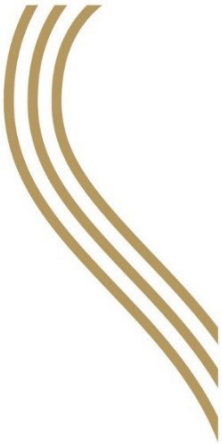


Dental implant recognition

by
Aviwe Kohlakala

Dissertation presented for the degree of
Doctor of Philosophy in Applied Mathematics
in the Faculty of Science at Stellenbosch University

Supervisor: Dr Johannes Coetzer

December 2023

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Name: Aviwe Kohlakala

Date: December 2023

Abstract

Deep learning-based frameworks have recently been steadily outperforming existing state-of-the-art systems in a number of computer vision applications, but these models require a large number of training samples in order to effectively train the model parameters. Within the medical field the limited availability of training data is one of the main challenges faced when using deep learning to create practical clinical applications in medical imaging. In this dissertation a novel algorithm for generating artificial training samples from triangulated three-dimensional (3D) surface models within the context of dental implant recognition is proposed. The proposed algorithm is based on the calculation of two-dimensional (2D) parallel projections from a number of different angles of 3D volumetric representations of computer-aided design (CAD) surface models. A fully convolutional network (FCN) is subsequently trained on the artificially generated X-ray images for the purpose of automatically identifying the connection type associated with a specific dental implant in an actual X-ray image. An ensemble of image processing and deep learning-based techniques capable of distinguishing between pixels that belong to an implant from those belonging to the background in an actual X-ray image is developed. Normalisation and preprocessing techniques are subsequently applied to the segmented dental implants within the questioned actual X-ray image. The normalised dental implants are presented to the trained FCN for classification purposes. Experiments are conducted on two data sets that contain the *simulated* and *actual* X-ray images in order to gauge the proficiency of the proposed systems. Given the fact that the novel systems proposed in this study utilise an ensemble of techniques that has not been employed for the purpose of dental implant classification/recognition on any previous occasion, the results achieved in this study are encouraging and constitute a significant contribution to the current state of the art, especially in scenarios where the proposed systems are combined with existing systems.

Opsomming

Diepleergebaseerde raamwerke het onlangs op bestaande staat-van-die-kuns stelsels in 'n aantal rekenaarvisietoepassings begin verbeter, maar hierdie modelle verg 'n groot aantal voorbeelde ten einde die modelparameters effektief af te rig. In die mediese veld is die beperkte beskikbaarheid van afrigdata een van die hoofuitdagings vir praktiese kliniese toepassings in mediese beeldvorming. In hierdie proefskrif word 'n nuwe algoritme vir die skep van kunsmatige afrigvoorbeelde vanuit driehoekgebaseerde driedimensionele (3D) oppervlakmodelle binne die konteks van tandimplantaatherkenning voorgestel. Die voorgestelde algoritme is op die berekening van tweedimensionele (2D) parallelle projeksies vanuit 'n aantal verskillende hoeke van 3D volumetriese voorstellings van rekenaarmatige ontwerp (CAD) oppervlakmodelle gebaseer. 'n Vol konvolusie-netwerk (FCN) word vervolgens op die kunsmatig-gegenereerde X-straalbeelde afgerig met die doel om die verbindingstipe geassosieer met 'n spesifieke tandimplantaat te identifiseer. 'n Ensemble van beeldverwerkings- en diepleergebaseerde tegnieke, wat in staat is om piksels wat tot 'n implantaat in 'n werklike X-straalbeeld hoort van dié wat tot die agtergrond hoort te onderskei, word ontwikkel. Normalisasie en voorverwerkingstegnieke word vervolgens op die gesegmenteerde tandimplantate in 'n bevraagtekende werklike X-straalbeeld toegepas. Die genormaliseerde tandimplantate word aan die afgerigte FCN voorgelê vir klassifikasiedoeleindes. Eksperimente word op twee datastelle, wat *gesimuleerde* en *werklike* X-straalbeelde bevat, toegepas ten einde die vaardigheid van die voorgestelde stelsels te beraam. Gegee die feit dat die nuwe stelsels voorgestel in hierdie studie van 'n ensemble van tegnieke gebruik maak wat nog nie voorheen vir die doel van tandimplantaatherkenning gebruik is nie, is die resultate behaal in hierdie studie baie bemoedigend en is dit 'n beduidende bydrae tot die huidige staat van die kuns, veral in scenarios waar die voorgestelde stelsels met bestaande stelsels gekombineer word.

Acknowledgements

I would like to express my sincere gratitude to the following people and organisations for enabling me to successfully complete this study:

- My supervisor, Dr Hanno Coetzer, for his invaluable insight, guidance, patience, unwavering support, immense knowledge and valuable critiques of this research work. This study would not have been possible without his input. It has been an absolute privilege to call him a mentor and be able to learn from his immense knowledge and expertise.
- The KU Leuven team, Jeroen Bertels and Professor Dirk Vandermeulen, for their valuable advice pertaining to this research and for also facilitating critical collaboration with industry (Medical Care NV) in Belgium.
- Medical Care NV and Nick Van Dooren for providing the anonymised database of X-ray images and valuable ideas concerning this work.
- The Postgraduate Funding Department of Stellenbosch University, for their support and financial assistance.
- The Ball family, for their financial support throughout my postgraduate studies.
- My family and close friends, for their unconditional love and support.

Lastly, I would like to dedicate this dissertation to my late parents, my mother, Bulie, and my father, Reuben. Thank you for your unconditional love.

Contents

List of Figures

List of Tables

List of Symbols

δ	Dirac delta function
η	Learning rate
γ	Momentum value
$\mathcal{R}_f$	Radon transform of function f
∇E	Gradient of the loss function
ϕ	Azimuth angle about the z axis
σ	Standard deviation
$\mathbf{b}$	Bias vector
$\mathbf{w}$	Weight vector
θ	Elevation angle about the x axis
θ_p	A single rotation about the pitch axis
θ_r	A single rotation about the roll axis
θ_y	A single rotation about the yaw axis
$d\ell$	Line integral along ℓ axis
p_j	Probability of jth class (output of softmax function)
x_i	Input for node i
y_i	Output for node i

List of Acronyms

2D	Two-dimensional
3D	Three-dimensional
CT	Computed tomography
CAD	Computer-aided design
CNN	Convolutional neural network
FCN	Fully convolutional network
RT	Radon transform
REC	Recall
ReLU	Rectified linear unit
ROI	Region of interest
PRE	Precision
SGD	Stochastic gradient descent
SGDM	Stochastic gradient descent with momentum

Chapter 1

Introduction

1.1 Background and motivation

The proficiency of deep learning algorithms to learn abstract and complex features for many computer vision applications such as object detection, image segmentation and image classification is well-documented. Recent advances in machine learning, especially with regard to deep learning, are assisting to identify, classify, and quantify patterns in medical images, therefore helping to diagnose and treat different diseases.

Deep learning-based algorithms in biomedical imaging have produced impressive diagnostic and predictive results in radiology and pathology research [3, 4]. A number of deep learning-based algorithms have also been investigated in various medical image analysis processes involving multiple organs, the brain, pancreas, breast cancer diagnosis and COVID-19 detection and diagnosis [5–10]. The well-documented success of deep learning in medical imaging has the potential for meeting dental implant recognition needs.

Dental implant recognition is crucial to multiple dental specialties, such as forensic identification and dental reconstruction of broken connections. Within the context of implant dentistry, implants provide promising prosthetic restoration alternatives for patients. In clinical practice where the dental records of a patient are not readily available, reliable categorisation of a dental implant previously inserted into the aforementioned patient's jaw is often challenging. Dentists often consider an X-ray image of the implant in question in order to discern the make, model, and dimensions of the implant. Based on this information, the connection type of the implant can be deduced. The dentist can subsequently order a suitable abutment and

artificial tooth to replace an existing one. Dentists may incur significant costs in scenarios where the wrong abutment or artificial tooth is ordered. A system that automates the classification of a dental implant based on an X-ray image of a patient's jaw may therefore be of great assistance to dental practitioners.

Deep learning-based models require a large number of training samples in order to effectively train the model parameters. Although large annotated image sets (like Caltech 256, PASCAL and ImageNet) exist, the generation and annotation of a large number of training images for a variety of new applications is labour intensive, expensive and requires many man-hours. Within the medical field collecting a large amount of image data from medical facilities can be difficult. The limited availability of training data with accurate annotations is one of the challenges faced when using deep learning to create practical clinical applications in medical imaging. Hence in this study a strategy of *artificially* generating a large number of training samples is investigated.

The purpose of this study is twofold. The first objective is to develop a proficient system that generates a large number of training samples efficiently within the context of dental implant recognition. The second objective is to develop a proficient deep learning-based dental implant recognition system, that is a system that automatically classifies a questioned dental implant within an actual X-ray image as one of the nine connection types investigated in this study. The proposed system should be proficient in the sense that it produces as few erroneous classifications as possible.

The remainder of this chapter is structured as follows: In Section 1.2, the background to the problem and a discussion of the key concepts of this study are supplied. An overview of the objectives of this study is presented in Section 1.3. This is followed by a brief synopsis of the proposed systems (see Section 1.4). In this dissertation, experiments are conducted on two data sets that contain the *simulated* and *actual* X-ray images (see Section 1.4.2). The abbreviated results are presented in Section 1.4.3, while the contributions of this study are listed in Section 1.5. An outline of this dissertation is provided in Section 1.6. Details about a publication that resulted from this research is provided in Section 1.7.

1.2 Key concepts

In this study, the development of a proficient end-to-end deep learning-based dental implant recognition system using synthetic X-ray images is proposed.

In this section, brief discussions on several of the fundamental concepts involved in such a development are presented.

1.2.1 Artificial generation of synthetic training samples

Synthetic (artificial) data is typically generated in scenarios where the availability of actual data is limited. Synthetic data should ideally imitate the original data by replicating important statistical properties, like the distribution of the data and the correlation between variables. Synthetic X-ray images are typically produced by projecting a three-dimensional (3D) computed tomography (CT) volume onto a two-dimensional (2D) image plane. A related but different and novel strategy for generating synthetic X-ray images will be followed in this study as summarised in Section 1.4.

1.2.2 Identification of dental implants

Dental implants have become a popular choice of treatment in replacing a patient's lost teeth or even entire dentitions. Each of the components used in this process are specific to the original implant down to the manufacturer, type and size, since most implant companies have a unique library of implant designs, sizes, and platforms. The amount of time it takes to complete an implant procedure from inception to final restoration can be as long as a year (for most patients) and can be costly.

There are a number of ways in which an unknown dental implant can be identified from a radiographic image, which includes the shape of the collar, groove and apex. The unique characteristics of an implant can be summarised as follows:

- **Dental implant interface.** The dental implant interface is the area associated with the implant connection where the prosthetic screw comes into contact with the abutment. There are a number of different types of abutment-implant interfaces, all of which are affected by the particular type of hardware used.

- **Dental implant body.** The dental implant body is surgically inserted into the jawbone of a patient to replace a tooth's root.

- **Dental implant screw.** The dental implant screw is the device used to secure the abutment to the implant. The screw is threaded and is typically tightened to the implant body through the gums to support the attached artificial tooth.

1.2.3 Region of interest detection

Image segmentation constitutes the process by which an image is subdivided into its constituent regions thereby delineating potentially meaningful areas for further processing, like object classification and detection. Image segmentation techniques can be based on either similarity or discontinuity properties associated with the image intensity values. Similarity-based thresholding is often achieved through thresholding, while machine learning algorithms that employ clustering are also based on this approach. Discontinuity-based thresholding is often based on edge detection [11].

1.2.4 Pattern recognition

Pattern recognition constitutes the automated assignment of input data to a specific pattern class, based on prominent features. From a machine-based perspective the pattern recognition process is generally subdivided into two key phases, namely *learning* and *testing*.

Learning

The learning phase, also referred to as training, is concerned with the construction of a mechanism that is able to interpret and analyse patterns associated with *known* input, in such a way that it should be equally adept at a similar interpretation and analysis of future patterns associated with a *questioned* input. This phase is characterised by learning features associated with the training set. There are two main learning paradigms, namely *supervised learning* which occurs in scenarios where the label of each training sample is known beforehand, and *unsupervised learning*, which occurs in scenarios where no class labels are available and the pattern classes have to be determined automatically, based on relative measures of their location and dispersion in feature space, which include clustering algorithms.

Testing

The testing phase is concerned with the automatic assignment of a class label to any questioned pattern submitted for classification. During the testing phase, a questioned pattern is presented to the trained model with the intention of extracting features for classification purposes. The feature set extracted from the questioned image is subsequently compared to the learned features acquired from the training data. The trained model typically outputs probability values for predicting the class of the queried pattern.

1.3 Objectives of this study

This research investigates the feasibility of deep learning for the purpose of automatically identifying the connection type associated with a specific dental implant within a questioned X-ray image. The main objectives of this study can be summarised as follows:

(i) Develop an efficient and proficient system for generating artificial training samples from triangulated 3D surface models within the context of dental implant recognition.

(ii) Develop a fully convolutional network (FCN) model, that is the *FCN-1* model, to be trained and validated on the artificially generated (simulated) X-ray images for the eventual purpose of assigning a questioned dental implant within an *actual* X-ray image to one of the *nine* connection types investigated in this study. The FCN-1 model therefore extracts (learns) the prominent *discriminative* features which describe the connection type associated with a specific dental implant from *simulated* X-ray images.

(iii) Develop efficient and proficient semi-automated and fully automated image segmentation systems capable of distinguishing between pixels in an actual X-ray image that belong to a dental implant from those belonging to the background. In the case of the semi-automated system suitable regions of interest (ROIs), that contain the dental implants, are manually specified (selected). Within the context of the fully automated system, that is a system that employs the *FCN-2* model, suitable ROIs are automatically detected through a deep learning-based technique.

(iv) Develop a system for the purpose of classifying a questioned dental implant within an actual X-ray image. The trained FCN-1 model is therefore applied to an *actual* X-ray image for the purpose of identifying the connection type associated with a questioned dental implant.

1.4 Overview of this study

In order to achieve the above mentioned objectives, a deep learning-based model is trained on a large number of *simulated* X-ray images generated through the proposed data generation strategy for the purpose of *automati-*

cally assigning a questioned dental implant within an *actual* X-ray image to a specific connection type.

In this section, the reader is provided with a brief overview of the proposed end-to-end deep learning-based dental implant recognition system. A discussion on the experiments conducted in order to evaluate the performance of the aforementioned system is provided and the experimental results are highlighted.

1.4.1 System design

The design of the system developed in this study is conceptualised in Figure 1.1. The proposed system can be divided into three parts, that is (i) the proposed strategy for artificially generating simulated X-ray images of dental implants, (ii) the strategy towards dental implant segmentation, and (iii) dental implant classification/recognition through machine learning.

Generation of simulated X-ray images

In this study, a *novel* algorithm for generating artificial training samples from triangulated 3D surface models within the context of dental implant recognition is proposed. The triangulated 3D surface coordinates of a specific dental implant are used to construct a 3D volumetric representation of the computer-aided design (CAD) surface model in question. Subsequently, 2D parallel projections of the volumetric representation are calculated from a number of different angles. The concept of X-ray CT for the purpose of reconstructing images from a series of projections [11–13] inspired the novel simulated X-ray data generation technique developed in this study. The implementation of the proposed data generation strategy is conceptualised in Figure 1.2.

Dental implant representation and modelling

A protocol is developed for extracting suitable features from simulated X-ray images for the purpose of identifying the connection type associated with a specific dental implant. The aforementioned protocol is based on extracting *learned* features through a deep learning-based technique. The deep learning-based model is therefore trained and validated on simulated X-ray images. The dental implant recognition protocol proposed in this study is conceptualised in Figure 1.3.

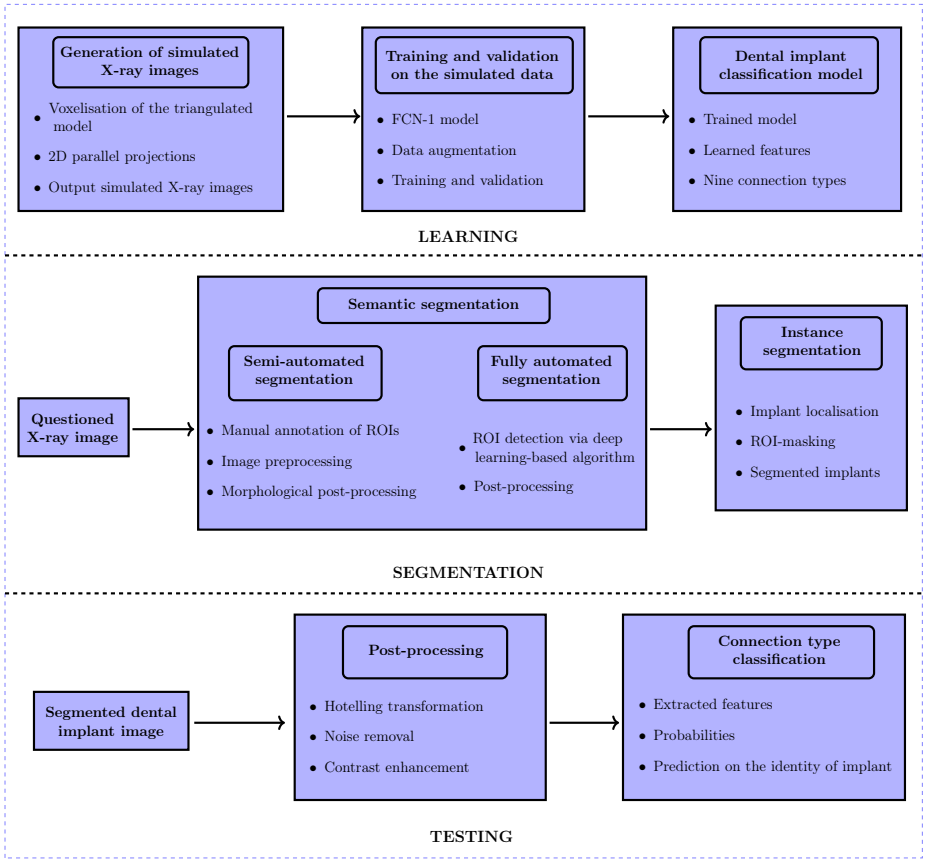


Figure 1.1: Schematic representation of the system developed in this study.

Image segmentation

A semantic segmentation protocol is developed for the purpose of detecting pixels associated with the dental implants in an actual X-ray image. The aforementioned semantic segmentation protocol can be dichotomised into semi-automated and fully automated systems. Semantic segmentation is performed on an actual X-ray image for the purpose of detecting pixels associated with the dental implants without differentiating implant instances. Instance segmentation is subsequently applied to the output mask image acquired through the aforementioned semantic segmentation algorithm in order to differentiate dental implant instances in an actual X-ray image. The image segmentation protocol developed in this study is depicted graphically in Figure 1.4.

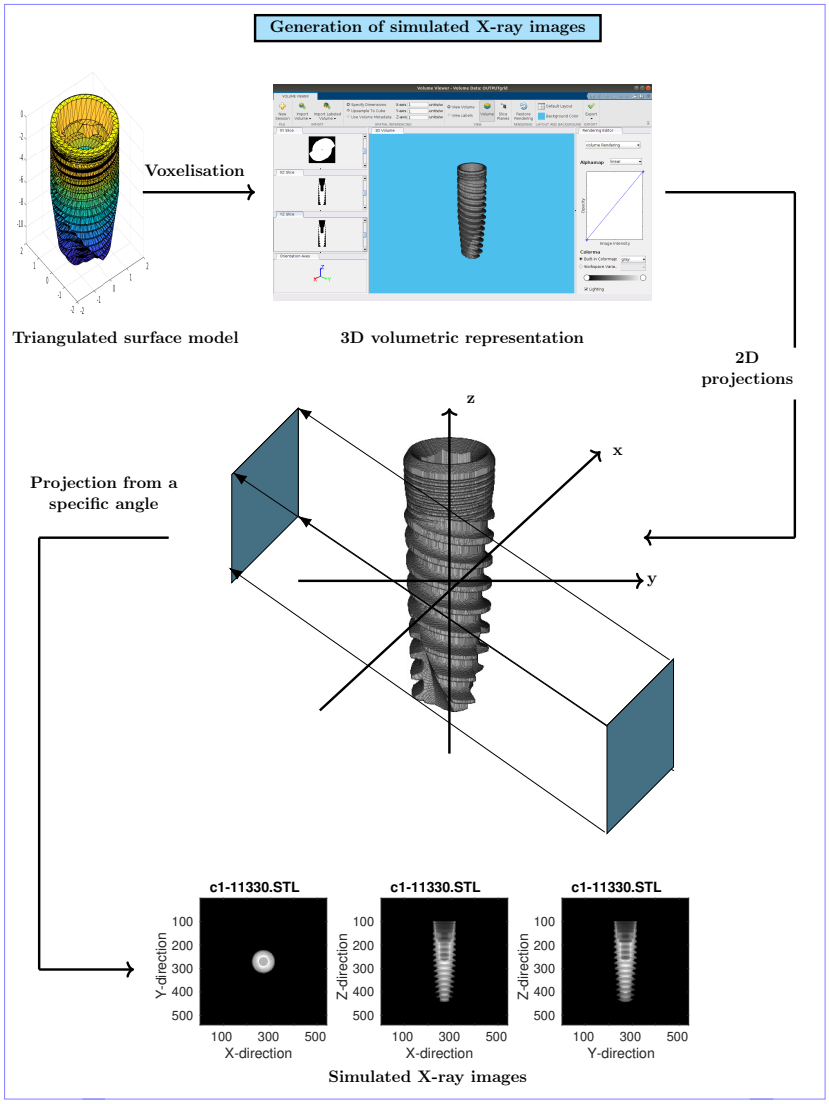


Figure 1.2: Framework of the simulated data generation protocol implemented in this study. The triangulated 3D surface model is converted into a 3D volumetric representation, where each voxel constitutes a cube with a value of one. Each simulated X-ray image is obtained by calculating a 2D projection of the volumetric representation from a specific angle.

Dental implant classification

The dental implant classification/recognition protocol developed in this study is based on a deep learning technique. This protocol learns discriminative

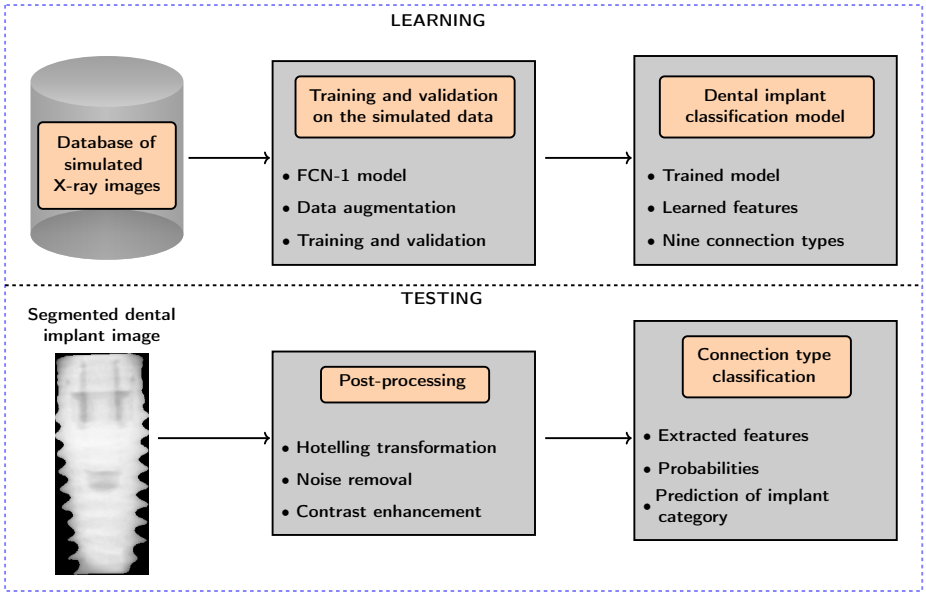


Figure 1.3: Schematic representation of the dental implant recognition protocol proposed in this study.

features for the purpose of distinguishing between different types of dental implants in X-ray images. The features are primarily based on the implant's external shape. A questioned actual X-ray image which contains only a single implant is presented to the trained FCN-1 model with the intention of extracting features for classification purposes. The aforementioned FCN-1 model therefore extracts features from the questioned dental implant for the purpose of predicting the connection type associated with the implant. This new set of features is subsequently compared to the learned features acquired from the artificially generated data. The proposed dental implant classification protocol is conceptualised in Figure 1.5.

1.4.2 Data

In this study, experiments are conducted on (1) an artificially generated data set and (2) an actual X-ray data set.

Artificially generated data set

In this study, the data set that contains simulated X-ray images of dental implants is generated from triangulated surface models, which constitute

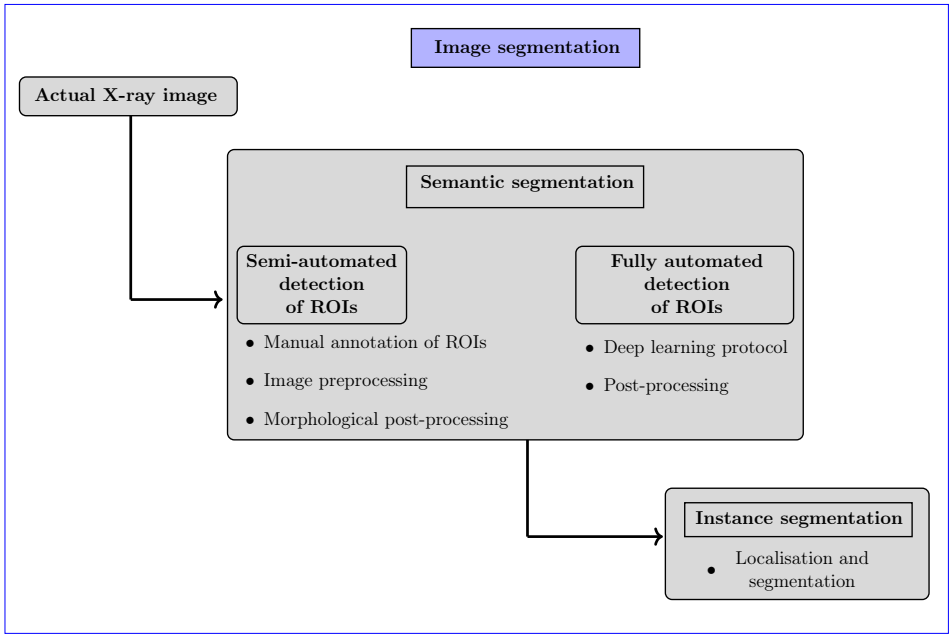


Figure 1.4: Conceptualisation of the image segmentation protocol implemented in this study.

standard triangle language (STL) files, engineered by the Make It Simple (MIS) manufacturing company. The triangulated 3D surface coordinates of a specific dental implant are used to construct a 3D volumetric representation of the model in question. A number of projections of the 3D volumetric representation (from different angles) are subsequently generated. Each projection is obtained by calculating a number of parallel ray-sums and constitutes a simulated X-ray image.

Actual X-ray images

The actual X-ray images considered in this study constitute a total of 483 labelled and unlabelled images, which contain implants inserted into either pig or human jaws. The database of X-ray images involving human jaws pertains to anonymous dental patients and was made available for this study by Medical Care NV. The database of X-ray images involving pig jaws was generated explicitly for this research by inserting the relevant dental implants into detached pig jaws obtained from butchers, after which the inserted implants were X-rayed with a similar device as the one used within the context

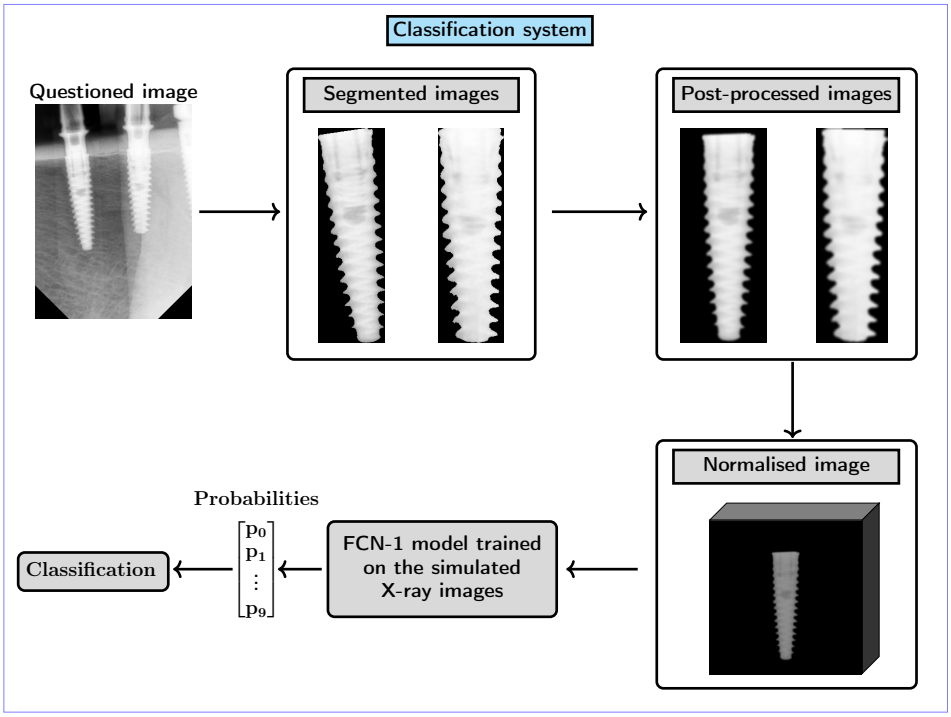


Figure 1.5: Conceptualisation of the proposed dental implant classification protocol. The input image is first segmented, after which the segmented implant is normalised and processed. Features are extracted from the input image through the proposed FCN-1 model and compared to the learned features acquired from the simulated data.

of the dental patients.

1.4.3 Abbreviated results

The systems developed in this study are subjected to a rigorous experimental evaluation protocol in order to gauge the proficiency of the proposed artificial data generation strategy, automated region of interest (ROI) detection algorithm, as well as the respective proficiencies of the semi-automated and fully automated systems within the context of dental implant classification/recognition.

Within the context of simulated X-ray images, a high accuracy of 96.56% and an average loss of 5.42% are achieved during validation across a 5-fold cross-validation protocol. This demonstrates that the proposed *FCN-1*

model (for automated connection type classification) effectively learns the prominent features associated with each artificially generated dental implant.

The performance of the proposed fully convolutional network (FCN) model, that is the *FCN-2* model for automated dental implant segmentation, is encouraging. The proposed system is able to classify the pixels associated with the dental implants (foreground) and those associated with the background with accuracies of 90.43% and 94.06% within the context of data set *one* (implants inserted into *pig jaws*) and data set *two* (implants inserted into *human jaws*) respectively. The aforementioned data sets are discussed in more detail in Section 7.2.

Within the context of implants inserted into *pig jaws*, accuracies of 74.63% and 71.72% are achieved for the proposed semi-automated and fully automated dental implant recognition systems respectively. Within the context of implants inserted into *human jaws*, accuracies of 69.76% and 68.67% are achieved for the proposed semi-automated and fully automated dental implant recognition systems respectively. Within the context of the semi-automated system the dental implants are *accurately* segmented from the actual X-ray images. This system therefore also serves as a benchmark in gauging the performance of the fully automated segmentation protocol.

1.5 Contribution of this study

This dissertation proposes a *novel* ensemble of techniques within the context of data generation and dental implant recognition. The feasibility of deep learning techniques for the purpose of automatically assigning a questioned dental implant within an *actual* X-ray image to a specific connection type is investigated. The key contributions of this dissertation can be summarised as follows:

- A novel framework for generating artificial training samples from triangulated 3D surface models within the context of dental implant recognition is proposed. The proposed algorithm is based on the calculation of 2D projections (from a number of different angles) of 3D volumetric representations of CAD surface models. The artificially generated X-ray images are subsequently employed for training and validating the FCN-1 model for the purpose of automatically identifying the connection type associated with a specific dental implant within an X-ray image.

- A framework of data augmentation strategies is implemented on the

simulated X-ray images in order to ensure that the artificially generated X-ray images are rendered more representative of the actual X-ray images. An ensemble of image preprocessing and normalisation strategies is performed on the actual X-ray images.

- A novel algorithm of object detection within the context of dental implant recognition which is capable of distinguishing between pixels in an actual X-ray image that belong to an implant from those belonging to the background is developed.
- Novel semi-automated and fully automated end-to-end deep learning-based systems for the purpose of classifying the connection type associated with a specific dental implant within an actual X-ray image are proposed.

1.6 Outline of this dissertation

The dissertation is structured as follows:

Chapter 2: Literature study. A concise overview of existing research within the context of the systems proposed in this dissertation is presented, that is a literature review on (i) strategies of artificially generating simulated data sets from three-dimensional models, (ii) strategies towards dental implant segmentation, and (iii) dental implant classification/recognition through machine learning.

Chapter 3: Generation of simulated X-ray images. The proposed algorithm for generating artificial training samples by calculating 2D projections (from a number of different angles) of a 3D volumetric representation of a dental implant is discussed. This 3D volumetric representation is inferred from a triangulated 3D CAD surface model of the implant in question.

Chapter 4: Image segmentation. The ensemble of object detection techniques proposed for the purpose of segmenting dental implants within an actual X-ray image is presented.

Chapter 5: Dental implant modelling. The proposed deep learning-based protocol for extracting suitable features from simulated X-ray images for the purpose of identifying the connection type associated with a specific dental implant in an actual X-ray image is described in detail.

Chapter 6: Dental implant classification. The ensemble of image processing and deep learning-based techniques employed within the context of dental implant classification/recognition in the process of identifying the connection type associated with a specific dental implant within a questioned actual X-ray image is discussed.

Chapter 7: Experiments. The data sets considered in this research and an outline of the experimental protocol employed in this dissertation are discussed. This is followed by exhaustive experiments that gauge and analyse the proficiency of the algorithms proposed in this dissertation.

Chapter 8: Conclusion and future work. The research conducted in this study, as well as the experimental results are analysed and placed into perspective, after which avenues for future research are explored.

1.7 Publication

The research presented in this dissertation was published in August 2022 in the *Journal of Medical and Biological Engineering and Computing* under the title *Deep learning-based dental implant recognition using synthetic X-ray images* [14].

The authors of this publication are Aviwe Kohlakala (the author of this dissertation), Johannes Coetzer (the supervisor of this dissertation), Jeroen Bertels (KU Leuven) and Dirk Vandermeulen (KU Leuven).

This research project was initiated in September 2018 when Johannes Coetzer visited KU Leuven. Jeroen Bertels is credited in this publication, since he introduced Johannes Coetzer to DeepVoxNet, a deep learning-based system maintained in KU Leuven for the purpose of analysing 3D medical images. Jeroen Bertels also provided valuable advice pertaining to the content of this publication and assisted with the editing process. Dirk Vandermeulen is credited in this publication, since he hosted Johannes Coetzer during his research visit in 2018 and engaged in valuable discussions. Dirk Vandermeulen also facilitated critical collaboration with industry (Medical Care NV) in Belgium.

The vast majority of the research presented in this publication (and in this dissertation) was conducted by the author of this dissertation, Aviwe Kohlakala.

Chapter 2

Literature study

2.1 Introduction

During the last few decades, deep learning has been adopted in a number of domains, which include medical image analysis. With the growing interest in deep learning-based algorithms and computational design in the biomedical field, the need for large, accessible and diverse data sets is increasing. The availability of large image data sets has been a crucial factor in the success of deep learning-based classification and detection methods. Within the medical field, the adoption of deep learning-based algorithms is often challenging, since large labelled data sets is rarely available, while privacy is another major concern.

The artificial generation of synthetic training samples is often crucial in evaluating algorithms during medical software development. Data scarcity in the medical field is mainly due to the fact that patient data is deemed private. Hospitals and physicians are therefore often not allowed to freely distribute this data, which forces system developers to rely on the acquisition of synthetic (artificially generated) data. Artificially generated images should be similar to the actual (real) images in terms of the features that can be potentially extracted from the objects being imaged. This implies that the artificially generated images should capture the same anatomical structures and characteristics as the actual (real) images, and that the features of interest in the actual images should be accurately represented.

In this chapter, a concise overview of existing research within the context of artificial generation of training samples from three-dimensional (3D) surface models and dental implant recognition is presented. The discussion provided on the aforementioned systems is in some way related to the

work presented in this dissertation. The systems are therefore categorised into (1) algorithms proposed for the artificial generation of synthetic training samples (see Section 2.2), (2) techniques proposed for the purpose of segmenting dental implants (see Section 2.3), and (3) proposed feature matching and verification paradigms for the purpose of dental implant classification/recognition (see Section 2.3).

Since most existing dental implant-based classification/recognition systems have not been evaluated on the same data sets than those considered in this dissertation (mainly due to privacy issues in the medical field), it is not possible to *directly* compare the reported proficiency of these systems to those proposed in this dissertation.

2.2 Generation of simulated data sets from three-dimensional models

The availability of large training sets is crucial in building proficient deep learning-based models. The use of synthetic data in a number of computer vision applications has provided a means of bridging the gap between simulated and actual training data. A number of algorithms for generating training samples from 3D models within the context of image classification have been investigated.

Tremblay et al. [15] proposed an approach of applying simulated data generated from 3D models to the problem of object detection. The proposed technique of transforming 3D models into two-dimensional (2D) images is based upon domain randomisation, that is through a process in which parameters like lighting, pose, and object textures are varied. The proposed framework involves the placing of a random number of the objects (such as cars) in a 3D scene at random positions and orientations. To better enable the network to learn to ignore objects in the scene that are not of interest (called *flying distractors*), a random number of geometric shapes are added to the scene. Random textures are subsequently applied to both the objects of interest and the flying distractors. A random number of lights of different types are inserted at random locations, while the scene is rendered from a random camera viewpoint, after which the result is composed over a random background image. The resulting images, with automatically generated ground truth labels (e.g. bounding boxes), are then used for training the neural network.

Rozantsev et al. [16] proposed a synthetic data generation technique

for the purpose of training object detectors. The proposed algorithm estimates the parameters required for generating training images from a given 3D model of the target object. These parameters can then be reused to generate an unlimited number of training images of the object of interest in arbitrary 3D poses, which can then be used to increase classification performance. Post-processing methods that involve object boundary and motion blurring, as well as the addition of noise and the variation of material properties, are implemented on the synthetic images in order to maximise their similarity to the actual (real) images. The effect of object boundary blurring is implemented by applying Gaussian blurring along the object boundaries after the object image has been overlaid on the background image. For motion blurring, anisotropic Gaussian blurring is applied to the pixels of the object in the direction of its motion. Gaussian noise is added to the synthetic images, while their material properties are varied by changing the weight of the diffuse reflection, which not only varies the color of the object, but also introduces some diffuse lighting effects.

Yu et al. [17] proposed a novel synthetic data generation technique for a vehicle data set. A 3D model of a specific make and model of a car is used to generate different views of the car in question. This is achieved by changing the camera distance, direction angle β within an interval $\beta \in [0°, 360°)$, and over angle α within an interval $\alpha \in [0°, 90°)$. The distance parameter refers to the distance between the camera and the 3D model. By changing the camera position, different views of the vehicle are generated. The direction angle β refers to the direction that the camera is facing in the horizontal plane. This parameter determines the horizontal orientation of the camera view. The over angle α determines the vertical orientation of the camera view in the vertical plane. By adjusting these three parameters, a wide range of different 2D images can be generated from the 3D model.

Teixeira et al. [18] proposed a novel algorithm for generating synthetic X-ray images of a patient by considering the patient's 3D surface geometry. The synthetic X-ray images are parameterised and can be manipulated by adjusting a set of body markers which are generated during the X-ray image prediction process. The proposed framework generates multiple synthetic X-ray images by varying the surface geometry and perturbing the aforementioned parameters. This is achieved by training a pair of networks inspired by the so-called U-Net network [5]. These networks are composed of four levels of convolutional blocks, where each block consists of three consecutive convolutional layers, followed by batch normalisation and ReLU layers. Each network contains 27 convolutional layers, with 32 filters in each layer. The

first network learns to generate the full image from a partial image and a set of parameters, while the second network learns to estimate the parameters given the full image.

Sui et al. [19] proposed a novel framework for generating synthetic training images of sewer pipes to improve the accuracy of automatic sewer pipe defect detection. The framework first creates a 3D model of the sewer pipe using computer-aided design (CAD) software. Once the 3D model has been created, synthetic images of the pipe are generated by simulating various imaging conditions, such as changes in lighting, camera position, and sensor settings. Different types of defects, with varying levels of severity, are included in the images. Additional data augmentation is performed on the synthetic images in order to generate a wide range of diverse images. The aforementioned augmentation techniques include the introduction of random noise and varying blur levels to simulate different types of real-world conditions such as cracks, corrosion, and deformations.

A novel technique for generating synthetic images from 3D volumetric data of the brain is proposed by Pinaya et al. [20]. The proposed technique is based on a generative latent diffusion model (LDM). The model is first trained on a large data set of actual (real) brain images to learn the statistical patterns and relationships within the data. The trained model is subsequently used to generate new synthetic brain images by randomly sampling from the model's learned distribution. Conditioning variables are used to effectively control the synthetic data generation protocol. These conditioning variables include information about the presence or absence of certain brain structures, the level of activity in specific brain regions, or other relevant features of the brain. By conditioning the generation process on these variables, the LDM is guided to generate synthetic images that represent different states or conditions of the brain. By employing this technique 100 000 synthetic brain images are generated.

Moreira et al. [21] proposed an algorithm for generating cone-beam computed tomography (CBCT) of dental implants from 3D surface models in order to accurately determine the *pose* of a dental implant. The proposed pose estimation algorithm is accomplished through a three-step approach: (i) a region of interest (ROI) is manually specified using two operator-defined points on the implant's main axis, (ii) a simulated CBCT volume of the known implant model is generated through Feldkamp-Davis-Kress (FDK) reconstruction and is coarsely aligned to the defined axis, after which (iii) a voxel-based rigid registration is performed to optimally align both patient and simulated CBCT data. The implant's pose is therefore extracted from

the optimal transformation. The framework is achieved as follows: Starting with a triangular polygon mesh of the implant model, a ray intersection method is used to obtain a binary volume of the implant of interest. The implant mesh is ray-traced in the x, y and z directions, after which a CBCT reconstruction algorithm is used to obtain the final simulated volume through the FDK algorithm.

The strategy proposed by Moreira et al. [21] is therefore based on the implementation of the FDK algorithm, which constitutes an approximation of filtered backprojection from *cone-beam* projections with a circular orbit about the X-ray source.

In *this* study the novel algorithm proposed for generating artificial dental implants is based on 2D projections that involve *parallel beams*. A more detailed description of the proposed strategy is provided in Chapter 3.

2.3 Dental implant detection

A number of semi-automated and fully automated systems have been proposed for the purpose of segmenting dental implants. The aforementioned segmentation algorithms are geared towards the detection of the ROIs which contain the dental implants. In this section an overview of the research that has been conducted on semi-automated and fully automated dental implant segmentation is presented.

Morais et al. [22] proposed a dental implant segmentation framework which uses an active contour protocol for optimally defining the dental implant's boundaries. The active contour protocol comprises of two steps: (i) initialisation of the desired model and (ii) evolution/propagation of the model. The optimal contour is estimated through the minimisation of an energy function that reflects the target shape. An approximate shape of the target is required to initialise the segmentation process, making the final implant boundary dependent on such an initialisation. The major axis of a dental implant is manually defined using a user-selected tip and base. Subsequently, 20 equally spaced lines perpendicular to the major axis are employed during the contour estimation step by implementing an optimal edge detection strategy. The final contour is achieved using the active contour protocol. In order to evaluate the proposed segmentation strategy, the semi-automatically detected contour is compared to a ground truth generated by a single expert observer. In their study, 125 different implant images of size 676×934 pixels are considered for the proposed segmentation strategy. The dice metric, mean absolute distance (MAD) and Hausdorff distance are employed to

quantify the differences between the contours (semi-automatic and manual), where a dice metric of 0.97±0.01 pixels, a MAD of 2.24±0.85 pixels and a Hausdorff distance of 11.12±6 pixels are respectively obtained.

A fully automated segmentation algorithm based on image preprocessing, followed by adjusted and trained active shape models is proposed by Cunha et al. [23]. The proposed segmentation method comprises of two steps: (i) image preprocessing which combines denoising filters, morphological operations and histogram threshold techniques and (ii) final segmentation, that involves adjusted and trained active shape models for detecting the precise location of the intended structures. The second phase of the segmentation framework involves the tuning and training of three active shape models, two of them devoted to the detection of the dental implant contour (one for each side of the implant) and the other one for detecting the bone line contour. In order to gauge the proficiency of the proposed segmentation strategy, the resulting measurements are compared to manual measurements made by experts on representative radiographs from patients. The calculated intraclass correlation coefficient is 75% and Bland-Altman analysis shows that 95% of the values are within the limits of agreement.

Pauwels et al. [24] proposed a fully automated dental implant segmentation technique based on contour detection and particle counting. The proposed technique is based on a three-step automated segmentation protocol. Firstly, images are pre-thresholded using T_{min}. Next, Sobel operators are applied for the purpose of contour detection. The edge image is subsequently thresholded with a second threshold value T_{edge} and converted into a binary mask. After thresholding with T_{edge}, the binary image is subjected to particle counting.

The aforementioned state-of-the-art algorithms for segmenting dental implants are mainly based on image processing techniques. The semi-automated segmentation algorithm proposed by Morais et al. [22] is, for example, based on an active contour protocol. Within the context of fully automated dental implant segmentation, an active shape model and contour detection technique are respectively investigated by Cunha et al. [23] and Pauwels et al. [24].

The *semi-automated* segmentation strategy proposed in *this* study is based on image processing techniques which include thresholding, connected component analysis and morphological post-processing. The *fully automated* segmentation strategy proposed in *this* study is based on a deep learning approach. A detailed description of the proposed segmentation algorithms is provided in Chapter 4.

2.4 Dental implant recognition

The identification of dental implants in X-ray images is often challenging due to the large number of different implant models. A certain degree of expertise is required to identify and distinguish between the various dental implant types available on the market. The accurate classification of an implant model is important in selecting a suitable replacement when the existing abutment and/or artificial tooth has been lost or damaged. An overview of the research that has been conducted on dental implant recognition is now presented.

A k-nearest neighbour (KNN) algorithm is proposed by Morais et al. [22]. In this approach the questioned implant is compared to all the labelled implants available in the reference database. The KNN classifier computes the distance between the geometric features of the questioned implant and all the available ones in the database. The reference implant associated with the smallest distance is considered to be the optimal solution. The proposed strategy is based upon a simple methodology, which employs an intensive search to directly compare the questioned implant with all labelled objects available in the reference database. However, the main disadvantage associated with this strategy is that it is time-consuming when a large number of features are considered.

A number of studies [25–28] have investigated deep convolutional neural networks (DCNN) for the purpose of classifying dental implants.

Lee et al. [25] employed the Neuro-T version 2.0.1 (Neurocle Inc., Seoul, Korea) tool for the purpose of automatically selecting the best performing model with optimal hyper-parameters. The aforementioned DCNN-based model architecture consists of 18 layers and does not contain any dropout layers. The adaptive momentum estimation (Adam) algorithm is used during network training, while L2 regularisation is employed as the baseline for transfer learning so as to optimise the weights and improve the output power by adjusting the hyper-parameters. The accuracy of the proposed system is compared against that of dental professionals using dental radiographic images collected from three dental hospitals. A total of 11 980 panoramic and periapical radiographic images with six different types of dental implant models are divided into a training and test set of 9 584 and 2 396 images, respectively. In order to evaluate the proposed DCNN-based model, 180 images are randomly selected from the test set, after which the performance metrics are compared to those of dental professionals. The area under the curve (AUC), Youden index, sensitivity, and specificity are employed as performance eval-

uation measures. The proposed network achieves an AUC of 95.4%, Youden index of 80.8%, sensitivity of 95.5%, and specificity of 85.3%.

Hadj et al. [26], Sukegawa et al. [27] and Kim et al. [28] investigated DCNN systems with transfer learning strategies for the purpose of classifying different dental implant models.

Hadj et al. [26] applied preprocessing techniques and transfer learning to a pretrained GoogLeNet Inception V3 network for the purpose of identifying the brand and model of a dental implant from a radiograph. In their study experiments are conducted on a total of 1 206 dental implant radiographic images of three different brands and six different models. The data is partitioned into training, validation and test sets, where 80% of the data is assigned to the training and validation sets, while 20% is assigned to the test set. The accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), receiver operating characteristic (ROC) curve, and AUC are employed as performance evaluation measures. The proposed network achieves a diagnostic accuracy of 93.8%, sensitivity of 93.5%, specificity of 94.2%, PPV of 92% and NPV of 91.5%.

Sukegawa et al. [27] proposed a DCNN system with transfer learning strategies for the purpose of classifying different dental implant brands. VGG networks are employed for transfer learning and fine-tuning purposes. Experiments are conducted on a total of 8 859 implant images of 11 implant systems from digital panoramic radiographs. The accuracy, precision, recall, ROC curve and F_1 score are employed as performance evaluation measures. The proposed network achieves a recall of 89.4%, a precision of 91.3%, an accuracy of 92.7% and an F_1 score of 90.2%.

Kim et al. [28] employed YOLOv3 for transfer learning and fine-tuning purposes in order to classify implant fixtures. Experiments are conducted on a total of 355 periapical radiographs of implant fixtures which constitute the Superline (Dentium Co. Ltd., Seoul, Korea), TS III (Osstem Implant Co. Ltd., Seoul, Korea), and Bone Level Implant (Institut Straumann AG, Basel, Switzerland) dental models acquired from patients who underwent dental implant treatment. The data was split into a training and test set at a ratio of 8 to 2. The performance of the network was evaluated on the test set, while the sensitivity, specificity, and accuracy were employed as performance evaluation measures. The network achieves a sensitivity of 94.4%, a specificity of 97.9%, and an accuracy of 96.7%.

In the aforementioned systems the classification of a dental implant is based on the type of dental implant *model*. The protocol proposed in *this* study delves deeper by investigating the classification of dental implant *con-*

nection types. The dental implant connection interface is a key feature to consider when choosing a replacement model for the abutment. The implant connection interface corresponds to the connection site where the dental implant body connects to the abutment. The accurate identification of the connection type of an implant is important in order to ensure the correct abutment is selected during the restorative phase. The dental implant connection interface can generally be described as either a conical, internal hexagonal or external hexagonal connection. The geometry of the connection can be further characterised as either a narrow, standard or wide platform. A more detailed description of the connection types investigated in this study is provided in Section 5.3.

Chapter 3

Generation of simulated X-ray images

3.1 Introduction

In scenarios where the dental records of a patient are not readily available, reliable categorisation of a dental implant previously inserted into the aforementioned patient's jaw is often challenging. In order to discern the connection type associated with a questioned dental implant, dentists often have to manually analyse an X-ray image of the implant. In scenarios where an incorrect abutment or artificial tooth is ordered significant costs may be incurred. Hence, a system that automates the classification of a dental implant based on an X-ray image of a patient's jaw may be of great assistance to dental practitioners. Based on the assumption that a large digital database of popular dental implant models is available, such an automated system may significantly reduce the number of possibilities which can save time and improve accuracy by eliminating the need for manual classification of the implant type.

The fact that a specific dental implant model may have a wide range of widths and lengths, while the connection type remains unchanged, further compounds the problem. Dental implant recognition in scenarios where the dental records of the patient in question are not readily available, is therefore currently a costly, time-consuming, and challenging task. The traditional protocol for dental implant recognition involves an entirely user-dependent process, where the expert compares a two-dimensional (2D) X-ray image of the dental implant with a generic database that consists of a large number of implant models with different designs. Due to a large number of available

implants and the potential similarity between them, semi-automated or fully automated frameworks that aid in the recognition of dental implant models are essential.

The accurate classification of an implant model is important in selecting a suitable replacement when the existing abutment and/or artificial tooth has been lost or damaged. The dentists involved in the initial stage of surgically inserting the dental implants into a patient's jaw, do not necessarily provide maintenance for dental implant fixtures. It is therefore essential that dentists and technicians can accurately identify the inserted implant model and its associated connection type.

In this study a novel algorithm for generating artificial training samples from triangulated three-dimensional (3D) surface models within the context of dental implant recognition is proposed. The triangulated 3D surface coordinates of a specific dental implant are used to construct a 3D volumetric representation of the computer-aided design (CAD) surface model in question. The concept of X-ray computed tomography (CT) for the purpose of reconstructing images from a series of projections [11–13] inspired the simulated X-ray data generation technique proposed in this study.

In this chapter the proposed protocol for generating artificial training data is discussed in detail. The aforementioned protocol is summarised in Figure 3.1. In Section 3.2, the principles of X-ray CT are discussed in detail. Subsequently, in Section 3.3, the voxelisation algorithm implemented in order to transform the triangulated surface model into a 3D volumetric representation of the model in question is discussed. In Section 3.4, the proposed strategy of calculating 2D projections of a 3D volumetric representation is described in detail.

3.2 X-ray computed tomography (CT)

Computed tomography (CT) is an imaging technique that involves propagating X-ray beams through a 3D object after which the attenuated X-ray intensities are registered by a sensor. Mathematical algorithms are subsequently employed to reconstruct a number of 2D cross-sectional images or "slices" of the 3D object in question. These slices are often referred to as tomographic images. During the tomographic reconstruction process X-ray radiology is employed for the purpose of generating a large number of radiographic projections, which are subsequently reconstructed using a mathematical algorithm [12, 29, 30] in order to produce a slice image of the object being scanned. These reconstructed slices can then be stacked to

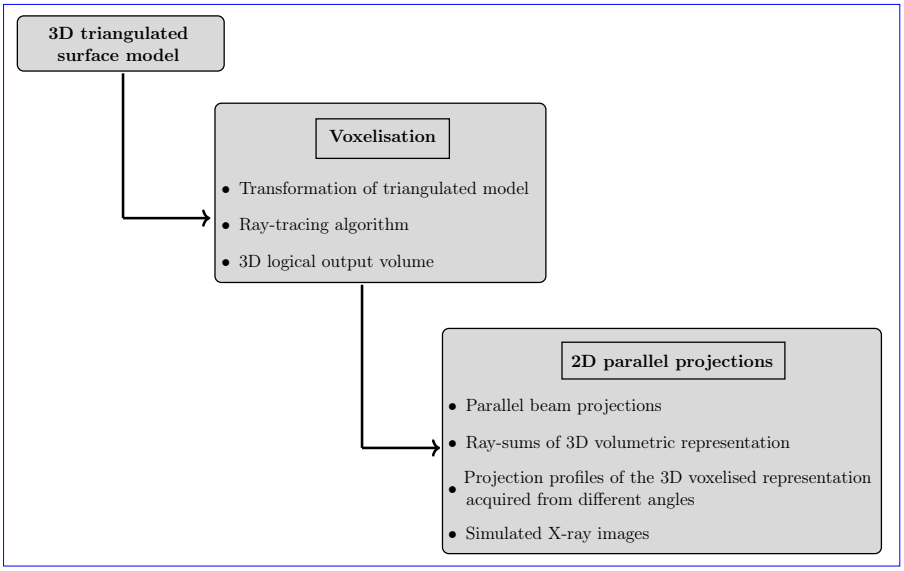


Figure 3.1: Overview of the proposed simulated X-ray data generation protocol.

form a 3D representation of the object that can be used in a wide range of applications [31–34]. The X-rays passing through the object are attenuated according to the Beer-Lambert law [13],

$$I(x) = I_0 \mathrm{e}^{-\mu x}, \tag{3.1}$$

where I is the recorded X-ray intensity after passing through the object, I_0 is the incident X-ray intensity, x is the length of the X-ray path through the object, and μ is the linear attenuation coefficient of the material for the X-ray energy being employed. Equation 3.1 assumes a homogeneous material and a monochromatic X-ray beam. For inhomogeneous objects like the human body, the attenuation of the X-rays may consequently be described by

$$I(x) = I_0 \mathrm{e}^{-\int \mu(x)\, dx}. \tag{3.2}$$

A CT image is obtained by rotating an X-ray source about the object with a detector positioned directly opposite the radiation source. The X-ray detector records attenuated intensity values after the X-ray beams penetrated the object (see Figure 3.2). The resultant intensity value of each detector pixel is a function of the attenuation coefficient μ and the path followed by the X-rays. Unlike traditional projection radiography, CT collects multiple

projections of the same object from different orientations by moving the X-ray source about the object. CT systems are outfitted with rows of digital detectors of which the signals serve as direct input to a computer. These signals are subsequently used to reconstruct one or more cross sections (slices) of the scanned object. In this way, although CT systems acquire projections that represent a "shadow" of the object, they generate truly tomographic images after reconstruction.

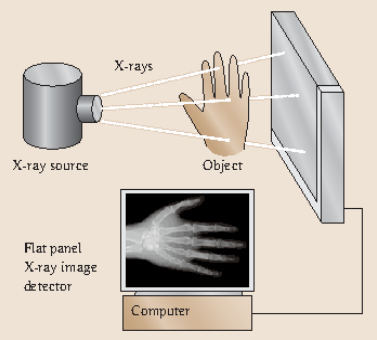


Figure 3.2: Acquisition of an X-ray image, with an X-ray source, object and an X-ray imaging detector [1].

The fundamental measurement required by a CT scanner is the estimation of X-ray attenuation along a line between an X-ray source and an X-ray detector [12, 13]. In order to reconstruct an image of a 2D cross section, a collection of such measurements are required along all lines within the cross section. The mathematical principles of CT were first investigated as early as 1917, when Johann Radon showed that a function can be reconstructed from an *infinite* set of its projections using the Radon transform (RT) [35]. The continuous RT in 2D constitutes an integral transform, that is the integral of a function along straight lines.

The RT of a 2D function $f(x, y)$ is defined as follows: Let A be a 2D object in $\Re^2$ and let $f(\mathbf{x})$ be the bi-dimensional binary function of A, defined as $f(\mathbf{x}) = 1$, when $\mathbf{x}$ lies within the 2D object's interior, and $f(\mathbf{x}) = 0$ otherwise [31, 36]. The 2D RT ($\mathcal{R}_f$) of the function $f(\mathbf{x})$ associates with each pair $(\boldsymbol{\eta}, \rho)$ the integral of $f(\mathbf{x})$ on the plane $P(\boldsymbol{\eta}, \rho) = \{\mathbf{x}|\mathbf{x}^t \cdot \boldsymbol{\eta} = \rho\}$ where $\boldsymbol{\eta} = (\cos\theta, \sin\theta)$. The plane P is normal to the direction $\boldsymbol{\eta}$ and a

distance ρ from the origin. The RT is therefore defined as

$$\begin{aligned}\mathcal{R}_f(\eta,\rho) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\delta(x\cos\theta + y\sin\theta - \rho)\,dxdy \\ &= \int_{-\infty}^{\infty} f(\rho\cos\theta - \ell\sin\theta, \rho\sin\theta + \ell\cos\theta)\,d\ell,\end{aligned} \tag{3.3}$$

where the Dirac delta function, δ, converts the 2D integral (with respect to x and y) to a line integral $d\ell$ along the line $x\cos\theta + y\sin\theta = \rho$ and $f(x,y)$ represents the X-ray attenuation coefficient, which is rotated by an angle θ with respect to the object location (x,y).

The ℓ axis is aligned with the direction of a given projection, that is, the lines (beams) along which attenuation information is accumulated through integration. In the particular case of CT, this coincides with the direction of the X-rays. For a fixed $\theta, \mathcal{R}_f$ represents a 1D projection of the 2D object $f(x,y)$ and $\{\mathcal{R}_f(\eta,\rho)|\theta \in [0,\pi)\}$ constitutes a complete collection of these 1D projections of the 2D object $f(x,y)$ [12].

In 3D, the RT of a function $f(x,y,z)$ is obtained by integrating along *planes*. The 3D RT constitutes a generalisation of Equation 3.3 and is defined as follows,

$$\mathcal{R}_f(\eta,\rho) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(\underline{x})\delta(x\cos\phi\sin\theta + y\sin\phi\sin\theta + z\cos\theta - \rho)\,d\underline{x} \tag{3.4}$$

where $\underline{x} = (x,y,z)$ and the unit vector $\boldsymbol{\eta}$ using spherical coordinates becomes $\boldsymbol{\eta} = (\cos\phi\sin\theta, \sin\phi\sin\theta, \cos\theta)$. The generation of a complete 2D RT-based data set only requires rotation about *one* axis, while the generation of a complete 3D RT-based data set requires a rotation that considers *two* angles, for example θ and ϕ that cover the entire unit sphere. In this way sets of planes that intersect the object at all possible angles are obtained. Figure 3.3 (a) depicts a straight line using its *normal* representation and Figure 3.3 (b) depicts the relevent 3D projection geometry. A more detailed description of the 3D RT can be found in [37, 38].

The proposed projection protocol *simulates* the acquisition of a CT scan that measures the X-ray attenuation along a line between an X-ray source and an X-ray detector. Each voxel within the 3D volumetric representation of a CAD surface model associated with a dental implant constitutes a *cube* which has a value of *one*. It is therefore assumed that the material of the dental implant is homogeneous and that all the attenuation coefficients are the same. The implementation of the data generation strategy (see Figure 3.4) is described in the following sections.

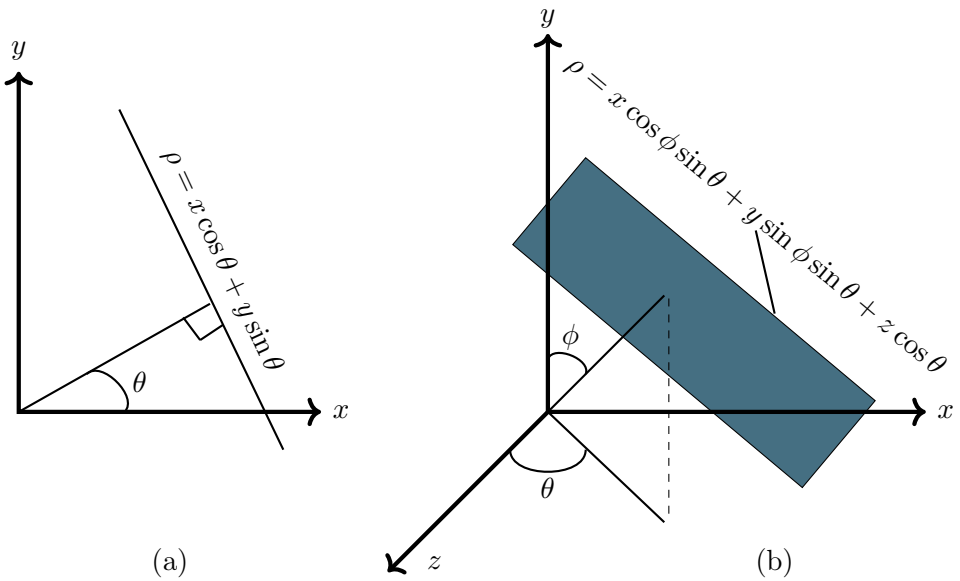


Figure 3.3: (a) Geometrical description of the path of integration along a line ℓ for the 2D problem. (b) Depiction of the analog integration process for the 3D problem.

3.3 Voxelisation

A triangulated 3D surface model constitutes a series of vertices, edges and faces stored in stereolithography (STL) format. Examples of triangulated 3D surface plots of the so-called C1 implant model with a length of 10 mm and external diameters of 3.30 mm, 3.75 mm and 5.00 mm are presented in Figure 3.5. The triangles associated with an STL file constitute the object surface, where each triangle is defined by its normal vector, n, and the coordinates of its three vertices, v_1, v_2, and v_3 (see Figure 3.6). The triangulated surface model describes *only* the outer shell or contour of a dental implant and not the entire volume.

Therefore, as a point of departure, a triangulated 3D surface model of a dental implant has to be converted into a so-called volumetric representation. This is achieved through a process known as *voxelisation*. It is proposed that a triangulated 3D surface model is converted into a voxel-based model using an algorithm conceptualised by Aitkenhead [39]. The aforementioned algorithm is based on ray-tracing as will be discussed shortly.

Aitkenhead's voxelisation algorithm produces a bounded logical grid rep-

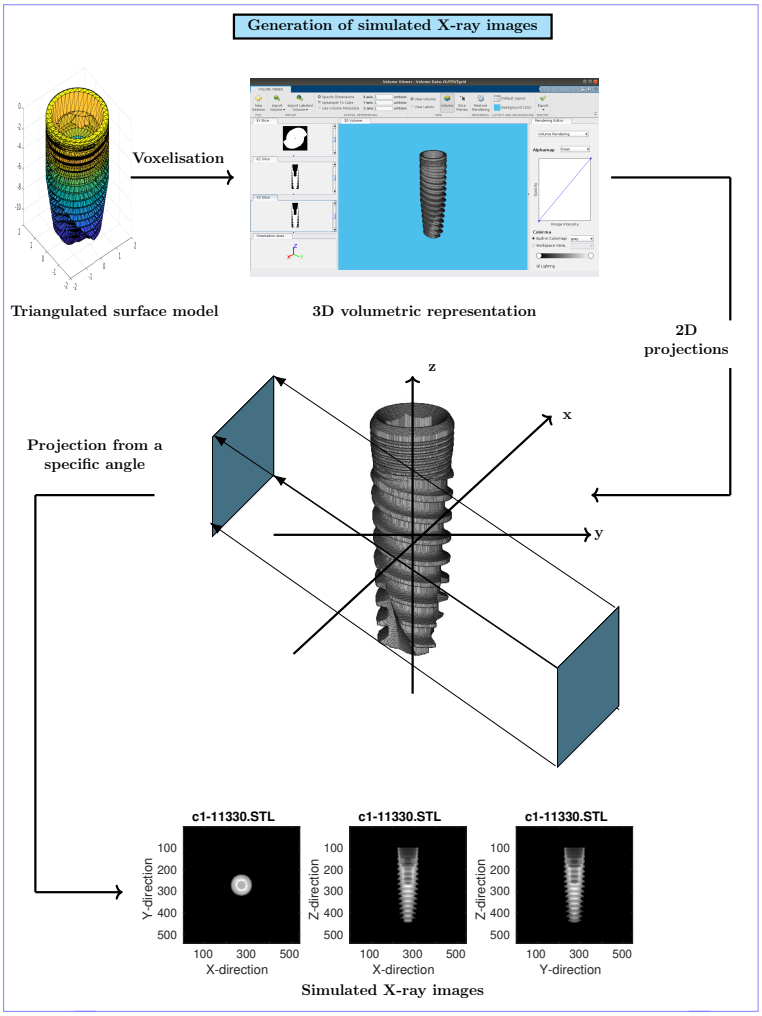


Figure 3.4: Framework of the proposed simulated data generation algorithm. The triangulated 3D surface model is converted into a 3D volumetric representation, where each voxel constitutes a cube with a value of one. Each simulated X-ray image is obtained by calculating a 2D projection of the volumetric representation from a specific angle.

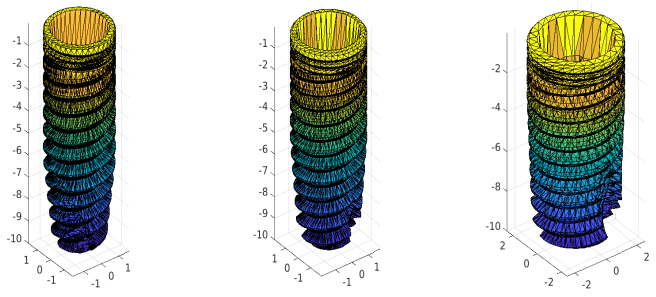


Figure 3.5: Examples of triangulated surface plots defining the geometry of dental implants. The depicted implants which belong to the so-called C1 model have a length of 10 mm and external diameters of (a) 3.30 mm, (b) 3.75 mm and (c) 5.00 mm.

resenting the implant shape. The output is a 3D logical volume containing the voxelised data, where a value of one is assigned to each cube-shaped voxel located inside the mesh, while a value of zero is assigned to each cube-shaped voxel located outside the mesh. More specifically, Aitkenhead's voxelisation algorithm operates by passing a set of rays through the object along a specific axis to compute the intersections with the facets. The basic principle of the ray-tracing algorithm is conceptualised in Figure 3.7.

The set of rays passes through voxel centres that are located in the grid perpendicular to the chosen axis. The ray-tracing algorithm therefore iterates through each voxel in the grid and considers a ray passing through the centre of each voxel. The iteration occurs along the two grid directions, typically the x and y axes, and starts from the minimum value associated with those directions and continues up to the maximum value. By considering a ray passing through the centre of each voxel, the algorithm is able to sample the entire 3D space represented by the voxel grid. Figure 3.8 depicts the voxelisation process for a shape in the form of a pyramid.

For each iteration, all the possible crossed mesh facets are found. For every one of these facets it is verified that the ray is on the same side of the opposing vertex with respect to every edge of the facet. On the other hand, if the ray crosses exactly on a vertex, it is verified that the nearby facets have normal components with the same or the opposite direction of the ray. The z coordinates of the points where the rays cross the mesh are subsequently individuated. The z coordinates are determined by solving the

(a) (b)

Figure 3.6: (a) A typical triangle associated with an STL file. (b) Example of an STL structure that comprises of three triangles.

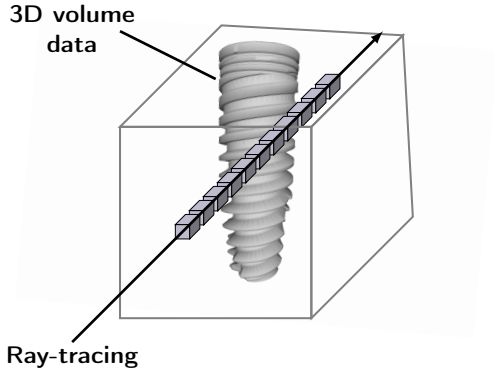


Figure 3.7: Conceptualisation of the ray-tracing algorithm implemented during the voxelisation process for the purpose of transforming a triangulated 3D surface model into a 3D volumetric representation.

following equation that describes the plane of the facet,

$$Ax + By + Cz + D = 0, \tag{3.5}$$

where

$$\begin{aligned}
A &= y_1(z_2 - z_3) + y_2(z_3 - z_1) + y_3(z_1 - z_2), \\
B &= z_1(x_2 - x_3) + z_2(x_3 - x_1) + z_3(x_1 - x_2), \\
C &= x_1(y_2 - y_3) + x_2(y_3 - y_1) + x_3(y_1 - y_2) \text{ and} \\
D &= x_1(y_3 z_2 - y_2 z_3) + x_2(y_1 z_3 - y_3 z_1) + x_3(y_2 z_1 - y_1 z_2).
\end{aligned}$$

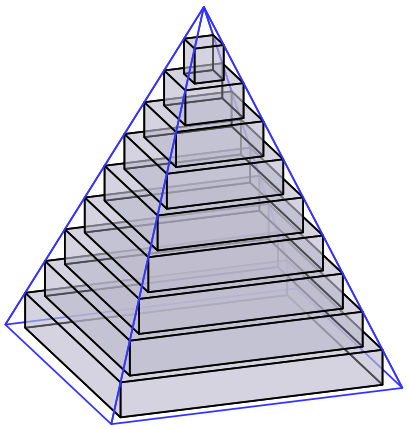


Figure 3.8: Conceptualisation of the voxelisation process within the context of a pyramid. The individual voxels in each horizontal layer are not shown.

Equation 3.5 is subsequently solved by substituting the x and y coordinates of the points in order to determine z. In order to reduce voxelisation errors, this procedure is repeated for each direction (x,y,z) so as to produce three intermediate volumes. These volumes are subsequently combined into a single 3D volumetric model by finding all voxels that intersect at least two of the intermediate volumes.

Linear interpolation is employed to fill in the rays that could not be voxelised. A ray for which the voxelisation algorithm did not give a clear result is therefore calculated by interpolating the adjacent rays. Each voxel in the volumetric representation constitutes a cube with a value of one.

The dimensions of the grid were chosen in such a way that each cube-shaped voxel has a side length of 1/30 mm. This spatial resolution is considered to be detailed enough to ensure that important features are not lost and coarse enough to avoid spatial inefficiency.

A dilation factor of 18 is implemented in all the directions in order not to deform the object. The initial dimensions of the model are therefore multiplied by 18 and rounded to the nearest integer.

The logical output grids containing three dental implant shapes are shown in Figure 3.9. Such a 3D output volume serves as input from which 2D projections are obtained for the purpose of generating simulated X-ray images.

For visual comparison purposes, a few examples of 3D triangulated surface models and the corresponding 3D volumetric representations are presented in Figure 3.10

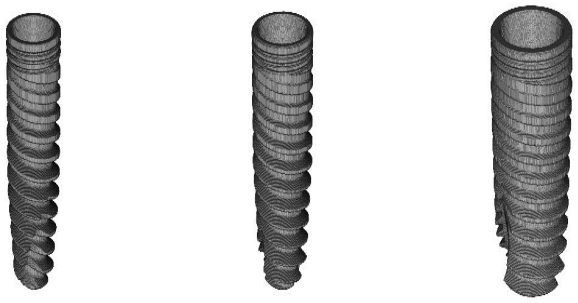


Figure 3.9: The 3D voxelised implant volumes corresponding to the triangulated surface plots depicted in Figure 3.5.

3.4 Two-dimensional projections

In this study the proposed projection strategy *imitates* the X-ray emission protocol in a CT scan by projecting parallel beams modelled by a set of lines across the 3D volumetric representation of a dental implant from different angles.

Two-dimensional projections of a 3D volumetric representation are therefore computed from a number of different angles. Each projection is obtained by calculating a number of parallel ray-sums of the 3D volumetric representation. The projections are acquired by rotating the 3D volumetric representation, while keeping the detector and source fixed. Note that in the case of clinical CT systems, on the other hand, the detector and the source rotate about the object (patient). The value of a specific pixel within the projection image is obtained by integrating (summing) the 3D volumetric representation along a beam in the direction perpendicular to the detector.

In 3D space the generic rotation matrix for the purpose of re-orientating a 3D function can be derived from basic plane rotations. Such a plane rotation is typically performed about one of the axes of the Cartesian system. The following three 3×3 rotation matrices rotate an arbitrary vector by an angle θ about the x, y and z axis respectively:

$$R_x(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}, R_y(\theta) = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix} \text{ and}$$

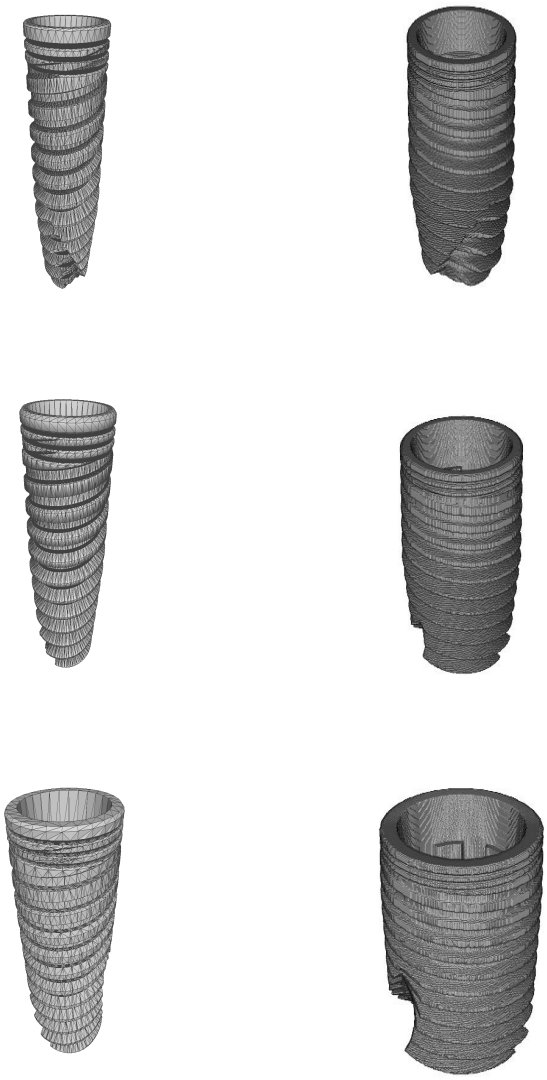


Figure 3.10: Triangulated surface models defining the geometry of a dental implant and their respective 3D voxelised representations. (**Left**) Zoomed-in versions of triangulated surface models. (**Right**) Zoomed-in versions of 3D volumetric representations corresponding to the triangulated surface models on the left.

$$R_z(\theta) = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

According to Euler's rotation theorem any arbitrary rotation in 3D space can be achieved by combining the above three rotation matrices through multiplication:

$$\begin{aligned} R &= R_z(\phi)R_y(\beta)R_x(\theta) \\ &= \begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix} \\ &= \begin{bmatrix} \cos\beta\cos\phi & \sin\theta\sin\beta\cos\phi - \cos\theta\sin\phi & \cos\theta\sin\beta\cos\phi + \sin\theta\sin\phi \\ \cos\beta\sin\phi & \sin\theta\sin\beta\sin\phi + \cos\theta\cos\phi & \cos\theta\sin\beta\sin\phi + \sin\theta\cos\phi \\ -\sin\beta & \sin\theta\cos\beta & \cos\theta\cos\beta \end{bmatrix}. \end{aligned}$$

The Euler-Rodrigues angle-axis representation principle may subsequently be employed, which consists of two components, that is, the unit vector defining the rotation axis and the angle of rotation [40].

In Figure 3.11 the proposed parallel projection protocol is conceptualised. The 3D volumetric representation resides in the world coordinate system (x, y, z), and its centre point coincides with the origin. The 2D projections are computed as the 3D volumetric representation rotates.

During the X-ray simulation process, each 3D volumetric representation of an implant undergoes axial, out-of-plane and in-plane rotations through a number of different angles, before its projection is generated (see Figure 3.12). The in-plane rotations are conducted during the data augmentation protocol that forms *part of training* the model proposed for the purpose of automatically classifying the connection type associated with a specific dental implant from an X-ray image (as will be discussed in detail in Chapter 5), while axial and out-of-plane rotations are conducted during the initial phase of the X-ray simulation process, that is *before* training. The transformation is the result of two consecutive rotations where the first rotation is through an azimuth angle ϕ about the z axis and the second rotation is through an elevation angle θ about the x axis after which the projected data is obtained through summation.

The Euler-Rodrigues principle (axis-angle rotation) is employed for the purpose of computing the rotation matrix about the relevant axis as follows

$$Q = \mathbf{I} + \boldsymbol{\Omega}\sin\theta + \boldsymbol{\Omega}^2(1 - \cos\theta), \tag{3.6}$$

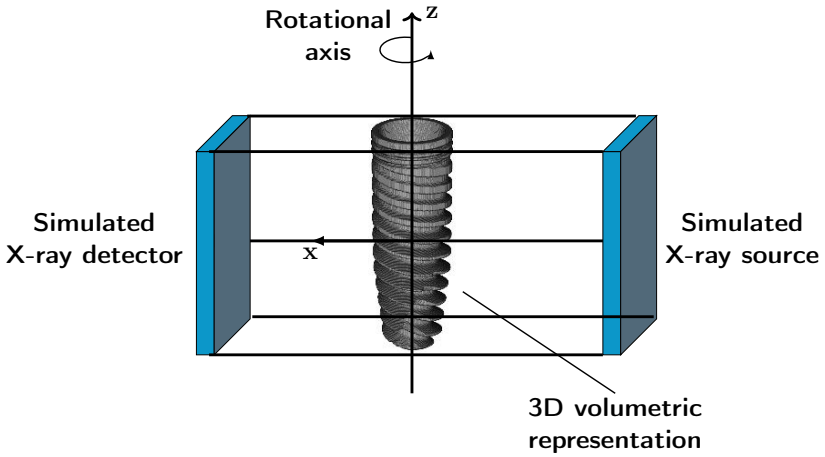


Figure 3.11: Conceptualisation of the proposed simulated X-ray data acquisition protocol that calculates 2D projections through the summation of voxels along parallel beams.

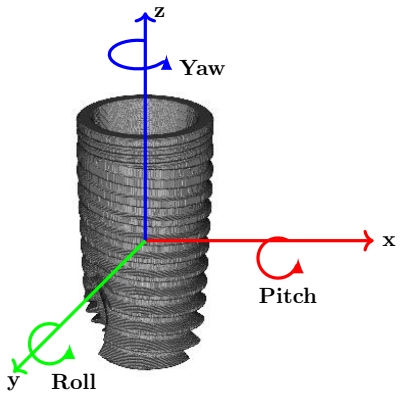


Figure 3.12: The angles of rotation about the x, y and z axes are referred to as pitch, roll and yaw, respectively. Assuming that the simulated X-ray source and sensor are perpendicular to the roll axis, that is, the scan beams are parallel to the roll axis, in-plane rotations are associated with rotations about the roll axis, while out-of-plane rotations are associated with rotations about the pitch axis. Axial rotations occur about the yaw axis.

where $\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and $\mathbf{\Omega}$ represents the cross-product matrix $\mathbf{\Omega} = \begin{bmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{bmatrix}$ with $\omega = [\omega_x, \omega_y, \omega_z]^T$ being the axis of rotation and θ the angle of rotation. Once the rotation matrix has been determined, the re-orientation of the 3D volumetric representation is performed. The 3D binary volume is rotated about the z axis using a complete azimuth range of $0°$ to $360°$ with increments of $5°$. The elevation rotations are computed within the intervals of $0°$ to $60°, 120°$ to $240°$ and $240°$ to $360°$ with increments of $5°$. Elevation angles outside the aforementioned intervals are not expected in practical scenarios. In order to eliminate wraparound effects, which may be due to trigonometric interpolation, zero-padding is implemented which ensures that the output is a cube. Zero-padding is implemented for each dimension by inserting zero-valued voxels before the first element and after the last element (see Figure 3.13).

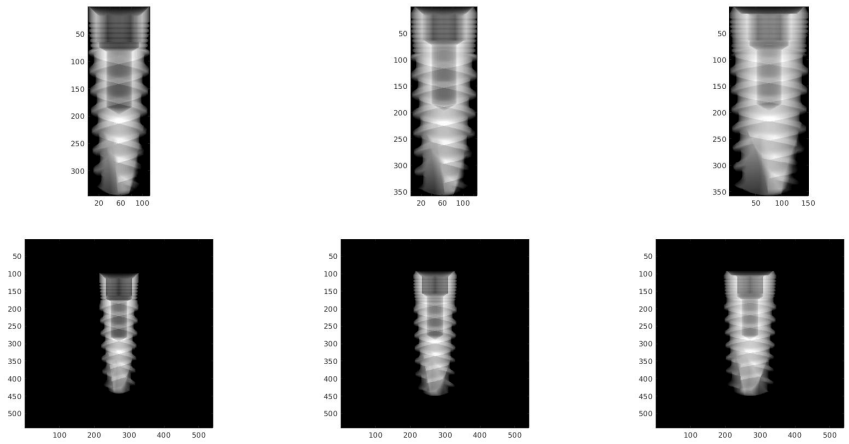


Figure 3.13: Examples of simulated X-ray images. **(Top)** Examples of simulated X-ray images *before* zero-padding. **(Bottom)** Examples of simulated X-ray images corresponding to the images on the top *after* zero-padding has been implemented.

For fixed angles θ and ϕ the beam-sum of voxels within the 3D volumetric representation is obtained. By iterating through all the parallel beams which intersect the 3D volumetric representation for fixed values for θ and ϕ a sin-

gle 2D projection profile is obtained. A projection profile of the 3D voxelised representation acquired from specific angles θ and ϕ therefore constitutes a simulated X-ray image. A particular pixel value within the simulated X-ray image is the beam-sum that constitutes the accumulated intensities of the 3D volumetric representation along the beam in question. By iteratively changing θ and ϕ and repeating the procedure, a large number of projections can be obtained. These projections can theoretically be acquired by considering rotations through the entire range of 360°.

Simulated X-ray images are therefore obtained when projections along parallel beams of a 3D voxelised representation of a dental implant are computed from specified directions. To generate a projection, a set of *parallel* beams are projected through the 3D volumetric representation, with each beam passing through a series of voxels at regular intervals. Along each beam, the contributions of the voxels intersected by the beam are subsequently summed to obtain a single value for a pixel in the final projected image. This resulting value therefore coincides with the intensity of the corresponding pixel in the 2D image. In Figure 3.14, the proposed parallel projection protocol is conceptualised.

As the 3D volumetric representation rotates about an axis, 2D parallel projections are calculated. For the purpose of illustrating the proposed out-of-plane rotation strategy implemented in generating simulated training samples, a number of examples of the dental implant model C1 with a conical narrow platform, external diameters of 3.30 mm and 5.00 mm and lengths of 10 mm and 16 mm are represented in Figure 3.15.

In addition to the out-of-plane and axial rotations (conducted *before* training), a number of in-plane rotations are conducted during the data augmentation protocol (*as part of* training) which are described in more detail in Chapter 5.

3.5 Concluding remarks

In this chapter an algorithm based on the calculation of 2D projections (from a number of different angles) of 3D volumetric representations for the purpose of generating artificial training samples from triangulated 3D surface models within the context of dental implant recognition was proposed. The artificially generated X-ray images serve as input for training and validating the proposed network for the purpose of identifying the connection type associated with a specific dental implant within a questioned X-ray image. This will be discussed in more detail in Chapter 5.

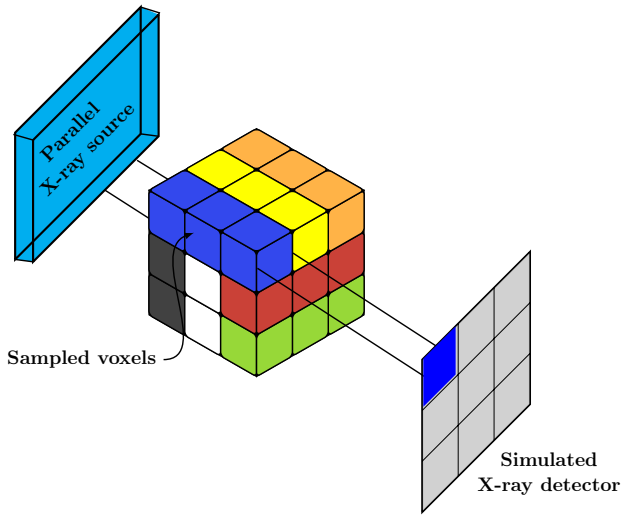


Figure 3.14: Conceptualisation of parallel beams being projected through a 3×3×3 volumetric representation in the form of a cube. Assuming that each blue voxel has a value of one, the blue pixel in the simulated X-ray detector will have a value of three. In this application all the voxels associated with a 3D volumetric representation of a dental implant is assumed to have a value of one. When the entire voxel does not lie within the path of a specific beam, a weight that coincides with the fraction of the volume that intersects the beam is used.

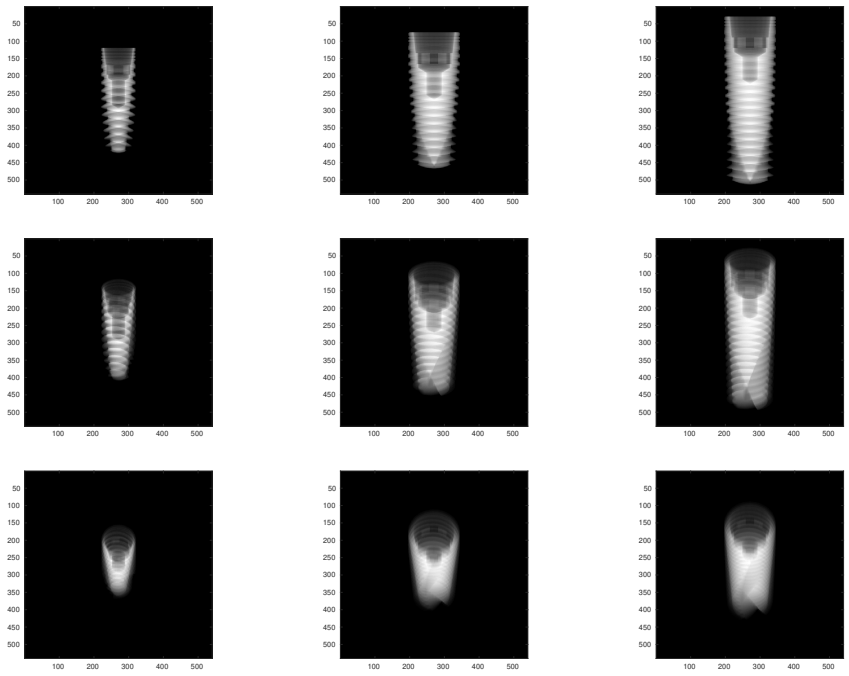


Figure 3.15: Examples of simulated X-ray images employed for training the proposed network. **(Top)** Examples of unrotated simulated X-ray images. **(Middle)** Examples of simulated X-ray images that underwent an out-of-plane rotation of 30°. **(Bottom)** Examples of simulated X-ray images that underwent an out-of-plane rotation of 60°.

Chapter 4

Image segmentation

4.1 Introduction

In this chapter, the novel semi-automated and fully automated image segmentation systems proposed in this study are discussed. In the case of the semi-automated system suitable regions of interest (ROIs), that contain the dental implants, are manually specified (selected). Within the context of the fully automated system, suitable ROIs are automatically detected through a deep learning-based technique. A semantic segmentation algorithm is proposed for the purpose of detecting pixels associated with the dental implants without differentiating implant instances in an *actual* X-ray image (see Section 4.2). Instance segmentation is subsequently applied to the output mask image acquired through the aforementioned semantic segmentation algorithm in order to differentiate dental implant instances in an *actual* X-ray image (see Section 4.3). The image segmentation protocol developed in this study is depicted graphically in Figure 4.1.

4.2 Semantic segmentation

In this section, the two types of image segmentation systems proposed in this study, that is, the *semi-automated* and *fully automated* systems, are discussed.

The proposed semi-automated system involves manual selection of suitable ROIs that contain dental implants. This means that a human operator needs to first identify and specify the areas where the dental implants are located. Once the ROIs are identified, the proposed system uses image processing techniques to perform semantic segmentation in order to detect the

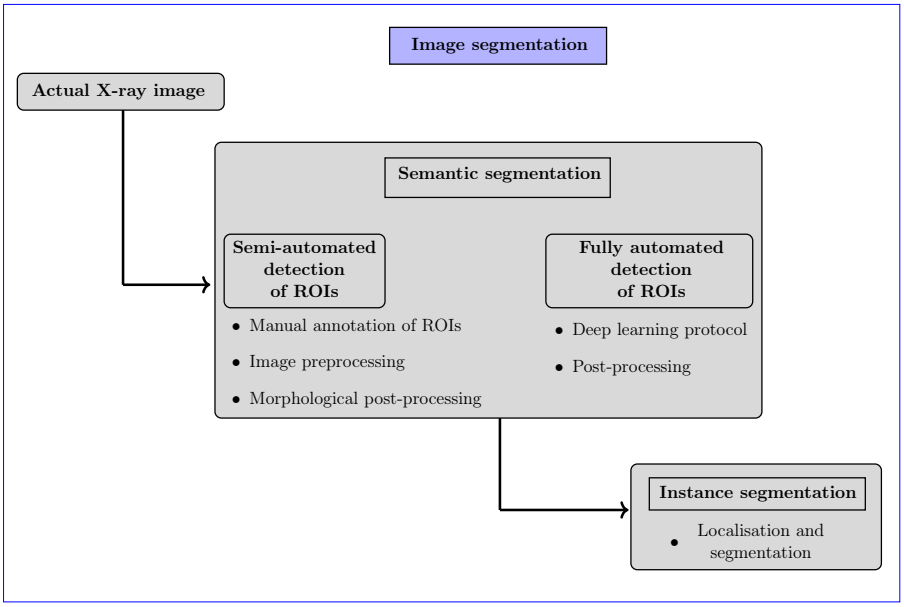


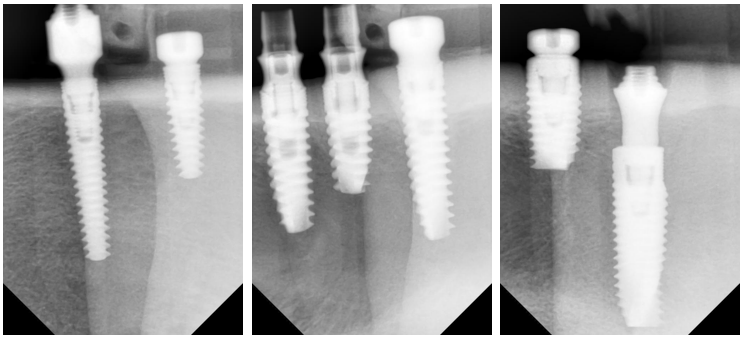
Figure 4.1: Conceptualisation of the proposed image segmentation protocol.

pixels associated with the dental implants. On the other hand, the proposed fully automated system uses deep learning-based techniques to automatically detect suitable ROIs without human intervention.

In both cases, the main purpose of the segmentation process is to identify the pixels that are associated with the dental implants and *those* associated with the background. The system does not differentiate between different implant instances, and only detects the pixels that are associated with implants.

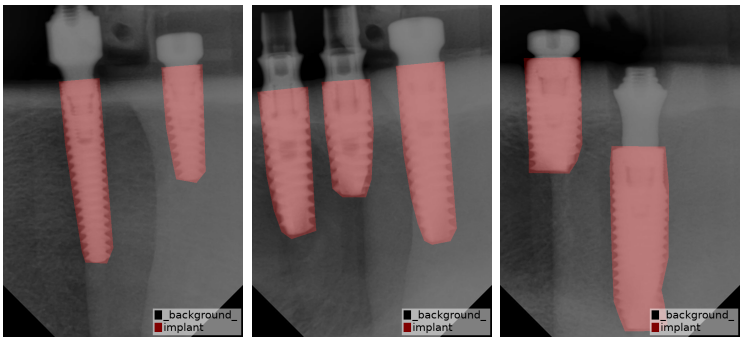
4.2.1 Semi-automated detection of the regions of interest

An ensemble of image processing techniques are implemented for the purpose of detecting pixels associated with the dental implants (foreground) and discarding those associated with the background without differentiating implant instances in the actual X-ray image. In Figure 4.2, a few examples of actual X-ray images are depicted. The suitable ROIs that contain the dental implants are manually selected within the actual X-ray images. Polygonal shapes are used to annotate the ROIs which contain the dental implants within a questioned image. In Figure 4.3 the manually selected ROIs are superimposed onto the dental implants.



(a) (b) (c)

Figure 4.2: Examples of actual X-ray images containing dental implants.

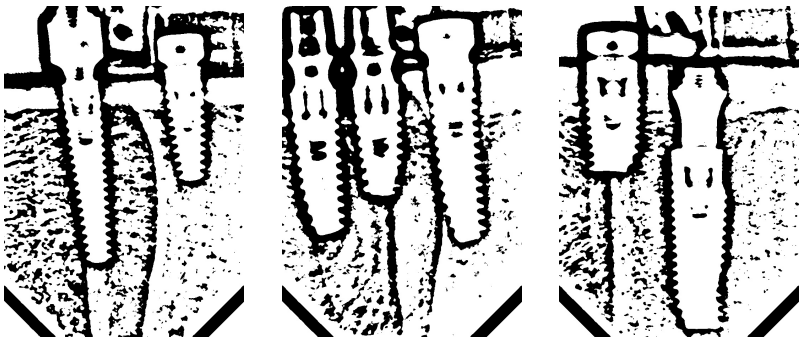


(a) (b) (c)

Figure 4.3: Examples of manually selected ROIs annotated using polygonal shapes, superimposed onto the actual X-ray images.

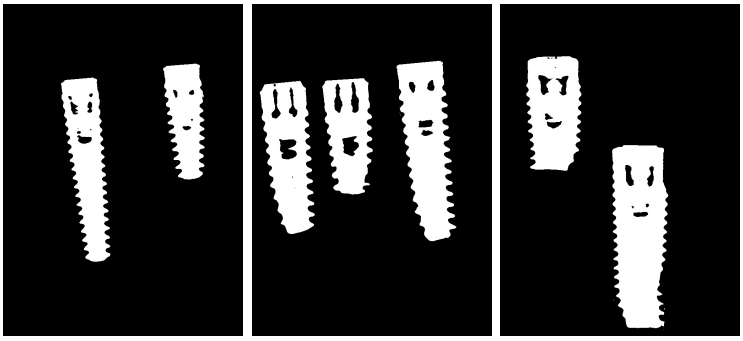
Local adaptive thresholding is applied to an actual X-ray image for the purpose of converting it from grey-scale to binary format (see Figure 4.4). The manually selected ROIs are subsequently employed as a mask image in order to remove the pixels *not* associated with the dental implants (see Figure 4.5).

In order to reduce the noise in the binary images depicted in Figure 4.5 and render the foreground boundaries more regular, the images are subjected to morphological post-processing. For this purpose a morphological hole-filling operation is applied which is subsequently followed by morphological closing. The aforementioned post-processing techniques are performed in order to eliminate noise, fill in the holes and enhance the binary mask image. The results are depicted in Figure 4.6. The post-processed binary masks also



(a) (b) (c)

Figure 4.4: Resulting binary images after local adaptive thresholding has been applied to the images depicted in Figure 4.2.



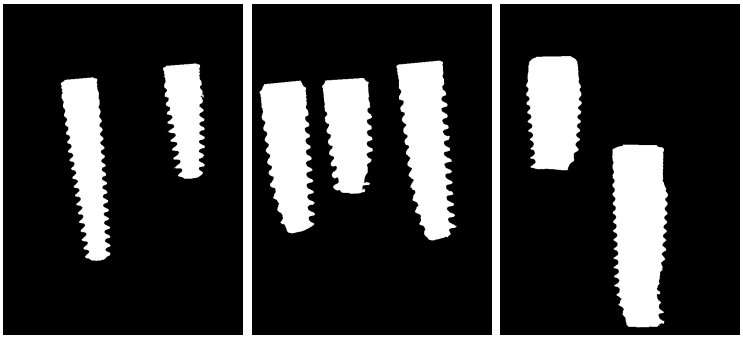
(a) (b) (c)

Figure 4.5: Detected ROIs after the implementation of ROI-masking.

serve as a ground truth for the purpose of *automatically* detecting suitable ROIs that contain the dental implants.

4.2.2 Fully automated detection of the regions of interest

In this study *two* independent fully convolutional network (FCN) models are investigated: (i) the first model (FCN-1) is proposed for the purpose of automatically classifying the connection type associated with a specific dental implant from an X-ray image, while (ii) the second model (FCN-2) is proposed for the detection of suitable ROIs that contain the dental implants in an actual X-ray image. Model FCN-1 will be described in more detail in Chapter 6.



(a) (b) (c)

Figure 4.6: Examples of mask images after the application of morphological post-processing techniques.

The FCN-2 model is proposed for the detection of suitable ROIs which contain the dental implants in an actual X-ray image. The FCN-2 model constitutes a two-dimensional (2D) convolutional neural network (CNN) within an encoder-decoder architecture, which performs semantic pixel-wise segmentation [5, 41, 42]. The architecture of the proposed FCN-2 model is depicted in Figure 4.7 and described in detail below.

Encoder

The input images (the actual X-ray images and the corresponding segmentation mapping) are fed through the encoder network so that downsampled feature maps are generated. The proposed encoder network consists of ten convolutional layers, where each of these layers is followed by a rectified linear unit (ReLU), batch normalisation (BN) and maximum pooling layer.

The encoder performs spatial convolution on the input image to produce a set of feature maps. An element-wise ReLU function is subsequently applied to the feature map which sets all of the negative pixel values in the feature map equal to zero. This process can be formulated as follows,

$$f(\mathbf{x}_i) = \mathrm{ReLU}(\mathbf{W}_i * \mathbf{X}_i + \mathbf{b}_i) \quad \text{for } i = 1, \ldots, N, \tag{4.1}$$

where N is the number of convolutional layers, $\mathbf{W}_i$ and $\mathbf{b}_i$ denote the weights and biases respectively, $*$ represents the convolution operator, $\mathbf{X}_i$ is the input image and $\mathrm{ReLU}(x) = \max(0, x)$ is the activation function.

Batch normalisation is subsequently applied to the output feature maps followed by max pooling. Max pooling is employed for the purpose of down-

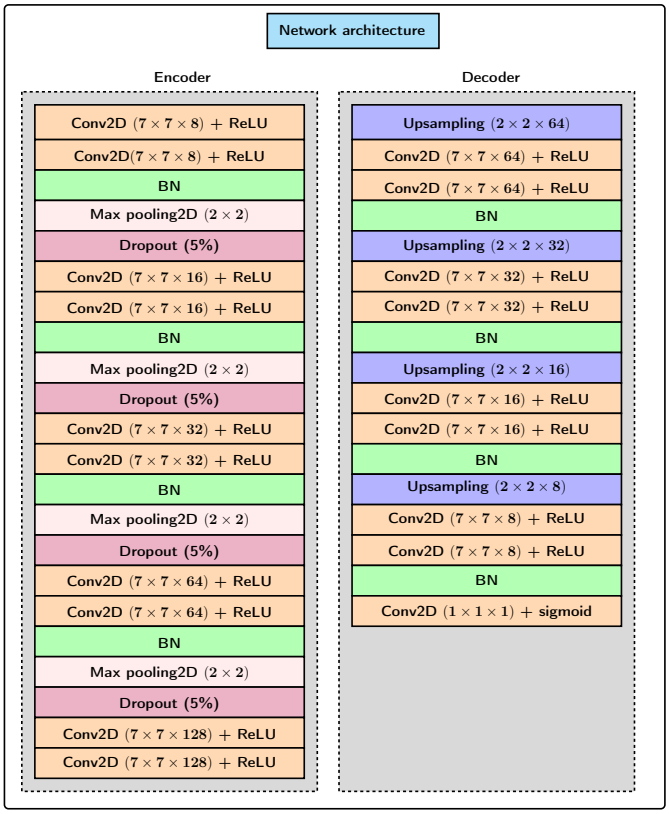


Figure 4.7: A depiction of the architecture of the proposed FCN-2 model which is employed for the purpose of automatically detecting suitable ROIs within an actual X-ray image.

sampling the feature maps in order to achieve translation invariance over small spatial shifts in the input image. A dropout rate of 5% is implemented which removes a random set of activations by setting them equal to zero, thereby essentially forcing the network to be redundant [43]. This implies that the network should be able to provide the correct classification or output even when certain activations are removed, therefore rendering the network less prone to overfitting, which in turn leads to better generalisation.

Decoder

The decoder network is implemented for the purpose of upsampling the feature maps to the same size as the original image. The upsampling layers are

followed by convolutional layers so as to generate dense feature maps. Each convolutional layer is followed by a ReLU layer. A batch normalisation step is then applied to each of the feature maps. A sigmoid function is applied to the final feature map in order to compute the probability distribution across the binary classes. The sigmoid function classifies each pixel independently. The output of the sigmoid classifier is a K-channel image of probabilities where K is the number of binary classes. The predicted segmentation corresponds to the class with maximum probability for each pixel.

Training

The input images and their corresponding segmentation maps are used to train the proposed model. The final binary masks acquired through the proposed semi-automated segmentation protocol serve as the ground truth. Typical examples of the training samples are presented in Figure 4.8.

During training, the actual X-ray images and the corresponding ground truth masks are augmented by applying geometric transformations. The following data augmentation techniques are implemented,

- rotations within an interval of $[0°, 50°)$,
- scaling variations within a range of [0.05, 0.2],
- translations within a range of [0, 2] pixels in the x and y directions, and
- horizontal and vertical flipping of the input images.

Since the images are augmented in real-time during training, a nearest neighbour interpolation technique is employed to fill in the missing information. Examples of output images acquired after the implementation of data augmentation are depicted in Figure 4.9.

During each forward pass of the training phase, information from the input images is propagated through the different layers of the network to obtain the predicted (network) output. The binary entropy between the model output and the ground truth (see Equation 4.2) serves as a loss function and is subsequently used during the backward pass to update the weights (parameters) within each layer of the network,

$$L = -\frac{1}{M}\sum_{i=1}^{M}[y_i \log(p_i) + (1 - y_i)\log(1 - p_i)], \tag{4.2}$$

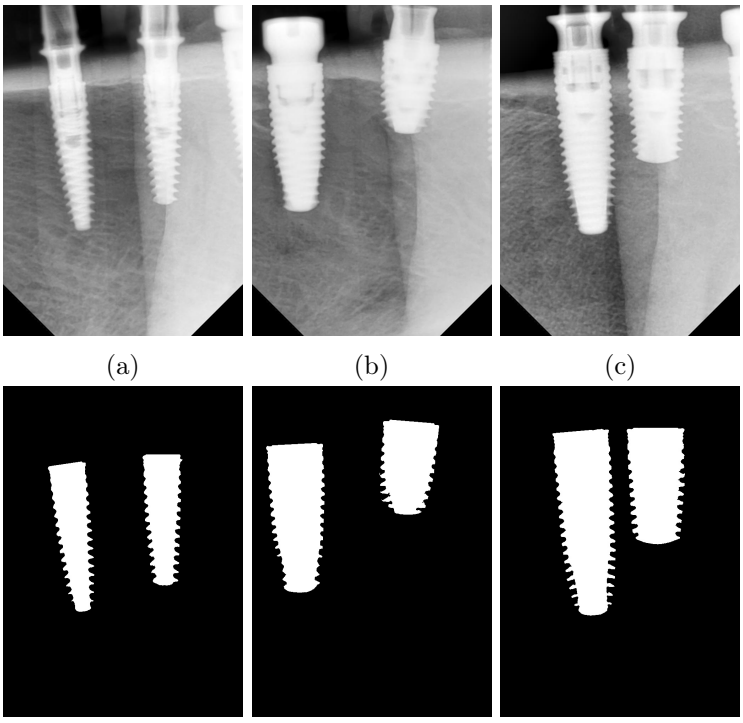


Figure 4.8: **(Top)** Examples of actual X-ray images containing dental implants. **(Bottom)** Binary versions of the corresponding images on the top after the application of the proposed semi-automated segmentation algorithm. These binary versions serve as a ground truth.

where M is the total number of pixels in the image and $y_i \in (0, 1)$ is the label of the i-th pixel where 0 indicates that the pixel belongs to the background and 1 indicates that the pixel belongs to the foreground. The network is trained by employing the adaptive momentum estimation (Adam) algorithm which optimises the model parameters by iteratively updating the weights using a batch of training data so as to minimise the error function [44].

Selected illustrational results and post-processing

Selected results illustrating the proficiency of the proposed FCN-2 model for the purpose of segmenting dental implant images into foreground and background regions are presented in Figure 4.10. The probability that a

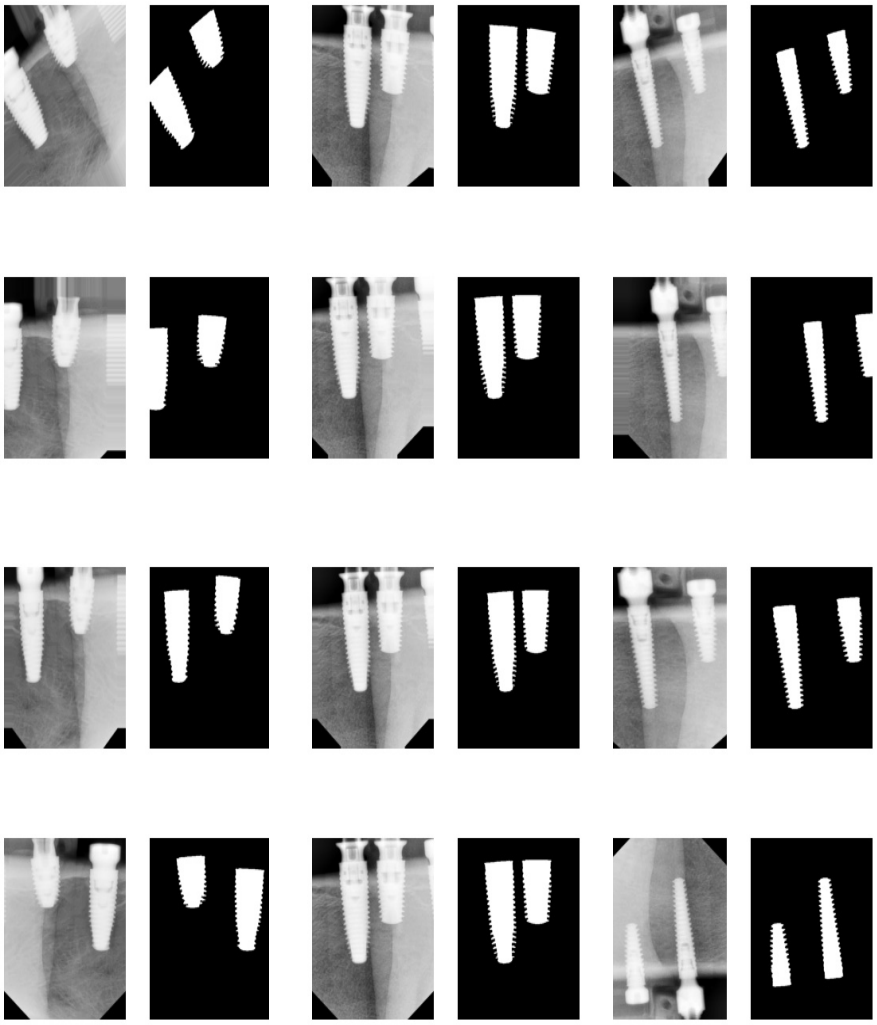


Figure 4.9: Visualisation of the augmented data. In each row different augmentation operations are depicted. The output images after the application of rotations, translations, variations in scale, as well as horizontal and vertical flipping are respectively depicted in the first row, second row, third row and fourth row.

pixel belongs to the foreground is denoted with a shade of red in Figures 4.10 (a), 4.10 (c) and 4.10 (e). The binarised images (using a probability threshold of 0.5) corresponding to the aforementioned probability matrices are depicted in Figures 4.10 (b), 4.10 (d) and 4.10 (f) respectively. The morphological post-processing techniques implemented within the context of the proposed semi-automated segmentation protocol are also conducted within the context of the proposed fully automated segmentation protocol, after which small connected components are removed. The results are depicted in Figure 4.11.

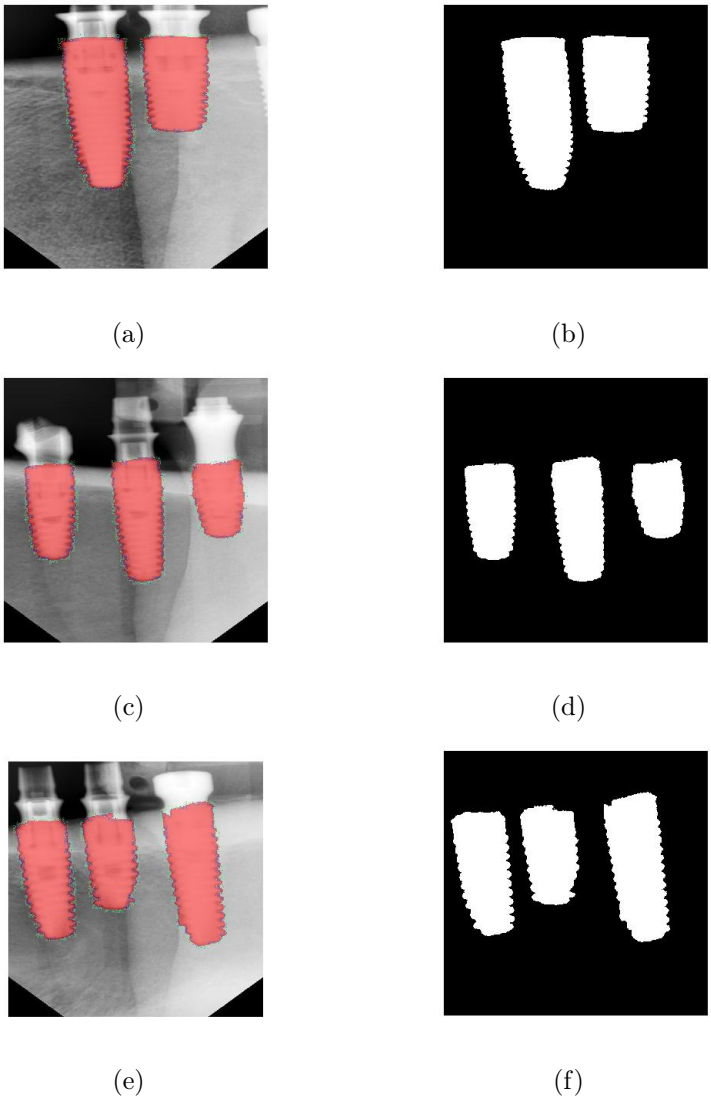
4.3 Instance segmentation

In this section instance segmentation is applied to the post-processed mask images acquired through the proposed semi-automated or fully automated algorithms in order to differentiate dental implant instances in an actual X-ray image. Each mask image (see Figure 4.11) is partitioned into its constituent components through connected component analysis and the detected dental implants are therefore localised and segmented.

A two-pass algorithm is employed for detecting the connected components and labelling each connected component within the binary image. A different label is therefore assigned to each dental implant (see Figure 4.12). Each component is delimited by a bounding box which is subsequently used to segment the actual X-ray image into its constituent dental implants for the purpose dental implant classification (see Figure 4.13).

ROI masking

During the final stage of segmentation an actual X-ray image is partitioned into its constituent components. The post-processed binary images acquired through the proposed semi-automated or fully automated algorithms are employed as mask images in order to isolate dental implants in the actual X-ray images (see Figure 4.14). Each masked X-ray image is subsequently segmented into its constituent dental implants by using the bounding boxes acquired through the connected component algorithm. Figure 4.15 depicts an actual X-ray image segmented into its constituent dental implants. Each segmented implant serves as input for the dental implant classification framework that attempts to accurately classify the connection type associated with the implant in question.

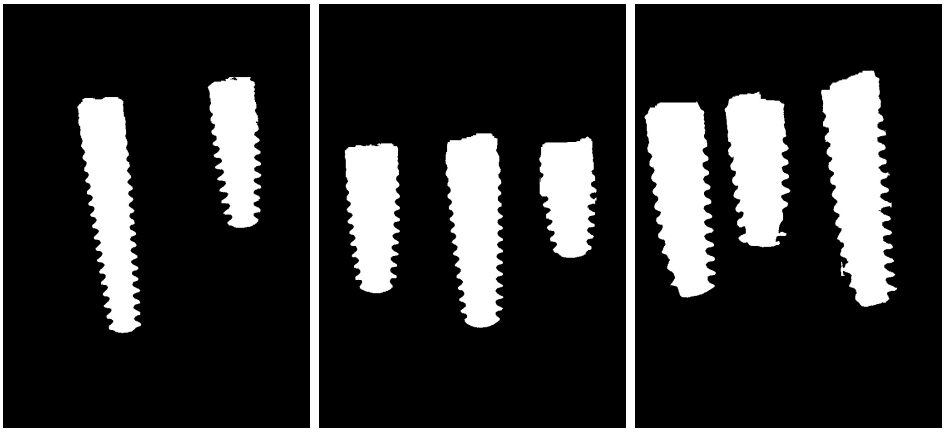


(a) (b)

(c) (d)

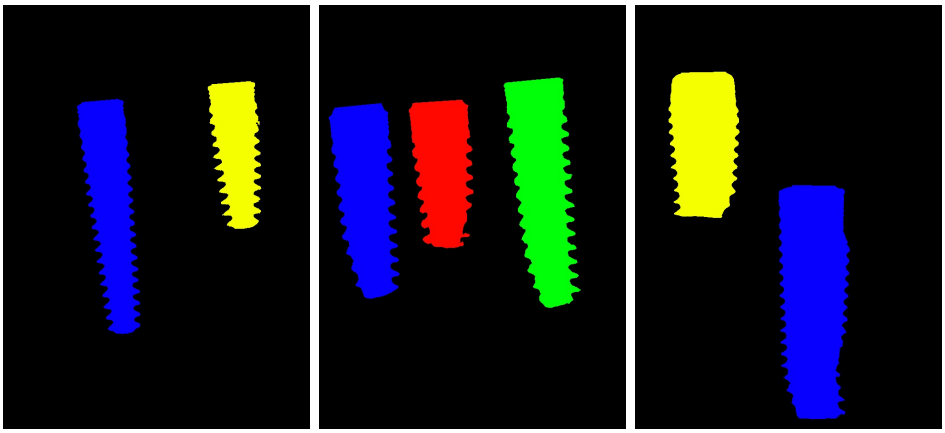
(e) (f)

Figure 4.10: **(Left)** Results of applying the proposed FCN-2 model for the purpose of automated ROI detection. **(Right)** Binary versions of the corresponding images on the left after a probability threshold of 0.5 has been applied.



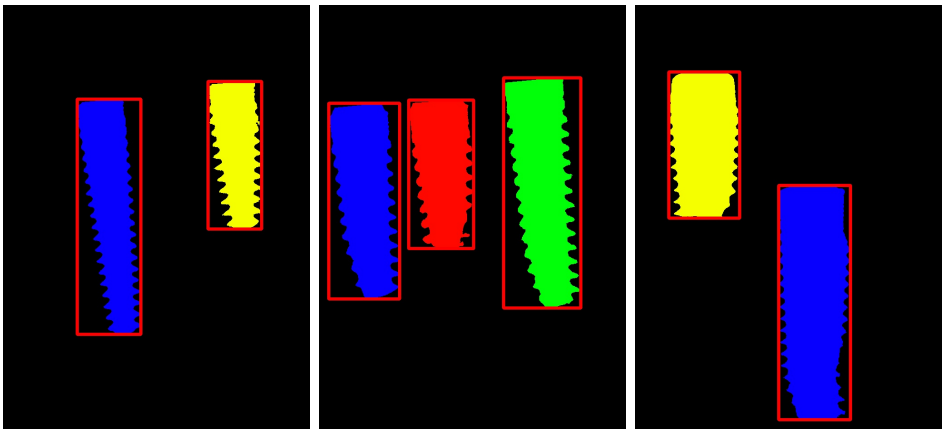
(a) (b) (c)

Figure 4.11: The automatically detected ROIs after post-processing operations have been applied to the binary images depicted in Figure 4.10.



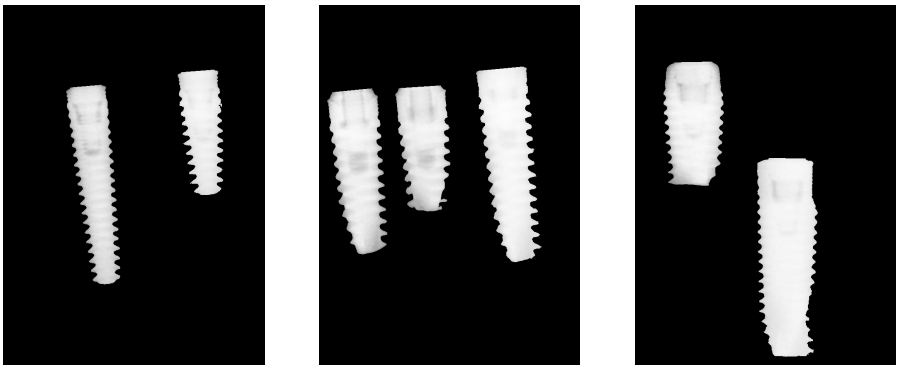
(a) (b) (c)

Figure 4.12: Examples of labelled connected components within the binary images, with each dental implant assigned a different label.



(a) (b) (c)

Figure 4.13: Examples of labelled connected components within the binary images, with each dental implant assigned a different label and delimited by a bounding box.



(a) (b) (c)

Figure 4.14: Actual X-ray images after the application of ROI-masking.

(a) (b) (c)

Figure 4.15: Segmented dental implant images corresponding to Figure 4.14 (b).

4.4 Concluding remarks

In this chapter the proposed semi-automated and fully automated image segmentation systems for the detection of ROIs associated with the dental implants within an actual X-ray image were discussed. Within the context of the semi-automated system suitable ROIs, that contain the dental implants, are manually specified (selected). For the fully automated system, suitable ROIs are automatically detected through a deep learning-based technique. A set of morphological post-processing techniques were conducted in order to eliminate noise and enhance the binary mask images. The post-processed binary masks were subsequently employed for ROI-masking purposes in order to segment the dental implants within an actual X-ray image into its constituent instances. The segmented dental implants serve as input for the dental implant classification framework discussed in Chapter 6.

Chapter 5

Dental implant modelling

5.1 Introduction

In this chapter a discussion is provided on the framework developed for the purpose of dental implant recognition. Within the context of this study, dental implant recognition is dichotomised as follows:

- Training: a model is trained on *simulated* X-ray images in order to learn a set of features associated with each of the *nine* different connection types investigated in this study.
- Testing: the trained model is applied to an *actual* X-ray image for the purpose of identifying the connection type associated with a questioned dental implant. The testing protocol will be discussed in Chapter 6.

A protocol is developed for extracting suitable features from simulated X-ray images for the purpose of identifying the connection type associated with a specific dental implant. The aforementioned protocol is based on extracting *learned* features through a deep learning-based technique. The deep learning-based model is trained on simulated X-ray images generated through the two dimensional (2D) parallel projection strategy discussed in Chapter 3.

The rationale behind employing a deep learning-based technique for the extraction of features within the context of dental implant recognition is the well-documented fact that deep learning algorithms have a powerful ability to learn abstract and complex features. Deep learning models, such as convolutional neural networks (CNNs), have been successfully applied to a wide range of image recognition tasks, including medical image analysis [5–10]. The well-documented success and proficiency of deep learning-based

algorithms in medical image analysis can potentially be extended to dental implant recognition as well. Overall, deep learning has the potential to transform and improve dental implant recognition by providing a highly accurate and efficient method for identifying and locating dental implants within X-ray images.

This chapter is organised as follows. In Section 5.2, the key concepts regarding dental implant recognition within the context of this study are presented. In Section 5.3, a description of the proposed fully convolution network (FCN-1) trained for the purpose of classifying the connection type associated with a specific dental implant within an actual X-ray image is provided. The dental implant recognition protocol proposed in this study is conceptualised in Figure 5.1.

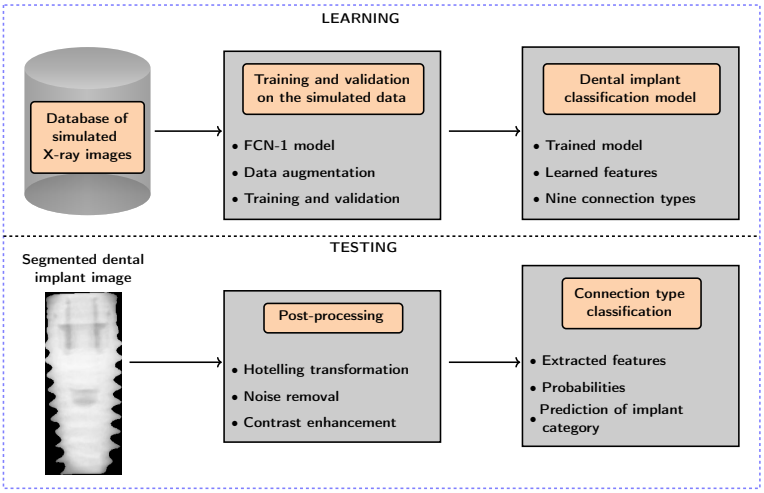


Figure 5.1: Schematic representation of the dental implant recognition protocol proposed in this study. The testing protocol will be discussed in detail in Chapter 6.

5.2 Overview

In this section the key aspects typically associated with dental implant representation and classification are discussed.

5.2.1 Dental implant characteristics

A dental implant is a medical device surgically inserted into the jawbone of a patient to replace a tooth's root. A typical dental implant consists of three main components, that is the dental implant body, the dental implant abutment and the dental implant crown (see Figure 5.2). The dental implant abutment is typically attached to the implant body by the abutment fixation screw and extends through the gums to support the attached artificial tooth.

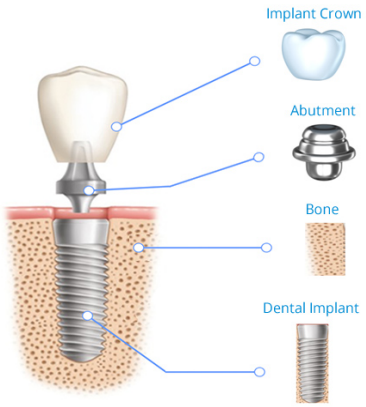


Figure 5.2: Structure of a dental implant [2].

The dental implant abutment connection can be categorised into three main types, that is an external connection, an internal connection, and a conical connection. An external connection constitutes a geometric feature that extends above the coronal surface of the implant (see Figure 5.3 (a)), while in the case of an internal connection the implant abutment connection is recessed into the body of the implant (see Figure 5.3 (b)). A conical connection constitutes a particular kind of internal connection with a conical, root-shaped geometry (see Figure 5.3 (c)). The dental implants depicted in Figure 5.3 are produced by the Make It Simple (MIS) manufacturing company.

The dental implants depicted in Figure 5.3 (a) have an *external* hexagon connection. This implant has apex grooves, and the apex has a dome shape. The implant body is tapered with buttress threads. The wide implant head constitutes the cervical part where the implant connects with the abutment. The dental implants depicted in Figure 5.3 (b) have an *internal* hexagon

connection. This implant has a sharp cutting tapered apex. The body of the implant is straight with a parallel V-shaped thread design. The implant head is cylindrical with a bevel geometry. The dental implants depicted in Figure 5.3 (c) have a *conical* connection. The implant connection is interthreaded with a six spline and morse taper inner geometry. The body of the implant is tapered with square threads and the head has microthreads at the crest module region. The implant apex has a flat shape with apical grooves.

The aforementioned geometrical attributes which describe the dental implants constitute the salient (discriminative) features learned during training for the purpose of assigning a questioned dental implant within an actual X-ray image to one of the *nine* different connection types investigated in this study.

(a) (b) (c)

Figure 5.3: Examples of dental implants. (a) External hexagon standard platform. (b) Internal hexagon standard platform. (c) Conical standard platform.

5.2.2 Modelling and feature extraction

In this study a fully convolutional network (FCN) model, the *FCN-1* model, is proposed for the purpose of classifying the connection type associated with a specific dental implant within an X-ray image. The FCN-1 model extracts (learns) the prominent *discriminative* features which describe the connection type associated with a specific dental implant. The aforemen-

tioned FCN-1 model constitutes a three dimensional (3D) convolutional neural network (CNN) which takes as input the simulated X-ray images acquired through the 2D parallel projection algorithm outlined in Chapter 3. The FCN-1 model employs DeepVoxNet [45] which is a deep learning processing framework for Keras developed in the Medical Imaging Research Center (MIRC) at KU Leuven for the efficient processing of 3D medical images.

The discriminative features learned during training are employed in the process of classifying the connection type associated with a specific dental implant within a questioned actual X-ray image. Morphological post-processing techniques are also applied to the questioned actual X-ray image during testing (see Chapter 6).

5.3 Feature extraction

In this section an in-depth description of the proposed feature extraction protocol which is based on deep learning is provided. The data, data augmentation protocol, model architecture and model training employed in this study are discussed in detail in the following sections.

Data

The simulated X-ray data set acquired through the proposed data generation protocol (see Chapter 3) is employed for training and validating the proposed FCN-1 model. The FCN-1 model is trained for the purpose of automatically identifying the connection type associated with a specific dental implant within an actual X-ray image.

The proposed strategy for generating the artificial training samples is based on the calculation of 2D projections (from a number of different angles) of a 3D volumetric representation of a dental implant. For each 3D volumetric representation of a dental implant three rotation axes are selected. These axes are orthogonal to the 3D volumetric representation and defined relative to it. The projections are acquired after rotating the 3D volumetric representation. The rotation of the 3D object is achieved by implementing three 3×3 rotation matrices R_p, R_y, R_r. Each matrix is defined by a single rotation angle θ_p, θ_y, and θ_r. These matrices affect rotations about the pitch, yaw, and roll axes respectively. After placing the 3D object in the initial position, 2D parallel projections are calculated as the 3D volumetric representation is rotated about one of the yaw, pitch or roll axes (see Figure 5.4).

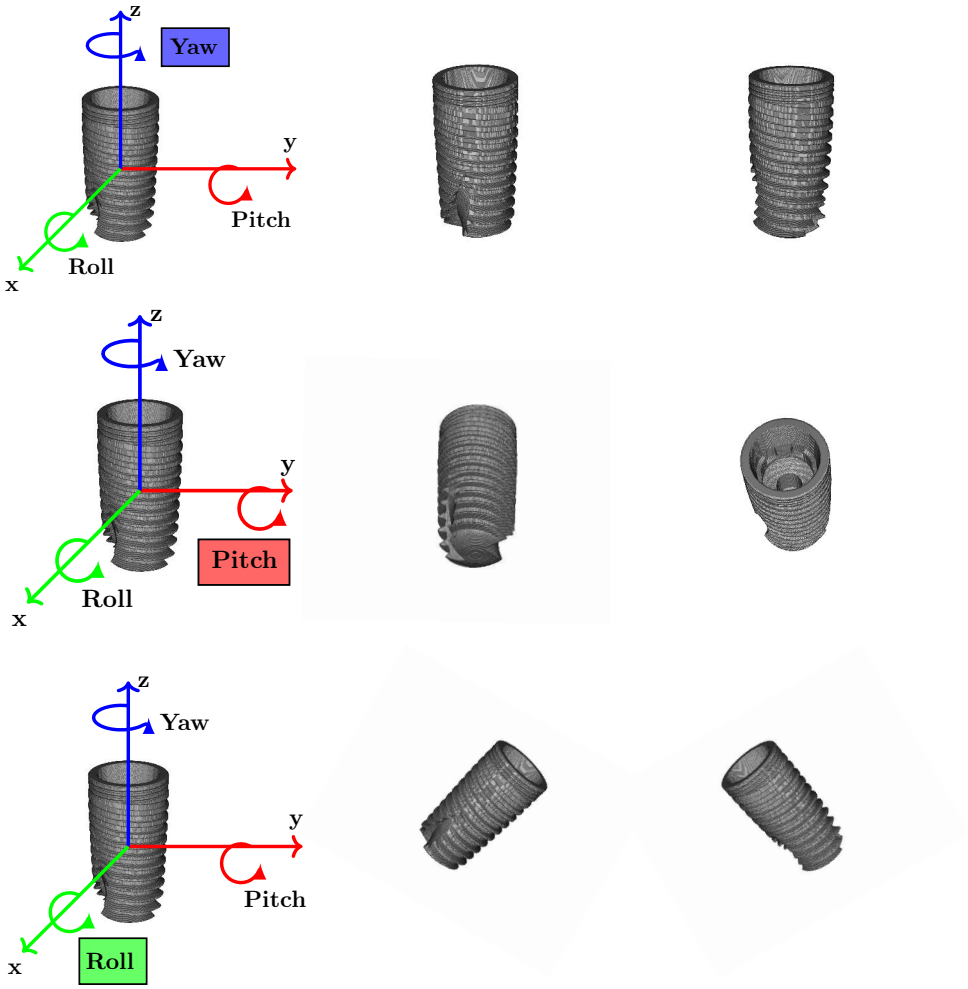


Figure 5.4: Conceptualisation of the rotation of a 3D object about the three different axes. The 3D volumetric representation resides in the world coordinate system (x, y, z), and its centre point coincides with the origin. This defines the relative zero rotation angle with respect to the initial position. *Rows*: Each row depicts a rotation of the 3D volumetric representation about a specific axis. The rotation axes are orthogonal to the 3D object as depicted by the arrows. The bounded box (in the first column) corresponds to the rotation axis associated with the relevant row. *First row*: Axial rotations which are associated with rotations about the yaw axis. *Second row*: Out-of-plane rotations which are associated with rotations about the pitch axis. *Third row*: In-plane rotations which are associated with rotations about the roll axis.

The 3D volumetric representation is rotated using the classical Euler sequence [46], that is the yaw–pitch–roll sequence. The yaw, pitch and roll rotation axes are orthogonal to the 3D volumetric representation and defined relative to it as depicted by the arrows in Figure 5.4. The 3D volumetric representation is first rotated about the pitch and the yaw axes and finally about the roll axis (during network training).

Similar to the Euler sequence, the *first* rotation is about the z axis, that is the yaw axis (see Figure 5.4). These rotations are also referred to as *axial* rotations. The *second* rotation is about y axis, that is the pitch axis (see Figure 5.4). These rotations are also referred to as *out-of-plane* rotations.

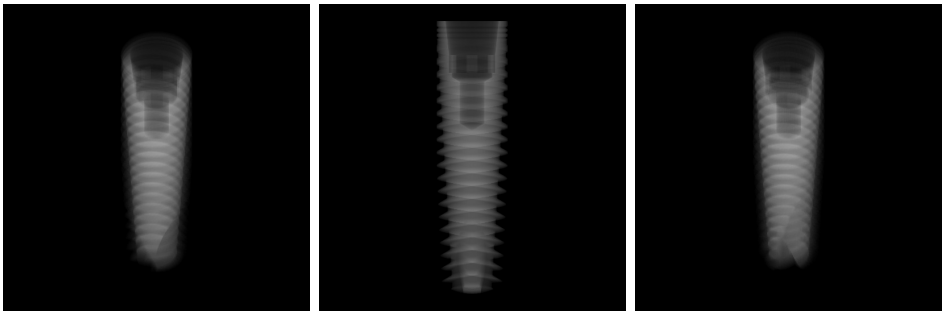
It is important to note that, after a 3D volumetric representation has been reorientated by implementing the above two rotations, a 2D projection of the reorientated 3D representation is calculated from a specific angle in order to obtain a simulated X-ray image. These rotations are conducted *statically*, that is *before* network training. In Figure 5.5, examples of the simulated X-ray images acquired after the application of axial and out-of-plane rotations are presented.

The resulting simulated X-ray image is finally rotated through an arbitrary angle. This is equivalent to rotating the 3D volumetric representation about the x axis or roll axis and is referred to as an *in-plane* rotation. These in-plane rotations are conducted *on the fly* during the data augmentation protocol that forms part of training the proposed model for the purpose of automatically classifying the connection type associated with a specific dental implant from an X-ray image.

In this study, dental implants with *nine* connection types are investigated. Experiments are conducted on the dental implants produced by MIS. The three main connection types investigated in this study, that is the external, internal and conical connections, can be further characterised as either a narrow, standard or wide platform.

Dental implants with an external diameter of 3.30 mm are referred to as "narrow platform" (NP), while implants with external diameters of 3.75 mm, 3.90 mm, 4.20 mm and 4.30 mm are referred to as "standard platform" (SP). Implants with external diameters of 5.0 mm and 6.00 mm are referred to as "wide platform" (WP). Only dental implants with the aforementioned diameters are considered in this study. The implant lengths considered in this study are 6 mm, 8 mm, 10 mm, 11 mm, 13 mm and 16 mm.

The dental implants are categorised into *nine* connection types according to the implant's characteristics as follows:



(a) (b) (c)

Figure 5.5: (a) A simulated X-ray image that underwent an out-of-plane rotation of 40°. (b) A simulated X-ray image that underwent an axial rotation of 40°. (c) A simulated X-ray image that underwent axial and out-of-plane rotations of 40°.

- *Four* connection types for dental implant models V3 and C1 (conical):
  - NP - implant model V3
  - NP - implant model C1
  - SP - implant models V3 and C1
  - WP - implant model C1
- *Three* connection types for dental implant models SEVEN and M4 (internal hexagon):
  - NP - implant models SEVEN and M4
  - SP - implant models SEVEN and M4
  - WP - implant models SEVEN and M4
- *Two* connection types for dental implant model LANCE (external hexagon):
  - SP - implant model LANCE
  - WP - implant model LANCE

A summary of the dental implants considered in this study is presented in Table 5.1. As depicted in Table 5.1, each unique connection type is assigned an integer value between 1 and 9. This process is called ordinal encoding or integer encoding. The classes are mutually exclusive and each implant is assigned to one (and only one) label.

Connection type	**Dental implant type**	**Length (mm)**
(1) Conical narrow platform (V3)	**V3:** External diameter 3.30 mm Internal diameter 2.75 mm	10, 11, 13, 16
(2) Conical narrow platform (C1)	**C1:** External diameter 3.30 mm Internal diameter 2.75 mm	10, 11, 13, 16
(3) Conical standard platform	**V3:** External diameter 3.90 mm Internal diameter 3.15 mm	8, 10, 11, 13, 16
	V3: External diameter 4.30 mm Internal diameter 3.15 mm	8, 10, 11, 13, 16
	V3: External diameter 5.00 mm Internal diameter 3.15 mm	8, 10, 11, 13, 16
	C1: External diameter 3.75 mm Internal diameter 3.15 mm	8, 10, 11, 13, 16
	C1: External diameter 4.20 mm Internal diameter 3.15 mm	8, 10, 11, 13, 16
(4) Conical wide platform	**C1:** External diameter 5.00 mm Internal diameter 4.00 mm	8, 10, 11, 13, 16
(5) Internal hex narrow platform	**SEVEN:** External diameter 3.30 mm Internal diameter 2.10 - 3.30 mm	10, 11, 13, 16
	M4: External diameter 3.30 mm Internal diameter 2.10 - 3.30 mm	10, 11, 13, 16
(6) Internal hex standard platform	**SEVEN:** External diameter 3.75 mm Internal diameter 2.45 - 3.75 mm	8, 10, 11, 13, 16
	SEVEN: External diameter 4.20 mm Internal diameter 2.45 - 3.75 mm	6, 8, 10, 11, 13, 16
	M4: External diameter 3.75 mm Internal diameter 2.45 - 3.75 mm	8, 10, 11, 13, 16
	M4: External diameter 4.20 mm Internal diameter 2.45 - 3.75 mm	6, 8, 10, 11, 13, 16
(7) Internal hex wide platform	**SEVEN:** External diameter 5.00 mm Internal diameter 2.45 - 4.50 mm	6, 8, 10, 11, 13, 16
	SEVEN: External diameter 6.00 mm Internal diameter 2.45 - 4.50 mm	6, 8, 10, 11, 13
	M4: External diameter 5.00 mm Internal diameter 2.45 - 4.50 mm	6, 8, 10, 11, 13, 16
	M4: External diameter 6.00 mm Internal diameter 2.45 - 4.50 mm	6, 8, 10, 11, 13
(8) External hex standard platform	**LANCE:** External diameter 3.75 mm Internal diameter 2.70 mm	10, 11.5, 13, 16
	LANCE: External diameter 4.20 mm Internal diameter 2.70 mm	8, 10, 11.5, 13, 16
(9) External hex wide platform	**LANCE:** External diameter 5.00 mm	8, 10, 11.5, 13, 16

Table 5.1: A summary of the connection type and corresponding geometrical features associated with each MIS dental implant investigated in this study. The boldfaced phrases are the names of the dental implant models.

The simulated X-ray data set is partitioned into training and validation sets for the purpose of assigning a questioned dental implant within an actual X-ray image to one of the nine connection types investigated in this study.

The training set (seen data) is used to learn the parameters (weights) for the proposed FCN-1 model, while the validation set is used for avoiding overfitting by enforcing a stopping criterion.

The proposed FCN-1 model is fed training samples (X, l) with X constituting the input images, each of size 540×540 pixels, and l the associated ground truth label, that is the class assigned to the connection type associated with a specific dental implant. Data preprocessing is implemented for the purpose of transforming the input X into a tuple of four dimensional (4D) arrays with three spatial dimensions and one feature dimension, that is a volume V of dimensions {width × height × depth}, where {width × height} represents the spatial dimensions of the input image and depth the number of colour channels. The input X therefore has a shape defined as $(540, 540, 1, 1)$.

A total of 403 200 simulated X-ray images are generated through the application of the axial and out-of-plane rotations. The background of the artificially generated X-ray images is set to black. A total of 322 560 simulated X-ray images (80%) are used for training purposes, while 80 640 simulated X-ray images (20%) are used for validation purposes. A few training samples from the artificially generated data set are presented in Figure 5.6.

Data augmentation

Data augmentation techniques in deep learning can be applied either statically or on the fly. Static data augmentation refers to appending the augmented data to the training data set and using the augmented data set for training the model. In this study, only in-plane rotations are conducted on the fly. The artificially generated data set is augmented in each batch while the model is being trained. On the fly augmentation is more efficient in terms of computational and storage requirements.

During training, the simulated X-ray images are augmented by in-plane rotations of maximally 60°. An in-plane rotation occurs about the so-called roll axis (see Figure 5.4) and is associated with rotations in the yz plane. A typical rotation about the x axis is depicted in Figure 5.7 and implemented by applying the following rotation matrix,

$$R_x(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}.$$

An affine transformation is applied for the purpose of rotating the simulated X-ray images. The above rotation matrix is therefore used to generate

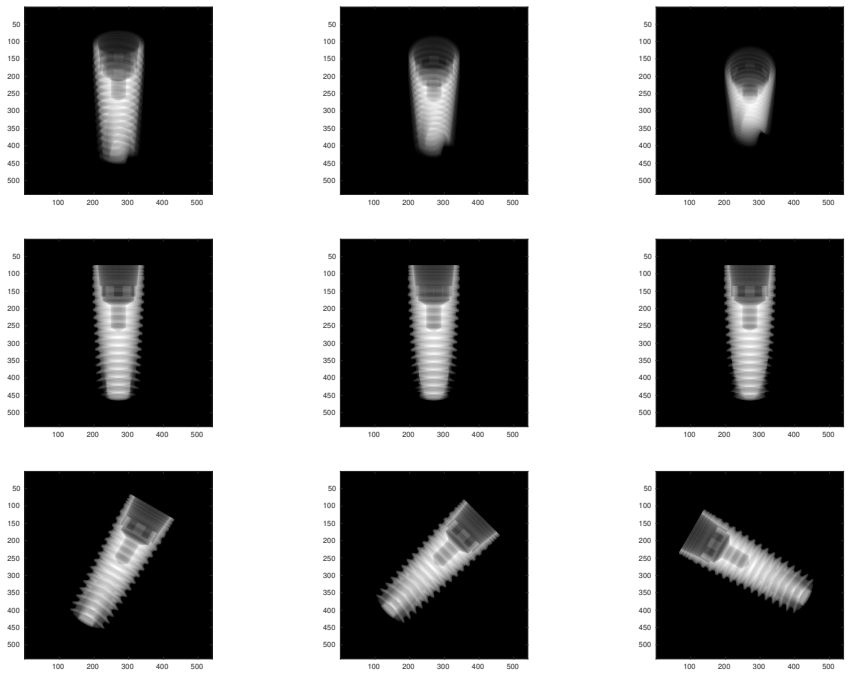


Figure 5.6: **(Top)** Examples of simulated X-ray images that underwent out-of-plane rotations of 30°, 45° and 60°. **(Middle)** Examples of simulated X-ray images that underwent axial rotations of 30°, 45° and 60°. **(Bottom)** Examples of simulated X-ray images that underwent in-plane rotations of 30°, 45° and 60°.

in-plane rotations of maximally 60°. In Figure 5.8, examples of the simulated X-ray images acquired after the application of axial, out-of-plane and in-plane rotations through a number of different angles are presented.

Architecture

The proposed FCN-1 model is trained on simulated X-ray images for the purpose of assigning a questioned dental implant within an actual X-ray image to one of *nine* different connection types. The aforementioned FCN-1 model consists of twelve convolutional layers, where each of these layers is followed by a rectified linear unit (ReLU) and maximum pooling layer. The first ten convolutional layers employ 32 kernels (filters) of size 3×3×1 pixels. Each pooling layer has a size of 2×2×1 pixels. A dropout layer

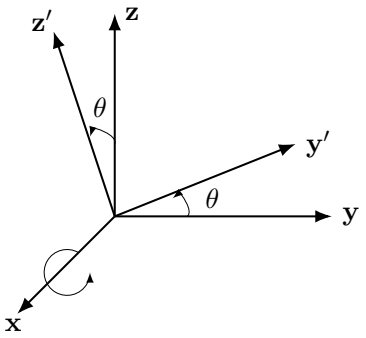
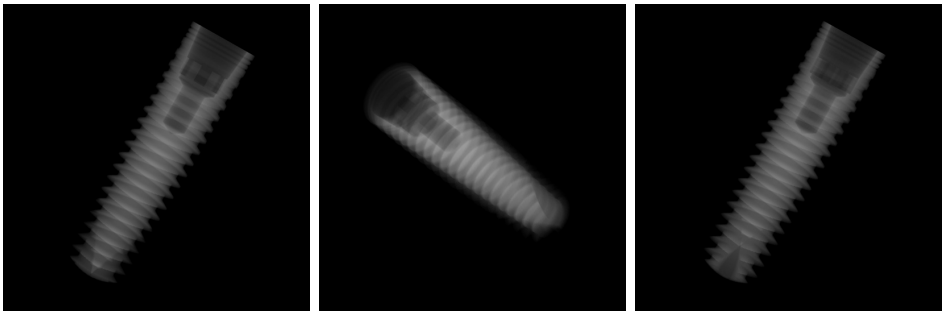


Figure 5.7: A typical rotation about the x axis.



(a) (b) (c)

Figure 5.8: (a) A simulated X-ray image that underwent an in-plane rotation of 40°. (b) A simulated X-ray image that underwent an out-of-plane rotation of 40° and an in-plane rotation of 60°. (c) A simulated X-ray image that underwent axial and in-plane rotations of 40°.

with a dropout rate of 50% is added before the final layer. The output of the final convolutional layer is fed to a softmax function which results in a probability distribution across the *nine* classes investigated in this study. The architecture of the proposed FCN-1 model is depicted in Figure 5.9.

The FCN-1 model takes as input the preprocessed 4D tensor data. The 3D convolution filter applies a convolutional operation (dot product of tensor) to produce the feature map. The 3D convolutional layer can be formulated as follows [47],

$$v_{ij}^{xyz} = b_{ij} + \sum_{m=1}^{M_{(i-1)}} \sum_{p=0}^{P_i-1} \sum_{q=0}^{Q_i-1} \sum_{r=0}^{R_i-1} w_{ijm}^{pqr} v_{(i-1)m}^{(x+p)(y+q)(z+r)}, \tag{5.1}$$

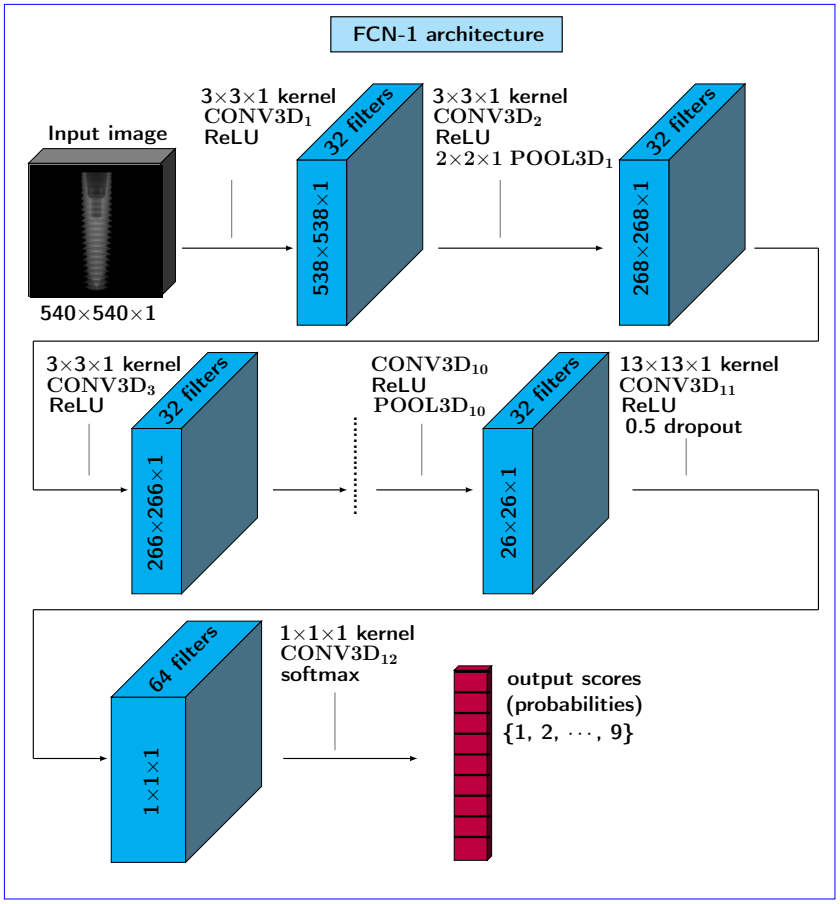


Figure 5.9: A depiction of the FCN-1 model architecture employed for the purpose of assigning a questioned dental implant within an actual X-ray image to one of *nine* different connection types investigated in this study.

where v_{ij}^{xyz} is the output at the position $(x,\ y,\ z)$, b_{ij} is the bias for the feature map, $M_{(i-1)}$ is the number of feature maps at the $(i-1)$-th layer, P_i, Q_i and R_i denote the size of the 3D filter at the i-th layer, w_{ijm}^{pqr} is the weight at position (p, q, r) of the kernel of the m-th feature map and $v_{(i-1)m}^{(x+p)(y+q)(z+r)}$ represents the value of the m-th feature map at the position $(x+p,\ y+q,\ z+r)$ in the $(i-1)$-th layer. The weights and biases of each filter are trained to extract the salient features associated with the dental implant connection types from the input data.

Stride, padding and filter size are a few common hyper-parameters defin-

ing convolutional operations. Stride denotes the step size that filters move at a time. Maximum pooling layers are added between successive convolutional layers in the model. The pooling layers progressively reduce the spatial size of the data through downsampling. The 3D pooling layer associated with the m-th feature map and a pooling kernel of size $k \times k \times k$ can be formulated as follows,

$$v_{ij}^{xyz} = \max_{p,q,r \in \{0,1,\ldots,k-1\}} v_{(i-1)m}^{(kx+p)(ky+q)(kz+r)}. \tag{5.2}$$

The activation layers are employed for the purpose of introducing nonlinearity into the model. The ReLU activation function takes the input tensor and performs an element-wise nonlinear transformation $g(v) = \max(0, v)$, where g is the activation function and v is the output feature map as defined in Equation 5.1.

A detailed summary of the proposed network architecture is provided in Table 5.2.

Layer	Activation map	Kernel size	Output shape	Trainable parameters
Input	$540 \times 540 \times 1 \times 1$			
CONV3D_1	$540 \times 540 \times 1 \times 32$	$3 \times 3 \times 1$	$538 \times 538 \times 1 \times 32$	320
CONV3D_2	$538 \times 538 \times 1 \times 32$	$3 \times 3 \times 1$	$536 \times 536 \times 1 \times 32$	9 248
POOL3D_1	$536 \times 536 \times 1 \times 32$	$2 \times 2 \times 1$	$268 \times 268 \times 1 \times 32$	
CONV3D_3	$268 \times 268 \times 1 \times 32$	$3 \times 3 \times 1$	$266 \times 266 \times 1 \times 32$	9 248
CONV3D_4	$266 \times 266 \times 1 \times 32$	$3 \times 3 \times 1$	$264 \times 264 \times 1 \times 32$	9 248
POOL3D_2	$264 \times 264 \times 1 \times 32$	$2 \times 2 \times 1$	$132 \times 132 \times 1 \times 32$	
CONV3D_5	$132 \times 132 \times 1 \times 32$	$3 \times 3 \times 1$	$130 \times 130 \times 1 \times 32$	9 248
CONV3D_6	$130 \times 130 \times 1 \times 32$	$3 \times 3 \times 1$	$128 \times 128 \times 1 \times 32$	9 248
POOL3D_3	$128 \times 128 \times 1 \times 32$	$2 \times 2 \times 1$	$64 \times 64 \times 1 \times 32$	
CONV3D_7	$64 \times 64 \times 1 \times 32$	$3 \times 3 \times 1$	$62 \times 62 \times 1 \times 32$	9 248
CONV3D_8	$62 \times 62 \times 1 \times 32$	$3 \times 3 \times 1$	$60 \times 60 \times 1 \times 32$	9 248
POOL3D_4	$60 \times 60 \times 1 \times 32$	$2 \times 2 \times 1$	$30 \times 30 \times 1 \times 32$	
CONV3D_9	$30 \times 30 \times 1 \times 32$	$3 \times 3 \times 1$	$28 \times 28 \times 1 \times 32$	9 248
CONV3D_{10}	$28 \times 28 \times 1 \times 32$	$3 \times 3 \times 1$	$26 \times 26 \times 1 \times 32$	9 248
POOL3D_5	$26 \times 26 \times 1 \times 32$	$2 \times 2 \times 1$	$13 \times 13 \times 1 \times 32$	
CONV3D_{11}	$13 \times 13 \times 1 \times 64$	$13 \times 13 \times 1$	$1 \times 1 \times 1 \times 64$	346 176
Dropout (50%)	$1 \times 1 \times 1 \times 64$		$1 \times 1 \times 1 \times 64$	
CONV3D_{12}	$1 \times 1 \times 1 \times 9$	$1 \times 1 \times 1$	$1 \times 1 \times 1 \times 9$	585

Table 5.2: The network architecture and hyper-parameters employed by the proposed FCN-1 model.

The first convolutional layer (CONV3D_1) processes an input image of size $540 \times 540 \times 1 \times 1$ with 32 different kernels (filters). Each kernel has a size

of 3×3×1 pixels. The layer therefore contains $540 \times 540 \times 1 = 291\ 600$ neurons, with $((3 \times 3 \times 1 \times 1) + 1) \times 32 = 320$ trainable parameters (weights). Non-linearity is introduced into the network through the application of ReLU layers.

The output feature map of the CONV3D_1 layer has a size of 538×538×1 for each of the 32 filters. The aforementioned feature map serves as the input of the second convolutional layer (CONV3D_2). The CONV3D_2 layer has $((3 \times 3 \times 1 \times 32) + 1) \times 32 = 9\ 248$ trainable parameters. The resulting activation maps, each of size 536×536×1 are subsequently subjected to max pooling (POOL3D_1) by considering kernels (filters) of size 2×2×1 pixels. This results in an output tensor of size 268×268×1×32.

The output of each convolutional layer constitutes the learned feature maps of the input. Only the convolutional layers have trainable parameters. The first ten convolutional layers, each employs 32 kernels (filters) of size 3×3×1 pixels. Each pooling layer has a size of 2×2×1 pixels.

The eleventh and twelfth convolutional layers (CONV3D_{11} and CONV3D_{12}) employ 64 and 9 different kernels (filters) respectively. Within the context of CONV3D_{11}, each kernel has a size of 13×13×1 pixels, while for CONV3D_{12}, each kernel has a size of 1×1×1 pixels. The CONV3D_{11} layer is followed by a ReLU layer, while no pooling layers are employed within the context of CONV3D_{11} and CONV3D_{12}.

A dropout layer with a dropout rate of 50% is added after the CONV3D_{11} layer. The output of the final layer, that is the CONV3D_{12} layer is transformed into a 1D vector of length nine since there are nine connection types. The activation function of the final layer is a softmax function, also referred to as the normalised exponential function which takes the 1D input vector and maps it to a probability distribution, with each value (probability) in the range (0,1), where the aforementioned values (probabilities) sum to 1. The softmax function is implemented as follows,

$$\tilde{y}_i = \frac{\exp(\mathbf{z}_i)}{\sum_{k=1}^{K} \exp(\mathbf{z}_k)} \qquad \text{for } i = 1, \ldots, 9, \tag{5.3}$$

where the vector $\mathbf{z} = [z_1, z_2, ..., z_k]$ is the output of the final convolutional layer CONV3D_{12}. The softmax function of the input vector $\mathbf{z} = [z_1, z_2, ..., z_9]$ is therefore also a vector,

$$\text{softmax}(\mathbf{z}) = \left[\frac{\exp(z_1)}{\sum_{k=1}^{K} \exp(z_k)}, \frac{\exp(z_2)}{\sum_{k=1}^{K} \exp(z_k)}, \ldots, \frac{\exp(z_9)}{\sum_{k=1}^{K} \exp(z_k)} \right]. \tag{5.4}$$

These probabilities are used to assign an input image to one of *nine* disjoint classes (during the evaluation phase) and for the computation of the loss function (during the training phase). The predicted class is decided by the maximum probability in the output. Subsequently, at each training iteration, categorical cross entropy is used to measure the difference between the predicted class and the true label. The loss function is then minimised and back propagated to adjust the weights.

Model training

When training the FCN-1 model, the predicted value and the ground truth associated with the training data set are used to calculate the loss value. During training the weights are updated through the stochastic gradient descent (SGD) optimisation technique. Back propagation is used to adjust the network parameters. During training the network loss is minimised until it converges. The categorical cross entropy ($\mathscr{L}_{\text{CCE}}$) between the network's predictions and the ground truth labels is adopted as a loss function and computed as follows,

$$\mathscr{L}_{\text{CCE}} = -\frac{1}{N}\sum_{n=1}^{N}\sum_{k=1}^{K} y_i^* \log(\tilde{y}_i), \tag{5.5}$$

where N and K denote the number of training samples and the number of classes respectively, y_i^* the i-th true class associated with the training sample and $\tilde{y}_i$ the predicted class.

The proposed FCN-1 model is trained using stochastic gradient descent with a momentum value of $\gamma = 0.9$ and an initial learning rate of $\epsilon = 0.001$. The kernels are updated as follows,

$$\mathbf{w}_{i+1} = \mathbf{w}_i - \epsilon\gamma_{i+1}. \tag{5.6}$$

A cross validation algorithm (see Chapter 7) is used to gauge the proficiency of the proposed model, as well as its capability to generalise to unseen data.

5.4 Concluding remarks

In this chapter, the training and validating protocol of the proposed FCN-1 model for the purpose of distinguishing between different types of dental implants in X-ray images, primarily based on the implant's external shape,

were presented. The proposed FCN-1 model was trained on artificially generated (simulated) X-ray images for the eventual purpose of assigning a questioned dental implant within an actual X-ray image to one of the nine connection types investigated in this study. A quantitative analysis of the proficiency of the proposed FCN-1 model will be conducted in Chapter 7.

Chapter 6

Dental implant classification

6.1 Introduction

In this study, the proposed dental implant classification/recognition protocol is based on a deep learning technique. The proposed dental implant classification/recognition protocol learns discriminative features for the purpose of distinguishing between different types of dental implants in X-ray images. The features are primarily based on the implant's external shape. In this study the task of distinguishing between different connection types uses *supervised learning*. Supervised learning refers to predictive modelling where a class label is predicted for given input data. The model learns to predict the output variables from the input variables using labelled data. Within the context of classification the model assigns the output variable to one of several discrete classes. The classification task can be binary, multi-class or multi-labelled. Multi-class refers to a classification task that has more than two class labels and multi-labelled implies that more than one class exists in the input data.

In this study the dental implant classification task is a multi-class problem. The connection types are mutually exclusive and each implant is assigned to one label. Within the context of the dental implant data set employed in this study, there exists a one-to-one correspondence between the external shape of an implant and the internal connection of the implant.

The proposed fully convolutional network (FCN) model, that is the *FCN-1* model (outlined in Chapter 5) which was trained and validated on artificially generated X-ray images is employed for testing purposes. The trained FCN-1 model is applied to a questioned actual X-ray image in order to predict the specific connection type associated with the questioned dental implant. The

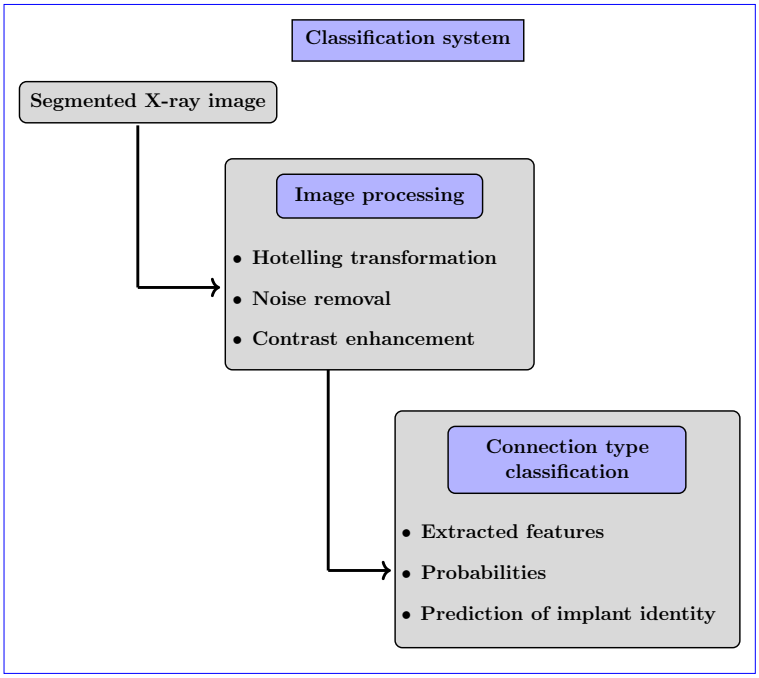


Figure 6.1: Conceptualisation of the proposed dental implant classification protocol.

proposed FCN-1 model considers an actual X-ray image which contains only a single dental implant as input. The actual X-ray images are first segmented through the segmentation protocol discussed in Chapter 4. During testing the segmented actual X-ray images are subjected to image processing techniques in order to render the segmented actual X-ray images similar to the simulated dental implants before being presented to the trained model.

In this chapter the testing protocol employed within the context of dental implant classification/recognition is presented. In Section 6.2 the image processing techniques employed for transforming the actual X-ray images are discussed. Each preprocessed questioned X-ray image is presented to the learned classification model which subsequently predicts the connection type (see Section 6.3). The proposed dental implant classification protocol is depicted in Figure 6.1.

6.2 Image processing

The segmented dental implants acquired through the proposed semi-automated or fully automated systems (see Chapter 4) serve as input for the dental implant classification framework. The following image processing techniques are applied to the segmented X-ray images in order to render the actual X-ray images similar to the simulated X-ray images.

Hotelling transform

The Hotelling transform [11] is applied to each questioned dental implant image for the purpose of eliminating in-plane rotations. The Hotelling transform, also referred to as the Karhunen–Loève transform [48, 49], eliminates rotational variations in the data. The Hotelling transform is therefore used to improve the robustness of an algorithm within the context of rotational and translational variations. The Hotelling transform is equivalent to principal components analysis (PCA), a technique which is widely used in image processing and data analysis.

The Hotelling transform is based on the statistical properties of a vector representation of the data. Given a population of random n-dimensional vectors [11],

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \tag{6.1}$$

the mean vector of the population is defined as

$$\mathbf{m_x} = E\{\mathbf{x}\}, \tag{6.2}$$

where E denotes the expected value of the population. The covariance matrix of the population is defined as

$$\mathbf{C_x} = E\{(\mathbf{x} - \mathbf{m_x})(\mathbf{x} - \mathbf{m_x})^T\}. \tag{6.3}$$

Element c_{ii} of $\mathbf{C_x}$ is the variance of x_i, that is the i-th component of the vectors in the population, and element c_{ij} is the covariance between components x_i and x_j. The matrix $\mathbf{C_x}$ is symmetric and real.

For a sample of K vectors from a random population, the mean vector and covariance matrix can be estimated as follows,

$$\mathbf{m_x} = \frac{1}{K}\sum_{k=1}^{K}\mathbf{x}_k, \tag{6.4}$$

$$\mathbf{C_x} = \frac{1}{K}\sum_{k=1}^{K}\mathbf{x}_k\mathbf{x}_k^T - \mathbf{m_x}\mathbf{m_x}^T. \tag{6.5}$$

Since $\mathbf{C_x}$ is symmetric and real it is always possible to find a set of n orthonormal eigenvectors. If $\mathbf{A}$ is a matrix of which the rows are the eigenvectors of $\mathbf{C_x}$, ordered in such a way that the first row of $\mathbf{A}$ is the eigenvector that corresponds to the largest eigenvalue of $\mathbf{C_x}$ and the last row of $\mathbf{A}$ is the eigenvector that corresponds to the smallest eigenvalue of $\mathbf{C_x}$, $\mathbf{A}$ and its transpose can be depicted as follows,

$$\mathbf{A} = \begin{bmatrix} \mathbf{e}_1^T \\ \mathbf{e}_2^T \\ \vdots \\ \mathbf{e}_n \end{bmatrix} \text{ and } \mathbf{A^T} = \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_n \end{bmatrix}, \tag{6.6}$$

where $\mathbf{e}_i$, $i = 1, 2, \ldots, n$ denote the eigenvectors of $\mathbf{C_x}$.

Now suppose that $\mathbf{A}$ is a transformation matrix that maps a vector $\mathbf{x}$ to a vector $\mathbf{y}$ using the following transformation,

$$\mathbf{y} = \mathbf{A}(\mathbf{x} - \mathbf{m_x}). \tag{6.7}$$

The above equation constitutes the *Hotelling transform*. The mean of the transformed $\mathbf{y}$ vectors is zero, while $\mathbf{C_x}$ and $\mathbf{C_y}$ have the same eigenvalues.

The Hotelling transform performs a linear transformation of a set of n-dimensional vectors that decorrelates the n coordinates. When the Hotelling transform is applied to an image, the resulting image is geometrically aligned along the principal axes. From the point of view of PCA, the first principal component (PC1) is the linear combination that accounts for the maximum variance. Geometrically, it corresponds to the direction of the longest axis of the object. That is, the geometric effect of the transformation given in Equation 6.7 on an object in an image constitutes both a translation and a rotation, so that the centroid of the object (the average of the x and y coordinates of all the object pixels) is translated to the origin and the object is rotated by an angle that minimises its moment of inertia [50]. Within the

context of image recognition this transformation is often important, since aligning the image with its principal axes eliminates the effects of translation and rotation in the analysis.

Figure 6.2 depicts examples of segmented actual X-ray images acquired through the proposed semi-automated or fully automated systems outlined in Chapter 4. Each segmented actual X-ray image contains a single dental implant. Figure 6.3 depicts examples of the resulting images after the Hotelling transform was applied to the images in Figure 6.2. Note that the general shape of the object is preserved, and that the original implant is only translated and rotated in a convenient way.

Spatial filtering and intensity transformation

Within the context of image recognition, the proficiency of the recognition system is significantly affected by the quality of the images. The three primary factors that determine the quality of an X-ray image are *noise*, *contrast* and *artefacts*.

An artefact in an X-ray image is a feature that does not correlate with the physical properties of the subject being imaged and may confound or obscure interpretation of the X-ray image [51]. In X-ray images noise may come into being as a result of the quantum noise properties of photons and the electronic noise associated with the detector system. The quantum noise is random as it relates to the number of photons detected, while the electronic noise originates from the X-ray detector system. Noise in X-ray images is often referred to as quantum mottle. Radiographic contrast constitutes the density difference in the signal or brightness between the structure of interest and its surroundings in a radiograph. High radiographic contrast is observed in radiographs where density differences are notably distinguished (black versus white). Low radiographic contrast is evident in radiographic images where adjacent regions have a low density difference (black versus gray).

Within the context of X-ray images, there are a number of different filters that can be applied for the purpose of noise reduction. The grayscale intensity can be adjusted using image post-processing techniques where pixel values are transformed in order to provide the expected range of contrast depending upon the specific clinical requirements. The contrast-to-noise ratio (CNR) and the signal-to-noise ratio (SNR) are the basic measures mostly considered in X-ray images. The CNR is used to estimate the contrast between the tissue of interest and the background (i.e. the neighbouring tissue), while the SNR estimates the ratio between the intensity of the signal and

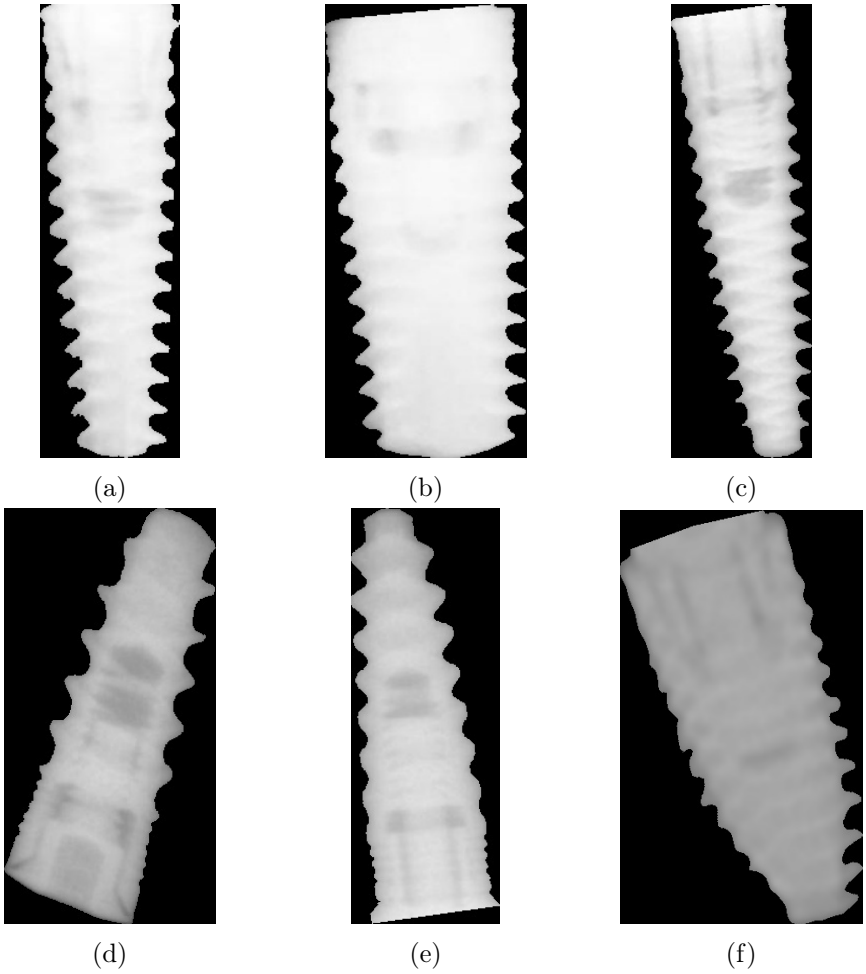


Figure 6.2: Examples of segmented actual X-ray images containing dental implants. (**Top**) Implants inserted into *pig jaws*. (**Bottom**) Implants inserted into *human jaws*.

the intensity of the background noise.

In this study an ensemble of grayscale intensity transformation and spatial filtering techniques are implemented for the purpose of enhancing the contrast and suppressing noise in the actual X-ray images.

Within the context of noise reduction, the application of a *Gaussian lowpass filter* [52] to the actual X-ray images is deemed appropriate. A Gaussian lowpass filter is therefore employed to smoothen each questioned

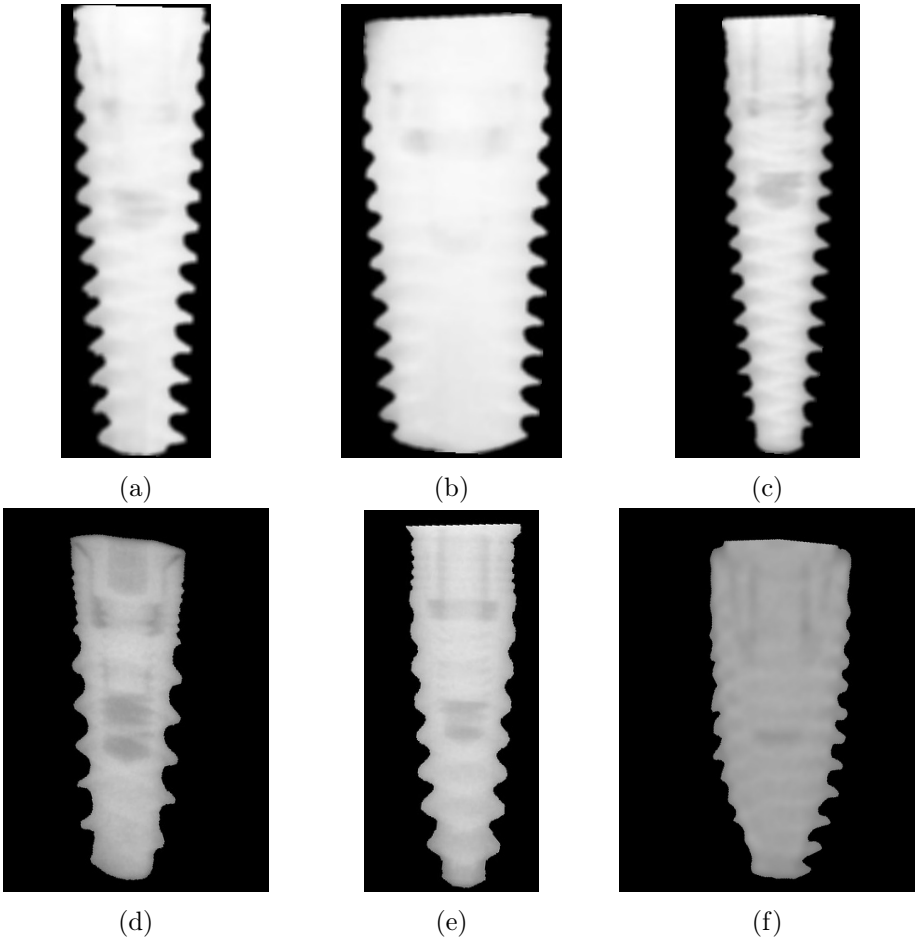


Figure 6.3: Examples of actual X-ray images after the Hotelling transform was applied to the images depicted in Figure 6.2. (**Top**) Implants within the context of the pig data set. (**Bottom**) Implants within the context of the human data set.

dental implant image.

Within the context of the actual X-ray images, a total of 483 labelled and unlabelled images, which contain implants inserted into either human or pig jaws, are considered. The aforementioned actual X-ray images are partitioned into two independent sets. The data partitioning protocol will be discussed in more detail in Chapter 7.

Different parameters for the Gaussian kernel size are employed for the

purpose of smoothing the pig and human implant images. Since the noise levels are more significant in the human implants than is the case for the pig implants, different parameters are deemed appropriate. An average SNR of 2.742 is estimated for the pig data set, while an average SNR of 1.273 is estimated for the human data set. The human data set therefore has a lower SNR which is typically associated with grainy images. Within the context of the pig implants, the smoothing of the X-ray images is achieved by applying a Gaussian kernel of size 5×5, with a standard deviation of $\sigma = 3$, to the image in question. In the case of the human implants, a Gaussian kernel of size 11×11, with a standard deviation of $\sigma = 3$, is employed. The aforementioned parameters were found to be optimal in removing a sufficient amount of noise while retaining the prominent features associated with a dental implant. Examples of smoothed actual X-ray images are presented in Figure 6.4.

A grayscale intensity transformation is implemented for the purpose of adjusting the dynamic range of the pixels in such a way that the dark pixels are significantly darkened while the bright pixels are only slightly darkened (see Figure 6.5). The aforementioned grayscale intensity transformation and spatial filtering techniques are implemented for the purpose of enhancing contrast and suppressing noise in the questioned image so as to render each actual X-ray image similar to the simulated dental implants.

The post-processed images are subsequently resized to a resolution of 540×540 pixels (see Figure 6.6). Finally, a suitable data transformation is implemented for the purpose of transforming the images into a tuple of four dimensional (4D) arrays with three spatial dimensions and one feature dimension, that is a volume V of dimensions $\{\text{width} \times \text{height} \times \text{depth}\}$, where $\{\text{width} \times \text{height}\}$ represents the spatial dimensions of the input image and depth the number of colour channels. The transformed test images therefore has a shape defined as $(540, 540, 1, 1)$ (see Figure 6.7).

6.3 Classification

The trained fully convolutional network (FCN) model, that is the *FCN-1* model discussed in Chapter 5, is presented with an actual X-ray image (see Figure 6.7) which contains only a single implant with the intention of extracting features for classification purposes. The aforementioned FCN-1 model therefore extracts features from the questioned dental implant for the purpose of predicting the connection type associated with the implant. The new set of features extracted from the questioned implant are subsequently com-

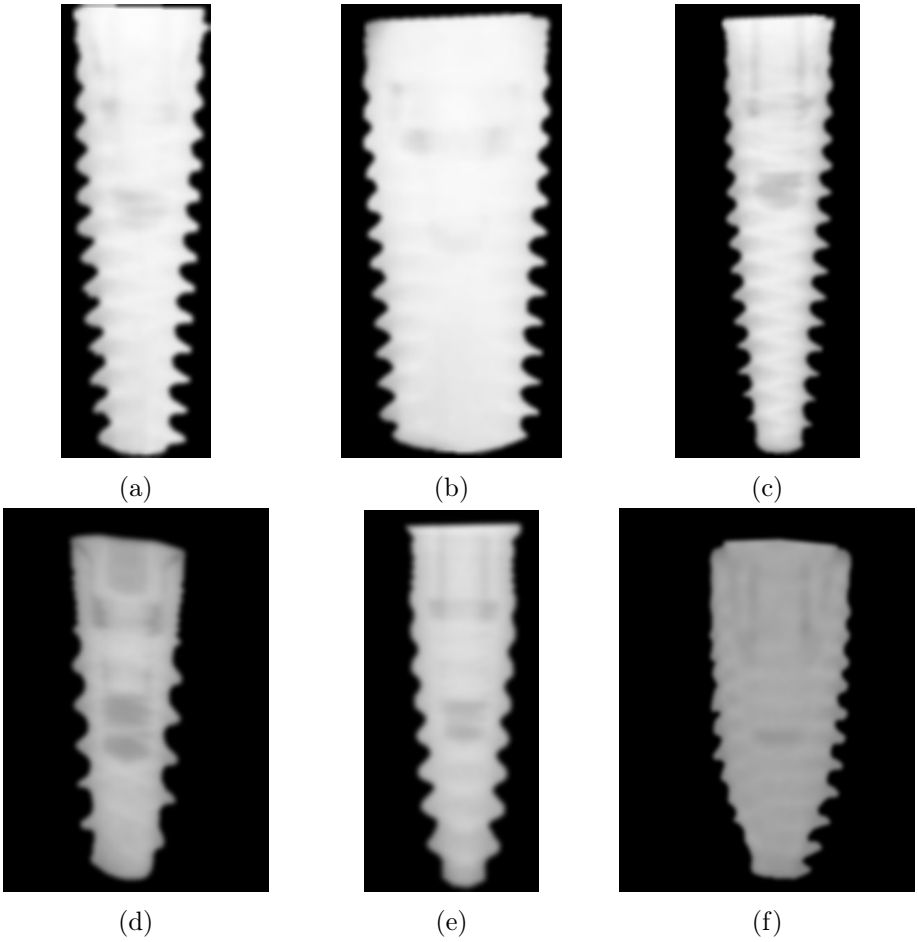


Figure 6.4: Noise reduction. Smoothed images after the application of a suitable Gaussian filter to the images depicted in Figure 6.3.

pared to the learned features in the artificially generated data. The model subsequently outputs probability values for predicting the implant category. The proposed dental implant classification protocol is conceptualised in Figure 6.8.

6.4 Concluding remarks

In this chapter, the ensemble of image processing techniques employed in the process of identifying the connection type associated with a specific dental

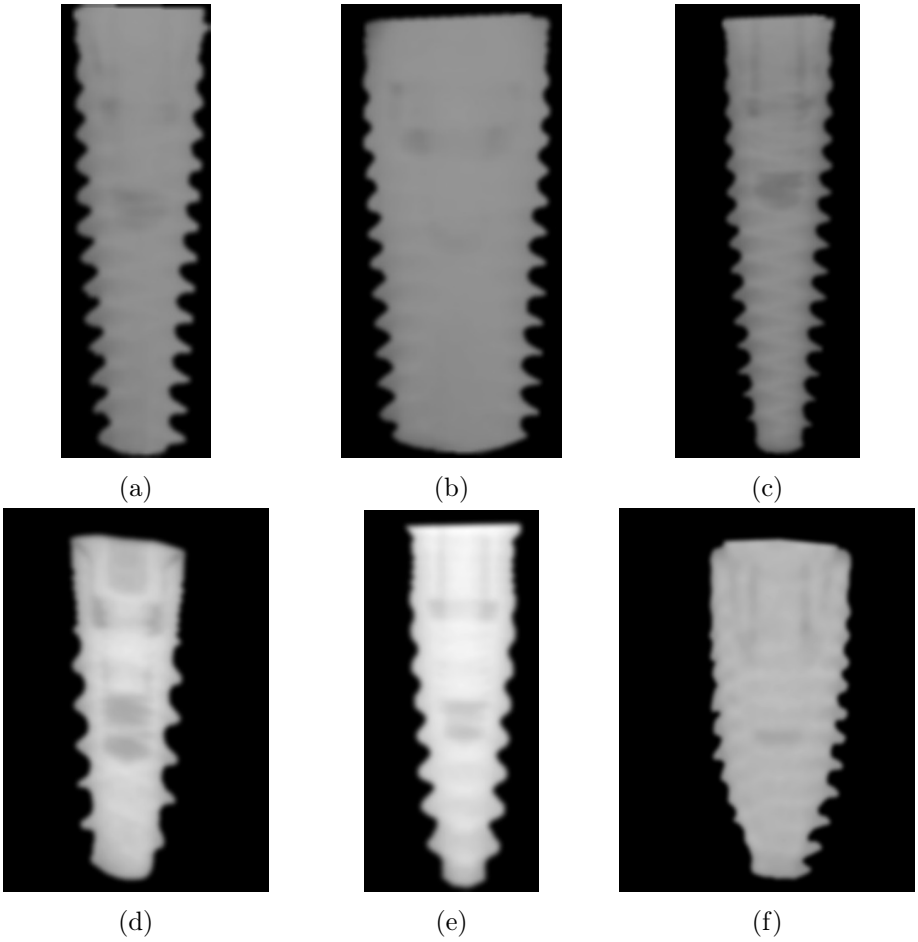


Figure 6.5: Contrast enhancement. Examples of actual X-ray images after a suitable grayscale intensity transformation was applied to the images depicted in Figure 6.4. (**Top**) Implants within the context of the pig data set. (**Bottom**) Implants within the context of the human data set.

implant within a questioned actual X-ray image were discussed. These image processing techniques were implemented in order to render each actual X-ray image similar to the simulated dental implants. The protocol for identifying the connection type from a questioned dental implant was presented.

In the next chapter the data considered in this research within the context of the *actual* X-ray images and the artificially generated *simulated* X-ray images, as well as the data partitioning protocol, are discussed in detail. This

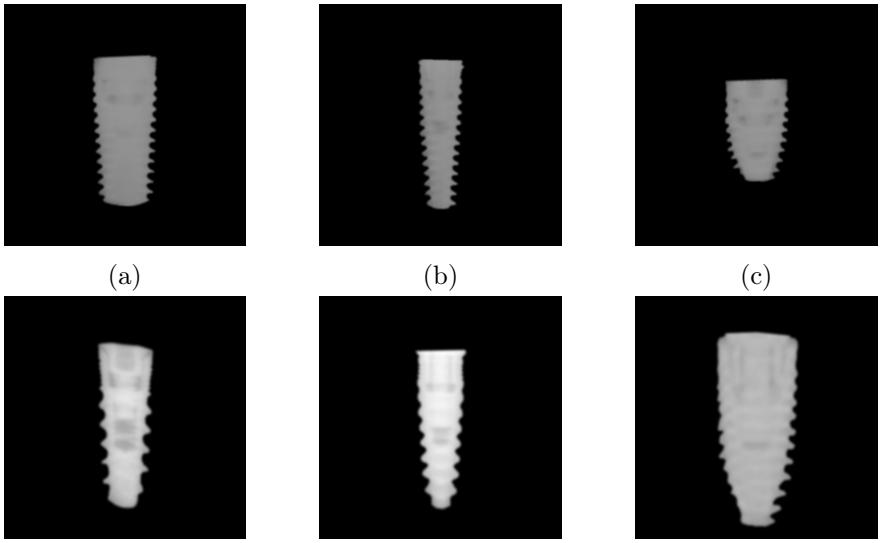


Figure 6.6: Examples of resized actual X-ray images. (**Top**) Implants within the context of the pig data set. (**Bottom**) Implants within the context of the human data set.

is followed by an outline of the experimental protocol. Exhaustive experiments are also conducted in order to evaluate the proposed artificial data generation strategy, automated region of interest (ROI)-detection strategy, as well as the proposed semi-automated and fully automated systems for the purpose of classifying a questioned dental implant within an actual X-ray image.

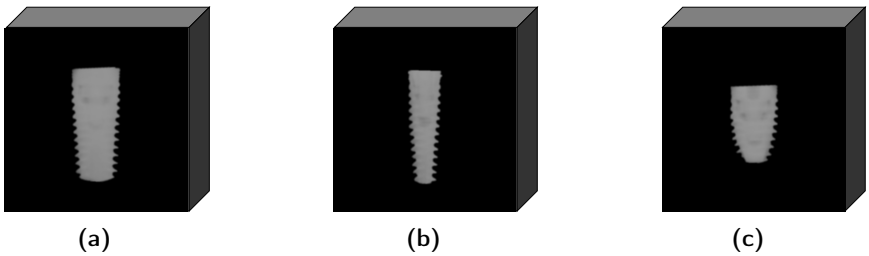


Figure 6.7: Examples of actual X-ray images of pig implants transformed into a tuple of 4D arrays.

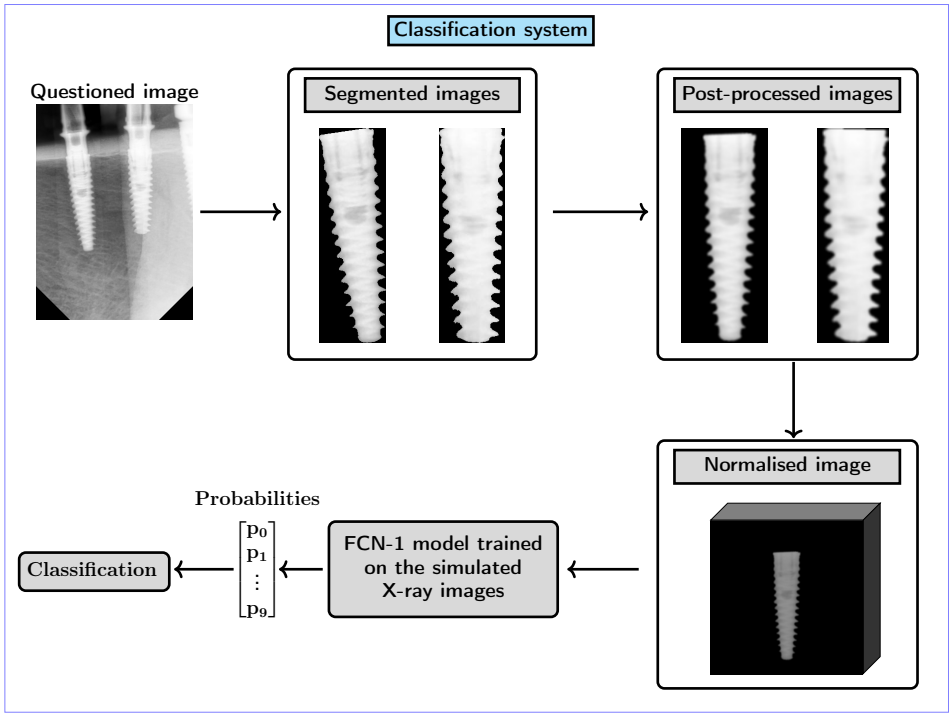


Figure 6.8: Conceptualisation of the proposed dental implant classification protocol. The input image is first segmented, after which the segmented implant is normalised and processed. Features are extracted from the input image through the proposed FCN-1 model and compared to the learned features in the simulated data.

Chapter 7

Experiments

7.1 Introduction

In this chapter exhaustive experiments are conducted in order to gauge the proficiency of the proposed artificial data generation strategy, automated region of interest (ROI) detection algorithm, as well as the respective proficiencies of the semi-automated and fully automated systems developed in this study within the context of classifying a questioned dental implant within an actual X-ray image. The aforementioned experiments are conducted on two data sets that contain the *simulated* and *actual* X-ray images. These data sets are described in detail in Section 7.2. The experimental protocol that is followed for each of the individual experiments is outlined in Section 7.3. The experimental results are presented and quantitatively analysed in Section 7.4. An overview of the software developed and hardware utilised in this study is presented in Section 7.5. Finally, an analysis of the results acquired for the experiments conducted in this study is presented in Section 7.6.

7.2 Data

In this study, experiments are conducted on (1) an artificially generated data set and (2) an actual X-ray data set. The aforementioned independent data sets are discussed in the following subsections.

7.2.1 Artificially generated X-ray images

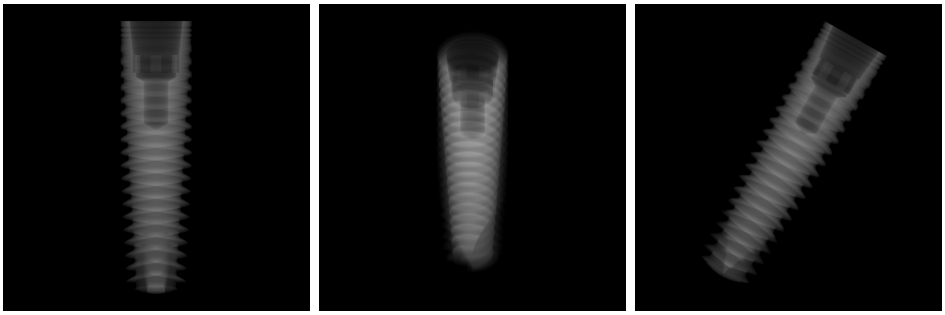
The data set that contains simulated X-ray images of dental implants is generated from triangulated surface models, which constitute standard triangle

language (STL) files, engineered by the Make It Simple (MIS) manufacturing company. The triangulated 3D surface coordinates of a specific dental implant are used to construct a 3D volumetric representation of the model in question. Subsequently, 2D parallel projections of the volumetric representation are calculated from a number of different angles. Each projection is obtained by calculating a number of parallel-ray sums of the 3D volumetric representation. Each projection profile constitutes a simulated X-ray image.

During the X-ray simulation process, each 3D volumetric representation of an implant undergoes axial, out-of-plane and in-plane rotations through a number of different angle combinations, after which a projection is generated. A total of 403 200 simulated X-ray images are generated in this way. The background of the artificially generated X-ray images is set to black. During training, the simulated X-ray images are augmented by in-plane rotations of maximally 60°.

A total of 322 560 simulated X-ray images (80%) are used for training purposes, while 80 640 simulated X-ray images (20%) are used for validation purposes. A number of training samples from the artificially generated data set are presented in Figure 7.1.

The simulated X-ray data set is employed for training and validating the proposed fully convolutional network (FCN) model, that is the *FCN-1* model, for the purpose of classifying the connection type associated with a specific dental implant within an actual X-ray image.



(a) (b) (c)

Figure 7.1: Examples of the artificially generated X-ray images employed for training the proposed FCN-1 model.

7.2.2 Actual X-ray images

Within the context of the actual X-ray images, a total of 483 labelled and unlabelled images, which contain implants inserted into either pig or human jaws, are considered (see Figure 7.2).

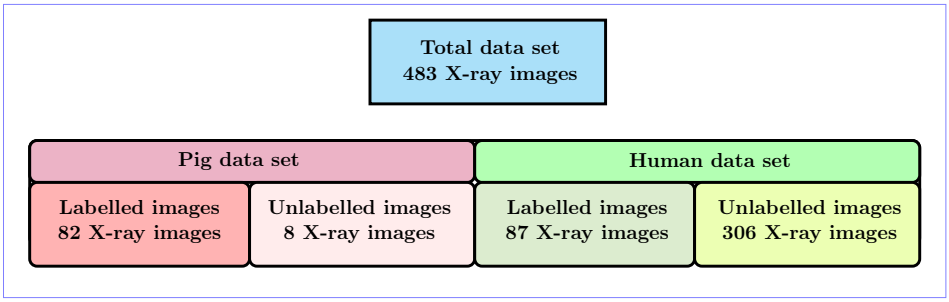
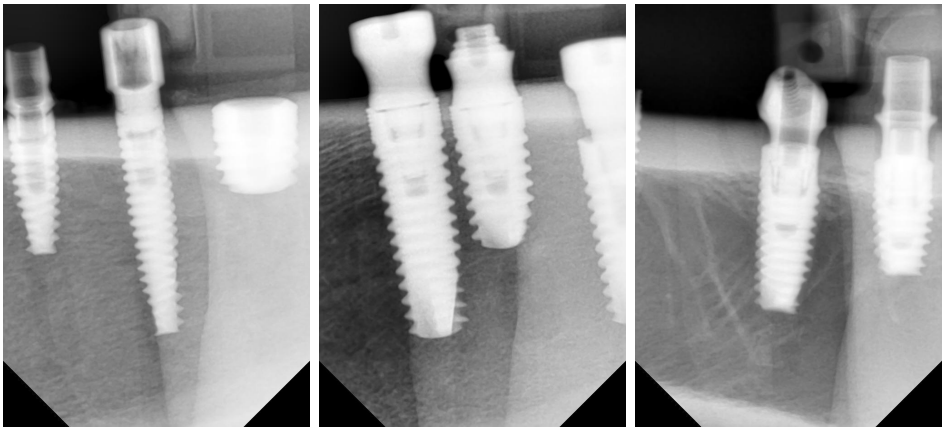


Figure 7.2: Total data set composition within the context of the actual X-ray images.

The database of X-ray images involving implants inserted into human jaws pertains to anonymous dental patients and was made available for this study by Medical Care NV. The database of X-ray images involving pig jaws was generated explicitly for this research by inserting the relevant dental implants into detached pig jaws obtained from butchers, after which the inserted implants were X-rayed with a similar device as the one used within the context of dental patients. Figures 7.3 and 7.4 depict examples of actual X-ray images within the context of the implants inserted into *pig jaws* and *human jaws*, respectively. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and national research committees and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study. All applicable international, national, and institutional guidelines for the care and use of animals were followed.

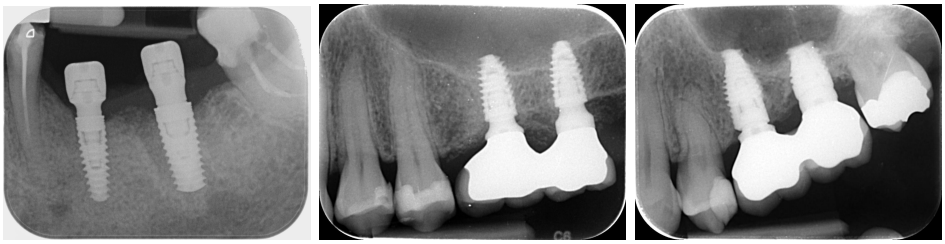
Within this context, *labelled* dental implant images refer to those X-ray images that are associated with a specific dental implant model and brand, while *unlabelled* X-ray images constitute dental implants with unknown models and brands. The labelled dental implant images comprise of four different brands (Anthogyr, Astra, MIS and Nobel Biocare).

The images are captured in grayscale format. Each of these images is resized to 512×512 pixels and saved in JPEG format. The data set (both



(a) (b) (c)

Figure 7.3: Examples of actual X-ray images within the context of the implants inserted into *pig jaws*.



(a) (b) (c)

Figure 7.4: Examples of actual X-ray images within the context of the implants inserted into *human jaws*.

the labelled and unlabelled X-ray images) are annotated for the purpose of training the proposed fully convolutional network (FCN) model, that is the FCN-2 model, to facilitate the automatic *detection* (segmentation) of the dental implants (discussed in Chapter 4). Within the context of semantic segmentation, the annotated data set contains the binary masks which separate the dental implants from the background in a pixel-wise fashion. The constructed data set therefore consists of the X-ray images *and* the corresponding set of masks that represent the ground truth associated with the segmentation protocol.

7.3 Experimental protocol

In this study, three main independent experiments are conducted for the purpose of investigating the proficiency of the proposed systems. The experimental protocol is categorised as follows:

(1) **Experiment 1.** This experiment investigates the proficiency of the proposed strategy of artificially generating simulated X-ray images of dental implants. The aforementioned data generation strategy is based on the calculation of 2D projections (from a number of different angles) of a 3D volumetric representation of a dental implant. Each projection is obtained by calculating a number of parallel ray-sums of the 3D volumetric representation.

(2) **Experiment 2.** This experiment investigates the proficiency of the proposed automated detection of regions of interest (ROIs) which contain the dental implants in actual X-ray images.

(3) **Experiment 3.** This experiment investigates the proficiency of the proposed network for the purpose of classifying a questioned dental implant within an actual X-ray image. This experiment is further dichotomised into two sub-experiments:

   (a) **Experiment 3A.** In this sub-experiment, a *semi-automated* segmentation strategy based on image processing techniques is implemented for the purpose of segmenting an actual X-ray image. This is followed by the classification/recognition of the connection type associated with a questioned dental implant.

   (b) **Experiment 3B.** In this sub-experiment, a *fully automated* segmentation strategy, which is based on deep learning, is employed for the detection of suitable ROIs that contain the dental implants in an actual X-ray image. This is followed by the classification/recognition of the connection type associated with a questioned dental implant.

In order to ensure that the results reported in this study represent a comprehensive and unbiased estimation of the proposed systems' performance, the experimental protocol incorporates a k-fold cross-validation protocol within the context of the proposed deep learning-based algorithms.

7.3.1 Experiment 1: Artificially generated X-ray images

In this section an experimental protocol is proposed to gauge the proficiency of the proposed fully convolutional network (FCN) model, that is the *FCN-1* model, which is trained and validated on artificially generated X-ray images for the purpose of assigning a questioned dental implant within an actual X-ray image to one of *nine* different connection types investigated in this study. In order to investigate the feasibility of the proposed FCN-1 model for the aforementioned purpose, the experiments are conducted on the simulated X-ray images. The experiments conducted in this section are used to gauge the proficiency of the proposed artificial data generation strategy.

The proposed FCN-1 model is trained from scratch. The filter weights are randomly initialised and optimally adjusted during training. The learned filters of the proposed FCN-1 model are visualised in Figure 7.5 where the corresponding filters for each network layer are depicted. Every filter is different so that different aspects of the image is enhanced in the output. When comparing some of the filters visually, interesting features can be noted. Although the input data is grayscale, the colours are used for visualisation purposes to depict how the different convolutional filters affect the output.

From Figure 7.5, it can be observed that the size of the first ten convolutional layers is $3\times3\times1$ pixels and that the final convolutional layer which is fed to the softmax layer for classification purposes is a one dimensional feature vector. Within the context of deep learning-based classification for feature extraction, the proposed FCN-1 model which constitutes a 3D CNN takes a batch of consecutively simulated X-ray images as input and generates different feature maps which are then transformed into a final one dimensional feature vector.

The simulated X-ray data is partitioned into training and validation sets, where 80% of the data is assigned to the *training* set, while 20% of the data is assigned to the *validation* set. A k-fold cross-validation protocol is employed during training for data splitting, which implies that the training set is divided into k different folds. One fold is held out as the validation set. The model is trained on the remaining $k-1$ folds and then applied to the validation set, after which the predictive performance is recorded. This process is repeated k times so that each fold has been used as a validation set once. This process ensures that different samples are used for the purpose of model training and validation, thereby avoiding potential overfitting, whilst still ensuring that all the samples represented in the data set are considered for evaluation. The recorded predictive performances are then averaged.

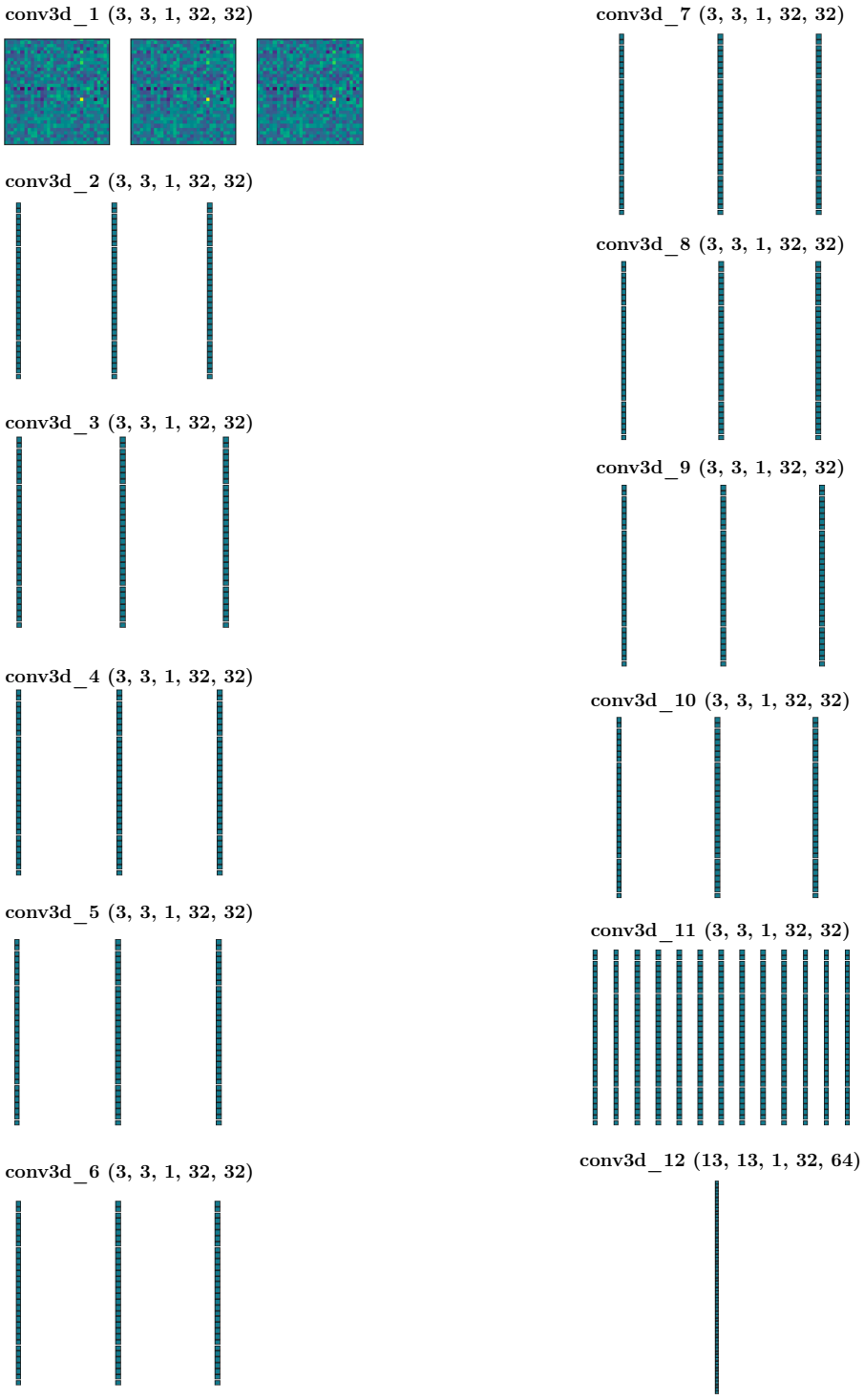


Figure 7.5: Visualisation of the filters of the twelve convolutional layers of the proposed FCN-1 model. The images reflect the filters at different layers.

7.3.2 Experiment 2: Automated ROI detection

In this experiment, the data set of X-ray images which consists of both labelled and unlabelled images is first partitioned into two independent sets. In *set one*, the *labelled* dental implants in pig jaws are employed for test purposes. In *set two*, the *labelled* dental implants in human jaws are employed for test purposes. For set one, the X-ray images in human jaws and the *unlabelled* X-ray images in pig jaws are used for training and validation purposes respectively. For set two, the X-ray images in pig jaws and the *unlabelled* X-ray images in human jaws are used for training and validation purposes respectively. The aforementioned data partitioning protocol is depicted in Figure 7.6.

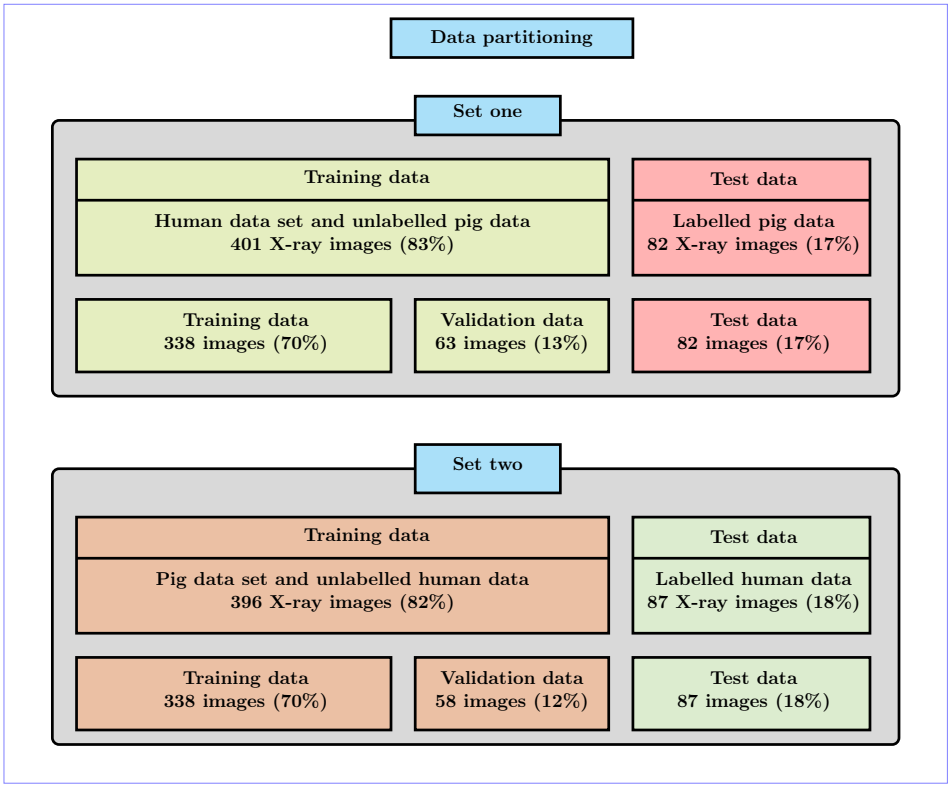


Figure 7.6: Conceptualisation of the proposed data partitioning protocol within the context of the actual X-ray images.

In order to conduct a robust analysis, a 5-fold cross-validation procedure is carried out. During network training, at the end of each epoch, the

validation sets are used to gauge the proficiency of the model.

In this section an experimental protocol is proposed to gauge the proficiency of the fully convolutional network (FCN-2) model proposed for the *automatic* detection of suitable ROIs which contain the dental implants in an actual X-ray image. The binary masks of the dental implants acquired through the proposed semi-automated segmentation system serve as the ground truth for the purpose of training the proposed FCN-2 model.

The output of the FCN-2 model is compared to the *manually* selected ROIs for the purpose of evaluating the proposed segmentation protocol. The pixels that are correctly classified as part of the implant are referred to as true positives, while those correctly classified as part of the background are referred to as true negatives. False positives constitute those pixels that are erroneously classified as part of the implant, while the pixels that are erroneously classified as part of the background are referred to as false negatives.

7.3.3 Experiment 3: Dental implant classification/recognition

In this section, experiments are conducted in order to investigate the proficiency of the proposed systems for the purpose of classifying a questioned dental implant within an actual X-ray image. In this experiment, the dental implants are extracted from actual X-ray images and presented to the trained model for evaluation purposes.

A questioned actual X-ray image is presented to the FCN-1 model trained and validated on artificially generated X-ray images in order to predict the specific connection type associated with the questioned dental implant. During the classification phase the trained FCN-1 model considers an actual X-ray image which contains only a single dental implant as input. The questioned dental implant is loaded and reshaped so as to match the input shape of the proposed FCN-1 model. Image processing techniques are implemented on the questioned implant image, after which the processed image is converted to an array and the dimensions of said array are expanded so that it represents a batch containing a single sample.

The redefined sample is fed to the proposed FCN-1 model which subsequently extracts features from the questioned dental implant for the purpose of predicting the connection type associated with the implant. The new set of features extracted from the questioned implant are subsequently compared to the learned features in the artificially generated data. The model subsequently outputs probability values for predicting the implant category.

Experiment 3A: Semi-automated dental implant recognition in actual X-ray images

In this sub-experiment the suitable ROIs that contain the dental implants in the actual X-ray images, are *manually* specified and segmented through the proposed semi-automated system which is based on image processing techniques. The segmented dental implants are presented to the FCN-1 model for the purpose of classifying each questioned dental implant within each actual X-ray image.

Experiment 3B: Fully automated dental implant recognition in actual X-ray images

In this sub-experiment the suitable ROIs which contain the dental implants in an actual X-ray image are *automatically* detected through the proposed FCN-2 model. This is followed by classification of the connection type associated with the questioned dental implant.

7.4 Results

In this section, the performance of the proposed systems is reported and a comprehensive analysis of the results pertaining to the experimental protocols outlined in the previous section, is presented. The relevant statistical performance measures employed in this study for the purpose of quantifying the proficiency of the proposed systems are listed and defined in Table 7.1.

Performance measure	Definition
Precision (PRE)	TP/(TP+FP)
Recall (REC)	TP/(TP+FN)
Accuracy (ACC)	(TP+TN)/(TP+FN+FP+TN)
F_1 score	2 * PRE * REC/(PRE+REC)

Table 7.1: The statistical performance measures employed in this study. The number of true positives, false positives, true negatives, and false negatives are denoted by TP, FP, TN, and FN, respectively.

7.4.1 Experiment 1: Training results for simulated X-ray images

The proposed FCN-1 model is trained on artificially generated X-ray images. During training the simulated data set is augmented by in-plane rotations of maximally 60°. The training algorithm is run for a maximum of 1000 epochs and validated across a 5-fold cross-validation protocol. The accuracy of the network is measured after each epoch, by employing the independent validation set.

During training, an average accuracy of 98.64% is achieved (see Figure 7.7 (a)), while an average loss of 3.56% is achieved (see Figure 7.8 (a)). An average accuracy of 96.56% and an average loss of 5.42% are achieved during validation across a 5-fold cross-validation protocol. The accuracy and loss achieved during validation are depicted in Figure 7.7 (b) and Figure 7.8 (b), respectively.

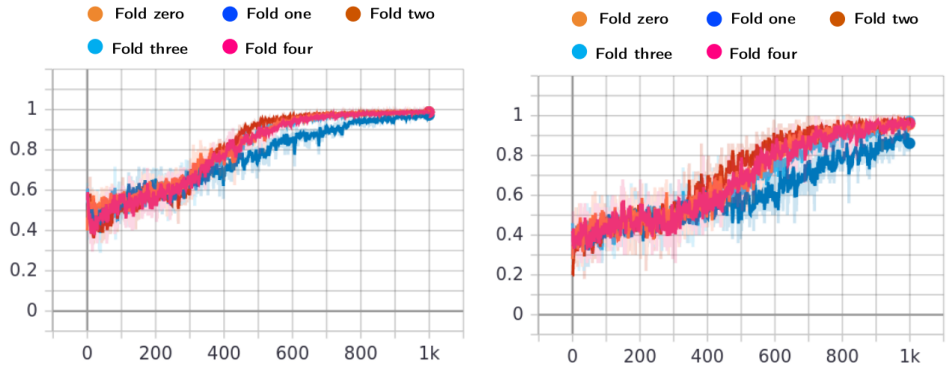


Figure 7.7: The *accuracy* achieved during the (a) training and (b) validation phases of the proposed network, when in-plane rotations of maximally 60° are employed and the model is trained for 1000 epochs across a 5-fold cross-validation protocol.

From Figure 7.7 it is clear that the training accuracy increases as the number of epochs increases, while the same trend is observed for the validation accuracy, especially after 500 epochs. The training and validation accuracies exceed 80% after 600 epochs. It is furthermore evident from Figure 7.8 that the training and validation loss decrease rapidly as the number of epochs increases. This decrease is more pronounced in the case of the training loss. It can therefore be concluded that the model converges as expected during training and performs well and robustly on the validation

data.

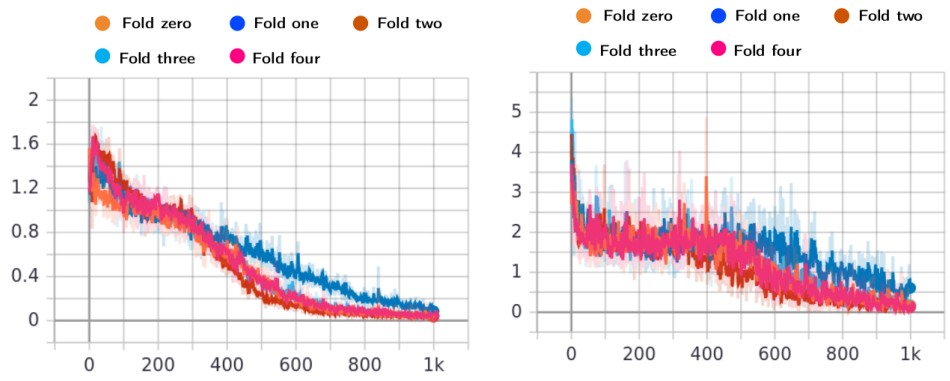


Figure 7.8: The *loss* achieved during the (a) training and (b) validation phases of the proposed network across a 5-fold cross-validation protocol.

7.4.2 Experiment 2: Results for dental implant detection in actual X-ray images

In order to conduct a robust analysis within the context of the proposed automatic ROI detection protocol, a 5-fold cross-validation procedure is carried out. During network training, at the end of each epoch, the validation sets are used to gauge the proficiency of the model. For sets one and two respectively, accuracies of 97.84% and 97.21% are achieved by analysing the segmentation performance in terms of pixel-wise accuracy during training. In order to evaluate the proficiency of the proposed ROI detection framework during the test phase, the evaluation is conducted on the predicted segmentation maps before post-processing is carried out. The precision, recall, accuracy, and F_1 score are employed as performance measures for both sets. The results achieved during testing are presented in Table 7.2.

Performance measure	Set one	Set two
PRE	74.38%	68.31%
REC	90.98%	78.64%
ACC	90.43%	94.06%
F_1 score	80.73%	84.48%

Table 7.2: Results for the proposed automated ROI detection protocol. The results constitute *averages* from a 5-fold cross-validation protocol.

Note that the precision metric is significantly lower than the accuracy and recall metrics. The proposed model therefore incorrectly classifies instances as positive on a number of occasions.

Selected results illustrating the proficiency of the proposed FCN-2 model for the purpose of segmenting dental implants into foreground and background regions are presented in Figure 7.9. The true positive, true negative, false positive and false negative pixels are depicted in white, black, green and pink respectively.

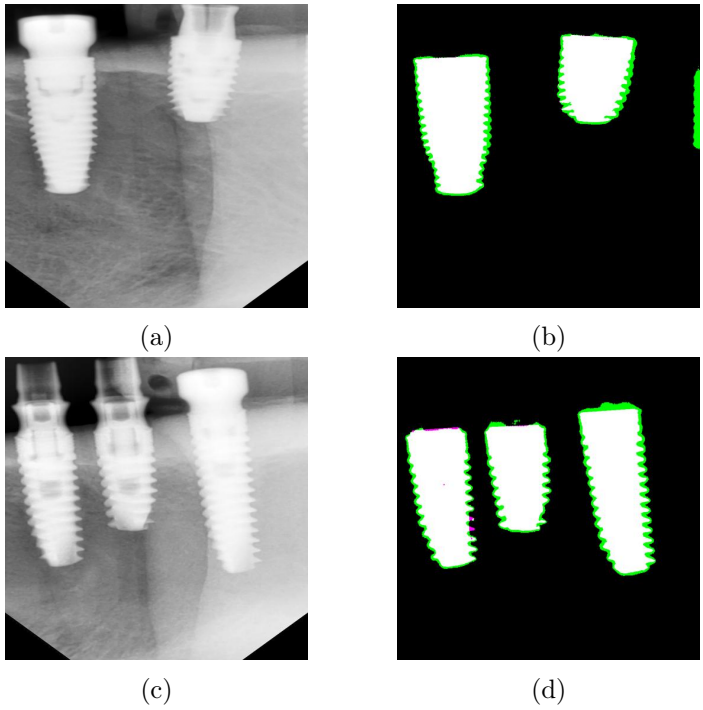


Figure 7.9: Qualitative depiction of the proficiency of the proposed automated ROI detection protocol. **(Left)** Examples of actual X-ray images. **(Right)** Comparison of manually selected and automatically detected ROIs. The true positive, true negative, false positive and false negative pixels are depicted in white, black, green and pink respectively.

7.4.3 Experiment 3: Results for dental implant recognition in actual X-ray images

In order to provide a more detailed perspective into the the proposed dental implant recognition protocol, confusion matrices are computed for the nine connection types across the five folds. These confusion matrices provide in-depth insight into the classification of each connection type within the actual X-ray images.

Experiment 3A: Results for the semi-automated dental implant recognition system

In this subsection the results for the proposed *semi-automated* dental implant classification system are presented. This system employs a protocol in which the ROI is *manually* selected through the proposed semi-automated framework. In Figures 7.10 and 7.11, the confusion matrices for the proposed semi-automated dental implant classification system within the context of implants inserted into pig jaws and human jaws are respectively depicted.

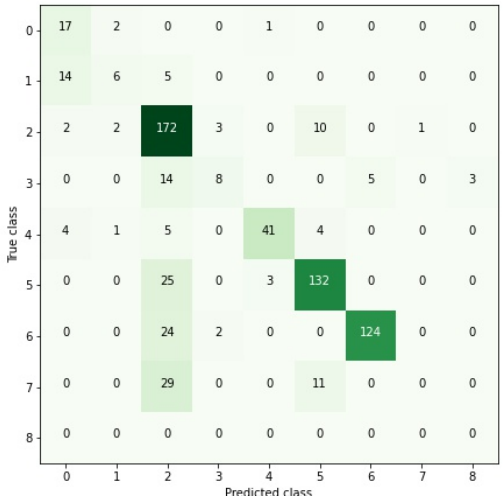


Figure 7.10: The confusion matrix of the aggregate across 5-fold cross-validation for the proposed *semi-automated* dental implant classification system when implants inserted into *pig jaws* are considered. The predicted and true classes represent the nine connection types described in Chapter 5 and depicted in Table 5.1.

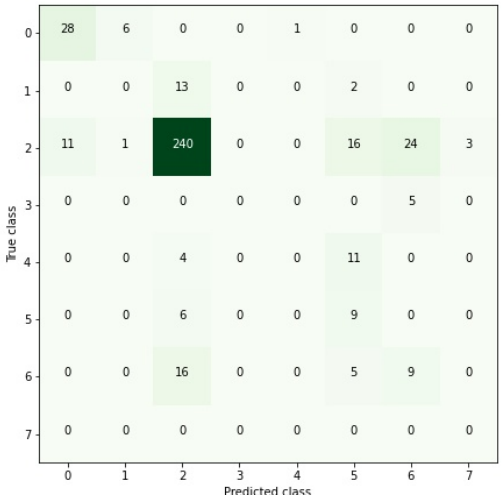


Figure 7.11: The confusion matrix of the aggregate across 5-fold cross-validation for the proposed *semi-automated* dental implant classification system when implants inserted into *human jaws* are considered. The predicted and true classes represent the nine connection types described in Chapter 5 and depicted in Table 5.1.

From Figure 7.10 which constitutes a confusion matrix across 5-fold cross-validation within the context of the dental implants inserted into *pig jaws*, it can be observed that classes *zero* and *one* which represent dental implants with a conical narrow connection type are underrepresented. From the aforementioned figure it can also be observed that there are no test samples for class *eight* which constitutes external hexagonal wide platform dental implants.

From Figure 7.11 which constitutes a confusion matrix across 5-fold cross-validation within the context of the dental implants inserted into *human jaws*, it can be observed that a high number of classes are underrepresented. From the aforementioned figure it can also be observed that a high number of test samples constitute class *two* which represents the conical standard platform connection type.

In order to further evaluate the proficiency of the proposed system, the precision, recall, F_1 score and accuracy are estimated from the confusion matrices. Recall that the data employed for dental implant recognition is imbalanced and that certain classes are underrepresented. In order to address the aforementioned data imbalance, *weighted average* metrics within

the context of the precision, recall and F_1 score are estimated from the confusion matrices.

In Table 7.3, results for the proposed semi-automated system are summarised for dental implants inserted into *pig jaws* and *human jaws*, respectively. The precision, recall, accuracy and F_1 score are employed as performance evaluation measures.

Performance measure	Pig set	Human set
PRE	73.04%	70.52%
REC	74.63%	69.76%
F_1 score	72.23%	69.70%
ACC	74.63%	69.76%

Table 7.3: Results for the proposed *semi-automated* dental implant recognition system within the context of implants inserted into *pig jaws* and *human jaws*. The results constitute *weighted averages* across the five folds.

Experiment 3B: Results for the fully automated dental implant recognition system

In this subsection the results for the proposed *fully automated* dental implant classification system are presented. The aforementioned system employs a protocol in which the ROI is *automatically* detected through deep learning.

Figures 7.12 and 7.13 depict the confusion matrices for the proposed fully automated dental implant classification system when implants inserted into pig jaws and human jaws are respectively considered.

From the confusion matrices, *weighted average* metrics within the context of the precision, recall and F_1 score are estimated for implants inserted into *pig jaws* and *human jaws*, respectively. The results are summarised in Table 7.4.

7.5 Software and hardware employed

The proposed framework for the purpose of artificially generating X-ray images is implemented in MATLAB^{TM}. The aforementioned framework employs a voxelisation algorithm [39] for the purpose of transforming the triangulated 3D surface coordinates of a specific dental implant into a 3D volumetric representation.

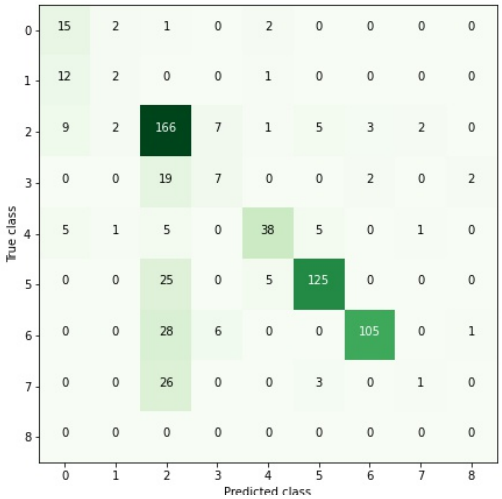


Figure 7.12: The confusion matrix of the aggregate across 5-fold cross-validation for the proposed *fully automated* dental implant classification system when implants inserted into *pig jaws* are considered. The predicted and true classes represent the nine connection types described in Chapter 5 and depicted in Table 5.1.

Performance measure	Pig set	Human set
PRE	73.12%	70.55%
REC	71.72%	68.67%
F_1 score	70.75%	67.60%
ACC	71.72%	68.67%

Table 7.4: Results for the proposed *fully automated* dental implant recognition system within the context of implants inserted into *pig jaws* and *human jaws*. The results constitute *weighted averages* across the five folds.

The neural network-based experimental protocols for the purpose of image segmentation and classification are implemented in TensorFlow [53], a premier open source library for handling complex back-end machine learning operations. TensorFlow offers a simple and powerful Python application program interface (API) to manage the computational graph implemented in `C++` and is an excellent tool for model diagnostics and evaluating performance giving it great reliability and flexibility. Keras [54] is used in conjunction with TensorFlow as a front-end API that is easy to use with

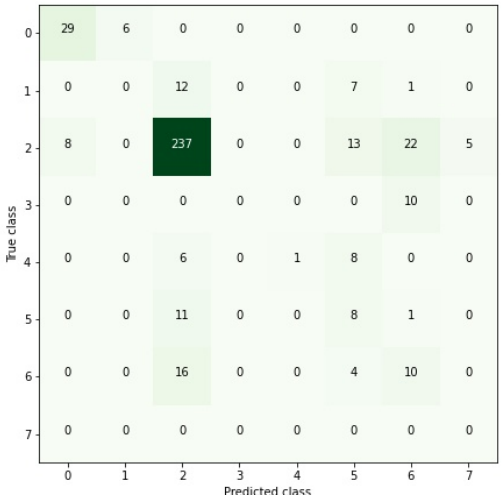


Figure 7.13: The confusion matrix of the aggregate across 5-fold cross-validation for the proposed *fully automated* dental implant classification system when implants inserted into *human jaws* are considered. The predicted and true classes represent the nine connection types described in Chapter 5 and depicted in Table 5.1.

TensorFlow as the powerful back-end. This combination of Keras and TensorFlow makes Python one of the most popular languages used when dealing specifically with neural networks. In addition to the aforementioned libraries employed within the context of neural network-based experimental protocols, the DeepVoxNet [45] module is employed for the efficient processing of the artificially generated X-ray images. DeepVoxNet is a deep learning processing framework for Keras developed in the Medical Imaging Research Center (MIRC) at KU Leuven for the efficient processing and analysis of 3D medical images.

The data preprocessing and post-processing techniques are implemented in *NumPy* which is a scientific computation library, as well as *Scikit* and *OpenCV* which are open source computer vision libraries within the Python environment.

The experimental protocols for the proposed deep learning-based models are conducted on the Nvidia Tesla V100 processor through the Kraken server. Model inspection and evaluation are performed in the Google collaboratory environment which offers free GPU usage for interactive sessions in a Jupyter Notebook-like environment.

7.6 Concluding remarks

Within the context of simulated X-ray images, a high accuracy of 96.56% and an average loss of 5.42% are achieved during validation across a 5-fold cross-validation protocol. This demonstrates that the proposed FCN-1 model (for automated connection type classification) effectively learns the prominent features associated with each artificially generated dental implant. Within this context, the simulated X-ray images are partitioned into a training and validation set. The proposed network uses the training data for learning prominent features, while the network is tested against the validation data after every epoch in order to prevent overfitting. Data augmentation is implemented in order to ensure that the model learns varied samples of the data so as to increase its capability to generalise on unseen data.

The performance of the proposed FCN-2 model (for automated dental implant segmentation) is encouraging. The proposed system is able to classify the pixels associated with the dental implants (foreground) and those associated with the background with accuracies of 90.43% and 94.06% within the context of sets one and two respectively. The proposed model is trained to perform semantic segmentation. Morphological post-processing techniques are applied to the output binary masks in order to remove noise and components *not* associated with the dental implants.

Within the context of implants inserted into *pig jaws*, classification accuracies of 74.63% and 71.72% are achieved for the proposed semi-automated and fully automated dental implant recognition systems respectively. Within the context of implants inserted into *human jaws*, classification accuracies of 69.76% and 68.67% are achieved for the semi-automated and fully automated dental implant recognition systems respectively. Recall that within the context of the semi-automated system the dental implants are accurately segmented from the actual X-ray images. This system therefore also serves as a benchmark in gauging the performance of the fully automated segmentation protocol. The classification accuracies within the context of the fully automated dental implant recognition system are therefore slightly lower compared to that of the semi-automated dental implant recognition system, which is expected. The proficiency of the proposed semi-automated and fully automated end-to-end systems within the context of dental implant recognition in *actual* X-ray images is significantly lower than the validation accuracy reported for the proposed FCN-1 system, which was trained and validated on the *artificially generated* X-ray images. This observation is also not surprising.

The proficiency of the proposed dental implant recognition systems is slightly lower for human implants than is the case for pig implants, which may be attributed to the presence of more significant noise levels in the human implants (as was discussed in more detail in Chapter 4). The signal-to-noise ratio (SNR) can be employed to measure the noise (e.g. random quantum mottle) in the actual X-ray images. An average SNR of 2.742 is estimated for the pig data set, while an average SNR of 1.273 is estimated for the human data set. The human data set has a lower SNR which is typically associated with grainy images.

The aforementioned results clearly demonstrate that the proposed protocol is proficient at generating artificial (simulated) X-ray images that closely resemble actual X-ray images. The ensemble of algorithms proposed in this study provides valuable insight into artificial data generation and automatic implant segmentation within the context of dental implant recognition.

A number of studies [22, 25–28, 55] have applied machine learning and especially deep learning algorithms for the purpose of classifying dental implants and achieved accuracies of 63% to 96%. It is important to reiterate that these systems considered different data sets than those considered in this study. In the aforementioned systems the classification of a dental implant is furthermore based on the *type* (brand or model) of the implant. The protocol proposed in *this* study delved deeper by investigating the classification of dental implant *connection types*. The dental implant connection interface is a vital feature to consider when choosing an abutment replacement model. The compatibility of the dental implant connection interface varies depending on the model. A number of implant models are incompatible with those of other brands [56]. It is therefore important to accurately classify the dental implant connection type. Since the system developed in *this* study is conceptually different from existing state-of-the-art systems, it is conceivable that the system proposed here will be *complimentary* to existing systems. When the proposed system is combined with any other existing state-of-the-art system, the combined system should therefore be superior to the existing system.

Chapter 8

Conclusion and future work

8.1 Conclusion

In this dissertation, novel deep learning-based dental implant recognition systems using synthetic X-ray images were proposed. Firstly, a novel algorithm for generating artificial training samples from triangulated three-dimensional (3D) surface models within the context of dental implant recognition was proposed. The proposed algorithm is based on the calculation of two-dimensional (2D) projections (from a number of different angles) of 3D volumetric representations of computer-aided design (CAD) surface models. The 2D projections were obtained by calculating a number of parallel ray-sums of each 3D volumetric representation. Each acquired 2D projection constitutes an artificially generated X-ray image. A fully convolutional network (FCN), that is the so-called *FCN-1* model, was subsequently trained on the artificially generated X-ray images for the purpose of automatically identifying the connection type associated with a specific dental implant in an actual X-ray image. Semi-automated and fully automated systems were proposed for the purpose of segmenting questioned dental implants from the background in actual X-ray images. Within the context of the semi-automated system suitable regions of interest (ROIs), that contain the dental implants, were manually specified (selected). However, as part of the fully automated system, suitable ROIs were automatically detected, which is facilitated by a deep learning-based model (the so-called *FCN-2* model) and morphological post-processing. Normalisation and preprocessing techniques were applied to the segmented dental implants within a questioned actual X-ray image. The normalised dental implants were presented to the trained *FCN-1* model for classification purposes.

In this dissertation, exhaustive experiments were conducted for the purpose of investigating the proficiency of the proposed systems, that is: (i) training and validation using artificially generated (simulated) X-ray images of dental implants, (ii) segmentation of dental implants within an actual X-ray image and (iii) dental implant classification/recognition.

The aforementioned experiments were conducted in order to gauge the proficiency of the proposed artificial data generation strategy, automated ROI detection algorithm, as well as the respective proficiencies of the semi-automated and fully automated systems developed in this study within the context of classifying a questioned dental implant within an actual X-ray image. The experiments were conducted on two data sets that contain the *simulated* and *actual* X-ray images. Within the context of the proposed artificial data generation protocol, the performance of the aforementioned system was demonstrated to be proficient at generating artificial (simulated) X-ray images that closely resemble actual X-ray images.

The ensemble of algorithms proposed in this dissertation provides valuable insight into artificial data generation in terms of biomedical imaging and automatic implant segmentation within the context of dental implant recognition.

8.2 Future work

Although the research conducted in this study provides valuable insight into numerous aspects relating to deep learning-based dental implant recognition using artificially generated training samples, the following alternative avenues can also be pursued and may represent interesting future work:

(1) A more in-depth investigation may be conducted into a model that is also capable of distinguishing between implants with the same external shape, but with different internal connection types. This may be the case in exceptional scenarios within the context of dental implants from Nobel Replace. Once it is established that the predicted implant type is associated with more than one connection type, the ROI that only contains the connection may be submitted to a different model that only differentiates between the connection types in question.

(2) The proposed strategy of generating 2D projections (from a number of angles) of 3D volumetric representations of CAD surface models is also applicable to a wide range of *other* objects. The CAD models for a

variety of objects such as vehicles, aircraft, and animals are either readily available or relatively easy to create within a short time-period.

Potential applications for this research include areas such as:

(i) vehicle detection and classification in traffic scenes,

(ii) the identification of aircraft, as well as

(iii) the categorisation of animals from aerial cameras.

(3) Within the context of simulated X-ray image generation, an investigation into a more realistic simulated X-ray acquisition process may be conducted. This can be achieved by also considering the physical attenuation process of the X-rays as they propagate through the material.

(4) Transfer learning may be employed by fine-tuning the proposed *FCN-1* model, which is trained and validated on the artificially generated (simulated) X-ray images, by *also* utilising a number of *actual* X-ray images for training and validation purposes. This technique should improve the proficiency of assigning a questioned dental implant within an *actual* X-ray image to one of the *nine* connection types investigated in this dissertation.

Bibliography

[1] M Zahangir Kabir and Safa Kasap. Photoconductors for x-ray image detectors. In *Springer Handbook of Electronic and Photonic Materials*, pages 1–1. Springer, 2017.

[2] Types of dental implants. https://qualitydentalcare.com.au/types-of-dental-implants-quality-dental-care-adelaide-south-australia, 2022. Accessed: 2022-09-17.

[3] Nathalie Lassau, Samy Ammari, Emilie Chouzenoux, Hugo Gortais, Paul Herent, Matthieu Devilder, Samer Soliman, Olivier Meyrignac, Marie-Pauline Talabard, Jean-Philippe Lamarque, et al. Integrating deep learning ct-scan model, biological and clinical variables to predict severity of covid-19 patients. *Nature Communications*, 12(1):1–11, 2021.

[4] Ravi Aggarwal, Viknesh Sounderajah, Guy Martin, Daniel SW Ting, Alan Karthikesalingam, Dominic King, Hutan Ashrafian, and Ara Darzi. Diagnostic accuracy of deep learning in medical imaging: A systematic review and meta-analysis. *NPJ Digital Medicine*, 4(1):1–23, 2021.

[5] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 234–241. Springer, 2015.

[6] Fausto Milletari, Nassir Navab, and Seyed-Ahmad Ahmadi. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In *2016 Fourth International Conference on 3D Vision (3DV)*, pages 565–571. IEEE, 2016.

[7] Mohammad Havaei, Axel Davy, David Warde-Farley, Antoine Biard, Aaron Courville, Yoshua Bengio, Chris Pal, Pierre-Marc Jodoin, and Hugo Larochelle. Brain tumor segmentation with deep neural networks. *Medical Image Analysis*, 35:18–31, 2017.

[8] Min Fu, Wenming Wu, Xiafei Hong, Qiuhua Liu, Jialin Jiang, Yaobin Ou, Yupei Zhao, and Xinqi Gong. Hierarchical combinatorial deep learning architecture for pancreas segmentation of medical computed tomography cancer images. *BMC Systems Biology*, 12(4):119–127, 2018.

[9] V Jasti, Abu Sarwar Zamani, K Arumugam, Mohd Naved, Harikumar Pallathadka, F Sammy, Abhishek Raghuvanshi, and Karthikeyan Kaliyaperumal. Computational technique based on machine learning and image processing for medical image analysis of breast cancer diagnosis. *Security and Communication Networks*, 2022, 2022.

[10] Eduardo Luz, Pedro Silva, Rodrigo Silva, Ludmila Silva, João Guimarães, Gustavo Miozzo, Gladston Moreira, and David Menotti. Towards an effective and efficient deep learning model for covid-19 patterns detection in x-ray images. *Research on Biomedical Engineering*, 38(1):149–162, 2022.

[11] Rafael C Gonzalez and Richard E Woods. *Digital image processing*. Pearson Education India, 2009.

[12] Jerry L Prince and Jonathan M Links. Medical imaging signals and systems. *Upper Saddle River: Pearson Prentice Hall*, Vol.37, 2006.

[13] Simone Carmignato, Wim Dewulf, and Richard Leach. *Industrial X-ray computed tomography*. Springer, 2018.

[14] Aviwe Kohlakala, Johannes Coetzer, Jeroen Bertels, and Dirk Vandermeulen. Deep learning-based dental implant recognition using synthetic x-ray images. *Medical & Biological Engineering & Computing*, 60(10):2951–2968, 2022.

[15] Jonathan Tremblay, Aayush Prakash, David Acuna, Mark Brophy, Varun Jampani, Cem Anil, Thang To, Eric Cameracci, Shaad Boochoon, and Stan Birchfield. Training deep networks with synthetic data: Bridging the reality gap by domain randomization. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 969–977, 2018.

[16] Artem Rozantsev, Vincent Lepetit, and Pascal Fua. On rendering synthetic images for training an object detector. *Computer Vision and Image Understanding*, 137:24–37, 2015.

[17] Shaoyong Yu, Yun Wu, Wei Li, Zhijun Song, and Wenhua Zeng. A model for fine-grained vehicle classification based on deep learning. *Neurocomputing*, 257:97–103, 2017.

[18] Brian Teixeira, Vivek Singh, Terrence Chen, Kai Ma, Birgi Tamersoy, Yifan Wu, Elena Balashova, and Dorin Comaniciu. Generating synthetic x-ray images of a person from the surface geometry. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 9059–9067, 2018.

[19] ChunFai Siu, Mingzhu Wang, and Jack CP Cheng. A framework for synthetic image generation and augmentation for improving automatic sewer pipe defect detection. *Automation in Construction*, 137:104213, 2022.

[20] Walter HL Pinaya, Petru-Daniel Tudosiu, Jessica Dafflon, Pedro F Da Costa, Virginia Fernandez, Parashkev Nachev, Sebastien Ourselin, and M Jorge Cardoso. Brain imaging generation with latent diffusion models. In *Deep Generative Models: Second MICCAI Workshop, DGM4MICCAI 2022, Held in Conjunction with MICCAI 2022, Singapore, September 22, 2022, Proceedings*, pages 117–126. Springer, 2022.

[21] António HJ Moreira, Sandro Queirós, Pedro Morais, Nuno F Rodrigues, André Ricardo Correia, Valter Fernandes, António CM Pinho, Jaime C Fonseca, and João L Vilaça. Voxel-based registration of simulated and real patient cbct data for accurate dental implant pose estimation. In *Medical Imaging 2015: Computer-Aided Diagnosis*, volume 9414, pages 852–857. SPIE, 2015.

[22] Pedro Morais, Sandro Queirós, António HJ Moreira, Adriano Ferreira, Ernesto Ferreira, Duarte Duque, Nuno F Rodrigues, and João L Vilaça. Computer-aided recognition of dental implants in x-ray images. In *Medical Imaging 2015: Computer-Aided Diagnosis*, volume 9414, pages 601–607. International Society for Optics and Photonics, 2015.

[23] Pedro Cunha, Miguel A Guevara, Ana Messias, Salomão Rocha, Rita Reis, and Pedro MG Nicolau. A method for segmentation of dental implants and crestal bone. *International Journal of Computer Assisted Radiology and Surgery*, 8 (5):711–721, 2013.

[24] Ruben Pauwels, Reinhilde Jacobs, Hilde Bosmans, Pisha Pittayapat, Pasupen Kosalagood, Onanong Silkosessak, and Soontra Panmekiate. Automated implant segmentation in cone-beam ct using edge detection and particle counting. *International Journal of Computer Assisted Radiology and Surgery*, 9(4): 733–743, 2014.

[25] Jae-Hong Lee, Young-Taek Kim, Jong-Bin Lee, and Seong-Nyum Jeong. A performance comparison between automated deep learning and dental professionals in classification of dental implant systems from dental imaging: A multi-center study. *Diagnostics*, 10(11):910, 2020.

[26] Mehdi Hadj Saïd, Marc-Kévin Le Roux, Jean-Hugues Catherine, and Romain Lan. Development of an artificial intelligence model to identify a dental implant from a radiograph. *International Journal of Oral & Maxillofacial Implants*, 35(6), 2020.

[27] Shintaro Sukegawa, Kazumasa Yoshii, Takeshi Hara, Katsusuke Yamashita, Keisuke Nakano, Norio Yamamoto, Hitoshi Nagatsuka, and Yoshihiko Furuki. Deep neural networks for dental implant system classification. *Biomolecules*, 10(7):984, 2020.

[28] Hak-Sun Kim, Eun-Gyu Ha, Young Hyun Kim, Kug Jin Jeon, Chena Lee, and Sang-Sun Han. Transfer learning in a deep convolutional neural network for implant fixture classification: A pilot study. *Imaging Science in Dentistry*, 52, 2022.

[29] Frank Natterer and Frank Wübbeling. *Mathematical methods in image reconstruction*. SIAM, 2001.

[30] Gengsheng Lawrence Zeng. *Medical image reconstruction: a conceptual tutorial*. Springer, 2010.

[31] Mario D'Acunto, Antonio Benassi, Davide Moroni, and Ovidio Salvetti. 3d image reconstruction using radon transform. *Signal, Image and Video Processing*, 10(1):1–8, 2016.

[32] Phlipp Henzler, Volker Rasche, Timo Ropinski, and Tobias Ritschel. Single-image tomography: 3d volumes from 2d cranial x-rays. In *Computer Graphics Forum*, volume 37, pages 377–388. Wiley Online Library, 2018.

[33] Liyue Shen, Wei Zhao, and Lei Xing. Patient-specific reconstruction of volumetric computed tomography images from a single projection view via deep learning. *Nature Biomedical Engineering*, 3(11):880–888, 2019.

[34] Hawraa H Al-waelly and Hameed M Abduljabbar. 3d image reconstruction from its 2d projection-a simulation study. *Progress in Industrial Ecology, an International Journal*, 14(1):10–18, 2020.

[35] Johann Radon. On the determination of functions from their integral values along certain manifolds. *IEEE Transactions on Medical Imaging*, 5(4):170–176, 1986.

[36] Gestur Ólafsson and Eric Todd Quinto. *The Radon Transform, Inverse Problems, and Tomography: American Mathematical Society Short Course, January 3-4, 2005, Atlanta, Georgia*, volume 63. American Mathematical Soc., 2006.

[37] Peter Toft. The radon transform. *Theory and Implementation (Ph.D. Dissertation)(Copenhagen: Technical University of Denmark)*, 1996.

[38] Stanley R Deans. *The Radon transform and some of its applications*. Courier Corporation, 2007.

[39] Adam Aitkenhead. Mesh voxelisation. https://www.mathworks.com/matlabcentral/fileexchange/27390-mesh-voxelisation, 2013.

[40] George Terzakis, Manolis Lourakis, and Djamel Ait-Boudaoud. Modified rodrigues parameters: an efficient representation of orientation in 3d vision and graphics. *Journal of Mathematical Imaging and Vision*, 60(3):422–442, 2018.

[41] Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(12):2481–2495, 2017.

[42] Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang. Unet++: A nested u-net architecture for medical image segmentation. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pages 3–11. Springer, 2018.

[43] Junsuk Choe and Hyunjung Shim. Attention-based dropout layer for weakly supervised object localization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2219–2228, 2019.

[44] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[45] David Robben, Jeroen Bertels, Siri Willems, Dirk Vandermeulen, Frederik Maes, and Paul Suetens. Deepvoxnet: voxel-wise prediction for 3d images. *Medical Image Computing (ESAT/PSI), KU Leuven, Belgium*, 2018.

[46] Howard Curtis. *Orbital mechanics for engineering students*. Butterworth-Heinemann, 2013.

[47] Shuiwang Ji, Wei Xu, Ming Yang, and Kai Yu. 3d convolutional neural networks for human action recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(1):221–231, 2012.

[48] A K Jain. A fast karhunen-loeve transform for a class of random processes. *IEEE Transactions on Communications*, 24(9):1023–1029, 1976.

[49] Michael Gastpar, Pier Luigi Dragotti, and Martin Vetterli. The distributed karhunen–loeve transform. *IEEE Transactions on Information Theory*, 52(12): 5177–5196, 2006.

[50] Joseph N Wilson and Gerhard X Ritter. *Handbook of computer vision algorithms in image algebra*. CRC press, 2000.

[51] Alisa I Walz-Flannigan, Kimberly J Brossoit, Dayne J Magnuson, and Beth A Schueler. Pictorial review of digital radiography artifacts. *Radiographics*, 38 (3):833–846, 2018.

[52] Ian T Young and Lucas J Van Vliet. Recursive implementation of the gaussian filter. *Signal Processing*, 44(2):139–151, 1995.

[53] Martín Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, et al. Tensorflow: a system for large-scale machine learning. In *Osdi*, volume 16, pages 265–283. Savannah, GA, USA, 2016.

[54] Francois Chollet. Keras. https://github.com/fchollet/keras, 2015.

[55] Veena Basappa Benakatti, Ramesh P Nayakar, Mallikarjun Anandhalli, et al. Machine learning for identification of dental implant systems based on shape–a descriptive study. *The Journal of Indian Prosthodontic Society*, 21(4):405, 2021.

[56] Matthias Karl and Ainara Irastorza-Landa. In vitro characterization of original and nonoriginal implant abutments. *International Journal of Oral & Maxillofacial Implants*, 33(6), 2018.