Breast Cancer Detection using MobileNetV2 and Inceptionv3 Deep Learning Techniques

Aastha Joshi[1]
Research Scholar, S.AT.I (D)
Vidisha, Madhya Pradesh, India

Nirmal Gaud[2]
Assistant Professor, S.AT.I (D)
Vidisha, Madhya Pradesh, India

Abstract—**Transfer learning is a key component of the research on restorative images, but it can be difficult to find high-quality training datasets for machine learning methods. Despite various research studies in order, there have not been many evaluation publications about the application of transfer learning in restorative image analysis to this point. Furthermore, no research has been done on the potential applications of transfer learning to ultrasonic breast imaging. This research reviews recent methodologies and highlights their advantages and disadvantages by reviewing earlier works on transfer learning-based breast cancer detection using ultrasound pictures. The future avenues of research for the use of transfer learning in ultrasonic imaging for breast diagnostic and therapeutic applications are also highlighted in this work. Deep learning methods MobileNetV2 and InceptionV3 have been applied to the detection of breast cancer. MobilenetV2 utilizes depth-wise separable convolution, whilst InceptionV3 uses regular convolution, which is the main contrast between the two. Therefore, MobileNetV2 is less parameterized than InceptionV3. However, there is a slight efficiency hit as a byproduct. We compared the results and assessed the accuracy of both models.**

Keywords: Transfer Learning, Convolutional Neural Networks, InceptionV3, MobileNetV2, Deep Learning, and Breast Cancer Detection

I. INTRODUCTION

Currently, breast cancer is the leading cause of mortality of women, causing the deaths of 12.5% of all females worldwide, regardless of their socioeconomic background [1]. Early detection of breast cancer is crucial, according to research in the past, since it can reduce mortality rates by up to 40% [2, 3]. In the show, ultrasound imaging has become a popular imaging technique for detecting breast cancer, particularly in young women with dense breasts [4]. In order to successfully extract tissue features, ultrasonic (US) imaging is frequently performed [5-7]. According to studies, utilising a range of modalities, including US imaging, decreased the incorrect negative acceptance rate in other breast symptomatic techniques, like biopsy and mammography (MG) [2]. Using ultrasonic imaging techniques to diagnose breast cancer can increase tumor finding by up to 17% [6]. Additionally, it is possible to reduce the number of unnecessary biopsies by around 40%, which would reduce the number of medications consumed [5]. Utilizing non-ionizing radiation, which has no adverse health consequences and just needs basic equipment, seems to be another advantage of ultrasonic imaging [7]. As a result, ultrasound scanners are more affordable and flexible than mammography machines [5-8]. However,

mammography and histological analyses are not a stand-alone approach for the detection of breast cancer [6,7], and they are integrated with ultrasound frameworks to demonstrate the results [8,]. Many studies have made use of modern technology to increase the ultrasonic imaging's demonstration power [9].

Machine learning, counting false positive rates, failing to recognise changes brought on by illness, a diminished appropriateness for treatment checking, and subjectivity have all been found to be solutions to many of the problems with ultrasound's categorization, discovery, and division of breast cancer [10–12]. Numerous machine learning methods, however, actually perform better when a particular presumption is true, i.e., when the training and test data sets come from the same feature space and have the same distribution. When the distribution changes, the bulk of the numerical values within the models must be completely rebuilt using fresh data that was gathered beforehand. [11–13].

It is challenging to push for the load demand information and to construct models in this way for therapeutic applications, such as breast ultrasound imaging [14].As a result, it is wise to lessen the work and requirements required to accelerate the processing of information [13,14].Details transmission from one task to the next would be ideal in these situations [15]. Transfer learning permits the use of a previously produced application on some other system as the learning target.[16]. In a nutshell, it takes less time and effort to acquire and organise learning resources [10–16].

The foundation of transfer learning is the notion that, in exceptional situations, previously learned knowledge may be connected to current challenges to address them more quickly and effectively [17,18]. Transfer learning ultimately requires tested machine learning techniques that can retain and make use of previously acquired data [19–21]. Exchange learning has most recently been connected to breast cancer imaging with the development of a few convolutional neural network models to tackle visual categorization tasks in image pairs published on shared picture datasets like ImageNet [22].

The WHO Worldwide Breast Cancer Activity (GBCI) aims to prevent 2.5 million breast cancer fatalities between 2020 and 2040 by reducing the yearly global breast cancer mortality rate by 2.5%.In order to prevent the disease from

coming back, every breast treatment of cancer should work to get rid of as much cancer from your body as feasible. We specifically looked for the following things to find:

- Obtain credit for breast cancer ultrasound images.
- Deploy the MobileNetV2 and InceptionV3 deep learning algorithms to it.
- Examine the outcomes.
- Find out which deep learning algorithms work best for identifying and categorizing breast cancer. (None, 512)
- Analyze the major incidents that led to breast cancer.
- Create a model that can filter photos according to the type of breast cancer.

The following describes the structure of this paper: In section I, we discussed the aim and objective of breast cancer detection. In section II, we did a literacy survey for the work. In section III, we have mentioned the methodology used. IV mentions the model development. Results analysis is provided in Section V. Section VI concludes the essay lastly.

II. LITERATURE REVIEW

Huynh et al. examined the effectiveness of incorporating features from pre-trained deep CNNs for computer-aided diagnosis (CADx) in their 2016 study on the primary usage of exchange learning in breast cancer imaging [23]. Then after, Byra et al. released a paper in which they proposed using neural exchange learning and ultrasonography to classify breast injury[24].The study by Yap et al., which supported the use of sophisticated neural learning algorithms for breast cancer diagnosis, was published right away [25]. Utilizing a pre-trained fully convolutional network called AlexNet, they looked into three alternative strategies: a patch-based LeNet approach, a U-Net demonstration, and an exchange learning technique. A few distributions on the usage of exchange learning for breast ultrasound imaging have been developed based on the distribution of these properties [26–29]. The aim of this work is to review papers on breast cancer imaging that provide contemporary methodologies and identify their benefits and drawbacks using exchange learning. Additionally, it provides immediate access to upcoming possibilities for ultrasound breast cancer imaging learning and sharing. The audit will be crucial in helping researchers discover emerging tactics that will advance as well as areas that seem to profit from more exchange learning-based ultrasound breast imaging study.

III. METHODOLOGY AND DATA

A. Data Source
- The dataset, which includes 693 ultrasound images of breast cancer from three distinct classes—benign, malignant, and normal—was collected from Kaggle.

- Jupyter Notebook as well as the Python libraries Numpy, Matplotlib, Pandas, Seaborn, Sklearn,

Tensorflow, and Keras are needed to create Python 3 programmes.

B. Feature Extraction
- The CNN model used in the feature extraction process has been successfully trained on a sizable dataset like ImageNet, making it an excellent extractor for the underserved target space, such as breast ultrasound imaging [72]. The well-trained CNN model resolves all of the convolutional layers, cleaning up the entirely associated layers in particular [31–39]. For unrelated applications, including diagnosing breast cancer, the convolution layers are used as a stable highlight extractor [41–45]. The classifier that can generate fully linked layers is then given the derived features [45]. Finally, though not entirely, the present classifier is built [51–53]. Instead, it is in what can be called its incipient stage.

- The feature extractor and fine-tuning are two transfer learning techniques that have been identified. They have the benefit of not requiring neural network preparation, which necessitates coordination of the eliminated characteristics when using image analysis forms at this time[72]. These two methods are popular and frequently used [73, 74, 75, 76]. To identify the strategy that yields the most notable results, numerous academics have carried out in-depth analyses. The three learning approaches suggested in include a CNN design created from scratch, a transfer learning technique using ultrasound images and an improved VGG16 CNN architecture, and a fine-tuned learning approach using deep learning parameters [24].

C. Benign Tumors
- The term "benign tumour" refers to a tumour that stays put and doesn't spread to other parts of the body. They don't spread to nearby structures or to far-off regions of the body. Benign tumours frequently have clear borders and grow slowly.

- The majority of benign tumours are harmless. However, they could grow big and put strain on other structures, resulting in discomfort or other health issues.As instance, a sizable benign lung tumour could block the trachea and make breathing difficult. It would require immediate surgical removal. After removal, benign tumours are unlikely to reoccur. Uterine fibroids and cutaneous lipomas are two examples of benign tumours.

- Some benign tumours have the capacity to develop into cancer. These are frequently observed and could need to be surgically removed. For illustration, colon polyps, another term for an abnormal clump of cells, are frequently surgically

removed since they have the potential to progress into cancer.

D. *Malignant Tumors*

- Some benign tumours have the capacity to develop into cancer. These are frequently observed and could need to be surgically removed. For instance, colon polyps, another term for an abnormal clump of cells, are frequently surgically removed since they have the potential to progress into cancer.

- Unchecked local or global dispersion of malignant tumour cell populations. Malignant tumours are among those that pose a threat (ie, they attack other destinations). They spread via the lymphatic or blood systems to distant locations. Metastasis is the medical term for this dispersion. The liver, lungs, brain, and bone are the most often affected organs by metastasis, while it can happen everywhere in the body.

FIGURE I: Ultrasound Images
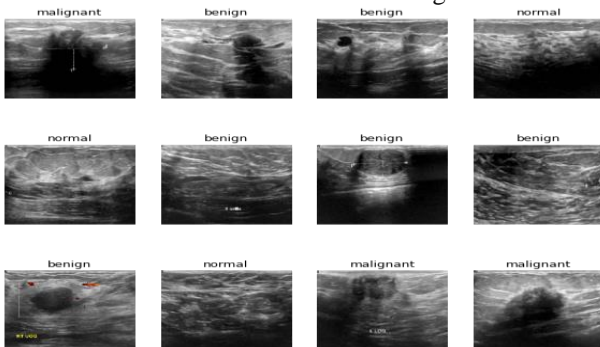
E. *Data Augmentation*

As fresh tagged data improves the CNN model's performance, the dataset has expanded. Images can be altered using techniques including rotation, translation, scaling, and layout flipping in the augmented process for visual data. The data augmentation of our dataset led to the following outcomes:

FIGURE II: Augmented Images
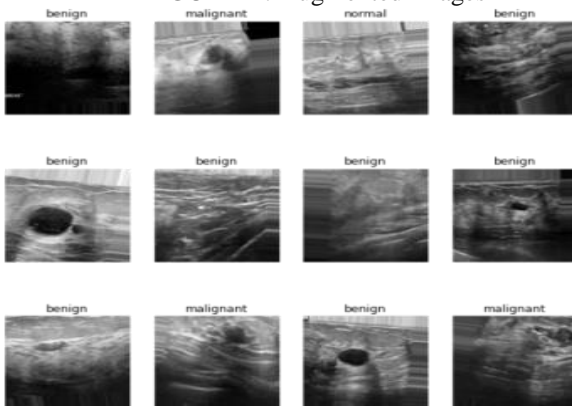
F. *Transfer Learning*

A common method for developing machine learning models without having to be concerned regarding how much information is presently accessible is transfer learning[30].Transfer learning may be able to assist with this issue since creating a thorough example may require some knowledge and computer abilities. A pre-made presentation may commonly be applied to multiple problems via exchange learning [31].For instance, a system that was created to perform one thing, like recognise various cell kinds, may be improved to do another, like classify malignancies. Transfer learning could be a helpful strategy for homework in computer vision. According to transfer learning theories, [31–33] abilities gained from using enormous image datasets like ImageNet are very adaptable to various picture identification tasks.

To move data from one show to another, there are two options. The method that is most usually employed is to insert a randomly initialized layer over the last one of the already-produced show [34]. The top-layer parameters, so to speak, are then prepared for the contemporary work while all other parameters are tidied up without being modified. The settled parcel operates as a highlight extractor, therefore this strategy might be viewed as the application of the swapping show as one. In the interim, the top layer maintains its normal, completely linked neuronal organizing function, making no special judgments about the input[34,35].

This kind of exchange learning might be the sole method for constructing a model without overfitting when there isn't a lot of data available [36]. Due to the reduced chance of overfitting when there are fewer parameters, this happens frequently. It is possible to unfreeze exchanged parameters and restructure the entire system [34–37] when actual information is widely available for scheduling, which is unusual in therapeutic situations. In this case, the characteristic values assigned to the parameters are switched [37]. Instead of just initialising the weights arbitrarily, using a pre-trained demonstration can give the new display a great start and speed up merging and fine-tuning.

It is customary to cut back the learning rate by a significant amount in terms of maintaining the baseline from pretraining [38, 39]. Using apparently frozen parameters, it is common practice to construct pretty randomly starting layers until they combine [40–44], restart all settings before adjusting the general configuration (Figure 1). When there is a sizable amount of data for one activity and only a passing familiarity with another that is related, or when a show has already been produced on such information, exchange learning is quite helpful [45]. Even when the destinations are disconnected and there is sufficient data to build a showcase from scratch, using a pre-trained show to establish the parameters is still preferable to erratic activation [46].

The key benefits of transfer learning include faster setup times, enhanced neural network performance,

and reduced input needs. [47–50]. Regardless of the exact job that a neural arrangement has been trained on, the fundamental layer parameters of neural systems that have been trained on a significant number of images remain similar [16, 47]. The most efficient CNN layers for recurrence frequently recognise edges, surfaces, and designs [31], and these layers catch the elements that are typically helpful for comprehending the distinctive photos [47]. Edge, corner, form, surface, and other feature-specific highlights can be thought of as universal highlight extractors and employed in a variety of contexts [30–33].

The layers tend to recall progressively exact features as harvest time approaches [48–50]. For that classification task, the final layer of a structure that has been prepped for classifying may be especially intriguing [49].One unit would have responded as it did to photographs of a specific tumour if the show had been instructed to exhibit abnormalities [23–28]. The most frequent case of transfer learning is when all layers excluding the topmost layer are swapped [17–20]. The first n levels of a pre-trained example may frequently be switched to a goal arrangement, with the remaining layers being initialized randomly [51].

The outermost layers may also undergo swaps [33].Contemplate a tumour recognition proof show that was created on grayscale photos in order to develop a tumour acknowledgment display that accepts inputs from both coloured and grayscale photos.It may not be possible to properly prepare an unused display due to a lack of informational points; in this situation, it may be more beneficial to replace the latter layers and re-learn the previous ones [52,53]. Exchange learning is thus extremely beneficial when a large quantity of prior knowledge may be applied to the target problem and a neural network cannot handle a new domain [47–53].

Information annotators are often highly helpful for managing medical images because they generally provide massive tagged datasets [24–29]. Furthermore, preparation time is limited because it could take less time to fully construct a modern deep neural network in the case of challenging objective problems [48,49].

Analysts working in the data-restricted field of restorative imaging now have the ability to understand the issue of limited test datasets and progress execution [13]. Exchange learning can be categorised as cross-domain or cross-modal if the source and target information come from the same field [54,55].

In the field of restorative ultrasound imaging study, cross-domain transfer learning may be an effective strategy for completing a range of tasks [9]. In machine learning, large sample datasets are frequently employed to pretrain models, and an abundance of training data guarantees good result; conversely, in the realm of restorative imaging, this is generally not the case [56]. When particularly in comparison to exchange learning from a neural network that has been pre-trained with considerable preparedness tests in another area, such as the generic picture database of

ImageNet, domain-specific models created from scratch can perform significantly better in the case of small preparation tests [57-60].

One of the reasons for this is that, before the progress hypothesis test can be performed, it is frequently challenging and requires a substantial amount of preparation to transform the pre-trained case from the raw image to the joint vectors required for a particular action, such as classification in the context of medicine [58–60]. Furthermore, for the quick-planning datasets frequently utilised in medical imaging, a highly well-organized little arrangement will be helpful [13,58,59]. In restorative scenarios, particularly in breast imaging, many imaging modalities, including magnetic resonance imaging (MRI), mammography, computed tomography, and ultrasound (US), are frequently used [63–65].

Mammography (i.e., X-rays) and ultrasound are the main diagnostic techniques for finding breast cancer, and both need far less processing than MRI and CT [66-68]. Building datasets and performing ground-truth interpretations are challenging with breast MRI due to its increased duration, cost, and frequent use to evaluate high-risk patients [29]. The best strategy in these circumstances is cross-modal transfer learning [69, 70]. A few experiments [29] have demonstrated the benefit of cross-modal transfer learning over cross-domain transfer learning for a particular task and in the case of smaller training datasets. The two methods of feature extraction and fine tuning are frequently used for transfer learning [71].

IV. MODEL DEVELOPMENT

The creation of models for the detection of breast cancer is the subject of this section. We'll create MobilenetV2 first, then InceptionV3, and finally.

A. *MobileNetV2*

MobilenetV2 is being utilized because of its compact design. It uses depth-aware independent convolutions, i.e., each color channel is subjected to a distinct convolution as opposed to being put and together smoothed. Each input route for MobileNets receives a single channel thanks to creative design. After pointwise convolution, an 11 convolution is used to merge the results of the depth-aware convolution. In a conventional convolution, the inputs are filtered and merged in one step to produce the outputs. This might be split into two layers using depth-aware differentiating convolution: a layer for merging and a layer for filtering. That both calculations and the display constraints are significantly reduced by this factorization.

A popular tool for classifying photos is Imagenet. Millions of smartphone photos are used each year in a competition with 1000 categories. Comparisons are made between the models' efficacy in performing imagenet categorization tasks. As a result, it offers an assessment of a photo classification model's effectiveness that is

Published by :

http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
**ISSN: 2278-0181**
Vol. 11 Issue 09, September 2022

"standard." Many different exchange learning demonstration models employ imagenet weights. You can add extra layers to your example to make it more applicable to your application if you're using exchange learning. Using the ImageNet weights is not needed, but it typically offers advantages because it expedites show joining.

Table I below shows the layered architecture of the transfer learning model based on MobileNetV2.

TABLE I: Model Architecture based on MobileNetV2

Layer (Type)	Output Shape	Activation Function	Param #
Mobilenet_1.00_224 (Functional)	(None, 7, 7, 1024)		3228864
flatten (Flatten)	(None, 50176)		0
flatten_1 (Flatten)	(None, 50176)		0
batch_normalization (BatchNormalization)	(None, 50176)		200704
dense (Dense)	(None, 512)	ReLU	25690624
dropout (Dropout)	(None, 512)		0
dense_1 (Dense)	(None, 256)	ReLU	131328
dropout_1 (Dropout)	(None, 256)		0
dense_2 (Dense)	(None, 128)	ReLU	32896
dropout_2 (Dropout)	(None, 128)		0
dense_3 (Dense)	(None, 3)	Softmax	387

The Figure III shows plots between accuracy and epochs and between training loss and epochs for Model based on MobileNetV2

FIGURE III: Loss/Epoch Graph for MobileNetV2 based Model
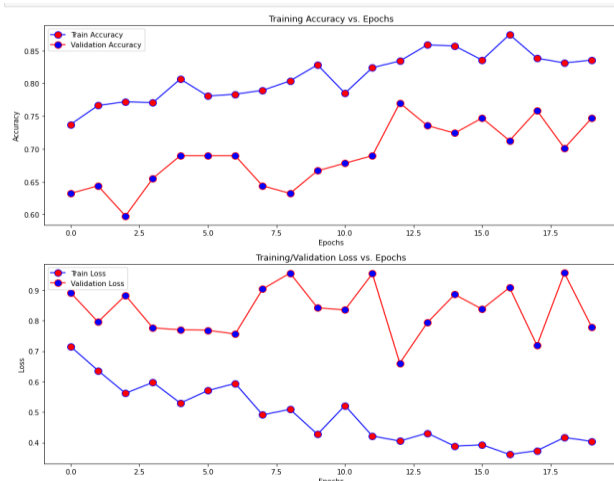
After the preparation of the test procedure for the display of test results. The test data's photos were categorized into benign, normal, and malignant in the results, which are displayed below.

The Figure IV shows plots between accuracy and epochs and between training loss and epochs

FIGURE IV: Loss/Epoch Graph for InceptionV3 based Model
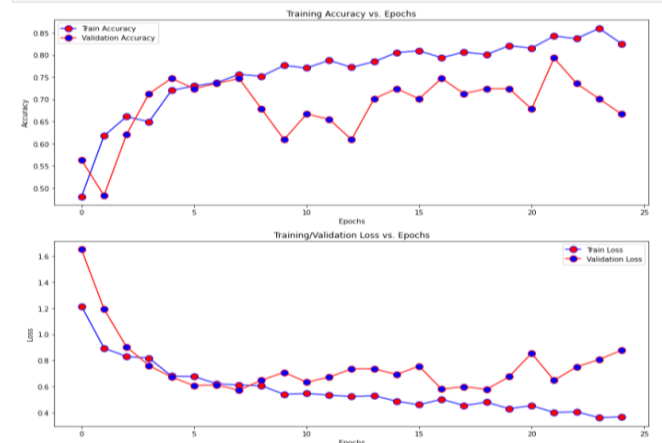
B. *InceptionV3*

Following MobileNetV2, we implemented the InceptionV3 architecture for the identification of breast cancer.

On the ImageNet dataset, it has been demonstrated that the image recognition model Inception v3 can attain optimal than 78.1% accuracy. The model is the result of numerous theories that various academics have created over time.

Convolutions, average pooling, max pooling, concatenations, dropouts, and entirely connected layers are a few of the asymmetric and symmetric building blocks that make up the model itself. The model makes heavy use of batch normalisation, which is also applied to the inputs for activation. The loss is computed using SoftMax.

The performance of the network's computing is decreased by this method since it needs fewer parameters. It also monitors the effectiveness of the network.

Using fewer, smaller convolutions instead of larger ones seems to speed up training. Think about a 25 parameter 5 5 filter. It can be replaced by two 3 3 filters, but then there are only 18 (3 3 + 3 3) parameters.

A 3-3 convolution could follow a 1-3 convolution in an uneven convolution, as opposed to the other way around. The parameter count would be slightly larger than the suggested asymmetric convolution if a 3 3 convolution were swapped out for a 2 2 convolution.

During training, a mini-CNN is added as an additional classifier. Any loss contributes to the overall network loss because of its interlayer placement

Table II below shows the layered architecture of the transfer learning model based on InceptionV3.

TABLE II: Model Architecture based on InceptionV3

Layer (Type)	Output Shape	Activation Function	Param #
inception_v3 (Functional)	(None, 5, 5, 2048)		21802784
flatten_9 (Flatten)	(None, 51200)		0
batch_normalization _210 (BatchNormalizatio n)	(None, 51200)		204800
dense_31 (Dense)	(None, 512)	ReLU	26214912
dropout_21 (Dropout)	(None, 512)		0
batch_normalization _211 (BatchNormalizatio n)	(None, 512)		2048
dense_32 (Dense)	(None, 256)	ReLU	131328
dropout_22 (Dropout)	(None, 256)		0
batch_normalization _212 (BatchNormalizatio n)	(None, 256)		1024
dense_33 (Dense)	(None, 128)	ReLU	32896
dropout_23 (Dropout)	(None, 128)		0
dense_34 (Dense)	(None, 3)	Softmax	387

Figure V shows the plots between accuracy and epochs, training loss and epochs.

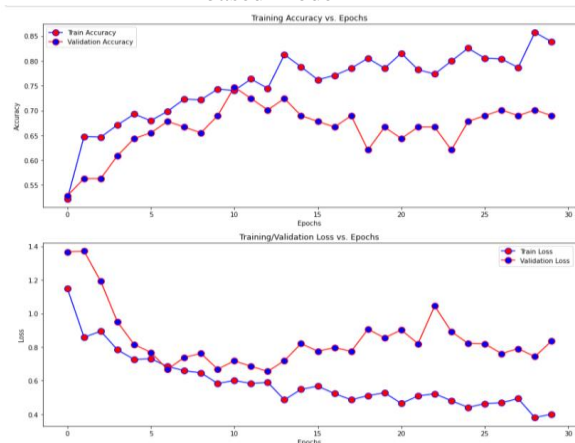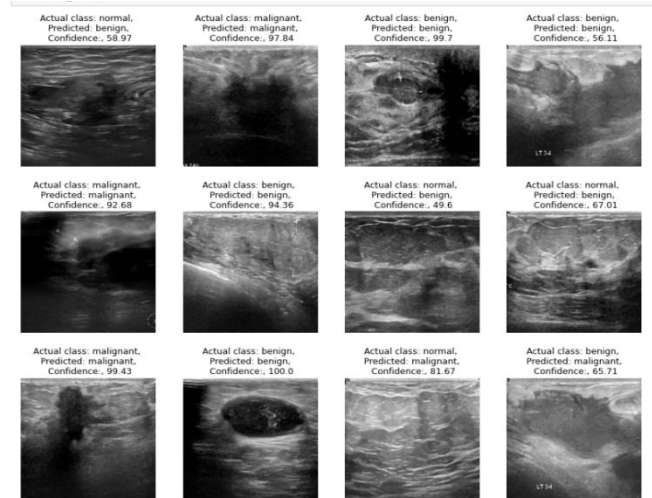FIGURE V: Loss/Epoch Graph for MobileNetV2 based Model

Figure VI shows the visualization of the test results using InceptionV3 architecture.

FIGURE VI: InceptionV3 Architecture Result Visualization

V. RESULT ANALYSIS:

Table III shows the accuracy, according to the techniques we used to find breast cancer:

TABLE III: Accuracy Results

Architecture	Accuracy
MobilenetV2	82.54%
InceptionV3	83.84%

InceptionV3 has the maximum accuracy of all the methods and can be used for breast cancer detection in conjunction with them, according to the data.

On the ImageNet dataset, it has been demonstrated that the image recognition model Inception v3 can achieve better than 78.1% accuracy. The model is the outcome of numerous theories that various academics have developed over time.

By lowering the amount of parameters used in a network, this technique decreases the network's computational performance. It also monitors the effectiveness of the network.

VI. CONCLUSION:

Breast cancer is the most prevalent and fatal disease, making detection a challenging endeavor. The condition gets worse every year and there is less chance of recuperation. To find breast cancer, deep learning and machine learning techniques are used. Our work and other research have shown that machine learning algorithms are more effective in their respective sectors. A variety of machine learning techniques were used in the prior work, along with some dataset augmentation and optimization for better performance. The recently developed deep learning technology is frequently employed in data science. The categorization of the breast cancer image data is done using CNN, a deep learning-based technique. The dataset from the image is primarily used by CNN. As a result, we used MobilenetV2 and InceptionV3 algorithms based on deep

learning for breast cancer screening and obtained useful results.

REFERENCES:

[1] Mutar, M.T.; Goyani, M.S.; Had, A.M.; Mahmood, A.S. Pattern of Presentation of Patients with Breast Cancer in Iraq in 2018: A Cross-Sectional Study. J. Glob. Oncol. 2019, 5, 1–6. [CrossRef] [PubMed]

[2] Coleman, C. Early Detection and Screening for Breast Cancer. Sem. Oncol. Nurs. 2017, 33, 141–155. [CrossRef] [PubMed]

[3] Smith, N.B.; Webb, A. Ultrasound Imaging. In Introduction to Medical Imaging: Physics, Engineering and Clinical Applications, 6th ed.; Saltzman, W.M., Chien, S., Eds.; Cambridge University Press: Cambridge, UK, 2010; Volume 1, pp. 145–197. [CrossRef]

[4] Gilbert, F.J.; Pinker-Domenig, K. Diagnosis and Staging of Breast Cancer: When and How to Use Mammography, Tomosyn thesis, Ultrasound, Contrast-Enhanced Mammography, and Magnetic Resonance Imaging. Dis. Chest Breast Heart Vessels 2019, 2019–2022, 155–166. [CrossRef]

[5] Jesneck, J.L.; Lo, J.Y.; Baker, J.A. Breast Mass Lesions: Computer-aided Diagnosis Models with Mammographic and Sonographic Descriptors. Radiology 2007, 244, 390–398. [CrossRef]

[6] Feldman, M.K.; Katyal, S.; Blackwood, M.S. US artifacts. Radiographics 2009, 29, 1179–1189. [CrossRef] [PubMed]

[7] Barr, R.; Hindi, A.; Peterson, C. Artifacts in diagnostic ultrasound. Rep. Med. Imaging 2013, 6, 29–49. [CrossRef]

[8] Zhou, Y. Ultrasound Diagnosis of Breast Cancer. J. Med. Imag. Health Inform. 2013, 3, 157–170. [CrossRef]

[9] Liu, S.; Wang, Y.; Yang, X.; Li, S.; Wang, T.; Lei, B.; Ni, D.; Liu, L. Deep Learning in Medical Ultrasound Analysis: A Review. Engineering 2019, 5, 261–275. [CrossRef]

[10] Huang, Q.; Zhang, F.; Li, X. Machine Learning in Ultrasound Computer-Aided Diagnostic Systems: A Survey. BioMed Res. Int. 2018, 7, 1–10. [CrossRef]

[11] Brattain, L.J.; Telfer, B.A.; Dhyani, M.; Grajo, J.R.; Samir, A.E. Machine learning for medical ultrasound: Status, methods, and future opportunities. Abdom. Radiol. 2018, 43, 786–799. [CrossRef]

[12] Sloun, R.J.G.v.; Cohen, R.; Eldar, Y.C. Deep Learning in Ultrasound Imaging. Proc. IEEE 2020, 108, 11–29. [CrossRef]

[13] Pan, S.J.; Yang, Q. A Survey on Transfer Learning. IEEE Trans. Knowl. Data Eng. 2010, 22, 1345–1359. [CrossRef]

[14] Khoshdel, V.; Ashraf, A.; LoVetri, J. Enhancement of Multimodal Microwave-Ultrasound Breast Imaging Using a Deep-Learning Technique. Sensors 2019, 19, 4050. [CrossRef] [PubMed]

[15] Day, O.; Khoshgoftaar, T.M. A survey on heterogeneous transfer learning. J. Big Dat. 2017, 4, 29. [CrossRef]

[16] Weiss, K.; Khoshgoftaar, T.M.; Wang, D. A survey of transfer learning. J. Big Dat. 2016, 3, 1–9. [CrossRef]

[17] Gentle Introduction to Transfer Learning. Available online: https://bit.ly/2KuPVMA (accessed on 10 November 2020).

[18] Taylor, M.E.; Kuhlmann, G.; Stone, P. Transfer Learning and Intelligence: An Argument and Approach. In Proceedings of the 2008 Conference on Artificial General Intelligence, Amsterdam, The Netherlands, 18–19 June 2008; pp. 326–337.

[19] Parisi, G.I.; Kemker, R.; Part, J.L.; Kanan, C.; Wermter, S. Continual lifelong learning with neural networks: A review. Neural Netw. J. Int. Neur. Net. Soci. 2019, 113, 54–71. [CrossRef] [PubMed]

[20] Silver, D.; Yang, Q.; Li, L. Lifelong Machine Learning Systems: Beyond Learning Algorithms. In Proceedings of the AAAI Spring Symposium, Palo Alto, CA, USA, 25–27 March 2013; pp. 49–55.

[21] Chen, Z.; Liu, B. Lifelong Machine Learning. Syn. Lect. Art. Intel. Machn. Learn. 2016, 10, 1–145. [CrossRef]

[22] Alom, M.Z.; Taha, T.; Yakopcic, C.; Westberg, S.; Hasan, M.; Esesn, B.; Awwal, A.; Asari, V. The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches. arXiv 2018, arXiv:abs/1803.01164.

[23] Huynh, B.; Drukker, K.; Giger, M. MO-DE-207B-06: Computer-Aided Diagnosis of Breast Ultrasound Images Using Transfer Learning from Deep Convolutional Neural Networks. Int. J. Med. Phys. Res. Prac. 2016, 43, 3705–3705. [CrossRef]

[24] Byra, M.; Galperin, M.; Ojeda-Fournier, H.; Olson, L.; O'Boyle, M.; Comstock, C.; Andre, M. Breast mass classification in sonography with transfer learning using a deep convolutional neural network and color conversion. Med. Phys. 2019, 46, 746–755. [CrossRef]

[25] Yap, M.H.; Pons, G.; Marti, J.; Ganau, S.; Sentis, M.; Zwiggelaar, R.; Davison, A.K.; Marti, R.; Moi Hoon, Y.; Pons, G.; et al. Automated Breast Ultrasound Lesions Detection Using Convolutional Neural Networks. IEEE J. Biomed. Health Inform. 2018, 22, 1218–1226. [CrossRef]

[26] Byra, M.; Sznajder, T.; Korzinek, D.; Piotrzkowska-Wroblewska, H.; Dobruch-Sobczak, K.; Nowicki, A.; Marasek, K. Impact of Ultrasound Image Reconstruction Method on Breast Lesion Classification with Deep Learning. arXiv 2018, arXiv:abs/1804.02119.

[27] Hijab, A.; Rushdi, M.A.; Gomaa, M.M.; Eldeib, A. Breast Cancer Classification in Ultrasound Images using Transfer Learning. In Proceedings of the 2019 Fifth International Conference on Advances in Biomedical Engineering (ICABME), Tripoli, Lebanon, 17–19 October 2019; pp. 1–4. [CrossRef] 2

[28] Yap, M.H.; Goyal, M.; Osman, F.M.; Martí, R.; Denton, E.; Juette, A.; Zwiggelaar, R. Breast ultrasound lesions recognition: End-to-end deep learning approaches. J. Med. Imaging 2019, 6, 1–7. [CrossRef]

[29] Hadad, O.; Bakalo, R.; Ben-Ari, R.; Hashoul, S.; Amit, G. Classification of breast lesions using cross-modal deep learning. IEEE 14th Intl. Symp. Biomed. Imaging 2017, 1, 109–112. [CrossRef]

[30] Transfer Learning. Available online: http://www.isikdogan.com/blog/transfer-learning.html (accessed on 20 November 2020).

[31] Chu, B.; Madhavan, V.; Beijbom, O.; Hoffman, J.; Darrell, T. Best Practices for Fine-Tuning Visual Classifiers to New Domains. In European Conference on Computer Vision; Springer: Cham, Switzerland, 2016; pp. 435–442.

[32] Transfer Learning. Available online: https://cs231n.github.io/transfer-learning (accessed on 19 November 2020).

[33] Yosinski, J.; Clune, J.; Bengio, Y.; Lipson, H. How transferable are features in deep neural networks? Adv. Neur. Inf. Proc. Sys. (NIPS). 2014, 27, 1–14.

[34] Huh, M.-Y.; Agrawal, P.; Efros, A.A.J.A. What makes ImageNet good for transfer learning? arXiv 2016, arXiv:abs/1608.08614.

[35] Li, Z.; Hoiem, D. Learning without Forgetting. IEEE Trans. Pattern Anal. Mach. Intell. 2018, 40, 2935–2947. [CrossRef]

[36] Building Trustworthy and Ethical AI Systems. Available online: https://www.kdnuggets.com/2019/06/5-ways-lack-datamachine-learning.html (accessed on 15 November 2020).

[37] Overfit and Underfit. Available online: https://www.tensorflow.org/tutorials/keras/overfit_and_underfit (accessed on 10 November 2020).

[38] Handling Overfitting in Deep Learning Models. Available online: https://towardsdatascience.com/handling-overfitting-indeep-learning-models-c760ee047c6e (accessed on 12 November 2020).

[39] Transfer Learning: The Dos and Don'ts. Available online: https://medium.com/starschema-blog/transfer-learning-the-dosand-donts-165729d66625 (accessed on 20 November 2020).

[40] Transfer Learning & Fine-Tuning. Available online: https://keras.io/guides/transfer_learning/ (accessed on 2 November 2020).

[41] How the pytorch freeze network in some layers, only the rest of the training? Available online: https://bit.ly/2KrE2qK (accessed on 2 November 2020).

[42] Transfer Learning. Available online: https://colab.research.google.com/github/kylemath/ml4aguides/blob/master/ notebooks/transferlearning.ipynb (accessed on 5 November 2020).

[43] A Comprehensive Hands-on Guide to Transfer Learning with Real-World Applications in Deep Learning. Available online: https://towardsdatascience.com/a-comprehensive-hands-on-guide-to-transfer-learning-with-real-world-applications-indeep-learning-212bf3b2f27a (accessed on 3 November 2020). Cancers 2021, 13, 738 14 of 15.

[44] Transfer Learning with Convolutional Neural Networks in PyTorch. Available online:

https://towardsdatascience.com/transferlearning-with-convolutional-neural-networks-in-pytorch-dd09190245ce (accessed on 25 October 2020).

[45] Best, N.; Ott, J.; Linstead, E.J. Exploring the efficacy of transfer learning in mining image-based software artifacts. J. Big Dat. 2020, 7, 1–10. [CrossRef]

[46] He, K.; Girshick, R.; Dollar, P. Rethinking ImageNet Pre-Training. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV), New York, NY, USA, 27 October–2 November 2019; pp. 4917–4926.

[47] Neyshabur, B.; Sedghi, H.; Zhang, C.J.A. What is being transferred in transfer learning? arXiv 2020, arXiv:abs/2008.11687.

[48] Liu, L.; Chen, J.; Fieguth, P.; Zhao, G.; Chellappa, R.; Pietikäinen, M. From BoW to CNN: Two Decades of Texture Representation for Texture Classification. Int. J. Comp. Vis. 2019, 127, 74–109. [CrossRef]

[49] Çarkacioglu, A.; Yarman Vural, F. SASI: A Generic Texture Descriptor for Image Retrieval. Pattern Recogn. 2003, 36, 2615–2633. [CrossRef]

[50] Yan, Y.; Ren, W.; Cao, X. Recolored Image Detection via a Deep Discriminative Model. IEEE Trans. Inf. Forensics Sec. 2018, 7, 1–7. [CrossRef]

[51] Imai, S.; Kawai, S.; Nobuhara, H. Stepwise PathNet: A layer-by-layer knowledge-selection-based transfer learning algorithm. Sci. Rep. 2020, 10, 1–14. [CrossRef]

[52] Zhao, Z.; Zheng, P.; Xu, S.; Wu, X. Object Detection with Deep Learning: A Review. IEEE Trans. Neur. Net. Learn. Sys. 2019, 30, 3212–3232. [CrossRef] [PubMed]

[53] Transfer Learning (C3W2L07). Available online: https://www.youtube.com/watch?v=yofjFQddwHE&t=1s (accessed on 3 November 2020).

[54] Zhang, J.; Li, W.; Ogunbona, P.; Xu, D. Recent Advances in Transfer Learning for Cross-Dataset Visual Recognition: A Problem Oriented Perspective. ACM Comput. Surv. 2019, 52, 1–38. [CrossRef]

[55] Nguyen, D.; Sridharan, S.; Denman, S.; Dean, D.; Fookes, C.J.A. Meta Transfer Learning for Emotion Recognition. arXiv 2020, arXiv:abs/2006.13211.

[56] Schmidt, J.; Marques, M.R.G.; Botti, S.; Marques, M.A.L. Recent advances and applications of machine learning in solid-state materials science. NPJ Comput. Mater. 2019, 5, 1–36. [CrossRef]

[57] D'souza, R.N.; Huang, P.-Y.; Yeh, F.-C. Structural Analysis and Optimization of Convolutional Neural Networks with a Small Sample Size. Sci. Rep. 2020, 10, 1–13. [CrossRef]

[58] Rizwan I Haque, I.; Neubert, J. Deep learning approaches to biomedical image segmentation. Inform. Med. Unlocked 2020, 18, 1–12. [CrossRef]

[72] Cheplygina, V.; de Bruijne, M.; Pluim, J.P.W. Not-so-supervised: A survey of semi-supervised, multi-instance, and transfer learning in medical image analysis. Med. Image Anal. 2019, 54, 280–296. [CrossRef] Cancers 2021, 13, 738 15 of 15

[73] Kensert, A.; Harrison, P.J.; Spjuth, O. Transfer Learning with Deep Convolutional Neural Networks for Classifying Cellular Morphological Changes. SLAS Discov. Adv. Life Sci. 2019, 24, 466–475. [CrossRef]

[74] Morid, M.A.; Borjali, A.; Del Fiol, G. A scoping review of transfer learning research on medical image analysis using ImageNet. Comput. Biol. Med. 2021, 128, 10–15. [CrossRef]

[75] Hesamian, M.H.; Jia, W.; He, X.; Kennedy, P. Deep Learning Techniques for Medical Image Segmentation: Achievements and Challenges. J. Dig. Imaging 2019, 32, 582–596. [CrossRef]

[76] Liu, L.; Ouyang, W.; Wang, X.; Fieguth, P.; Chen, J.; Liu, X.; Pietikäinen, M. Deep Learning for Generic Object Detection: A Survey. Int. J. Comput. Vis. 2020, 128, 261–318. [CrossRef]

[59] Azizi, S.; Mousavi, P.; Yan, P.; Tahmasebi, A.; Kwak, J.T.; Xu, S.; Turkbey, B.; Choyke, P.; Pinto, P.; Wood, B.; et al. Transfer learning from RF to B-mode temporal enhanced ultrasound features for prostate cancer detection. Int. J. Comp. Assist. Radiol. Surg. 2017, 12, 1111–1121. [CrossRef] [PubMed]

[60] Amit, G.; Ben-Ari, R.; Hadad, O.; Monovich, E.; Granot, N.; Hashoul, S. Classification of breast MRI lesions using small-size training sets: Comparison of deep learning approaches. Proc. SPIE 2017, 10134, 1–6.

[61] Tajbakhsh, N.; Jeyaseelan, L.; Li, Q.; Chiang, J.N.; Wu, Z.; Ding, X. Embracing imperfect datasets: A review of deep learning solutions for medical image segmentation. Med. Image Anal. 2020, 63, 1–30. [CrossRef] [PubMed]

[62] Yamashita, R.; Nishio, M.; Do, R.K.G.; Togashi, K. Convolutional neural networks: An overview and application in radiology. Insights Imaging 2018, 9, 611–629. [CrossRef] [PubMed]

[63] Calisto, F.M.; Nunes, N.; Nascimento, J. BreastScreening: On the Use of Multi-Modality in Medical Imaging Diagnosis. arXiv 2020, arXiv:2004.03500v2. [CrossRef]

[64] Evans, A.; Trimboli, R.M.; Athanasiou, A.; Balleyguier, C.; Baltzer, P.A.; Bick, U. Breast ultrasound: Recommendations for information to women and referring physicians by the European Society of Breast Imaging. Insights Imaging 2018, 9, 449–461. [CrossRef]

[65] Mammography in Breast Cancer. Available online: https://bit.ly/2Jyf8pl (accessed on 20 November 2020).

[66] Eggertson, L. MRIs more accurate than mammograms but expensive. CMAJ 2004, 171, 840. [CrossRef]

[67] Salem, D.S.; Kamal, R.M.; Mansour, S.M.; Salah, L.A.; Wessam, R. Breast imaging in the young: The role of magnetic resonance imaging in breast cancer screening, diagnosis and follow-up. J. Thorac. Dis. 2013, 5, 9–18. [CrossRef]

[68] A Literature Review of Emerging Technologies in Breast Cancer Screening. Available online: https://bit.ly/37Ccmas (accessed on 20 October 2020).

[69] Li, W.; Gu, S.; Zhang, X.; Chen, T. Transfer learning for process fault diagnosis: Knowledge transfer from simulation to physical processes. Comp. Chem. Eng. 2020, 139, 1–10. [CrossRef]

[70] Zhong, E.; Fan, W.; Yang, Q.; Verscheure, O.; Ren, J. Cross Validation Framework to Choose amongst Models and Datasets for Transfer Learning. In Proceedings of the Machine Learning and Knowledge Discovery in Databases, Berlin, Heidelberg, Germany, 12–15 July 2010; pp. 547–562.

[71] Baykal, E.; Dogan, H.; Ercin, M.E.; Ersoz, S.; Ekinci, M. Transfer learning with pre-trained deep convolutional neural networks for serous cell classification. Multimed. Tools Appl. 2020, 79, 15593–15611. [CrossRef]