

## Classification of Histopathological Images in Automatic Detection of Breast Cancer with Deep Learning Approach

Yasin KIRELLİ<sup>1\*</sup>, Gizem AYDIN<sup>2</sup>

<sup>1</sup>Kütahya Dumlupınar University, Management Information Systems Department, Kütahya-Türkiye

\* **Corresponding Author: Email:** yasin.kirelli@dpu.edu.tr – **ORCID:** 0000-0002-3605-8621

<sup>2</sup>Ereteam, İstanbul-Türkiye

**Email:** gaydin@ereteam.com – **ORCID:** 0000-0001-6353-0648

### **Article Info:**

DOI: 10.22399/ijcesen.1332504

Received : 25 July 2023

Accepted : 13 November 2023

### **Keywords**

Deep Learning

VGG16

Inceptionv3

Resnet50

BreakHis

### **Abstract:**

Convolutional neural networks have emerged as an essential tool for image classification and object detection. In the health field, these tools are a crucial factor in saving time and minimizing the margin of error for the health system and employees. Breast cancer is the most common type of cancer in women worldwide. In many cases, it can threaten human life, resulting in death. Although methods have been developed for the early diagnosis of this health problem, its support with digital systems remains incomplete. In diagnosis, histopathological images are examined with microscope methods. In cases where the number of pathologies is insufficient, delay problems may occur and the error rate increases in manual controls. The study aims to design a deep-learning object detection method for the pre-detection of breast cancer. The publicly published BreakHis dataset is used as the dataset. Model results that generated with VGG16, InceptionV3 and ResNet50 deep learning architectures have been compared. The highest accuracy rate have been obtained with the proposed model as 85%. Accuracy, AUC, precision, recall, F-score performance metrics have been analysed for each model. A decision support system screen design has been created using the proposed model weight file. With the study, the computer-assisted clinical support system makes clinicians' life more manageable and recommends early diagnosis.

## 1. Introduction

With the development of artificial intelligence day by day, image classification problems have become a popular topic. It is expected to make people's lives easier with machine learning methods used in many areas in our daily life. In recent years, machine learning techniques have been used in cancer detection [1]. In order to prevent human errors, it has become more important to move machine learning processes to automated processes. Various ML-based methods have been presented to contribute to early diagnosis in computer-aided diagnosis systems, and their reliability and effectiveness have been proven [2,3]. Deep learning has had great success in medical image diagnostics. Convolutional networks perform better in medical image analysis than other machine learning techniques. However,

large amounts of labelled data are needed for these networks to be effective [4,5].

Breast cancer is the most common type of cancer in women and the highest risk of death in the world. A report by the American Cancer Society shows that breast cancer accounts for an estimated 31% of new cancer cases in women [6]. In 2012, it was observed that approximately 12% of newly diagnosed cancer patients were women, but it has now been found that it is rarely seen in men. The most effective and essential way to prevent this situation is the early diagnosis of cancer [5,7]. Three different models, frequently preferred in the literature, have been used in the study. The user interface is designed with a proposed VGG16-based model with high accuracy. Performance metrics obtained from the models have been evaluated with AUC metrics and a confusion matrix.

## 2. Artificial intelligence (ai) studies in breast cancer

### 2.1. Breast Cancer

Breast cancer is a disease caused by cell mutations. It usually begins in the breast tissue but may rarely occur in the chest wall, breastbone, or lungs. Breast cancer, a malignant tumour that can be seen in both men and women, is the most common type of cancer affecting many women worldwide [13]. In addition, side effects during the treatment can affect the quality of life. Therefore, regular screening and early diagnosis are important in the fight against breast cancer [14].

According to statistics, breast cancer is one of the leading causes of cancer-related deaths in women. According to the World Health Organization, approximately 2.3 million women worldwide were diagnosed with breast cancer in 2020, and about 685,000 women died [14]. However, with regular screening tests and early detection, death rates from breast cancer can be reduced. According to the National Cancer Institute, diagnosing breast cancer in its early stages increases treatment options and control of the disease. Therefore, regular screening tests are essential for diagnosing breast cancer, especially in high-risk individuals (e.g., those with a family history) [13]. Breast cancer treatment may include different approaches such as surgery, chemotherapy, radiotherapy, and hormone therapy. However, treatment options are determined by the patient's age, general health, stage of the disease, and other factors. Breast cancer treatment can also cause various side effects, so it's crucial for patients to follow their treatment plan and consult their doctor to manage side effects. In conclusion, breast cancer is a significant health problem, and regular screening tests and early diagnosis are vital for treatment and maintenance of quality of life [15].

### 2.2. Early Detection of Breast Cancer

Early detection of breast cancer is vital as it is the most common type in women. Early breast cancer diagnosis involves detecting cancer cells before they multiply rapidly [16]. Thanks to early diagnosis, the progression of the disease is prevented by diagnosing cancer in treatable stages. Breast examination, mammography, and other imaging tests detect breast cancer early in women [17].

Women can monitor themselves by checking themselves regularly, but self-examination in this way is not enough. It is imperative to have regular screening tests. Traditional methods for the early diagnosis of this cancer take a long time and do not always give accurate results [18]. Therefore, automated machine learning systems are needed for breast cancer prediction.

### 2.3. Use of Artificial Intelligence in Breast Cancer Diagnosis

Artificial intelligence has been used widely in breast cancer diagnosis in recent years. Artificial intelligence and its deep learning subfield are practical tools for analysing and interpreting breast cancer data. This technology can help to obtain faster and more accurate results in screening tests, cancer diagnosis, and treatment processes. Artificial intelligence is accepted to automatically recognize and diagnose tumour lesions by quickly learning the image data with correctly labelled data [19]. In this way, it is expected to alleviate the workload of doctors to a great extent [20].

Mammography images create large datasets and artificial intelligence technologies such as machine learning and AI networks can be used to analyse these data. These technologies are also used to detect early stages of cancer in screening tests, which can help make treatment more successful [21]. In addition, predicting the risk of cancer spreading by AI tools. For example, an artificial intelligence algorithm that considers many factors such as tumour size, type, extent of spread, personal and family history, and genetic factors can predict a patient's risk of cancer progression. This can help make treatment more effective and personalized [22]. Related studies and AI methods that used in the subject are shown comparatively in the Table 1.

## 3. Materials and Methods

### 3.1. BreakHis Dataset

BreakHis is a database called Breast Cancer Histopathological Image Classification. Different size factors are used in this database as Table 2. It consists of 9,109 microscopic breast tumour images obtained from 82 patients.

**Table 1.** Literature review on the subject

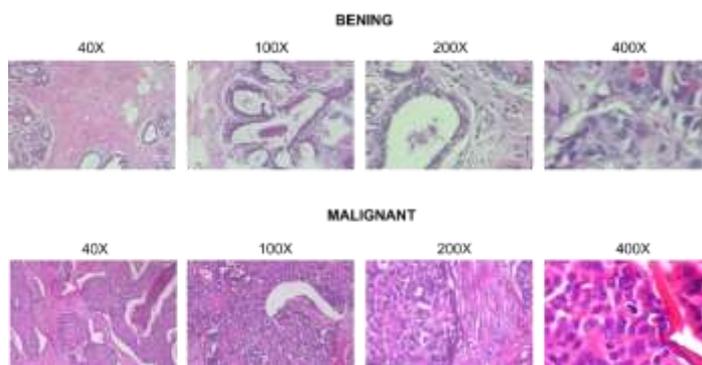
Author	Data Size	Method	Result - AUC
Farjana Parvin, Md. Al Mehedi Hasan[4]	2013	VGG-16	%86
		LeNet-5	%50
		AlexNet	%85
		ResNet-50	%50
Xujuan Zhou, Yuefeng Li v.d.[8]	10389	Inception-v1	%94
		The Proposed Model	%84
Karthiga R, Narasimhan K[9]	300	GoogleNet	%54
		SVM	%78
		Linear SVM	%93
		Fine Gaussian SVM	%92
		Weighted KNN	%91
Zhan Xiang, Zhang Ting v.d.[10]	4960	Fine KNN	%90
		AlexNet	%83
Mahesh Gour, Sweta Jain, T.Sunil Kumar[11]	2013	AlexNet	%84
		GoogleNet	%83
		VGG16	%83
Karan Gupta, Nidhi Chawla[12]	1617	VGG16	%82
		VGG19	%83
		Xception	%78
		ResNet50	%84

- Size factors of images include 40X, 100X, 200X, and 400X.
- These images contain a total of 2,480 samples, both benign and malignant.
- The features of the images are represented by 700X460 pixels, 3-channel RGB color space, and each channel has 8 bits of depth.
- The format of the images is PNG [23].

**Table 2.** BreakHis image sets and features

Magnification	Benign	Malignant	Total
40X	652	1,370	1,995
100X	644	1,437	2,081
200X	623	1,390	2,013
400X	588	1,232	1,820
Total of Images	2,480	5,429	7,909

The image below is image of some examples in the BreakHis dataset. The upper row shows benign tumours of different sizes, and the lower row shows malignant tumours of various sizes as Figure 1.



**Figure 1.** BreakHis image set samples

- 800 images from the BreakHis image database have been used in the model.
- Of these, 640 images have been used as a training set and 160 as a test dataset.
- Training and test data sets are separated by 80% - 20%, respectively.
- Within the training and test data, 200x images have been used in the dataset.

### 3.2. Proposed VGG Based Model

VGG16 is a CNN architecture introduced in 2014 by Karen Simonyan and Andrew Zisserman in their article titled Very Deep Convolutional Networks for Large-Scale Image Recognition [24]. There are two forms of the VGG Neural Network model. These are the VGG16 and VGG19 models. VGG16 is a 16-layer neural network and is one of the most popular deep learning architectures, not counting the maximum pooling and softmax layers. VGG19 is a slightly deeper version of VGG16 and consists of 19 layers. The Keras library includes pre-trained models VGG16 and VGG19 for both Theano and TensorFlow backends. These models have pre-trained weights for use in tasks such as image classification and are frequently preferred for transfer learning [25]. The VGG16 architecture is as Figure 2.

- VGG16 uses only 3x3 size filters in convolutional and pooling layers. These filters perform convolution by acting on the input data at 3x3 size and then downscaling it to a 2x2 layer for pooling.
- Convolutional layers use 3x3 filters in VGG16. This small kernel size is used to create feature maps that are more complex and can generalize better.
- VGG16 has 1x1 convolution layers to transform the input linearly. These layers perform dimensional transformations by changing the input data channels without interacting with each other.

- In convolutional layers of the network, the step size is 1 pixel. This helps preserve spatial resolution because filters are applied to each pixel, thus helping to maintain detail.
- ReLU (Rectified Linear Unit) activation function is used in all hidden layers of VGG16. ReLU adds non-linearity by directly passing positive inputs while reducing negative inputs to zero.

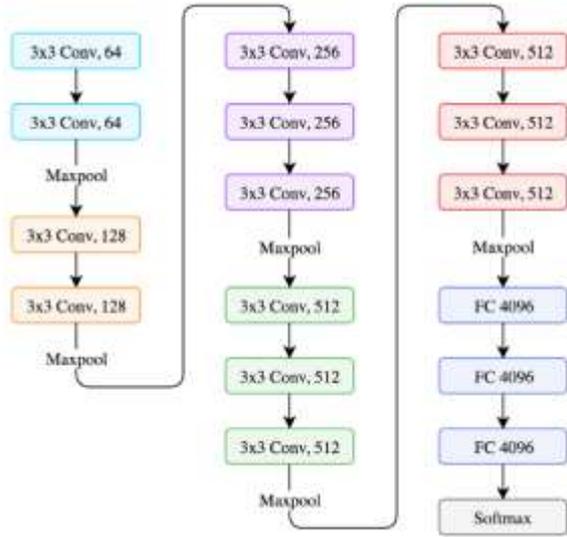


Figure 2. VGG16 Architecture [25]

The model has been developed using Python language and Keras deep learning library. As shown in Figure 3, the Vgg16 model architecture is the first layer and offers a relearning model with a previously learned model. As shown in the figure, the number of parameters at 14 million in the Vgg16 layer starts with a previously pre-trained parameter number.

Layer (type)	Output Shape	Param
vgg16 (Functional)	(None, 7, 7, 512)	14714
flatten (Flatten)	(None, 25088)	0
dense (Dense)	(None, 256)	64227
dense_1 (Dense)	(None, 1)	257
-----		
Total params: 21,137,729		
Trainable params: 6,423,041		
Non-trainable params: 14,714,688		



Figure 3. Proposed VGG16 based architecture

The result of a 70-epoch training cycle is shown in accuracy and loss graphs with training and validation curves. According to the charts, it is seen in the graph that the accuracy rate is 85% on average, and it performs successful classification as shown Figure 4.

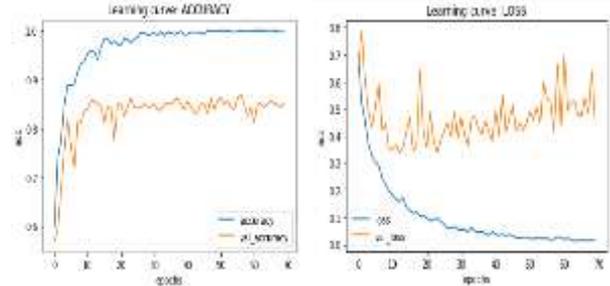


Figure 4. Proposed model accuracy and loss chart

Accuracy, precision, recall, and f1 score values according to the general and individual metrics of the classification metrics are given in the table. According to the metrics, a satisfactory classification has been achieved with 85% accuracy. According to the classification metrics of the separate classes, a classification rate has been obtained as more successful than benign, with a success rate of 86% and 83% for malignant as shown Table 3.

Table 3. Performance metrics of proposed model

	Accuracy	Precision	Recall	F1-Score
General	0.850	0.858	0.850	0.849
Class 'benign'	-	0.912	0.775	0.838
Class 'malignant'	-	0.804	0.925	0.860

As another metric in the classification performance obtained with the test data set, the ROC curve is given in the image below. The upward trend of the curve indicates a successful classification efficiency with an AUC ratio of 0.93 as shown Figure 5.

### 3.3. ResNet50 Architecture Model

ResNet-50 is a 50-layer convolutional neural network and backbone architecture used in computer



**Asymmetric convolutions:** This technique helps capture features of different scales and orientations using various filter sizes in other channels.

**Auxiliary classifiers:** These classifiers are added to the intermediate layers of the network to help improve gradient flow and regularization.

**Efficient grid dimension reduction:** This technique combines maximum pooling and convolutional layers to reduce the spatial dimensions of feature maps, which helps minimize computational load while preserving crucial spatial information. These architectural improvements have helped Inception V3 achieve state-of-the-art results on various image classification benchmarks. For example, Inception V3 has reached a top 5 error rate of 23.1% in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2015 [29-33].

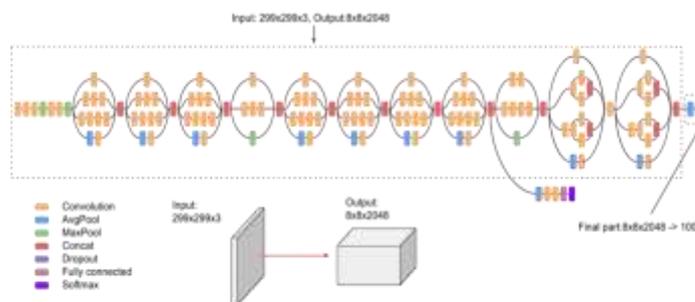


Figure 9. InceptionV3 architecture

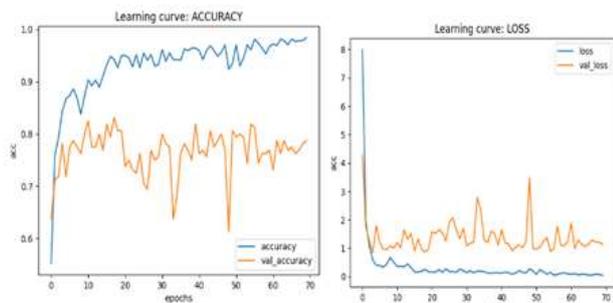


Figure 10. InceptionV3 model accuracy and loss chart

According to the graphs, it is seen in the chart that the accuracy rate is 78% on average, and the ROC curve, as another metric in the classification performance obtained with the test data set, is given in the image below with an AUC ratio of 0.86 as shown Figure 10 and Figure 11. Accuracy, precision, recall, and f1 score values according to the general and individual metrics of the classification metrics are given in the table. It has been observed that with 78% accuracy according to the metrics. According to the classification metrics of the separate classes

Table 5. Performance metrics of InceptionV3 model

	Accuracy	Precision	Recall	F1-Score
General	0.787	0.788	0.787	0.787
Class 'benign'	-	0.774	0.812	0.793
Class 'malignant'	-	0.803	0.762	0.782

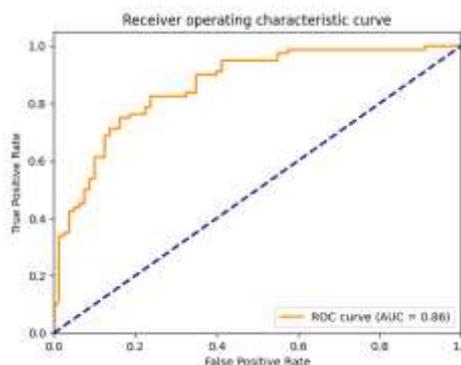


Figure 11. InceptionV3 model ROC curve

a classification rate has been obtained as more successful than benign as shown Table 5. A screen for the end user has been developed using the Python Tkinter library. The Keras weight file has been obtained due to the application's model training. The flowchart shows that the weight file of the presented model that provides the most successful classification among the three models has been selected as shown Figure 12. Thus, as shown in the screenshots as Figure 13, as a result of a fast classification, the classification process from the selected image takes less than one second. It creates a decision support system for physicians. In future studies, it has shed light on using different layered model architectures to increase the classification success.

#### 4. Results

In this study proposed vgg-based deep learning method has been developed to classify histopathological images of breast cancer. The presented model has achieved successful results in the classification of pathological images, and accordingly, it shows that it can be an essential help to doctors in clinical applications. The results obtained with the presented model emphasize the

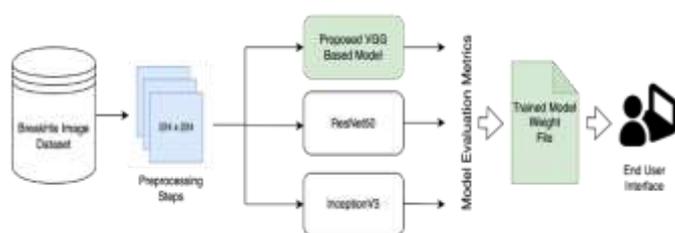


Figure 12. Clinical decision support system workflow chart

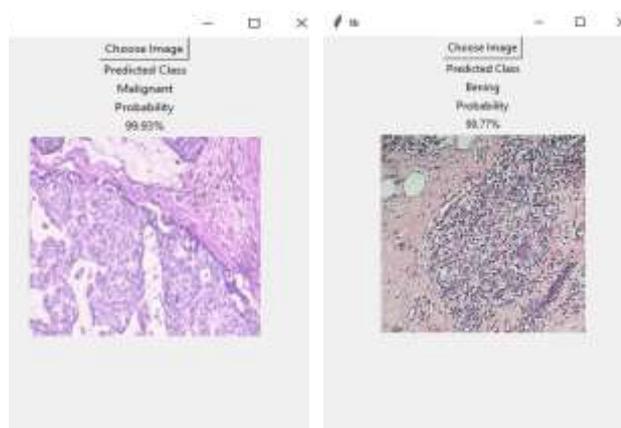


Figure 13. User interface and sample result screens

importance of making a quick and accurate decision to diagnose the disease and provide a decision support system with the interface created. With the proposed model, the result presented as the study output, false positive or false negative classification in pathological classification, can help reduce the effects and make the treatment process more effective. In addition, the high sensitivity obtained as a result of the presented model classification can reduce the risk of exposure of patients and physicians to the workload of patients and physicians when potentially invasive interventions are not needed. In conclusion, this study demonstrated the potential of using deep learning techniques, user interfaces, and practical applications to diagnose breast cancer. The developed VGG-based model offers high accuracy and sensitivity, close to 90 percent, which can be used as a reliable tool to classify histopathological images. This study may contribute to the further development of artificial intelligence-based methods for disease diagnosis with histopathological images by inspiring future research.

#### Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.

- **Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper
- **Acknowledgement:** The authors declare that they have nobody or no-company to acknowledge.
- **Author contributions:** The authors declare that they have equal right on this paper.
- **Funding information:** The authors declare that there is no funding to be acknowledged.
- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

#### References

- [1] Abdar, M., Zomorodi-Moghadam, M., Zhou, X., Gururajan, R., Tao, X., Barua, P. D., & Gururajan, R. (2020). A new nested ensemble technique for automated diagnosis of breast cancer, *Pattern Recognition Letters*, vol. 132, pp. 123–131, doi: 10.1016/j.patrec.2018.11.004.
- [2] Liu, M., Hu, L., Tang, Y., Wang, C., He, Y., Zeng, C., ... & Huo, W. (2022). A Deep Learning Method for Breast Cancer Classification in the Pathology Images, *IEEE Journal of Biomedical Health Informatics*, vol. 26, no. 10, pp. 5025–5032, doi: 10.1109/JBHI.2022.3187765.
- [3] Nahid, A. A., & Kong, Y. (2017). Involvement of machine learning for breast cancer image classification: a survey. *Computational and mathematical methods in medicine*, 2017., doi: 10.1155/2017/3781951.
- [4] Parvin, F., & Hasan, M. A. M. (2020, June). A comparative study of different types of convolutional neural networks for breast cancer histopathological image classification. In *2020 IEEE Region 10 Symposium (TENSYP)* (pp. 945-948), doi: 10.1109/TENSYP50017.2020.9230787.
- [5] Jiang, Y., Chen, L., Zhang, H., & Xiao, X. (2019). Breast cancer histopathological image classification using convolutional neural networks with small SE-ResNet module. *PloS one*, 14(3), e0214587, doi: 10.1371/journal.pone.0214587.
- [6] Giaquinto, A. N., Miller, K. D., Tossas, K. Y., Winn, R. A., Jemal, A., & Siegel, R. L. (2022). Cancer statistics for African American/black people 2022. *CA: a cancer journal for clinicians*, 72(3), 202–229, doi: 10.3322/caac.21718.
- [7] Kakde, A., Arora, N., & Sharma, D. (2019). A comparative study of different types of cnn and highway cnn techniques. *Global Journal of Engineering Science and Research Management*, 6(4), 18-31, doi: 10.5281/zenodo.2639265.

- [8] Zhou, X., Li, Y., Gururajan, R., Bargshady, G., Tao, X., Venkataraman, R., ... & Kondalsamy-Chennakesavan, S. (2020, November). A new deep convolutional neural network model for automated breast Cancer detection. In *2020 7th International Conference on Behavioural and Social Computing (BESC)* (pp. 1-4), doi: 10.1109/BESC51023.2020.9348322.
- [9] Karthiga, R., & Narasimhan, K. (2018, March). Automated diagnosis of breast cancer using wavelet based entropy features. In *2018 Second international conference on electronics, communication and aerospace technology (ICECA)* (pp. 274-279), doi: 10.1109/ICECA.2018.8474739.
- [10] Xiang, Z., Ting, Z., Weiyan, F., & Cong, L. (2019, June). Breast cancer diagnosis from histopathological image based on deep learning. In *2019 Chinese Control And Decision Conference (CCDC)* (pp. 4616-4619), doi: 10.1109/CCDC.2019.8833431.
- [11] Gour, M., Jain, S., & Sunil Kumar, T. (2020). Residual learning based CNN for breast cancer histopathological image classification. *International Journal of Imaging Systems and Technology*, 30(3), 621-635, doi: 10.1002/ima.22403.
- [12] Gupta, K., & Chawla, N. (2020). Analysis of histopathological images for prediction of breast cancer using traditional classifiers with pre-trained CNN. *Procedia Computer Science*, 167, 878-889, doi: 10.1016/j.procs.2020.03.427.
- [13] "Breast Cancer Treatment," National Cancer Institute, 2023. <https://www.cancer.gov/types/breast/patient/breast-treatment-pdq>.
- [14] "Breast Cancer," World Health Organization, 2021. <https://www.who.int/news-room/fact-sheets/detail/breast-cancer>.
- [15] DeSantis, C. E., Bray, F., Ferlay, J., Lortet-Tieulent, J., Anderson, B. O., & Jemal, A. (2015). International variation in female breast cancer incidence and mortality rates. *Cancer epidemiology, biomarkers & prevention*, 24(10), 1495-1506, doi: 10.1158/1055-9965.EPI-15-0535.
- [16] Anderson, B. O., Braun, S., Lim, S., Smith, R. A., Taplin, S., & Thomas, D. B. (2003). Early detection of breast cancer in countries with limited resources. *The breast journal*, 9, S51-S59, doi: 10.1046/j.1524-4741.9.s2.4.x.
- [17] Kösters, J. P., Götzsche, P. C., & Cochrane Breast Cancer Group. (1996). Regular self-examination or clinical examination for early detection of breast cancer. *Cochrane Database of Systematic Reviews*, 2010(1), doi: 10.1002/14651858.CD003373.
- [18] Spanhol, F. A., Oliveira, L. S., Petitjean, C., & Heutte, L. (2016, July). Breast cancer histopathological image classification using convolutional neural networks. In *2016 international joint conference on neural networks (IJCNN)* (pp. 2560-2567), doi: 10.1109/IJCNN.2016.7727519.
- [19] Zheng, D., He, X., & Jing, J. (2023). Overview of artificial intelligence in breast cancer medical imaging. *Journal of Clinical Medicine*, 12(2), 419, doi: 10.3390/jcm12020419.
- [20] Evans, K. K., Birdwell, R. L., & Wolfe, J. M. (2013). If you don't find it often, you often don't find it: why some cancers are missed in breast cancer screening. *PloS one*, 8(5), e64366., doi: 10.1371/journal.pone.0064366.
- [21] Cheng, J. Z., Ni, D., Chou, Y. H., Qin, J., Tiu, C. M., Chang, Y. C., ... & Chen, C. M. (2016). Computer-aided diagnosis with deep learning architecture: applications to breast lesions in US images and pulmonary nodules in CT scans. *Scientific reports*, 6(1), 24454.
- [22] Esteva, A., Kuprel, B., Novoa, R. et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 115-118 (2017). <https://doi.org/10.1038/nature21056>.
- [23] Spanhol, F. A., Oliveira, L. S., Petitjean, C., & Heutte, L. (2015). A dataset for breast cancer histopathological image classification. *Ieee transactions on biomedical engineering*, 63(7), 1455-1462, doi: 10.1109/TBME.2015.2496264.
- [24] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- [25] Sewak, M., Karim, M. R., & Pujari, P. (2018). *Practical convolutional neural networks: implement advanced deep learning models using Python*. Packt Publishing Ltd.
- [26] ERDEM, E., & AYDIN, T. (2021). Göğüs Kanseri Histopatolojik Görüntü Sınıflandırması. *Bilişim Teknolojileri Dergisi*, 14(1), 87-94, doi: 10.17671/gazibtd.746673.
- [27] A. Rastogi, "ResNet50." <https://blog.devgenius.io/resnet50-6b42934db431>.
- [28] S. Mukherjee, "The Annotated ResNet-50." <https://towardsdatascience.com/the-annotated-resnet-50-a6c536034758>.
- [29] Al Husaini, M. A. S., Habaebi, M. H., Gunawan, T. S., Islam, M. R., & Hameed, S. A. (2021, June). Automatic breast cancer detection using inception V3 in thermography. In *2021 8th International Conference on Computer and Communication Engineering (ICCCE)* (pp. 255-258), doi: 10.1109/ICCCE50029.2021.9467231.
- [30] Cao, J., Yan, M., Jia, Y., Tian, X., & Zhang, Z. (2021). Application of a modified Inception-v3 model in the dynasty-based classification of ancient murals. *EURASIP Journal on Advances in Signal Processing*, 2021(1), 1-25., doi: 10.1186/s13634-021-00740-8.
- [31] Mujahid, M., Rustam, F., Álvarez, R., Luis Vidal Mazón, J., Diez, I. D. L. T., & Ashraf, I. (2022). Pneumonia classification from X-ray images with inception-V3 and convolutional neural network. *Diagnostics*, 12(5), 1280, doi: 10.3390/diagnostics12051280.
- [32] Mednikov, Y., Nehemia, S., Zheng, B., Benzaquen, O., & Lederman, D. (2018, July). Transfer representation learning using Inception-V3 for the detection of masses in mammography. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 2587-2590). IEEE., doi: 10.1109/EMBC.2018.8512750.

- [33] Liu, Z., Yang, C., Huang, J., Liu, S., Zhuo, Y., & Lu, X. (2021). Deep learning framework based on integration of S-Mask R-CNN and Inception-v3 for ultrasound image-aided diagnosis of prostate cancer. *Future Generation Computer Systems*, 114, 358-367, doi: 10.1016/j.future.2020.08.015.