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

26.12.23

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

26.12.2023

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

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

अध्यक्ष एवं विभागाध्यक्ष
Professor & Head
संगणक विज्ञान एवं अभियांत्रिकी विभाग
Department of Computer Sc. & Engg
भारतीय प्रौद्योगिकी संस्थान
Indian Institute of Technology
(बनारस हिन्दू यूनिवर्सिटी)
(Banaras Hindu University)
वाराणसी–२२१००५ / 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

List of Figures

List of Tables

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
θ	All the trainable parameter
P	The predicted segmented map output
✓	Applicable
G	Ground truth set of input image
θ_{TBAB1}	All the trainable parameters for module TBAB1
θ_{TBAB2}	All the trainable parameters for module TBAB2
$\theta_{SegBlock}$	All the trainable parameters for module SegBlock
μ_{y_i}	Mean of predicted teeth segmented map
μ_{g_i}	Mean of ground truth map
σ_{y_i}	Standard deviations predicted teeth segmented map
σ_{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.