Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 10, No. 01, March 2024.

USING RESNET18 IN A DEEP-LEARNING FRAMEWORK AND ASSESSING THE EFFECTS OF ADAPTIVE LEARNING RATES IN THE IDENTIFICATION OF MALIGNANT BREAST MASSES IN MAMMOGRAMS

Soumia Benbakreti[1], Samir Benbakreti[2], Kadda Benyahia[3] and Mohamed Benouis[4]

(Received: 12-Nov.-2023, Revised: 8-Jan.-2024, Accepted: 25-Jan.-2024)

ABSTRACT

Breast cancer is a prevalent disease that primarily affects women globally, but it can also affect men. Early detection is crucial for better treatment outcomes and mammography is a common screening method. Recommendations for mammograms vary by age and country. Early breast-cancer screening is vital for timely interventions. This paper aims to introduce artificial-intelligence methods through deep-learning approaches utilizing pre-trained CNN-based models for the diagnosis of masses depicted in breast images. These masses may be either malignant or benign, necessitating distinct management strategies for each scenario. The experiments conducted on pre-trained models (AlexNet, InceptionV3 and ResNet18) are designed to underscore the significance of selecting the batch size and adaptive learning rate in influencing the results, ultimately facilitating a notable enhancement in classification rates. Pre-trained models applied to a merged dataset comprising three datasets (Inbreast+MIAS+DDSM) yielded an accuracy of 93.7% for InceptionV3 and 88.9% for AlexNet. However, the most favorable outcome was observed with ResNet18, achieving an accuracy of 95% (with precision, recall and F1-score of 94.90%, 94.91% and 94.91%, respectively).

KEYWORDS

Malignant breast cancer, Lesion classification, Transfer learning, Residual network (ResNet18), Adaptive learning rate.

1. INTRODUCTION

Breast cancer remains a significant concern in global public health, impacting millions of women annually and profoundly affecting their well-being. In 2018, the International Agency for Research on Cancer documented a global 18,078,957 cancer cases. Among these cases, breast cancer constituted 2,088,849 instances, accounting for 11.6% of all cancers and ranking as the second most prevalent type. Notably, breast cancer emerged as the most frequently diagnosed cancer among women, making up 24.2% of the cases [1]. Despite substantial progress in medical research, breast cancer poses ongoing complexities in terms of early detection, effective treatment and long-term management. This multifaceted disease exhibits considerable variations in clinical presentations, disease progression and responses to therapies [2]. Consequently, gaining a deeper understanding of its underlying mechanisms, risk factors and innovative diagnostic and treatment approaches is essential for enhancing the clinical outcomes of breast-cancer patients. CNNs have played a crucial role in enhancing breast-cancer diagnosis [3][4][5]. CNN models, such as ResNet and Inception, have been adapted for the analysis of mammographic images, enabling more accurate detection of suspicious masses and micro-calcifications. There are several methods for diagnosing breast cancer, including mammography [6], which is commonly used in women over 40 clinical-breast examination [7] performed by a healthcare professional to detect anomalies, breast ultrasound that uses sound waves to visualize breast tissues [8] and breast MRI, which is more sensitive than mammography and often used in high-risk women [9]. Biopsy, involving the collection of a tissue sample for laboratory analysis, remains the most precise method for diagnosing breast cancer and determining its type [10]. Additionally, a sentinel lymph node biopsy may be necessary to assess the spread of cancer to lymph nodes. For women with a family history of breast cancer, genetic tests, such as BRCA1 and BRCA2 mutations, can be performed [11]. Finally, in some cases, advanced imaging tests, like CT scans or bone scans, are needed to evaluate the extent

1. S. Benbakreti is with Laboratory of Mathematic, University of Djillali Liabes, Sidi Bel Abbes, Algeria. Email: souben2223@gmail.com
2. S. Benbakreti is with ENSTTIC, Department of Speciality, Street of Senia, Oran, 31000, Algeria. Email: samir.benbakreti@ensttic.dz
3. K Benyahia is with LTC Laboratory, University of Tahar Moulay, Saida, Algéria. Email: benyahiak@gmail.com
4. M Benouis is with University of Augsburg, Augsburg, Germany. Email: mhbenouis@gmail.com

of cancer [12]. The choice of method depends on individual factors and should be discussed with a healthcare professional. On the other hand, the classification of breast tumors into malignant (cancerous) and benign (non-cancerous) categories is a critical task in both medical imaging and oncology. This classification presents specific challenges, notably the visual similarity between certain malignant and benign tumors in medical images (see Figure 1), which complicates their differentiation.

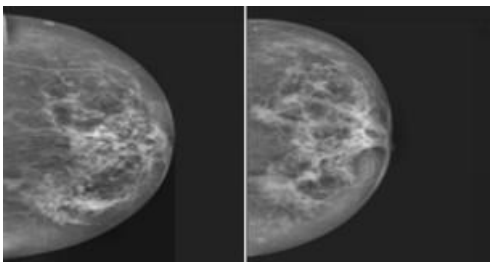


Figure 1. a- Benign, b- Malignant.

Ultimately, our aim is to contribute to the improvement of patient care for those grappling with breast cancer by showcasing the latest scientific and clinical breakthroughs, while also underscoring the persistent challenges warranting specific focus within the medical and research communities. This research focused on constructing a convolutional neural network (CNN) model to differentiate between benign and malignant breast-cancer tumors. Notably, the study emphasizes three key contributions. Firstly, the careful selection of the batch size is highlighted for its substantial impact on image-classification tasks within a CNN. Secondly, the adoption of an adaptive learning rate is advocated to address issues, such as model stability and overfitting. Lastly, the integration of three reference datasets (Inbreast + DDSM + MIAS) is underscored, enhancing the model's adaptability to diverse dataset characteristics and resulting in a more robust and efficient classification model.

The structure of the present document is as follows: An introduction is presented in Section 1 and research works on the early identification of breast cancer are discussed in Section 2. The proposed model and dataset are described in Sections 3 and 4. Section 5 details the experiments performed on the different architectures and the conclusions that were reached.

2. RELATED WORKS

Artificial intelligence (AI) is transforming the medical landscape through its provision of advanced diagnostic solutions. In brain-cancer early detection [26], AI systems meticulously analyse medical images to identify initial signs, thereby enhancing the chances of successful treatment. In neurological diagnostics, AI adeptly interprets brain images, simplifying the detection of diseases, like Alzheimer's [37]. Additionally, AI algorithms analyse ECG data for cardiac assessment [38], enabling the diagnosis of cardiac issues and facilitating preventive interventions. In [39], the authors employ deep-learning algorithms to analyse fundus images, distinguishing early signs and allowing differentiation between multiple ocular diseases. In [42], a system integrating advanced deep-learning networks (EfficientNet, Xception, MobileNetV2, InceptionV3 and Resnet50) through the innovative method of adaptive consensus weighting (CAW) was proposed. The method enables dynamic adjustment of multiple deep networks, thereby enhancing the system's detection capabilities. Evaluations on various datasets, including DDSM, demonstrated quite interesting performances. On the other hand, the study of [43] explored the potential of artificial intelligence (AI) and machine learning (ML) in providing precise diagnoses, enhancing outcome predictions and recognizing disparities in the treatment of CLTI (Chronic Limb-Threatening Ischemia). The authors underscored the importance of AI/ML approaches in patient management and highlighted how existing data could be leveraged for computer-guided interventions. These applications underscore the growing significance of AI in medical diagnosis, paving the way for swifter and personalized treatments. Subsequently, we will delve into advancements related to breast cancer in this context.

In [13], the most recent machine learning-based models for detecting and classifying breast cancer are examined through a comparative study. Additionally, a compilation of widely accessible and well-received datasets is offered to facilitate future experiments and comparisons. The results of the analysis

highlight You Only Look Once (YOLO) and RetinaNet as the most accurate recent models for both detection and classification. But, the YOLO model can be tailored to specific requirements; however, it does exhibit limitations when it comes to handling closely positioned objects. In particular, it struggles to achieve high accuracy with small-sized objects and may occasionally make errors in localization. Similarly, RetinaNet has certain shortcomings, like model overloading and complexity, which make training so difficult. The study of Ismail & Sovuthy made a comparative analysis of breast-cancer detection using two deep-learning model networks. The entire process encompasses image preprocessing, classification and performance assessment [14]. Specifically, the performance of two deep-learning model networks, VGG16 and ResNet50, is assessed for distinguishing between normal and abnormal tumors utilizing the IRMA dataset. The findings indicate that VGG16 outperforms ResNet50 in terms of accuracy, achieving a 94% accuracy rate compared to ResNet50's 91.7%. Although the study was conducted on a database containing a very limited number of samples (1515 images), the results obtained are excellent, but cannot be generalized. In article [15], the authors aim to conduct a comprehensive examination of machine-learning techniques and their practical applications in diagnosing and predicting outcomes for breast cancer in the context of British Columbia. They begin by offering an overview of various ML techniques, including artificial neural networks (ANNs), support vector machines (SVMs), decision trees (DTs) and k-nearest neighbors (k-NNs). Subsequently, they apply these techniques specifically to the breast cancer scenario in British Columbia, utilizing primary data sourced from the Wisconsin Breast Cancer Database (WBCD) as a benchmark for comparing results obtained through different algorithms. Lastly, they also introduce a healthcare-system model derived from their recent-research endeavors. However, traditional-machine learning techniques sometimes need human-feature extraction and have difficulties in processing complex, high-dimensional data. In the provided framework, features are extracted from images using pre-trained CNN architectures, specifically GoogLeNet, Visual Geometry Group Network (VGGNet) and Residual Networks (ResNet). These extracted features are then input into a fully connected layer for classifying malignant and benign cells, employing average pooling for classification. To assess the effectiveness of this framework, experiments were conducted on well-established benchmark datasets. The results demonstrated that the proposed framework achieves an impressive accuracy rate of 97.52% [16]. However, we will note that the framework requires a lot of hardware resources, which are sometimes lacking. The study of [17] introduces an all convolutional-network method for categorizing screening mammograms, surpassing previous approaches. When tested on digitized film mammograms (CBIS-DDSM), the best model achieved an AUC of 0.88 and combining four models improved the AUC to 0.91. Similarly, on a distinct set of full-field digital mammography images (INbreast database), the top model achieved an AUC of 0.95 and averaging results from four models raised the AUC to 0.98.

Furthermore, the research demonstrates that a classifier trained using their method on CBIS-DDSM mammograms can be successfully adapted to INbreast FFDM images with minimal additional data for fine-tuning. These findings highlight the potential of deep-learning methods in enhancing clinical tools and reducing errors in mammography screening. Wang et al. proposed a strategy that utilizes SqueezeNet with fire modules and a complex bypass to extract informative features from mammography images [34]. Subsequently, these extracted features are employed to train a Support Vector Machine (SVM) for the classification of mammography images. The model, known as SNSVM, which integrates SqueezeNet guidance with SVM, exhibited promising results when tested on the MIAS dataset, achieving an accuracy of 94.10%. In [35], the authors devised a novel framework for breast-cancer diagnosis, employing entropy-controlled deep learning and flower-pollination optimization based on mammogram images. Within this proposed framework, a contrast-enhancement method is developed using filter fusion. The pre-trained ResNet50 model undergoes refinement and training through transfer learning on both the original and enhanced datasets, employing various data augmentation techniques. While the results obtained from individual datasets appear highly promising, it's worth noting that the pretrained model, due to its depth, can be resource-intensive during processing. Similar research has been conducted on ResNet101 [36], incorporating feature fusion through the proposed highly corrected function-controlled canonical-correlation analysis approach and optimal feature selection using the Newton-Raphson algorithm controlled by Satin Bowerbird Optimization. The experiments within the devised framework were carried out utilizing the CBIS-DDSM dataset, yielding a top accuracy of 94.5%. Han et al. introduced a novel deep-learning model for the multi-classification of breast cancer. The structured deep learning model has delivered

impressive results, achieving an average accuracy rate of 93.2% on a substantial dataset [18]. Nevertheless, we believe that the exploitation of this deep-structured model can be improved, as well as the results obtained.

3. PROPOSED MODEL

In this research, we will investigate three distinct approaches for detecting masses within breast images. These pre-trained neural network architectures, specifically drawn from CNN [19], include AlexNet, ResNet18 and InceptionV3. We selected these three models considering that resource consumption varies among pre-trained models based on their complexity and parameter count. Typically, deeper and more complex models, like ResNet152 [40] or advanced Inception variants [41], may demand more resources in terms of GPU memory and computing power for training and inference. Thus, our choice aimed to find a compromise between performance and resource consumption. In this study, transfer learning is implemented with deep-learning models to detect suspicious masses in mammograms. The deep-learning model is characterized as a pre-trained model, trained using a Convolutional Neural Network, with features generated through transfer learning. The proposed model is shown in Figure 2.

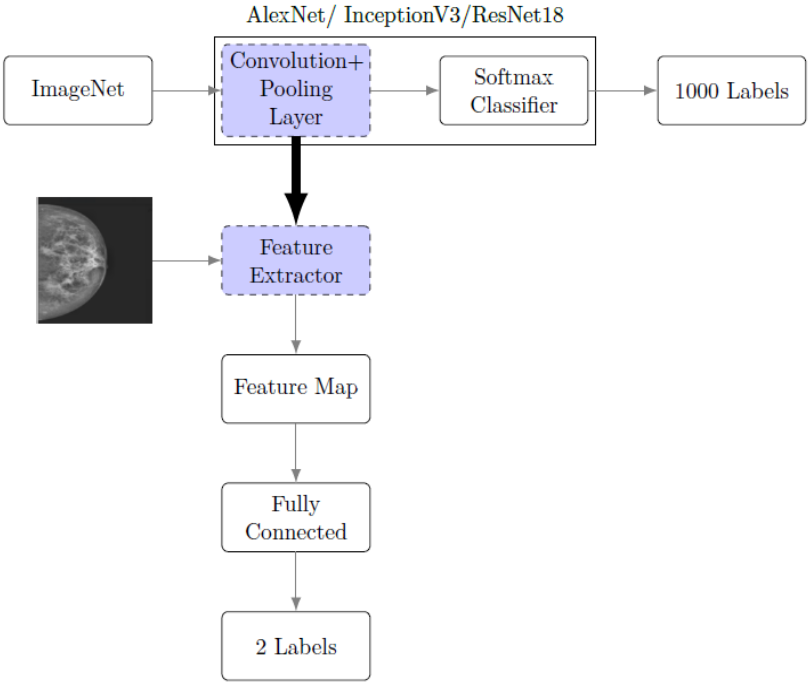


Figure 2. Proposed model.

Utilizing the extracted features, a fully connected layer is employed to construct the final model. The breast-mass identification scheme is illustrated in Figure 2. During the transfer learning and retraining of the deep-learning model, the extracted features are employed to train the new model. The input comprises images and all convolutional and pooling layers are reused in the training of the new model. In essence, ResNet18, InceptionV3 or AlexNet serves as a feature extractor, subsequently transmitted to an FC layer.

3.1 AlexNet

AlexNet is an eight-layer convolutional neural network [20] composed of five convolutional layers and three fully-connected layers. The first two convolutional layers are followed by a normalization layer and a max-pooling layer, while the third and fourth layers are directly connected. The fifth convolutional layer is succeeded by a max-pooling layer. The resulting output undergoes a sequence of two fully-connected layers, where the second fully-connected layer contributes to a Softmax classifier. AlexNet uses ReLU as the activation function, differing from the conventional sigmoid and tanh functions used in previous neural networks. ReLU is a non-saturating activation function that not only significantly accelerates the model's training time, but also more effectively addresses the issues

of gradient disappearance and explosion, making it simpler to train a deeper network. The standard input size for the AlexNet model is 227×227×3. Table 1 summarizes the AlexNet architecture.

Table 1. The AlexNet architecture.

Model	Alexnet	
	Filter Size/Stride	Output Size
Conv1	11×11/4	96×55×55
Pool1	3×3/2	96×27×27
Conv2	5×5/1	256×27×27
Pool2	3×3/2	256×13×13
Conv3	3×3/1	384×13×13
Conv4	3×3/1	384×13×13
Conv5	3×3/1	256×13×13
Pool5	3×3/2	256×6×6
Fc5	-	4096
Fc6	-	4096
Fc7	-	1000

3.2 InceptionV3

InceptionV3 contains 50 layers and is a convolutional neural network (CNN) [21]. The algorithm, named "going deeper with convolutions" was developed and trained by Google. The development of the Inception module, which consists of a series of 1-by-1 convolutional layers/blocks used for dimensionality reduction and feature aggregation, is the key component of GoogleNet/Inception architecture. This model had 9 inception modules and a total of 22 layers. Up to 1000 objects can be classified using the pre-trained version of InceptionV3 with the ImageNet dataset weights [22]. This network's image input was 299×299 pixels in size.

The InceptionV3 stands out by incorporating a crucial element known as an inception module. This module utilizes receptive kernels of various sizes, maintaining a consistent output size for the convolution operation through the use of zero padding. The ultimate feature maps are obtained by concatenating the filters. The inception operation plays a significant role in extracting more comprehensive features from the input image. Refer to Figure 3 for a visual representation of this module.

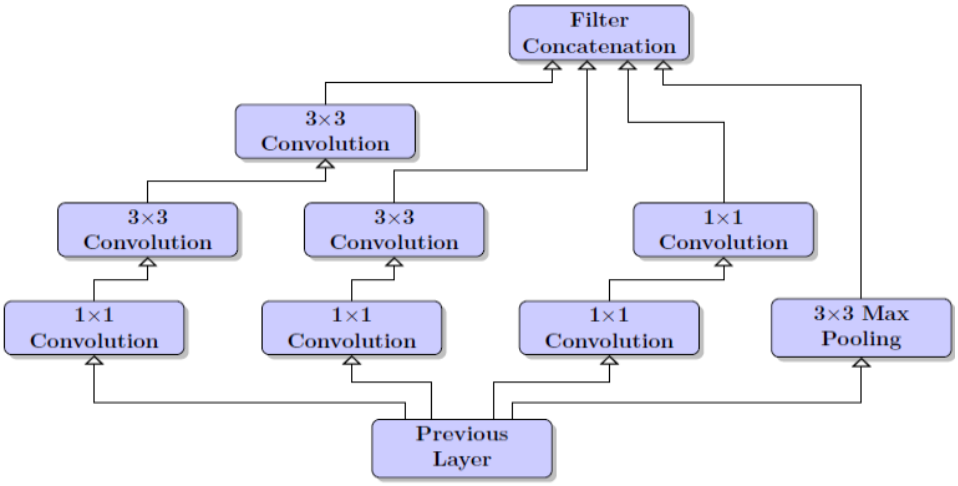


Figure 3. Representation of inception module.

3.3 ResNet18

A convolutional neural network with 18 layers in depth is called ResNet18. Deep Residual Learning for Image Recognition, as it is known, was developed and trained by Microsoft in 2015 ([23]). To address the issue of vanishing gradient that may affect the weightage change in neural networks, ResNet architectures

introduced the use of residual layers and skip connections. This made training easier and allowed neural networks to get much deeper with greater performance. More than a million images from the ImageNet dataset were used to train this model. The network was trained on colored images with a resolution of 224×224 pixels and can categorize up to 1000 objects.

Given its comparatively modest architecture, ResNet18 was selected for its balanced trade-off between depth and performance. It comprises a 7×7 convolutional layer, 2 pooling layers, 5 residual blocks and an FC layer. Each residual block is composed of two 3×3 convolutional layers, followed by a batch-normalization layer and a ReLU activation function (see Table 2).

The network achieves high classification accuracy by employing bottleneck residual blocks, batch normalization for adjusting input layers and identifying connections to mitigate the risk of vanishing gradients.

Figure 4 illustrates the architecture of the ResNet18 network, detailing the configuration of the employed residual blocks (Res Block2 represents a ResNet block with a 1×1 convolution). The notation "FC" denotes a fully-connected layer with two outputs corresponding to malignant and benign classifications. Table 3 gives a summary of the pre-trained models used in this study and their main characteristics.

Table 2. ResNet18 architecture.

Layer Name	Output Size	Description
conv1	112×112×64	7×7, 64, stride 2
conv2_x	56×56×64	3×3 max pool, stride 2
		$\begin{bmatrix} 3\times 3, 64 \\ 3\times 3, 64 \end{bmatrix} \times 2$
conv3_x	28×28×128	$\begin{bmatrix} 3\times 3, 128 \\ 3\times 3, 128 \end{bmatrix} \times 2$
conv4_x	14×14×256	$\begin{bmatrix} 3\times 3, 256 \\ 3\times 3, 256 \end{bmatrix} \times 2$
conv5_x	7×7×512	$\begin{bmatrix} 3\times 3, 512 \\ 3\times 3, 512 \end{bmatrix} \times 2$
average pool	1×1×512	7× 7 average pool
fully connected	1000	512×1000 fully connections
softmax	1000	

Table 3. Summary of pre-trained models.

Pre-trained Model	Depth	Parameters (Millions)	Image Input	Specific feature
AlexNet	8	60	227×227	Deeper
ResNet18	18	11.7	224×224	Residual block
InceptionV3	50	23.9	299×299	Inception module and size kernel

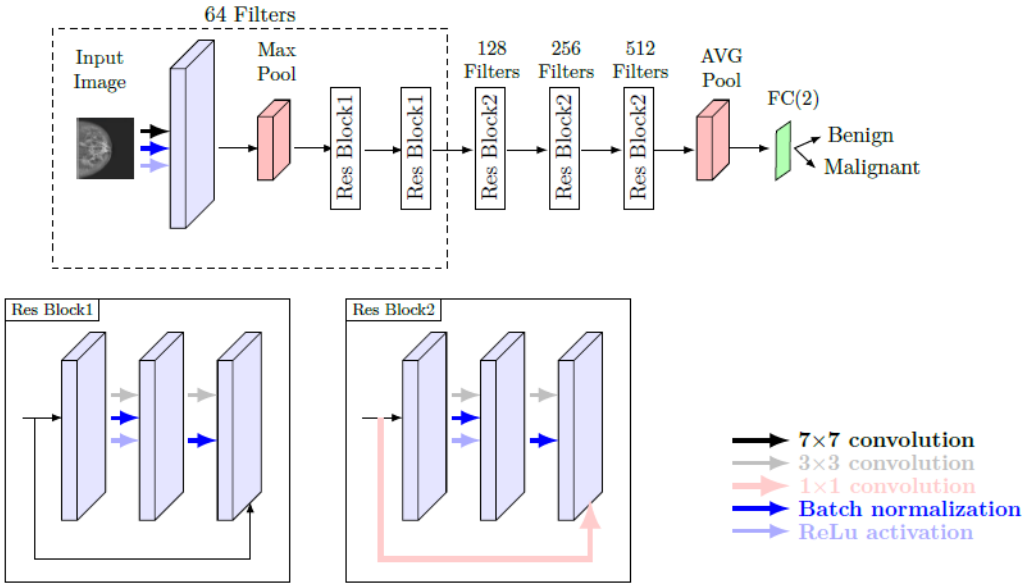


Figure 4. ResNet18 architecture used in the proposed model.

4. The Used Dataset

The dataset comprises mammograms featuring both benign and malignant masses, resulting from the amalgamation of several datasets [24]. Specifically, it incorporates 7,632 images from the Inbreast dataset, 3,816 images from the MIAS dataset and 13,128 images from the DDSM dataset. All of these images have been uniformly resized to dimensions of 227×227 pixels. The total number of images is 24576. Table 4 displays details about images from the datasets that were used. Matlab 2021 was used for the training. It was installed on a workstation running Windows 10 Pro, 64-bit, with 24 GB of RAM, an Intel(R) Xeon(R) CPU E5-2620 v3 @ 2.40GHz and a NVIDIA Quadro K620 GPU. The majority of the dataset, 80%, is used for training, while 20% is utilized for testing. We refrained from employing data augmentation for two primary reasons.

Firstly, we harnessed the power of transfer learning [25]. Our models were originally trained on extensive datasets encompassing a diverse range of images, such as ImageNet [22]. Consequently, these pretrained models served as our initial framework for the task at hand, which involved classifying breast masses. Transfer learning is especially advantageous when dealing with limited data, as was the case with our dataset, as it allows the model to leverage the knowledge it acquired during previous training. Secondly, we considered existing literature, which suggested that data augmentation does not universally guarantee improved performance, particularly when the introduced transformations lack relevance [26].

Table 4. The used datasets (Inbreast+MIAS+DDSM).

Dataset	Total number of images	Benign	Malignant
Inbreast	7632	2520	5112
MIAS	3816	2376	1440
DDSM	13128	5970	7158
Inbreast+MIAS+DDSM	24576	10866	13710

5. Results and Discussion

To evaluate the performance of our models, we used the following metrics:

$$Accuracy = \frac{TP+TN}{TP+FN+TN+FP} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$\text{F1} - \text{score} = \frac{2 * Precision * Recall}{Precision + Recall} \quad (4)$$

Where:

- TP (True Positive): It represents the number of correctly predicted positive samples by the model. This means that the model has correctly identified these samples as positive.
- TN (True Negative): It represents the number of correctly predicted negative samples by the model. This means that the model has correctly identified these samples as negative.
- FP (False Positive): It represents the number of negative samples incorrectly predicted as positive by the model. This means that the model has identified these samples as positive when they are actually negative.
- FN (False Negative): It represents the number of positive samples incorrectly predicted as negative by the model. This means that the model has identified these samples as negative when they are actually positive.

The initial parameters of the three architectures are mentioned in Table 5.

Table 5. The training parameters of the pre-trained models.

Parameter	**Value**
Initial learning rate	0.001
Optimizer	SGDM
Max. epoch	1
Mini-batch size	20
Activation function	Softmax
Validation frequency	10

When:

- Initial learning rate: It is the initial value of the learning rate in a machine-learning algorithm, especially in neural networks.
- SGDM: the stochastic gradient descent with momentum solver [27].
- Mini-batch size: The stochastic gradient descent technique uses a portion of the training data for each iteration to evaluate the gradient and update the parameters. At each iteration, a distinct subset, known as a mini-batch, is used to assess the gradient of the loss function and update the weights.
- Max. epoch: It is the maximum number of epochs for training.
- Activation function: Transfer functions calculate a layer's output from its net input.
- The validation-frequency value indicates how many iterations there are between evaluations of validation metrics.

The results obtained are shown in Table 6.

Table 6. The obtained results.

Models	**The used metrics**			
	Accuracy (%)	**Precision** (%)	**Recall** (%)	**F1-score** (%)
AlexeNet	83.5	85.5	82	83.72
InceptionV3	**89.8**	**90.86**	**88.87**	**89.86**
ResNet18	88.24	88.13	88	88.07

As depicted in Table 6, it is evident that the three models have produced satisfactory outcomes, with InceptionV3 emerging as the top performer (89.8 % of accuracy). To achieve a more comprehensive and accurate evaluation of the models' performance, we have incorporated supplementary metrics, including

precision, recall and F1-score, alongside accuracy. This approach proves to be particularly valuable in the context of imbalanced class distribution within this scenario.

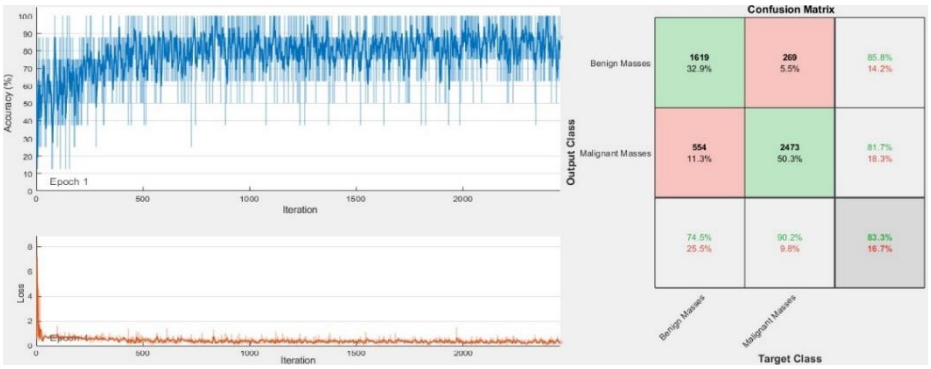
5.1 Experiment 1: Impact of Batch Size on Outcomes

In this experiment, we will explore how altering the batch size influences the results of classification. These results are presented in Table 7.

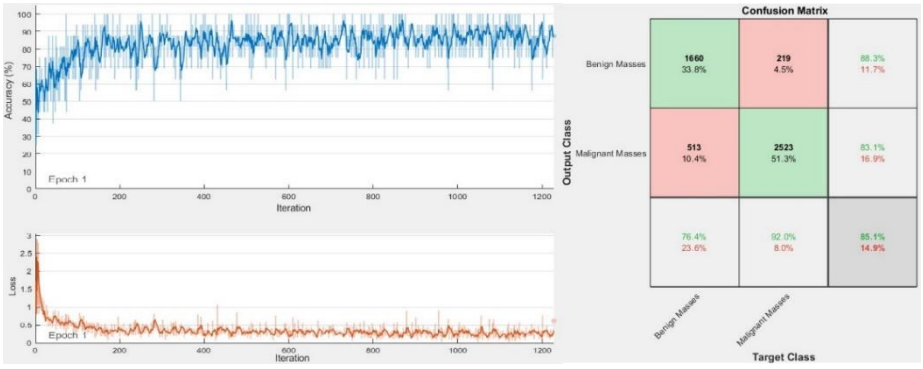
Table 7. The obtained results showing the impact of batch size.

Batch size	Models	The used metrics			
		Accuracy (%)	**Precision** (%)	**Recall** (%)	**F1-score** (%)
8	AlexeNet	80.2	83.5	87.2	80.75
	InceptionV3	**83.3**	**83.73**	**82.35**	**83.03**
	ResNet18	83.1	23.95	83.4	83.17
16	AlexeNet	82.4	82.66	82.73	82.19
	InceptionV3	**85.1**	**85.72**	**84.20**	**84.96**
	ResNet18	84.8	84.78	84.35	84.56
32	AlexeNet	85.1	84.87	84.84	84.86
	InceptionV3	90.89	90.76	91.21	91.02
	ResNet18	**91.8**	**91.64**	**91.74**	**91.74**
64	AlexeNet	86	85.94	86.42	86.18
	InceptionV3	88.5	88.64	89.14	88.89
	ResNet18	**91.6**	**91.47**	**92**	**91.73**

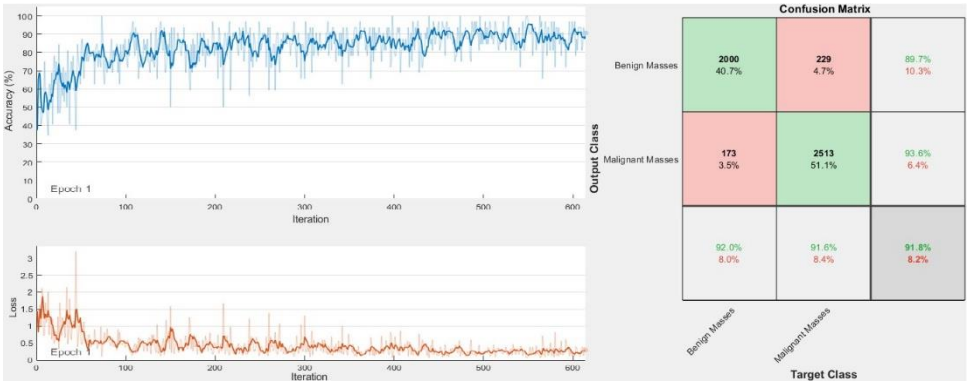
Upon reviewing the curves in Figure 5, we can observe that, apart from the notably quicker convergence seen with a smaller batch size, a larger batch size tends to diminish noise in weight updates (Figure 5(d)). This reduction in noise can enhance training stability and reduce vulnerability to random fluctuations, which proves advantageous, particularly when dealing with noisy datasets or unstable gradients. However, it is important to note that increased stability may not always result in overall improved performance. Conversely, a smaller batch size introduces noise into weight updates, effectively aiding in the prevention of overfitting (Figure 5(a)) [28].



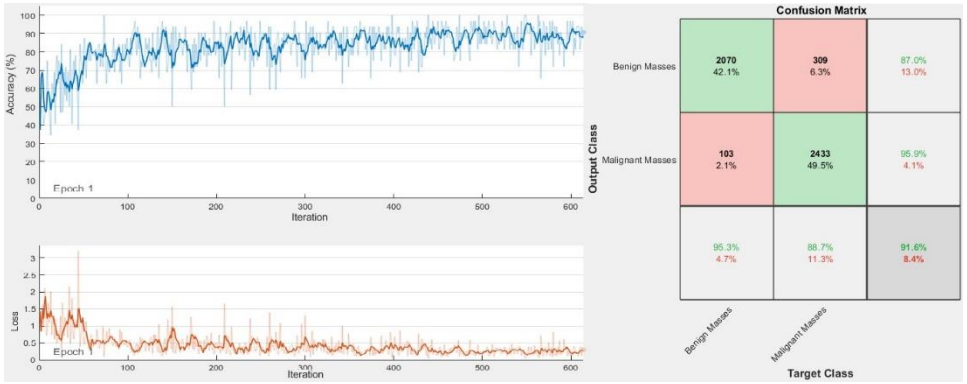
(a)



(b)



(c)



(d)

Figure 5. (a) Best result for breast-mass classification (InceptionV3) with BS=8. (b) Best result (InceptionV3) with BS=16, (c) Best result (ResNet18) with BS=32, (d) Best result (ResNet18) with BS=64.

Overfitting, a situation where the model overly adapts to the training data and struggles to generalize to validation or test data, is mitigated by the use of a smaller batch size, albeit at the cost of requiring more training iterations. Furthermore, it is worth mentioning that the selection of an appropriate batch size is intricately connected with other model hyper-parameters, such as the learning rate.

This often necessitates an iterative exploration of these hyper-parameters to identify the optimal combination for a specific classification task [29]. The ResNet18 model with a batch size of 32 demonstrated the highest performance in building a transfer-learning model for classifying mammogram-detected breast lesions. Nevertheless, there is not a universally ideal batch size; typically, it necessitates experimentation to determine the best configuration for a specific problem.

5.2 Experiment 2: Impact of Adaptive Learning Rate

In this experimental setup, we maintained the parameters as detailed earlier in Table 5 in particular the same optimizer: SGDM, with the exception of adjusting the batch size to 32 (yielding the best result) and modifying the number of iterations. This modification was made to specifically investigate the influence of the adaptive learning rate. We will utilize a variable learning rate that adapts throughout the training process within a pre-trained model, employing an adaptive learning rate adjustment method. This can enhance the convergence of the learning process.

Table 8. The additional training parameters of the pre-trained models.

Parameter	Value
Max. epoch	4
Learning-rate schedule	Piecewise
Learning-rate drop period	1
Learning-rate drop factor	0.1

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 10, No. 01, March 2024.

In Table 8, the 'LearnRateSchedule' is configured as 'piecewise,' indicating the use of an adaptive learning rate. The 'LearnRateDropPeriod' determines the frequency of learning-rate adjustments, for instance, every 1 epoch and the 'LearnRateDropFactor' specifies the magnitude of reduction at each adjustment, set at 0.1. The classification results of the three pre-trained models are mentioned in Table 9.

Table 9. Adaptive learning-rate results.

Models	The used metrics			
	Accuracy (%)	**Precision** (%)	**Recall** (%)	**F1-score** (%)
AlexeNet	88.9	88.60	88.73	88.86
InceptionV3	93.7	93.55	93.5	93.57
ResNet18	**95**	**94.90**	**94.91**	**94.91**

When we compare Tables 7 and 9, it becomes evident that the implementation of adaptive learning rates has notably enhanced the results for all three models. Nevertheless, ResNet18 continues to stand out as the network that achieved the highest classification rate, reaching 95%. (see Figure 6).

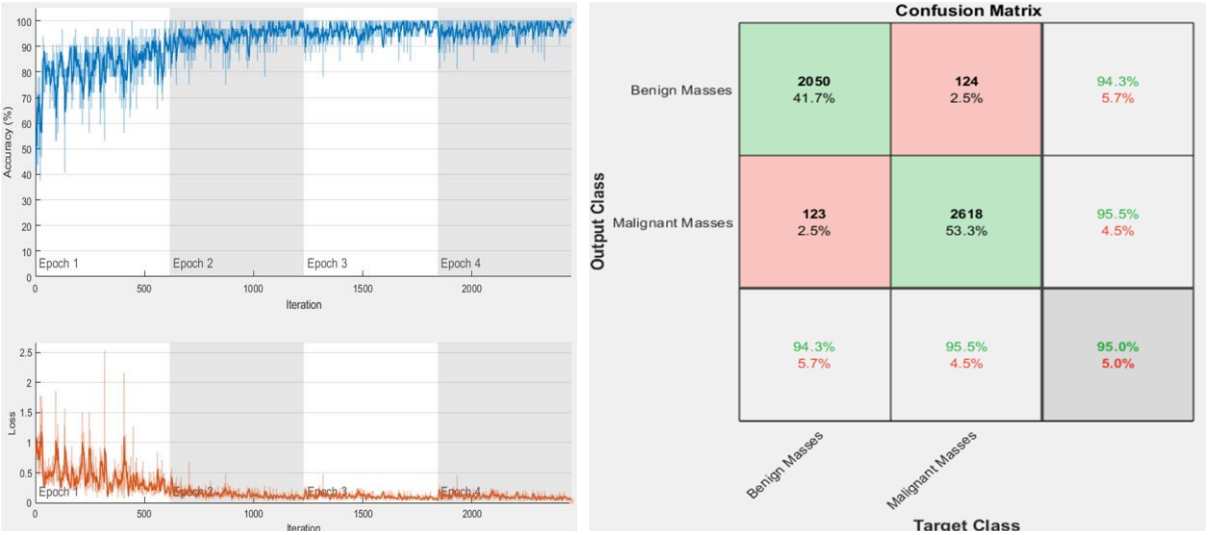


Figure 6. Best result for breast mass classification (ResNet18) with an adaptive learning rate.

Comparing Figure 5(d) and Figure 6 reveals that an adaptive learning rate offers enhanced stability, whereas a fixed learning rate can lead to issues like divergence or learning stagnation.

The adjustment of the learning rate based on model performance helps prevent these problems and ensures a stable learning process. Adaptive learning rates increase when the model is distant from convergence, facilitating exploration of the search space. They subsequently decrease the learning rate as convergence nears, allowing for more precise exploration around the optimal solution. Another noteworthy advantage is that adjusting the learning rate can prevent the model from overfitting the training data, thereby improving its capacity to generalize to new data. Additionally, this approach reduces the need for manual fine-tuning of the learning rate, which can be both labor-intensive and intricate.

Acknowledging the phenomenon where gradients diminish as convolutional neural networks deepen, ResNet addresses this challenge through shortcut connections, allowing the network to achieve greater depths. Moreover, ResNet networks provide the advantage of deep scaling, enabling the adjustment of architecture to the specific classification task at hand. This adaptability results in outstanding classification performance across various levels of image detail, including finer details, underscoring effectiveness in the analysis of medical images.

Table 10 provides a comparison of recent studies focusing on mammogram analysis. Our model stands out with a superior classification rate. In [30], the utilization of machine-learning algorithms raises concerns about potential divergence when dealing with a substantial volume of images. The authors of [32] employ a restricted sub-set of the Inbreast dataset (200 images), potentially resulting in a model tailored to the training data, but lacking robust generalization to new data. While surpassing the

achievements of [34] and [36], it is important to note that ResNet101 may pose challenges due to its resource-intensive nature. Moreover, the majority of contributions in the literature, as stated in the state of the art, use a single dataset. Our goal in summing the three datasets was to make the task more difficult. As shown in Table 10, the suggested model performs well, especially when the comparison is consistent [33], as a result of the authors' combination of the three datasets (MIAS+Inbreast and DDSM).

Table 10. Comparison between our model and related works.

Reference	Year	Used Dataset	Classification model	Accuracy
Zhang et al. [30]	2020	DDSM	SVM, Naïve Bayes, KNN	90.91%
Ting et al. [31]	2019	MIAS	CNNI+BCC	90.5%
Chougrad et al. [32]	2018	Inbreast (200 images)	VGG16	95%
Khartiga et al [33]	2022	Inbreast+DDSM+MIAS	CNN	92.27%
Wang et al [34]	2023	MIAS	SNSVM	94.1%
Fatima et al [36]	2023	CBIS+DDSM	ResNet101	94.5%
Proposed model	**2023**	**Inbreast+DDSM+MIAS**	**ResNet18**	**95%**

6. CONCLUSIONS

The goal of this research was to construct a convolutional neural network (CNN) model to distinguish between benign and malignant breast-cancer tumors. The used pre-trained models are AlexNet, InvceptionV3 and ResNet18. After training on 80% of the dataset and testing on 20%, ResNet18 produced best classification result with an accuracy of 95%. The effectiveness of this model can be attributed to the incorporation of residual connections in ResNet18, which effectively addresses the issue of vanishing gradients during the training of deep networks. Furthermore, ResNet18 strikes a well-balanced equilibrium between classification performance and model complexity. It maintains a leaner profile compared to deeper architectures like ResNet50 while still delivering commendable performance. We also note that InceptionV3 achieved very good results (93.7%). This can be attributed to InceptionV3's architecture, which incorporates techniques, such as regularization and normalization, to reduce overfitting. Consequently, it minimizes the risk of over-adapting to the training data and improves generalization to new data. In addition, InceptionV3 employs factorized convolution, which reduces the number of parameters while maintaining high performance. This enhances its efficiency in terms of memory usage and computational resources. In summary, InceptionV3 is highly regarded for its ability to combine outstanding performance with a relatively lightweight architecture, making it a versatile choice for a variety of image-processing applications. We have also shown that adaptive learning rates enhance the efficiency, stability and performance of deep-learning models, making them a valuable asset in the training of neural networks, particularly CNN-based models.

In addition to investigating the impact of batch size and adaptive learning rate on the outcomes, this study provides insights into several key aspects. Firstly, it challenges the common notion that preprocessing and data augmentation are indispensable for achieving satisfactory results, suggesting that, under certain circumstances, these steps may not be imperative. Secondly, the integration of three reference datasets not only introduces a novel challenge, elevating the complexity of the classifier, but also opens avenues for exploring innovative approaches to enhance results. This approach aims to pave the way for more effective models that can truly serve as robust decision-support tools in real-world scenarios. Diagnostic-aid applications, especially those focused on identifying benign masses, prove to be invaluable tools in the medical domain. Their effectiveness is evident in their swiftness, precision and capacity to handle extensive datasets, such as mammograms. This enhances medical decision-making, facilitates early disease diagnosis and supports more targeted and efficient interventions. Additionally, these applications assist healthcare professionals in optimizing their time, especially during peak pandemic periods and help mitigate the risk of human errors. Nevertheless, it is essential to emphasize that the use of such applications should complement human expertise and not replace the clinical judgment of healthcare professionals.

REFERENCES

[1] F. Bray, J. Ferlay, I. Soerjomataram, R.L. Siegel, L.A. Torre and A. Jemal, "Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries,"

CA: A Cancer Journal for Clinicians, vol. 68, no. 6, pp. 394–424, DOI: 10.3322/caac.21492, 2018.

[2] Y. Tan et al., "Tumor-derived Exosomal Components: The Multifaceted Roles and Mechanisms in Breast Cancer Metastasis," Cell Death and Disease, vol. 12, no. 6, DOI: 10.1038/s41419-021-03825-2, 2021.

[3] M. Desai and M. Shah, "An Anatomization on Breast Cancer Detection and Diagnosis Employing Multi-layer Perceptron Neural Network (MLP) and Convolutional Neural Network (CNN)," Clinical eHealth, vol. 4, pp. 1–11, DOI:10.1016/j.ceh.2020.11.002, 2021.

[4] J. Zuluaga-Gomez et al., "A CNN-based Methodology for Breast Cancer Diagnosis Using Thermal Images, Computer Methods in Biomechanics and Biomedical Engineering," Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, vol. 9, no. 2, pp. 131–145, 2021.

[5] C. Gonçalves, J. Souza and H. Fernandes, "CNN Architecture Optimization Using Bio-inspired Algorithms for Breast Cancer Detection in Infrared Images," Computers in Biology and Medicine, vol. 142, DOI: 10.1016/j.compbiomed.2021.105205, 2022.

[6] D. A. Zebari et al., "Systematic Review of Computing Approaches for Breast Cancer Detection Based Computer Aided Diagnosis Using Mammogram Images," Applied Artificial Intelligence, vol. 35, no. 15, pp. 2157–2203, DOI: 10.1080/08839514.2021.2001177, 2021.

[7] World Health Organisation, "Breast Cancer: Early Diagnosis and Screening," [Online], Available: http://www.who.int/cancer/prevention/diagnosis-screening/breast-cancer/en/, 2018.

[8] W. Al-Dhabyani, M. Gomaa, H. Khaled and A. Fahmy, "Dataset of Breast Ultrasound Images," Data Brief, vol. 28, pp. 104863, DOI: 10.1016/j.dib.2019.104863, 2020.

[9] D. Sheth and M. L. Giger, "Artificial Intelligence in the Interpretation of Breast Cancer on MRI," Journal of Magnetic Resonance Imaging, vol. 51, no. 5, pp. 1310–1324, DOI: 10.1002/jmri.26878, 2020.

[10] T. K. Y. Tay and P. H. Tan, " Liquid Biopsy in Breast Cancer: A Focused Review," The Archives of Pathology & Laboratory Medicine, vol. 145, no. 6, pp. 678–686, 2021.

[11] J. Kotsopoulos et al., "Tamoxifen and the Risk of Breast Cancer in Women with a BRCA1 or BRCA2 Mutation," Breast Cancer Research and Treatment, vol. 201, no. 2, pp. 257–264, DOI: 10.1007/s10549-023-06991-3, 2023.

[12] Z. M. Colbert and P. Ramachandran, "Auto-segmentation of Thoracic Organs in CT Scans of Breast Cancer Patients Using a 3D U-net Cascaded into 2D PatchGANs", Biomedical Physics & Engineering Express, vol. 9, no. 5, DOI:10.1088/2057-1976/ace631, 2023.

[13] G. Hamed et al.," Deep Learning in Breast Cancer Detection and Classification," Proc. of the Int. Conf. on Artificial Intelligence and Computer Vision (AICV2022), pp. 322–333, DOI: 10.1007/978-3-030-44289-7_30, Cairo, Egypt, 2020.

[14] N. S. Ismail and C. Sovuthy, "Breast Cancer Detection Based on Deep Learning Technique," Proc. of 2019 Int. UNIMAS STEM 12th Eng. Conf. (EnCon), pp. 89–92, DOI: 10.1109/EnCon.2019.8861256, Kuching, Malaysia, 2019.

[15] J. Wu and C. Hicks, "Breast Cancer Type Classification Using Machine Learning," Journal of Personalized Medicine, vol. 11, no. 2:61, DOI: 10.3390/jpm11020061, 2021.

[16] S. Khan, N. Islam, Z. Jan, I. U. Din and Joel J. P. C. Rodrigues, "A Novel Deep Learning Based Framework for the Detection and Classification of Breast Cancer Using Transfer Learning," Pattern Recognition Letters, vol. 125, pp. 1–6, DOI: 10.1016/j.patrec.2019.03.022, 2019.

[17] L. Shen, L. R. Margolies, J. H. Rothstein, E. Fluder, R. McBride and W. Sieh, "Deep Learning to Improve Breast Cancer Detection on Screening Mammography," Scientific Reports, vol.9, no. 1, p. 12495, DOI: 10.1038/s41598-019-48995-4, 2019.

[18] Z. Han, B. Wei and Y. Zheng, "Breast Cancer Multi-classification from Histopathological Images with Structured Deep Learning Model," Scientific Reports, vol. 7, no. 1, p. 4172, DOI: 10.1038/s41598-017-04075-z, 2017.

[19] D. H. Hubel and T. N. Wiesel, "Receptive Fields and Functional Architecture of Monkey Striate Cortex," Journal of Physiology, vol. 195, no. 1, pp. 215–243, DOI: 10.1113/jphysiol.1968.sp008455, 1968.

[20] S. Li, L. Wang, J. Li and Y. Yao, "Image Classification Algorithm based on Improved AlexNet," Journal of Physics: Conference Series, vol. 1813, no. 1, p. 012051, 2021.

[21] C. Szegedy et al., "Going Deeper with Convolutions," Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), pp. 1–9, DOI: 10.1109/CVPR.2015.7298594, Boston, MA, USA, 2015.

[22] A. Krizhevsky, I. Sutskever and G. E. Hinton, "Imagenet Classification with Deep Convolutional Neural Networks," Proc. of the Advances in Neural Information Processing Systems (NIPS 2012), vol.25, DOI: 10.1145/3065386, Lake Tahoe, Nevada, United States, 2012.

[23] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), pp. 770–778, DOI: 10.1109/CVPR.2016.90, Las Vegas, NV, USA, 2016.

[24] Kaggle, "Breast Mammography Images with Masses," [Online], Available: https://www.kaggle.com/datasets/tommyngx/breastcancermasses, 04/10/2023.

[25] F. Zhuang et al., "A Comprehensive Survey on Transfer Learning," Proc. of the IEEE, vol. 109, no. 1, pp. 43–76, DOI: 10.1109/JPROC.2020.3004555, 2021.

[26] S. Benbakreti, M. Benouis, A. Roumane and S. Benbakreti, "Impact of the Data Augmentation on the Detection of Brain Tumor from MRI Images Based on CNN and Pre-trained Models," Multimedia Tools and Applications, DOI: 10.1007/s11042-023-17092-0, 2023.

[27] E. M. Dogo et al., "A Comparative Analysis of Gradient Descent-based Optimization Algorithms on Convolutional Neural Networks," Proc. of the Int. Conf. on Computational Techniques, Electronics and Mechanical Systems, pp. 92–99, DOI: 10.1109/CTEMS.2018.8769211, Belagavi, India, 2018.

[28] C. F. G. D. Santos and J. P. Papa, "Avoiding Overfitting: A Survey on Regularization Methods for Convolutional Neural Networks," ACM Computing Surveys, vol. 54, no. (10s), pp. 1-25, DOI: 10.1145/3510413, 2022.

[29] K. Shankar, Y. Zhang, Y. Liu, L. Wu and C.H. Chen, "Hyper-parameter Tuning Deep Learning for Diabetic Retinopathy Fundus Image Classification," IEEE Access, vol. 8, pp. 118164–118173, DOI: 10.1109/ACCESS.2020.3005152, 2020.

[30] X. Zhang et al., "Classification of Mammographic Masses by Deep Learning," Proc. of the 2017 56th Annual Conf. of the Society of Instrument and Control Engineers of Japan (SICE), pp. 793–796, DOI: 10.23919/SICE.2017.8105545, Kanazawa, Japan, 2017.

[31] F. F. Ting, Y. J. Tan and K. S. Sim, "Convolutional Neural Network Improvement for Breast Cancer Classification," Expert Systems with Applications, vol. 120, pp. 103–115, 2019.

[32] H. Chougrad, H. Zouaki and O. Alheyane, "Deep Convolutional Neural Networks for Breast Cancer Screening," Computer Methods and Programs in Biomedicine, vol. 157, pp. 19-30, 2018.

[33] R. Karthiga, K. Narasimhan and R. Amirtharajan, "Diagnosis of Breast Cancer for Modern Mammography Using Artificial Intelligence," Mathematics and Computers in Simulation, vol. 202, pp. 316-330, DOI: /10.1016/j.matcom.2022.05.038, 2022.

[34] J. Wang, M.A. Khan, S. Wang and Y. Zhang, "SNSVM: SqueezeNet-Guided SVM for Breast Cancer Diagnosis," Computers, Materials & Continua, vol. 76, no. 2, DOI: 10.32604/cmc.2023.041191, 2023.

[35] S. U. Rehman et al., "BRMI-Net: Deep Learning Features and Flower Pollination-controlled Regula Falsi-based Feature Selection Framework for Breast Cancer Recognition in Mammography Images," Diagnostics, vol. 13, no. 9, p.1618, DOI: 10.3390/diagnostics13091618, 2023.

[36] M. Fatima, M. A. Khan, S. Shaheen, N. A. Almujally and S. H. Wang, "$B^2C^3NetF^2$: Breast Cancer Classification Using an End-to-end Deep Learning Feature Fusion and Satin Bowerbird Optimization Controlled Newton Raphson Feature Selection," CAAI Transactions on Intelligence Technology, vol.8, no.4, pp. 1374-1390, DOI: 10.1049/cit2.12219, 2023.

[37] D. Shigemizu et al., "Classification and Deep Learning–based Prediction of Alzheimer Disease Sub-types by Using Genomic Data," Translational Psychiatry, vol.13, no. 1, p. 232, DOI: 10.1038/s41398-023-02531-1, 2023.

[38] X. Liu, H. Wang, Z. Li and L. Qin, "Deep Learning in ECG Diagnosis: A Review," Knowledge-based Systems, vol. 227, p.107187, DOI:10.1016/j.knosys.2021.107187, 2021.

[39] S. Al-Fahdawi et al., "Fundus-DeepNet: Multi-label Deep Learning Classification System for Enhanced Detection of Multiple Ocular Diseases through Data Fusion of Fundus Images," Information Fusion, vol. 102, p.102059, DOI:10.1016/j.inffus.2023.102059, 2024.

[40] P. Roy, M. M. O. Chisty and H. A. Fattah, "Alzheimer's Disease Diagnosis from MRI Images Using ResNet-152 Neural Network Architecture," Proc. of the 5th Int. Conf. on Electrical Information and Communication Technology (EICT), pp. 1-6, DOI:10.1109/EICT54103.2021.9733507, 2021.

[41] R. Patnaik, P. S. Rath, S. Padhy and S. Dash, "Histopathology Colorectal Cancer Image Classification by Using Inception V4 CNN Model," Proc. of the Int. Conf. on Robotics, Control, Automation and Artificial Intelligence, pp. 1003-1014, DOI:10.1007/978-981-99-4634-1-79, 2022.

[42] M. Dehghan Rouzi et al., "Breast Cancer Detection with an Ensemble of Deep Learning Networks Using a Consensus-adaptive Weighting Method," Journal of Imaging, vol.9, no. 11, p. 247, DOI: 10.3390/jimaging9110247, 2023.

[43] A. B. Bagheri et al., "Potential Applications of Artificial Intelligence (AI) and Machine Learning (ML) on Diagnosis, Treatment, Outcome Prediction to Address Healthcare Disparities of Chronic Limb-threatening Ischemia (CLTI)," Seminars in Vascular Surgery, vol. 36, no. 3, pp. 454-459, 2023.

ملخص البحث:

سرطان الثّدي مرض واسع الانتشار عالمياً، وهو مرض يصيب النّساء في الغالب، لكنّه قد يصيب الرّجال كذلك. ويعدّ الكشْف المبكّر عن الإصابة بالمرض حاسماً في نجاعة العلاج؛ علماً بأنّ التّصوير الشّعاعي هو من بين الطّرق المستخدمة في التّشخيص. وتختلف التّوصيات بشأن الصّور الشّعاعية المستخدمة للكشْف عن الإصابة بسرطان الثّدي حسب العمر، كما تختلف من بلدٍ لآخر.

تهدف هذه الورقة إلى استخدام نماذج تقوم على التّعلُّم العميق والشّبكات العصبية الالتفافية بغرض تشخيص الإصابة بسرطان الثّدي بناءً على صُور الكُتَل الملتَقَطة للثّدي. والجدير بالذّكر أنّ تلك الكُتل قد تكون إمّا خبيثةً أو حميدةً، مما يتطلّب استراتيجياتِ تدخُّلٍ لكلّ سيناريو.

وتبيّن النّماذج المقترحة في هذه الدّراسة أهمية اختيار كلٍّ من حجْم الحُزمة ومعدّل التّعلُّم التّكيُّفي في التّأثير على النّتائج وتحسين دِقّة التّصنيف. وقد جرى استخدام مجموعة بياناتٍ مُدْمجة من ثلاث مجموعات بيانات، مع حساب الدّقّة عند استخدام كلٍّ منها، إلى جانب الدّقّة عند استخدام مجموعة البيانات المجمَّعة. وكانت أفضل دِقّة تم الحصول عليها (95%) لمجموعة البيانات (RESNET18)، الأمر الّذي يؤكّد ملاءَمة النّظام المقترح لتشخيص الإصابة بسرطان الثّدي.