CERTIFICATE

It is certified that the work contained in this thesis entitled " COMPLEX SALIENT OBJECT DETECTION USING PROBABILISTIC MODELLING AND DEEP LEARNING" by "Surya Kant Singh" has been carried out under my supervision and that it has not been submitted elsewhere for a degree.

It is further certified that the student has fulfilled all the requirements of Comprehensive, Candidacy and SOTA.

24.5.2022

Prof. Rajeev Srivastava

Professor

Department of Computer Science and Engineering

Indian Institute of Technology

(Banaras Hindu University)

Varanasi-221005

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, Surya Kant Singh , 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 from JULY-2017 to MAY-2022, at the Department of Computer Science and Engineering, Indian Institute of Technology (Banaras Hindu University) 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 willfully 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 : 24/05/2022

Place : Varanasi

Signature of the Student

(Surya Kant Singh)

CERTIFICATE BY THE SUPERVISOR

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

24.5.2022

Signature of Supervisor

(Prof. Rajeev Srivastava)
Department of Computer Science and Engineering
Indian Institute of Technology
(Banaras Hindu University)
Varanasi-221005

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)

आचार्य व 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: COMPLEX SALIENT OBJECT DETECTION USING PROBABILISTIC MODELLING AND DEEP LEARNING

Name of the Student: Surya Kant Singh

Copyright Transfer

Date : 24/05/2022

Place : Varanasi

(Surya Kant Singh)

To

My Beloved Parents

Mrs. JANKEE SINGH Mr. NANDLAL SINGH

and

My Better Half
Mrs. NEETU SINGH

ACKNOWLEDGEMENTS

I want to take this opportunity to express my deep sense of gratitude to all who helped me directly or indirectly during this thesis work. Firstly, I would like to thank my supervisor, **Prof. Rajeev Srivastava**, for being a great mentor and the best adviser I could ever have. His advice, encouragement, and critics are a source of innovative ideas, inspiration, and causes behind the successful completion of this Thesis work. The confidence shown on me by him was the most significant source of inspiration for me. It has been a privilege working with him for several years. I am highly obliged to all the faculty members of the Computer Science and Engineering Department for their support and encouragement. I express my sincere thanks to Prof. K. K. Shukla 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 Ph.D. Research Scholars of Indian Institute of Technology (BHU) Varanasi India. I express my gratitude to Director, Registrar, Deans, Heads, Students and Alumni of the Indian Institute of Technology (BHU) Varanasi.

My memory of the study period at IIT (BHU) can never be complete without mentioning my fellow research scholars. Special thanks to Dr. Vibhav Prakash Singh, Dr. Ankit Jaiswal, Dr. Gargi Srivastava, Mr. Saurabh Arora and Mr. Santosh Kumar Tripathy for their great help and cooperation.

I extend special thanks to the non-teaching staff in the department, particularly the Mr. Manoj Kumar Singh, Mr. Ravi Kumar Bharti, Mr. Prakhar Kumar, Mr. Ritesh Singh, and Mr. Shubham Pandey for their consistent support.

My parents, Mrs. Jankee Singh and Mr. Nandlal Singh, who gave me the power and brain to work out on this research and their help at every level, made me see this success. I owe thanks to my wife, Mrs. Neetu Singh Solanki, for his continued and unfailing love, support, and understanding during my pursuit of Ph.D. degree that made the completion of my thesis possible. You were always around at times

I thought that it is impossible to continue, you helped me to keep things in perspective. I greatly value his contribution and sincerely appreciate his belief in me. I want to share my source of happiness during the up and down in the Ph.D. journey are the smiles of my daughter Namya and son Shresth.
I extend my thanks to my uncle Baban singh, my sister Shivangi Singh and my brother Dhan jee, whose blessings gave me the internal power to do hard work.
Words are insufficient to express my profound sense of gratitude to respected Prof. Shreesh Chaudhary and my friends Dr. Roshan Singh, Mr. Akash kumar singh, and Mr. Vipul Kumar Jha, whose encouragement gave me physical and moral strength throughout my career as well in the present research. I extend my thanks to my father in-law Dr. D.V Singh and my Mother in-law Mrs. Bindu Singh who are part of my inspiration. Finally, I would like to wind up by paying my heartfelt thanks and prayers to the Almighty for his unbound love and grace.

- Surya Kant Singh

Contents

List of Figures

List of Tables

Abbreviations

GR	**G**raphical **R**anking
ASNet	**A**ttentive **S**aliency **N**etwork
AUC	**A**rea **U**nder **C**urve
MCDL	**M**ulti **c**ontext **D**eep *L*earning
CNN	**C**onvolutional **N**eural **N**etwork
ESL	**E**nhance and **S**uppress **L**uminance
ECSSD	**E**xtended **C**omplex **S**cene **S**aliency **D**ataset
FCN	**F**ully **C**onvolutional **N**eural Networks
FN	**F**alse **N**egative
FP	**F**ixation **P**rediction
GAN	**G**enerative **A**dversarial **N**etwork
HED	**H**olistically-Nested **E**dge **IoU**
Intersection **o**ver **U**nion	
MAE	**M**ean **A**bsolute **E**rror
MDB	**M**inimum **D**irectional **B**backgroundness
MDC	**M**inimum **D**irectional **C**ontrast
MSRA	**M**icrosoft **R**esearch **A**sia
MSRC	**M**icrosoft **R**esearch **C**ambridge
PR	**P**recision **R**ecall
PC	**P**oisson **P**robabilistic **C**ontrast **RBD**
Robust **B**ackground **D**etection	
R-CNN	**R**egion-**C**onvolutional **N**eural **N**etwork

textbfRGB	**R**ed **G**reen **B**lue
RGBD	**R**ed **G**reen **B**lue **D**epth
ROC	**R**eceiver **O**perating **C**haracteristics
RPN	**R**egion **P**roposal **N**etwork
SOD	**S**alient **O**bject **D**etection
VGG	**V**isual **G**eometry **G**roup
GU	Global **U**niqueness
HC	**H**istogram-based global **C**ontrast
RC	**R**egional **C**ontrast
DoG	**D**ifference of **G**aussian
MSS	**M**aximum **S**ymmetric **S**urround
MBD	**M**inimum **B**arrier **D**istance
AMC	**A**bsorbing **M**arkov **C**hain
MST	**M**inimum **S**panning **T**ree
FPR	**F**alse **P**ositive **R**ate
TPR	**T**rue**P**ositive **R**ate
ILG	**I**terative **L**aplacian of **G**aussian
FIT	**F**eature **I**ntegration **T**heory
DHS	**D**eep **H**ierarchical **S**aliency
GTS	**G**lobal **T**opographical **S**urface
MADW	**M**utual **A**ttention based **D**istinguished **W**indow
SDD	**S**elf **L**earning-Based **D**ense **D**ecoder
JL-DCF	**J**oint **L**earning and **D**ensely **C**ooperative **F**usion
S2MA	**S**elf-**S**elective **M**utual **A**ttention
CFF	**C**oss-**C**omplementary **F**usion
CSA-Net	**D**eep **C**ross-**C**omplementary **S**elf **A**ttention
MSF	**M**odality-**S**pecific **S**aliency **F**usion Model
CFF	**C**ross-**C**omplementary **F**eatures **F**usion

IFF	**I**ntra-**C**omplementary **F**eatures **F**usion
OSS	**O**ptimal **S**elective **S**aliency **F**usion **M**odel

Symbols

PDF	Poisson probability distribution
T_f	Enhance and Suppress Luminance-(ESL) coefficient
dd	Luminance contrast normalization coefficient
d	Angular threshold
λ	Represents the complex image indicator parameter
AD_j	Regional areal density
$Dis_0(r_k, r_j)$	Spatial distance weight
σ	Controlling parameter
IF_w	Gaussian weighted background suppression
μ	Mean
$Areal_k(r)$	Spatial regional saliency
$S_{RC}(r_k)$	Reginol distance based saliency
$S_{RCC}(r_k)$	Reginol color saliency
$IFactor_w(r_k)$	Integrating factor
S_{Center}	Central saliency
$\psi(...)$	Up-sampling function
$\varphi(...)$	Convolution operation
$Sig(...)$	Sigmoid function
$Con(i, j)$	Convolution operation
$Max - pool(i, j)$	Max-pooling operation
E_ψ	E-measure

M_{att}	Mutual attention feature maps
f_g	Global Localized feature
f_h^i	High level sementic features
f_c^i	Cross-complemntary features

PREFACE

The phenomenal use of digital media in our day-to-day life and society brings us to the corner of another world. This transformation appeals and inspires me to innovate and add something to this revolution. Images play an essential role in online shopping and purchasing to online education. It empowers us in various ways, like photo and information sharing, medical imaging, simulation, and military handling of the pandemic. Together with the empowerment, it is necessary to organize these images and use them for various computer vision purposes. One of the most critical computer vision tasks is salient object detection. Salient object detection is to identify the most prominent and relevant object in images. Humans can do this task automatically. The computer can learn to do this by understanding the human visual attention mechanism. It has diverse applications in image captioning, image segmentation, object recognition, content-based image retrieval, image and video compression, and video summarization. This thesis aims to develop computer algorithms, methods, and models for salient object detection, which can detect salient objects similar to human beings.

The main focus of the thesis is to generate probabilistic models for salient object detection to simulate structure, shape, size, and background uncertainty. While designing these models, an attempt is made to discover the current research gaps in existing algorithms, methods, and models. For this, a comprehensive literature survey is performed. Research gaps of these fields are identified. After that, statistical, probabilistic, and deep learning-based models are developed to address the identified research gaps.

The five models have been developed for salient object detection in 2D and 3D modalities. Among these models, three are based on probabilistic approaches and rest are based on deep learning based approaches. The first two models propose probabilistic contrast and edge enhanced Global Topographical Surface. This surface encloses the prominent object with all its structural and spatial information. Then it is used as a reference plane for regional depth, color, and spatial saliency integration. Most saliency methods generate prominent objects from 2D information, while human attention systems are 3D perception mechanisms. Inspired by this perception, the following models are proposed based on 3D saliency.

The third proposed models is utilized additional depth information from RGBD to robustly and correctly detect the salient object in a complex and cluttered background. To distinguish the salient object in complex and cluttered background, the Poisson probabilistic contrast space is proposed. This process produces a global concave reference surface. The various regional saliencies are integrated into this global concave reference surface to detect the salient object correctly. Background estimation and central saliency integration will thoroughly remove the background. This algorithm generates a robust conspicuous object.

The thesis is presented in a way so that the conventional, as well as newer models like deep learning, are covered. The exiting models incur inconsistency or distribution loss of salient points and regions. These drawbacks are targets to design two deep network models. The fourth model is CSA-Net, which produces essential features: non-complementary, cross-complementary, intra-complementary, and deep localized improved high-level features. The designed 2×3 encoder and decoder streams produce these essential features and assure modality-specific saliency preservation. The cross and intra- complementary fusion are deeply guided by proposed novel, cross-complementary self-attention to produce fused saliency. The attention map is computed by two-stage additive fusion based on a Non-Local network.

Finally, in fifth model, a composite backbone network-based deep CNN framework is designed to produce more acute saliency in a challenging and complex scenario. This composite backbone produces enhanced features that are fused with a self-learning-based dense decoder. A comparative analysis of the proposed models are summarized and discussed. The models are compared against state-of-the-art algorithms and evaluated on benchmark databases. Possible future works and research directions are also included in the last chapter.