Breast Cancer Detection with Thermal Images and using Deep Learning

Amit Sarode[1†] and Vibha Bora[2††],

[1,2]G H Raisoni College of Engineering, Nagpur, India

Abstract
According to most experts and health workers, a living creature's body heat is little understood and crucial in the identification of disorders. Doctors in ancient medicine used wet mud or slurry clay to heal patients. When either of these progressed throughout the body, the area that dried up first was called the infected part. Today, thermal cameras that generate images with electromagnetic frequencies can be used to accomplish this. Thermography can detect swelling and clot areas that predict cancer without the need for harmful radiation and irritational touch. It has a significant benefit in medical testing because it can be utilized before any observable symptoms appear. In this work, machine learning (ML) is defined as statistical approaches that enable software systems to learn from data without having to be explicitly coded. By taking note of these heat scans of breasts and pinpointing suspected places where a doctor needs to conduct additional investigation, ML can assist in this endeavor. Thermal imaging is a more cost-effective alternative to other approaches that require specialized equipment, allowing machines to deliver a more convenient and effective approach to doctors.

Keywords:
Breast Cancer, Thermal images, Deep Learning, VGG16 and Inception V3.

1. Introduction

The frequently diagnosed cancer in women is breast cancer due to lack of awerness and its mortlity is highest compared to any other cancer [7]. These tumours eventually grow into a malignant lumpy tumour. According to the World Health Organization (WHO), India has at least 10 lakh breast cancer patients each year. Of the 10 lakh cases reported, five lakh people died, and by 2025, this number might be five times higher, with a 23 percent increase in women and a 19 percent increase in men [8]. According to the World Health Organization, it also identifies breast cancer at a last stage in a large number of females, reducing chances of survival of patitents [9]. The breast cancer is widely studied in both developed and developing countries. Breast cancer survival rates differ greatly over the world, ranging from 80% to roughly 60% in high-income countries to fewer than 40% in emerging and low-income countries. Not only breast cancer is assumed to have only happened in rich countries, it is worrying that low-income and developed countries account for roughly 50% of cases and 58 percent of deaths[10]. Early detection of symptoms lowers treatment expenses significantly. The majority of the people, on the other hand, will not be cautious until more serious symptoms appear. Exams are avoided because of the expense and complexity of the screening, the dread of agony and aggravation, and fact that most examinations are lengthy. It identifies excessive density regions in the breasts using x-rays.

. Breast cancer screening should be done every one to two years for women around 45. Sadly, the number of mistaken positives in mammography is 49.1% every ten years. If a tumour is identified early on, the expense of therapy can be greatly decreased [3]. Regardless proper screening for breast cancer incidence, cost of screening inhibits women in developing nations from carrying out regular examinations. In differentiation, the intrusive nature of these exams sometimes contributes to the misconception that mammograms are not accepted unless a doctor precribes them. However, MRI has its own limitations. As a result, other diagnostic approaches such as thermography must be studied by researchers and organisations. Because breast thermography is inexpensive, it can compensate for the shortcomings of these diagnostic imaging modalities, has few side effects, is radiation-free, and can detect cancers before they become palpable. CNN has advanced rapidly in the field of medical imaging as a result of advances in computer vision and deep learning models. Clinical observations and CAD systems are frequently utilised to diagnose many diseases. The most extensively used field of CAD today is breast cancer screening using thermal imaging. The research is primarily focused on enhancing tumour detection and classification accuracy. After the clay was applied, the section that dried out the fastest

Manuscript received August 5, 2023
Manuscript revised August 20, 2023
https://doi.org/10.22937/IJCSNS.2023.23.8.11

was designated the afflicted part. A lesser-known but useful characteristic is temperature as a proxy for diagnosing probable ailments. Using contemporary technology, such as thermography, which allows for the visualisation of infrared diseased part in order to map an body's temperature. These scanners produce scan temperature patterns a wide such as lymph enlargement and inflammatory regions, can be detected considerably earlier than the conventional indicators and without any contact or radiation. The usage of thermal components can be a great approach to identify potential reasons before hazardous symptoms appear. It is substantially less expensive than a mammogram because it simply requires an thermal camera to capture photos. As a result, numerous researchers have been doing research on thermography for breast cancer scanning by utilising machine learning for the past two decades. Some researchers focus on the location and size of vanishing tumours as well as simulation models, studying aspects breast segmentation, menstruation & acquisition techniques. Order to diagnose breast cancer, researchers have turned to deep learning in recent years. Deep learning accurately identifies photos by abstracting the high dimensional representation of images through convolution and other techniques. In any instance, the key to designing the model is to identify the obvious aspects of the image that can be seen with the naked eye. Several deep learning models for medical image applications have recently been created and implemented [7]. CNN is the most famous deep learning model of image classification because of its rapid and less tedious implementation, ability to extract less amounts of information from big amounts of data, ability to learn instinctively significant feature ranking to raw data given for training to it, and many other advantages. CNNs are highly useful for detecting the most essential or obvious area of the images features identification. Researchers had adopted learning algorithms known as CNN which shows good results compared to traditional strategies to rely on tactical salient feature building [15,16]. Many researchers have successfully employed CNNs as breast cancer detection on various imaging modalities in recent decades, Unluckily, only a handful of them are employed CNN in the past for cancer detection with thermographic. The primary cause for this could be CNN key flaws of radiology: the lack in a big dataset of thermography & the problem of overfitting. Nonetheless, in last 2 years, CNN have become the utmost popular model for dataset and image recognition since it can dramatically enhance identification rates while also lowering the requirements for original picture quality and the amount of training samples. As a result, we are creating a CNN based model for cancer detection using a database of infrared pictures in this study [8]. For performance evaluation, we compare VGG16, InceptionV3 and Sequential models, which are pre trained CNN models**.** In addition, for screening of breast cancer datasets, an examination of multiple factors such as complexity, augmentation, and performance tuning of parameter in CNN is undertaken to improve the performance of our CNN model.

2. Methodology

Jupyter Notebooks The research combines a largely used diagnostic technique with a deep learning system to assist researcher in developing model which can discriminate between healthy & sick persons. Python is one of the most common language for machine learning projects. This project will make use of the Python programming language and the Jupyter Notebook, an experimental notebook. Jupyter Notebooks can be used alone or in conjunction with the Anaconda installation, which is a full open-source software distribution intended specifically for science computing. Python predefined libraries are useful in developing deep learning models. Python is one the most widely used programming language. Working with the picture collection can be challenging. Widely used packages in computational programming is Numpy. Pandas is a free and open-source framework for information structuring & analysis. Matplotlib and Open Source CV that allows users to construct 2 D graphs & diagrams such as among other things. These algorithms recognises a wide range of objects and are very efficient. Keras, is used in this project.

VGG16

VGG16 is CNN pre trained model that has been pre-trained to recognise huge images. In ImageNet, covers a collection over million images, tests demonstrate 92.7 percent precision. Many 3x3 filters one after the other in VGG16 are used in place of Large filters (Fig. 3) compared to the AlexNet network. NVIDIA

TITAN BLACK video cards were used to train the VGG16 n/w for the competition for several months. InceptionV3 model InceptionV3 is a Google-developed version of the effective CNN concept. Placing it in the top 5 in the ILSVRC-2015. Figure 3 depicts the topology of the neural network InceptionV3. The core concepts of developing the architecture were developed during designing InceptionV3.

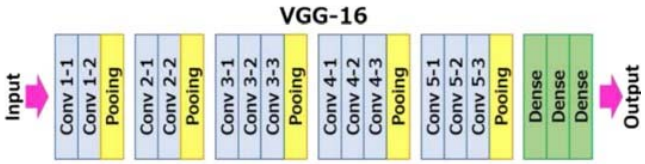

Fig.3 VGG16 Model

InceptionV3 model

InceptionV3 is a Google-developed version of the effective CNN concept. Placing it in the top 5 in the ILSVRC-2015 [51]. Figure 4 depicts the topology of the neural network InceptionV3. The core concepts of developing the architecture were developed during designing InceptionV3.

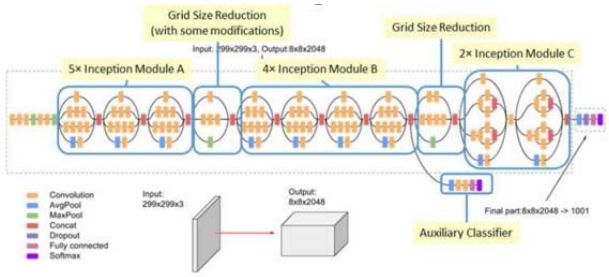

Fig.2 Inception V3 Model

3. Dataset

Dataset Infrared imaging has long been used to detect breast cancer as a screening approach. The thermal pictures used in this study were obtained from PROENG, an online project database.

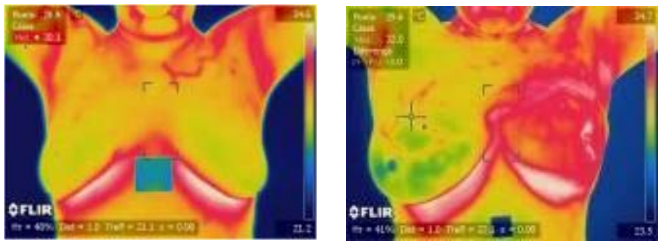

Fig.3 Healthy & Sick Patient [6]

Dataset structure

The model has 2039 photos in total, however owing to similarities in thermal images, a few of them must be eliminated to avoid model training complications.

Table 1 Classification of data

	Training	Testing	Validation
Healthy	714	248	55
Sick	712	255	55

Results and Discussions

We can observe that the augment1 algorithm performs better for the InceptionV3 model than the augment2 algorithm, so we'll use augment1 as the default model in the following tests. The classification model accuracy results for VGG16 and InceptionV3 are shown in Table 2. As can be seen, InceptionV3 with model3 has greater accuracy results than VGG16. As a result, in the next experiments, we fine-tuned model4 and InceptionV3.

Table 2 Result of Fine-tuning hyperparameter

Ex. No.	Model	Learning rate	Decay	# augment Training Images	Drop rate1	Drop Rate1	Batch	InceptionV3
1	model1	0.001	0.8	630	0.7	NA	NO	89.3
2	by keeping	0.001	0.5	15000	0.6	NA	NO	90.5

3	last convol ayer	0.0009	0.7	9000	0.6	NA	NO	91.3
4	unfroze n	0.0005	0.5	630	0.8	0.6	YES	93.4
5		0.0005	0.8	20000	0.6	NA	YES	94.3
6		0.005	0.8	8000	0.5	0.6	NO	97.4
1	model4	0.001	0.7	630	0.7	0.5	NO	87.9
2	by keepin g	0.001	0.5	630	0.6	0.7	NO	89.3
3	last convla yer	0.0009	0.7	9000	0.6	0.5	YES	92.4
4	unfroze n	0.0005	0.7	20000	0.7	NA	YES	94.3

Table 3 Model Accuracy.

Model	VGG 16	InceptionV3
1	58.8	71.3
2	73.2	77.9
3	84.9	89.6
4	90.9	95.6

References

[1] Hemalatha K, VVetriselvi. “Deep Learning based Classification of Cervical Cancer using Transfer Learning’, 2022International Conference on Electronic Systems and Intelligent Computing (ICESIC), 2022

[2] Agravat RR, Raval MS (2019) Prediction of overall survival of brain tumor patients. In: TENCON 2019- 2019 IEEE Region10 Conference (TENCON). IEEE, pp 31–35

[3] Cano F, Madabhushi A, Cruz-Roa A (2018) A comparative analysis of sensitivity of convolutional neural networks for histopathology image classification in breast cancer. In: Romero E, Lepore N, Brieva J (eds) 14th international symposium on medical information processing and analysis. International society for optics and photonics, vol 10975. SPIE, pp 277–284. [Online]. Available: https://doi.org/10.1117/12.2511647

[4] Erdogan H (2019) Trajectory Analysis and control using recurrent neural networks.

[5] Samir S. Yadav, Shivajirao M. Jadhav. “Thermal infrared imaging based breast cancer diagnosis using machine learning techniques”, Multimedia Tools and Applications, 2020

[6] Elias T (2019) Visual lab - a methodology for breast disease computer-aided diagnosis. [online] visual. ic.uff.br. http://visual.ic.uff.br/en/proeng/thiagoelias/

[7] Lunenfeld B, Stratton P (2013) The clinical consequences of an ageing world and preventive strategies. Best Pract Res Clin Obstet Gynaecol 27(5):643–659

[8] de Oliveira NPD, dos Santos Siqueira CA, de Lima KYN, de Camargo Cancela M, de Souza DLB (2020) Association of cervical and breast cancer mortality with socioeconomic indicators and availability of health services, vol 64

[9] W.H. Organization (2019) Global action plan on physical activity 2018-2030: more active people for a healthier world. World Health Organization, 2019

[10] Acharya UR, Ng EY-K, Tan J-H, Sree SV (2012) Thermography based breast cancer detection using texture features and support vector machine. J Med Syst 36(3):1503– 1510

[11] Schaefer G, Zavi ´ sek M, Nakashima T (2009) Thermography based breast cancer analysis using statistical ˇ features and fuzzy classification. Pattern Recogn 42(6):1133–1137

[12] Etehadtavakol M, Emrani Z, Ng E (2019) Rapid extraction of the hottest or coldest regions of medical thermographic images. Med Biol Eng Comput 57(2):379–388.

[13] Qi H, Head JF (2001) Asymmetry analysis using automatic segmentation and classification for breast cancer detection in thermograms. In: 2001 conference proceedings of the 23rd annual international conference of the IEEE engineering in medicine and biology society, vol 3. IEEE, pp 2866–2869.

[14] Gerasimova E, Audit B, Roux SG, Khalil A, Gileva O, Argoul F, Naimark O, Arneodo A (2014) Waveletbased multifractal

[15] Fernandez-Ovies FJ, Alf ´ erez-Baquero ES, de Andr ´ es-Galiana EJ, Cernea A, Fern ´ andez-Mu ´ niz Z, ˜ Fernandez-mart´ ´ınez JL (2019) Detection of breast cancer using infrared thermography and deep neural networks. In: International work-conference on bioinformatics and biomedical engineering. Springer, pp 514–523

[16] Mambou SJ, Maresova P, Krejcar O, Selamat A, Kuca K (2018) Breast cancer detection using infrared thermal imaging and a deep learning model. Sensors 18(9):2799

Amit Sarode done Bachelor of Engineering in Electronic and Telecommunication from Bharati Vidyapeeth’s College of Engineering, Pune in 2020. I am Currently Masters Student at G H Raisoni College of Engineering, Nagpur.

Email: amit.sarode.vlsi@ghrce.raisoni.net

Dr.Vibha Rajesh Bora done his Bachelor of Engineering in Electronic Communication Engineering from Government College of Engineering, Amravati in 1996, Master of Technology in Electronics from Yeshwantrao Chavan College of Engineering, Nagpur in 2016. She is presently working as Professor & Incharge of Biomedical Engineering & Technology incubation Center, G H Raisoni College of Engineering, Nagpur (BETiC-GHRCE) which is in collaboration with BETiC-IIT Bombay. Currently she of chairman of Interdisciplinary Board of Studies. She is member of IETE, ISTE.

Email: vibha.bora@raisoni.net