Transfer Learning and Data Augmentation for Improved Breast Cancer Histopathological Images Classifier

Rania Maalej[1][0000-0003-1657-3324], Anis Mezghani[2], Mohamed Elleuch[3], Imen Mohamed ben ahmed[4] and Monji Kherallah[5]

[1]Medical School of Sfax, University of Sfax, Tunisia
rania.maalej@medecinesfax.org
[2]Higher Institute of Industrial Management, University of Sfax, Sfax, Tunisia
anis.mezghani@gmail.com
[3]National School of Computer Science (ENSI), Manouba University
mohamed.elleuch.2015@ieee.org
[4]Faculty of Sciences, University of Gafsa, Tunisia
[5]Faculty of Sciences, University of Sfax, Tunisia
monji.kherallah@fss.usf.tn

***Abstract*: Nowadays, Breast cancer is a massive health problem worldwide. To fight against this disease, we propose a high-performance Computer-Aided Diagnosis system using deep learning. Specifically, we focus on the classification of histopathological images of breast cancer into two classes (benign and malignant). For that, we present a Mobilenet-based breast cancer classification model. This model is trained with a new extended Breakhis dataset, which is created by applying some data augmentation techniques. According to the experiments, our proposed model gives a very competitive result and the accuracy reaches 0.9. This proposed model outperforms two others proposed models based on Inception and Inception-Resnet.**

***Keywords*: Breast cancer classification, MobileNet, Deep learning, Data Augmentation, BreakHis dataset**

I. Introduction

Breast cancer is a disease due to the uncontrolled growth of certain cells in the breast. It is a major health issue and the leading cause of female cancer deaths in the world [33].
To fight against this disease and to improve the survival rate, the development of automatic medical imaging processing is becoming a necessity. Indeed, this field is a rapidly expanding area where the problem of automatic interpretation of medical images is a pressing need. And given the large number of medical imaging devices, it becomes tedious to process this huge amount of information, hence the need for artificial intelligence, especially in the deep learning field. In this context, a multitude of datasets are collected and a variety of deep neural networks have been proposed.

In this work, we are interested in the classification of histopathological images of breast cancer into two binary classes (benign and malignant). To achieve this goal, and to obtain a robust model, we use the Breakhis dataset [1] as a model validator and the pre-trained network Mobilenet [2] as the features extractor.

This paper is organized as follows: section 2 presents the field of automated Breast Cancer Diagnosis. Section 3 describes the proposed materials and methods. In Section 4, we give and analyze experiment results, and at the end, conclusions and future works are given in Section 5.

II. Automated Breast Cancer Diagnosis

A. Breast cancer imaging modalities

Since cancer detection at an early stage of its development considerably increases the chances of successful treatment, several breast cancer screening methods are proposed (Ethiopian Cancer Association 2016) and among which we cite:

- Screen-film mammography (SFM)
- Digital mammography (DM)
- Ultrasound (US)
- Magnetic resonance imaging (MRI)
- Digital breast tomosynthesis (DBT)
- A genome Sequencing diagnostic (Gene)
- Histology images diagnostics (Histology)

In the table 1, we present some advantages and disadvantages of each modality.

Table 1. Breast Cancer modalities

Modality	Advantages	Disadvantages	Breast cancer dataset samples
SFM	Detect lesions at an early stage, High sensitivity in detecting breasts with fatty tissue.	Low sensitivity in detecting dense breast tissue, In dense breast tissue, 20% of tissue breast cancers are not visualized, Not digital imaging modality.	VinDr-Mammo [3].
DM	Effective imaging modality for early detection, Image improvement for better contrast is possible, Improves radiologist's sensitivity compared to SFM.	High imaging cost relative to SFM, High probability to miss cancer for overlapping breast tissue, High false positive results.	InBreast [4] Mammography Image Analysis Society (MIAS)[5] CBIS-DDSM Digital Database for Screening Mammography (DDSM)[6].
US	Achieved high accuracy in detecting and classifying benign and malignant, High sensitivity to identify abnormalities in dense breasts.	Breast lesions detection is possible only with the help of an operator, Ultrasound image interpretation is not straightforward.	Breast US Image [7].
MRI	Used for clinical diagnosis and monitoring of breast cancer, Has high sensitivity for breast cancer diagnosis compared to US and DM.	It has low specificity, It is expensive compared to US and DM.	DCE-MRI[8].
DBT	Improves the effect of overlapping breast tissue using reconstructed volume, Significantly decreases screening recall and decreased women's compliance, Reduces the number of false positives and false negatives, Detects more invasive cancers, malignant and benign lesions compared to 2D DM, Achieves high-resolution images with limited-angle tomography, Produces multiple 3D images with a single screening, Improves accuracy, sensitivity, and specificity compared to DM, Benign visibility is superior in DBT than DM.	X-ray dose for tomosynthesis image is similar to DM, It is very expensive compared to SFM, US,DM, and MRI, Calcification detection using DBT is questionable, Benign visibility is higher but mimic malignant mass-like appearance.	the Duke University Health System [9].
Gene	Provide a precise diagnosis, Particularly helpful in proactively treating cognitive or behavioral disorders before people show signs of disease.	Generate a large volume of data, although most of the data could be misleading or useless, It is also less efficient at predicting some conditions because there are conditions it is not screened for, It may uncover information that an individual may not be looking for and does not want to know.	NKI Breast Cancer Data [10], Breast Cancer Proteomes [11].
Histology	Provide more comprehensive information for diagnosis and the diseases are analyzed by detecting tissue and cells in lesions.	Preparation of the slides using the paraffin technique can be time-consuming; frozen slides are faster to prepare, but this can affect the resolution, especially when using light microscopy.	Bioimaging Challenge 2015 Breast Histology Dataset [12], The BACH dataset [13], BreaKHis [1].

B. Histological Breast cancer image datasets

Over the last few decades, a lot of datasets were produced and published in different repositories and some of them were publicly available for use. Histological Breast cancer image analysis has mainly used these databases. The summary of the most cited and recently updated histological breast cancer databases is presented in Table 2.

Table 2. Breast Cancer Images Datasets.

Name	Classes	Images
Bioimaging Challenge 2015 Breast Histology Dataset [12].	4	249
The BACH dataset[13].	4	500
BreaKHis [1].	2	7909

In Bioimaging Challenge 2015 Breast Histology Dataset [12], images are composed of high-resolution (2040 × 1536 pixels). All the images are digitized with the same acquisition conditions, with a magnification of 200×. Each image is labeled with one of four classes:

1.   normal tissue,
2.   benign lesion,
3.   insitu carcinoma,
4.   invasive carcinoma.

However, the BACH dataset[13] is composed of microscopy images annotated by two expert pathologists from the Institute of Molecular Pathology and Immunology of the University of Porto (IPATIMUP) and the Institute for Research and Innovation in Health (i3S). This dataset is composed of 400 training and 100 test images, with four classes. Images, where there was disagreement between the Normal and Benign classes, were discarded. The provided images are in RGB .tiff format and have a size of 2048 × 1536 pixels. The labels of the images were provided in .csv format.

The BreaKHis [1] is an open-source dataset available. It consists of 7909 clinical breast tumor histopathological images of 700 x 460-pixel size, respectively, including 2480 benign tumors (adenoma, fibroadenoma, trichome tumors, and tubular adenoma) images and 5429 malignant tumors (ductal carcinoma, lobular carcinoma, mucinous carcinoma, and papillary carcinoma) images at four magnifications of 40, 100, 200, and 400 as shown in Table 3.

Table 3. The Breakhis image distribution by the magnification factor.

Magnification factors	Class Benign	Class Malignant	Total
40×	625	1370	1995
100×	644	1437	2081
200×	623	1390	2013
400×	588	1232	1820

Figure 1 illustrates four different magnification images of breast tissue sections containing malignant tumors from the BreakHis dataset.In this work, we focus on the BreakHis dataset because it has the highest number of images.
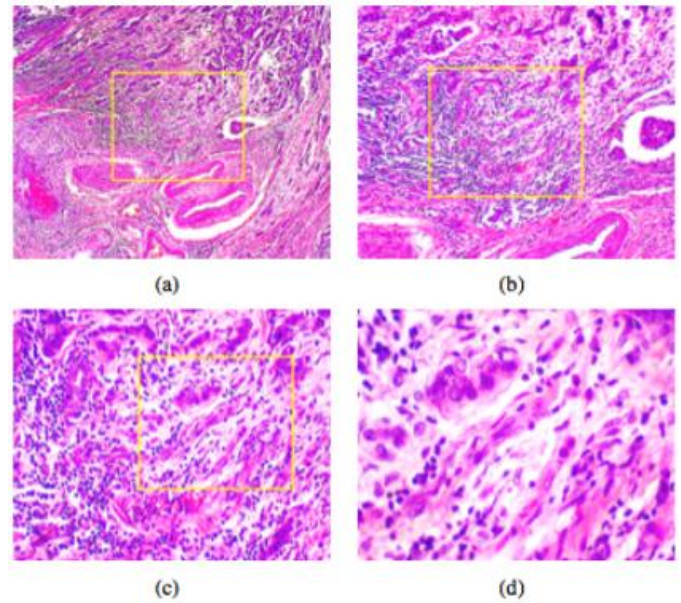

Figure 1. A slide of a breast malignant tumor seen in different magnification factors: (a) 40×, (b) 100×, (c) 200×, and (d) 400×[1]

C. Previous works on BreakHis

With the importance of breast cancer classification in histopathological imaging, as shown in Table 4, there are several studies in the existing literature, and the most recent ones are based on the deep learning field. The most common works used the Convolutional Neural Network [14] or the pre-trained models such as AlexNet [15], GoogleNet [16], ResNet [17], and VGG16 [18]. Indeed, F. A. Spanhol et al. [19] presented a novel strategy for training the Alexnet-CNN architecture, based on the extraction of patches obtained randomly or by a sliding window mechanism, to process high-resolution textured images. Experimental results obtained on the BreaKHis dataset showed high performance and the accuracy reached 85.6 %.

In [20], H. Seo, et al. proposed a novel Primal-Dual Multi-Instance SVM classification method, which allows scaling to a wide range of features. Histopathological images are segmented into patches. The feature vector is extracted through the Parameter Free Threshold Statistics (PFTAS) method for each patch. The PFTAS method extracts texture features by counting the number of black pixels in the neighborhood of a pixel. Experiment results on the BreaKHis dataset showed an improved accuracy of 89.8%. In [21] a Deep Convolution Generative Adversial Network (DCGAN) is applied to give the number of consistent images in the minority class (benign) with that in the majority class (malignant). In addition, the pre-trained DenseNet201 model is used, and features are extracted from the lower layers of DensNet201, via a global average pooling (GAP). These features are passed throw the SoftMax layer to classify breast cancer. The proposed architecture was evaluated using histopathological images from the BreakHis database and showed promising results with 96% with the 40× magnification.

M. Saini et S. Susan [22] used the Deep Convolutional Generative Adversarial Network (DCGAN) for minority data augmentation in the initial phase of their experiments. DCGAN is used to generate high-quality synthetic fake

images from the available distribution of minority data. Then, the new dataset, with a balanced class distribution, is got across the deep transfer network. To enhance the network performance, the proposed VGG16 deep transfer architecture is followed by the batch normalization, 2D convolutional (CONV2D) layer, 2D Global Average Pooling, dropout, and dense layers. The model is evaluated using a two-class BreaKhis provided at four magnification levels and the best accuracy was 96.5%.

A. M. Ibraheem et al. [23] proposed three parallel CNN branches (3PCNNBNet) for breast cancer classification through histopathological images. This network offered several advantages, such as learning the high-level and low-level features by considering local and global features simultaneously. They also deployed deep residual blocks using skip connections to help the proposed model overcome the vanishing gradient problem and to improve training and testing. The proposed 3PCNNB-Net architecture was evaluated using histopathological images from the BreakHis database. The 3PCNNB-Net architecture achieved promising results, including a maximum accuracy of 97.04% with a 200× magnification.

Y. Zou et al. [24], introduced a novel attention high-order deep network (AHoNet) by simultaneously embedding attention mechanism and high-order statistical representation into a residual convolutional network, this technique allows the network to capture more discriminant deep features for breast cancer pathological images. Experiments on the benchmark BreakHis dataset for different magnification factors: 40X, 100X, 200X, and 400X validate the effectiveness of the proposed deep network in terms of the high scores obtained compared to the state-of-the-art deep networks, in fact, AHoNet gets the optimal image-level classification accuracies of 99.09%.

III. Materials and Methods

A. Data Augmentation on BreakHis histopathological breast cancer dataset

As the size of the dataset plays a very important role in achieving excellent performance in a deep-learning model, data augmentation enhances network performance and overcomes the overfitting problem [25].
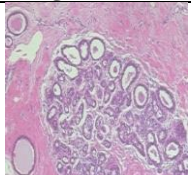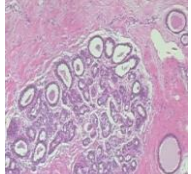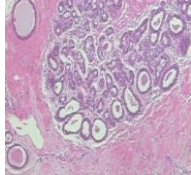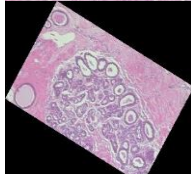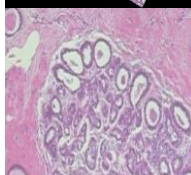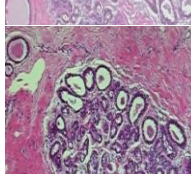
Data augmentation is a process of artificially increasing the amount of data by generating new data points from existing data. Augmented data are derived from original images with some sort of minor geometric transformations to increase the diversity of the training set.

In this study, multiple data augmentation techniques are applied, creating multiple versions of each image. Moreover, the horizontal flip transformation is used to generate synthetic high-quality fake images from the available distribution of minority data (benign). So, the class distribution is balanced. Then some other techniques are used to increase the total number of images in this new extended dataset we use the vertical flip, rotation, cropping, and sharpening techniques so, as shown in Table 5. As a consequence, the total number of images becomes 51945 (24800 Benign and 27145 Malignant) as illustrated in Figure 2.

Table 4. Compared results with CNN-based methods on the BreakHis dataset at the image level

System	Data Augmentation	Features extraction	Classification	Magnification	Accuracy %
F. A. Spanhol et al. [19]	-	AlexNet CNN	Softmax layer	All	85.60
H. Seo, et al. [20]	-	PFTAS method.	Primal-Dual Multi-Instance SVM	200X	89.80
H. Djouima,et al. [21]	DCGAN + rotation ; shear; zoom; horizontal flip; fill mode; width shift; height shift.	DenseNet201.	Softmax layer	40X	96.00
M. Saini et S. Susan [22]	DCGAN.	VGG16+ CNN.	Softmax layer	40X	96.50
A. M. Ibraheem et al. [23]	-Three zoom ranges, -Rotation at 90°, -Horizontal and vertical flipping.	Three parallel CNN branches + residual blocks.	Softmax layer	200X	97.04
Y. Zou, et al. [24]	- simple cropping, -horizontal and vertical flipping, -rotation (90, 120, and 180), -Cutmix method for data amplification.	ResNet1825 + attention mechanism + high-order statistical representation.	Softmax layer	200X	99.09

Table 5. Data augmentation techniques were applied to image "SOB_B_A-14-22549AB-40-017" from the Breakhis dataset.

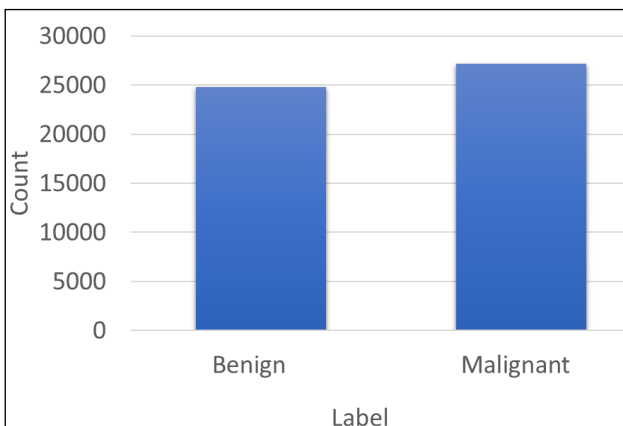| Image	Transformation
	Raw image
	Horizontal flip
	Vertical flip
	Rotation
	Crop
	Sharpen

Figure 2. The balanced class distribution in the new extended Breakhis dataset

B. MobilNet for features extraction

MobileNet V3 is a family of neural network architectures for efficient on-device image classification and related tasks, originally published by A. Howard et al. [2]

Similar to earlier MobileNets, MobileNet V3 adjusts the accuracy vs. latency tradeoff by multiplying the depth (number of features) in the convolutional layers. In order to adapt the network to use cases requiring low or high resource levels, MobileNet V3 is available in two sizes, small and large. All pre-trained checkpoints were provided with the exact 224 × 224 input resolution, even though V3 networks, like other Mobilenets, can be constructed with different input resolutions. This TF Hub model uses the TF-Slim implementation of mobilenet_v3 as a large network with a depth multiplier of 0.75.

This implementation of Mobilenet V3 (Figure 3) rounds feature depths to multiples of 8. Depth multipliers less than 1.0 are not applied to the last convolutional layer (from which the module takes the image feature vector).

The model contains a trained instance of the network, packaged to get feature vectors from images. the full model including the classification it was originally trained for use. The checkpoint's weights were originally obtained by training on the ILSVRC-2012-CLS dataset for image classification ("Imagenet").
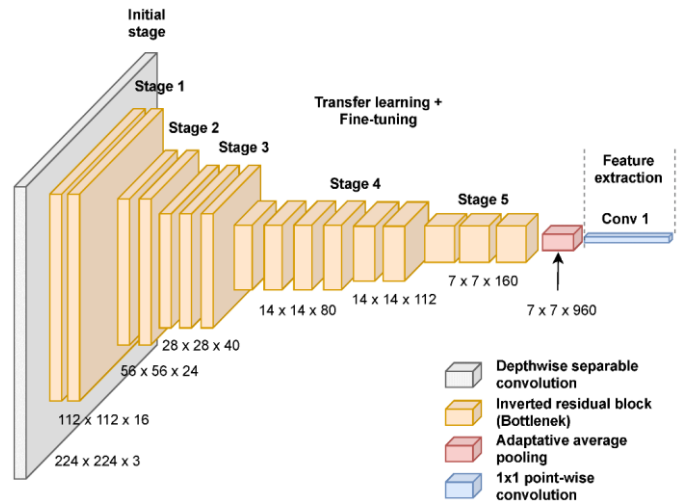

Figure 3. The Mobilenet architecture for features extraction.

C. Fully connected network for classification

A fully connected layer refers to a neural network in which each neuron applies a linear transformation to the input vector through weights matrix. As a result, all possible connections layer-to-layer are present, meaning every input of the input vector influences every output of the output vector.

High-level features in the data are represented in the convolutional layer output. It is possible to flatten that output and connect it to the output layer, but adding a fully connected layer provides a (typically) less expensive technique to learn non-linear combinations of these features.

D. BatchNormalization layer

The BatchNormalization layer normalizes the activation of the previous layer at each batch and by default, it is using the following values:

- Momentum defaults to 0.99
- The hyperparameter ε defaults to 0.001
- The hyperparameter β defaults to an all-zeros vector
- The hyperparameter γ defaults to an all-ones vector

E. Dropout

In a deep network, overfitting may occur as a result of the huge number of parameters. To protect the network against this problem, dropout [26] was applied in different layers. This technique is highly effective although it may increase the training time since it consists in temporarily removing randomly selected units from the network during the training step. This regularization method improves network performance and significantly reduces the error rate.

F. ReLU

As the activation functions can change any linear classifier into a non-linear one, they are regarded as a vital component of neural networks. These activation functions have also witnessed high performance across several tasks in recent years. Different activation functions like tanh or sigmoid can approximate arbitrary continuous functions, so, theoretically, they are often considered equivalent; however; practically they often show very diverse behaviors. For example, the sigmoid was revealed to be less suitable for learning despite its highly important activation functions in neural networks. This different behavior is caused by its small derivative which may result in the vanishing gradients phenomenon. In this framework, the ReLU function has proven its greater suitability because it has an identity derivative in the positive region and is thus claimed to be less susceptible to vanishing gradients.

G. System architecture

First, The Breakhis dataset is randomly shuffled to be divided into training, testing, and validation sets. Then, the images are resized to 224×224 pixels to comply with the input size requirements of MobileNet V3[2] used for feature extraction. The MobileNet aims to use depthwise separable convolutions to make lighter deep neural networks, so the computational cost is less than the regular convolutional networks. However, MobileNet V3 uses a depth (number of features) multiplier in the convolutional layers to adjust the tradeoff between accuracy and latency.

For classification, we propose a fully connected network with four dense layers, three Batch normalization layers, and three dropout layers with a rate of 0.5.

Our proposed model to classify the histopathological images of breast cancer into benign and malignant classes is presented in Figure 4.
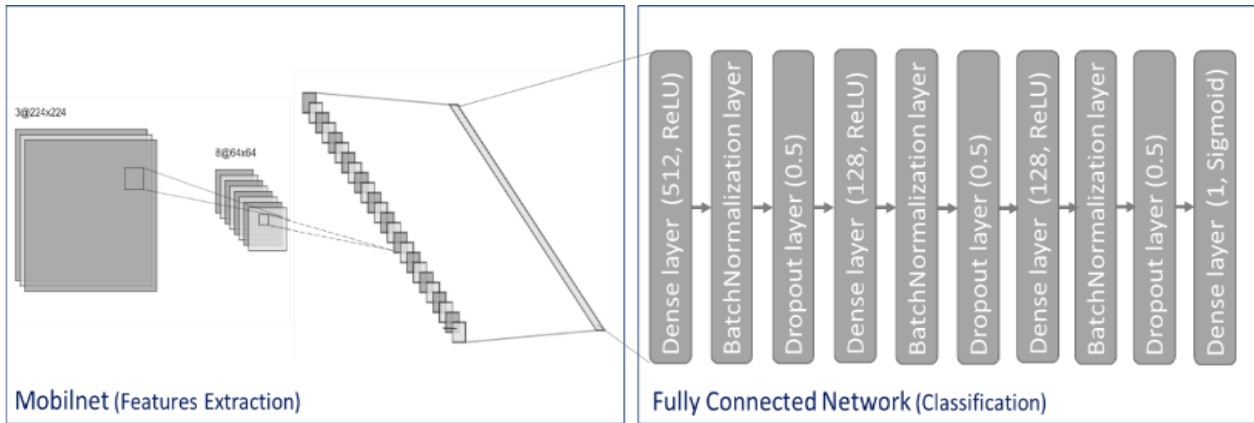

Figure 4. The proposed system architecture

IV. Experiment results and comparison

A. Experiment results

To train our model, we choose the Adam optimizer [27] and use the 5-Fold Cross Validation Technique, we perform categorical cross entropy [28] as a cost function and we set the batch size parameter as 32.

The best accuracy during the training is 0.99. Figure 5 visualizes the model performance history during the training and validation steps [29].

In addition, to analyze the recognition results on the testing set, we make the confusion matrix for our proposed model, as shown in Figure 6. It's pretty obvious that our method has a very good discriminative effect on the histopathological images of breast cancer [34].

To evaluate the results, we follow the metrics proposed in [30] as it is a classification task. The proposed models are evaluated by Accuracy A, Precision P, Recall R, and F-score F.

These metrics are computed according to Equations 1,2,3 and 4 where TP is the number of True Positive, TN is the number of True Negative, FN is the number of False Negative, and False Positive (FP).

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

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

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

$$F = \frac{2*P*R}{P+R} \quad (4)$$

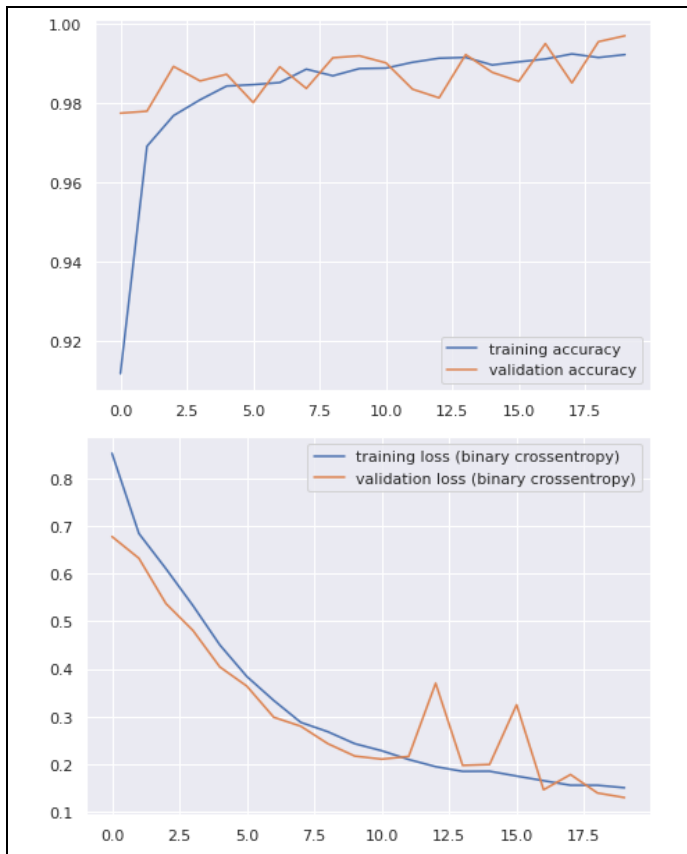$$A = 0.902 \; ; \quad P = 0.904 \; ;$$
$$R = 0.897 \; ; \quad F = 0.900$$

Figure 5. Accuracy and Loss during training and validation.

Figure 6. Confusion matrix.

B. Comparison

We train the proposed system with other deep models using other features extraction techniques such as Inception V3 and Inception_resnet V2 presented below:

1) Inception V3

Inception V3 is a convolutional neural network, which uses the mathematical operations pooling and convolutions. The inception module is illustrated in Figure 7.

A typical pooling and convolutional layer stem, followed by pooling layers in between inception modules, make up the architecture. The Inception V3 architecture also uses reduction modules, which are conceptually equivalent to inception modules but aim to condense the input's dimensions. Inception V3 has roughly 24M parameters in total. It's also important to note that the V3 employs an RSMProp optimizer and requires an input of 299×299×3 by default.

2) Inception_resnet V2

A convolutional neural architecture called Inception-ResNet-v2 expands on the Inception family of architectures while incorporating residual connections, replacing the filter.

A convolutional neural architecture called Inception-ResNet-v2 expands on the Inception family of architectures while incorporating residual connections, replacing the filter concatenation stage of the Inception architecture, as shown in Figure 8.

We note that Inception-ResNet-v2 is a variation of the Inception V3 model, and it is considerably deeper than the previous Inception V3 but the Mobilenet uses Depthwise separable convolution while Inception V3 uses standard convolution. This results in a lesser number of parameters in MobileNet compared to InceptionV3.

To evaluate the results, we follow the same metrics proposed before, and both proposed models are evaluated by Accuracy A, Precision P, Recall R, and F-score F. Experimental results are listed in Table 6.

Table 6. Obtained metrics by proposed models validated on the Breakhis dataset.

Metrics	Mobilenet_V3	Inception_v3 [31]	Inception resnet[32]
Accuracy	0.902	0.899	0.896
Precision	0.904	0.900	0.898
Recall	0.897	0.893	0.891
F1-score	0.900	0.897	0.894

V. Conclusions

This paper presents a Mobilenet-based model for breast cancer classification from histopathological images. This model is trained with a new extended Breakhis dataset, which is created by applying some data augmentation techniques. Based on the experiments, our proposed model gives a very competitive result. Indeed, the accuracy reaches 0.9. This proposed model outperforms two others proposed models based on Inception and Inception-Resnet.

In Future work, we propose to improve the network by increasing its depth to classify the subclasses of the BreastHis dataset.
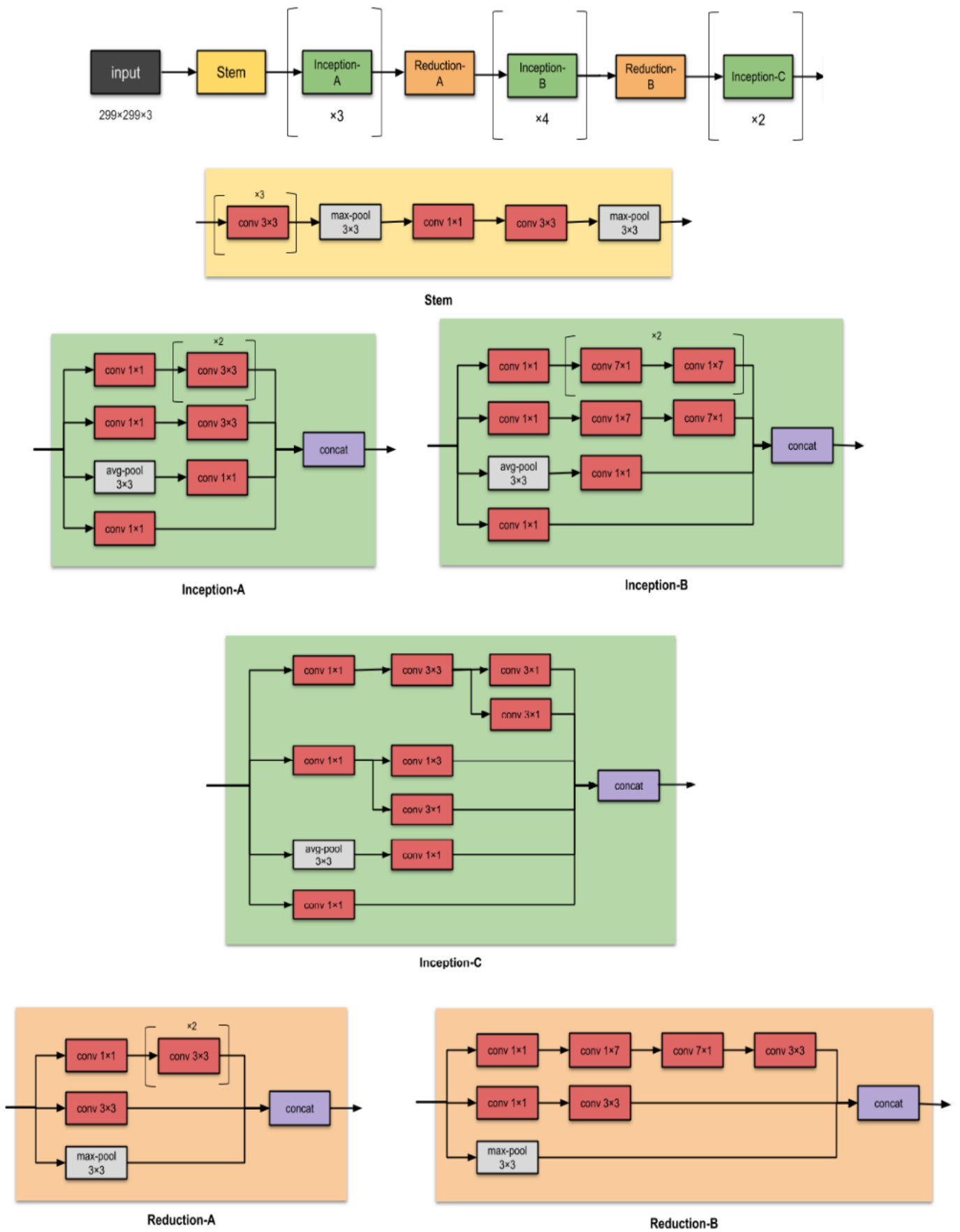
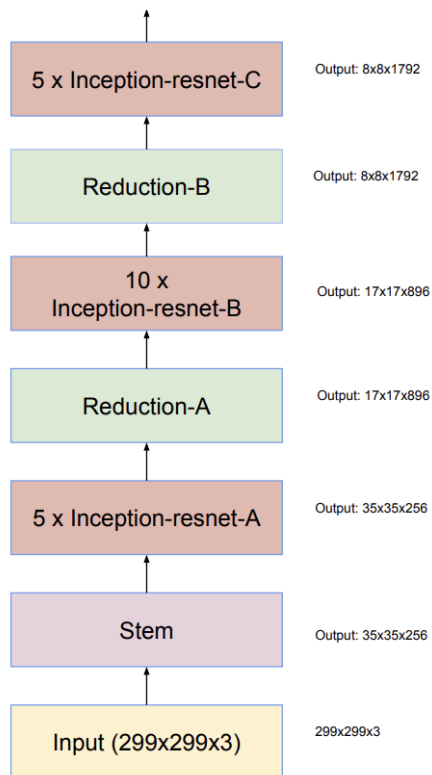**Figure 7.** The Inception V3 architecture for features extraction

Figure 8. The Inception Resnet V2 architecture for features extraction.

References

[1] F. A. Spanhol, L. S. Oliveira, C. Petitjean, et L. Heutte, « A Dataset for Breast Cancer Histopathological Image Classification », *IEEE Transactions on Biomedical Engineering*, vol. 63, nº 7, p. 1455-1462, juill. 2016.

[2] [A. Howard, M. Sandler, G. Chu, L. C. Chen, B. Chen, M. Tan, ... & H. Adam. Searching for mobilenetv3. In *Proceedings of the IEEE/CVF international conference on computer vision* (pp. 1314-1324) (2019).

[3] H. T. Nguyen, H. Q. Nguyen, H. H. Pham, K. Lam, L. T. Le, M. Dao & V. Vu. VinDr-Mammo «A large-scale benchmark dataset for computer-aided diagnosis in full-field digital mammography» *MedRxiv*, p. 2022.03. 07.22272009. (2022-03).

[4] I. C. Moreira, I. Amaral, I. Domingues, A. Cardoso, M. J. Cardoso, et J. S. Cardoso, « INbreast: toward a full-field digital mammographic database » *Acad Radiol*, vol. 19, no 2, p. 236-248, févr. 2012.

[5] J. Suckling, J. Parker, D. Dance, S. Astley, I. Hutt, C. Boggis, ... & J. Savage, « Mammographic Image Analysis Society (MIAS) database v1.21 ». 28 août 2015.

[6] M. H. PUB, K. Bowyer, D. Kopans, R. Moore, et P. K. Jr, « The digital database for screening mammography ». In: *Proceedings of the Third International Workshop on Digital Mammography*, Chicago, IL, USA. p. 9-12. 1996.

[7] W. Al-Dhabyani, M. Gomaa, H. Khaled, et A. Fahmy, « Dataset of breast ultrasound images », *Data in Brief*, vol. 28, p. 104863, févr. 2020.

[8] A. Saha et al., « A machine learning approach to radiogenomics of breast cancer: a study of 922 subjects and 529 DCE-MRI features », *Br J Cancer*, vol. 119, no 4, p. 508-516, août 2018.

[9] M. Buda, A. Saha, R. Walsh, S. Ghate, N. Li, A. Święcicki, ... & M. A. Mazurowski, « A Data Set and Deep Learning Algorithm for the Detection of Masses and Architectural Distortions in Digital Breast Tomosynthesis Images » *JAMA network open*, vol. 4, no 8, p. e2119100-e2119100, 2021.

[10] P. Y. Lum, G. Singh, A. Lehman, T. Ishkanov, M. Vejdemo-Johansson, M. Alagappan, ... & G. Carlsson, « Extracting insights from the shape of complex data using topology » *Scientific reports*, 2013, vol. 3, no 1, p. 1-8. (2013).

[11] P. Mertins, D. R. Mani, K. V. Ruggles, M. A. Gillette, K. R. Clauser, P. Wang, ... & Nci Cptac « Proteogenomics connects somatic mutations to signalling in breast cancer » *Nature*, vol. 534, no 7605, p. 55-62. juin 2016.

[12] T. Araújo, G. Aresta, E. Castro, J. Rouco, P. Aguiar, C. Eloy, ... & A. Campilho « Classification of breast cancer histology images using Convolutional Neural Networks », *PLoS ONE*, vol. 12, no 6, p. e0177544, juin 2017.

[13] G. Aresta, T. Araújo, S. Kwok, S. S. Chennamsetty, M. Safwan, V. Alex, ... & P. Aguiar « BACH: Grand Challenge on Breast Cancer Histology Images » *Medical Image Analysis*, vol. 56, p. 122-139, août 2019.

[14] F. F. Ting, Y. J. Tan, et K. S. Sim, « Convolutional neural network improvement for breast cancer classification », *Expert Systems with Applications*, vol. 120, p. 103-115, avr. 2019.

[15] S. A. A. Hassan, M. S. Sayed, M. I. Abdalla & M. A. Rashwan «Breast cancer masses classification using deep convolutional neural networks and transfer learning» *Multimedia Tools and Applications*, vol. 79, p. 30735-30768. (2020).

[16] X. Yao, X. Wang, Y. Karaca, J. Xie, & S. Wang, « Glomerulus Classification via an Improved GoogLeNet », *IEEE Access*, vol. 8, p. 176916-176923, 2020.

[17] Y. Jiang, L. Chen, H. Zhang & X. Xiao «Breast cancer histopathological image classification using convolutional neural networks with small SE-ResNet module» *PLoS one*, vol. 14, no 3, p. e0214587. (2019).

[18] D. Albashish, R. Al-Sayyed, A. Abdullah, M. H. Ryalat, et N. Ahmad Almansour, « Deep CNN Model based on VGG16 for Breast Cancer Classification », in *2021 International Conference on Information Technology (ICIT)*, p. 805-810, juill. 2021.

[19] F. A. Spanhol, L. S. Oliveira, C. Petitjean, et L. Heutte, « Breast cancer histopathological image classification using Convolutional Neural Networks », *in 2016 International Joint Conference on Neural Networks (IJCNN)*, Vancouver, BC, Canada, p. 2560-2567, juill. 2016.

[20] H. Seo, L. Brand, L. S. Barco, et H. Wang, « Scaling multi-instance support vector machine to breast cancer detection on the BreaKHis dataset »,

Bioinformatics, vol. 38, no Supplement_1, p. i92-i100, juin 2022.

[21] H. Djouima, A. Zitouni, A. C. Megherbi, et S. Sbaa, « Classification of Breast Cancer Histopathological Images using DensNet201 », in *2022 7th International Conference on Image and Signal Processing and their Applications (ISPA)*, Mostaganem, Algeria, p. 1-6, mai 2022.

[22] [M. Saini et S. Susan, « Deep transfer with minority data augmentation for imbalanced breast cancer dataset », *Applied Soft Computing*, vol. 97, p. 106759, déc. 2020.

[23] A. M. Ibraheem, K. H. Rahouma, et H. F. A. Hamed, « 3PCNNB-Net: Three Parallel CNN Branches for Breast Cancer Classification Through Histopathological Images », *J. Med. Biol. Eng.*, vol. 41, no 4, p. 494-503, août 2021,

[24] Y. Zou, J. Zhang, S. Huang, et B. Liu, « Breast cancer histopathological image classification using attention high-order deep network », *International Journal of Imaging Systems and Technology*, vol. 32, no 1, p. 266-279, 2022.

[25] R. Maalej et M. Kherallah, « New MDLSTM-based designs with data augmentation for offline Arabic handwriting recognition », *Multimedia Tools and Applications*, vol. 81, no 7, p. 10243-10260, mars 2022.

[26] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. «Dropout: a simple way to prevent neural networks from overfitting». *The journal of machine learning research*, 15(1), 1929-1958 (2014).

[27] D. P. Kingma et J. Ba, « Adam: A Method for Stochastic Optimization ». *arXiv preprint arXiv*:1412.6980, 2014.

[28] R. E. Hoskisson, M. A. Hitt, R. A. Johnson, et D. D. Moesel, « Construct validity of an objective (entropy) categorical measure of diversification strategy », *Strategic Management Journal*, vol. 14, no 3, p. 215-235, 1993.

[29] I. ahmed ben Mohamed, R. Maalej, et M. Kherallah, « MobileNet-based model for Histopathological breast cancer image classification », *In 22nd International Conference on Hybrid Intelligent Systems (HIS 2022)*, held December 13-15, 2022, Springer International Publishing, 2023.

[30] Q. Kong, Y. Xu, I. Sobieraj, W. Wang, et M. D. Plumbley, « Sound event detection and time–frequency segmentation from weakly labelled data » *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 27, no 4, p. 777-787, 2019.

[31] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. «Rethinking the inception architecture for computer vision» *In Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2818-2826) (2016).

[32] Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. «Inception-v4, inception-resnet and the impact of residual connections on learning» *In Proceedings of the AAAI conference on artificial intelligence*, Vol. 31, No. 1, (2017).

[33] L Dora, S Agrawal, R Panda, A Abraham, Optimal breast cancer classification using Gauss–Newton representation based algorithm, Expert Systems with Applications, 97: 134-145, 2017.

[34] A Abraham, Intelligent systems: Architectures and perspectives, Recent advances in intelligent paradigms and applications, Springer, 1-35, 2003.