

# Utility of Deep Learning in breast cancer imaging in India

Krithika Rangarajan



School of Information Technology  
Indian Institute of Technology, Delhi  
January 2025

**© Indian Institute of Technology Delhi (IITD), New Delhi, 2025**

# Utility of Deep Learning in breast cancer imaging in India

*submitted by*

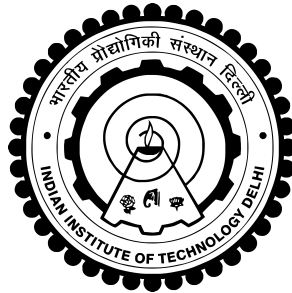
**Krithika Rangarajan**  
**ANZ198277**

School Of Information Technology (SIT)

*in fulfilment of the requirements of the degree of*

*of*

Doctor of Philosophy



INDIAN INSTITUTE OF TECHNOLOGY DELHI

**January 2025**

## THESIS CERTIFICATE

This is to certify that the thesis titled **Utility of Deep Learning in breast cancer imaging in India**, submitted by **Krithika Rangarajan (2019ANZ8277)**, to the Indian Institute of Technology, Delhi, for the award of the degree of **Doctor of Philosophy**, is a bona fide record of the research work done by her under our supervision. The contents of this thesis, in full or in parts, have not been submitted to any other Institute or University for the award of any degree or diploma.

**Chetan Arora**  
Professor  
Dept. of Computer Science  
IIT-Delhi, 600 036

Place: New Delhi

Date: 31 January 2025

**Subhashis Banerjee**  
Professor  
Dept. of Computer Science  
IIT-Delhi, 600 036

Place: New Delhi

Date: 31 January 2025

## ACKNOWLEDGEMENTS

*At the outset I would like to thank the faculty of the School of Information Technology, particularly **Dr Chetan Arora and Dr Subhashis Banerjee**, who guided me and who were extremely mindful of my clinical responsibilities while devising the curriculum and the research plan of this thesis. I am also particularly thankful to **Dr Prem Kalra**, the head of SIT when I joined, who responded to my email and worked out the logistics for me to be a post doctoral researcher in the department, which set the path for my subsequent PhD*

*I express my heartfelt gratitude to all the student collaborators, without whose patient guidance, support, time and effort this thesis would have been impossible. **Pranjal Agarwal, Dhruv Kumar Gupta, Aman Gupta, Saptarishi Majumdar, Suransh Chopra, Vishwas Lathi, Anvit Mangal, Surabhi S Nath, Gaurav Jain, Shivang Chopra and Kshitij Jain** for all having been involved in long calls through the day and night attempting to understand medical imaging and teaching machine learning in return. I am particularly thankful to **Rohan Dhanakshirur** who also patiently helped with my coursework apart from the research in the thesis. My sincere gratitude also to **Hema Malhotra**, whose meticulous work and efficiency allowed us to formulate specific standard operating procedures for handling such large amounts of data*

*My heartfelt gratitude towards Professor **Sanjay Thulkar**, Head of Radiology in IRCH, who ensured all administrative hurdles were easily tackled. I am particularly indebted to my colleagues in my department in AIIMS Delhi, who adjusted their work schedules around mine, particularly **Dr Sanjay Thulkar, Dr Mukesh Yadav, Dr Chandrashekhara and Dr Ekta Dhamija**. I thank **Dr Smriti Hari, Dr Sanjay Thulkar** for always having encouraged and mentored me, and **Dr Arun Kumar Gupta and Dr Raju Sharma**, for making the path to this thesis possible.*

*I began the PhD course with a 1 year old child, and had another through the course of this work. I am most indebted to these 2 lovely children, **Maya and Krish**, for putting up with my schedule, while I balanced clinical work with the research in the course of this thesis. My heartfelt gratitude towards my husband, **Dr Mayank Bansal**, also a surgeon (and pursuing his own PhD) for always encouraging me through the thick and thin, and making this journey full of fun. But most of all, I am deeply indebted to my parents, **Sujatha Rangarajan and Rangarajan Raghavan**, who supported me in every way possible, by making space for my working schedule, relentlessly pushing the boundaries of what parents and grandparents can possibly achieve. No words can express my gratitude towards my family who have borne the brunt of the effects of long hours that have gone into this work*

# ABSTRACT

**KEYWORDS:** Deep Learning; Breast Cancer; Mammography

Deep Learning techniques have revolutionised computer vision and image processing. The applications of deep learning have been seen in everyday life, in social media, personalised advertisements, organisation of emails and photographs etc, but can potentially have far-reaching consequences in medicine. In oncology, a wide variety of applications of deep learning have been seen, ranging from detection of cancer, to prediction of response to therapy, and guiding management.

The practice of medicine is extremely complex. While over the last 5 years several deep learning neural networks have been trained and tested, very few have made it to clinical practice. A large part of the reason, is the complexity of the profession. There is thus a need for medical professionals to actively participate in all aspects of network design, not just in the data generation process.

Breast cancer is the leading cause of cancer related mortality in India. It is also a cancer which is amenable to screening for early detection. It has been proven through many randomised control trials with many decades of follow-up that early detection of cancer with an X-ray based modality called mammography can result in significant mortality benefit.

Most western countries implement population level screening programs with annual or biennial screening with mammography for all women above a certain age. The interpretation of mammography is an art and science that requires great training and practice, and also suffers from the possibility of misinterpretation due to human fatigue. Thus there has been a concerted effort towards development of networks to act as second readers, triage mammograms to ensure the most suspicious images are seen well, as well as for automated interpretation.

In India, due to paucity of resources, such population based screening is not possible. As a result, breast cancer is detected in later stages (Stage 3 or 4 in comparison to Stage 1 and 2 in the West), and therefore has a higher mortality and morbidity in comparison to the West. If computer vision techniques can help triaging mammograms, or perform an initial automated read, population screening may become a possibility.

This thesis has been written both- for a medical audience as well as engineers. Therefore we describe the basics of these fields extensively. We begin with describing the various appli-

cations of artificial intelligence in the field of medicine, we explain some basic concepts and terms related to machine learning and deep learning for the benefit of medical professionals, we perform a systematic review of literature on deep learning solutions available for mammography, and then describe our work, the networks we built and tested on Indian data for detection of breast cancer on mammograms. We then describe a generative network based tool we built for training of residents in interpretation of mammograms, and present results of a randomised controlled trial we performed to evaluate the same. Finally, we present details of a rapid reporting tool, which can be used for generating conventional radiology reports within a few clicks, we present results of our clinical study where we showed the benefit of use of such a tool in routine clinical practice.

## थीसिस सारांश (हिंदी)

डीप लर्निंग तकनीकों ने कंप्यूटर विज्ञान और इमेज प्रोसेसिंग में क्रांति ला दी है। डीप लर्निंग के अनुप्रयोगों को सोशल मीडिया, व्यक्तिगत विज्ञापन, ईमेल और फोटोग्राफ्स के संगठन जैसी दैनिक जीवन की गतिविधियों में देखा गया है, लेकिन यह चिकित्सा क्षेत्र में व्यापक प्रभाव डाल सकते हैं। ऑन्कोलॉजी में, डीप लर्निंग का उपयोग कैंसर का पता लगाने, उपचार के प्रति प्रतिक्रिया की भविष्यवाणी करने और प्रबंधन का मार्गदर्शन करने जैसे कई उद्देश्यों के लिए किया गया है।

चिकित्सा का अभ्यास अत्यंत जटिल है। पिछले 5 वर्षों में कई डीप लर्निंग न्यूरल नेटवर्क का प्रशिक्षण और परीक्षण किया गया है, लेकिन बहुत कम ही नैदानिक अभ्यास में उपयोग किए जा रहे हैं। इसका एक प्रमुख कारण इस पेशे की जटिलता है। इसलिए, यह आवश्यक है कि चिकित्सा पेशेवर केवल डेटा उत्पन्न करने की प्रक्रिया में ही नहीं, बल्कि नेटवर्क डिज़ाइन के सभी पहलुओं में सक्रिय रूप से भाग लें।

भारत में स्तन कैंसर कैंसर से संबंधित मृत्यु दर का प्रमुख कारण है। यह एक ऐसा कैंसर है जिसे शुरुआती पहचान के लिए स्क्रीनिंग के माध्यम से नियंत्रित किया जा सकता है। कई रैंडमाइज्ड कंट्रोल ट्रायल्स और दशकों के फॉलो-अप से यह साबित हुआ है कि एक्स-रे आधारित तकनीक, जिसे मैमोग्राफी कहते हैं, के माध्यम से कैंसर की प्रारंभिक पहचान से मृत्यु दर में महत्वपूर्ण कमी लाई जा सकती है।

अधिकांश पश्चिमी देशों में, एक निश्चित आयु से ऊपर की सभी महिलाओं के लिए वार्षिक या द्विवार्षिक मैमोग्राफी स्क्रीनिंग कार्यक्रम लागू किए जाते हैं। मैमोग्राफी की व्याख्या एक कला और विज्ञान है, जिसके लिए गहन प्रशिक्षण और अभ्यास की आवश्यकता होती है। यह मानव थकान के कारण गलत व्याख्या की संभावना से भी प्रभावित हो सकती है। इस कारण से, ऐसे नेटवर्क विकसित करने के लिए एक सघन प्रयास किया गया है जो दूसरे पाठक के रूप में काम करें, सबसे संदिग्ध छवियों को प्राथमिकता दें, और स्वचालित व्याख्या प्रदान करें।

भारत में संसाधनों की कमी के कारण, ऐसी जनसंख्या आधारित स्क्रीनिंग संभव नहीं है। इसके परिणामस्वरूप, स्तन कैंसर का पता अधिकतर उन्नत चरणों (चरण 3 या 4) में लगता है, जबकि पश्चिमी देशों में यह चरण 1 या 2 में ही पता चल जाता है। इस वजह से भारत में मृत्यु और बीमारी की दर पश्चिम की तुलना में अधिक है। यदि कंप्यूटर विज्ञान तकनीकें मैमोग्राफि को प्राथमिकता देने या प्रारंभिक स्वचालित पढ़ाई में मदद कर सकती हैं, तो जनसंख्या स्क्रीनिंग संभव हो सकती है।

यह शोध प्रबंध चिकित्सा पेशेवरों और इंजीनियरों, दोनों के लिए लिखा गया है। इसलिए, हम इन क्षेत्रों की बुनियादी बातें विस्तार से बताते हैं। हम चिकित्सा क्षेत्र में कृत्रिम बुद्धिमत्ता के विभिन्न अनुप्रयोगों का वर्णन करते हैं, चिकित्सा पेशेवरों के लाभ के लिए मशीन लर्निंग और डीप लर्निंग से संबंधित कुछ बुनियादी अवधारणाओं और शब्दों को समझाते हैं, मैमोग्राफी के लिए उपलब्ध डीप लर्निंग समाधानों पर साहित्य का व्यवस्थित समीक्षा करते हैं, और फिर हमारे कार्य, भारतीय डेटा पर स्तन कैंसर का पता लगाने के लिए विकसित और परीक्षण किए गए नेटवर्क का वर्णन करते हैं। इसके बाद, हम एक जनरेटिव नेटवर्क आधारित उपकरण का वर्णन करते हैं जिसे हमने मैमोग्राम की व्याख्या में निवासियों के प्रशिक्षण के लिए बनाया और परीक्षण किया है। अंत में, हम एक तेज़ रिपोर्टिंग टूल के विवरण प्रस्तुत करते हैं, जो कुछ ही क्लिक में पारंपरिक रेडियोलॉजी रिपोर्ट तैयार करने के लिए उपयोगी हो सकता है। हम अपने नैदानिक अध्ययन के परिणाम प्रस्तुत करते हैं, जिसमें हमने नियमित नैदानिक अभ्यास में इस टूल के उपयोग के लाभ दिखाए हैं।

# Contents

<b>ACKNOWLEDGEMENTS</b>	ii
<b>ABSTRACT</b>	iii
<b>LIST OF FIGURES</b>	xii
<b>LIST OF TABLES</b>	xiii
<b>ABBREVIATIONS</b>	xiv
<b>1 INTRODUCTION</b>	1
1.1 Breast Cancer in India . . . . .	1
1.2 Contributions of the Thesis . . . . .	2
<b>2 BACKGROUND</b>	5
2.1 Classification . . . . .	5
2.2 Detection . . . . .	7
2.3 Segmentation and quantification . . . . .	7
2.4 Image generation . . . . .	8
2.5 Natural language processing . . . . .	9
2.6 Predictive modelling, radiomics, and radiogenomics . . . . .	9
2.6.1 Genomic data analysis . . . . .	10
2.7 Supervised, Semi-supervised, Weakly supervised and Unsupervised Learning	12
2.8 Machine Learning: Essential concept . . . . .	13
2.9 Neural Networks and Deep Learning . . . . .	18
2.10 Convolutional Neural Network (CNN) . . . . .	20
2.11 Common terminology and their meaning . . . . .	21
2.11.1 Data related . . . . .	22
2.11.2 Model related . . . . .	22

---

2.11.3 Performance assessment	23
2.12 Challenges in evaluation of AI algorithms	25
2.12.1 Data related challenges	25
2.12.2 Explainability	25
<b>3 Principles of Breast Imaging</b>	<b>26</b>
3.1 Screening for Breast cancer	26
3.2 Evaluation of a patient with suspected breast cancer	27
<b>4 Review of Literature: Deep Learning in Breast Imaging</b>	<b>30</b>
4.1 Method of systematic review	32
4.1.1 Information sources	32
4.1.2 Eligibility criteria	33
4.1.3 Data extraction and quality assessment	33
4.1.4 Clinical Study Design	38
4.1.5 Performance of AI	38
4.2 Results	39
4.2.1 Literature search and study selection	39
4.2.2 Study Analysis	39
4.2.3 Analysis of generalisability	40
4.2.4 Clinical Study Design	50
4.2.5 Common practices used for AI model design and training:	52
4.2.6 Performance of AI	55
4.3 Discussion	55
4.3.1 Common practices for model design	55
4.3.2 Common practices for clinical design	56
4.3.3 Overall assessment of position of AI in breast imaging:	57
<b>5 Cancer Detection: Detection of small masses</b>	<b>58</b>
5.1 Detection of small cancers	59
5.1.1 Introduction	59
5.1.2 Materials and Methods	61
5.1.3 Results and Experiments	64

---

5.1.4 Discussion . . . . .	66
<b>6 Detection of isodense, obscure masses in dense breasts</b>	<b>68</b>
6.1 Introduction . . . . .	68
6.1.1 Methods . . . . .	69
6.1.2 Results . . . . .	75
6.1.3 Discussion . . . . .	79
<b>7 GAN based Simulation Training module for Resident Education</b>	<b>81</b>
7.1 Background . . . . .	81
7.2 Materials and Methods . . . . .	82
7.2.1 Training the Neural Network for generation of mammograms . . . . .	82
7.2.2 Development of User Interface (UI) for Simulation Tool . . . . .	87
7.2.3 Fidelity Testing of tool . . . . .	89
7.2.4 Test of Simulation tool in resident performance- Randomised Control Trial (RCT) . . . . .	89
7.2.5 Outcome measures and Statistical Analysis . . . . .	90
7.3 Results . . . . .	90
7.3.1 Fidelity testing of the generated mammograms . . . . .	91
7.4 Discussion . . . . .	93
<b>8 Tool for image annotation, rapid report generation and patient interac- tion</b>	<b>96</b>
8.1 Background . . . . .	96
8.2 Design of Reporting Tool . . . . .	98
8.3 Clinical Test of Tool: Study Design . . . . .	100
8.4 Clinical Test of Tool: Results . . . . .	102
8.4.1 Assessment of tool by reporting radiologist . . . . .	102
8.4.2 Assessment of reports by Consultant Radiologist . . . . .	103
8.4.3 Time Taken for reporting mammograms with reporting tool . . . . .	103
8.4.4 Assessment of Search-ability of reports with reporting tool . . . . .	103
<b>9 Conclusion</b>	<b>104</b>
9.1 Future Work . . . . .	105

9.1.1 Papers in Refereed Journals	118
9.1.2 Presentations in Conferences	118
9.1.3 Biodata	119

## List of Figures

1.1	Mammography is an Xray machine which produces low energy Xrays to image the breast . . . . .	2
1.2	Mammography involves compression of the breast, a 3 dimensional structure, into a 2 dimensional format for imaging. 2 Views are taken - CC and MLO. As shown in steps 5 and 6 in the figure, a true mass in the breast must be seen on both views, in order to be considered a mass, versus overlapping breast tissue which can masquerade as a mass in a dingle view . . . . .	3
1.3	From Left to Right- The first image shows a mammogram of a fatty breast has more "black" regions while the second image is a dense breast which has many "white" regions. The arrow in the third image shows a cancer which is well seen since it is in the fatty part of the breast, while the arrow in the fourth image shows a cancer which is harder to detect because it is seen hiding behind fibroglandular tissue. . . . .	3
2.1	Applications of AI in oncology . . . . .	6
2.2	By helping transform the workflow, AI can potentially free up the doctor's time to allow effective communication with the patient. This workflow shows how AI can potentially help generate a traditional radiology report, a report for the physician/ surgeon as well as a report in simple layman language for the patient, as well as allow the patient to interact with the radiologist, thus transforming a field which currently always functions behind closed curtains. . . . .	11
2.3	Essential difference between traditional software engineering and machine learning . . . . .	14
2.4	Training, validation and testing a machine learning model . . . . .	15
2.5	Optimization of a machine learning model. Assume the task is separation of different colored dots from each. The initial model misclassifies a large number of dots. The per- formance gradually improves over a few iterations, and reaches a point where only a small number are misclassified (Optimally trained). With further training, it achieves complete accuracy on the training data; this is the stage of "overfitting." When applied to the test data, the "overfitting" model performs worse than the model which was optimally trained. Created with Bio- Render.com. . . . .	16
2.6	Gradient Descent . . . . .	17
2.7	Similarity of the neural network with the biological neuron, where some input signals result in a pattern of activations, resulting in the output signal . . .	18

2.8	Many hidden layers within a deep neural network architecture	19
2.9	Convolutional Neural networks attempt to learn "filters" which represent essential features of an image	20
2.10	Saliency maps help in understanding where the network is focussing in order to make a decision	24
4.1	Method of selection of studies for systematic review	32
4.2	Risk of bias and applicability: QUADAS-2 checklist	33
5.1	Importance of resolution, scale and Image-context in detection of small cancers on mammography.	60
5.2	Proposed Architecture for detection of small cancers. The image is rescaled to 3 diferent scales, and fixed sized crops are systematically obtained to provide inputs with variable resolution, scale and context to the network. Predictions from all 3 scales are combined at test time.	63
5.3	Performance of the proposed network on the DM dataset (a), SM dataset (b) and the SC dataset (c). Results of baseline architecture are plotted for comparison.	64
5.4	Efect of resolution (a), scale (b) and context (c) on performance of network	65
5.5	Visualization of our results of our network on small masses<1 cm. Yellow represents the ground truth bounding box, green represents predictions of our network.	67
6.1	Description of various datasets used in the study. Number of dense breasts and cancers classified as isodense/ obscure is mentioned separately	69
6.2	Images thresholded just above the fat density	72
6.3	Preprocessing steps to obtain TI ans CABD images	73
6.4	Proposed bilateral network architecture. Note that "Right" and "Left" are for representation only	73
6.5	Histogram analysis of the training dataset. Histograms of ACR A (a) ACR B (b) ACR C (c ) and ACR D (d) are shown separately	77
6.6	FROC curves on the DM dataset (a) shows performance of our proposed network on the entire dataset (b) shows performance on dense breasts (ACR C and D) and (c ) shows performance on cancers classified as isodense and obscure. Baseline network (FRCNN) performance is plotted for comparison	77
6.7	FROC curves on the SM dataset (a) shows performance of our proposed network on the entire dataset (b) shows performance on dense breasts (ACR C and D) and (c ) shows performance on cancers classified as isodense and obscure. Baseline network (FRCNN) performance is plotted for comparison	78

7.1 Schematic diagram representing training of the ALAE frame-work, and using it for generation of mammograms with specific characteristics for simulation training (created with biorender.com)	83
7.2 Training of E,D and G for mammogram data	84
7.3 PCA over latent vectors at train time to obtain the principle components to represent the most important mammographic features	85
7.4 At Inference time for a given image, the latent vector is represented by its principle components. The alteration of coefficients of these components then result in a new latent vector which is fed to the generator to give an image, which differs from the original image in a specific mammographic character represented by the chosen principle component.	86
7.5 Row 1 shows original images, row b-e show images generated by other methods (VAE-Variational AutoEncoders)	86
7.6 Figures a through e show generated mammograms with the same cancer blended into different regions of the image, using Poisson Blend technique	87
7.7 Steps of simulation training (figure created with biorender.com)	88
7.8 Recruitment of patients for RCT	90
7.9 Fidelity testing of mammogram images	92
7.10 Comparison of differences in mean values of various calculated parameters between Test 1 and Test 2 in the practice group and non-practice group	93
8.1 An open source application "Labellmg" was used to build the tool. This Figure shows a DICOM image of a mammogram opened on this tool, and demonstrates the inbuilt features of the tool	98
8.2 PACS server can be selected, Cases can be Queried and retrieved into Database. Patients in the local database can be searched with labels such as patient name, id etc	99
8.3 All 4 images of a patient are opened together. Labels for right breast and left breast can be selected in the panel on the right through drop down menus	99
8.4 A bounding box can be drawn around an abnormality and characteristics of the lesion selected on drop down menus	100
8.5 A free-text report is generated automatically based on options selected in the drop down menus	101
8.6 The tool allows searching reports/ images based on labels such as "Impression", "BIRADS category", " Mass", "Architectural Distortion", "Microcalcification" etc	101

## List of Tables

3.1 Breast Imaging Reporting and Data System Lexicon . . . . .	28
5.1 Selection of a fully convolutional network as baseline for our proposed network. FPI refers to false positives per image. Sensitivity of network at specific FPI values have been represented in the table. . . . .	64
5.2 Sensitivity on different datasets (sensitivity of proposed/ baseline has been given) . . . . .	65
5.3 Analysis of detection performance on individual scales prior to WBF . . . . .	66
6.1 Description of various datasets used in the study. Number of dense breasts and cancers classified as isodense/ obscure is mentioned separately . . . . .	75
6.2 Performance of state of art Object detection networks on our dataset,along with a comparison with their performance in dense breasts and isodense/ obscure masses. Performance of our proposed network has also been shown for comparison . . . . .	76
6.3 Performance of each individual technique- thresholding, contrast stretching and bilateral analysis on our isodense/ obscure cancer dataset . . . . .	78
6.4 Performance of the proposed network on the publicly available INBreast dataset. Performance of other investigators who have presented their results on this dataset has also been shown. . . . .	80
7.1 Number of Junior and Senior residents in both groups . . . . .	91
7.2 Caption . . . . .	91
7.3 Comparison of median values of various calculated parameters between test 1 and test 2 . . . . .	92
7.4 P values for testing significance in change in performance . . . . .	93
7.5 Summary of similar studies on simulation in diagnostic radiology . . . . .	95

## ABBREVIATIONS

<b>IITD</b>	Indian Institute of Technology, Delhi
<b>AIIMS</b>	All India Institute of Medical Sciences New Delhi
<b>CC</b>	Cranio Caudal view
<b>MLO</b>	Medio Lateral Oblique View
<b>CNN</b>	Convolutional Neural Network
<b>GAN</b>	Generative Adversarial Network
<b>ML</b>	Machine Learning
<b>DL</b>	Deep Learning