ISSN: 2710-2165

Enhancing Breast Cancer Classification using a Modified GoogLeNet Architecture with Attention Mechanism

Alaa Hussein Abdulaal [*], Morteza Valizadeh [**], Baraa M. Albaker [***], Riyam Ali Yassin [****], Mehdi Chehel Amirani [*****], A. F. M. Shahen Shah [******]

[*] Department of Electrical Engineering, Al-Iraqia University
Department of Communication Engineering, Urmia University
Email: Engineeralaahussein@gmail.com
https://orcid.org/0000-0003-2316-2822

[**] Department of Communication Engineering, Urmia University
Email: m.valizade@gmail.com
https://orcid.org/0000-0003-2204-3303

[***] Department of Electrical Engineering, Al-Iraqia University
Email: baraamalbaker@ymail.com
https://orcid.org/0000-0002-6030-3121

[****]Department of Communication Engineering, Urmia University
Email: RiyamAliYassin@gmail.com
https://orcid.org/0009-0004-1202-0380

[*****]Department of Communication Engineering, Urmia University
Email: m.amirani@urmia.ac.ir
https://orcid.org/0000-0002-5179-9831

[******] Department of Electronics and Communication Engineering, Yildiz Technical University
Email: shah@yildiz.edu.tr
https://orcid.org/0000-0002-3133-6557

Abstract

Breast cancer incidence has been soaring sharply, and this is causing grave concern worldwide due to its high mortality rates. It should be correctly diagnosed in the early stages in order to achieve better patient outcomes. Over the last decade, there has been a great demand for diagnosis systems based on AI that could be used in breast cancer detection and classification. These computerized devices utilize deep learning algorithms to analyze medical scans thereby allowing for subtle abnormality recognition and distinguishing malignant from benign tumors. Computer-aided diagnosis named CAD systems can assist radiologists and pathologists to be more precise with their diagnoses while at the same time increasing productivity. Furthermore, recent advances in CNN architectures coupled with attention mechanisms have further improved CAD systems for breast cancer diagnosis. Attention-based CNN models focus on crucial regions hence enhancing classification accuracy and reliability. In this study, we introduce a new approach that improves the classification of breast cancer using GoogLeNet architecture modified by an attention mechanism that is based on image regions. The modified GoogLeNet has a spatial transformer network (STN), which allows it to focus on significant areas in breast histopathology images using selective attention. Through the attention mechanism, the model becomes better at learning discriminatory features that indicate different subtypes of breast cancer. In order to evaluate the effectiveness of this method, we implemented experiments using the BreaKHis dataset for classifying breast carcinomas. This dataset has been intentionally collected under various magnifications so as to facilitate binary and multiple classification tasks. These outcomes clearly show that modified GoogLeNet with attention outperforms the original GoogLeNet architecture in terms of accuracy. For binary classification, the proposed model demonstrated an accuracy rate of 98.08%, whereas GoogLeNet's rate was 94.99%. For multi-class classification, at 100x magnification, this model achieved an accuracy of 94.63% while the accuracy of the original GoogLeNet was 85.06%. It is evident from these findings that the efficiency of breast cancer diagnosis significantly improved under this proposed approach. The findings of the study show that incorporating the modified GoogLeNet framework with an attention mechanism in CAD systems for breast cancer detection can improve their performance. Combining deep learning and attention models together can lead to more accurate treatment decisions and better patient results. More efforts are needed to further develop CAD systems in this area to assist ongoing endeavors towards upgrading them and ultimately contribute to saving many lives in the fight against breast cancer.

Keywords- Breast cancer, computer-aided diagnosis, deep learning, attention mechanism, spatial transformer network.

I. INTRODUCTION

The most common type of invasive cancer among women worldwide is breast cancer, and it is indeed a serious health concern. It is the second commonest cause of death among females as a result of cancer after lung cancer [1]. The International Agency for Research on Cancer (IARC) which is a branch of the World Health Organization (WHO) reported about 8.2 million cancer-related deaths within 2012 only [2].
In terms of the incidence rate, there will be more cases in the future. Generally, it is estimated that there will be over 27 million new cases by 2030 [3]. Timely detection and treatment of breast cancer are crucial in the management, diagnosis or even prevention of the disease. This is why deep learning research, which may include computer vision, is very important for fighting this rampant and devastatingly hazardous condition. Such research might form the basis for more efficient diagnostic tools along with improved breast cancer treatment planning and patient care.

A large number of women are still affected by breast cancer which remains a major issue in global health. If effective treatment and patient outcomes matter, then early detection and accurate diagnosis cannot be overemphasized [4-5]. Recently, doctors have found that computer-aided diagnostic (CAD) systems which use convolutional neural networks (CNNs) could provide valuable help to them in detecting and classifying breast cancer [6].

Of all the techniques for examining mammograms, ultrasound scans, and histopathological slides, the CAD system is believed to be one of the best [6-7]. CNNs are just algorithms that work specifically on visual inputs using a deep learning approach. They have a great capacity to identify complex patterns as well as structures of images thus making them good at analyzing breast cancer. CNN-based CAD systems detect subtle abnormalities in medical imaging or classify breast tumors using large datasets and complex architectures [8]. This is because CNNs have the ability to learn multiple levels of representation, which makes them capable of distinguishing between benign and malignant tumors as well as identifying the subtypes of breast cancers.

The integration of CAD systems into daily clinical practice has shown great promise in improving the accuracy and efficiency of radiologists and pathologists. Thus, these systems can be applied for tumor localization, segmentation, and classification providing valuable decision-support tools to reduce human errors.

In addition, CNN architectures and attention mechanisms have been continuously improving, which has further enhanced the CAD system capabilities in breast cancer diagnosis. Attention mechanisms can lead to the selective focus of the model on the relevant areas of medical images thus helping in accurate analysis while reducing unwanted noise or unnecessary information in those images. As such, breast cancer classification can be more reliable and accurate with attention-based CNN models that turn out critical regions revealing malignant tumor parts. The interplay between breast cancer research, CAD systems, and CNNs has transformed the world of breast cancer diagnosis. Incorporating attention mechanisms and involving deep learning techniques such as deep neural networks can enable significant strides towards improved early detection, faster diagnosis, etc., that may save lives [8-9].
Continuous research and developments are aimed at enhancing CAD systems based on CNNs by leveraging on this. By incorporating these advances, we can provide medical professionals with tools that support accurate breast cancer diagnosis and thus enable them to make more informed decisions when it comes to treatment for the sake of better outcomes for patients with breast cancer.

In this study, we suggest a new way of improving the performance of breast cancer classification using a modified GoogLeNet architecture and an attention mechanism based on image regions. Breast cancer detection is crucial for early recognition and effective therapy, while deep learning approaches that involve GoogLeNet showed encouraging outcomes in this area.

GoogLeNet, which is also referred to as Inception-v1, is a deep convolutional neural network architecture that has shown its effectiveness in image classification tasks. Nonetheless, to enhance its performance for breast cancer classification, we propose incorporating an attention mechanism based on image regions.

STN as implemented by an attention mechanism makes it possible for the model to focus on particular areas within the BreaKHis dataset which is a collection of breast histopathology images. Focused on relevant areas, the attention mechanism can make the model capture fine-grained patterns and cues revealing different types of breast cancer. BreaKHis dataset classification accuracy would be improved by the GoogleNet model, as it seeks to make the models more discriminative by this particular approach based on attention. This STN module is a GoogLeNet modified with an integrated STN module that would allow it to understand how to transform the input image in space and emphasize important parts of it, which are detected after the convolutional layer. Our belief is that by identifying essential parts of images, the modified GoogLeNet with an attention mechanism will make it possible to give better predictions, higher precision, and reliability in breast cancer identification. The task of the attention mechanism is to allow the model to allocate its computing power toward the most crucial zones by making irrelevant details less influential.

By means of this approach, our aim is to contribute towards CAD systems for breast cancer diagnosis. Combining popular GoogLeNet architecture with an attention mechanism specifically designed for BreaKHis dataset could result in better accuracy and efficiency in classifying breast cancers thus enabling accurate treatment decisions that would lead to better patient outcomes at long last. The effectiveness of the proposed approach in improving the accuracy of breast cancer classification can be evaluated by comparing and evaluating the modified GoogLeNet model against the original GoogLeNet architecture. The researchers' findings from this research can light up key areas for more studies for practitioners and researchers in breast cancer diagnosis as well as contribute to an ongoing endeavor of developing efficient CAD systems for accurate breast cancer detection and classification.

II. LITERATURE REVIEW

Machine learning in biomedical engineering has evolved and thus many studies have used handcrafted features-based approaches for histopathology image classification of breast cancer. As an illustration, other CNN models which include VGG16, VGG19, inceptionV3, and ResNet50 were compared for classifying benign and malignant images of breast cancer [10]. A model of CNN was proposed to extract features from histopathology images of breast cancer and classify them using support vector machines (SVM) [11]. The authors recently suggested a new deep neural network that uses the clustering method and CNN model, Long-Short-Term-Memory (LSTM), or a combination of CNN and LSTM models [12]. An AlexNet as a feature extractor; an SVM as a classification model; A hybrid CNN model composed of AlexNet, MobileNetV2, and ResNet50 [15].

It is known that different types of nucleus-guided feature extraction framework based on the CNN approach have been fronted for histopathological image classification [16], as well as automated segmentation of glandular epithelium on Hematoxylin and eosin (H&E) stained images with immunohistology compatibility (IHC) stains [17]. A weak supervised learning approach using multiple instance learning (MIL) was proposed in another study for comparison of several MIL methods like Axis-Parallel Rectangle (APR), diverse density, MIL-support vector machines, k-nearest neighbor; and reported that non-parametric approach employing MIL-CNN deep learning model outperformed the other methods [18]. To automatically segment and classify the epithelial and stromal regions from the microarray images of digital tumor tissue, a DL-based CNN approach was suggested. Most approaches are based on low-level image features such as color, texture, and local binary patterns (LBP) when classifying the two regions. The deep CNN feature extractor is directly learned from the raw pixel intensity value of epithelial and stromal tissues, unlike the low-level image feature-based approaches which are task-dependent representations [19]. Another study proposed a segmentation method to delineate cells using a Gaussian-based hierarchical voting and repulsive balloon model and classify adenocarcinoma and squamous carcinoma [20]. Four texture features were utilized by Al-Kadi (two statistical features and two model-based) who found that combined Gaussian Markov random Feld and run-length matrix texture measures with a Bayesian classifier performed better at classifying meningioma tissue [21]. In order to compare the performance of different pre-trained models like VGG16, Inception [22], ResNet, and NASNet [23], the authors employed transfer learning while applying the collective dataset as well.

We also reviewed some background or related works to this paper. These include modeling transformations with neural networks [24-26], learning and analyzing transformation-invariant representations [27-32] as well as attention and detection mechanisms for feature selection [33-37]. Hinton's early work [24] on assigning canonical frames of reference to object parts is reminiscent of [25], which presented a generative model made up of transformed parts that were created by modeling 2D affine transformations. The inputs in the generative training scheme are images that have been transformed themselves, with the network being provided with an additional input representing a transformation from input images to targets (as shown in Fig. 1). Thus, this is a generative model that can learn to generate transformed images of objects through composing parts. Tieleman extends the concept of a composition of transformed parts by explicitly affinely transforming learnt parts using the predicted transform from the network. Generative capsule models such as these can learn discriminative features for classification from transformation supervision.
Invariants and equivariances of CNN representations through the input image transformations can be determined by estimating linear relationships between original images and transformed ones [31]. Cohen & Welling [28] study this behavior under symmetry groups, which also leads to a model proposed by Gens & Domingos [29] having more invariant feature maps with respect to symmetry groups. Other examples include scattering networks [27], and CNNs that build filter banks of transformed filters [30, 32]. Stollenga et al. [39] use a policy based on the activations of a network to gate the responses of any filters of the network on subsequent forward passes of the same image, thus enabling attention to specific features. In [40] some invariant representations through data manipulation instead of feature extractors were proposed using a clustering algorithm. Neural networks with selective attention accomplish this by cropping parts of the image (i.e., they do not consider whole images) and hence are able to learn translation invariance. Reinforcement learning is used for training purposes in cases such as these where differentiable attention mechanisms have been avoided [33, 37], non-differentiable attention mechanisms are used instead [36] by using Gaussian kernels in generative models. It has been shown that a region proposal algorithm can be used as a form of attention in the work by Girshick et al. [35] and Jaderberg et al. show in [34] that CNNs can be used to regress salient regions.

Jaderberg et al. [41] propose a new learnable module called the Spatial Transformer to address these disadvantages of Convolutional Neural Networks (CNNs) in achieving spatial invariance to input data efficiently. It is conceivable that neural networks, by incorporating this module into existing CNN architectures, gain the power to actively manipulate data spatially on the basis of feature maps without any additional training supervision or modification to the optimization process. The use of Spatial Transformers by the authors results in translation, scale, rotation, and more involved warping invariant models with top performance on different benchmarks and various transformations.

III. PROPOSED METHOD

A. Dataset:

The BreaKHis (Breast Cancer Histopathological Image Classification) dataset [42] is an extensively used reference dataset for breast cancer classification and research. It is highly beneficial in the development and evaluation of CAD systems for detecting and classifying breast cancer. Several important features make the BreaKHis dataset a valuable resource for breast cancer research.

The optical microscopy images of the histopathology slides in the BreaKHis dataset were taken. Breast tumor samples were stained with hematoxylin and eosin (H&E), which is one of the most common staining techniques in histopathology [42]. The images presented here demonstrate the cellular and tissue architecture of breast tumors at high magnification, which make them good candidates for detailed analysis. Annotations serve as ground truth labels for training the classifiers and evaluative measures. This dataset also has annotations on benign, malignant cases, and other subdivisions based on tumor characteristics. Such exhaustive annotation enables researchers to develop and validate computer aided diagnostic tools that can solve binary classification problems like benign versus malignant, or multi-class classification problems like subtypes of breast cancer.

BreaKHis dataset covers different variants of breast cancer such as benign, malignant ductal carcinoma in situ (DCIS), malignant invasive ductal carcinoma (IDC), malignant invasive lobular carcinoma (ILC), and others as given in figures 1 and 2. This dataset's heterogeneity enables examination of the performance of CAD systems in making distinctions between different types of breast cancer and consequently facilitates more tailored treatment decisions.

Furthermore, BreaKHis dataset has a substantial number of images which makes it a big data set for training and evaluation purposes. Concerning this, it contains roughly 7,909 histopathological images which were split into training and test data sets as shown in Table 1. The large scale enables robust model training and testing hence ensuring reliable performance evaluation of CAD systems [42].

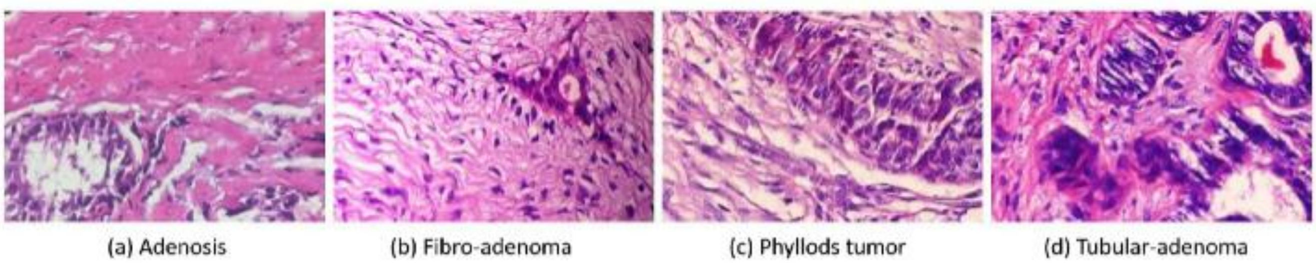

Figure 1. Benign Sub-Classes [43]

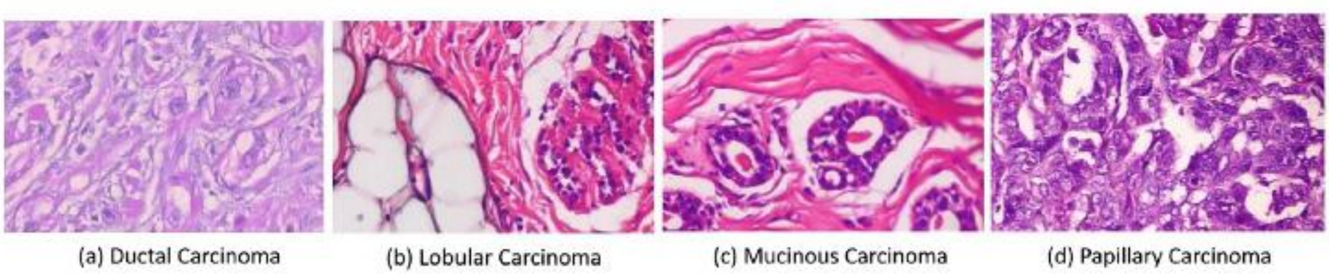

Figure 2. Malignant Sub-Classes [43]

The BreaKHis dataset is a challenge and an opportunity to researchers. In real situations, histopathological images can differ in staining, image quality and tissue presentation. These challenges provide opportunities for creating powerful CAD systems that can

ISSN: 2710-2165

handle different image characteristics. Moreover, the dataset is richly annotated and has large scale that enables the application of modern deep learning methods like convolutional neural networks (CNNs) and attention mechanisms for enhancing the accuracy of breast cancer classification.

TABLE 1. BreaKHis Dataset

Class	Types	Abb.	40X	100X	200X	400X
Benign	Adenosis	A	114	113	111	106
	Fibro	F	253	260	264	237
	Tubular	TA	109	121	108	115
	Phyllodes	PT	149	150	140	130
Malignant	Ductal	DC	864	903	896	788
	Lobular	LC	156	170	163	137
	Mucinous	MC	205	222	196	169
	Papillary	PC	145	142	135	138
Total			1995	2081	2013	1820

To sum up, BreaKHis dataset is an important resource for the study of breast cancer that facilitates development and evaluation of CAD systems for accurate detection and characterization of breast cancer. The large sizes, multiple subtypes and complex descriptions have been employed to push the frontiers further making it possible to develop more effective diagnostic tools and individualized strategies of overcoming this threat.

B. Modified GoogLeNet Architecture:

The GoogLeNet architecture, as described by [44], is improved by the addition of an attention mechanism. Inception modules and parallel convolutional pathways are some of the features that make GoogLeNet architecture known for excelling in image classification tasks. These features enable this model to “see” images from different scales and obtain useful information (Figure 3). However, the model itself has to be modified so as to become attentive on certain areas in images, hence leading to more efficient discrimination in its tasks like object recognition or image classification [45]. This combination offers a powerful technique for integrating GoogLeNet architecture with an attention mechanism already in place, thus enabling the accurate and efficient solution of challenging visual problems.

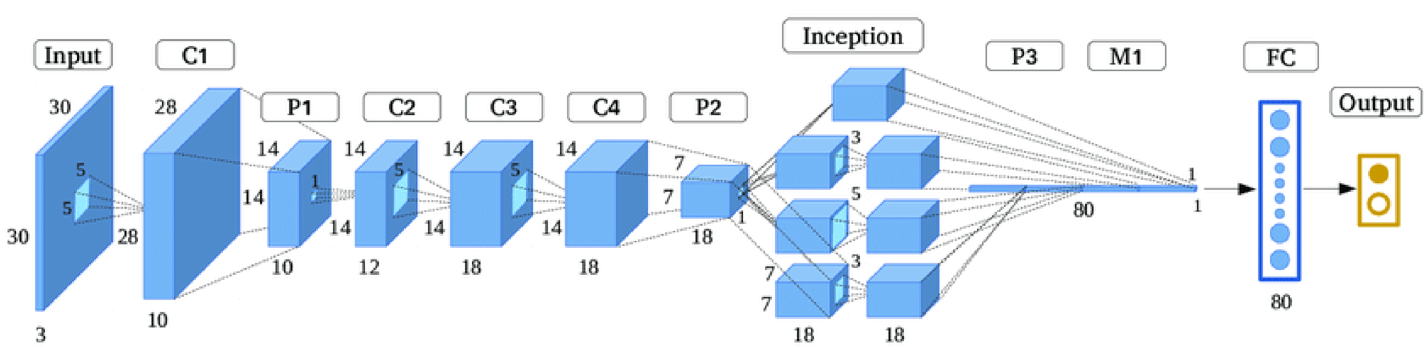

Figure 3. Google Net architecture

C. Attention Mechanism using Spatial Transformer Network (STN):

The attention mechanism has been introduced to the modified GoogLeNet network through seamless integration with the spatial transformer network (STN) [46] to enhance its abilities in image understanding tasks. The STN module is strategically placed after a convolutional layer as shown in Figure 4, it consists of three main sections: localization network, grid generator, and sampler.

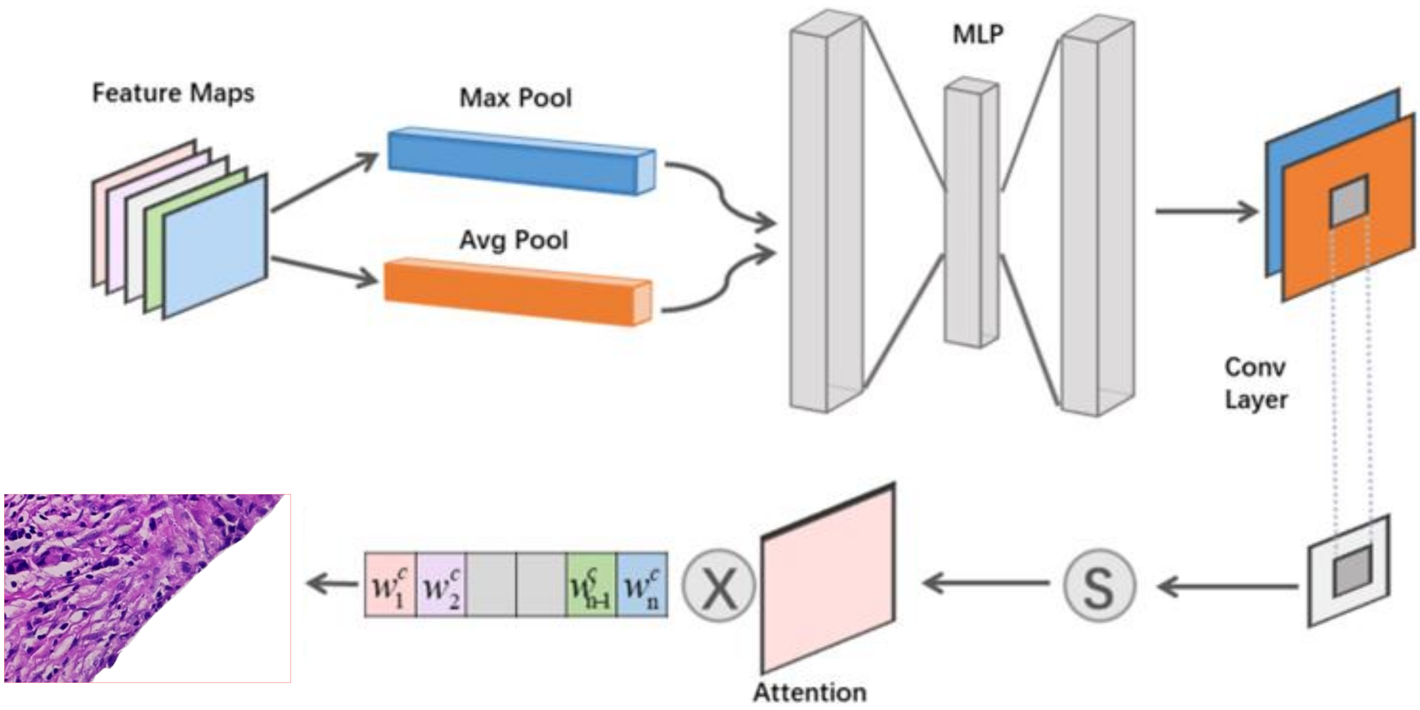

Figure 4. Attention mechanism.

The convolutional layers in the localization network form a localization network that projects transformation parameters that are vital for spatial managements like translation, rotation, and scaling [47]. The model acquires the competence to rotate, dynamically align, and attend to any particular spot of attention within the input image by determining these parameters.

After being predicted by the localization network, the grid generator component uses them as its inputs to create a grid of sampling points. Through this grid, it is possible to ascertain the manner in which to change an input image hence letting this model easily concentrate on different sections of the image depending on its characteristics.

The sampler module works in conjunction with both the input image and the generated grid so as to perform mapping operations through images deformation based on the described grid [48]. For example, this process allows the model to selectively focus on certain informative areas thus improving its feature extraction capacity and its overall performance can be seen in Figure 5.

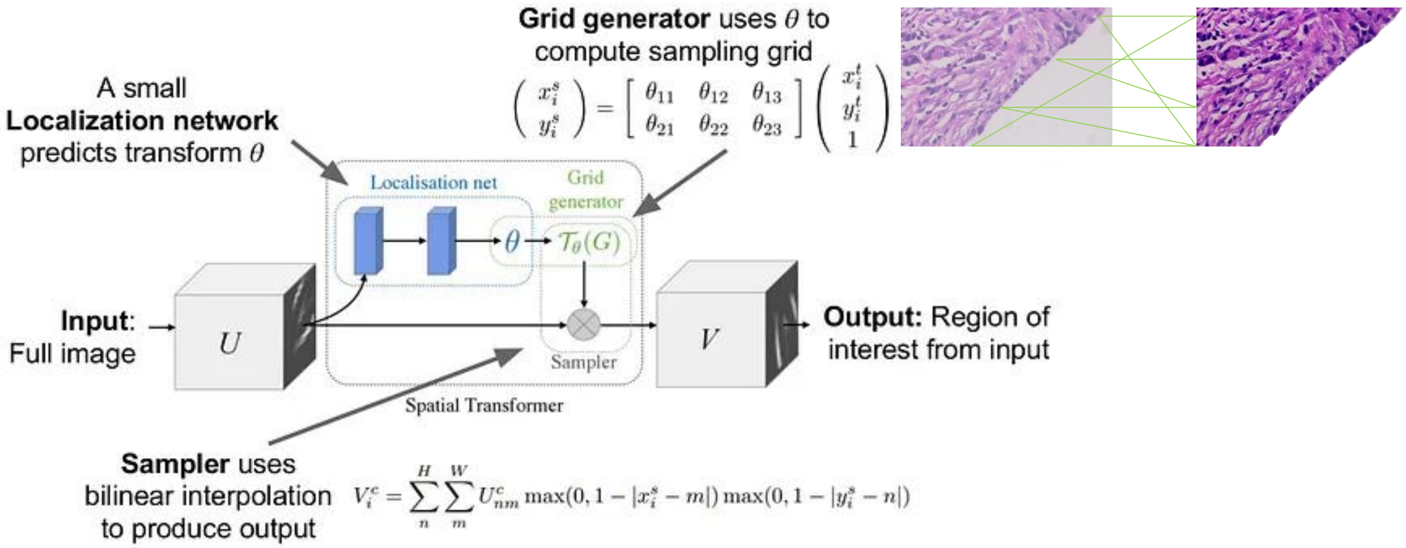

Figure 5. Spatial Transformer Network (STN) workflow

A remarkable modification of the architecture of GoogLeNet is the integration of an attention mechanism through a spatial transformer network. Figure 6 displays how the modified GoogLeNet architecture gains an exceptional feature that dynamically and selectively concentrates on important regions of the input image. The adaptive attention mechanism extends the model's ability to capture relevant features thus improving accuracy as well as robustness during various image understanding tasks among them: object detection, semantic segmentation, and image classification. Combining the established GoogLeNet architecture along with an attention mechanism via STN makes a powerful way of addressing complex visual problems effectively and moving the state-of-the-art in computer vision research forward.

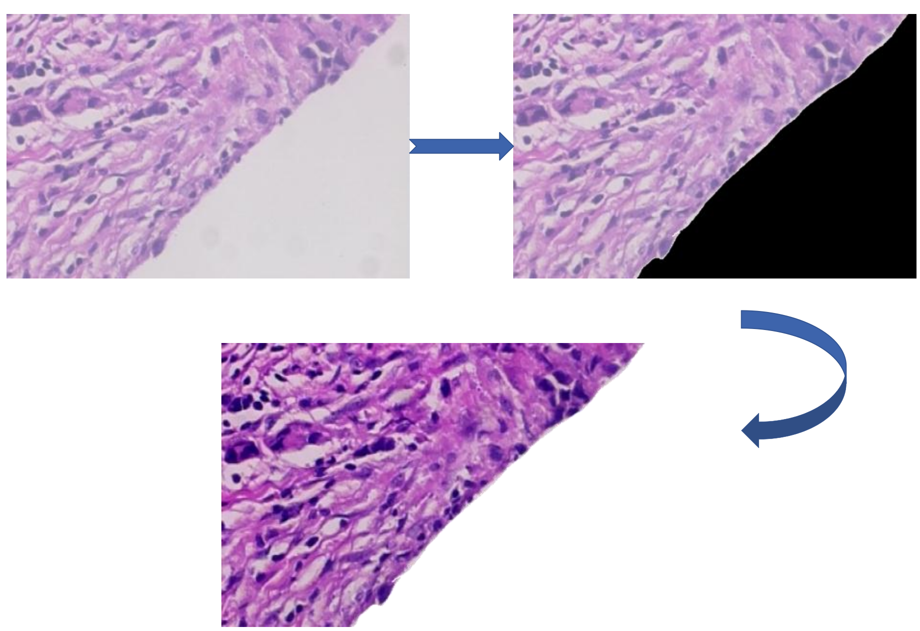

Figure 6. The process of selectively focusing on significant regions

D. Training and Optimization:

The modified GoogLeNet architecture is equipped with the integrated attention mechanism and was trained using the BreaKHis dataset, which is a popular benchmark dataset for breast cancer classification [49]. The training process involves many steps to optimize the model performance.

As input to the network, the labeled histopathological images from the BreaKHis dataset are given. This includes images for which information on the subtype or classification of breast cancer has been provided.

While training, loss is computed by comparing predicted output with ground truth labels using an appropriate loss function like cross-entropy. Therefore, this function allows us to measure how far our model is from predicting accurately.

Optimization algorithms such as stochastic gradient descent (SGD) are utilized to improve the parameters of our model and enhance its performance [50]. In other words, SGD updates (adjusts) network parameters iteratively in a way that minimizes the loss function. Until there is an acceptable performance level or convergence level, this process will be continued iteratively.

More importantly, training of the attention mechanism incorporated within the modified GoogLeNet architecture takes place concurrently with that of the rest of the network. Within images, the attention module has to be able to learn which parts are important for accurate breast cancer classification. These parameters can be matched with those of the rest of the network to jointly optimize them so that the model could adaptively select relevant regions in an end-to-end fashion and thus increasing its discriminative power while improving on classification accuracy.

This training process will make the modified GoogLeNet architecture with the attention mechanism become specialized at detecting and classifying breast cancer subtypes based on histopathological images. It exploits the diverse BreaKHis dataset that is also annotated, hence enabling the model to effectively attend to significant areas resulting in improved results in breast cancer classification tasks.

E. Evaluation and Performance Metrics:

Several investigations have been carried out over the years to ascertain the suitability of this method. In order to determine how effective, it is, a trained model is tested with respect to being successful on a different test set from the BreaKHis dataset. Thus, we can conclude that this method can be used for breast cancer classification by demonstrating performance metrics such as Accuracy, Precision, Recall, ROC Curve, Confusion Matrix, and F1 Score [51].

The accuracy metric shows us the accuracy of the models' predictions as a ratio of correctly classified samples (predicted) in relation to all samples in the test set. Precision measures true positive instances relative to positively predicted cases indicating its ability to correctly recognize breast cancer incidences. Also called sensitivity, recall calculates true positive cases divided by all actual positive cases hence underlining its capacity for properly finding breast cancer incidences. The F1 score combines precision and recall into one metric and offers an equal measure of the model’s performance [51-52].

$$Accuracy = (tp + tn)/total\ samples\ (1)$$

$$Sensitivity = tp/(tp + fn)\ (2)$$

$$Specificity = tn/(tn + fp)(3)$$

$$Precision = tp/(tp + fp)(4)$$

$$F1\ score = 2 * (precision * sensitivity) / (precision + sensitivity)(5)$$

The ROC curve is a graphic representation of the model’s false positive rate (1-specificity) against the true positive rate (sensitivity) at different classification thresholds. It helps to evaluate the performance of the model under different threshold settings as well as its ability to discriminate [52].

The confusion matrix is an array that demonstrates the accuracy of a machine-learning algorithm in classifying problems. The true positives, true negatives, false negatives, and false positives are given by it, enabling an in-depth analysis of how well a model performs [52].

By comparing the performance of modified GoogLeNet architecture with attention mechanism to the performance of original GoogLeNet architecture, one can assess whether there is any additional value in the attention mechanism in improving breast cancer classification accuracy. These assessments help in understanding if focusing only on crucial parts of images would improve their ability to identify cases of breast cancer correctly.

In short, this method is aimed at improving GoogLeNet architecture's classification through embedding spatial transformation-based attention mechanisms within. The proposed method aims to improve the precision and dependability of breast cancer classification, which can lead to more effective diagnosis and treatment decisions by selectively attending to pertinent image regions.

F. Binary Classification for BreaKHis:

The modified GoogLeNet architecture for breast cancer classification with an attention mechanism was proposed and can be employed for both binary and multi-class classification tasks using the BreaKHis dataset. The BreaKHis dataset contains annotations on different types of breast cancer, making it possible to examine the performance of the devised method under various classification settings.

In binary breast cancer classification [53-54], training of the modified GoogLeNet architecture which has the attention mechanism is done on the BreaKHis dataset that contains annotated histopathological images classified as benign and malignant. The model's parameters are optimized for this purpose during training to ensure accurate classifications of the images as either benign or malignant. For this reason, we have developed a model that captures significant features that differentiate benign from malignant tumors. This is then complemented by attention mechanisms in order to identify informative regions in images. In this way, we are able to focus on slight variations and any other important part that will aid us in making precise classifications.

Different image areas are assigned importance scores or weights by the attention mechanism indicating their relevance for classification. This in turn allows the model to flexibly focus on those regions which are most discriminative, hence improving its ability to discriminate between benign and malignant tumors. With the integration of an attention mechanism into the modified GoogLeNet architecture, it becomes more effective at capturing and utilizing crucial image patches for distinguishing features. Consequently, this in turn enhances its discriminative power thereby leading to a better percentage of accurately classified breast tumor samples as either benign or malignant.

G. Multi-class Classification:

The BreakHis dataset has annotations for various subtypes of breast cancer, making it possible to classify histopathological images into different classes. Such subtypes may include benign, ductal carcinoma in situ (DCIS), invasive ductal carcinoma (IDC) and others [55-57]. For Binary and multi-class classification, GoogLeNet architecture has been modified to include attention mechanism. This model comes with one extra output node per subtype that is designed to capture these types precisely. In other words, during training, the model was given information that helped it to differentiate between different subtypes of breast cancer. Therefore, for changes in the attention mechanisms that require a focus on informative regions within images differentiating against other subtypes of cancer were added. The error of misclassification is minimized and the accuracy of assigning correct image type is improved by optimizing its parameters. The training starts with images going through the network, loss being computed based on predicted class probabilities and model parameters updated via backpropagation. Hence at the end of training, one can examine how the model performed on a separate test set in terms of multi-class breast cancer classification. This will involve computing performance metrics like accuracy, precision, recall and F1 score to establish how well the proposed model correctly classifies histopathological images into their respective subtypes.

For BreaKHis dataset, if we use this technique for binary and multi-class classification, it will enable us tell how efficient our modified GoogLeNet architecture is in classifying breast cancer subtypes with an attention mechanism integrated into it. Such evaluations indicate its versatility in different classification situations that improve the strength and flexibility of CAD systems used for early detection of breast cancer.

IV. RESULTS

This part shows the findings of binary and multi-class classification tests on breast cancer using histopathological pictures. Two versions of GoogleNet deep neural networks that were included in these tests are pre-trained GoogleNet and modified GoogleNet with attention mechanism. BreaKHis dataset was employed for training and testing purposes, which is available to the public. The data was split in such a way as to have unbiased assessment and stable performance evaluation: 80% went to training set and 20% remained in test set. The 5-fold cross-validation strategy is used which involves dividing data into five folds or subsets, whereby each model is trained and assessed five times with different folds serving as test sets while others are train sets. This technique ensures that possible bias in the data does not affect how robustly these models are evaluated for performance.

Transfer Learning (TL) techniques were used during training of deep neural networks. By using this approach, it becomes possible to leverage knowledge obtained from previously trained models on diverse yet similar tasks or datasets. In this case, the fine-tuning process involved taking weights from pre-trained GoogleNet model. For testing purposes and 5-fold cross-validation, two GoogleNet architectures are used: pre-trained GoogleNet and modified GoogleNet with an attention mechanism. In evaluating the performances of these networks, several performance measures such as precision, recall, F1 score, etc., depending on the evaluation criteria employed in this study. These analyses and evaluations help in determining the effectiveness of pre-trained GoogleNet and modified GoogleNet with attention mechanism in binary and multi-class breast cancer classification using histopathology images. Model performance assessment of the BreaKHis dataset is adequately done using results from the 5-fold cross-validation method.

A. Performance analyzing

We show in this section how two Deep Neural Network (DNN) classifiers; GoogleNet pre-trained and modified GoogleNet with attention mechanism, performed. Testing of these classifiers involved the utilization of BreaKHis dataset which consists of histopathology images related to breast cancer. Meanwhile, the performance of the DNN classifiers is evaluated by comparing tables 2, 3, and 4 for both binary and multi-class classification. Furthermore, more details are provided in Figures 7, 8, 9, 10 and 11.

TABLE 2. Evaluation Measures for the Google Net and Proposed Model (Binary Classification)

	CNN Type	factors	Accuracy	Sensitivity	Specificity	Precision	F1 score
1	Google Net	40X	94.99	94.39	94.39	94.01	94.20
		100X	93.51	92.74	92.74	92.21	92.46
		200X	93.95	93.17	93.17	92.60	92.88
		400X	94.51	94.17	94.17	93.43	93.78
2	Proposed Model	40X	97.24	96.69	96.69	96.89	96.79
		100X	96.88	96.67	96.67	96.09	96.37
		200X	97.77	97.50	97.50	97.29	97.40
		400X	98.08	97.92	97.92	97.71	97.81

TABLE 3. Evaluation Measures for the Google Net and Proposed Model (Multi-Class classification with 100x)

	CNN Type	Sub-Class	Accuracy	Sensitivity	Specificity	Precision	F1 score
1	Google Net	Adenosis	77.78	91.30	98.98	84	87.50
		Fibro	90.74	94.23	99.45	96.08	95.15
		Tubular	84.62	91.67	99.49	91.67	91.67
		Phyllodes	81.82	90	99.22	90	90
		Lobular	88.89	94.12	99.48	94.12	94.12

ISSN: 2710-2165

		Ductal	97.25	97.79	99.57	99.44	98.61
		Mucinous	81.25	88.64	98.92	90.70	89.66
		Papillary	78.13	92.59	98.71	83.33	87.72
		Average	85.06	92.54	99.22	91.17	91.80
2	Proposed Model	Adenosis	91.67	95.65	99.74	95.65	95.65
		Fibro	94.44	98.08	99.45	96.23	97.14
		Tubular	95.83	95.83	100	100	97.87
		Phyllodes	96.67	96.67	100	100	98.31
		Lobular	94.29	97.06	99.74	97.06	97.06
		Ductal	97.81	98.90	99.15	98.90	98.90
		Mucinous	93.48	97.73	99.46	95.56	96.63
		Papillary	92.86	96.30	99.74	96.30	96.30
		Average	94.63	97.03	99.66	97.46	97.23

TABLE 4. Evaluation Measures Average for the Google Net and Proposed Model (Multi-Class classification with 100x)

	CNN Type	**Accuracy**	**Sensitivity**	**Specificity**	**Precision**	**F1 score**
1	Google Net	85.06	92.54	99.22	91.17	91.80
2	Proposed Model	94.63	97.03	99.66	97.46	97.23

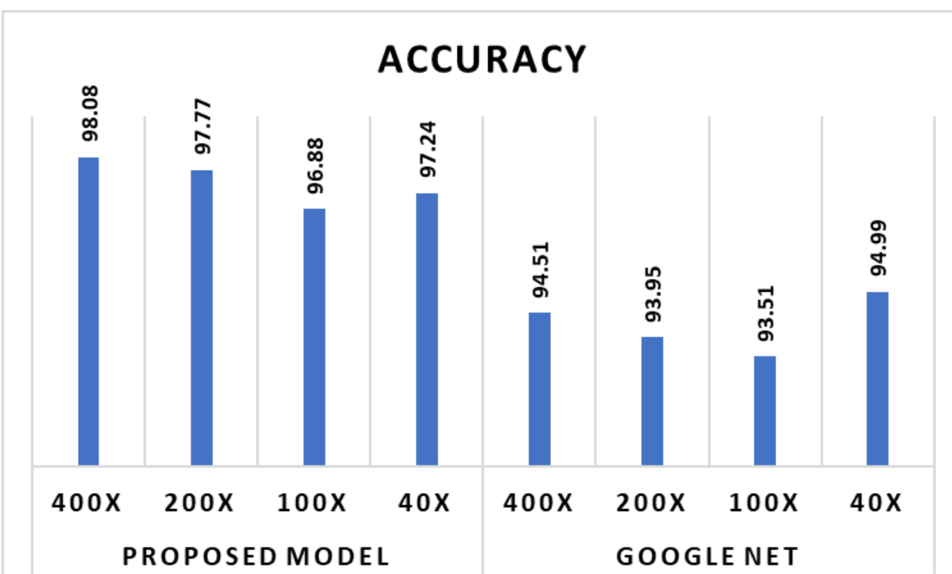

Figure 7. DNN Accuracy at Multiple Magnifications for Google Net and Proposed Model (Binary classification)

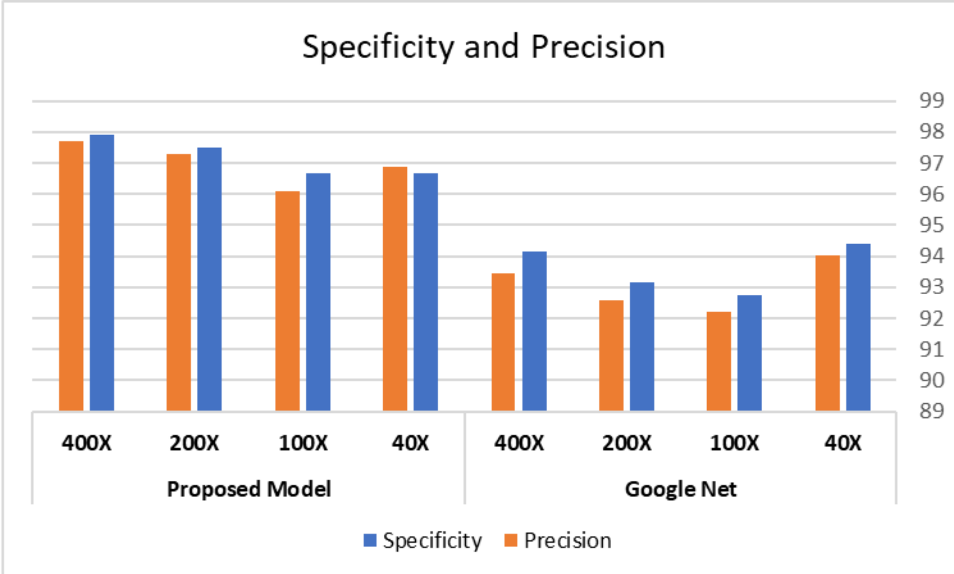

Figure 8. Specificity and Precision at Multiple Magnifications for Google Net and Proposed Model (Binary classification)

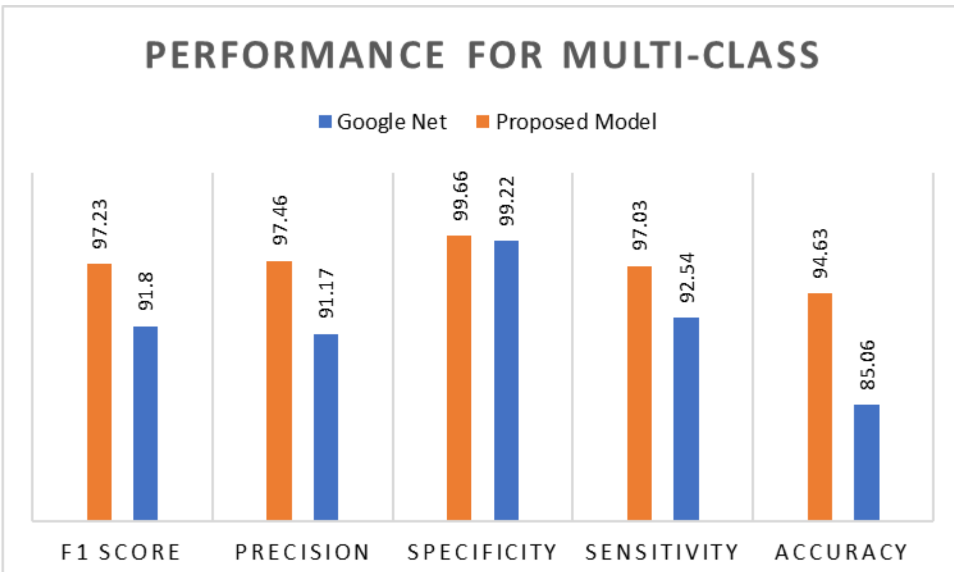

Figure 9. Performance analysis for Google Net and Proposed Model (Multi-Class classification)

It is worth noting that the attention mechanism of the modified GoogleNet has been shown to have the highest performance in all measured metrics. This suggests that addition of attention mechanism in this model helps it to focus on informative regions within images thus improving its classification performance. However, it is important to remember that the pre-trained network mentioned earlier acts as a feature extractor for these classification tasks. These DNN classifiers take advantage of the learned representations and contribute to improving their overall efficacy by using pre-trained weights and knowledge from a model trained on a related task or dataset.

These are details of Figures 10 and 11 which show the summary of how well classical GoogleNet and modified GoogleNet with attention mechanism did on the BreaKHis dataset. These results imply that deep learning methods can be employed in breast cancer classification based on histopathology images.

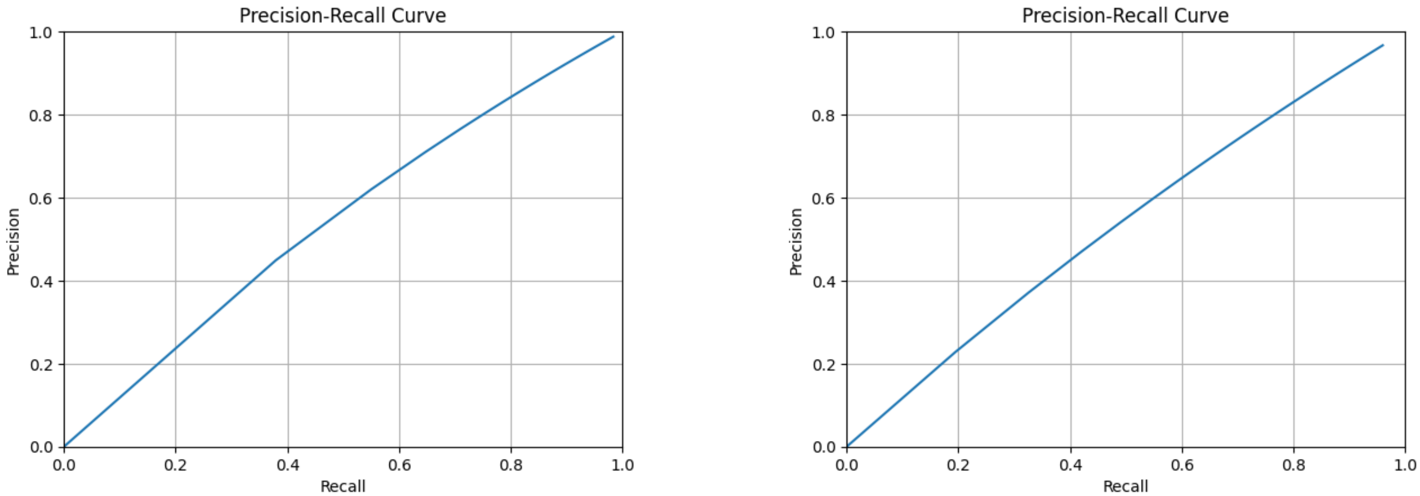

Figure 10. Precision-Recall Curves for Proposed Model (400X) and Google Net (40x)

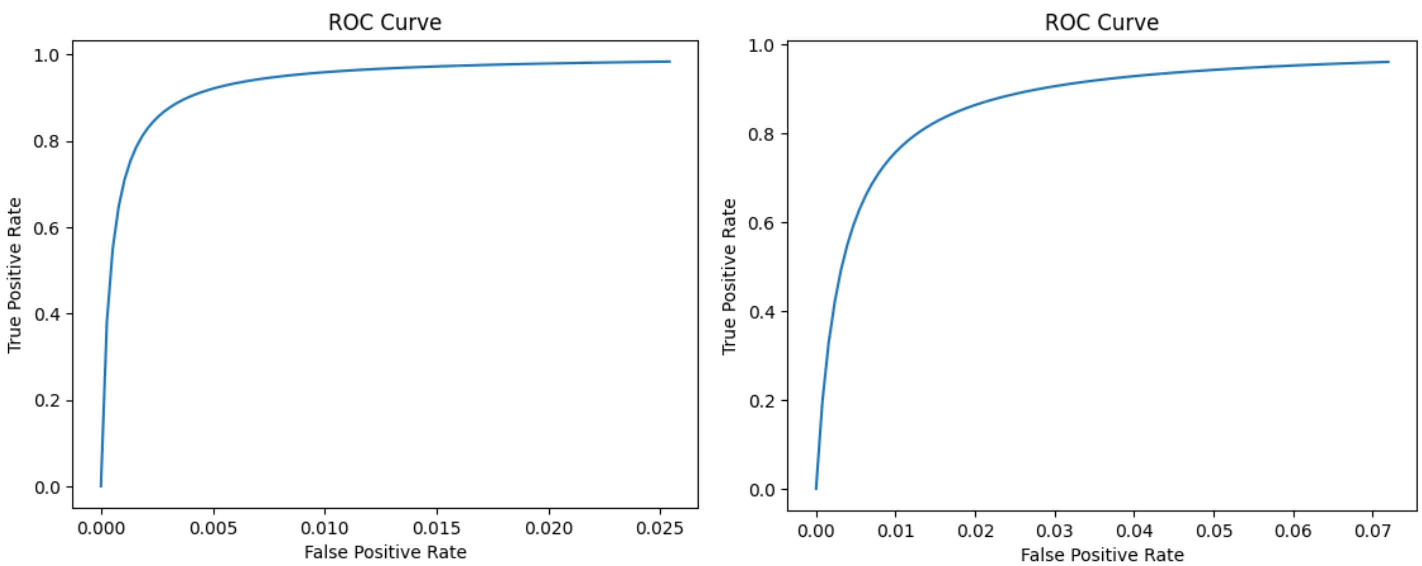

Figure 11. ROC Curves for Proposed Model (400x) and Google Net (40x)

B. Confusion matrices

Machine learning uses confusion matrices extensively to assess classification model accuracy and performance. They provide a tabular summary comparing actual and predicted classes, to give insights about the model's performance. For the binary and multi-class classification, such as those in Deep Neural Networks (DNNs)'s context, confusion matrices are used to evaluate their performance. The two DNNs' confusions matrices in Tables 5 and 6 represent confusion matrices tailored specifically for binary and multi-class classifications respectively. Each matrix indicates how correct and incorrect predictions are distributed across various classes, therefore offering detailed insights into the performance of each DNN.

Investigating the cells that make up the confusion matrix reveals the level of success in identifying a given class to which an observation belongs. A binary classification confusion matrix has four entries.

There are True Negatives (TN), which represent the number of correctly predicted instances that belong to the negative class, False Positives (FP) indicating incorrectly predicted negative cases as positives, False Negatives (FN) representing positive instances incorrectly predicted as negative cases, and True Positives (TP) indicating correctly predicted positive instances.

TABLE 5. Confusion Matrix for Binary Classification

	CNN Type	Factors	TP	FP	FN	TN	A%
1	Google Net	40X	116	9	11	263	94.99
		100X	117	12	15	272	93.51
		200X	114	11	14	264	93.95
		400X	110	8	12	234	94.51
2	Proposed Model	40X	119	6	5	269	97.24
		100X	124	5	8	279	96.88
		200X	121	4	5	273	97.77
		400X	115	3	4	242	98.08

In multi-classification, the confusion matrix is expanded to cover more classes. In the matrix, each row represents the real class while each column represents the predicted class. Entries in the matrix give the number of samples from a particular actual class that were predicted as belonging to a certain predicted class.

TABLE 6. Confusion Matrix for Multi-Class Classification with 100x

	CNN Type	Sub-Class	**A**	F	T	Ph	**L**	**D**	**M**	**Pa**
1	Google Net	Adenosis	21	0	1	0	0	0	1	0
		Fibro	1	49	0	0	1	0	0	1
		Tubular	0	1	22	0	0	0	1	0
		Phyllodes	1	0	1	27	1	0	0	0
		Lobular	0	0	0	0	32	0	1	1
		Ductal	1	0	0	1	0	177	1	1
		Mucinous	0	1	0	1	0	1	39	2
		Papillary	1	0	0	1	0	0	0	25
2	Proposed Model	Adenosis	22	0	0	0	0	0	1	0
		Fibro	0	51	0	0	0	0	0	1
		Tubular	0	0	23	0	0	1	0	0
		Phyllodes	0	0	0	29	1	0	0	0
		Lobular	0	1	0	0	33	0	0	0
		Ductal	1	0	0	0	0	179	1	0
		Mucinous	0	1	0	0	0	0	43	0
		Papillary	0	0	0	0	0	1	0	26

Insight can be gained into the DNN strengths and weaknesses by examining confusion matrices. It helps to know how models behave in terms of correct or wrong predictions across various classes as this identifies the areas where the most improvement can be made in classification performance. In totality, the confusion matrices give a full evaluation of the DNNs' performance through binary and multi-class classification tasks, which include accuracy and misclassifications.

V. CONCLUSION

All in all, the findings from this study are consistent with those of earlier investigations. This way, the breast cancer diagnosis success rate can be increased by using a modified GoogLeNet architecture and attention mechanism on image regions. To this end, a modified GoogLeNet architecture with image region attention based on a modified GoogLeNet architecture was proposed.

In conclusion, the proposed method yielded promising results for improving the accuracy of breast cancer classification using the BreaKHis dataset. The performed experiments have demonstrated that it is an effective approach.

It is worth mentioning that binary classification was achieved with 98.08% accuracy of the modified GoogLeNet with attention mechanism which outperformed the original GoogLeNet achieving 94.99% accuracy. At 100x magnification, GoogLeNet achieved 85.06% accuracy while the proposed model had an accuracy of 94.63% for multi-class. These results show significant improvement in breast cancer classification using the proposed method.

Subtle patterns indicating various subtypes of breast cancer can be best detected when we include attention mechanisms into the modified GoogLeNet at the object level including salient regions only by which the model gets more discriminative power to distinguish between benign and malignant tumors. The major benefit of such an approach is its ability to identify benign tumors and reveal some patterns indicative for particular forms of malignancy among them as well as for varying degrees of tumor aggressiveness.

The research found in this study established the possibility of using attention-based CNN models as a way of improving the accuracy and efficiency of computer-aided diagnosis systems for breast cancer. The combination of deep learning algorithms and attention mechanisms helps to detect slight abnormalities and support radiologists and pathologists respectively in making more accurate and informed diagnostic choices.

Although the study has shown promise in the BreaKHis dataset, there is a need for further investigation and validation on bigger and more varied datasets to be able to evaluate its generalizability as well as robustness. Moreover, there is also a need to explore how well it performs on other histopathological datasets by comparing it with current state-of-the-art methodologies.

In summary, this new methodology, through CAD systems, represents a significant step forward in the classification of breast cancer. Such an approach could potentially lead to better patient outcomes by aiding healthcare professionals in making more precise diagnoses and treatment plans. More work needs to be done on deep learning techniques and attentional mechanisms in this field so that CAD becomes more effective in clinical practice for breast cancer diagnosis.

REFERENCES

[1] M. C. Chun. (2018). Breast Cancer: Symptoms, Risk Factors, and Treatment, Medical News Today. Accessed: Mar. 10, 2018. [Online]. Available: https://www.medicalnewstoday.com/articles/37136.php

[2] World Health Organization. Accessed: Mar. 10, 2018. [Online]. Available: http://www.who.int/en/

[3] P. Boyle and B. Levin. (2008). World Cancer Report. [Online]. Available: http://www.iarc.fr/en/publications/pdfs-online/wcr/2008/wcr2008.pdfB. Smith, “An approach to graphs of linear forms (Unpublished work style),” unpublished.

[4] [World Health Organization. Breast Cancer. Available online: https://www.who.int/news-room/fact-sheets/detail/breast-cancer (accessed on 15 June 2022).

[5] Wilkinson, Louise, and Toral Gathani. "Understanding breast cancer as a global health concern." The British Journal of Radiology 95.1130 (2022): 20211033.

[6] M. L. Giger, H. Chan, and J. Boone, “Anniversary Paper: History and status of CAD and quantitative image analysis : The role of medical physics and AAPM,” Med. Phys., vol. 35, no.12, pp. 5799–5820, 2008, doi: 10.1118/1.3013555.]

[7] Q. Li and R. M. Nishikawa, “Computer aided detection and diagnosis in medical imaging: A review of clinical and educational applications,” in Proceedings of the Fourth International Conference on Technological Ecosystems for Enhancing Multiculturality - TEEM ’16, 2016, pp.1–425, doi: https://doi.org/10.1145/3012430.3012567].

[8] M. Kashif, K. R. Malik, S. Jabbar, and J. Chaudhry, Application of machine learning and image processing for detection of breast cancer. Elsevier Inc., 2020.]

[9] [P. Taylor, J. Champness, K. Johnston, and H. Potts, “Impact of computer-aided detection prompts on screening mammography,” 2005. doi: 10.3310/hta9060.

[10] Zhang H, Han L, Chen K, et al. Diagnostic efciency of the breast ultrasound computer-aided prediction model based on convolutional neural network in breast cancer. J Digit Imaging. 2020;33:1218–23.

[11] Araújo T, Aresta G, et al. Classifcation of breast cancer histology images using convolutional neural networks. PLoS ONE. 2017;12:e0177544.

[12] Nahid AA, Mehrabi MA, Kong Y. Histopathological breast cancer image classifcation by deep neural network techniques guided by local clustering. Biomed Res Int. 2018;2018:2362108.

[13] Arevalo J, Gonza´lez FA, Ramos-Polla´n R, Oliveira JL, Lopez MAG. Representation learning for mammography mass lesion classifcation with convolutional neural networks. Comput Methods Programs Biomed.2016;127:248–57.

[14] Huynh BQ, Li H, Giger ML. Digital mammographic tumor classifcation using transfer learning from deep convolutional neural networks. J Med Imaging. 2016;3:034501.

[15] Yeşim E, Muhammed Y, Ahmet C. Convolutional neural networks based classifcation of breast ultrasonography images by hybrid method with respect to benign, malignant, and normal using mRMR. Comput Biol Med. 2021;133:104407.

[16] Zheng Y, et al. Feature extraction from histopathological images based on nucleus-guided convolutional neural network for breast lesion classifcation. Pattern Recogn. 2017;71:14–25.

[17] Van Eycke YR, et al. Segmentation of glandular epithelium in colorectal tumours to automatically compartmentalise IHC biomarker quantifcation: a deep learning approach. Med Image Anal. 2018;49:35–45.

[18] Sudharshan PJ, et al. Multiple instance learning for histopathological breast cancer image classifcation. Expert Syst Appl. 2019;117:103–11.

[19] Xu J, Luo X, Wang G, Gilmore H, Madabhushi A. A deep convolutional neural network for segmenting and classifying epithelial and stromal regions in histopathological images. Neurocomput. 2016;191:214–23.

[20] Zhang X, et al. High-throughput histopathological image analysis via robust cell segmentation and hashing. Med Image Anal. 2015;26:306–15.

[21] Al-Kadi OS. Texture measures combination for improved meningioma classifcation of histopathological images. Pattern Recogn. 2010;43:2043–53.

[22] Szegedy C, Vanhoucke V, Iofe, Shlens J, Wojna Z. Rethinking the Inception Architecture for Computer Vision. arXiv 2015; reprint arXiv:1512.00567.

[23] Barret Z, Vasudevan V, Shlens J, Quoc VL. Learning transferable architectures for scalable image recognition. arXiv 2017; preprint arXiv:1707.07012.

[24] [12] G. E. Hinton. A parallel computation that assigns canonical object-based frames of reference. In IJCAI, 1981.

[25] [13] G. E. Hinton, A. Krizhevsky, and S. D. Wang. Transforming auto-encoders. In ICANN. 2011.

[26] [27] T. Tieleman. Optimizing Neural Networks that Generate Images. PhD thesis, University of Toronto, 2014

[27] [3] J. Bruna and S. Mallat. Invariant scattering convolution networks. IEEE PAMI, 35(8):1872–1886, 2013.

[28] [5] T. S. Cohen and M. Welling. Transformation properties of learned visual representations. ICLR, 2015.

[29] [8] R. Gens and P. M. Domingos. Deep symmetry networks. In NIPS, 2014.

[30] [17] A. Kanazawa, A. Sharma, and D. Jacobs. Locally scale-invariant convolutional neural networks. In NIPS,2014.

[31] [19] K. Lenc and A. Vedaldi. Understanding image representations by measuring their equivariance and equiv alence. CVPR, 2015.

[32] [25] K. Sohn and H. Lee. Learning invariant representations with local transformations. arXiv:1206.6418,2012.

[33] [1] J. Ba, V. Mnih, and K. Kavukcuoglu. Multiple object recognition with visual attention. ICLR, 2015.

[34] [6] D. Erhan, C. Szegedy, A. Toshev, and D. Anguelov. Scalable object detection using deep neural networks.

[35] [9] R. B. Girshick, J. Donahue, T. Darrell, and J. Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In CVPR, 2014.

[36] [11] K. Gregor, I. Danihelka, A. Graves, and D. Wierstra. Draw: A recurrent neural network for image generation. ICML, 2015.

[37] [23] P. Sermanet, A. Frome, and E. Real. Attention for fine-grained categorization. arXiv:1412.7054, 2014.

[38] [27] T. Tieleman. Optimizing Neural Networks that Generate Images. PhD thesis, University of Toronto, 2014.

[39] [26] M. F. Stollenga, J. Masci, F. Gomez, and J. Schmidhuber. Deep networks with internal selective attention through feedback connections. In NIPS, 2014.

[40] [7] B. J. Frey and N. Jojic. Fast, large-scale transformation-invariant clustering. In NIPS, 2001.

[41] Jaderberg, Max, Karen Simonyan, and Andrew Zisserman. "Spatial transformer networks." Advances in neural information processing systems 28 (2015).

[42] Spanhol F. A., Oliveira L. S., Petitjean C., Heutte L., "Breast cancer histopathological image classification using Convolutional Neural Networks," 2016 International Joint Conference on Neural Networks (IJCNN), 2016, pages 2560-2567.

[43] [39] Sharma, S., Mehra, R., & Kumar, S. (2021). Optimised CNN in conjunction with efficient pooling strategy for the multi-classification of breast cancer. IET Image Processing, 15(4), 936-946.]

[44] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, “Going deeper with convolutions,” CVPR, 2015.]

[45] Yang, Le, et al. "GoogLeNet based on residual network and attention mechanism identification of rice leaf diseases." Computers and Electronics in Agriculture 204 (2023): 107543.

[46] Jaderberg, Max, Karen Simonyan, and Andrew Zisserman. "Spatial transformer networks." Advances in neural information processing systems 28 (2015).

[47] Zhu, Xizhou, et al. "An empirical study of spatial attention mechanisms in deep networks." Proceedings of the IEEE/CVF international conference on computer vision. 2019.

[48] Liu, Dichao, Yu Wang, and Jien Kato. "Supervised Spatial Transformer Networks for Attention Learning in Fine-grained Action Recognition." VISIGRAPP (4: VISAPP). 2019.

[49] Sun, Jiamei, and Alexander Binder. "Comparison of deep learning architectures for H&E histopathology images." 2017 IEEE Conference on Big Data and Analytics (ICBDA). IEEE, 2017.

[50] Hardt, Moritz, Ben Recht, and Yoram Singer. "Train faster, generalize better: Stability of stochastic gradient descent." International conference on machine learning. PMLR, 2016.

[51] D. Martin and W. Powers, “Evaluation: From precision, recall and F-measure to ROC, informedness, markedness & correlation,” J. Mach. Learn. Technol., vol. 2, no. 1, pp. 37–63, 2011, doi: 10.9735/2229-3981

[52] Kelleher, John D., Brian Tierney, and Aoife D'Arcy. "Chapter 8 - Evaluating Models." Data Science An Introduction, 2nd ed., Elsevier, 2018, pp. 189-214.

[53] Zerouaoui, Hasnae, and Ali Idri. "Deep hybrid architectures for binary classification of medical breast cancer images." Biomedical Signal Processing and Control 71 (2022): 103226.

[54] Abdulaal, Alaa Hussein, et al. "A self-learning deep neural network for classification of breast histopathological images." Biomedical Signal Processing and Control 87 (2024): 105418.

[55] Nawaz, Majid, Adel A. Sewissy, and Taysir Hassan A. Soliman. "Multi-class breast cancer classification using deep learning convolutional neural network." Int. J. Adv. Comput. Sci. Appl 9.6 (2018): 316-332.

[56] Heenaye-Mamode Khan, Maleika, et al. "Multi-class classification of breast cancer abnormalities using Deep Convolutional Neural Network (CNN)." Plos one 16.8 (2021): e0256500.

[57] Mi, Weiming, et al. "Deep learning-based multi-class classification of breast digital pathology images." Cancer Management and Research (2021): 4605-4617.