International Standard

ISO/IEC 3532-2

Information technology — Medical image-based modelling for 3D printing —

Part 2: Segmentation

Technologies de l'information — Modélisation médicale à base d'images pour l'impression 3D —

Partie 2: Segmentation

First edition
2024-02

Reference number
ISO/IEC 3532-2:2024(en)

Contents

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives or www.iec.ch/members_experts/refdocs).

ISO and IEC draw attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO and IEC take no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO and IEC had not received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents and https://patents.iec.ch. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see www.iso.org/iso/foreword.html. In the IEC, see www.iec.ch/understanding-standards.

This document was prepared by Joint Technical Committee ISO/IEC JTC 1, Information technology.

A list of all parts in the ISO/IEC 3532 series can be found on the ISO and IEC websites.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html and www.iec.ch/national-committees.

Introduction

This document was developed in response to the need for customization of 3D printing technology in the medical industry through the use of information and communication technology (ICT).

There are many points where the existing standards for additive manufacturing (AM) do not match the requirements of the medical industry. From medical images to 3D printing, medical device development is quite a complex journey with complicated management of multiple pieces of software.

With the emerging market for medical 3D printed parts, there are many points requiring standardization.

There is currently no standardized process for the creation of protocols and validation procedures to ensure that medical imaging data can be consistently and accurately transformed into a 3D printed object.

For medical 3D printing, segmentation techniques should be optimized and combined according to the characteristics of the medical images and corresponding body parts to get an optimal 3D model.

In particular, during medical image segmentation, identification of the pixels of organs or lesions from raw data such as computed tomography (CT) or magnetic resonance (MR) images, is one of the most challenging analysis tasks.

For example, segmentation of the orbital bone is necessary for orbital wall reconstruction in cranio-maxillofacial surgery to support the eye globe position and restore the volume and shape of the orbit. However, orbital bone segmentation is challenging as the orbital bone is composed of cortical bone with a high intensity value, and trabecular and thin bone with low intensity values, similar to soft tissue.

The human bone is delineated and extracted by segmentation techniques, and a 3D skeletal model is built from this segmentation. The minimization of errors during segmentation of relevant body parts of interest is critical. As there are several known critical issues for this segmentation, a verification process is made before proceeding.

Not only single segmentation techniques but also combinations of those techniques should be adopted for accurate extraction of a target body part. However, this process depends heavily on the operator. For minimization of errors during this job, operators should know which segmentation technique is most used in their imaging software and possess the necessary skills for that technique.

Thresholding techniques which are provided by a default Hounsfield unit (HU) range do not completely recover true bony structure.[1] An operator should typically adjust the extent of the segmentation manually. The problem is usually under-segmentation. However, over-segmentation will also be problematic for further designing processes, especially for surgical implants. Various techniques have been suggested to reduce human error and improve performance and consistency for segmentation issues.[2]

This document proposes a standardized process for the optimization of segmentation.

Information technology — Medical image-based modelling for 3D printing —

Part 2: Segmentation

1 Scope

This document provides an overview of the segmentation process for medical image-based modelling of human bone. This document specifies a standardized process to improve the performance of human bone segmentation.

This document is also applicable to medical 3D printing systems that include medical 3D modelling capabilities.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 15708-1, *Non-destructive testing — Radiation methods for computed tomography — Part 1: Terminology*

ISO/IEC 2382, *Information technology — Vocabulary*

ISO/ASTM 52950, *Additive manufacturing — General principles — Overview of data processing*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/ASTM 52950, ISO 15708-1, ISO/IEC 2382 and the following apply.

ISO and IEC maintain terminology databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at https://www.electropedia.org/

3.1
image acquisition
scanning of the structure of interest using computed tomography (CT), magnetic resonance (MR) imaging or other three-dimensional imaging technology

3.2
image annotation
process of attaching labels to an image

3.3
label
classifying phrase or name applied to a target

3.4
learning
<machine learning> process by which a biological or an automatic system gains knowledge or skills that it may use to improve its performance

[SOURCE: ISO/IEC 2382:2015, 2122966, modified — Notes to entry have been removed.]

3.5
segmentation
process of separating the objects of interest from their surroundings

Note 1 to entry: Segmentation can be applicable to 2D, 3D, raster or vector data.

3.6
ground-truth label
correct answer of the training set for segmentation based on supervised learning

3.7
region of interest
ROI
specified boundary as defined in the image

3.8
machine learning
ML
process of optimizing model parameters through computational techniques, such that the model's behaviour reflects the data or experience

[SOURCE: ISO/IEC 22989:2022, 3.3.5]

3.9
labelled data
group of data that have been tagged with one or more labels

3.10
hyperparameter
characteristic of a machine learning algorithm that affects its learning process

Note 1 to entry: Hyperparameters are selected prior to training and can be used in processes to help estimate model parameters.

Note 2 to entry: Examples of hyperparameters include number of network layers, width of each layer, type of activation function, optimization method, learning rate for neural networks; the choice of kernel function in a support vector machine; number of leaves or depth of a tree; the K for K-means clustering; the maximum number of iterations of the expectation maximization algorithm; the number of Gaussians in a Gaussian mixture.

[SOURCE: ISO/IEC 22989:2022, 3.3.4]

3.11
medical image
type of images generated by medical imaging devices

4 Abbreviated terms

AI	artificial intelligence
AIM	annotation and image markup
CAD	computer-aided diagnosis
CCL	connected component labelling
CNN	convolutional neural network
CRF	conditional random field
CT	computed tomography
DICOM	digital imaging and communications in medicine
DL	deep learning
DSC	dice similarity coefficient
FCN	fully convolutional network
FOV	field of view
FN	false negative
FP	false positive
HU	Hounsfield unit
IoU	intersection over union
MIoU	mean intersection over union
ML	machine learning
MR	magnetic resonance
MRI	magnetic resonance imaging
NIfTI	neuroimaging informatics technology initiative
PET	positron emission tomography
SPECT	single-photon emission computerized tomography
TN	true negative
TP	true positive
US	ultrasonography

5 Objective of segmentation

5.1 Background

The purpose of segmentation is to extract a specific region or organ from a patient's CT/MR medical image and use it to create a 3D model.

Segmentation is the process of partitioning an image into different meaningful segments. For medical images, segmentation techniques should be optimized and combined according to the characteristics of image acquisition modalities and body parts to get an ideal 3D visualization. The human bone is delineated and extracted by segmentation techniques, and a 3D skeletal model is built from this segmentation. The minimization of errors during segmentation of relevant anatomy is critical.

Modelling for medical 3D printing requires optimized segmentation to provide better overlaying and matching processes for human tissues and organs. However, most of the commercially available image software cannot segment human bone effectively.

For improved medical image based 3D modelling, formalization and standardization of these procedures is required.

5.2 Types of segmentation methods

Several methods have been investigated that segment human bone from CT images.[18][19][22][23]

Semi-automatic segmentation: Thresholding is the simplest method of image segmentation, and thresholding can be used to create binary images. This method replaces each pixel in an image with an object pixel if the pixel intensity is greater than a specific human bone threshold value, or replaces it with a background pixel if the pixel intensity is less than a specific human bone threshold value.

Deformable model-based segmenation: Deformable models are curves or surfaces for segmentation in the image domain, which deform under the influence of internal and external forces to delineate the object boundaries. The internal forces are defined such that they preserve the shape smoothness of the model, while the external forces are defined by the image features to drive the model toward the desired region boundary. By constraining extracted boundaries to be smooth and incorporating other prior information about the shape, deformable models offer robustness to both image noise and boundary gaps.[35][36] However, there is a limit due to the difficulty of segmentation at the weak object boundary of a thin bone with a low intensity value similar to soft tissue.

CNN-based segmentation: The FCN is a network that does not contain any dense layers. Instead, it contains 1x1 convolutions that perform the task of fully connected layers. The U-Net network, which is commonly used for medical image segmentation, is based on a fully convolutional network and consists of a contracting path and an expansive path, which gives it the U-shaped architecture.

6 Overall segmentation process

6.1 General

The overall segmentation process consists of seven steps in total, as described in 6.2 to 6.8. Figure 1 shows the overall process flow of segmentation, where the numbers in parenthesis refer to clauses of this document.

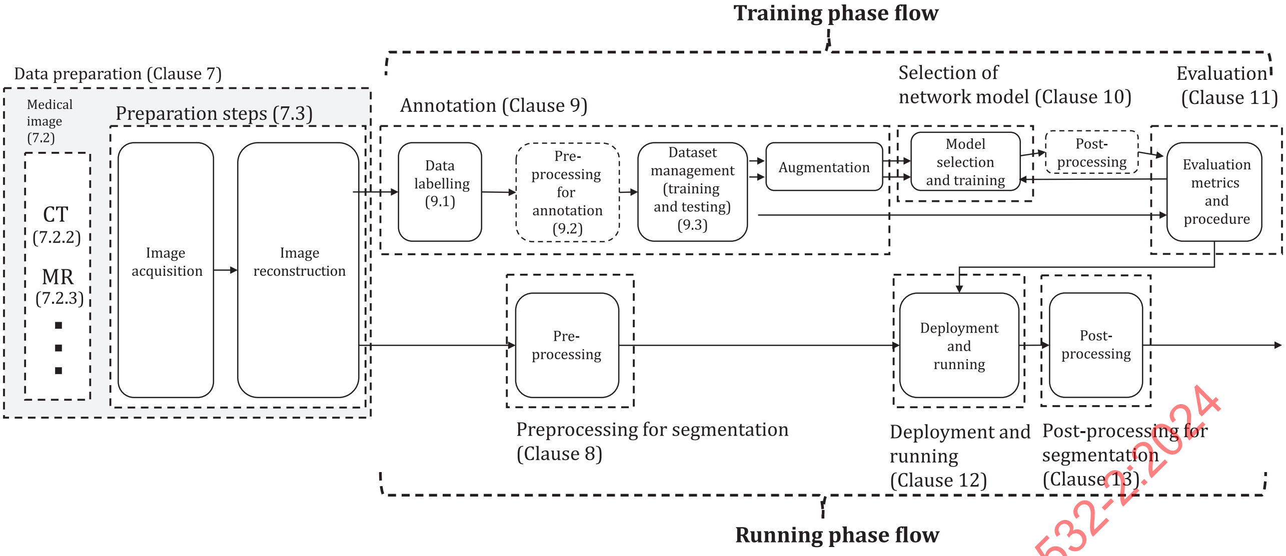


Figure 1 — Overall process flow of segmentation

The software developer should implement the optimized segmentation process for medical image-based modelling.

The considerations of the optimized segmentation process for medical image-based modelling should be referenced from this document.

6.2 Step 1: data preparation

The objective of the data preparation stage is to transform the raw data so that the segmentation algorithm can be applied. Detailed information is provided in Clause 7.

6.3 Step 2: preprocessing for segmentation

The objective of the preprocessing for segmentation stage is to normalize the data quality and induce consistent segmentation results. This stage is an important step that affects the ability of the network model to learn. Detailed information is provided in Clause 8.

6.4 Step 3: annotation

The objective of the annotation stage is to make the labelled training data fit for the learning of ML/DL-based segmentation network models. Detailed information is provided in Clause 9.

6.5 Step 4: selection of segmentation network model

The objective of the segmentation network model selection stage is to select the optimal segmentation network model. Detailed information is provided in Clause 10.

6.6 Step 5: performance evaluation

The objective of the performance evaluation stage is to calculate the agreement between the result of applying the segmentation technique and the ground-truth label. Detailed information is provided in Clause 11.

6.7 Step 6: model deployment and running

The objective of the deployment and running stage is to apply the optimally trained deep learning network model to the 3D printing modelling system for achieving the maximized segmentation performance in a real world environment. Detailed information is provided in Clause 12.

6.8 Step 7: post-processing for segmentation

The objective of the post-processing for segmentation stage is to refine the incorrect region after applying the segmentation method. Detailed information is provided in Clause 13.

7 Data preparation

7.1 General

The objective of the data preparation stage is to transform the raw data so that the segmentation algorithm can be applied.

7.2 Medical image

7.2.1 General

In the image acquisition phase, medical images are produced from devices, such as CT, MRI, PET/SPECT, US, and optical scanners. Normally, they are reconstructed from "raw" or source data produced by detectors. Generated medical image data from various devices should be stored in a standardized format (e.g. DICOM) for medical image processing.

7.2.2 CT scan

Computed tomography (CT) is a medical imaging technique that uses X-rays to generate 2D slice images of the body. These 2D slice images are typically stored in the DICOM file format. The intensity range of each pixel is [−1024, 3072], because most CT images use 12 storage bits per pixel. Each pixel has its own value according to the degree of transmitted radiation that passes through the body.

A CT scanner takes multiple radiographic projections and then uses an image reconstruction technique to generate a series of 2D image slices covering a specific portion of the human body. The resolution of the CT image is mainly dependent on pixel spacing and number of pixels, and partially dependent on the slice thickness of the 2D images.

For obtaining better quality medical images, standardized CT scanning parameter conditions should be used when scanning human bone. For example, in the case of the orbital bone, the minimum recommendation of CT scanning conditions is defined in Annex A.

7.2.3 MR image

Magnetic resonance imaging (MRI) equipment is medical electrical equipment which is generally intended for in vivo magnetic resonance examination of a patient. MRI is a medical imaging technique to form pictures of the anatomy and physiological processes of the body. MRI scanners use strong magnetic fields, magnetic field gradients and radio waves to create images of internal patient anatomy.

Standardized MRI scanning conditions or MRI scanning profiles should be considered.

7.3 Preparation steps

7.3.1 General

Since the data preparation stage consists of two steps, image acquisition and image reconstruction, the essential considerations for each step should be reflected.

7.3.2 Image acquisition

The image acquisition step can be defined as the action of acquiring a set of image data from a hardware source.

In order to create a good quality medical image-based 3D model for 3D printing, the following should be considered during the image acquisition step.

1) High-quality medical imaging equipment should be used as much as possible and images should be stored in the highest image quality.
2) Low-quality medical imaging equipment and protocols should not be used.
3) The acquisition of good quality medical images enables the production of good results.
4) Standardized image acquisition protocols should be used.
5) The image enhancement function provided by the device or manufacturer should be avoided as far as possible.

NOTE Specific CT scanning protocols are recommended for delicate or complex structures, such as the orbital wall, in order to optimize visualization of the entire structure. Annex A describes the CT scanning conditions for orbital bone segmentation as an example of a specialized CT scanning protocol. See also Annex B.

7.3.3 Image reconstruction

The image reconstruction step can be defined as the mathematical process that generates composite images from signals (or raw data) obtained during the image acquisition step.

The following points should be considered during the image reconstruction phase.

1) Image reconstruction can affect image quality.
2) There are many differences in the type and performance of CT image reconstruction kernels provided by equipment manufacturers which should be considered.

NOTE The CT image reconstruction kernel, also known as a convolution algorithm, refers to the process used to modify the frequency contents of projection data prior to back projection during image reconstruction in a CT scanner. This process corrects the image by reducing blurring. The kernel affects the appearance of image structures by sharpening the image. Different kernels have been developed for specific anatomical applications including soft tissue (standard kernel) and bone (bone kernel).[5]

8 Preprocessing for segmentation

8.1 General

The objective of the preprocessing stage is to normalize the image quality and induce consistent segmentation results.

The following should be considered during the preprocessing step.

1) The potential sequence of preprocessing steps may be considered as: denoising, interpolation, registration, organ windowing, followed by normalization, and potentially zero-padding to improve the performance of segmentation.

2) Inconsistent intensity range and pixel spacing of an image can have a significant impact on the performance of the segmentation method.

3) The normalization method may be required to ensure that the image training and testing data has a consistent intensity range and pixel spacing.

8.2 Intensity normalization

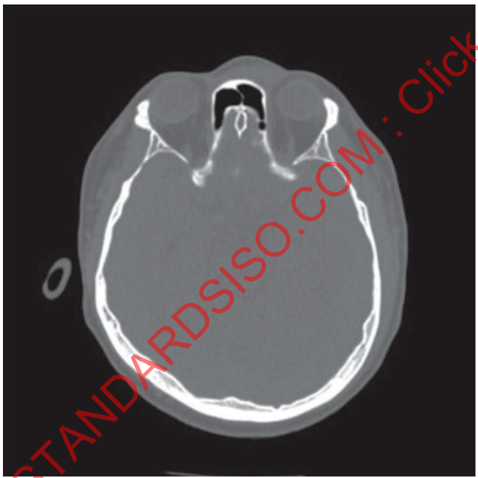
This subclause specifically refers to images acquired by CT as an example. Intensity normalization is also applicable, in principle, to images acquired using different technologies (see Figure 2).

Most CT images use 12 storage bits per pixel, but images used in DL approaches typically use 8 storage bits per pixel. Considering the intensity range of the human bone, intensity normalization is performed so that the HU intensity range [–200, 300] in the 12-bit CT image is transformed into the intensity range [0, 255], as shown in Formula (1):

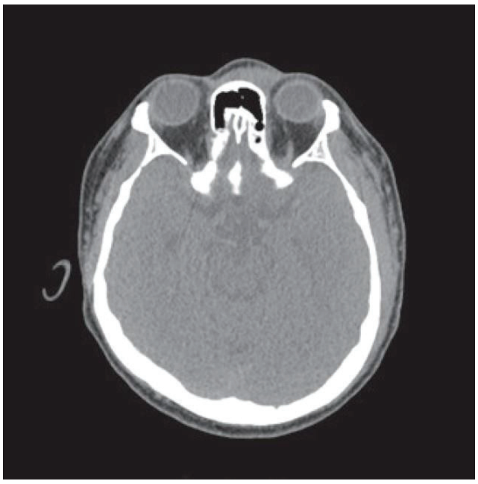
$$I_{\text{new}} = (I - I_{\min}) \frac{I_{\max,n} - I_{\min,n}}{I_{\max} - I_{\min}} + I_{\min,n} \tag{1}$$

where

I	is the pixel intensity value of the CT image;
$I_{\min}$	is the minimum intensity value of the orbit bones;
$I_{\max}$	is the maximum intensity value of the orbit bones;
$I_{\min,n}$	is the minimum intensity value of the new image;
$I_{\max,n}$	is the maximum intensity value of the new image.



a) Original 12-bit CT image (level: 400 HU, width: 1000)



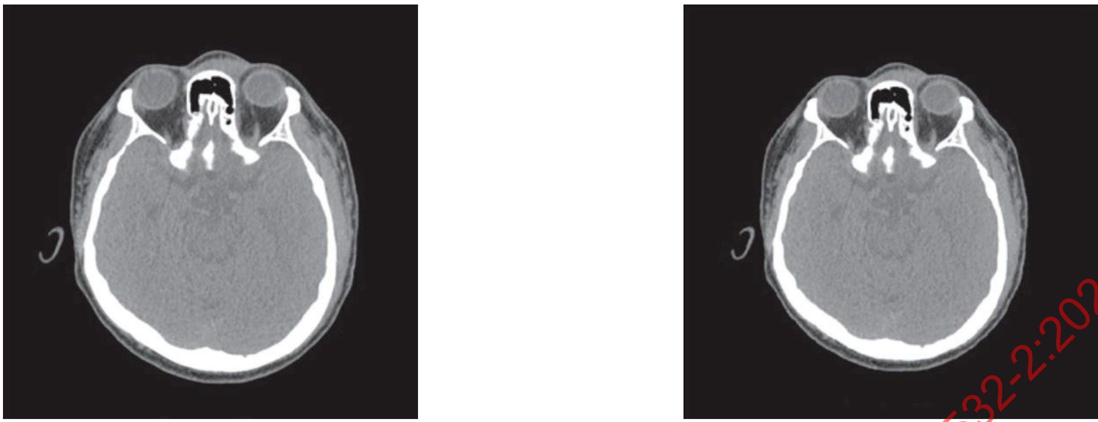
b) Transformed 8-bit CT image after intensity normalization

Figure 2 — Intensity normalization in head-and-neck CT image

8.3 Spacing normalization

This subclause specifically refers to images acquired by CT as an example. Spacing normalization is applicable, in principle, also to images acquired using different technologies (see Figure 3).

CT images can have different pixel spacing depending on patients. The difference between the pixel spacing means that the interval considered by one pixel is different. To normalize the pixel spacing, the pixel spacing normalization is performed using the maximum or minimum spacing of the dataset



a) 8-bit CT image with original pixel spacing (0.49 mm)

b) Transformed 8-bit CT image after spacing normalization (0.53 mm)

Figure 3 — Spacing normalization in head-and-neck CT images

9 Annotation

9.1 Data labelling

Labelled data is used as input data in the training stage to find the optimal parameters of the segmentation network.

Machine learning algorithms typically fall within the domain of supervised artificial intelligence and are designed to "learn" from annotated data. Machine learning models require large, diverse training datasets for optimal model convergence. This means that before an ML algorithm can be trained and tested, the ground truth (annotation) needs to be defined and linked to the image.[31]

For training of segmentation network model preparation, the following should be considered during the labelling data step.

1) Well-annotated datasets should be prepared. Well-annotated datasets are crucial to training accurate, generalizeable ML algorithm models.

2) There are many types of image annotation (e.g. closed curve, curve, ellipse, freehand, line, point, pointer, polygon, polyline, rectangle, text, text pointer, etc.). Any type of image annotation should be used in an interoperable way.

3) Annotation (or labelling) format should support annotation interoperability, such as AIM (annotation and image markup), NIfTI-1 data format or DICOM.

4) Consistent pixel values should be used for labelling. For example, the pixel value of the segmented human bone region can be set as 1 and the pixel value of the other region can be set as 0.

9.2 Preprocessing for annotation

The objective of the preprocessing stage is to normalize the data quality and to enhance the training ability of the segmentation network model.

The considerations for this stage are nearly the same as those described in Clause 8.

9.3 Dataset management (training and testing)

The objective of the data management stage is to effectively manage all data (training, validation, test) needed to optimize the actual performance of the segmentation network model.

Data management procedures should be conducted in the following order.

a) Collect the raw medical image data.

b) Apply data labelling.

c) Select a data sampling strategy for training/validation/test.

d) Split the datasets for training/validation/test.

The following should be considered during the data management step.

1) Collect as much high-quality data as possible. Performance increases logarithmically based on volume of training data.[32]

2) A well-designed dataset should be prepared which increases the quality of the resulting segmentation network model.

3) Even the most sophisticated segmentation network model cannot be trained with poor quality data.

4) A management process should be used for data quality.

5) Data should be assigned separately for training, validation and test purposes, as follows.

   — Training dataset: The set of data used to fit the model.

   — Validation dataset: The set of data used to provide an unbiased evaluation of a trained model fitting on the training dataset while tuning model hyperparameters.

   — Test dataset: The set of data used to provide an unbiased evaluation of a final trained model fitting on the training dataset.

9.4 Augmentation

The objective of the augmentation stage is to generate the augmented dataset which increases the amount of data by adding slightly modified copies of already existing data or newly created synthetic data from existing data.

There are different types of data augmentation techniques.[37] Increasing the dataset size by data warping or oversampling is generally used. Data warping augmentations transform existing images such that their labels are preserved, which includes augmentations such as geometric and color transformations, random erasing, adversarial training, and neural style transfer. Oversampling augmentations create synthetic instances and add them to the training set, which includes image mixing, feature space augmentations, and generative adversarial networks.

Data augmentation is useful as it improves the performance and outcomes of machine learning models by forming new and different examples for datasets. Data augmentation can provide benefits such as solving small dataset problems, preventing data scarcity, reducing data overfitting, and increasing the generalization ability.

10 Selection of network model

10.1 General

The objective of the network model selection stage is to find and select the optimal network model and model parameters for segmentation.

The broad success of DL has prompted the development of new image segmentation approaches leveraging ML/DL models. More than 100 DL-based image segmentation models have been developed (see Annex C) which have differences in network architecture selection, training data, loss function, training strategy and evaluation metrics.

The overall steps for finding and selecting the best segmentation model should proceed in the following order.

1) Divide the available data into training, validation and test dataset.

2) Select a model and set the training hyperparameters.

3) Train the model using the training set.

4) Evaluate the model using the validation dataset.

5) Repeat steps 2) through 4) using different model and training parameters.

6) Select the best model and train it using data from the training and validation dataset.

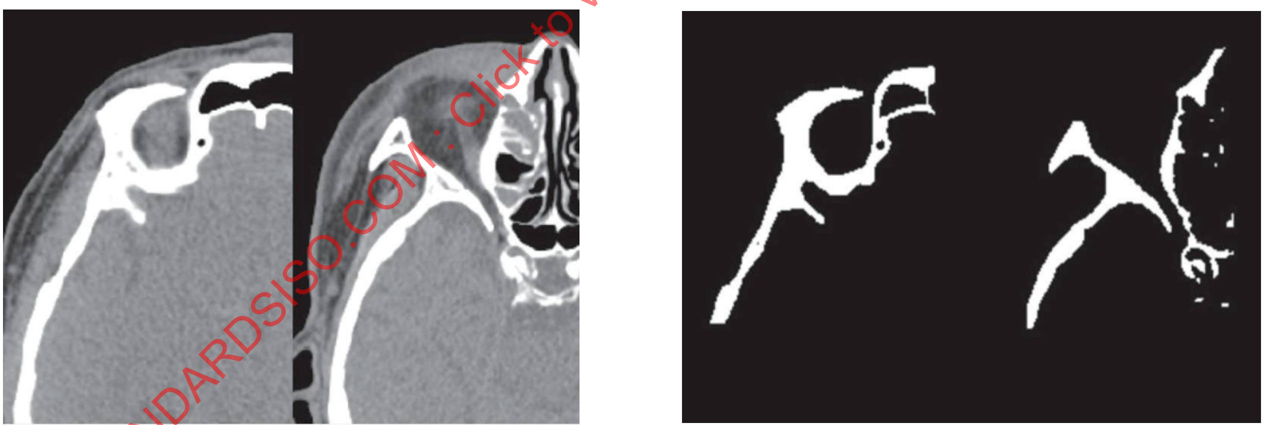
7) Assess this final model using the test dataset.

8) If external validation is required, assess the final model using the external validation dataset or open dataset

The evaluation method for finding and selecting the best segmentation model shall use the procedure and metrics defined in Clause 11.

10.2 Input patch

A 2D image patch is usually a .jpg image file format with 8 storage bits per pixel. The input patch consists of an image patch and a mask patch. The image patch is used in the training and test stages. The mask patch is used only in the training stage (see Figure 4).



a) Image patch for CT images **b) Mask patch for labelling data**

Figure 4 — Input patch

11 Evaluation

11.1 General

The objective of the evaluation stage is to assess the differences between the results of applying the trained segmentation model and the ground-truth labels.

Comparing results to evaluate the quality of segmentation is an essential part of finding the best segmentation model and algorithm. Annex D provides an example of how to consider overall segmentation performance in the case of the orbital bone.

There are different quality aspects in medical image segmentation based on which types of segmentation errors can be defined. Evaluation metrics for segmentation can be selected to indicate these errors, depending on the data and the target.[32][33]

Several classification metrics are derived from the following elements:

— true positive, T_P, is the number of pixels correctly identified as positive;

— true negative, T_N, is the number of pixels correctly identified as negative;

— false positive, F_P, is the number of pixels incorrectly identified as positive;

— false negative, F_N, is the number of pixels incorrectly identified as negative.

11.2 Evaluation metrics

Software developers should select the evaluation metrics that can best evaluate the quality and performance of their segmentation models, and use those evaluation metrics to select the best segmentation model.

Performance measurements for segmenting the human bone are calculated using sensitivity, specificity, accuracy, DSC and IoU. T_P, F_P, T_N and F_N refer to the detected number of pixels which are true positives, false positives, true negatives and false negatives, respectively.

Sensitivity is the ability to determine the labelled human bone areas correctly. Sensitivity measures the portion of true positive in the labelled human bone area of the ground truth, and is defined as:

$$s = \frac{T_P}{T_P + F_N}$$

where s is the calculated sensitivity value.

Specificity is the ability to determine the background correctly. Specificity measures the portion of true negative in the background of ground truth, and is defined as:

$$S = \frac{T_N}{T_N + F_P}$$

where S is the calculated specificity value.

Accuracy is the ability to differentiate the labelled human bone areas and background correctly. Accuracy measures the proportion of true positive and true negative in all ROI, and is defined as:

$$A = \frac{(T_P + T_N)}{(T_P + F_P + T_N + F_N)}$$

where A is the calculated accuracy value.

Dice similarity coefficient (DSC) is the most widely used metric when validating medical volume segmentation, and is defined as:

$$D = \frac{2T_P}{2T_P + F_P + F_N}$$

where D is the calculated dice similarity coefficient value.

Intersection over union (IoU) is the Jaccard index and is defined as:

$$i = \frac{T_P}{(T_P + F_P) + (T_P + F_N) - T_P}$$

where i is the calculated intersection over union value.

Additional evaluation metrics can exist which are not described here.

NOTE 1 Evaluation metrics are used in the evaluation procedure.

NOTE 2 ISO/IEC TS 4213 provides evaluation metrics for classification tasks.

NOTE 3 Evaluation metrics are also often used as the loss function for deep learning segmentation.[33][34][38][39]

11.3 Evaluation procedure

The evaluation procedure for assessing the performance of segmentation techniques should be performed in three steps using the evaluation metrics.

First, the ground-truth label should be obtained from human experts.

Second, the segmentation result should be obtained by applying the segmentation technique.

Third, the discrepancy between the segmentation result and the ground-truth label should be measured.

The method of measuring discrepancies should be counted as the difference between a segmented result and a correctly segmented image (ground truth) using evaluation metrics.

NOTE Often, the best method is to compare the segmentation result with the expert inter-observer variability. Comparison with the state-of-the-art is also important, as is the use of open source image data and reproducible code in order to allow objective comparison between methods. Properties of the segmented shape can also be evaluated against population norms (e.g. volume, curvature, etc.).

12 Deployment and running

The objective of the deployment and running stage is to apply the optimally trained DL network model to the 3D printing modelling system for achieving the maximized segmentation performance in a real work environment.

The following should be considered during the deployment and running stage.

1) For successful deployment and running, the risks and considerations should be eliminated early in the total segmentation process. Eliminating more risks and problems in the early stages of model selection and training reduces failure in the deployment and running stage.

2) It is highly recommended that the deployment process be standardized to ensure smooth testing and integration at multiple points.

3) Preparing the complete environment to access the right data via well-trained models on the right ML running environment is essential to the success of ML-based segmentation.

4) ML-based segmentation running environments should be tested and closely monitored in real time.

5) In a sophisticated experimental system, it should be possible to send test results back to the evaluation phase so that the model can be updated.

6) The system should always check monitored data quality and model performance.

13 Post-processing for segmentation

The objective of the post-processing stage is to refine the incorrect region after applying the segmentation method.

If over-segmentation/under-segmentation occurs in the segmentation result, post-processing techniques can be applied to refine the final result of the segmentation model.

For post-processing, 2D/3D connected component labelling (CCL) may be used to remove outliers, and a conditional random field (CRF) model may be used to further recover the boundary of a thin structure or weak edge.

Annex A
(informative)

CT scanning conditions for orbital bone segmentation

High-quality CT scanning is one of the key preconditions in obtaining accurate human bone segmentation. Table A.1 shows the CT scanning parameters and conditions for segmentation of the orbital bone and how to acquire the medical image data. These CT scan conditions are the minimum recommendations considered for the most difficult orbital bone segmentation. In the cases of other human bones, they can be adjusted to appropriate values.

Table A.1 — CT scanning conditions (minimum recommendations)

Term		**Condition**
Basic conditions	Region of interest	For orbital bone, the whole head can be scanned, and an optimal ROI can be extracted from 1 cm above the supraorbital foramen to infraorbital foramen.
	Bony position	Supine position
	Image resolution	At least 512 × 512
Scanning conditions	Tube voltage	80 kV to 120 kV (100 kV recommended)
	Tube current	250 mA to 400 mA (CT-AEC recommended)
	Scan slice thickness	Below 1.0 mm
	Scan (rotation) time	0.5 s to 1.5 s
Reconstructing conditions	FOV	100 mm to 150 mm
	Slice thickness	Below 1.25 mm
	Slice spacing	Below 1.0 mm
	Reconstruction kernel	Standard
	Output format	DICOM (digital imaging and communications in medicine) format, as most medical 3D modelling software is compatible with the DICOM format.

NOTE ISO 19233-1 recommends CT scanning conditions for the orthopedic joint prosthesis.[21]

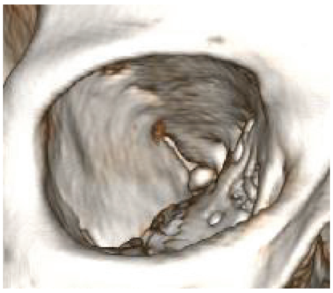
Annex B
(informative)

Characteristics of orbital bone segmentation from CT

The orbital wall is a four-sided pyramidal bony structure that includes the orbital roof, the orbital medial wall, the orbital lateral wall, and the orbital floor to protect the eyes and neighbouring nerves and consists of high-intensity cortical bones and low-intensity trabecular and thin bones [see Figure B.1 a)].

Orbital wall fracture is one of the most common fractures caused by trauma, especially in the orbital medial wall and orbital floor. Orbital bone segmentation is a prerequisite for orbital wall reconstruction of cranio-maxillofacial surgery to support the eye globe position and to restore the volume and shape of the orbit.[6]

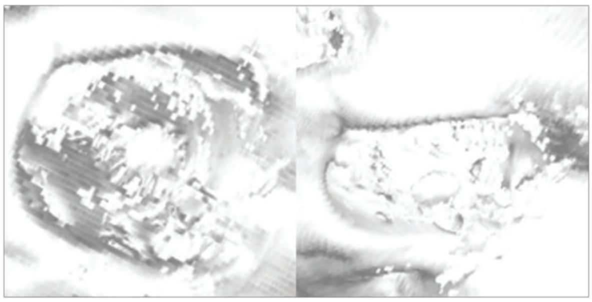
CT is a key imaging modality because it can provide detailed information about the orbital bone structure. However, it is difficult to segment the orbital bone with the conventional thresholding technique because the thin bone of the orbital medial wall and orbital floor have low intensity values similar to those of neighbouring soft tissues [see Figure B.1 b)].[7][8] Figure B.1 c) shows that when a thresholding technique with a threshold of 300 HU is applied to segment high-intensity cortical bone, low-intensity trabecular and thin bones are barely segmented. Figure B.1 d) shows that leakage into soft tissue around the orbital bone occurs when the threshold is set to 50 HU to segment low-intensity trabecular and thin bones.



a) Orbital bone structure

b) Cortical and thin bones in orbital floor and medial wall

c) Enlarged volume rendering view 1 (300 HU, segmenting cortical bone)



d) Enlarged volume rendering view 2 (50 HU, segmenting thin bones of orbital medial wall and orbital floor)

Figure B.1 — Characteristics of orbital bones in CT images

Several methods have been investigated to segment the orbital bone in head-and-neck CT images.

a) Semi-automatic segmentation: Thresholding is the simplest method of image segmentation, and thresholding can be used to create binary images. This method replaces each pixel in an image with an object pixel if the pixel intensity is greater than a specific orbital bone threshold value, or replaces it with a background pixel if the pixel intensity is less than a specific orbital bone threshold value. Because of the low intensity of trabecular and thin bones, manual editing is further required to segment the trabecular and thin bones. However, manual editing is time-consuming and causes inter-observer variability.

— Nystrom et al.[9] applied hysteresis thresholding to the orbital bone for segmenting the cortical bone and the trabecular bone connected to the cortical bone.

— Jesper et al.[10] applied semi-automatic orbit cavity segmentation using the subtraction of bone and air density masks, which used 400 HU for bone density mask and −600 HU for air density mask.

b) Deformable model based segmentation: Deformable models are curves or surfaces for segmentation in the image domain, which deform under the influence of internal and external forces to delineate the object boundary. The internal forces are defined such that they preserve the shape smoothness of the model, while the external forces are defined by the image features to drive the model towards the desired region boundary. However, there is a limit due to the difficulty of segmentation at the weak object boundary of a thin bone with a low intensity value similar to soft tissue.

— Maximilian et al.[11] segmented the orbital cavity for orbital volume measurement using model-based segmentation. In model-based segmentation, the orbital model is manually placed inside the orbital border and performs both growth and shrinkage processes using three basic forces (a growing force, a smoothing force and an image force).

— Lamecker et al.[12] generated an orbital statistical shape model (SSM) and segmented the orbit. However, the shape and size of the orbit varies, which limits clinical application.

c) CNN based segmentation: CNN based on cSegNet,[13] RefineNet[14] and U-Net[15] have been applied to medical image segmentation. The U-Net network, which is commonly used for medical image segmentation, is based on a fully convolutional network and consists of a contracting path and an expansive path, which gives it the U-shaped architecture. However, this method tends to ignore the thin bones of the medial wall of the orbit during learning because the orbital bone is mainly composed of cortical bone with high intensity value and thin bones with low intensity value occupy a small part.

— Zhao et al.[16] proposed a method of cranial segmentation after transferring MR images to CT images through a cascaded generative adversarial network with a deep-supervision discriminator (Deep-supGAN). However, the Deep-supGAN did not segment the trabecular and thin bones, but only the cortical bone.

Annex C
(informative)

Deep learning techniques

ML and DL have progressed rapidly in recent years. Various techniques of ML and DL have played important roles in the medical field, such as medical image processing, CAD, image analysis and image segmentation.

ISO/IEC 23053 defines an AI and ML framework for describing a generic AI system using ML technology. The framework describes the system components and their functions in the AI ecosystem.

In ISO/IEC 23053, ML is defined as the process of optimizing model parameters through computational techniques, such that the model's behaviour reflects the data or experience.

The broad success of DL has prompted the development of new image segmentation approaches leveraging ML/DL models.[24]-[30] According to a research survey result (see Reference [26]), there have been more than 100 approaches within 11 method categories proposed to date. These approaches differ in terms of network architecture selection, training data, loss function, training strategy and evaluation metrics. Each have their features, strengths and weaknesses. The 11 method categories are as follows:

1) fully-convolutional networks;

2) convolutional models with graphical models;

3) encoder/decoder-based models;

4) multiscale and pyramid network-based models;

5) R-CNN based models (e.g. segmentation);

6) dilated convolutional models and DeepLab family;

7) recurrent neural network-based models;

8) attention-based models;

9) generative models and adversarial training;

10) convolutional models with active contour models;

11) other models.

DL-based image segmentation is also now firmly recognized as a robust method of image segmentation in the medical field. References [24] and [27]–[30] analyze medical image segmentation using various DL techniques. Reference [24] investigates the types of network architectures and network training techniques used in medical image segmentation and summarizes the major challenges associated with training DL models. In other references within the Bibliography, detailed investigations and studies on the use of U-net and its variants,[27] multi-modality fusion,[28] and U-shaped networks[30] have been conducted for the segmentation of the medical images.

Annex D
(informative)

Considerations for overall segmentation performance

D.1 Overview

Figure D.1 shows an overview of orbital segmentation based on a convolutional neural network consisting of three major steps: data preprocessing, training and testing of multi-grey-level U-Nets for segmentation of orbital bone, and whole orbital bone mask generation integrating the segmented masks of cortical and thin bones.

Due to the wide range of intensities of the orbital bone, single-grey-level U-Net-based segmentation shows under-segmentation in low intensity thin bones. To overcome the under-segmentation of thin bones, the problem of segmenting the whole orbital bone is transformed into the problem of segmenting the cortical and thin bones and integrating two segmented bone masks.

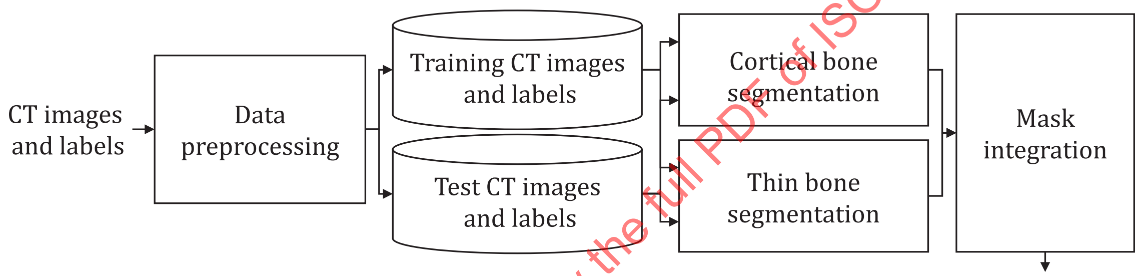
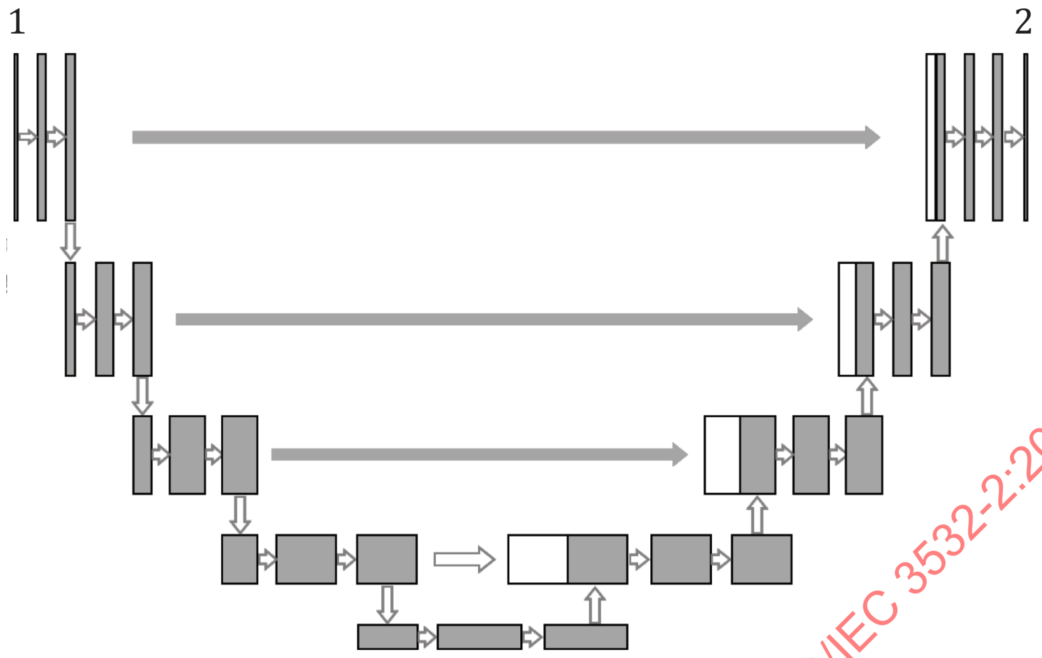


Figure D.1 — The pipeline of proposed multi-grey-level U-Nets for orbital bone segmentation from 3D head-and-neck CT images

Figure D.2 shows the basic structure of 2D U-Net to segment the orbital bones containing cortical bone, trabecular bone, and thin bone with different intensity values.

A 2D U-Net consists of a contracting path and an expansive path, which gives it the U-shaped architecture. The contracting path is a typical convolutional network that consists of repeated application of convolutions, each followed by a rectified linear unit (ReLU) and a maximum pooling operation. During the contraction, the spatial information is reduced while feature information is increased. The expansive path combines the feature and spatial information through a sequence of up-convolutions and concatenations with high-resolution features from the contracting path.



Key

1 input image

2 output segmentation map

Figure D.2 — U-Net structure

D.2 Data preparation

To adjust the image properties due to different acquisition protocols in each dataset, intensity and pixel spacing normalization is performed on the input images.

To consider the intensity range of the orbital bone, an intensity normalization is performed in which the intensity range [-200 HU, 300 HU] of the 12-bit CT image is transformed into the intensity range [0, 255] of the 8-bit image where a pixel with an intensity greater than 300 HU means a definite cortical bone and a pixel with an intensity less than -200 HU means low-intensity region such as air and fat.

To equalize the data of different pixel sizes, the pixel size of all images is changed to the maximum or minimum pixel size.

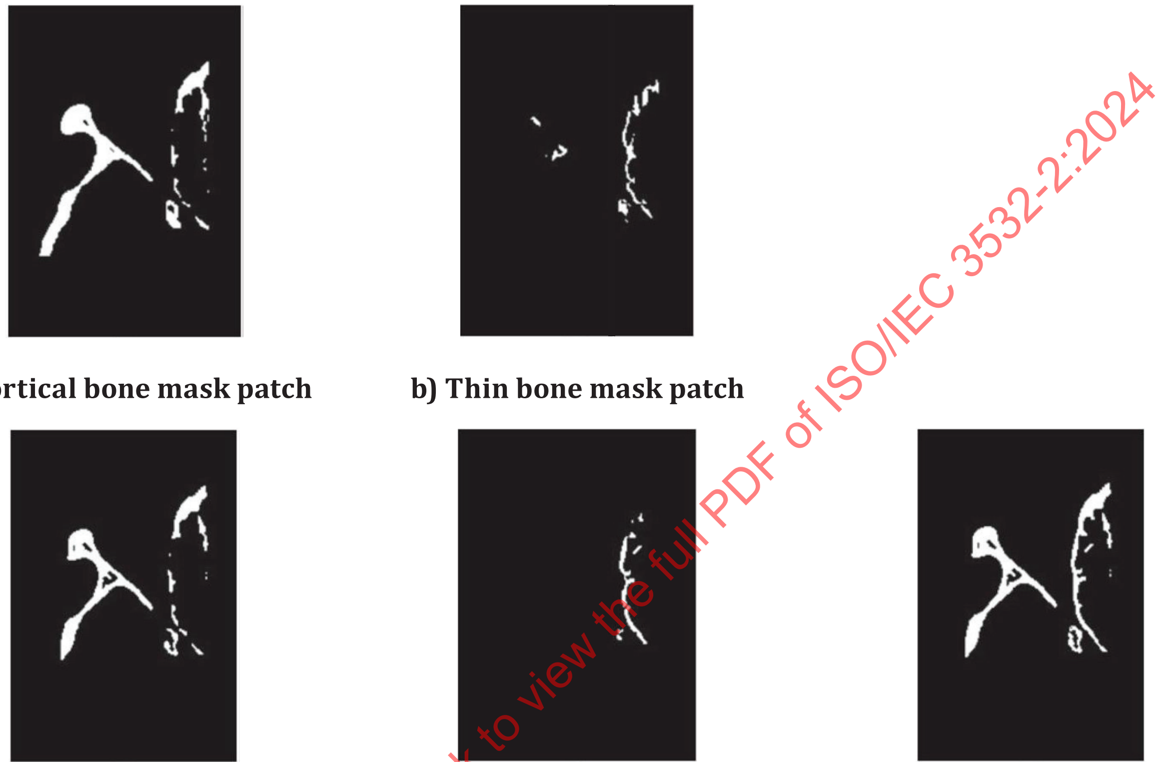
D.3 Multi-grey-level FCN-based segmentation and mask integration

To overcome under-segmentation of thin bones, a single orbital bone mask is divided into cortical and thin bone masks. In the training stage, two single-grey-level U-Nets are individually learned in these masks and each cortical and thin bone segmentation result generated in the test stage is integrated to obtain a whole orbital bone segmentation result.

In general, the cortical bone of the orbital bone has an intensity of 300 HU or more, and the thin bone of the orbital bone has an intensity of 200 HU or less. Thus, in a single orbital bone mask, the cortical and thin bone masks are constructed by selecting pixels with intensity greater than 200 HU and less than 300 HU and are constructed by superimposing the intensities to complement each other in the two masks.

Multi-grey-level U-Nets consist of two single-grey-level U-Nets using the same structures as a 2D U-Net architecture. A single-grey-level U-Net consists of a concatenating path and an expanding path, each path consisting of 6 layers. The concatenating path can help to extract more advanced features but it also reduces the size of feature maps, while the expanding path can help to give the localization information from the contraction path to the expansion path. The final output feature map of a single-grey-level U-Net is generated with only two classes (object, background) by a softmax function which gives the probabilities of the segmentation labels. In the training step, the loss function is calculated by the sum of the pixel-wise cross-entropy and the dice coefficient. To optimize multi-grey-level U-Nets, the Adam optimizer was adopted with a learning rate of 0.000 1. The epoch and the batch size were set to 100 and 20, respectively.

The two single-grey-level U-Nets are learned individually in two kinds of dataset for cortical bone masks (cortical bone, background) and thin bone masks (thin bone, background). With the two trained single-grey-level U-Nets, segmentation of the cortical bone and the thin bone is performed in the test image and the whole orbital bone mask is constructed by integrating the segmented cortical bone and thin bone areas.



a) Cortical bone mask patch

b) Thin bone mask patch

c) Cortical bone segmentation result

d) Thin bone segmentation result

e) Combined orbital bone segmentation result

Figure D.3 — Orbital bone segmentation results

D.4 Experimental plan

For the segmentation of orbital bones, head-and-neck CT images were acquired from a GE MEDICAL SYSTEMS Revolution EVO and SIEMENS Sensation 64[1)] with 0.41 mm to 0.59 mm pixel size, 0.63 mm to 1.00 mm slice thickness and 512×512 image resolutions. For the experiment, the datasets were divided into 30 datasets for training (1 776 images), 13 datasets for validation, and 19 datasets for test according to CT acquisition date.

The multi-grey-level U-Nets (Method C) was compared with two different methods: thresholding (Method A) and single-grey-level U-Net (Method B).

— Thresholding (Method A): Thresholding is the simplest method of image segmentation, and thresholding can be used to create binary images. This method replaces each pixel in an image with an object pixel if the pixel intensity is greater than a specific orbital bone threshold value (300 HU), or replaces it with a background pixel if the pixel intensity is less than a specific orbital bone threshold value (300 HU).

1) These trademarks are provided for reasons of public interest or public safety. This information is given for the convenience of users of this document and does not constitute an endorsement by ISO or IEC.