

***DESIGN AND DEVELOPMENT OF SOME  
METHODS AND MODELS FOR DENTAL IMAGE  
SEGMENTATION USING DEEP LEARNING  
APPROACHES***

गहन शिक्षण दृष्टिकोण का उपयोग करके दंत छवि विभाजन के लिए कुछ तरीकों और मॉडलों का डिजाइन और विकास



**Thesis submitted in partial fulfilment**

**for the award of degree**

***DOCTOR OF PHILOSOPHY***

**By**

**SAURABH ARORA**

**सौरभ अरोड़ा**

**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**

**INDIAN INSTITUTE OF TECHNOLOGY**

**(BANARAS HINDU UNIVERSITY)**

**VARANASI – 221 005**

**Roll No: 17071505**

**December, 2023**

## Certificate

It is certified that the work contained in the thesis titled “**DESIGN AND DEVELOPMENT OF SOME METHODS AND MODELS FOR DENTAL IMAGE SEGMENTATION USING DEEP LEARNING APPROACHES**” has been carried out under my supervision and that this work has not been submitted elsewhere for a degree.

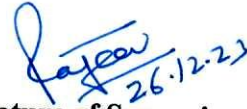
It is further certified that the student has fulfilled all the requirements of Comprehensive, Candidacy and State-of-the-art seminar.



**Signature of Co-Supervisor**

Dr. Ruchir Gupta

Associate Professor, Department of  
Computer Science and Engineering  
Indian Institute of Technology (BHU),  
Varanasi



**Signature of Supervisor**

Prof. Rajeev Srivastava

Professor, Department of Computer  
Science and Engineering  
Indian Institute of Technology (BHU),  
Varanasi

पर्यवेक्षक/Supervisor  
संगणक विज्ञान एवं अभियांत्रिकी विभाग  
Department of Computer Sc. & Engg  
भारतीय प्रौद्योगिकी संस्थान  
Indian Institute of Technology  
(काशी हिन्दू विश्वविद्यालय)  
(Banaras Hindu University)  
वाराणसी Varanasi-221005

## Declaration by the Candidate

I, **Saurabh Arora** certify that the work embodied in this thesis is my own bonafide work and carried out by me under the supervision of **Prof. Rajeev Srivastava (Supervisor)** and **Dr. Ruchir Gupta (Co-Supervisor)** from **December 2017 to December 2023**, at the **Department of Computer Science and Engineering**, Indian Institute of Technology (BHU), Varanasi. The matter embodied in this thesis has not been submitted for the award of any other degree/diploma. I declare that I have faithfully acknowledged and given credits to the research workers wherever their works have been cited in my work in this thesis. I further declare that I have not wilfully lifted up any other's work, paragraphs, text, data, results, etc., reported in journals, books, magazines, reports dissertations, theses, etc., or available at websites and included them in this thesis and cited as my own work.

Date: 26/12/2023

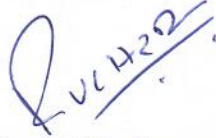
Place: IIT (BHU), Varanasi



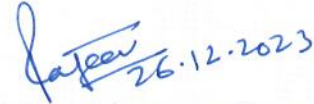
(Saurabh Arora)

## CERTIFICATE BY THE SUPERVISOR

It is certified that the above statement made by the student is correct to the best of my/our knowledge.

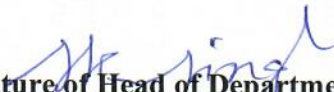


**Dr. Ruchir Gupta (Co-Supervisor)**  
Associate Professor, Department of  
Computer Science and Engineering  
Indian Institute of Technology (BHU),  
Varanasi



**Prof. Rajeev Srivastava (Supervisor)**  
Professor, Department of Computer  
Science and Engineering  
Indian Institute of Technology (BHU),  
Varanasi

प्राध्यापक/Supervisor  
समन्वयक विज्ञान एवं अभियांत्रिकी विभाग  
Department of Computer Sc. & Engg  
भारतीय प्रौद्योगिकी संस्थान  
Indian Institute of Technology  
(काशी हिन्दू विश्वविद्यालय)  
(Banaras Hindu University)  
वाराणसी Varanasi-221005

  
Signature of Head of Department  
(Prof. Sanjay Kumar Singh)

अधीक्षक एवं विभागाध्यक्ष  
Professor & Head  
समन्वयक विज्ञान एवं अभियांत्रिकी विभाग  
Department of Computer Sc. & Engg  
भारतीय प्रौद्योगिकी संस्थान  
Indian Institute of Technology  
(बनारस हिन्दू यूनिवर्सिटी)  
(Banaras Hindu University)  
वाराणसी-221005 / Varanasi-221005

## Copyright Transfer Certificate

---

**Title of the Thesis:** *Design and Development of Some Methods and Models for Dental Image Segmentation using Deep Learning Approaches.*

**Name of the Student:** Saurabh Arora

### Copyright Transfer

The undersigned hereby assigns to the Indian Institute of Technology (Banaras Hindu University), Varanasi all rights under copyright that may exist in and for the above thesis submitted for the award of the DOCTOR OF PHILOSOPHY.

Date: 26/12/2023

Place: IIT (BHU), Varanasi



Signature of the Student

(Saurabh Arora)

**Note:** However, the author may reproduce or authorize others to reproduce material extracted verbatim from the thesis or derivative of the thesis for author's personal use provided that the source and the Institute's copyright notice are indicated.

# Acknowledgement

---

I would like to express my deepest gratitude to all those people who helped me directly or indirectly to complete this thesis work. Firstly, I would like to express my heartfelt gratitude to my supervisor, **Prof. Rajeev Srivastava** and co-supervisor Dr. Ruchir Gupta, for being a great mentor and the best adviser I could ever have. Their advice, encouragement and critics are source of innovative ideas, inspiration are causes behind the successful completion of this Thesis work. The confidence shown on me by them was the biggest source of inspiration for me. It has been a privilege working with them from several years.

I would like to express my deepest appreciation to my research progress evaluation committee members Dr. K. Lakshmanan of the Department of Computer Science and Engineering and Prof. Subir Das, Department of Mathematical Sciences IIT (BHU), for providing continuous support, encouragement, and advice.

I express my sincere thanks to all the Professors, Deans, office staff, supporting staff and PhD Research Scholars of Indian Institute of Technology (BHU) Varanasi India. I express my gratitude to Director, Registrars, Deans, Heads, and Student Alumni of the Indian Institute of Technology (BHU) Varanasi.

Special thanks to Dr. Subodh Srivastava (Assistant Professor, NIT, Patna) for his continuous support during my study. I'd like to extend my gratitude to Dr. Santosh Kumar Tripathy (Assistant Professor, VIT Bhopal University), Dr. Roshan Singh (System Analyst IIT BHU Varanasi), Dr. Vibhav Prakash Singh (Assistant Professor MNNIT Allahabad), Dr. Sumit Kumar (Assistant Professor, Galgotia University, Greater Noida).

I also wish to thank my lab members Rajat Arya, Shraddha Jain, Shweta Singh and Vipin Maurya for their consistent support and help during my research work.

I extend special thanks to the non-teaching staff in the Department, particularly, Mr. Manoj Singh, Mr. Ravi Bharti, Mr. Subham Pandey, Mr. Prakhar for their consistent support.

My parents, Smt. Madhu Arora and Shri Pradeep Kumar Arora, who gave me the power and brain to work out on this research and their help at every level made me to see this success.

Words are insufficient to express my profound sense of gratitude to my loving sister Jyotika. Her encouragement and support gave me physical and moral strength throughout my career as well in the present research. I extend my thanks to my brother-in-law Pratyush who is my part of inspiration. Finally, I would like to wind up by paying my heartfelt thanks and prayers to the Almighty for his unbound love and grace.



Saurabh Arora

# Table of content

---

---

<b>Certificate .....</b>	<b>ii</b>
<b>Declaration by the Candidate .....</b>	<b>iii</b>
<b>Copyright Transfer Certificate .....</b>	<b>iv</b>
<b>Acknowledgement .....</b>	<b>v</b>
<b>Preface.....</b>	<b>vii</b>
<b>Table of content.....</b>	<b>x</b>
<b>List of Figures.....</b>	<b>xiii</b>
<b>List of Tables .....</b>	<b>xvi</b>
<b>List of Symbols .....</b>	<b>xvii</b>
<b>List of Abbreviations .....</b>	<b>xviii</b>
<b>Chapter 1 Introduction.....</b>	<b>1</b>
1.1. Background .....	1
1.2 Dental Image Segmentation .....	3
1.3 Motivation .....	4
1.4 Problem Statement .....	5
1.5 Research Objectives .....	6
1.6 Contributions to the Thesis.....	6
1.7 Thesis Organization.....	7
<b>Chapter 2 Literature Review .....</b>	<b>9</b>
2.1 Literature Review for Dental Image Segmentation .....	9
2.1.1 Conventional Approaches.....	9
2.1.2 Deep Learning Techniques.....	14
2.1.3 Literature for Capsule Network for medical image segmentation .....	24
2.1.4 Advances made by industry for Dentistry .....	29
2.2 Research Gaps .....	30
2.3 Dataset used for Experimental Analysis.....	31
2.3.1 UFBA_USEC Dataset .....	32
2.3.2 Tufts Dental Database .....	33
2.4 Performance Metrics .....	33
2.5 Conclusion.....	34

<b>Chapter 3 Teeth image segmentation methods using Multimodal CNN architecture and Attention Guided deep network.....</b>	<b>35</b>
3.1 Background .....	36
3.2 Related Work .....	36
3.3 Proposed Methods.....	37
3.3.1 Exploiting Multimodal CNN architecture for automated teeth segmentation on dental panoramic X-ray images .....	37
3.3.2 Cascaded Deep Neural Network with Attention Guidance for Teeth Segmentation on Dental Panoramic Radiographs .....	48
3.4 Conclusion .....	70
<b>Chapter 4 Enhancing teeth segmentation using multifusion deep neural net in panoramic X-rays</b>	<b>71</b>
4.1 Background .....	71
4.2 Related Work .....	72
4.3 Proposed Methods.....	75
4.3.1 Overview .....	75
4.3.2 Detailed Architecture of Proposed Network .....	76
4.4 Experiments and Results.....	80
4.4.1 Datasets .....	80
4.4.2 Experimental Setup .....	80
4.4.3 Result and Discussion .....	81
4.4.4 Ablation Study.....	86
4.5 Conclusion .....	90
<b>Chapter 5 TeethCaps: A dental image segmentation model based on capsule network .....</b>	<b>92</b>
5.1 Background.....	92
5.2 Related Work .....	94
5.3 Proposed Method .....	96
5.3.1 Overview .....	96
5.3.2 Details of the Proposed Architecture TeethCaps.....	97
5.4 Experiments and Results.....	101
5.4.1 Datasets .....	101
5.4.2 Experimental Setup .....	101
5.4.3 Result and Discussion .....	101
5.5 Conclusion .....	105
<b>Chapter 6 Investigation of Transformer-based neural network for dental image segmentation ...</b>	<b>107</b>
6.1 Background.....	107
6.2 Related Work .....	108
6.3 Proposed Method .....	109
6.3.1 Proposed Network Architecture .....	109
6.4 Experiment and Results.....	111
6.4.1 Datasets .....	111

6.4.2 Experimental Setup.....	111
6.4.3 Result and Discussion.....	112
6.5 Conclusion.....	116
<b>Chapter 7 Conclusion and Future Scope .....</b>	<b>118</b>
7.1 Conclusions .....	118
7.2 Suggestion for Future Research.....	122
<b>References.....</b>	<b>124</b>
<b>Appendix A.....</b>	<b>133</b>
<b>List of Publications.....</b>	<b>135</b>

## List of Figures

---

Figure 1.1 General deep learning framework for dental image segmentation.....	3
Figure 2.1 Sample from Datasets of dental panoramic X-ray images .....	32
Figure 3.1 Architecture of the proposed model .....	40
Figure 3.2 Training versus Validation Loss. The plot shows the training and validation loss of the model for each epoch.....	42
Figure 3.3 Qualitative results of the predicted segmentation maps. (a) Original Images (b) Ground truth segmentation maps of corresponding maps (c) Predicted segmentation maps .....	44
Figure 3.4 Qualitative results of the predicted segmentation maps of the modules. (a) Modules (b) Original Images (c) Ground truth (d) Predicted segmentation maps .....	47
Figure 3.5 Block Diagram of the presented methodology.....	50
Figure 3.6 The detailed architecture of the proposed model. ....	51
Figure 3.7 TBAB: Teeth boundary attention block .....	54
Figure 3.8 The example of noisy and blurry segmented map predicted by Model-1 marked by yellow lines whereas predicted segmented map by Model-2 removes corrupted pixels. (a) Predicted segmented predicted by Model-1. (b) Segmented Map of Teeth Area predicted by Model-2. (c) The ground truth of the correspondent images.....	55
Figure 3.9 Visual results obtained from the proposed model (a) Original Images from the dataset (b) Ground Truth map of the respective image (c) Predicted Segmented Teeth Area by the proposed model. ....	63
Figure 3.10 Visual results of the ablation study of different modules (a) Original Image of Dataset (b) Ground truth of respective image (c) Predicted segmented image by each module. ....	69

Figure 4.1 The detailed architecture of the proposed model. ....78

Figure 4. 2 Losses of proposed model for both datasets. (a) The plot demonstrates the training loss and validation loss for dataset 1 it seems the model converges approximately at 55<sup>th</sup> epoch. (b) This plot demonstrates training loss and validation loss for dataset 2 which seems to be converged at epoch 60.....81

Figure 4.3 Comparison of visual results of proposed model with different segmentation based deep models on Dataset 1. The first two rows represent original image and its corresponding ground truth while the remaining rows represents the predicted segmented teeth map by different methods including the proposed model. ....84

Figure 4.4 Comparison of visual results of proposed model with different segmentation based deep models on Dataset 2. The first two rows represent original image and its corresponding ground truth respectively while rest shows segmented teeth region mask predicted by various deep models. ....85

Figure 4.5 Training vs Validation loss for Dataset 1. (a) This plot represents the training and validation loss for module without skip connections. (b) The loss for module 2 having long skip connections. (c) This plot shows loss for the third module having only short skip connections. (d) This represents the training and validation loss for proposed model. All modules including proposed model is converged at epoch 60.....87

Figure 4.6 Training vs Validation loss for Dataset 2. (a) This plot represents the loss for first module with no skip connections. (b) The training and validation loss for second module having long skip connections is shown. (c) This plot shows loss for the third module having only short skip connections. (d) This represents the training and validation loss for proposed model. All modules including proposed model is converged at epoch 60.....88

Figure 4.7 Comparison of visual results from ablation experiment on Dataset 1. The first two rows represent original image and its ground truth respectively. The next three rows demonstrate the predicted segmented map by different modules while the last row represents the predicted segmented map by proposed model.....	90
Figure 5.1 Architecture details of the proposed TeethCaps .....	98
Figure 5.2 Visual results comparative analysis of different methods for Dataset 1 with proposed capsule-based model .....	104
Figure 5.3 Visual results comparative analysis of different methods for Dataset 2 with proposed capsuled-based model.....	105
Figure 6.1 Architecture of the proposed model with transformer.....	111
Figure 6.2 Visual results of the transformer based deep models for dataset 1. ....	113
Figure 6. 3 Visual results of the transformer based deep models for dataset 1. ....	114
Figure 6.4 Visual results of the transformer based deep models for dataset 2. ....	115
Figure 6.5 Visual results of the transformer based deep models for dataset 2. ....	116

## List of Tables

---

Table 2.1 Summary of conventional approaches for dental image segmentation .....	12
Table 2.2 Summary of deep learning techniques for dental image segmentation .....	22
Table 2.3 Summary of medical image segmentation using Capsule Network .....	28
Table 3.1 The details of the layers used in proposed model.....	40
Table 3.2 Performance comparison with state-of-the-art deep learning approaches .....	44
Table 3.3 Performance comparison with conventional approaches .....	45
Table 3. 4 Comparison of several modules in ablation study .....	46
Table 3.5 Quantitative result comparison with state-of-the-art deep learning methods..	62
Table 3.6 Quantitative result comparison with conventional methods .....	64
Table 3.7 Quantitative results of ablation study of different modules. ....	66
Table 4.1 Performance summary of the proposed model compared with deep segment state-of-the-art methods for Dataset 1 .....	82
Table 4.2: Performance summary of the proposed model compared with deep segment state-of-the-art methods for Dataset 2. ....	82
Table 4 3 Ablation experiment comparison of different modules on Dataset 1. ....	88
Table 4.4 Ablation experiment comparison of different modules on Dataset 2. ....	89
Table 5.1 Comparative analysis of different methods for Dataset 1 with proposed capsule bed model.....	102
Table 5. 2 Comparative analysis of different methods for Dataset 2 with proposed capsule based model .....	103
Table 6.1 Performance comparison of deep methods using a transformer structure for dataset 1 .....	112
Table 6.2 Performance comparison of deep methods using a transformer structure for dataset 2 .....	112
Table 7.1 Comparison of all the proposed techniques .....	119

## List of Symbols

---

$\cup$	Union
$\cap$	Intersection
$+$	Addition
$T_{TP}$	The number of correctly forecasted pixels of teeth region
$F_{FP}$	The number of wrongly forecasted pixels of teeth region
$T_{TN}$	The number of correctly forecasted pixels of background region
$F_{FN}$	The number of wrongly forecasted pixels of background region
$\theta$	All the trainable parameter
$P$	The predicted segmented map output
$\checkmark$	Applicable
$G$	Ground truth set of input image
$\theta_{TBAB1}$	All the trainable parameters for module TBAB1
$\theta_{TBAB2}$	All the trainable parameters for module TBAB2
$\theta_{SegBlock}$	All the trainable parameters for module SegBlock
$\mu_{y_i}$	Mean of predicted teeth segmented map
$\mu_{g_i}$	Mean of ground truth map
$\sigma_{y_i}$	Standard deviations predicted teeth segmented map
$\sigma_{g_i}$	Standard deviations teeth ground truth
$C_f$	Channel attention map
$\odot$	Represents elementwise multiplication
$S_f$	Spatial attention map
$F_m$	Feature map
$ct_i^m$	Capsule-type in layer m
$\hat{u}_{a,b ct_i^m}$	Prediction vector of the convolutional capsules
$u_{a,b ct_i^m}$	Activation vector of the child capsule
$d_{ct_i^m a,b}$	Coupling coefficient between child capsule and parent capsules
$k_{ct_i^m a,b}$	log prior probabilities

## List of Abbreviations

---

CBCT	Cone-Beam Computed Tomography
CNN	Convolutional Neural Network
GPU	Graphics processing unit
CBA	Crowd Behaviour Analysis
SVM	Support Vector Machine
DCNN	Deep Convolutional Neural Network
RCNN	Region-Based Convolutional Neural Network
FCNN	Fully Connected Neural Network
FCN	Fully Convolutional Neural Network
GCN	Graph Convolutional Network
IOS	Intraoral Scan
MRIs	Magnetic Resonance Imaging
GAN	Generative Adversarial Neural Network
CT	Computed Tomography
OCT	Optical Coherence Tomography
PDDCA	Public Domain Database For Computational Anatomy
ReLU	Rectified Linear Units
TBAB	Tooth Boundary Attention Blocks
SSIM	Structural Similarity Index Measure
BCE	Binary Cross Entropy
CAM	Channel Attention Module
SAM	Spatial Attention Module
IoU	Intersection Over Union
IVUS	Intravascular Ultrasound

# Preface

---

Dental image segmentation is an important task in dental image analysis and plays a crucial role in various dental applications, such as diagnosis, treatment planning, and clinical research. With the advent of different imaging modalities, there has been a significant increase in the complexity and variability of dental images, thus making the dental image segmentation task more challenging. Recent advancements in deep learning techniques have led to significant improvements in dental image segmentation accuracy and efficiency.

This thesis mainly focuses on studying and analysing state-of-the-art dental image segmentation techniques, finding their advantages and limitations, and proposing new methods and models to accomplish the objectives. This thesis aims to conduct a comprehensive literature review on segmentation models for dental images. The various methods and models concerning current research trends have been analysed by mentioning their pros and cons and identifying possible research gaps. Various models using deep learning approaches have been proposed in this thesis to fulfil the research gap in segmenting dental images.

The first contribution in the thesis is related to the task of teeth area segmentation from dental panoramic X-ray images, where two models using deep learning approaches have been proposed. The first proposed model exploits the Multimodal CNN architecture for automated segmentation of teeth region. This model consists of multiCNN encoder and decoder. The encoder has three different streams: conventional CNN, dilated CNN and separable CNN. The main objective behind the Multimodal CNN architecture is to combine the advantages of different CNN streams for rich dental information to improve

segmentation performance. The second proposed model for teeth area segmentation is a novel cascaded deep architecture with attention guidance for tooth boundaries. This model comprises of two deep models. The first model has an attention-guided encoder-decoder network while the second one is the deep convolutional architecture. The purpose of this model is to generate a contextual rich feature map with a focus on tooth boundaries using the proposed attention block TBAB. The second model maintains the semantic information and improves the quality of the predicted segmented map removing blurry pixels and generating accurate and clear segmentation maps.

The second contribution to the thesis is to enhance teeth segmentation by using multifusion deep neural network. This deep neural network comprises of four major components: encoder, decoder, skip connections and hybrid loss function. The encoder gathers detailed teeth information whereas decoder is responsible for reconstructing the predicted segmented maps. The skip connections are of two types short and long which maintain texture information in an encoder and preserve low-level information in a decoder. The model is trained using a combination of binary cross entropy loss and dice loss to maintain structural information of teeth region. The objective of this model is to generate precise and enhanced teeth segmentation map with clear teeth region boundaries.

The third contribution is TeethCaps a deep segmentation model based on a capsule network for segmenting teeth regions from panoramic radiographs. This model has three modules, a feature extractor, an attention block and a capsule block. The feature extractor has convolution and dilated convolutions and is responsible for preserving the most discriminant features. The second module is an attention block having CBAM block which gives attention to channel and spatial features and generates refined feature maps. The third is the capsule block consists of convolutional and deconvolutional capsule layers which maintain the spatial information while performing segmentation.

The fourth contribution of the thesis is to segment the precise teeth region by a novel deep encoder-decoder architecture utilizing the transformers. The transformer structure is fused with the encoder-decoder architecture to handle spatial information in panoramic images and leverages its attention mechanism for capturing contextual dependencies. Skip connections have been made in the model to retain the features of the original image, ensuring adequate reconstruction of the image. Various deep segmentation models using transformers were also implemented and analysed for teeth region segmentation on dental images.

The proposed model has been implemented and validated on publicly available benchmark datasets. The results and analysis reveal that the proposed models outperform state-of-the-art methodologies presented in the literature. Extensive ablation of some proposed models has also been conducted to demonstrate the effect of various components of the proposed models.

The thesis concludes with an overall conclusion of the proposed research work, followed by a discussion of future research directions in the field of dental image segmentation.