



# Breast Cancer Diagnosis Based On Thermography Images Using Pre-Trained Networks

<sup>1</sup>Hasan Ucuzal , <sup>2</sup>Muhammet Baykara , <sup>3</sup>Zeynep Kucukakcali 

<sup>1</sup>Inonu University Department of Biostatistics and Medical Informatics, Faculty of Medicine, Malatya, Turkey. (e-mail: [hasan.ucuzal@inonu.edu.tr](mailto:hasan.ucuzal@inonu.edu.tr)).

<sup>2</sup>Firat University, Faculty of Technology, Department of Software Engineering, Elazığ, Turkey. (e-mail: [mbaykara@firat.edu.tr](mailto:mbaykara@firat.edu.tr)).

<sup>3</sup>Inonu University Department of Biostatistics and Medical Informatics, Faculty of Medicine, Malatya, Turkey. (e-mail: [zeynep.tunc@inonu.edu.tr](mailto:zeynep.tunc@inonu.edu.tr)).

## ARTICLE INFO

**Received:** Aug.,03.2021

**Revised:** Oct,21.2021

**Accepted:** Nov.,17.2021

### Keywords:

Classification  
Breast cancer  
Deep learning  
Image processing  
Thermography

**Corresponding author:** Zeynep

Küçükakçalı

✉ [zeynep.tunc@inonu.edu.tr](mailto:zeynep.tunc@inonu.edu.tr)

☎ +90 536 424 32 06

**ISSN:** 2548-0650

**DOI:** <https://doi.org/10.52876/jcs.990948>

## ABSTRACT

Breast cancer is the leading cause of death among women around the world. Because of its low cost and the fact that it does not emit hazardous radiation, infrared thermography has emerged as a viable approach for diagnosing the condition in young women. This study aims to create a computer-aided diagnostic system that can process thermographic breast cancer images and classify breast cancer with pre-trained networks in order to use thermography as a diagnostic method. In this study, an open-access data set consisting of thermographic breast cancer images was used for diagnostic purposes. The data set consists of 179 healthy images and 101 images from patients. The images were converted from .txt format to .jpeg format. The data set is acquired from <http://visual.ic.uff.br/dmi/>. In this study, various pre-trained networks were used to reduce the training time. Different metrics were employed to assess the performance of the models. The images obtained during the modeling phase were used to display both breasts in the image without distinguishing the right and left breasts, that is, without fragmenting the images. According to the results of the different pre-trained network models after the data preprocessing stages, the best classification performance was achieved for the ResNet50V2 model with an accuracy value of 0.996. In this study, a computer-aided diagnosis system was created by developing an interface for breast cancer classification from thermographic images in addition to experimental findings. The web software based on the proposed models has provided promising predictions of breast cancer from thermographic images. The developed software can help medical and other healthcare professionals easily spot breast cancer.

## 1. INTRODUCTION

CANCER is one of the most difficult and deadly diseases in the world. Every year, it claims the lives of millions of people. Cancer claimed the lives of 8.8 million people in 2015, according to the World Health Organization (WHO) [1].

Because of somatic cell changes caused by epigenetic or genetic changes, malignant neoplasm or cancer is considered a genetic illness [2]. Cancer affects people of all ages, including men and women, the old and the young, but it is more prevalent in the elderly than in the youth. Different statistics exist for the most common forms of malignancies, and both geographic location and gender have a factor in their occurrence. Breast cancer is one of the most often diagnosed cancers among women and is the leading cause of death among women around the world. It is critical to discover the disease early to improve the odds of people being treated and cured. Because of its low cost and the fact that it does not emit hazardous radiation, infrared thermography has emerged as a

viable approach for diagnosing the condition in young women. As a result, research is being conducted to find both diagnostic and therapy methods. In general, the causes of breast cancer are unknown. As of today, doctors and specialists do not have an exact justification for the incidence of breast cancer in some women over others [3]. It is critical to detect breast cancer early to prevent mortality and morbidity [4].

Mammography, ultrasound, and magnetic resonance imaging are the most often used imaging modalities for early breast cancer detection. However, constraints such as x-rays, cost, dense tissue at a young age, false positives (FP), and false-negative (FN) rates prompted researchers and institutions to do an extensive study into other techniques such as thermography. Contrary to other methods, thermography is a non-invasive, non-inclusive, radiation-free, and low-cost technique [5]. Thermography has experienced a surge in popularity in recent years, particularly for breast cancer screening [6]. This is owing to the allure of its low-risk technology, as well as the possibility of future advancements

with modern technological innovation. The ultimate goal of current research in this subject is to develop a more accurate and confirmed tumor diagnosis that can be used as a gold standard for breast cancer screening [3]. These innovative approaches, such as thermography, are frequently used in conjunction with computer-aided diagnostic (CAD) systems.

One of the most common applications in breast cancer screening although is thermography, which has not been accepted as the standard procedure. Thermographic imaging is yet to be accepted as the gold standard for this application. Furthermore, even though mammography is not a completely risk-free procedure, physicians prefer it over thermography results. As a result, if thermographic breast cancer screening improves to a satisfactory level, it can be a replacement candidate. The major problem to solve here is the image-processing task. For this reason, computer-aided diagnosis systems should be constructed by applying image processing methods to thermographic breast cancer images and obtaining valid results by processing them [3].

In a clinical study, the approaches for computer-aided diagnosis of breast cancer utilizing thermal imaging are presented in a paper using various Convolutional Neural Networks (CNNs). The accuracy, precision, recall, F1 measure, and Matthews Correlation coefficient of the developed nets were evaluated on a benchmarking dataset. The results reveal that architecture with pre-trained convolutional layers and training freshly added fully connected layers outperforms other architectures. Using transfer learning methods and CNN, 94.3 percent accuracy, 94.7 percent precision, and 93.3 percent recall were obtained [7]. Another work uses convolutional neural networks (i.e., AlexNet, GoogLeNet, ResNet-18, VGG-16, and VGG-19) to identify 440 infrared photos of 88 individuals into two categories: normal and pathological demonstrate that combining deep learning techniques with infrared imagery to help breast cancer diagnosis has a lot of promise for clinical purposes [8]. Additionally, the use of a deep convolutional neural network with transfer learning to automatically classify thermograms into two classes (normal and abnormal) is proposed in a research. A total of 311 female individuals were used to assess the CNN's performance, with one research using a balanced class distribution and the other using a typical screening cohort with a low proportion of aberrant thermograms. The ResNet-101 model exhibited a sensitivity of 92.3 percent and a specificity of 53.8 percent, respectively. These findings imply that the proposed model can accurately categorize abnormal thermograms, indicating that infrared thermography can be used as an auxiliary tool for breast cancer screening [9].

This study aims to create a computer-aided diagnostic system that can process thermographic breast cancer images and classify breast cancer with pre-trained networks in order to use thermography as a diagnostic method.

## 2. MATERIAL and METHODS

### 2.1. Dataset

In this study, an open-access data set consisting of thermographic breast cancer images was used for diagnostic purposes. The data set includes a total of 5862 images, including 3504 healthy and 2457 diseased images. Images are in 640\*480 jpeg format. The data set consists of 179 healthy images and 101 images from patients. The images were converted from .txt format to .jpeg format. The data set is acquired from <http://visual.ic.uff.br/dmi/>. Sample images of the data set are given in Figure 1.

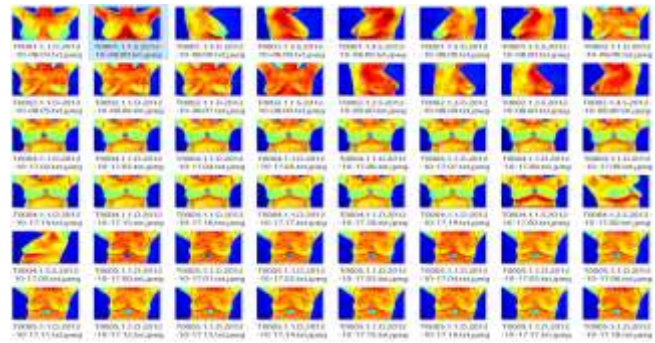


Fig. 1. Sample images of the data set

### 2.2. Deep Learning

The Deep learning, a subfield of machine learning (ML), has grown in prominence in recent years as computing power has increased dramatically and massive new data sets are being generated daily. Deep learning has produced ground-breaking results in a range of hard tasks, including image classification, object detection, speech recognition, language translation, natural language processing, and games. Deep learning models have outperformed many traditional machine learning approaches in terms of modeling big data sets due to their ability to function on the graphics processing unit. In the processing of medical image and video data, CNN deep learning architecture is commonly utilized. Deep learning algorithms such as CNN have recently been used as a decision support system for breast cancer detection on mammograms, segmentation of liver metastases with computed tomography (CT), brain tumor segmentation with magnetic resonance (MR) imaging, and classification of high-resolution chest CT images of interstitial lung patients [10]. Multiple layers, pooling, local connections, and shared weights distinguish CNN from a typical neural network. The primary idea behind CNN is that input data can be interpreted as images. This can reduce the number of parameters used, resulting in faster processing. Convolutional layers (CLs), pooling layers (PLs), and the rectified linear unit (ReLU) are the layers that makeup CNN architecture (ReLU). CLs learn the convolutions and deliver the best data categorization performance. Overfitting is controlled by PLs, which allow for stable conversion and enhance computing performance by lowering the number of structures produced by convolutions. Through the ReLU activation function, ReLU improves the network's nonlinear features. Various study structures have been devised and introduced based on the type of data, image, and objective [11, 12]. In image processing studies with deep learning methods, the model training phase is time-consuming. Furthermore, training a deep learning model requires a lot of data as well as processing time. It is often useful to use pre-trained networks on big datasets lasting days or weeks to train and enhance use-case training. Therefore, in this study, many different pre-trained networks were used to reduce the training time.

#### 2.2.1. Image Processing

Image processing is the process of digitizing an image using various methods and techniques to extract usable information and produce enhanced images. In other words, image processing is a set of computer studies aiming at altering digital picture data using a computer or software in order to meet a certain criterion [13].

### 2.2.2. Thermography

At certain temperatures, objects inherently generate thermal signals. The type of signals transmitted or radiated, and the temperature range, are determined by the object's features. Under normal circumstances, the human body, like other objects, emits Infrared (IR) signals [14]. Because of the heat, the transmitted IR signals differ from one part to the next. This notion is commonly used in medical tests, particularly for breast cancer screenings. Any malignant growth in the breast is linked to the development of inflammation and blood vessels, both of which have a higher temperature profile. Thermography, often known as thermal imaging, is a non-invasive way of collecting the heat map of a specific limb in medicine. The approach is non-contact, non-destructive, and does not emit radiation, making it suitable for repeated usage. The procedure here uses a thermographic camera to capture a heat map of the breasts and their surroundings to highlight any anomalies [3]. Examples of diseased and healthy breast images are given in Figure 2.

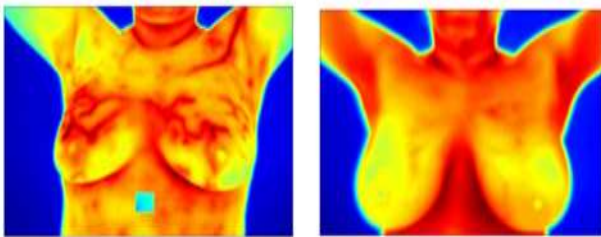


Fig. 2. Examples of the diseased and healthy breast images

Thermography has been recently revived in medical applications. It is evaluated together with image processing methods, especially for use in the diagnosis of breast cancer. Breast thermography takes advantage of the variation in the heat map beneath the skin between healthy and cancerous breasts. The presence of a breast tumor raises the warmth of the tissues that surround it [15]. Asymmetric study of healthy and diseased breasts is generally used by specialists. The process for using thermography to screen for breast cancer is relatively simple. It begins with a physical inspection of the surface of the breasts. This allows the doctor to correlate any unexpected presence to the heat map. Then, the individual needs to stay at room temperature for 15 minutes to acclimate to the environment. This procedure is done in a room where both humidity and temperature are controlled. At the same time, the person will need to undress the upper part of their body from waist to chin. After the body temperature reaches equilibrium, the person is asked to take his hands on his right side so that the relevant surfaces can be viewed. Then the imaging procedure starts to complete the process [3].

### 2.3. Developed Web-Based Software

The developed web-based interface is designed to classify breast cancer from thermographic breast cancer images. During the development of the interface, Python programming language and additionally Flask [16], TensorFlow [17], Keras [18], Pandas [19], NumPy [20], Scikit-learn [21] libraries were used. A general working diagram is given in Figure 3.

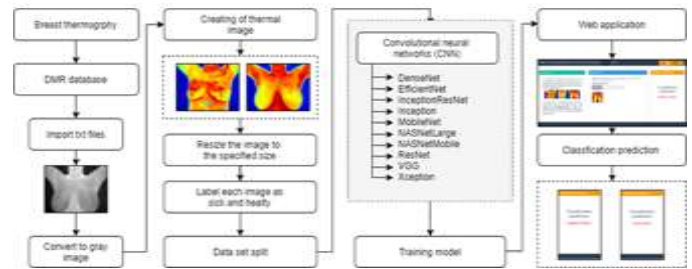


Fig. 3. A general working diagram

### 2.4. An Illustrative Application in the Developed Web-Based Software

A sample image has been uploaded to the system to show the usability of the developed web-based software. The screen showing classification prediction results of the loaded image is given in figure 4 and it was classified as healthy.



Fig. 4. The screen showing classification prediction results

## 3. RESULTS

Before classifying the thermographic breast cancer images, the text files were converted into images for visual analysis. For clear and identifiable images, the image pre-processing phase is critical. For this reason, all images were converted to 640x480 jpg format and taken into modeling. The images obtained during the modeling phase were used to display both breasts in the image without distinguishing the right and left breasts, that is, without fragmenting the images. The results of different pre-trained network models after the data preprocessing stages are given in Table 1.

TABLE I  
ACCURACY RESULTS FROM MODELS

Model name	Validation accuracy
DenseNet121	0.989
DenseNet169	0.995
DenseNet201	0.995
InceptionResNetV2	0.980
InceptionV3	0.984
MobileNet	0.994
MobileNetV2	0.993
NASNetLarge	0.994
NASNetMobile	0.981
ResNet101V2	0.995
ResNet152V2	0.992
ResNet50V2	0.996
VGG16	0.937
VGG19	0.931
Xception	0.991

According to the results in Table 1, an accuracy value above 0.90 was obtained for all models. The confusion matrix of the

ResNet50V2 model, which gives the highest accuracy of 0.996, is given in Figure 5. After ResNet50V2, the models with the highest accuracy values were DenseNet169, DenseNet201, and ResNet101V2.

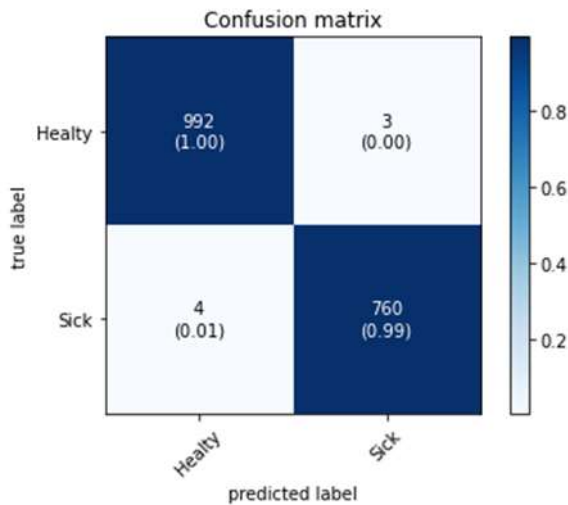


Fig. 5. The confusion matrix of the ResNet50V2 model

The confusion matrices of the next best models are  $\begin{bmatrix} 993 & 6 \\ 3 & 757 \end{bmatrix}$ ,  $\begin{bmatrix} 991 & 9 \\ 5 & 754 \end{bmatrix}$ ,  $\begin{bmatrix} 991 & 4 \\ 5 & 759 \end{bmatrix}$  respectively.

Performance metrics for the ResNet50V2 model are given with 95 percent confidence intervals in Table 2.

TABLE II  
PERFORMANCE METRICS FOR THE RESNET50V2 MODEL

Metrics	Value	Testing value(%) (95%CI)	
		Confidence interval lower limit	Confidence interval upper limit
Accuracy	0.996	0.993	0.999
Sensitivity	0.996	0.990	0.999
Specificity	0.996	0.989	0.999
F1-Score	0.996	0.994	0.999
MCC	0.992	0.988	0.996
G-Mean	0.996	0.993	0.999

#### 4. DISCUSSION

Early identification of breast cancer can help prevent the disease from spreading and becoming lethal. Due to budgetary constraints, lack of convenience, and pain involved with standard screenings such as mammograms, most women neglect the importance of regular breast cancer screening. Thermography can help with some of these issues because it is inexpensive, simple to use, and causes little discomfort to the patient because it is non-invasive and painless. Thermography is a more advanced alternative to mammography, which is known to induce pain and is unsuccessful at detecting breast cancer. The requirement for a mammography exam would be avoided if the thermographic applications could identify cells without cancer. In addition, if the thermogram indicates the presence of a tumor,

mammography can be used to validate or refute the allegation. This option may encourage more women to have their breasts evaluated [3, 22]. For these reasons, there is a need for computer-aided diagnostic systems to be obtained by combining thermography with image processing methods. This study aims to demonstrate the feasibility of a computer-aided system through image processing methods to show the usability of thermography instead of the expensive and painful known methods used in the diagnosis of breast cancer. For this purpose, unlike some previous studies, it is aimed to reduce the training time and increase the model performance by using pre-trained networks for image processing in this study [3]. In previous imaging studies with thermographic breast cancer, the breast cross-section was divided into right and left breasts and processed in the studies. However, in this study, image processing was performed by using images of the chest area to cover both breasts, without distinguishing between right and left breasts [3, 22, 23, 24].

According to the results obtained from the different models used in this study, the highest classification performance was obtained with the ResNet50V2 model as 0.996. When previous studies were examined, the highest accuracy was found to be 0.989 in a study that classified thermographic breast cancer images with CNN [3]. In the breast cancer classification study from thermographic images with another CNN method, the highest accuracy was obtained as 0.90 [22].

Finally, unlike other studies, in this study, a computer-aided diagnosis system was created by developing an interface for breast cancer classification from thermographic images in addition to experimental findings. The web software based on the proposed models has provided promising predictions of breast cancer from thermographic images. The developed software helps breast cancer to be easily diagnosed by medical professionals and other healthcare professionals. Thus, the workload of medical professionals can be reduced and a faster consultation system can be created. With the consultation system created, the patient can be directed to a more advanced diagnosis system which can pave the way for early diagnosis and preventing deadly cancer.

#### REFERENCES

- [1] Brookman-May, S. D., Rodriguez-Faba, O., Langenhuijsen, J. F., Akdogan, B., Linares, E., Minervini, A., ... & Marszalek, M. (2017). Challenges, hurdles and possible approaches to improve cancer care in developing countries—A short breakdown of the status quo and future perspective. *Advances in Modern Oncology Research*, 3(5), 204-212.
- [2] Chakraborty, S., & Rahman, T. (2012). The difficulties in cancer treatment. *ecancermedicalscience*, 6.
- [3] Ekici, S., & Jawzal, H. (2020). Breast cancer diagnosis using thermography and convolutional neural networks. *Medical hypotheses*, 137, 109542.
- [4] Li, T., Sun, L., Miller, N., Nicklee, T., Woo, J., Hulse-Smith, L., ... & Boyd, N. (2005). The association of measured breast tissue characteristics with mammographic density and other risk factors for breast cancer. *Cancer Epidemiology and Prevention Biomarkers*, 14(2), 343-349.
- [5] Zuluaga-Gomez, J., Zerhouni, N., Al Masry, Z., Devalland, C., & Varnier, C. (2019). A survey of breast cancer screening techniques: thermography and electrical impedance tomography. *Journal of medical engineering & technology*, 43(5), 305-322.
- [6] Kandlikar, S. G., Perez-Raya, I., Raghupathi, P. A., Gonzalez-Hernandez, J. L., Dabydeen, D., Medeiros, L., & Phatak, P. (2017). Infrared imaging technology for breast cancer detection—Current status, protocols and new directions. *International Journal of Heat and Mass Transfer*, 108, 2303-2320.

- [7] Cabioğlu, C., & Oğul, H. (2020, May). Computer-Aided Breast Cancer Diagnosis from Thermal Images Using Transfer Learning. In International Work-Conference on Bioinformatics and Biomedical Engineering (pp. 716-726). Springer, Cham.
- [8] Chaves, E., Gonçalves, C. B., Albertini, M. K., Lee, S., Jeon, G., & Fernandes, H. C. (2020). Evaluation of transfer learning of pre-trained CNNs applied to breast cancer detection on infrared images. *Applied Optics*, 59(17), E23-E28.
- [9] Torres-Galván, J. C., Guevara, E., Kolosovas-Machuca, E. S., Ocegüera-Villanueva, A., Flores, J. L., & González, F. J. (2021). Deep convolutional neural networks for classifying breast cancer using infrared thermography. *Quantitative InfraRed Thermography Journal*, 1-12.
- [10] Chartrand, G., Cheng, P. M., Vorontsov, E., Drozdal, M., Turcotte, S., Pal, C. J., ... & Tang, A. (2017). Deep learning: a primer for radiologists. *Radiographics*, 37(7), 2113-2131.
- [11] Nair, V., & Hinton, G. E. (2010, January). Rectified linear units improve restricted boltzmann machines. In *Icml*.
- [12] Albawi, S., Mohammed, T. A., & Al-Zawi, S. (2017, August). Understanding of a convolutional neural network. In 2017 International Conference on Engineering and Technology (ICET) (pp. 1-6). Ieee.
- [13] İnce, Ö., Şenel, İ. K., & Yılmaz, F. (2020). Image Processing and Analysis in Health: Advantages, Challenges, Threats and Examples. *Archives of health science and research (Online)*, 7(1), 66-74.
- [14] Tan, J. H., Ng, E. Y. K., Acharya, U. R., & Chee, C. (2009). Infrared thermography on ocular surface temperature: a review. *Infrared physics & technology*, 52(4), 97-108.
- [15] Ibrahim, A., Mohammed, S., & Ali, H. A. (2018, February). Breast cancer detection and classification using thermography: a review. In *International Conference on Advanced Machine Learning Technologies and Applications* (pp. 496-505). Springer, Cham.
- [16] Grinberg, M. (2018). *Flask web development: developing web applications with python*. " O'Reilly Media, Inc."
- [17] Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., ... & Zheng, X. (2016). Tensorflow: A system for large-scale machine learning. In 12th {USENIX} symposium on operating systems design and implementation ({OSDI} 16) (pp. 265-283).
- [18] Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M. (2019). *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*.
- [19] McKinney, W. (2011). pandas: a foundational Python library for data analysis and statistics. *Python for high performance and scientific computing*, 14(9), 1-9.
- [20] Van Der Walt, S., Colbert, S. C., & Varoquaux, G. (2011). The NumPy array: a structure for efficient numerical computation. *Computing in science & engineering*, 13(2), 22-30.
- [21] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *the Journal of machine Learning research*, 12, 2825-2830.
- [22] Zuluaga-Gomez, J., Al Masry, Z., Benaggoune, K., Meraghni, S., & Zerhouni, N. (2021). A CNN-based methodology for breast cancer diagnosis using thermal images. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 9(2), 131-145.
- [23] Baykara, M. (2021). Performance Analysis of Various Classification Algorithms for Computer-Aided Breast Cancer Diagnosis System Using Thermal Medical Images, *Turkish Journal of Science and Technology*, vol. 16, no. 1, pp. 65-84,
- [24] Kaya, M. O. (2021). Computer-Aided Model For The Classification Of Acute Inflammations Via Radial-Based Function Artificial Neural Network. *The Journal of Cognitive Systems*, 6(1), 1-4.

**Zeynep Küçükakçalı** obtained her BSc. degree in mathematics from Çukurova University in 2010. She received MSc. degree in biostatistics and medical informatics from the Inonu University in 2018. She currently continues Ph.D. degrees in biostatistics and medical informatics from the Inonu University. In 2014, she joined the Department of Biostatistics and Medical Informatics at Inonu University as a researcher assistant. Her research interests are cognitive systems, data mining, machine learning, deep learning.

## BIOGRAPHIES

**Hasan Ucuzal** obtained her BSc. degree in software engineering from Firat University. He received MSc. degree in biostatistics and medical informatics from the Inonu University in 2020. He is currently working in the information processing department at Inonu University. His research interests are cognitive systems, data mining, image processing, machine learning, deep learning.

**Muhammet Baykara** was born in Elazig, Turkey. He received his BS and MSc. in Computer Engineering from Firat University in 2006, 2009 respectively. He received his Ph.D. in Software Engineering from Firat University in 2016. Currently, he is an assistant professor in the Department of Software Engineering at Firat University. His research interests are Information Security, Honeypots, Intrusion Detection and Prevention Systems, image processing and deep learning.