify all image voxels, which significantly slows down training and increases segmentation times.

In order to overcome this challenge, Ronneberger et al published a variation of the traditional CNN architecture, known as U-Net.[85] U-Net can directly provide a segmented image as output (*i.e.* semantic segmentation) which increases its computational efficiency compared to patch-based CNNs. As a result, this U-Net architecture has ever since been applied for several CT bone segmentation tasks. For example, Klein et al,[48,56] and Noguchi et al,[67] applied U-net to segment bone in whole-body CT scans and reported that the U-Net performed significantly better than the standard segmentation procedure, *i.e.* global thresholding combined with morphological operations. Furthermore, U-Net was used to segment vertebrae,[54,59,68,71] teeth,[55,72] pelvic bones,[62] orbital bones,[57] cochleae[69] and cranial bones.[70] In a different study,[45] a CNN architecture very similar to the U-Net, namely SegNet,[87] was used to perform edge detection and multiscale segmentation of the femur in CT scans. Finally, Lessmann et al extended the standard U-Net architecture in order to both segment and identify an a priori unknown number of vertebrae.[49,58] Their proposed architecture was able to automatically identify the individual vertebrae, whilst having comparable segmentation performance as the standard U-Net.

Since high segmentation performances have been achieved by U-net across a large number of studies, U-Net is currently considered the state-of-the-art for CT bone segmentation. Nevertheless, alternative CNN architectures for medical image segmentation have also been

widely employed in the reviewed papers. For example, Men et al proposed a deep dilated CNN (DDCNN), in which the first and last layers perform dilated convolutions in order to extract multiscale features.[47] Such a dilated convolution refers to the inflation of a convolution kernel while leaving sparse spaces between its elements. This dilation thus increases the receptive field of the kernel without increasing the number of model parameters. Torosdagli et al[60] segmented the mandible using a fully convolutional DenseNet, which is comparable to U-Net, but uses dense blocks instead of regular convolutional layers. These dense blocks comprise multiple densely connected convolutional layers, which means that each convolutional layer within the dense block is connected to all other layers in the block. Similarly, a UDS-Net[64] has been proposed, consisting of a U-Net with a dense block and spatial dropout, to segment teeth. Furthermore, 3D-DSD net has been developed which consists of a U-Net with a dense block and additional skip connections.[65] The use of dilated convolutions and dense connections was further exploited in the mixed-scale dense CNN (MS-D network),[88] which was used to segment the mandible in cone-beam CT scans.[4] In an MS-D network, each convolutional layer performs a dilated convolution and is densely connected to all other layers of the network. It was found that these properties allow the MS-D network to achieve comparable segmentation performances as U-Net, while using far fewer trainable parameters.[4] Finally, Zhou et al developed the so-called N-Net, which is similar to U-Net but has an additional stream of downsampling layers,[53] whereas Rao et al modified the U-Net by replacing the normal convolutions with so-called Deep Bottleneck Architectures.[73]

In this section, different NN approaches used for bone segmentation in CT scans were reviewed. Similar to the NN approaches used for CT image reconstruction, bone segmentation seems to be increasingly performed using CNNs in favor of alternative NNs. The CNN approaches identified in this review can be roughly categorized into patch-based approaches and semantic segmentation approaches. The patch-based approaches allow extracting a large number of patches from relatively few CT scans, which facilitates CNN training. Semantic segmentation approaches, on the other hand, annotate each voxel of an image during a single forward pass through the CNN, which is typically far more computationally efficient than patch-based approaches. As a result, current state-of-the-art CNN approaches usually perform semantic segmentation. Examples of such widely used CNN architectures are U-Net,[85] ResNet[89] and MS-D network.[88]

Surgical planning

The final step in the maxillofacial CAS workflow is surgical planning. This step typically involves a combination of computer-simulated bone reconstruction and subsequent designing of appropriate patient-specific implants. For example, a simple method to reconstruct fractured bones in the skull is to mirror the bony structures from the contralateral healthy side.[90] However, this technique is often constrained to small defects on one side of the skull. For larger defects, a complex procedure involving various 3D-modelling software packages is required. The success of NNs in image reconstruction and image segmentation calls for the application of similar techniques during surgical planning and implant design. In this section, we therefore review different NN approaches that have been applied for surgical planning.

An interesting area for automated surgical planning is the reconstruction of skull plates because of the simplicity of the local anatomical geometry. Already in 2001, efforts were made to automatically design skull implants.[74,75] The authors of these papers used a single-layer MLP in order to reconstruct cranial bone from CT images of patients with skull defects. In order to optimize training of the MLP, an approach relying on orthogonal functions (Legendre polynomials) was employed. The presented MLP was able to significantly speed up and improve the design of skull implants. However, while mathematically interesting, this particular approach is unlikely to generalize to larger defects on the side of the skull, since Legendre polynomials are insufficient to correctly approximate such complex defects. A similar approach was proposed by Knopf and Al-Naji,[76] which also approximates anatomical features using an MLP with a single layer of polynomial functions, specifically Bernstein polynomials. After being trained on a set of segmented CT slices, the network was able to reproduce the anatomical structure of healthy bone. The main advantage of this approach is that the output of the MLP is in the shape of curves that describe the bony structures, which can be easily loaded into CAD software. Nevertheless, it must be noted that most approaches based on curve fitting are currently restricted to the neurocranium where the geometry of the skull can be reasonably approximated by smooth curves.

A different way of employing NNs for surgical planning is to optimize parametrized implant designs. The advantage of working with parametrized designs is that the NNs do not require imaging data. In addition, since implant designs can often be described using a few parameters, relatively simple NN architectures can be used to optimize the parametrization of the design. This approach was, *e.g.* taken by Chanda et al,[80] who optimized parameterized hip implants based on the effects of initial micromotion, stress shielding, and interface stress. Using a combination of a single-hidden-layer MLP, a genetic algorithm and finite element analysis (FEA), the authors deduced that the standard implant design can be significantly improved. A similar combination of an MLP, a genetic algorithm and FEA was also used in a different study in order to find the best combination of material properties and geometry to generate patient-specific dental molar implants.[81] This

approach was also used by Biswas et al,[82] who optimized the design of patient-specific spine implants based on the bone condition, body weight and implant diameter.

Another well-attended problem in surgical planning is to predict how bone adapts to different loads. This bone adaptation can be caused by surgical implants and may lead to complications after surgery.[91] To date, bone adaptation prediction has been commonly performed using FEA. However, researchers recently found that a combination of the well-established FEA method and an MLP can significantly speed up the computations.[77] The inverse problem has also been studied, specifically the estimation of load parameters for a given bone porosity inferred from CT images. For example, Campoli et al[78] showed how a single-layer MLP can solve this inverse problem and compute the load for a femur. The network was specific to a single femur and was trained using simulated data. After the introduction of noise to the training data, the network was still able to successfully estimate the load parameters for the femur.

An interesting NN approach that does not rely on the geometric properties of bone was employed by Morais et al.[92] They trained a CNN, specifically a deep convolutional autoencoder, to reconstruct fractured or missing parts of the skull. This autoencoder learned a representation of a healthy skull and was able to subsequently reconstruct a portion of the skull that was artificially removed. However, the accuracy and applicability of this deep learning model to high-resolution data remains challenging due to the computing power necessary for training and validation. Moreover, the proposed method was only validated on MRI data and not yet on CT data.

A relatively new approach towards reconstructing tissue morphologies is to apply generative adversarial networks (GANs). GANs consist of two networks: an generative network for generating new, fictive images based on input images, and a discriminator for distinguishing the generated images from real images. In the context of morphology reconstruction, a GAN can be used to generate images of healthy tissues based on images of fractured or diseased tissues. The generated images can then be hardly, if not at all, distinguished from real images of healthy tissues. From a clinical perspective, such generated healthy images can be extremely useful as they provide a surgeon a view of what the result should resemble. Furthermore, generating healthy images can be particularly helpful when constructing patient-specific implants. An example of a GAN applied in such a setting can be found in the study by Liang et al,[84] who developed a GAN to reconstruct the morphology of the mandible based on CT images of patients suffering from ameloblastoma or gingival cancer. Similarly, Kodym et al[83] used a GAN to reconstruct the shape of defective skulls. Even though interesting approaches have been developed, relatively few studies have described NNs for surgical planning (Table 1). The majority of surgical planning studies

included in this review implemented MLPs for applications such as dental implant design[80] and prediction of bone adaptation.[77,78] A possible explanation for the relatively few studies describing NN-based surgical planning might be that it is extremely difficult to develop a single network to account for all variations that clinicians face during surgical planning. For example, there are numerous possibilities of designing implants or surgical tools, and choosing an adequate design heavily depends on the available software tools on the market, the type of imaging data used, and the personal preferences of medical engineers. As a consequence, it is almost impossible to effectively train a NN to design implants if no constraints are imposed. Although a few studies solved this problem by using a parametrized implant design to reduce the degrees of freedom,[80,81] this leads to more generic and less patient-specific implants. Hence, automating personalized surgical planning and implant design using NNs remains difficult. Nevertheless, the field is still currently active, as demonstrated by the recent AutoImplant Challenge[93] that was created to motivate participants to develop automated methods for cranial implant design.

Current challenges and future research

In order to further develop and validate NN approaches for the three main steps of the maxillofacial CAS workflow, challenges remain that need to be overcome. One of these challenges is the quantitative performance evaluation. To date, the SSIM and MSE have been commonly used to assess image quality in the CT image reconstruction step, whereas the DSC is commonly used to assess segmentation performances (Figure 4). These generic metrics, however, do not always represent clinical relevance. For example, maxillofacial surgeons often assess the surface of the bony structures in order to create a treatment plan and design implants. Therefore, surface-based performance metrics might be preferred over the generic metrics. One possible way of evaluating segmentation performance would be to convert segmented CT scans into virtual 3D models and subsequently calculating geometrical distances between a gold-standard virtual model and the NN-based virtual models.[94] This approach enables the quantification of the surface quality of bony structures, and also allows the visualization and interpretation of the differences between the two virtual models.

Another well-known limitation of most NN approaches is the need for large amounts of paired training data, *i.e.* input and target.[95] Although an extremely large number of medical images are acquired on a daily basis that can be used as input to train a NN, they commonly lack appropriate annotations. Annotating such medical images requires a high level of domain-specific expertise, in contrast to natural images that can be easily annotated through crowdsourcing. In order to avoid the challenge of acquiring annotated target images, an interesting research direction might

be to develop semi- and unsupervised NN training approaches, which do not depend on annotated data sets to learn.

To date, the development of NN approaches has been mainly performed in academic settings. Although many different NNs have shown to improve the efficiency, accuracy and consistency in which clinical tasks can be performed, none of these NN approaches has, to the best of our knowledge, been approved for maxillofacial CAS by the United States Food & Drug Administration. As a consequence, the application of NNs in routine clinical tasks remains very limited. In order to allow for large-scale use of NNs in clinical settings, additional research is necessary that focus on the robustness of NNs when faced with large anatomical variations and different imaging characteristics. For example, two recent studies have already shown that a single CNN is able to accurately segment CT scans acquired using various CT scanners.[50,68]

The reconstruction of CT scans in the maxillofacial CAS workflow is typically optimized for visual interpretation by clinicians. This optimization is to be expected since visual assessment is often fundamental to establish correct diagnoses. However, a CT scan optimized for human interpretation might not be the ideal input for CNNs to perform the subsequent segmentation and surgical planning tasks. Namely, CNNs solely extract patterns from the CT scans without taking visual aspects into account. A possible way of improving the quality of the CAS workflow may therefore be to jointly train two CNNs to simultaneously reconstruct and segment a CT scan. Such joint training approaches have already shown promising results in lung nodule detection,[96] and may open up promising avenues when used in the maxillofacial CAS workflow.

One of the limitations that was encountered in the reviewed publications is that the heterogeneity of the data used to train the NNs makes it particularly challenging to draw generic conclusions. NN models have been trained on CT images of various different anatomical regions or phantoms that have been acquired using a plethora of different scanners, or through simulation. In order to validate a methodology, benchmark data sets should be developed which would enable a better comparison between the NN approaches. Furthermore, any work performed on simulated data and/or phantoms should be validated on clinical data, as it is challenging to assess the clinical applicability of NN approaches that have been solely validated on synthetic data.

Finally, it must be noted that research on artificial intelligence, and deep learning in particular, is evolving at an incredible rate due to the unprecedented interest and resource investment these fields. As a consequence, the state-of-the-art on this topic will naturally evolve rapidly. Nevertheless, this does not mean that a review cannot be useful. In fact, literature reviews may be arguably even more important when considering such rapidly evolving topics, as they help to identify bottlenecks, and to suggest new lines of research to advance the field.

Conclusion

This scoping review describes different NN approaches used in the three main steps of the maxillofacial CAS workflow, *i.e.* CT image reconstruction, bone segmentation and surgical planning. In recent years, CNNs have rapidly become the most popular NN approach for CT image reconstruction and image segmentation, whereas MLPs remain the most common approach in surgical planning. Although CT image reconstruction and bone segmentation have been widely explored fields of research, additional research is required on the application of NNs for surgical planning. In order to reach the full potential of NNs for maxillofacial CAS, future research should focus on overcoming the challenges addressed in this review.

REFERENCES

1. Swennen GRJ, Mollemans W, Schutyser F. Three-dimensional treatment planning of orthognathic surgery in the era of virtual imaging. *J Oral Maxillofac Surg* 2009; **67**: 2080–92. https://doi.org/10.1016/j.joms.2009.06.007
2. Harris BT, Montero D, Grant GT, Morton D, Llop DR, Lin W-S. Creation of A 3-dimensional virtual dental patient for computer-guided surgery and CAD-CAM interim complete removable and fixed dental prostheses: A clinical report. *J Prosthet Dent* 2017; **117**: S0022-3913(16)30281-5: 197–204: . https://doi.org/10.1016/j.prosdent.2016.06.012
3. Berman NB, Durning SJ, Fischer MR, Huwendiek S, Triola MM. The role for virtual patients in the future of medical education.

Acad Med 2016; **91**: 1217–22. https://doi.org/10.1097/ACM. 0000000000001146

4. Minnema J, van Eijnatten M, Hendriksen AA, Liberton N, Pelt DM, Batenburg KJ, et al. Segmentation of dental cone-beam CT scans affected by metal artifacts using a mixed-scale dense convolutional neural network. *Med Phys* 2019; **46**: 5027–35. https://doi.org/10.1002/mp.13793

5. Su S, Moran K, Robar JL. Design and production of 3D printed bolus for electron radiation therapy. *J Appl Clin Med Phys* 2014; **15**: 4831. https://doi.org/10.1120/jacmp.v15i4.4831

6. van Eijnatten M, van Dijk R, Dobbe J, Streekstra G, Koivisto J, Wolff J. CT image segmentation methods for bone used in medical additive manufacturing. *Med Eng Phys* 2018; **51**: S1350-4533(17)30263-1: 6–16: . https://doi.org/10.1016/j.medengphy. 2017.10.008

7. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. *Med Image Anal* 2017; **42**: S1361-8415(17)30113-5: 60–88: . https://doi.org/10.1016/j.media.2017.07.005

8. . Kratsios A. The Universal Approximation Property: Characterizations, Construction, Representation and Existence. Ann Math Artif Intell 2021; 89:435-469.

9. Esteva A, Robicquet A, Ramsundar B, Kuleshov V, DePristo M, Chou K, et al. A guide to deep learning in healthcare. *Nat Med* 2019; **25**: 24–29. https://doi.org/10.1038/s41591-018-0316-z

10. Hu M, Guo P, Lyu MR. COMPARATIVE STUDIES ON THE CT IMAGE RECONSTRUCTION BASED ON THE RBF NEURAL NETWORK. In: Zhu Q, ed. 11th Joint International Computer Conference - JICC 2005; Chongqing, China. Chongqing, China: World Scientific; October 2005. pp. 948–51. https://doi.org/10.1142/9789812701534_0213

11. Guo P, Hu M, Jia Y. eds. 2006 International Conference on Computational Intelligence and Security; Guangzhou, China. Guangzhou, China: IEEE; November 2006. pp. 1865–68. 10.1109/ICCIAS.2006.295389.. doi: https://doi.org/10.1109/ICCIAS.2006. 295389

12. Cierniak R. A 2D approach to tomographic image reconstruction using A hopfield-type neural network. *Artif Intell Med* 2008; **43**: 113–25. https://doi.org/10.1016/j.artmed.2008.03.003

13. Cierniak R. A new approach to image reconstruction from projections using A recurrent neural network. *International Journal of Applied Mathematics and Computer Science* 2008; **18**: 147–57. https://doi.org/10.2478/v10006-008-0014-y

14. Cierniak R. New neural network algorithm for image reconstruction from fan-beam projections. *Neurocomputing* 2009; **72**: 3238–44. https://doi.org/10.1016/j.neucom.2009.02.005

15. Cierniak Ret alA Statistical Tailored Image Reconstruction from Projections Method. In: Phillips-Wren G, Jain LC, Nakamatsu K, et al., eds. *Advances in Intelligent Decision Technologies*. Berlin, Heidelberg: Springer; 2010, pp. 181–90.

16. Cierniak R. A Statistical Appraoch to Image Reconstruction from Projections Problem Using Recurrent Neural Network. In: Diamantaras K, Duch W, Iliadis LS, eds. *ICANN 2010*. Berlin, Heidelberg: Springer; 2010, pp. 138–41.

17. Cierniak R. Neural network algorithm for image reconstruction using the "grid-friendly" projections. *Australas Phys Eng Sci Med* 2011; **34**: 375–89. https://doi.org/10.1007/s13246-011-0089-x

18. Cierniak R, Lorent A. A Neuronal Approach to the Statistical Image Reconstruction from Projections Problem. In: Nguyen N-T, Hoang K, Jędrzejowicz P, eds. *ICCCI 2012*. Berlin, Heidelberg: Springer; 2012, pp. 344–53.

19. Würfl T, Ghesu FC, Christlein V, Maier A. et alDeep Learning Computed Tomography. In: Ourselin S, Joskowicz L, Sabuncu MR, et al., eds. *MICCAI 2016*. Berlin, Heidelberg: Springer International Publishing; 2016, pp. 432–40.

20. Adler J, Öktem O. Solving ill-posed inverse problems using iterative deep neural networks. *Inverse Problems* 2017; **33**: 124007. https://doi.org/10.1088/1361-6420/aa9581

21. Adler J, Oktem O. Learned primal-dual reconstruction. *IEEE Trans Med Imaging* 2018; **37**: 1322–32. https://doi.org/10.1109/TMI.2018.2799231

22. Chen H, Zhang Y, Chen Y, Zhang J, Zhang W, Sun H, et al. LEARN: learned experts' assessment-based reconstruction network for sparse-data CT. *IEEE Trans Med Imaging* 2018; **37**: 1333–47. https://doi.org/10.1109/TMI.2018.2805692

23. Gupta H, Jin KH, Nguyen HQ, McCann MT, Unser M. CNN-based projected gradient descent for consistent CT image reconstruction. *IEEE Trans Med Imaging* 2018; **37**: 1440–53. https://doi.org/10.1109/TMI.2018.2832656

24. Han Y, Ye JC. Framing U-net via deep convolutional framelets: application to sparse-view CT. *IEEE Trans Med Imaging* 2018; **37**: 1418–29. https://doi.org/10.1109/TMI.2018.2823768

25. Liang K, Xing Y, Yang H, Kang K, Chen G-H, Lo JY, et al. Improve angular resolution for sparse-view CT with residual convolutional neural network. In: Chen G-H, Lo JY, Gilat Schmidt T, eds. Physics of Medical Imaging; Houston, United States. Houston, United States: SPIE. Vol. 105731K; 15 March 2018. https://doi.org/10.1117/12.2293319

26. Wurfl T, Hoffmann M, Christlein V, Breininger K, Huang Y, Unberath M, et al. Deep learning computed tomography: learning projection-domain weights from image domain in limited angle problems. *IEEE Trans Med Imaging* 2018; **37**: 1454–63. https://doi.org/10.1109/TMI.2018.2833499

27. Dong J, Fu J, He Z. A deep learning reconstruction framework for X-ray computed tomography with incomplete data. *PLoS ONE* 2019; **14**(11): e0224426. https://doi.org/10.1371/journal.pone. 0224426

28. Fu L, De Man B, Matej S, Metzler SD. A hierarchical approach to deep learning and its application to tomographic reconstruction. In: Matej S, ed. The Fifteenth International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine; Philadelphia, United States. Philadelphia, United States: SPIE; 2 June 2019. pp. 1107202. https://doi.org/10.1117/12.2534615

29. He J, Yang Y, Wang Y, Zeng D, Bian Z, Zhang H, et al. Optimizing a parameterized plug-and-play ADMM for iterative low-dose CT reconstruction. *IEEE Trans Med Imaging* 2019; **38**: 371–82. https://doi.org/10.1109/TMI.2018.2865202

30. Lee D, Choi S, Kim H-J. High quality imaging from sparsely sampled computed tomography data with deep learning and wavelet transform in various domains. *Med Phys* 2019; **46**: 104–15. https://doi.org/10.1002/mp.13258

31. Li Z, Cai A, Wang L, Zhang W, Tang C, Li L, et al. Promising generative adversarial network based sinogram inpainting method for ultra-limited-angle computed tomography imaging. *Sensors (Basel)* 2019; **19**: E3941: 3941: . https://doi.org/10.3390/s19183941

32. Shen L, Zhao W, Xing L, Bosmans H, Chen G-H, Gilat Schmidt T. Harnessing the power of deep learning for volumetric CT imaging with single or limited number of projections. In: Bosmans H, Chen G-H, Gilat Schmidt T, eds. Physics of Medical Imaging; San Diego, United States. San Diego, United States: SPIE; 21 March 2019. pp. 1094826. https://doi.org/10.1117/12.2513032

33. . Wu D, Kim K, El Fakhri G, Li Q. Computational-efficient cascaded neural network for CT image reconstruction. In: Bosmans H, Chen G-H, Gilat Schmidt T (Eds). Medical Imaging 2019: Physics of Medical Imaging. San Diego, United States: SPIE, 2019, 109485Z

34. . Zhang J, Zuo H. Iterative CT image reconstruction using neural network optimization algorithms. In: Bosmans H, Chen G-H, Gilat Schmidt T (Eds). Medical Imaging 2019: Physics of Medical Imaging. San Diego, United States: SPIE, 2019, 1094863

35. Zhang Y, Huang X, Wang J. Advanced 4-dimensional cone-beam computed tomography reconstruction by combining motion estimation, motion-compensated reconstruction, biomechanical modeling and deep learning. *Vis Comput Ind Biomed Art* 2019; **2**: 23. https://doi.org/10.1186/s42492-019-0033-6

36. Chen G, Hong X, Ding Q, Zhang Y, Chen H, Fu S, et al. AirNet: fused analytical and iterative reconstruction with deep neural network regularization for sparse-data CT. *Med Phys* 2020; **47**: 2916–30. https://doi.org/10.1002/mp.14170

37. Wang J, Liang J, Cheng J, Guo Y, Zeng L. Deep learning based image reconstruction algorithm for limited-angle translational

computed tomography. *PLoS ONE* 2020; **15**(1): e0226963. https://doi.org/10.1371/journal.pone.0226963

38. Baguer DO, Leuschner J, Schmidt M. Computed tomography reconstruction using deep image prior and learned reconstruction methods. *Inverse Problems* 2020; **36**: 094004. https://doi.org/10.1088/1361-6420/aba415

39. Ma G, Zhu Y, Zhao X. Learning image from projection: A full-automatic reconstruction (FAR) net for computed tomography. *IEEE Access* 2020; **8**: 219400–414. https://doi.org/10.1109/ACCESS.2020.3039638

40. Wang W, Xia X-G, He C, Ren Z, Lu J, Wang T, et al. An end-to-end deep network for reconstructing CT images directly from sparse sinograms. *IEEE Trans Comput Imaging* 2020; **6**: 1548–60. https://doi.org/10.1109/TCI.2020.3039385

41. Xie H, Shan H, Cong W, Liu C, Zhang X, Liu S, et al. Deep efficient end-to-end reconstruction (DEER) network for few-view breast CT image reconstruction. *IEEE Access* 2020; **8**: 196633–46. https://doi.org/10.1109/ACCESS.2020.3033795

42. Zhang D, Sonka M, Fitzpatrick JM, Valentino DJ. Segmentation of anatomical structures in x-ray computed tomography images using artificial neural networks. In: Sonka M, Fitzpatrick JM, eds. Medical Imaging 2002; San Diego, CA. San Diego, United States: SPIE; 2002. pp. 4684. https://doi.org/10.1117/12.467133

43. Gassman EE, Powell SM, Kallemeyn NA, DeVries NA, Shivanna KH, Magnotta VA, et al. Automated bony region identification using artificial neural networks: Reliability and validation measurements. Skeletal Radiol 2008; 37:313–319.

44. Kuo C-FJ. Three-dimensional Reconstruction System for Automatic Recognition of Nasal Vestibule and Nasal Septum in CT Images. J Med Biol Eng 2014; 34:574-580.

45. Chen F, Liu J, Zhao Z, Zhu M, Liao H. Three-dimensional feature-enhanced network for automatic femur segmentation. *IEEE J Biomed Health Inform* 2019; **23**: 243–52. https://doi.org/10.1109/JBHI.2017.2785389

46. Ibragimov B, Xing L. Segmentation of organs-at-risks in head and neck CT images using convolutional neural networks. *Med Phys* 2017; **44**: 547–57. https://doi.org/10.1002/mp.12045

47. Men K, Dai J, Li Y. Automatic segmentation of the clinical target volume and organs at risk in the planning CT for rectal cancer using deep dilated convolutional neural networks. *Med Phys* 2017; **44**: 6377–89. https://doi.org/10.1002/mp.12602

48. Klein A, Warszawski J, Hillengaß J, Maier-Hein KH. et alTowards Whole-body CT Bone Segmentation. In: Maier A, Deserno TM, Handels H, et al., eds. *Bildverarbeitung für die Medizin 2018*. Berlin, Heidelberg: Springer; 2018, pp. 204–9.

49. Lessmann N, Išgum I, Ginneken B. Iterative convolutional neural networks for automatic vertebra identification and segmentation in CT images. In: Angelini ED, Landman BA, eds. *Medical Imaging 2018: Image Processing*. Houston, United States: SPIE; 2018, pp. 1057408.

50. Minnema J, van Eijnatten M, Kouw W, Diblen F, Mendrik A, Wolff J. CT image segmentation of bone for medical additive manufacturing using a convolutional neural network. *Comput Biol Med* 2018; **103**: S0010-4825(18)30311-1: 130–39: . https://doi.org/10.1016/j.compbiomed.2018.10.012

51. Furqan Qadri S, Ai D, Hu G, Ahmad M, Huang Y, Wang Y, et al. Automatic deep feature learning via patch-based deep belief network for vertebrae segmentation in CT images. *Applied Sciences* 2018; **9**: 69. https://doi.org/10.3390/app9010069

52. Yan M, Guo J, Tian W, Yi Z. Symmetric convolutional neural network for mandible segmentation. *Knowledge-Based Systems* 2018; **159**: 63–71. https://doi.org/10.1016/j.knosys.2018.06.003

53. Zhou W, Lin L, Ge G. *N* -net: 3D fully convolution network-based vertebrae segmentation from CT spinal images . *Int J Patt Recogn Artif Intell* 2019; **33**: 1957003. https://doi.org/10.1142/S0218001419570039

54. Dutta S, Das B, Kaushik S, Bak PR, Chen P-H. Assessment of optimal deep learning configuration for vertebrae segmentation from CT images. In: Bak PR, Chen P-H, eds. Imaging Informatics for Healthcare, Research, and Applications; San Diego, United States. San Diego, United States: SPIE. Vol. 109541A; 21 March 2019. https://doi.org/10.1117/12.2512636

55. Gou M, Rao Y, Zhang M, Sun J, Cheng K. Automatic Image Annotation and Deep Learning for Tooth CT Image Segmentation. In: Zhao Y, Barnes N, Chen B, Westermann R, Kong X, Lin C (Eds). ICIG 2019: Springer International Publishing, 2019, pp 519–28.

56. Klein A, Warszawski J, Hillengaß J, Maier-Hein KH. Automatic bone segmentation in whole-body CT images. Int J CARS 2019; 14:21–29.

57. Lee MJ, Hong H, Shim KW, Park S. M-N. Orbital Bone Segmentation from Head and Neck CT Images Using Multi-Graylevel-Bone Convolutional Networks. In: *ISBI 2019*. Venice, Italy: IEEE; 2019, pp. 692–95.

58. Lessmann N, van Ginneken B, de Jong PA, Išgum I. Iterative fully convolutional neural networks for automatic vertebra segmentation and identification. *Medical Image Analysis* 2019; **53**: 142–55. https://doi.org/10.1016/j.media.2019.02.005

59. Rehman F, Ali Shah SI, Riaz MN, Gilani SO, R. F. A region-based deep level set formulation for vertebral bone segmentation of osteoporotic fractures. *J Digit Imaging* 2019; **33**: 191–203. https://doi.org/10.1007/s10278-019-00216-0

60. Torosdagli N, Liberton DK, Verma P, Sincan M, Lee JS, Bagci U. Deep Geodesic Learning for Segmentation and Anatomical Landmarking. IEEE Trans Med Imaging 2019; 38:919–931.

61. Vania M, Mureja D, Lee D. Automatic spine segmentation from CT images using convolutional neural network via redundant generation of class labels. *Journal of Computational Design and Engineering* 2019; **6**: 224–32. https://doi.org/10.1016/j.jcde.2018.05.002

62. Wang C, Connolly B, Oliveira Lopes PF, Frangi AF, Smedby Ö. et alPelvis Segmentation Using Multi-pass U-Net and Iterative Shape Estimation. In: Vrtovec T, Yao J, Zheng G, et al., eds. *MSKI*. Granada, Spain: Springer International Publishing; 2019, pp. 49–57.

63. Yazdani A, Stephens NB, Cherukuri V, Ryan T, Monga V. Domain-Enriched Deep Network for Micro-CT Image Segmentation. In: Matthews MB (ed). 2019 53rd Asilomar Conference on Signals, Systems, and Computers. Pacific Grove, United States: IEEE, 2019 pp 1867–1871.

64. Lee S, Woo S, Yu J, Seo J, Lee J, Lee C. Automated CNN-based tooth segmentation in cone-beam CT for dental implant planning. *IEEE Access* 2020; **8**: 50507–18. https://doi.org/10.1109/ACCESS.2020.2975826

65. Li X, Gong Z, Yin H, Zhang H, Wang Z, Zhuo L. A 3D deep supervised densely network for small organs of human temporal bone segmentation in CT images. *Neural Netw* 2020; **124**: S0893-6080(20)30007-1: 75–85: . https://doi.org/10.1016/j.neunet.2020.01.005

66. Yin X, Li Y, Shin B. Automatic Segmentation of Human Spine with Deep Neural Network. In: Park JJ, Park D-S, Jeong Y-S, Pan Y (Eds). Advances in Computer Science and Ubiquitous Computing. Singapore: Springer Singapore, 2020 pp 202–207.

67. Noguchi S, Nishio M, Yakami M, Nakagomi K, Togashi K. Bone segmentation on whole-body CT using convolutional neural network with novel data augmentation techniques. *Comput Biol Med* 2020; **121**: S0010-4825(20)30140-2: 103767: . https://doi.org/10.1016/j.compbiomed.2020.103767

68. Bae H-J, Hyun H, Byeon Y, Shin K, Cho Y, Song YJ, et al. Fully automated 3D segmentation and separation of multiple cervical vertebrae in CT images using a 2D convolutional neural network. Comput Biol Med 2020; 184:105119.

69. Heutink F, Koch V, Verbist B, van der Woude WJ, Mylanus E, Huinck W, et al. Multi-scale deep learning framework for cochlea localization, segmentation and analysis on clinical ultra-high-resolution CT images. *Comput Methods Programs Biomed* 2020; **191**: S0169-2607(19)32023-1: 105387: . https://doi.org/10.1016/j.cmpb.2020.105387

70. Zhang J, Liu M, Wang L, Chen S, Yuan P, Li J, et al. Context-guided fully convolutional networks for joint craniomaxillofacial bone segmentation and landmark digitization. *Med Image Anal*

2020; **60**: S1361-8415(18)30209-3: 101621: . https://doi.org/10.1016/j.media.2019.101621

71. Xia L, Xiao L, Quan G, Bo W. 3D cascaded convolutional networks for multi-vertebrae segmentation. *Curr Med Imaging* 2020; **16**: 231–40. https://doi.org/10.2174/1573405615666181204151943

72. Chen Y, Du H, Yun Z, Yang S, Dai Z, Zhong L, et al. Automatic segmentation of individual tooth in dental CBCT images from tooth surface map by a multi-task FCN. *IEEE Access* 2020; **8**: 97296–309. https://doi.org/10.1109/ACCESS.2020.2991799

73. . Rao Y, Wang Y, Meng F, Pu J, Sun J, Wang Q. A Symetric Fully Convolutional Residual Network with DCRF for Accurate Tooth Segmentation. IEEE Access 2020; 8:92028-92038

74. Hsu J-H, Tseng C-S. Application of orthogonal neural network to craniomaxillary reconstruction. *J Med Eng Technol* 2000; **24**: 262–66. https://doi.org/10.1080/030919000300037221

75. Hsu J-H, Tseng C-S. Application of three-dimensional orthogonal neural network to craniomaxillary reconstruction. *Comput Med Imaging Graph* 2001; **25**: 477–82. https://doi.org/10.1016/s0895-6111(01)00019-2

76. Knopf GK, Al-Naji R. Adaptive reconstruction of bone geometry from serial cross-sections. *Artificial Intelligence in Engineering* 2001; **15**: 227–39. https://doi.org/10.1016/S0954-1810(01)00006-1

77. Hambli R. Application of neural networks and finite element computation for multiscale simulation of bone remodeling. *J Biomech Eng* 2010; **132**(11): 114502. https://doi.org/10.1115/1.4002536

78. Campoli G, Weinans H, Zadpoor AA. Computational load estimation of the femur. *J Mech Behav Biomed Mater* 2012; **10**: 108–19. https://doi.org/10.1016/j.jmbbm.2012.02.011

79. Liebschner MAK. Computer-Aided Tissue Engineering. In: Liebschner MAK, ed. *A Neural Network Technique for Remeshing of Bone Microstructure*. Totowa, NJ: Humana Press; 2012., pp. 135–41. https://doi.org/10.1007/978-1-61779-764-4

80. Chanda S, Gupta S, Pratihar DK. Effects of interfacial conditions on shape optimization of cementless hip stem: an investigation based on a hybrid framework. *Struct Multidisc Optim* 2015; **53**: 1143–55. https://doi.org/10.1007/s00158-015-1382-1

81. Roy S, Dey S, Khutia N, Roy Chowdhury A, Datta S. Design of patient specific dental implant using FE analysis and computational intelligence techniques. *Applied Soft Computing* 2018; **65**: 272–79. https://doi.org/10.1016/j.asoc.2018.01.025

82. Biswas JK, Dey S, Karmakar SK, Roychowdhury A, Datta S. Design of patient specific spinal implant (pedicle screw fixation) using FE analysis and soft computing techniques. *CMIR* 2020; **16**: 371–82. https://doi.org/10.2174/1573405614666181018122538

83. . Kodym O, Španěl M, Herout A. Skull shape reconstruction using cascaded convolutional networks. Comput Biol Med 2020; 123:103886.

84. Liang Y, Huan J, Li J-D, Jiang C, Fang C, Liu Y. Use of artificial intelligence to recover mandibular morphology after disease. *Sci Rep* 2020; **10**: 16431. https://doi.org/10.1038/s41598-020-73394-5

85. . Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. 2015;234–241.

86. Krizhevsky A, Sutskever I, Geoffrey E. H. ImageNet classification with deep convolutional neural networks. *NIPS'12* 2012; **1**: 1097–1105.

87. Badrinarayanan V, Kendall A, Cipolla R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans Pattern Anal Mach Intell* 2017; **39**: 2481–95. https://doi.org/10.1109/TPAMI.2016.2644615

88. Pelt DM, Sethian JA. A mixed-scale dense convolutional neural network for image analysis. *Proc Natl Acad Sci U S A* 2018; **115**: 254–59. https://doi.org/10.1073/pnas.1715832114

89. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); Las Vegas, NV, USA. United States: IEEE; June 2016. pp. 770–78. https://doi.org/10.1109/CVPR.2016.90

90. Scolozzi P. Maxillofacial reconstruction using polyetheretherketone patient-specific implants by "mirroring" computational planning. *Aesthetic Plast Surg* 2012; **36**: 660–65. https://doi.org/10.1007/s00266-011-9853-2

91. Comenda M, Quental C, Folgado J, Sarmento M, Monteiro J. Bone adaptation impact of stemless shoulder implants: a computational analysis. *J Shoulder Elbow Surg* 2019; **28**: S1058-2746(19)30183-1: 1886–96: . https://doi.org/10.1016/j.jse.2019.03.007

92. . Morais A, Egger J, Alves V. Automated Computer-aided Design of Cranial Implants Using a Deep Volumetric Convolutional Denoising Autoencoder. In: Rocha Á, Adeli H, Reis LP, Costanzo S (Eds). New Knowledge in Information Systems and Technologies. La Toja Island, Spain: Springer International Publishing, 2019, pp 151–160.

93. Egger J, Li J, Chen X, Schäfer U, Campe GO, Krall M, et al. Towards the automatization of cranial implant design in cranioplasty. 2020.

94. van Eijnatten M, Koivisto J, Karhu K, Forouzanfar T, Wolff J. The impact of manual threshold selection in medical additive manufacturing. *Int J Comput Assist Radiol Surg* 2017; **12**: 607–15. https://doi.org/10.1007/s11548-016-1490-4

95. Maier A, Syben C, Lasser T, Riess C. A gentle introduction to deep learning in medical image processing. *Z Med Phys* 2019; **29**: S0939-3889(18)30120-X: 86–101: . https://doi.org/10.1016/j.zemedi.2018.12.003

96. . Wu D, Kim K, Dong B, Fakhri GE, Li Q. End-to-End Lung Nodule Detection in Computed Tomography. In: Shi Y, Suk H-I, Liu M (Eds). Machine Learning in Medical Imaging. Granada, Spain: Springer International Publishing, 2018, pp 37–45.