


Breast Cancer Histopathological Image Classification Using Stochastic Dilated Residual Ghost Model

Ramgopal Kashyap, Amity School of Engineering and Technology, Amity University, Raipur, India

 <https://orcid.org/0000-0002-5352-1286>

ABSTRACT

A new deep learning-based classification model called the stochastic dilated residual ghost (SDRG) was proposed in this work for categorizing histopathology images of breast cancer. The SDRG model used the proposed multiscale stochastic dilated convolution (MSDC) model, a ghost unit, stochastic upsampling, and downsampling units to categorize breast cancer accurately. This study addresses three primary issues. First, strain normalization was used to manage color divergence. Data augmentation with several factors was used to handle the overfitting. The second challenge is extracting and enhancing tiny and low-level information such as edge, contour, and color accuracy; it is done by the proposed multiscale stochastic and dilation unit. The third contribution is to remove redundant or similar information from the convolution neural network using a ghost unit. According to the assessment findings, the SDRG model scored overall 95.65% accuracy rates in categorizing images with a precision of 99.17%, superior to state-of-the-art approaches.

KEYWORDS

Channel Dilated Convolution Unit, Deep Learning, Dilation Convolution, Multiscale Stochastic Dilated Convolution Model, Spatial Dilated Convolution Unit, Stochastic Dilated Residual Ghost Model

1. INTRODUCTION

Breast Cancer (BC) is the primary reason behind the death of women globally and is answerable for numerous deaths every year (Afifi et al., 2019). BC cases are rising daily, and patients are increasing each day; health sciences are still troubled by the upper and proper prognosis of cancer and deep learning. Machine learning is attempting to supply technical advancements. As per the medical specialty reports, carcinoma breast cancer is hard to hunt out at the initial stage; the survival rate can increase if detected at the initial stage. Numerous methods and models have developed, and deep learning involvement gives rise to cancer's accurate and fast prognosis (Pacilè et al., 2020). There are several risk factors for cancer like increasing age, family history, biological time, internal secretion. BC has also reported at younger age women if having a primary child at an older age and women have never been pregnant. BC arises because the breast cell gets started growing unusually, and cells split quickly from the usual structure, frame a lump, and unfold to various parts (Yıldırım, 2018). Types of breast carcinoma like adenosis, fibroadenoma, phyllodes, cylindrical adenoma, mucinous carcinoma are shown in figure 1.

DOI: 10.4018/IJIRR.289655

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

Figure 1. Four types of benign (first row) (a) adenosis (b) fibroadenoma (c) phyllodes tumor (d) tubular adenoma and malignant (second row) (a) ductal carcinoma (b) lobular carcinoma (c) mucinous carcinoma (d) papillary carcinoma tumor images from the Breast Cancer Histopathological Image Classification (BreakHis) dataset (Spanhol et al., 2016a).

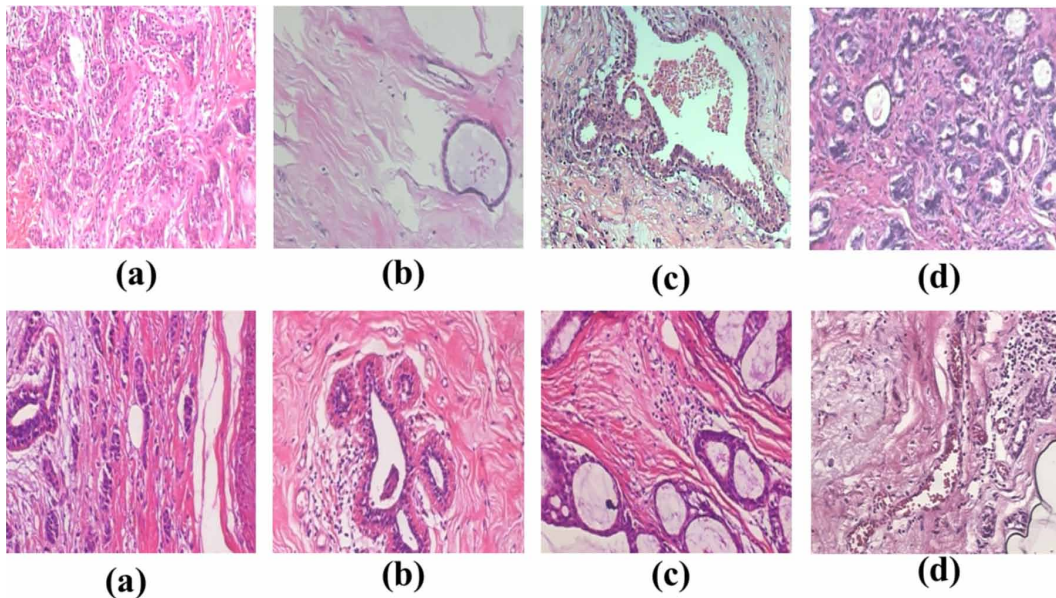


Image classification has shown performance improvements using deep learning models (LeCun et al., 2015). The AlexNet model provides better results than conventional machine learning methods in detecting benign and malignant cancer cells (Spanhol et al., 2016b). State of the art has shown that deep learning models providing 83% accuracy (Bayramoglu N. et al., 2016), the pre-trained convolution neural network (CNN) model extracted the deep convolutional activation features from the network achieved 90% (Spanhol F. et al., 2017), and that BiCNN model up to 97% (Wei B. et al., 2017) accuracy in breast cancer detection. Deep learning still offers good precision with a large volume of data and shows its presence in detecting mitosis (Chen H. et al., 2016) and metastasis (Wang D. et al., 2016). The classification of histological images is very ambitious since invasive carcinoma and stroma have fragile borders, and the core element for the detection of cancer is tissue structures (Monaco J. et al., 2012) (Liu F. and Yang L., 2015). Hematoxylin and eosin (H&E) images show a variable appearance even with the same malignancy tissue biopsies (Giannakeas N. et al., 2017). Other issues with breast cancer detection with deep learning models are the availability of limited datasets, which cause overfitting (Pang et al., 2018) (Gavrilo et al., 2018); it also causes a lack of feature evaluation and results in cell imaging desegregation (Dapson et al., 2010). The VGG model uses VGG16 and a neural network to extract the properties (Gultom et al., 2018); the VGG16, VGG19, and ResNet50 models use a layered architecture but do not work well with smaller training datasets (Shallu & Mehra, 2018). AlexNet booths with large color images, LeNet5 has reduced training time and exceeds performance (Huang et al., 2015). The Inception v3 model based on deep learning was used to classify lung cancer and be a reliable model (Coudray et al., 2018). The contribution of this research is to design a model for the precise detection of nuclei, which remains a challenge as H&E images use the blue/purple color to represent the nuclei and the pink color to represent the cytoplasm very complicated (Veta et al., 2014). The main challenge is the presence of the maximum core in high-resolution images, noise, and variations in the core's texture, intensity, and shape (Fuyong Xing et al., 2014) (Chanho Jung & Changick Kim, 2010).

There are some significant challenges, and their solutions are given here.

Challenge 1: CNN overfit because of the limited number of images available and affect the actual detection of breast cancer.

Solution: To increase the size of the dataset, the proposed model performs preprocessing, which contains two steps. First, it combines two datasets BreakHis and Breast Cancer Histopathological Annotation and Diagnosis (BreCaHAD) (Aksac et al., 2019). The second step is to perform augmentation with many parameters scaling, rotation, shifting, gamma value, flip and resize to increase the size of datasets to solve the overfitting issue.

Challenge 2: The structure of H&E images reveals volatile expression, even with the same malignancy level.

Solution: The strain normalization process resolve this issue; it takes augmented images to improve the contrast and brightness without affecting the image information with contrast limited adaptive histogram equalization method.

Challenge 3: Accurately detect the small objects in the histological image for the prognosis of breast cancer.

Solution: The proposed Multiscale Stochastic Dilated Convolution Model accurately enhances small and low-level features like edge, contour, and color. The proposed MSDC model combines spatial dilated convolution units: channel dilated convolution and stochastic dilated convolution units.

Challenge 4: Removal of redundant and similar features in the model, making the process complex.

Solution: Removing redundant or similar features using the ghost unit uses linear transformations to create feature maps. The ghost unit removes the redundant and similar features by applying a series of cheap linear operations on each intrinsic feature to generate ghost features.

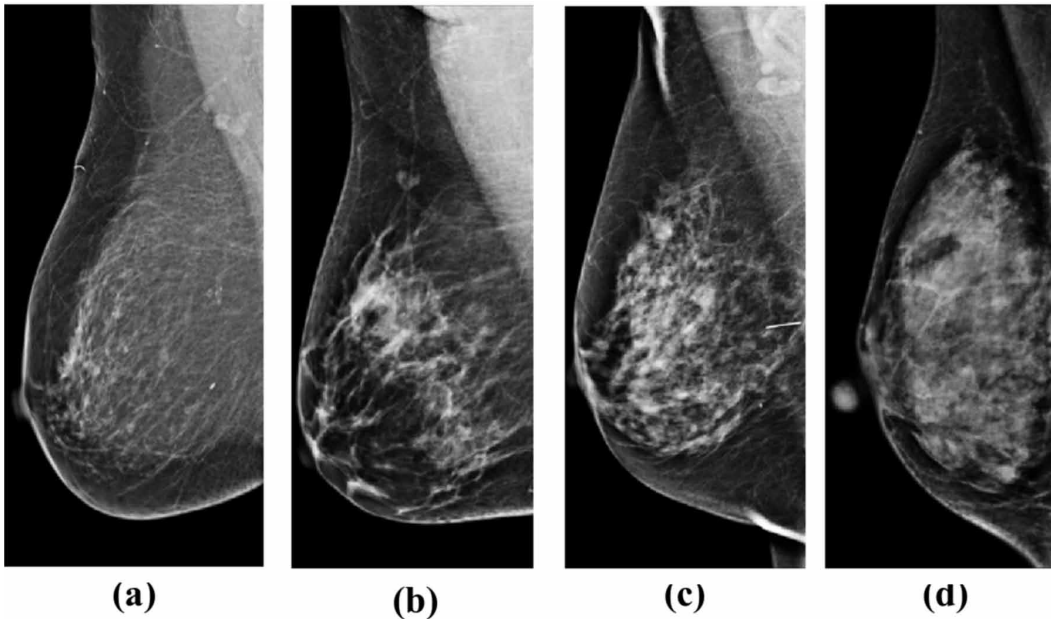
This paper proposes a deep learning model for histopathology images of breast cancer that takes images and detects cancer accurately with enhanced accuracy. The main contribution of our research is summarized below: We designed a deep learning-based model for the accurate detection of breast cancers on histopathological images. We are solving the overfitting and color divergence problem that affects breast cancer classification. The data expansion was implemented and delivered promising results. The proposed multiscale stochastic and dilation model detects small, low-level features such as edge, contour, and color accuracy without increasing the complexity. Eliminate redundancy and reduce model complexity using the Ghost unit. The model performance improves with the receptive field of the dilated convolution without losing image information. The Ghost unit for maintaining intrinsic feature maps and upsampling stochastic groupings reduces feature dimensions and improves performance. Some challenges are still open for future work as the comprehensive data set for the deep learning model is even smaller. Feeding the data into the model in parallel should improve performance in terms of accuracy. The proposed stochastic dilated residual ghost Network has 14 layers, which integrates the proposed multiscale stochastic dilated convolution model, stochastic downsampling, ghost unit, and stochastic upsampling operation very effectively to get the accurate result.

The following are the remaining sections of this paper. Background information is included in Section 2; the research materials and methods utilized to perform this study are given in Section 3. Sections 4 and 5 depict the experimental setup and outcomes. Finally, Section 6 summarises the study's findings and future directions.

2. BACKGROUND

Breast cancer treatment can be very effective with the involvement of technology and pathology in medical science. BC tumors can be detected by imaging techniques like mammography, positron emission tomography, biopsy, ultrasonography, and magnetic resonance imaging. A pathologist assesses these tissues and reports whether it is benign or malignant and may vary each time because of the complicated structure of H & E images. The hazardously grade 3 cancer cells, grade 1 to 4 of breast cancer depict in figure 2, and their prognosis can help a higher survival rate.

Figure 2. Mammographic images contain breast density from fatty (left) to highly dense (right) (“Breast Density,” 2021)



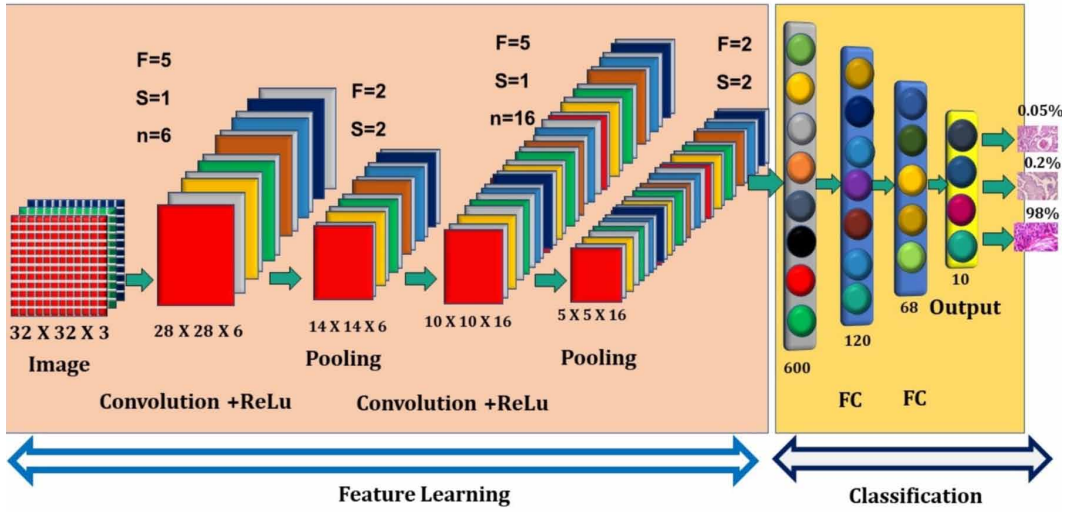
CNN adds value to deep learning techniques that have shown to be particularly effective in breast cancer detection and classification. A CNN model and its layered architecture recognize cancer and support medical science with the best outcomes (Choi et al., 2020). The convolution layer extracts an image’s local features, then the pooling layer is used for the dimensionality reduction, and the fully connected layer gives the intended output. The convolution’s excellent performance uses brighter pixels to show the edges in the image and highlights the local image features for further processing. The convolutional, pooling, fully connected, normalization layers are the network base; convolution applies a filter to an image and generates the image as an output. An image of (128x128x3) dimension in the CNN layer represents an image width of 128, the height of 128 pixels, and three channels, and the dimension that needs to process is 49,152 (128x128x3). When the image datasets are small, the traditional neural network may cause overfitting; when the image size is large, the conventional neural network may take up too much memory and require complicated calculations. A convolution operation plays a keen role in the CNN, as shown in equation 1:

$$C = \sum_1^i \sum_1^j I_{ij} \cdot F_{ij} \quad (1)$$

C is a feature map, I and F images and filters for the convolution operation, and i and j are rows and columns. The convolution operation is nothing more than the element-wise multiplication of the image by the kernel, a diagram of a traditional convolutional neural network is shown in Figure 3.

The left side of figure 3 shows feature learning; it takes an input image of 32x32, then passes the image to the convolution layer. The convolution layer applies a filter to generate a feature map that encapsulates the valuable features for the computationally simple model thanks to the rectified linear activation function, or ReLU operation, which covers negative to zero values. The pooling layer controls overfitting by reducing feature maps, training parameters, and computation costs. Figure 3 shows a convolutional neural network that can repeat one or more times. The right side of the figure

Figure 3. Convolution neural network architecture



shows classification; it performs by fully connected (FC) layers depending on the characteristics retrieved by the preceding layers.

Convolution transforms an image into a perceptron, which creates a local receptive field flattened on a feature map of size $(y) \times (z)$. Every layer has a set of x filters, each of which discovers features. The C_i^L of layer, L result to made up of x_1^L feature maps of size $(y)^L \times (z)^L$ with the i^{th} feature map denoted as C_i^L being computed as

$$C_i^L = B_i^L + \sum_{j=1}^{x^{(L-1)}} F_{i,j}^L * C_j^{(L-1)} \quad (2)$$

Where B_i^L is bias matrix and $F_{i,j}^L$ is the filter of size $(2h_1^L + 1) \times (2h_2^L + 1)$ connecting the layer's j^{th} feature map. Every channel kernel matrix in a convolution neural network may be different; for example, the blue channel may be minus one value, whereas the other channel may be positive or zero. The convolution procedure can employ several kernels; the dot product generates the convolution value using the kernel and channel matrices. The convolution procedure reduces the size of an image. For example, if (Is) is the image size and (Fs) is the kernel size, the resultant image size will be determined by

$$Cs = ((Is - Fs) / S) + 1 \quad (3)$$

Is , Fs , and S are 6, 3, and 1, respectively, in the picture, hence $Cs = (6 - 3 + 1) = 4$. The following formula computes the output dimension:

$$\left(\frac{Width - Fs + 2P}{S} \right) + 1 \quad (4)$$

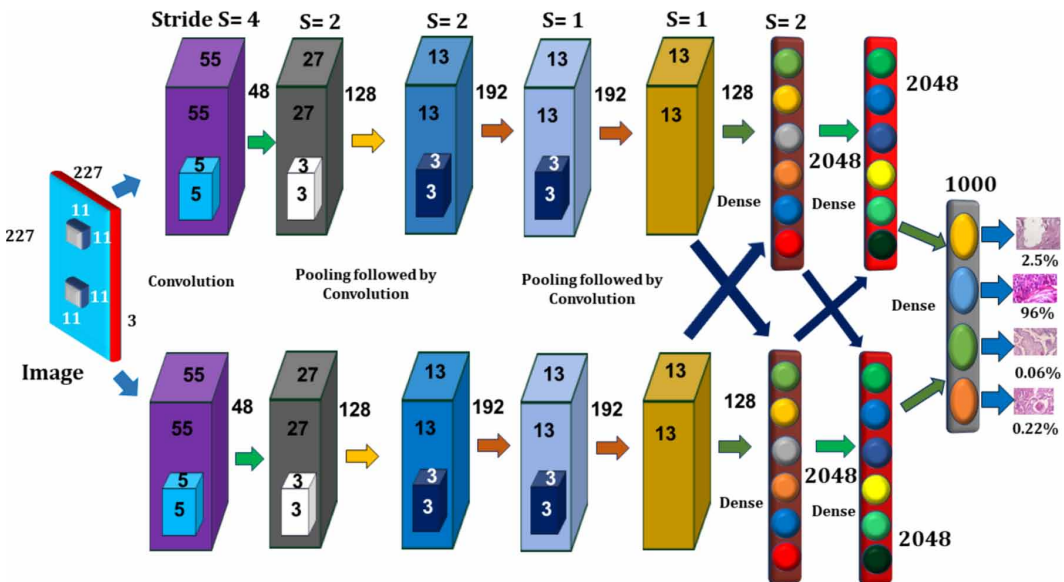
Fs stands for Filter Size, P for Padding, S for Stride, and equation 4 for the Floor Value.

As demonstrated in Figure 3, CNN models execute convolution on each picture before pooling. The FC layer converts to a 1×1 convolution layer, categorizing breast cancer using the activation function. Several CNN models have contributed to object categorization and detection, and a few are covered below.

2.1 AlexNet

AlexNet solves the gradient dispersion problem by using ReLU non-linearity. AlexNet also demonstrates that deep CNNs can be trained faster with saturating activation functions like Tanh or Sigmoid. Fully connected layers incorporate the dropout, which changes the architecture of the neurons to avoid overfitting, and max-pooling, which eliminates blur effects with feature enhancement. The AlexNet's data augmentation enhances the training datasets, which now total 62.3 million parameters. After padding, the input picture size is $227 \times 227 \times 3$ (Krizhevsky et al., 2017), and the architecture depicted in Figure 4 is used for parallel processing. Finally, the 4096 neurons and Softmax activation function produce 1000 values, which aid in improved categorization.

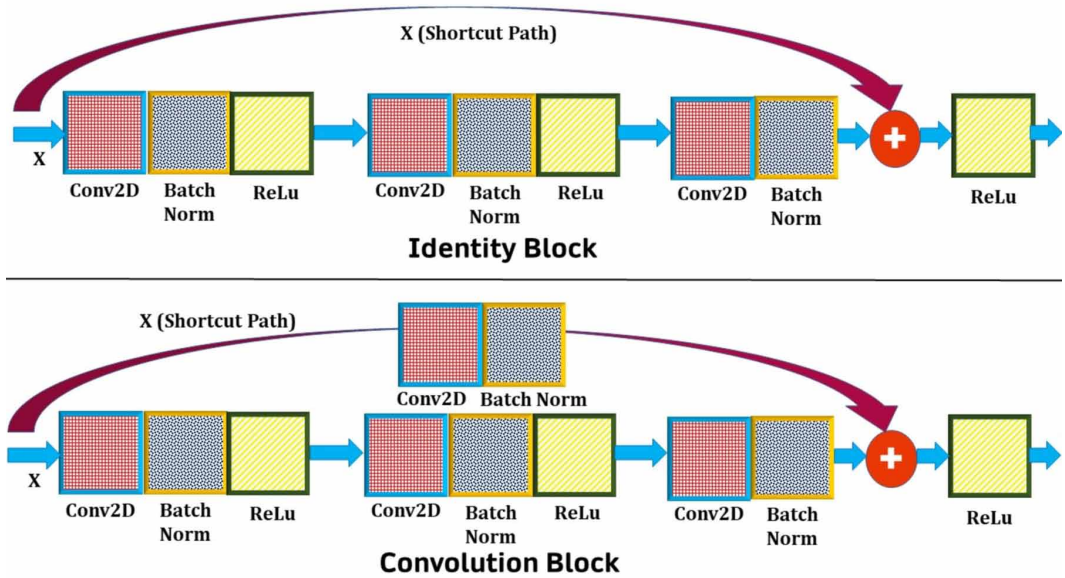
Figure 4. AlexNet Architecture



2.2 VGG16

The VGG16 network comprises 16 layers, is extremely simple to construct, and has a 7.7% error rate. This network utilizes a relu activation function without local response normalization (Simonyan, K., & Zisserman, A., 2015). The main difference between VGG and AlexNet is the kernel size; VGG uses 3×3 with stride 1, which provides better prediction, whereas AlexNet uses 11×11 with stride 4 (Shen et al., 2019). VGG16 model design consists of five convolution units followed by a max-pooling layer to reduce image size. It contains three ultimately linked layers: the first layer generates 4096 neurons and reduces the picture size while increasing the number of channels; the second and last layers each generate 100 neurons.

Figure 5. Residual learning a. identity block (Top) B. convolution block (Bottom)



2.3 ResNet

Residual Network (ResNet) bypasses the input information and retains the integrity of the data throughout the process. The network's depth and learning rate affect the deep layers by losing gradients; the learning rate problem is solved by adding a layer to the network that improves performance (Koné & Boulmane, 2018).

$$H(i) = f(W * i) + b \quad (5)$$

The equation without skipping the network is represented by equation 6, where f is the activation function; the output would be wherever the network takes the shortcut path.

$$H(i) = f(i) + I \quad (6)$$

The number of deep layers and the network employing these two identities and convolution blocks are the key differences between the ResNet varieties. Residual blocks considerably enhance performance (Zhou et al., 2020). The specifics of these ResNet variations are represented in Table 1. The 18 layers of ResNet must train using 11.17 million parameters and 1.8 billion flops. From the convolutional layers 2 through 5, each filter has 64-512 output channels. Similarly, the ResNet 50 and 152 employ 23.52 and 25.5 million trained neurons, respectively, and have 3.8 and 4 flops. The fundamental difference between ResNeXt-50 and ResNet 152 is that ResNeXt-50 uses 32 routes for cardinality, whereas ResNet152 does not. ResNeXt-50 employs 128-2048 output channels per filter, whereas ResNet-152 utilises 64-512 filter channels output.

The ResNeXt-101 has the lowest error rates between the top 1 and 5 of 20.4 percent and 5.3 percent when compared to ResNet-101, ResNet-200, ResNeXt-50, and ResNet-50. The primary distinction is between training and testing mistakes; it is less than RestNet and has less complexity in RestNeXt.

Table 1. Error percentage of resnet network on imagenet

S.N.	Model	Structure	Top-1 error	Top-5 error
1	ResNeXt-101	64x4d	20.4	5.3
2	ResNeXt-101	2x64d	20.7	5.5
3	ResNeXt-101	32x4d	21.2	5.6
4	ResNet-101	1x100d	21.3	5.7
5	ResNet-200	1x64d	21.7	5.8
6	ResNet-101	1X64d	22	6
7	ResNeXt-50	32x4d	24.4	6.6
8	ResNet-50	1X64d	27.1	8.2

3. RESEARCH MATERIALS AND METHODS

3.1 Datasets

The datasets for this investigation were Breast Cancer Histopathological Image Classification (BreakHis) and Breast Cancer Histopathological Annotation and Diagnosis (BreCaHAD). The BreakHis contains 2,480 benign and 5,429 malignant pictures from 82 patients with magnification factors of 700X460 pixels (40X, 100X, 200X, and 400X). While BreCaHAD contains 162 breast cancer images with resolution 1360×1024 pixels.

3.2 Proposed Model

The proposed model is divided into three parts. In the first part, we discuss data augmentation and strain normalization. In contrast, the second part discusses the MSDC model, which is instrumental in the proposed SDRG model for detecting small features. The third part shows the Proposed SDRG model for accurate detection of breast cancer. The complete model is shown in figure 12.

3.2.1 Strain Normalization

Color divergence is still a problem in histopathological pictures created with various slide scanners, biopsy materials, and staining procedures (Anghel et al., 2019) information after deformation normalization, as shown in Figure 6.

3.2.2 Data Augmentation

Data augmentation will enhance the number of images in a dataset, resulting in a more robust dataset for coaching and resolving overfitting issues caused by dataset size constraints (Zhu et al., 2019). The suggested technique combines two datasets, BreakHis and BreCaHAD, then performs operations such as flipping, rotating, shifting, resizing, and gamma worth table 2 lists the parameters evaluated for the augmentation of various parameters and values. Figure 7 shows the outcomes after image augmentation first row contains a histological image from the dataset and strain normalized image followed by a rotated image. The second row shows flipped image, an image after applying gamma values 0.3 and 1.2.

Image augmentation performs flipping, scaling, rotation, and gamma-correction. The scaled images were created by applying scaling factors of 0.5, 0.8, and 1.2 to each image. Using the horizontal and vertical reflections generates flipped images and the angles 40, 80, 120, and 180 degrees to rotate the image (Shahidi et al., 2020). The gamma correction used in the augmentation of the pictures shown in Figure 7 (a-f) ranges from 0.3 to 1.2. After adding 19 parameters for magnification factors (40x,

Figure 6. Effects of Strain Normalization A: An Original Image B: An Image after Strain Operation

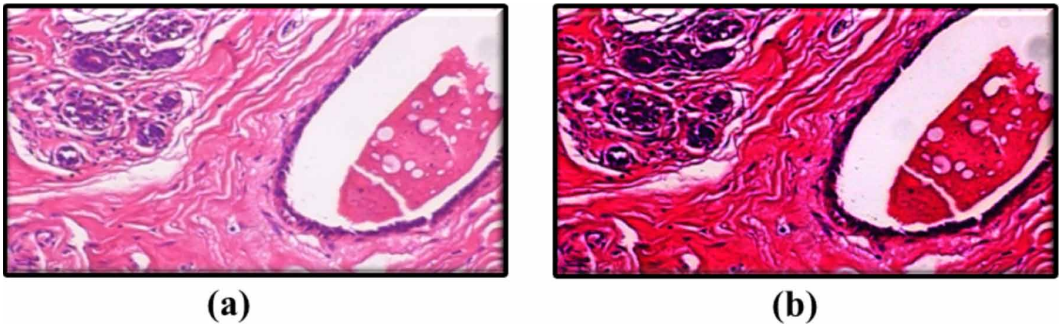
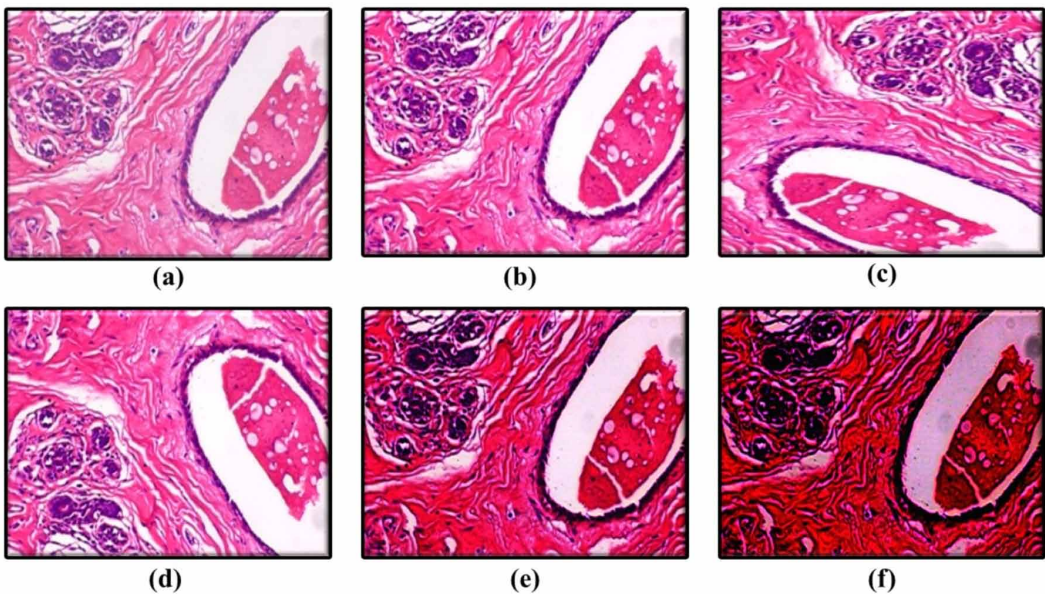


Table 2. Parameters of data augmentation

S.N.	Parameters of Image Augmentation	Number of Parameters	Value Used
1	Image Scale	3	0.5, 0.8, 1.2
2	Image Rotation	4	40,80,120,180
3	Width Shift	1	0.3
4	Height Shift	1	0.3
5	Resize	4	224, 112,64,32
6	Flip	2	Horizontal, Vertical
7	Gamma Value	4	0.3,0.6,0.9 and 1.2
Total		19	19

Figure 7. Data Augmentation Results: a) Histological Image, b) Image after Strain Normalization, c) Rotated image, d) Flipped image, e) Gamma Value 0.3, f) Gamma Value 1.2



100x, 200x, and 400x) to 7909 samples, the total number of samples is 1,53,349 contains 1,50,271 BreakHis images and 3,078 BreCaHAD images.

3.2.3 Proposed multiscale stochastic dilated convolution model

The first proposed model is the Multiscale Stochastic Dilated Convolution Model, constructed and integrated into three units to determine the outcome. The units are proposed spatial dilated convolution unit, proposed channel dilated convolution unit, and stochastic dilated convolution unit, which increases tiny and low-level characteristics such edge, contour, and color correctly (Wang et al., 2017). The feature extraction methods are affected by the variable intensity levels of the pixels in the histology images. This model considers the strain normalized images and enhanced images to detect tiny cells better. Figure 8 shows the feature extraction model in action. After convolution, the unit accepts an image as input and sends the convoluted pictures to max-pooling and stochastic pooling in parallel. Max pooling detects microscopic cancer cells and thin boundary detection by obtaining the most present sections of the cancer cells. Stochastic pooling selects a value at random from each receptive field.

The probability P_n for the value X_n can be identified at position n in the receptive field using equation 7.

$$P_n = \frac{x_n}{\sum_{k \in z} x_z} \quad (7)$$

and the stochastic pooling represented as:

$$P_1 = x_m \text{ where } m \sim P (p_1, \dots, p_m) \quad (8)$$

Stochastic pooling may choose any value and classify it as a heterogeneous approach since it can select tiny and large values. The dilation unit discovers the cells without increasing the complexity by concatenating the max pooling, and stochastic pooling results in a good feature map subsequently processed by the dilation unit (Perone et al., 2018). Dilation procedures increase the depth of the network while increasing the size of the filters $(2d+1) \times (2d+1)$ may be used to calculate the depth, with values ranging from 1, 2, 3, 4, and so on. 3×3 , 5×5 , 7×7 , and 9×9 are the receptive fields, accordingly. The dilated convolution expresses in the following way:

$$D(x,y) = \sum_{i=1}^X \sum_{j=1}^Y I(x + S_{xi}, y + S_{yj}) F_{i,j} \quad (9)$$

The proposed spatial dilated convolution unit uses different filters for each channel, followed by element-wise convolution, as illustrated in figure 8. The dilated convolution is $D(x,y)$, the filter is $F(i,j)$, the input image is $I(x,y)$, and the sparse kernel is s . When $s=1$, the outcome of dilation and standard convolution is the same. Dilated features may be computed easily (Lin & Wu, 2020).

$$D(W,y) = \sum_{i=1}^X \sum_{j=1}^Y W(i, j) \odot y(i+x, j+y) \quad (10)$$

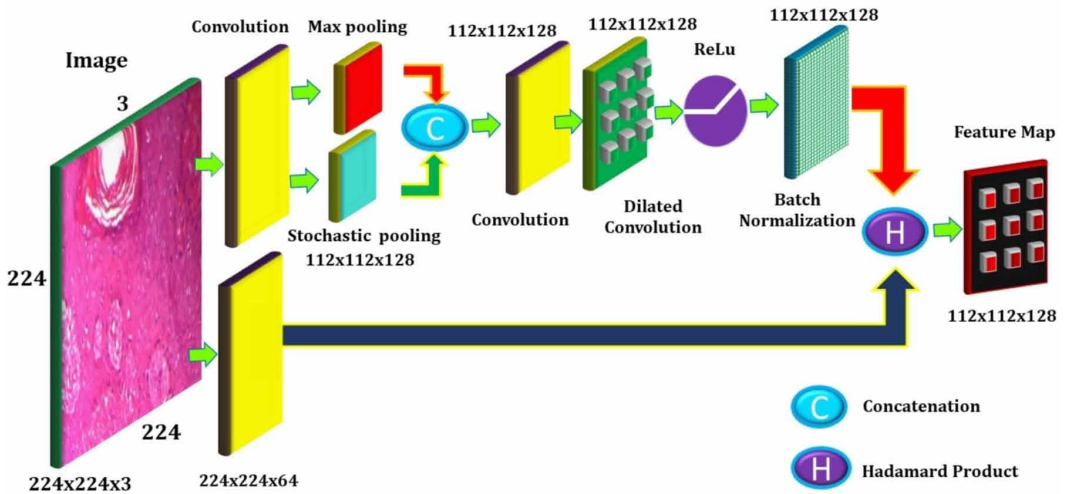
$$D(W,y) = \sum_{m=1}^M W(m) \odot y(i, j, m) \quad (11)$$

Here dilated convolution operation uses depth-wise convolution with dilation rates ranging from 1 to 8. After the relu operation (Wang et al., 2020), the extracted features are processed for batch normalization, eliminating the gradient vanishing problem and speeding up the network. The Hadamard product combines the convoluted image features with the outcome of the created feature map to provide spatially dilated features. Dilated spatial convolution processes all channels equally and displays low-level elements such as edges, contours, and colors. When combined with the dilated characteristics, a feature map contains essential cell information that aids in the correct detection of cancer cells (Han et al., 2017). The suggested spatial dilation attention unit broadens the receptive field and feature map.

$$C_i^L = B_i^L + \sum_{j=1}^{x^{(L-1)}} F_{i,j}^L * C_j^{(L-1)} \quad (12)$$

$$\{H\sigma D[C[P_1(C(I)) \parallel P_2(C(I))]]\} \quad (13)$$

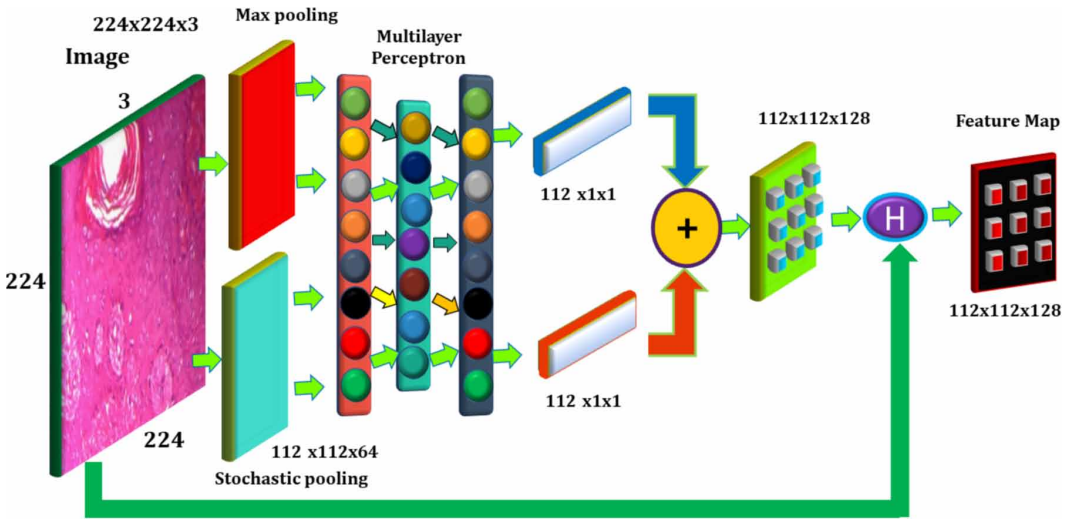
Figure 8. The proposed spatial dilated convolution unit



The input image is I , while the convolution, stochastic pooling, and max pooling procedures are C , P_1 , and P_2 , respectively. The suggested stochastic dilated convolution unit takes an input picture, conducts convolution, and then creates maximum pooling and stochastic pooling concatenated outputs for convolution and dilation convolution. The overfitting difficulties in the network are handled by overlapping the pool layers, which also improves accuracy.

The dilated channel attention unit was initially employed to give channel dimension information. It uses max pooling and stochastic pooling, followed by a fully connected layer to assist the classification process represented in Figure 9. The 1×1 convolution increases the depth of the components; these features concatenate, and dilation convolution occurs, accentuating the narrow cell borders (Zhang

Figure 9. The proposed channel dilated convolution unit



et al., 2020). The feature map is generated by these two descriptors using dilated convolution and the Hadamard product.

$$\{H[D(FC(P_1(I)))\parallel FC(P_2(I))]\} \quad (14)$$

After completely connected layers, the Hadamard product is applied to the dilated result of max pooling and stochastic pooling.

$$D(P_1(C(I))) \quad (15)$$

An image is convoluted to create the feature map displayed in Figure 10, then stochastically pooled and dilated. It assists in the storing of data with helps in accurate cancer cell detection.

The proposed model, shown in figure 11, consists of a spatial dilated convolution unit, channel dilated convolution unit, and stochastic dilated convolution unit.

The model's output is produced by concatenating equations 13, 14, and 15, which properly delivers the missing information for the pixels.

$$F_{out} = \{H\sigma D[C[P_1(C(I))\parallel P_2(C(I))]]\parallel \{H[D(FC(P_1(I)))\parallel FC(P_2(I))]\parallel \{D(P_1(C(I)))\} \} \quad (16)$$

D and H are Dilated convolution and Hadamard product, respectively, whereas relu and concatenation are represented by and \parallel . To overcome the fundamental depth problem that impacts the network's efficiency and accuracy, the proposed MSDC model identifies missing features and small features illustrated in Figure 11.

3.2.4 Proposed stochastic dilated residual ghost model

The second model is the Proposed Stochastic Dilated Residual Ghost model, which incorporates the proposed MSDC model, proposed stochastic downsampling unit, and the upsampling unit. The ghost

Figure 10. The stochastic dilated convolution unit

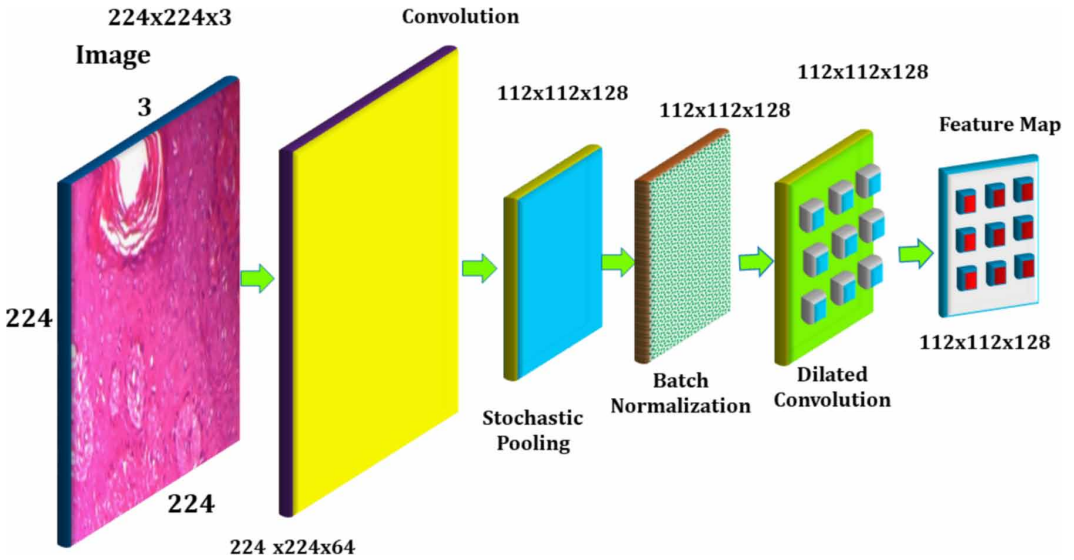
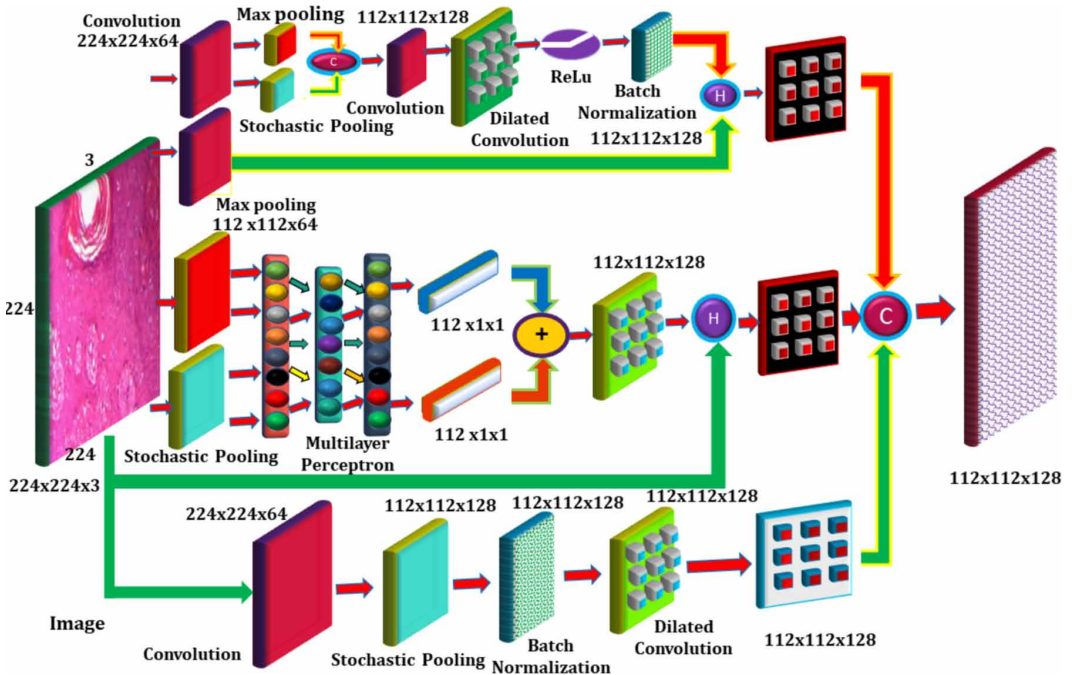
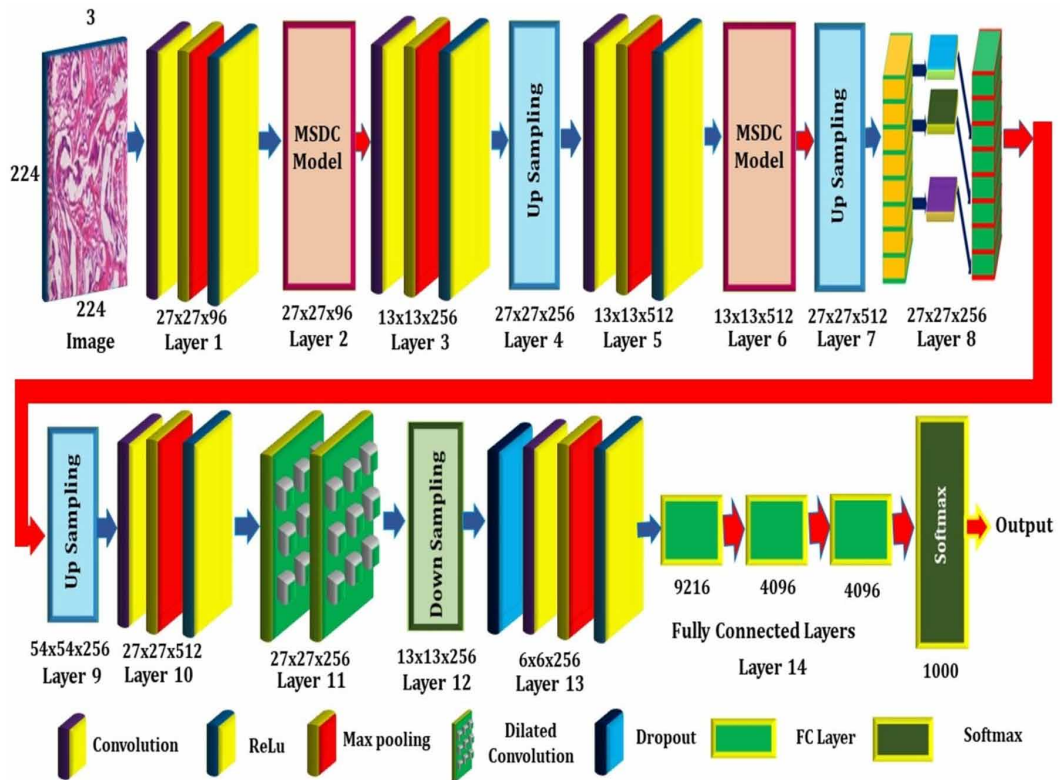


Figure 11. The proposed multiscale stochastic dilated convolution model



unit effectively and extracts features accurately. It has 14 deep layers; the first layer takes an input image with $224 \times 224 \times 3$ pixels. It is convoluted using 3×3 filters, resulting in a dimension of $224 \times 224 \times 64$. The convolutional kernel divides a picture into small slices called receptive fields. When an image is broken into tiny parts and then pooled using a filter size of 2×2 and a stride of 2, extracting feature motifs becomes easier. As a result, the image size will be $112 \times 112 \times 128$. Convolutional networks use

Figure 12. The proposed stochastic dilated residual ghost (SDRG) model



pooling layers to reduce the model’s size, improve the calculation speed, and improve the robustness of the extracted features.

The first layer is the convolutional layer, which has 96 distinct 11x11-size receptive filters to conduct convolution and max-pooling using Local Response Normalization (LRN) or relu. Three filters with a stride size of 2 are used to achieve the maximum pooling operations. The second layer contains the proposed MSDC model, takes 27x27x96, extracts missing and additional features from the image. In the third layer, identical actions of layer one are carried out with 5x5 filters. The fourth, seventh, and ninth layers are upsampling layers that upscaled the feature maps up to 27x27x256, 27x27x512, and 54x54x256. These layers generate random values and multiply with the transposed features received from the previous layer.

$$Xu1 = \sigma(W1^T X) \tag{17}$$

The convoluted features cascaded with the convoluted features seen in equations 18-22 after the output function convoluted.

$$Xu2 = \sigma(W1Xu1) \tag{18}$$

$$Xu3 = [\sigma(W2Xu2) + Xu2] \tag{19}$$

$$Xu4 = [\sigma(W3Xu3) + Xu3] \tag{20}$$

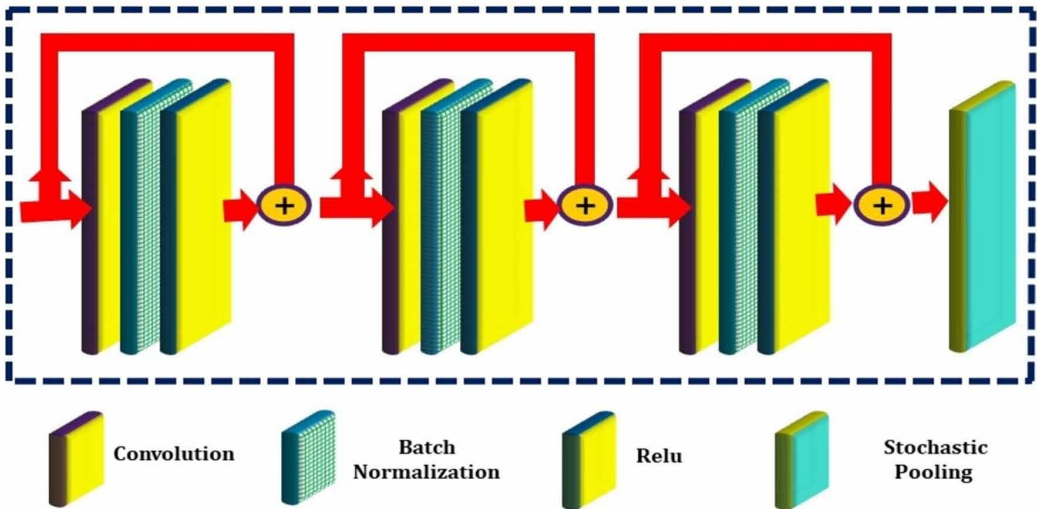
$$Xu5 = [\sigma(W4Xu4) + Xu4 + Xu2] \tag{21}$$

$$Xu6 = \sigma(W5Xu5) \tag{22}$$

Here, X is the random value generated by the unit ($W1^T$) is the transposed of the features received from the previous layer, $Xu1$ to $Xu6$ are the upscaled features, and σ is the element-wise multiplication operation.

The proposed stochastic upsampling unit employs a skip layer to improve classification accuracy, as shown in figure 13.

Figure 13. The proposed stochastic upsampling unit



Filters 3x3 with 512, 512, and 256 feature maps are utilized in the fifth, tenth, and thirteenth convolutional layers. Layer 6 is the proposed MSDC model again, which takes a feature map of 13x13x512 and extracts missing and additional features from the image. Pointwise convolutions (Ma et al., 2018) (Zhang et al., 2018) (Sandler et al., 2018) are used to analyze features across channels before utilizing depth-wise convolution to handle spatial data in current techniques. As mentioned in layer 8, the ghost unit (Han et al., 2020) employs ordinary convolution to produce a few intrinsic feature maps before supplementing the features and enlarging the channels using low-cost linear operations. Feature maps produced by convolutional layers commonly contain. There is a great deal of redundancy, and some of the items may be like one another. It's worth noting that with a large number of duplicate feature maps, it's not necessary to construct one by one by FLOPS and parameters. Let us pretend the output feature is Maps are "replicas" of a few key feature maps. The ghost unit uses a series of low-cost linear operations on each intrinsic feature to generate s ghost features that follow the formula:

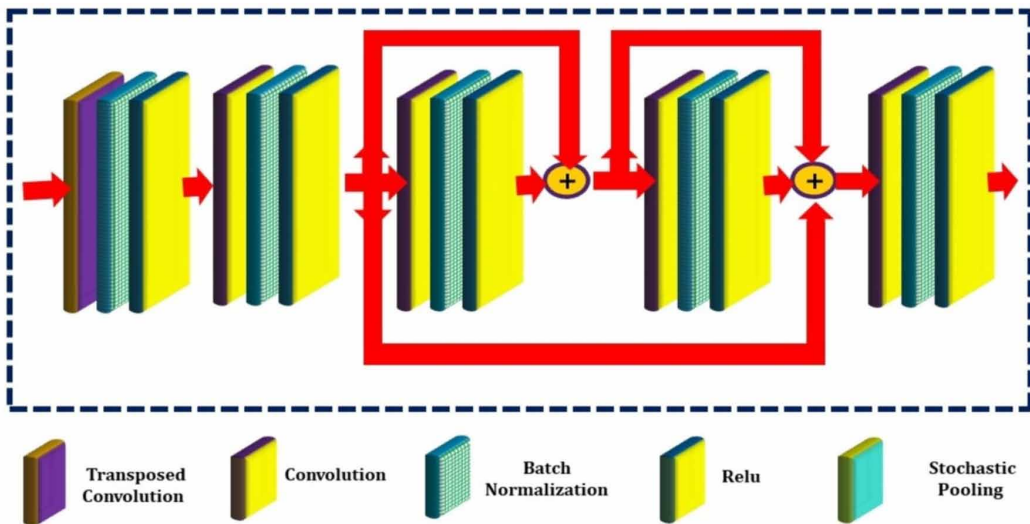
$$X_{(i,j)} = \phi_{\{i,j\}}(X_{(i)})' \tag{23}$$

Where $(X_{(i)})$ is the intrinsic function map and $\phi_{\{i,j\}}$ are the linear operations.

Unlike the units (Zhang et al., 2018), which make considerable use of 11 pointwise convolutions, the ghost unit's main convolution can have a variable kernel size. Processing each feature map in previously efficient systems was limited to depth-wise convolution or shift operations, while linear functions in the ghost module can offer many possibilities. The identity mapping is paralleled with linear transformations in the ghost unit to preserve the inherent feature mappings.

Layer 11 performs cascaded dilation convolution to extract features and reduce the complexity of the network. The features in the 1x one convolution converted at a faster rate, and the elements concatenated to provide better features for the minor cell observation. Here it generates a feature map of size 27x27x256. Layer 2 is a stochastic down-sampling unit shown in figure 14 used to boost the performance of the network.

Figure 14. The proposed stochastic down sampling unit



$$Xd1 = \sigma(W1X) \tag{24}$$

As a cascade operation, the output of the first layer is connected to the input of the stochastic downsampling block, resulting in the output Xd2.

$$Xd2 = \sigma(W2(Xd1)+X) \tag{25}$$

Similarly, the outputs Xd1 and Xd2 and the input of the down sampling block are connected to create a feature map, which is then reduced in dimensionality using a 1x1 convolution to boost performance.

$$Xd3= W3(Xd1+Xd2+X) \quad (26)$$

A dropout layer is employed, in which a few neurons are removed from the neural network during the training process, resulting in a smaller model. Layer 13 is a dropout unit with the convolutional operation followed by a pooling and relu activation function on several channels or feature maps to reduce dimensions and extract features. Layer 14 is a series of three FC layers followed by a Softmax layer. The weights, biases, and neurons make up the FC layer, linking the neurons between two layers. Here FC layers are converting 9216 into 1000 neurons with softmax activation function, which generates the classification factors for the better classification of benign and malignant cancer cells.

4. EXPERIMENTAL RESULTS

In this section, we evaluated the performances of the proposed deep learning model. We began by emphasizing the performance measures of the proposed model with other models.

4.1 Experimental Setup

The model was implemented with Python version 3.9, TensorFlow 2, and Keras version 2.4. The Xeon 2.3 GHz CPU configuration, 12 GB memory, supports two Nvidia GeForce GTX 2070 graphics processors (GPUs) and 16 MB cache memory.

4.2 Results of SDRG Model

The proposed model is trained and evaluated using image data from the BreakHis dataset, containing 7909 actual samples (pictures) from 82 individuals, separated into two subgroups of 2,480 benign and 5,429 malignant samples and 162 breast cancer images from BreCaHAD. Total 7909 BreakHis samples and 3,078 BreCaHAD images were augmented with 19 parameters for magnification factors (40x, 100x, 200x, and 400x), for a total of 1,53,349. In the next sections, the findings of binary and multiclass classification will be explained, and the findings will be compared to those of earlier research. The proposed multiscale dilated convolution model detects small objects accurately. Conventional models relied on max pooling to improve the receptive field, followed by a convolution operation. It is replaced by dilated convolution in the proposed MSDC model. The proposed Stochastic Dilated Residual Ghost (SDRG) Model classifies benign and malignant cells accurately. The proposed model employs the proposed MSDC model, ghost unit, stochastic downsampling, stochastic up-sampling units, and other convolution layers. Table 3 displays the outcomes of the proposed model's evaluation.

With the augmented images seen in table 3, increased specificity, accuracy, precision-recall, and F1-score values have observed, resulting in better performance. BreCaHAD and BrekHis amplified images with 100 magnification factors; the SDRG model has an accuracy of 98.60 percent and 98.41 percent, respectively. The proposed model with the enhanced images has a better receiver operating characteristic shown in figure 15 than the previous model for fold five is 0.9432.

True positive and true negative values can detect malignant and healthy cells, respectively. The tests' preparation and research data ratios were 90-10, 80-20, and 70-30. Precision, recall, f1 score, accuracy, and average precision score used to compile the results, as seen in Table 4. The proposed models' efficiency compared to VGG16, VGG19, ResNet50, and ResNeXt-101 models.

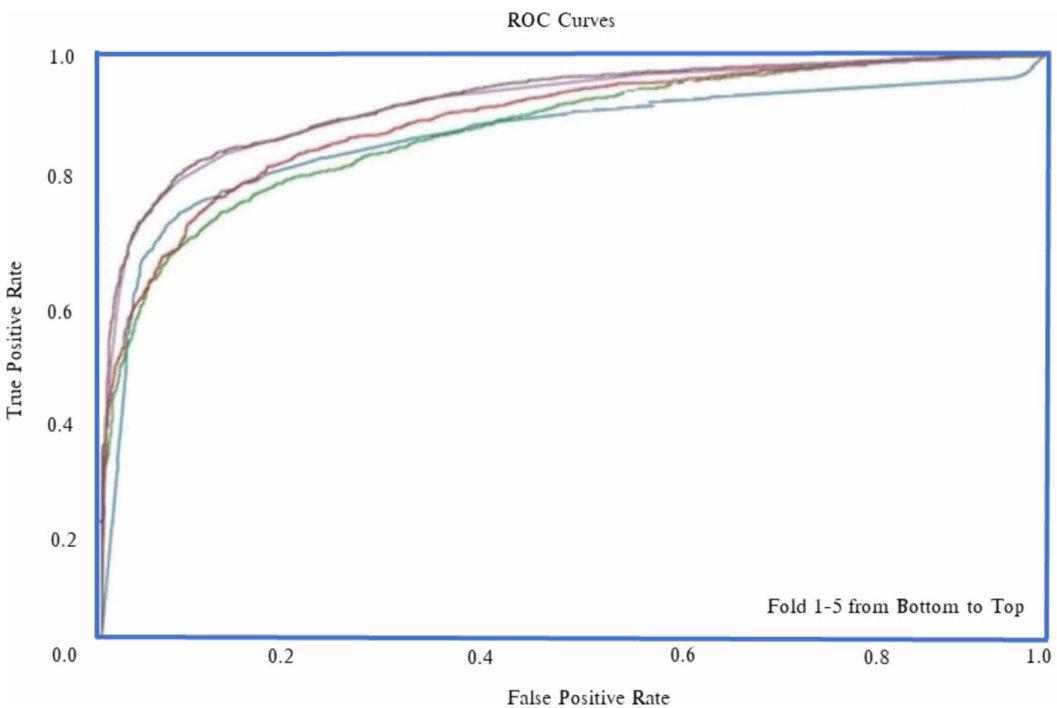
On the given datasets, the VGG16 model performed better than VGG19, and ResNeXt101 outperformed ResNet50, as seen in table 4. The proposed model outperforms standard models in terms of accuracy, sensitivity, and precision.

The top-5 testing error rate for the VGG19 and VGG16 is 8.9% and 9.2%, respectively. The top-5 validation errors for the ResNeXt-101 and ResNet-50 are 6.4 and 7.8 percent, respectively, while the proposed SDRG model achieves the lowest top-5 validation error of 5.1 percent, as seen in table 5.

Table 3. Evolution metrics of the proposed model

Images	Image Type	Specificity	Accuracy	Precision	Recall	F1-score
Magnification Factor 40 X	Original	98.50 ± 0.14	93.13 ± 4.36	99.07 ± 0.04	92.59 ± 0.04	97.49 ± 0.04
	Augmented	96.88 ± 0.07	98.41 ± 1.00	98.17 ± 0.05	94.19 ± 0.04	97.94 ± 0.04
Magnification Factor 100 X	Original	91.90 ± 0.09	93.99 ± 5.13	95.60 ± 0.05	93.18 ± 0.04	96.20 ± 0.05
	Augmented	94.42 ± 0.10	98.37 ± 1.52	95.88 ± 0.07	97.74 ± 0.05	96.58 ± 0.05
Magnification Factor 200 X	Original	86.13 ± 0.12	91.97 ± 8.01	92.26 ± 0.14	98.90 ± 0.05	94.57 ± 0.10
	Augmented	96.73 ± 0.07	96.26 ± 2.30	98.24 ± 0.03	99.71 ± 0.03	99.45 ± 0.04
Magnification Factor 400X	Original	99.87 ± 0.12	92.94 ± 5.15	92.60 ± 0.08	96.55 ± 0.04	95.38 ± 0.05
	Augmented	92.63 ± 0.09	97.38 ± 2.18	94.84 ± 0.06	96.94 ± 0.05	96.75 ± 0.04
BreCaHAD	Original	97.40 ± 0.07	95.23 ± 4.38	98.07 ± 0.04	97.39 ± 0.06	98.49 ± 0.04
	Augmented	96.98 ± 0.10	98.60 ± 0.99	99.17 ± 0.05	98.99 ± 0.05	99.94 ± 0.04

Figure 15. Receiver operating characteristic curve of the proposed model on the breakhis dataset with the magnification factor 40X fold (1-5)



5. DISCUSSION

We will discuss the performance of our proposed SDRG model in this section. As previously stated, the proposed model incorporated the proposed MSDC model, a ghost unit, stochastic upsampling, and downsampling units. The proposed model addresses the research challenge of creating an accurate model that can detect missing and tiny features, color divergence, and overcome network difficulties. The image was brightened and contrasted using contrast-limited adaptive histogram equalization.

Table 4. Performance analysis of model VGG16, VGG19, ResNet50, ResNeXt101, and the Proposed SDRG model

Model	Training and Testing Data Ratio	Precision	Sensitivity	F1 Score	Accuracy	APS
VGG16	90-10%	0.77	0.78	0.78	74.50	88.80
	80-20%	0.80	0.78	0.79	75.30	88.59
	70-30%	0.78	0.78	0.78	73.14	88.34
VGG19	90-10%	0.64	0.81	0.78	64.91	73.99
	80-20%	0.79	0.74	0.73	72.21	76.11
	70-30%	0.80	0.64	0.49	61.68	69.55
ResNet50	90-10%	0.93	0.93	0.92	89.21	92.45
	80-20%	0.94	0.92	0.92	89.31	92.45
	70-30%	0.94	0.94	0.94	90.44	93.87
ResNeXt-101	90-10%	0.95	0.95	0.95	90.42	92.54
	80-20%	0.97	0.94	0.94	91.32	92.54
	70-30%	0.98	0.96	0.95	91.53	93.65
Proposed SDRG	90-10%	0.98	0.97	0.93	95.65	96.56
	80-20%	0.99	0.98	0.94	95.65	97.72
	70-30%	0.98	0.98	0.93	94.63	94.94

Table 5. Error rate comparison of the proposed method with traditional method

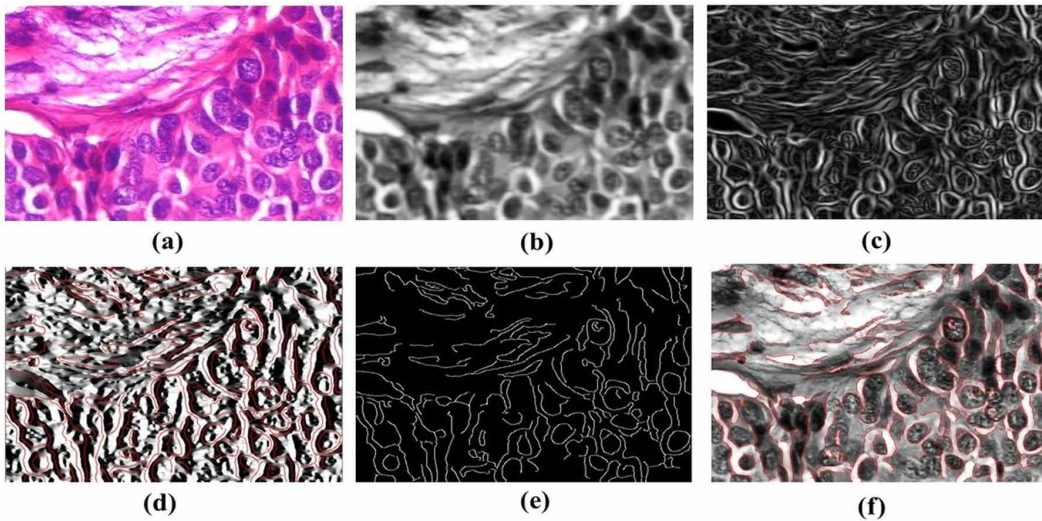
S.N.	Method	Top-1 Val Error Rate (%)	Top-5 Val Error Rate (%)	Top-5 Test Error Rate (%)
1	VGG16	25.6	9.3	9.2
2	VGG19	25.4	9.1	8.9
3	ResNet-50	24.1	7.8	7.7
4	ResNeXt-101	21.8	6.4	6.3
5	Proposed SDRG	17.7	5.1	3.9

Overfitting posed challenges for deep learning models, which were overcome in the experiment by integrating the two datasets (BreakHis) and (BreCaHAD) and data augmentation. The proposed spatial dilated convolution unit, proposed channel dilated convolution unit, and stochastic dilated convolution unit make up the multiscale stochastic dilated convolution model, which increases tiny and low-level properties such as tip, contour, and color precision. In this context, the BreakHis and BreCaHAD histological breast cancer datasets were subjected to a thorough experimental study. We compared our proposed approach to some past research, and the findings are shown in Table 4. We can observe that our proposed model can obtain the best accuracy, precision, sensitivity, and average precision. Furthermore, we can observe from Table 4 that prior studies were not accurate in breast cancer diagnosis. Figure 15 shows the ROC curves in the experimental study.

The proposed stochastic dilated residual ghost (SDRG) model comprises 14 layers and successfully incorporates the suggested multiscale stochastic dilated convolution model, stochastic downsampling, ghost unit, and stochastic upsampling operation to get an accurate result. The receptive field of the

dilated convolution enhances model performance without sacrificing picture information. The ghost unit decreases feature dimensions and increases performance by retaining intrinsic feature mappings and stochastic upsampling groups. Because the complete data set for the deep learning model is significantly smaller, several obstacles remain for future research. The accuracy of the model should increase if the data is fed into it simultaneously. The proposed SDRG model detects tiny abnormalities without increasing the system's complexity. Figure 16 depicts segmented histopathology depictions.

Figure 16. Object Detection (a) Histopathological Image (b) Smoothing (c) Gradient in x direction (d) Gradient in y direction (e) Non-Max Suppression (f) Segmented Image



We can derive from the empirical findings based on breast cancer datasets that our proposed methodology is the most suited method for breast cancer detection. It can deliver fantastic results while requiring just a little computing cost to solve the breast cancer classification problem. Our proposed methodology has the potential to be applied to the diagnosis of various illnesses.

6. CONCLUSION

This paper proposed stochastic dilated residual ghost model, multiscale stochastic dilated convolution model, stochastic upsampling, and stochastic downsampling units. The experiments were conducted using BreakHis and BreCaHAD histopathological breast cancer images. Strain normalization was done to improve the image quality and solve the color divergence problem in H & E images. We considered different criteria such as magnification factor, resized sample inputs, rotated, gamma correction, and flipped parameters were applied to get augmented images. The proposed multiscale stochastic and dilation convolution model detects small, low-level features such as edge, contour, and color accuracy without increasing the complexity. The model performance improves with the receptive field of the dilated convolution without losing image information. The ghost unit for maintaining intrinsic feature maps eliminates redundancy. The proposed stochastic dilated residual ghost model integrates the proposed multiscale stochastic dilated convolution model, stochastic downsampling, ghost unit, and stochastic upsampling units to get accurate results. The proposed method improves average recognition accuracy by 4.33 percent on the BreakHis and BreCaHAD datasets using the current and enhanced

residual deep network ResNeXt101. The proposed approach outperforms multiple state-of-the-art method's accuracies by 95.65 percent. In both datasets, our model beat existing state-of-the-art models. The suggested model has high sensitivity and specificity for various cases, which is valuable for pathologists and researchers working in cancer detection utilizing histological images. We want to test the model's performance on different datasets with various cancer cases in the future. The suggested model has high sensitivity and specificity for different situations, which is valuable for pathologists and researchers working in cancer detection utilizing histological images. We will also investigate which feature map combinations are most likely to be essential for categorization. Researchers can investigate which assortments of feature maps are likely to be most important for classification. They can test the model's performance on different datasets with various cancer cases in the future.

REFERENCES

- Afifi, A., Saad, A., Al-Husseini, M., Elmeharth, A., Northfelt, D., & Sonbol, M. (2019). Causes of death after breast cancer diagnosis: A US population-based analysis. *Cancer*, *126*(7), 1559–1567. doi:10.1002/cncr.32648 PMID:31840240
- Aksac, A., Demetrick, D., Ozyer, T., & Alhaji, R. (2019). BreCaHAD: A dataset for breast cancer histopathological annotation and diagnosis. *BMC Research Notes*, *12*(1), 82. Advance online publication. doi:10.1186/s13104-019-4121-7 PMID:30755250
- Anghel, A., Stanisavljevic, M., Andani, S., Papandreou, N., Rüschoff, J., Wild, P., Gabrani, M., & Pozidis, H. (2019). A High-Performance System for Robust Stain Normalization of Whole-Slide Images in Histopathology. *Frontiers In Medicine*, *6*, 193. Advance online publication. doi:10.3389/fmed.2019.00193 PMID:31632974
- Bayramoglu, N., Kannala, J., & Heikkilä, J. (2016). Deep learning for magnification independent breast cancer histopathology image classification. *23rd International Conference on Pattern Recognition (ICPR)*, 2440-2445. doi:10.1109/ICPR.2016.7900002
- Breastlink. (2021). *Breast Density*. Retrieved 26 May 2021, from <https://www.breastlink.com/breast-cancer-101/risk-factors/breast-density/>
- Chanho, J., & Changick, K. (2010). Segmenting Clustered Nuclei Using H-minima Transform-Based Marker Extraction and Contour Parameterization. *IEEE Transactions on Biomedical Engineering*, *57*(10), 2600–2604. doi:10.1109/TBME.2010.2060336 PMID:20656653
- Chen, H., Dou, Q., Wang, X., Qin, J., & Heng, P. (2016). Mitosis detection in breast cancer histology images via deep cascaded networks. *Proc. 30th AAI Conf. Artif. Intell.*, 1160–1166.
- Choi, J., Kim, H., Kim, W., Lim, I., Lee, I., Byun, B., Noh, W. C., Seong, M.-K., Lee, S.-S., Kim, B. I., Choi, C. W., Lim, S. M., & Woo, S.-K. (2020). Early prediction of neoadjuvant chemotherapy response for advanced breast cancer using PET/MRI image deep learning. *Scientific Reports*, *10*(1), 21149. Advance online publication. doi:10.1038/s41598-020-77875-5 PMID:33273490
- Coudray N., Ocampo P., Sakellaropoulos T., Narula N., Snuderl M., Fenyö D., Moreira A., Razavian N., & Tsirigos A. (2018). Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nature Med.*, *24*(10), 1559-1567. .10.1038/s41591-018-0177-5
- Dapson, R., Horobin, R., & Kiernan, J. (2010). Hematoxylin shortages: Their causes and duration, and other dyes that can replace hemalum in routine Hematoxylin and eosin staining. *Biotechnic & Histochemistry*, *85*(1), 55–63. doi:10.3109/10520290903048400 PMID:19562570
- Fuyong, X., Hai, S., Neltner, J., & Yang, L. (2014). Automatic Ki-67 Counting Using Robust Cell Detection and Online Dictionary Learning. *IEEE Transactions on Biomedical Engineering*, *61*(3), 859–870. doi:10.1109/TBME.2013.2291703 PMID:24557687
- Gavrilov, A., Jordache, A., Vasdani, M., & Deng, J. (2018). Preventing Model Overfitting and Underfitting in Convolutional Neural Networks. *International Journal of Software Science and Computational Intelligence*, *10*(4), 19–28. doi:10.4018/IJSSCI.2018100102
- Giannakeas, N., Tsiplakidou, M., Tsiouras, M., Manousou, P., Forlano, R., & Tzallas, A. (2017). Image Enhancement of Routine Biopsies: A Case for Liver Tissue Detection. *IEEE 17th International Conference on Bioinformatics and Bioengineering (BIBE), 2017*, 236-240. doi:10.1109/BIBE.2017.00-49
- Gultom, Y., Arymurthy, A., & Masikome, R. (n.d.). Batik classification using deep convolutional network transfer learning. *J. Ilmu Komputer Inf.*, *11*(2), 59. .10.21609/jiki.v11i2.507
- Han, K., Wang, Y., Tian, Q., Guo, J., Xu, C., & Xu, C. (2020). GhostNet: More Features from Cheap Operations. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 1580-1589). IEEE.
- Han, Z., Wei, B., Zheng, Y., Yin, Y., Li, K., & Li, S. (2017). Breast Cancer Multi-classification from Histopathological Images with Structured Deep Learning Model. *Scientific Reports*, *7*(1), 4172. Advance online publication. doi:10.1038/s41598-017-04075-z PMID:28646155

- Huang, Y., Wu, R., Sun, Y., Wang, W., & Ding, X. (2015). Vehicle Logo Recognition System Based on Convolutional Neural Networks with a Pretraining Strategy. *IEEE Transactions on Intelligent Transportation Systems*, 16(4), 1951–1960. doi:10.1109/TITS.2014.2387069
- Koné, I., & Boulmane, L. (2018). Hierarchical ResNeXt Models for Breast Cancer Histology Image Classification. *Lecture Notes in Computer Science*, 10882, 796–803. doi:10.1007/978-3-319-93000-8_90
- Krizhevsky, A., Sutskever, I., & Hinton, G. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84–90. doi:10.1145/3065386
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. doi:10.1038/nature14539 PMID:26017442
- Lin, Y., & Wu, J. (2020). A Novel Multichannel Dilated Convolution Neural Network for Human Activity Recognition. *Mathematical Problems in Engineering*, 2020, 1–10. doi:10.1155/2020/5426532
- Liu, F., & Yang, L. (2015). A novel cell detection method using deep convolutional neural network and maximum-weight independent set. *Proc. MICCAI*, 2015, 349–357.
- Ma, N., Zhang, X., Zheng, H. T., & Sun, J. (2018). ShuffleNet V2: Practical Guidelines for Efficient CNN Architecture Design. In V. Ferrari, M. Hebert, C. Sminchisescu, & Y. Weiss (Eds.), *Lecture Notes in Computer Science: Vol. 11218. Computer Vision – ECCV 2018. ECCV 2018*. Springer. doi:10.1007/978-3-030-01264-9_8
- Monaco, J., Hipp, J., Lucas, D., Smith, S., Balis, U., & Madabhushi, A. (2012). Image segmentation with implicit color standardization using spatially constrained expectation maximization: Detection of nuclei. *Proc. MICCAI*, 45, 365–372. doi:10.1007/978-3-642-33415-3_45
- Pacilè, S., Lopez, J., Chone, P., Bertinotti, T., Grouin, J., & Fillard, P. (2020). Improving Breast Cancer Detection Accuracy of Mammography with the Concurrent Use of an Artificial Intelligence Tool. *Radiology. Artificial Intelligence*, 2(6), e190208. doi:10.1148/ryai.2020190208 PMID:33937844
- Pang, Y., Sun, M., Jiang, X., & Li, X. (2018). Convolution in Convolution for Network in Network. *IEEE Transactions on Neural Networks and Learning Systems*, 29(5), 1587–1597. doi:10.1109/TNNLS.2017.2676130 PMID:28328517
- Perone, C., Calabrese, E., & Cohen-Adad, J. (2018). Spinal cord gray matter segmentation using deep dilated convolutions. *Scientific Reports*, 8(1), 5966. Advance online publication. doi:10.1038/s41598-018-24304-3 PMID:29654236
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. (2018). MobileNetV2: Inverted Residuals and Linear Bottlenecks. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 4510–4520. doi:10.1109/CVPR.2018.00474
- Shahidi, F., Mohd Daud, S., Abas, H., Ahmad, N., & Maarop, N. (2020). Breast Cancer Classification Using Deep Learning Approaches and Histopathology Image: A Comparison Study. *IEEE Access : Practical Innovations, Open Solutions*, 8, 187531–187552. doi:10.1109/ACCESS.2020.3029881
- Shallu & Mehra, R. (2018). Breast cancer histology images classification: Training from scratch or transfer learning? *ICT Express*, 4(4), 247–254. .10.1016/j.ict.2018.10.007
- Shen, L., Margolies, L., Rothstein, J., Fluder, E., McBride, R., & Sieh, W. (2019). Deep Learning to Improve Breast Cancer Detection on Screening Mammography. *Scientific Reports*, 9(1), 12495. Advance online publication. doi:10.1038/s41598-019-48995-4 PMID:31467326
- Simonyan, K., & Zisserman, A. (2015). Very Deep Convolutional Networks for Large-Scale Visual Recognition. *International Conference on Learning Representations*.
- Spanhol, F., Oliveira, L., Cavalin, P., Petitjean, C., & Heutte, L. (2017). Deep features for breast cancer histopathological image classification. *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, 1868–1873.
- Spanhol, F., Oliveira, L., Petitjean, C., & Heutte, L. (2016a). A Dataset for Breast Cancer Histopathological Image Classification. *IEEE Transactions on Biomedical Engineering*, 63(7), 1455–1462. doi:10.1109/TBME.2015.2496264 PMID:26540668

Spanhol, F., Oliveira, L., Petitjean, C., & Heutte, L. (2016b). Breast cancer histopathological image classification using convolutional neural networks. *Proc. Int. Joint Conf. Neural Netw.*, 2560–2567. doi:10.1109/IJCNN.2016.7727519

Veta, M., Pluim, J., van Diest, P., & Viergever, M. (2014). Corrections to “Breast cancer histopathology image analysis: A review”. *IEEE Transactions on Biomedical Engineering*, 61(11), 2819–2819. doi:10.1109/TBME.2014.2351051

Wang, B., Zhang, X., Zhou, X., & Li, J. (2020). A Gated Dilated Convolution with Attention Model for Clinical Cloze-Style Reading Comprehension. *International Journal of Environmental Research and Public Health*, 17(4), 1323. doi:10.3390/ijerph17041323 PMID:32092861

Wang, D., Khosla, A., Gargeya, R., Irshad, H., & Beck, A. (2016). *Deep learning for identifying metastatic breast cancer*. Available: <https://arxiv.org/abs/1606.05718>

Wang, S., Lv, Y., Sui, Y., Liu, S., Wang, S., & Zhang, Y. (2017). Alcoholism Detection by Data Augmentation and Convolutional Neural Network with Stochastic Pooling. *Journal of Medical Systems*, 42(1), 2. Advance online publication. doi:10.1007/s10916-017-0845-x PMID:29159706

Wei, B., Han, Z., He, X., & Yin, Y. (2017). Deep learning model-based breast cancer histopathological image classification. *Proc. IEEE Int. Conf. Cloud Comput. Big Data Anal.*, 348–353.

Yıldırım, N. (2018). Mucinous Breast Carcinoma: A Single-Center Experience. *Turkish Journal of Oncology*. doi:10.5505/tjo.2018.1727

Zhang, J., Lin, S., Ding, L., & Bruzzone, L. (2020). Multi-Scale Context Aggregation for Semantic Segmentation of Remote Sensing Images. *Remote Sensing*, 12(4), 701. doi:10.3390/rs12040701

Zhang, X., Zhou, X., Lin, M., & Sun, J. (2018). ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 6848–6856. doi:10.1109/CVPR.2018.00716

Zhou, X., Li, C., Rahaman, M., Yao, Y., Ai, S., Sun, C., Wang, Q., Zhang, Y., Li, M., Li, X., Jiang, T., Xue, D., Qi, S., & Teng, Y. (2020). A Comprehensive Review for Breast Histopathology Image Analysis Using Classical and Deep Neural Networks. *IEEE Access : Practical Innovations, Open Solutions*, 8, 90931–90956. doi:10.1109/ACCESS.2020.2993788

Zhu, C., Song, F., Wang, Y., Dong, H., Guo, Y., & Liu, J. (2019). Breast cancer histopathology image classification through assembling multiple compact CNNs. *BMC Medical Informatics and Decision Making*, 19(1), 198. Advance online publication. doi:10.1186/s12911-019-0913-x PMID:31640686

Ramgopal Kashyap is associated with Amity University Chhattisgarh and his area of interest are image processing, pattern recognition, and machine learning. He has published many research papers in international journals and conferences like Springer, Inderscience, Elsevier, ACM, and IGI-Global indexed by Science Citation Index (SCI) and Scopus (Elsevier) and many book chapters. He has Reviewed Research Papers in the Science Citation Index Expanded, Springer Journals and Editorial Board Member and conferences programme committee member of the IEEE, Springer international conferences and journals held in countries: Czech Republic, Switzerland, UAE, Australia, Hungary, Poland, Taiwan, Denmark, India, USA, UK, Austria, and Turkey. He has written many book chapters published by Elsevier, Springer and IGI Global, USA.