

Empowering Histopathological Breast Cancer Diagnosis through Convolutional Neural Networks

Thyagaraj T.¹, Keshava Prasanna², Hariprasad S. A.³

Submitted: 31/12/2023 Revised: 07/02/2024 Accepted: 15/02/2024

Abstract: Recent advancements in Convolutional Neural Networks (CNNs) have significantly supported the field of breast cancer discovery using medical imaging. This paper explains a custom CNN framework for breast cancer detection employing histopathological images. We examine improved CNN performances in the classification of breast cancer highlighting the effectiveness of convolutional neural networks. The paper also discusses the impact of specific data augmentation like picture in picture on CNN performance and the disadvantage of training standard CNN models by curated datasets. With the highest accuracy documented as 90.50% for standard histopathological data, the performance of the model when validated on data sourced from oncology hospital is presented. The hyperparameters like learning rate of the model that present the optimal performance is compared and presented. A comparative analysis of the CNN model performance for different architectures is presented and the best validation accuracy of the model obtained is 91.31%. A summary of key findings and future research directions, emphasizing the need of custom CNN models in breast cancer detection is provided.

Keywords: convolutional neural networks, breast cancer, histopathological images

1. Introduction

One of the biggest worldwide public health concerns is still breast cancer, particularly affecting women. Its early detection and accurate classification are vital for effective treatment strategies and ultimately improving survival rates. The advent of Convolutional Neural Networks (CNNs) in medical imaging, especially in analyzing histopathological images, has revolutionized the field of oncology diagnostics, creating new pathways for the accurate and effective detection of breast cancer. The prevalence of breast cancer in India stands at 92.6 per lakh population, with a mortality rate of 12.7 per lakh population [1]. Notably, the incidence rate of breast cancer in India is 25.8 per 100,000 women, with a higher prevalence in urban areas (30.5) compared to rural areas (19.8) [2].

The traditional method of diagnosing breast cancer relies heavily on the meticulous examination of histopathological images by pathologists. While this method has been the cornerstone of breast cancer diagnostics, it is fraught with challenges, such as the subjectivity of interpretations and potential for human error. These limitations can lead to variability in diagnoses, affecting the consistency and reliability of breast cancer detection. In this context, the

emergence of machine learning (ML) techniques and deep learning techniques (DL) especially the advancement of CNN as a diagnostic aid presents an opportunity to augment the accuracy and objectivity in interpreting histopathological images.

Recent studies have demonstrated the significant impact of CNNs in improving the precision of breast cancer diagnoses from histopathological images. For instance, the use of transfer learning with convolutional neural networks has shown to enhance diagnosis accuracy substantially, surpassing current best results [3]. This advancement indicates a paradigm shift from traditional methods towards more technologically advanced, data-driven approaches. In addition to improving diagnostic accuracy, CNNs have also proven useful in the classification of images with histopathology into benign and malignant subtypes. Novel CNN architectures, such as small SE-ResNet modules combined with new learning rate schedulers, have proven to be particularly effective in this regard, enhancing both binary and multi-class classification accuracies [4]. The application of deep learning models with transfer learning and feature concatenation has further refined the classification of breast cancer tumor cells, achieving impressive accuracies. A study in 2022 highlighted an average accuracy of 98.76% in such classifications [5]. This level of precision is significant, particularly when considering the complex nature of cancer cells and their varying manifestations in histopathological imagery.

While CNNs have proven remarkably effective in controlled study environments, their application to real-time datasets from hospitals presents unique challenges. Standard CNN models, when trained on well-curated datasets, often

¹Department of Electronics and Communication, BMS Institute of Technology and Management, Visvesvaraya Technological University, Belagavi, India

thyagaraj_tanjavur@bmsit.in

²College of Horticulture Keladi Shivappa Nayaka University of Agricultural and Horticultural Sciences, Shivamogga, India
keshavaprasanna2013@gmail.com

³Faculty of Engineering and Technology JAIN deemed to be university, Bengaluru, India
sa.hariprasad@jainuniversity.ac.in

struggle to maintain the same level of performance when applied to clinical data. This discrepancy arises due to the variability and complexity inherent in real-world medical data, which may not be adequately represented in curated datasets. Real-time hospital datasets often include a wide span of image qualities, variations in staining techniques, and diverse manifestations of cancerous cells, which can differ substantially from the images in curated datasets. As a result, CNN models trained on these curated datasets might not generalize well to real-world data, leading to reduced accuracy and reliability in clinical diagnoses.

To tackle this challenge, the research is focused on a custom CNN framework specifically constructed for breast cancer detection in histopathological images. By incorporating a broader range of data variability and complexity into the training process, these custom models are better equipped to handle the nuances of real-time hospital data.

2. Literature Survey

A. Advances in CNN

Recent studies have demonstrated the effectiveness of deep learning models and depth-wise convolutional neural networks in recognizing breast cancer. For instance, Chorianopoulos et al. (2020) reported a remarkable accuracy of 96.82% on ultrasounds and 88.23% on breast histology using these advanced CNNs, highlighting their potential in various medical imaging modalities [6]. Fernández-Ovies et al. (2019) explored the use of CNNs in conjunction with infrared thermography for early breast cancer recognition, achieving a predictive accuracy of 100% in blind validation. This study underscores the versatility of CNNs in integrating with different imaging technologies[7]. A significant development in CNN architecture was introduced by Medjeded et al. (2019), who developed a new triplet convolutional neural network. This architecture, comprising three subnetworks, achieved an accuracy of 93.13%, sensitivity of 96%, and specificity of 90.25% for breast cancer detection on mammograms, marking a sizable improvement over prevailing models [8]. Agarwal et al. (2020) demonstrated the improved performance of the Faster R-CNN model trained on Hologic images in detecting masses in Full-Field Digital Mammograms. Their findings suggest a potential for these models in routine breast cancer screening [9]. Another noteworthy contribution was made by Tan, Sim, and Ting (2017), who developed the BCDCNN model. This model significantly improved the precision of breast cancer detection in mammogram images, demonstrating the effectiveness of CNNs in distinguishing between benign and malignant masses [10]. The study by Alanazi et al. (2021) presented a convolutional neural network method that achieved an 87% accuracy rate in detecting breast cancer, surpassing the performance of traditional machine learning algorithms by 9%. This advancement is indicative of the superior capabilities of

CNNs in clinical image analysis [11]. The CNNI-BCC algorithm developed by Ting, Tan, and Sim (2019) further exemplifies the evolution of CNNs in medical imaging. This algorithm can classify breast cancer medical images into malignant, benign, and healthy categories with high sensitivity, accuracy, and specificity, marking a significant step forward in automated medical diagnostics[12]. Table 1 contains a compilation of key findings listing the accuracy of classification achieved by models from literature.

TABLE I. CLASSIFICATION ACCURACY FOR VARIOUS DEEP LEARNING TECHNIQUES

Author	Modality and methods	Key findings
Chorianopoulos et al. (2020)	Ultrasound and histopathology images, CNN	Accuracy - 96.82% (ultrasounds) Accuracy – 88.23% (breast histology)
Fernández-Ovies et al. (2019)	Infrared thermography, CNN	Predictive accuracy of 100% in blind validation
Medjeded et al. (2019)	Triplet CNN mammogram analysis	Accuracy of 93.13%, Sensitivity of 96%, Specificity of 90.25%
Agarwal et al. (2020)	Faster R-CNN model for Full-Field Digital Mammograms	Sensitivity – 87.00%, Specificity – 73.00%
Tan, Sim, & Ting (2017)	BCDCNN mammogram image classification	Sensitivity – 82.68%, Specificity – 82.73% Accuracy – 82.71%,
Alanazi et al. (2021)	CNN approach for breast cancer detection	Accuracy 87%.
Ting, Tan, & Sim (2019)	CNN Improvement intended for Breast Cancer Classification	Sensitivity - 89.47%, Specificity - 90.71% Accuracy - 90.50%,

B. Database

Data from two sources are used in this work. One the publicly available database: BreakHis (Breast Cancer Histopathological Database), LC25000 (Lung and Colon Histopathological Image Dataset), CoCaHis (Colon Histopathological Dataset) and another from local hospitals. Developed by Spanhol et al. (2016), the BreakHis database

contains 7,909 images collected from 82 patients, providing a substantial resource for the development and validation of image classification systems. The dataset includes images at various magnifications (40x, 100x, 200x, and 400x), reflecting the real-world scenarios encountered in pathological assessments. The BreakHis database has been integral in fostering research in automated breast cancer diagnosis. By providing a diverse range of histopathological images, it allows for the training of robust deep learning models that can accurately classify breast cancer subtypes. The diversity in magnification levels further challenges algorithms to maintain consistency in diagnosis across different resolutions [13], [14]. Numerous studies have focussed on benchmarking the performance of classification algorithms on BreakHis, comparing conventional machine learning methods with state-of-the-art deep learning approaches. Challenges, such as class disparity and variations in staining techniques exist among all datasets.

The images from public repositories are shown in Fig. 1.

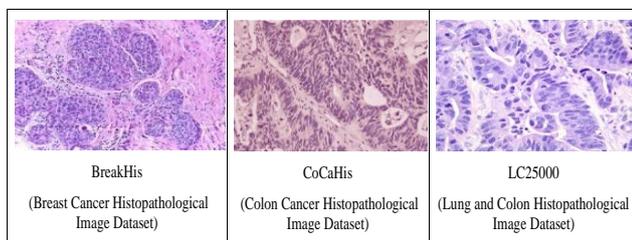


Fig. 1 Sample images from public repositories.

The CoCaHis database contains 82 microscopic images of H&E-stained sections of frozen human specimens of metastatic colon cancer in a liver collected intraoperatively from 19 patients. Thereby, 32.75% of the pixels represent cancer class. The images were stained by H&E and immune-stained to diagnosis-relevant antigens. Images were collected through a clinical study from March 2017 to February 2020 at the Department of Pathology and Cytology in Clinical Hospital Dubrava, Zagreb, Croatia [18].

Lung and Colon cancer Histopathological Image Dataset (LC25000) consists of 25,000 colour images in five classes. All images are 768 x 768 pixels in size and are in jpeg file format. Each class contains 5,000 images of the following categories: colon adenocarcinoma, benign colonic tissue, lung squamous cell carcinoma, lung adenocarcinoma, and benign lung tissue [19].

The second set of data used was acquired from oncology hospitals and were reserved for validating the models. This database consisted of histopathological images of tissue samples. Fig. 2 shows the sample of images. The dataset curated henceforth will be referred to as Oncy-B dataset. The Oncy-B dataset contained a total of 96 whole slide images (WSI) of breast tissue extracted from the tumour. Of the 96 WSI, 21 were labelled as benign and 75 images were labelled as malignant. Hematoxylin-Eosin (H&E) staining

was used on the tissue sample before imaging. While staining reveal cellular components, counter stains were used to provide contrast. The images in the Oncy-B dataset were a 200x magnification of the original WSI. Whole-slide images are on .svs format and have the specifications: Color model: RGB, Size: variable. These images were further split into patches of size 224x224 pixels. This exercise created 216 images which consisted of 196 images in the benign category and 486 images under the malignant category.

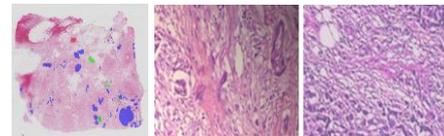


Fig. 2 Sample WSI image and segmented images of Oncy-B dataset

3. Methodology

The methodology utilized is depicted in Fig. 3. The steps followed include preprocessing, development of custom CNN model and modification of various layers and creation of four variations of Custom CNN model i.e., CuCNN-1, CuCNN-2, CuCNN-3 and CuCNN-4. Initially, the high-resolution histopathological images from the dataset are preprocessed to enhance their quality and reduce noise. Subsequently, Data augmentation techniques, such as rotation, flipping, and scaling, are applied to augment the training set and improve model generalization. Finding a suitable architecture being the challenging task, a CNN architecture is designed, comprising convolutional layers for feature extraction, pooling layers for spatial down-sampling, and fully connected layers for classification. The model is then fine-tuned to adapt its representations specifically to the BreakHis breast cancer histopathology dataset. Parameters like learning rate, batch size, and dropout rates significantly impact model performance are ascertained. Fine-tuning convolutional layer parameters, such as kernel size and number of filters, that helps capture relevant features in histopathological images are determined. Regularization techniques are employed for preventing overfitting, enhancing the generalization ability of CNNs. The impact of each model hyperparameter on the convolutional is determined. Evaluation is performed using metrics such as accuracy, precision, recall, and F1 score, and the trained CNN is validated on separate test and validation sets to assess its performance in classifying benign and malignant breast tumor samples.

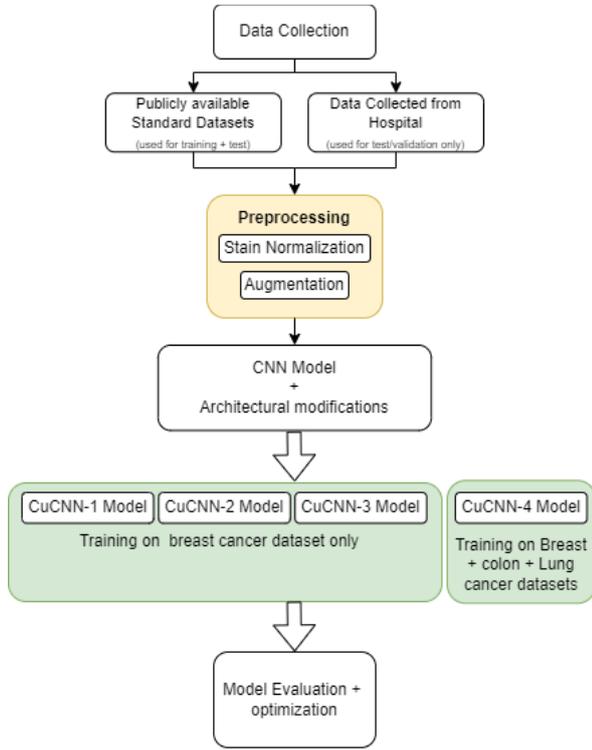


Fig. 3 Overall Methodology

C. Preprocessing

The preprocessing was carried out on datasets sourced from the two different sources. Techniques such as stain normalization, augmentation and cropping were performed for all the images. Fig. 4 shows the preprocessing techniques followed in the work.

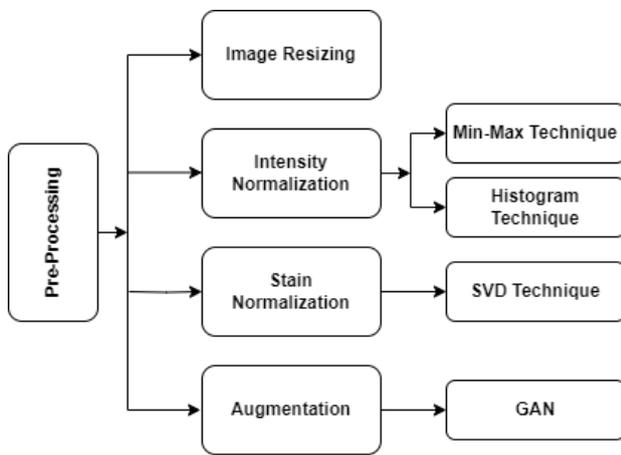


Fig. 4 Preprocessing steps

The images were resized to a consistent resolution. This is imperative to ensure that all images in the dataset have the same dimensions, allowing them to be fed into a neural network. Image size for input is 224x224 pixels. Bilinear interpolation procedure was used for image resizing. The pixel values were estimated based on the weighted average of neighboring pixels. The interpolation technique generated images wherein the details of the images were preserved. Resizing biopsy images to a standard pixel size is required for consistency and comparability. Having a common pixel size, the classification models tend to be more robust and

generalizable, as they learn patterns and features consistently across various datasets. Standardization enhances model interpretability. Fig. 5 shows the images of cropping and resizing.

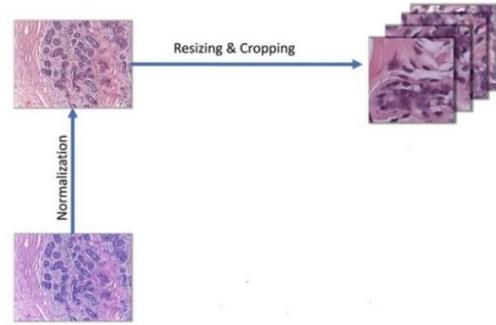


Fig 5 Image resizing and cropping.

To align the image intensity levels with a standard scale, Methods like min-max scaling and z-score normalization are used. This reduces the effect of differences in visual brightness and contrast. In the Min-max scaling method, every value in a column is replaced with a new value using a formula in (1), where x , x_{\max} , and x_{\min} provide the input image's pixel, maximum, and minimum intensity levels, respectively.

$$y = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (1)$$

If the image intensity $Y_{ij}(P)$ at each pixel P is expressed in arbitrary units and measured for field i at j with a specific mode, Normalization is transformation of the type $R_{ij}(P) = \sum_i N_{ij}(P)$. It is useful to conceptualize the histogram of intensities $N_{ij}(P)$ as a mixture of densities where $f_{ijk}(x)$ is the mode specific intensity densities of empty space and benign tissues, like fluid, white matter. The weights $w_{ijk} \geq 0$ sum to 1 and represent the relative weights of components $k = 1, \dots, K$. The densities $f_{ijk}(x)$ and weights w_{ijk} are indirectly logged but are estimated by segmenting the images and estimating w_{ijk} using the proportion of the image data in the k^{th} tissue class and $f_{ijk}(x)$ by the histogram of intensities in that tissue. Fig. 6 depicts the changes in visual appearance of the images before and after the transformation.

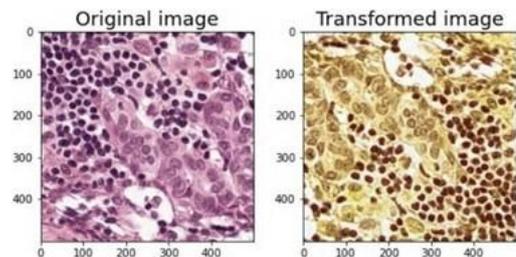


Fig. 6 Image before and after transformation.

One of the main pre-processing steps for whole-slide image

(WSI) analysis is color normalization of stained tissue samples. Fluctuations in antigen concentration, incubation time, and temperature contribute to significant variances in staining. Such fluctuations in color and intensity have an adverse effect on the accuracy and performance of Machine learning models. Stain normalization strategies produce images with consistent characteristics for variation of stains. Stain normalization is the procedure of shifting the color distribution to fit the target image. The stain normalization relied on a single reference image and used pixel-by-pixel color mapping approach. The outcome of the transformation can be seen in Fig. 7.

D. Steps involved in stain normalization.

Step 1: Optical density (OD) from RGB

Step 2: Remove pixels with exceedingly low optical densities

Step 3: Use the largest two values from the SVD application to create the SVD plane for the OD tuples.

Step 4: Project the data onto the plane and normalize it to unit length.

Step 5: Determine each point's angle (ϕ) with respect to the first (or second) SVD direction.

Step 6: Determine the angle ϕ 's robust extremes, or the α^{th} and $(100-\alpha)^{\text{th}}$ percentiles.

Step 7: Determine how to project the extreme values back into OD space in step seven.

Step 8: Utilize the optical density matrix (ODM) in this projection.

Step 9: Use ODM's inverse to determine each stain's concentration (Ch and Ce).

Step 10. Determine the robust maximum ($(100 - \alpha)^{\text{th}}$ percentiles) for each stain concentration separately.

Step 11: Using an H&E template, normalize and convert concentrations to OD space and back to RGB.

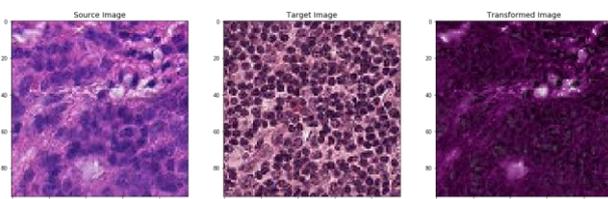


Fig. 7 Image before and after stain normalisation

E. Data Augmentation

The BreakHis dataset has total of 7,909 images of which 2,480 belong to the category of benign and 5,429 belong to malignant class. The number of images available for training the CuCNN-1 model being low, data augmentation techniques were applied to artificially increase the quantity of samples in the training database.

With the objective to swell the number of images available for training and testing, arithmetic manipulation of data was conducted.

Instance 1: Both training and test data were manipulated in such a way that every picture to be added with the same picture in smaller size on the top left corner. Fig. 8 depicts a sample image inserted in its miniature size.

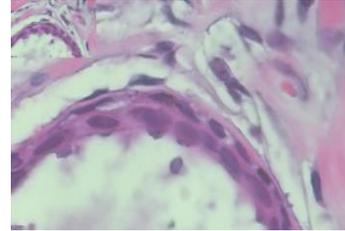


Fig 8 Image manipulation by insertion of image in image

Instance 2: Both training and test data were manipulated in such a way that every picture to be added with the other picture in smaller size on the top left corner but of the same type (i.e., benign image in benign and so forth).

The images available in the BreakHis dataset were used for augmentation. The images were rotated, flipped, and cloned. Each image was manipulated to generate 100 new images with all combinations of the geometric modifications. This increased the volume of the dataset by 70 percent. The number of images in the dataset increased from 7909 to 553630 s [7].

4. Proposed CNN Model Development

F. Configuring CuCNN-1

From a CNN with one input and output layer model, after experimenting with the different number of hidden layers a base model was ascertained, this model henceforth would be referred to as CuCNN-1. The CuCNN -1 model has the structure indicated in Fig. 9.

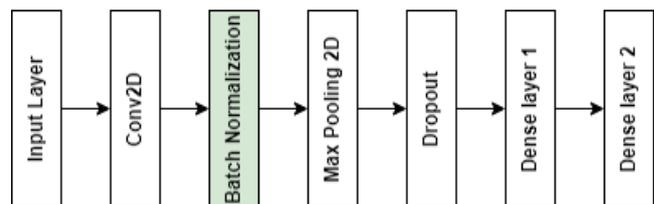


Fig 9 Internal architecture of CuCNN-1 model.

Batch normalization makes neural networks faster and more stable by adding extra layers in the network. The new layer performs the standardizing and normalizing operations on the input, which is coming from a previous layer. Equation (2) presents the neuron computation without batch normalization.

$$p = k(w, x) + b; a = f(p) \quad (2)$$

Where $k()$ is the linear transformation of the neuron, w the weights of the neuron, b the bias of the neurons, and $f()$ the

activation function. The model learns the parameters w and b .

Equation (3) presents the neuron computation after batch normalization.

$$p = k(w, x); P^N = \left(\frac{P - m_P}{S_P} \right) \gamma + \beta; a = f(P^N) \quad (3)$$

Where P^N is the yield of batch normalization, m_P is mean of the neurons output, S_P the standard deviance of the output of the neurons, and γ gamma and β beta, learning parameters of batch normalization. Note that the bias of the neuron ‘b’ is removed. This is because as we subtract the mean m_P , any constant over the values of P such as b can be ignored as it will be subtracted by itself.

The parameters β beta and γ gamma shift the mean and standard deviation, respectively. Thus, the outputs of Batch normalization over a layer, results in a distribution with a mean β beta and a standard deviation of γ gamma. These values are learned over epochs and the other learning parameters, such as the weights of the neurons, aiming to decrease the loss of the model. By lowering internal covariate shift, batch normalization allows for higher learning rates, accelerates convergence, and improves generalization performance. The CuCNN-1 model was trained using the augmented BreakHis dataset and performance was validated using the data obtained from hospital.

The level of accuracy of the model’s classification capability is expected to be high as this model is being proposed for a medical diagnosis application. Inadequate performance is considered as the inability to extract sufficient features from the image. Since the model did not perform to the expectations, Architectural modifications were done to CuCNN-1. The Batch normalization in the CuCNN-1 model was substituted with a convolution layer and a dense layer. This model is henceforth denoted to as CuCNN-2. The internal structure of the framework is shown in Fig. 10

G. Configuring CuCNN-2

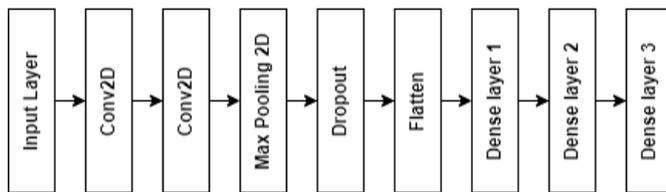


Fig 10 Internal architecture of CuCNN-2 model.

The convolutional layer detects a local association of features from the previous layer and helps in mapping their appearance to a feature map. The size of the feature map is controlled by three parameters, Depth, Stride and padding.

Calculating total error at the output layer with summation over all classes, the total error is calculated using (4)

$$\text{Total Error} = \sum \frac{(\text{target probability} - \text{output probability})^2}{2} \quad (4)$$

Because of the added convolution of neuronic networks, the image is separated into perceptrons, building local receptive fields, and finally compacting the perceptrons into feature maps.

The CuCNN-2 model was compiled with different learning rates (Lr-rt) in a loop and the outcomes were tabulated. The weights of the network were examined. It was discovered that the weights had large values, these large values of weights are an indicator of a network that has overfit the training data. A simple and effective regularization method is to probabilistically dropout nodes in the network.

The dropout layer is a mask that nullifies the contribution of a few neurons towards the next layer. The values of performance indicators of the model had scope for improvement, hence the architectural organization of the model CuCNN-2 was modified to include more dropout layers. Fig. 11 shows the modified structure which contained additional dropout layers in between the dense layers. This model is henceforward referred to as CuCNN-3.

H. Configuring CuCNN-3

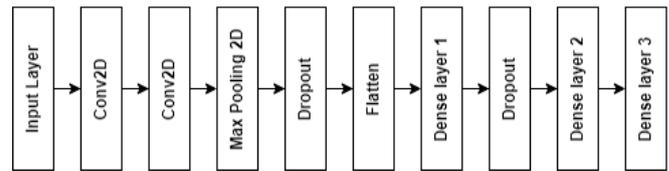


Fig 11 Internal architecture of CuCNN-3 model.

The issue in regularization is due to the phenomenon of co-adaptation wherein the predictive capability of the connections become deterministic in choosing and rejecting weights. The dropout overcomes this issue of co-adaptation. For model estimation, the loss function is to be minimized. For a normal linear layer and the dropout layer the least square loss is given by (5)

$$E_N = \frac{1}{2} \left(t - \sum_{i=1}^n w' \cdot i_i \right)^2; E_D = \frac{1}{2} \left(t - \sum_{i=1}^n \delta w' \cdot i_i \right)^2 \quad (5)$$

Where E_N , E_D represent the least square loss of a normal and dropout layer. The dropout rate δ is equal to 1 with probability p and 0 otherwise. Since the gradient descent approach is used for backpropagation in network training, the relationship between the gradients of the normal layer and the dropout layer is explored. The derivative of normal and dropout loss is presented by (6) and (7).

$$\frac{\partial E_N}{\partial w_i} = -t p_i I_i + w_i p_i^2 I_i^2 + \sum_{j=1, j \neq i}^n w_j p_j I_j I_j \quad (6)$$

$$\frac{\partial E_D}{\partial w_i} = -t \partial_i I_i + w_i \partial_i^2 I_i^2 + \sum_{j=1, j \neq i}^n w_j \partial_i \partial_j I_i I_j \quad (7)$$

If $w' = p * w$, then the expectation of the gradient is given by (8)

$$E \left[\frac{\partial E_D}{\partial w_i} \right] = \frac{\partial E_N}{\partial w_i} + w_i p_i (1 - p_i) I_i^2 \quad (8)$$

The expectation of the gradient in the dropout layer is equal to the gradient of regularized regular network. Hence the selection of dropout rate of $p=0.5$ is expected to provide maximum regularization to the model. The CuCNN-3 performance in terms of training and validation loss as plot is shown in Fig. 12.

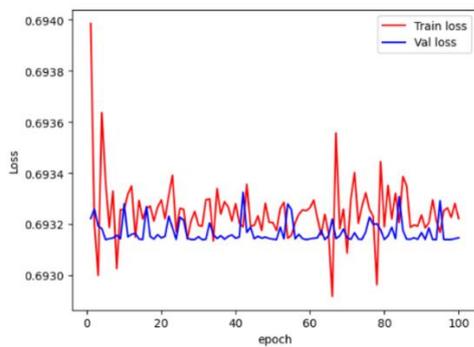


Fig 12 Epoch v/s Loss plot of CuCNN-3

It is found that the performance had degraded with the modification. The possible reasons for the degradation could be due to the dropout rate being high for the input layers. This could have eliminated a set of neurons which aided for a better feature map.

Hence, the structure of the CuCNN-3 model was modified to include batch normalization layers and the dropout rate was reduced to 0.2. Fig. 13 shows the modified structure which contained an additional layer between the convolution layers. This new structure will henceforth be referred to as CuCNN-4.

I. Configuring CuCNN-4

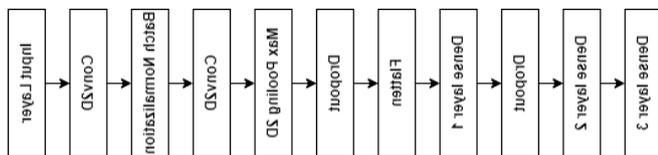


Fig 13 Internal architecture of CuCNN-4 model.

The CuCNN-4 model training was on the BreakHis breast cancer image dataset and the validation was performed on the Oncy-B dataset, the CuCNN-4 model's performance. The best classification performance recorded in literature for accuracy was 96.82%, this being the results of the framework being trained, tested, and validated on data sourced from the same dataset. The model, when validated using the Oncy-B

data sourced from hospital and not utilized for training, produced an accuracy of 86.87%. The accuracy values were low in contrast to the published results in literature. CNN models perform well when large training data are available. It was hypothesized that by training the model on various datasets, additional generalization capabilities would be added. The CuCNN-4 model utilized three datasets belonging to lung, colon, and breast cancer for training. The rationale behind the trial was that although cancer affects different organs of human physiology, the growth pattern of the cells are alike).

5. Results

The compiled graph and results of validation accuracy and loss for distinct learning rates for CuCNN-2 model is displayed in Fig. 14 and Table 2.

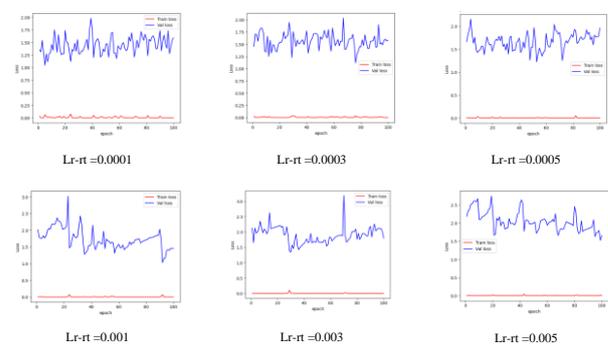


Fig 14 Epoch v/s Loss plot of CuCNN-2 for different learning rates.

TABLE II. PERFORMANCE OF CUCNN-2 FOR DIFFERENT LEARNING RATES

Sl. No.	Learning Rate (Lr-rt)	Validation loss	Validation Accuracy
1.	0.0001	1.5916	0.8547
2.	0.0003	1.5798	0.8531
3.	0.0005	1.9611	0.8415
4.	0.001	1.196	0.8679
5.	0.003	1.339	0.8593
6.	0.005	1.5208	0.8555

It is perceived that for different learning rates, the validation accuracy of model was similar. The model performance with the Lr-rt of 0.001 produced the best results. However, all the results have very high loss values which is not preferred. The continuous fluctuation of Validation loss value suggests that the models are not converging.

The CuCNN-4 model was investigated with Adam, RMSprop and SGD as optimizers, and the effect on the model's training and validation loss is captured in the epoch versus loss plots displayed in Fig. 15. Table 3 shows the

validation loss and accuracy values of optimizers for 100 epochs.

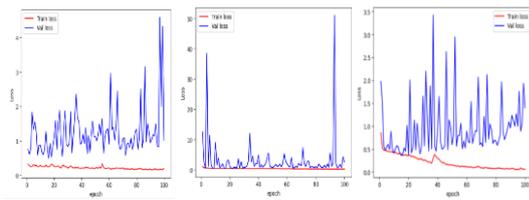


Fig 15 Epoch v/s Loss plots of Adam, RMSprop and SGD optimizers

TABLE III. PERFORMANCE OF CuCNN-4 FOR DIFFERENT OPTIMIZERS

Sl. No.	Optimizer	Epochs	Validation loss	Validation Accuracy
1.	Adam	100	1.0146	85.94%
2.	RMSprop	100	2.6557	83.61%
3.	SGD	100	1.3999	86.87%

It is evident that the SGD optimizer produced the highest validation accuracy for CuCNN-4 model.

The performances of CNN model variants regarding the training and validation accuracy is tabulated in Table 4.

TABLE IV. COMPARISON OF PERFORMANCE OF CuCNN MODELS

(Training on BreakHis and validation on hospital data)	Training accuracy (%)	Validation accuracy (%)
CuCNN - 1	99.88	80.83
CuCNN - 2	98.35	81.67
CuCNN - 3	99.19	50.19
CuCNN - 4	93.24	86.87

The validation accuracy of the CuCNN-4 model is the highest at 86.87%. As further modifications attempted lead to the decrease in the performance accuracy, the CuCNN-4 model was retained.

The training data for the CuCNN-4 model was augmented with the data from CoCaHis and LC25000 dataset containing histopathological images of Colon and Lung cancer. The CuCNN-4 model’s training and validation loss and accuracy plots and values are presented in Fig. 16 and Table 5 respectively.

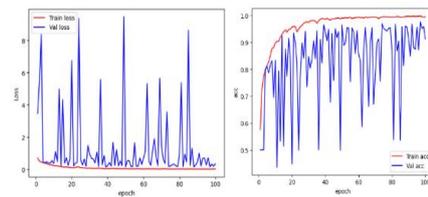


Fig 16 Loss v/s Epoch and Accuracy v/s Epoch plots

TABLE V. PERFORMANCE OF CuCNN-4 TRAINED ON DIFFERENT DATASETS

Cancer Data	Trainin g accurac y	Train Loss	Validation Accuracy	Validation Loss
Colon + Lung + Breast	0.9948	0.0116	0.9131	0.3285

The training of the CuCNN-4 models on three different datasets and validation on the hospital data yielded an improvement in validation accuracy from 86.87% to 91.31%. This significant increase in the model’s performance can be attributed to the null hypothesis that the training of the model on related data improves the model’s performance.

6. Conclusion

In this paper, a CNN model built for breast cancer classification is presented. The model’s performance was evaluated with changes in architecture. The CuCNN-4 model emerged as the best performer. The scarcity of data for training was negated by the application of image augmentation on original images to generate new images. Some of the augmentation techniques like picture in picture, rotation, inclination was validated for its performance. It was found that although rotation and inclination when performed on the images alone did not contribute to any significant improvement, the cloning process of adding a picture in picture added value to the training with trials indicating an improvement in performance. The CuCNN-4 models’ performance with different learning rates was compared and the Lr-rt of 0.001 produced the best validation accuracy of 86.79%. The CuCNN-4 model produced a validation accuracy of 86.87% when the SGD optimizer was used. Further, the training of CuCNN-4 on other cancer datasets like Lung and Colon cancer along with breast cancer to test the hypothesis of training the model on similar domain images led to the improvement in the model’s performance to 91.31%.

References

- [1] P. Dyavarishetty and S. Kowli, “Prevalence of risk factors for breast cancer in women aged 30 years and above in Mumbai,” *Int. J. Community Med. Public*

- Health, vol. 5, p. 647, 2018, doi: 10.18203/2394-6040.IJCMPh20180244.
- [2] D. Purkayastha, C. Venkateswaran, K. Nayar, and U. Unnikrishnan, "Prevalence of Depression in Breast Cancer Patients and its Association with their Quality of Life: A Cross-sectional Observational Study," *Indian J. Palliat. Care*, vol. 23, pp. 268–273, 2017, doi: 10.4103/IJPC.IJPC_6_17.
- [3] W. Zhi, H. W. F. Yeung, Z. Chen, S. M. Zandavi, Z. Lu, and Y. Y. Chung, "Using Transfer Learning with Convolutional Neural Networks to Diagnose Breast Cancer from Histopathological Images," pp. 669–676, 2017, doi: 10.1007/978-3-319-70093-9_71.
- [4] Y. Jiang, L. Chen, H. Zhang, and X. Xiao, "Breast cancer histopathological image classification using convolutional neural networks with small SE-ResNet module," *PLoS ONE*, vol. 14, 2019, doi: 10.1371/journal.pone.0214587.
- [5] A. M. Hassan, M. El-Mashade, and A. M. Aboshosha, "Deep learning for cancer tumor classification using transfer learning and feature concatenation," *Int. J. Electr. Comput. Eng. IJECE*, 2022, doi: 10.11591/ijece.v12i6.pp6736-6743.
- [6] A. Choriantopoulos, I. Daramouskas, I. Perikos, F. Grivokostopoulou, and I. Hatzilygeroudis, "Deep Learning Methods in Medical Imaging for the Recognition of Breast Cancer," 2020 11th Int. Conf. Inf. Intell. Syst. Appl. IISA, pp. 1–8, 2020, doi: 10.1109/IISA50023.2020.9284373.
- [7] F. J. Fernández-Ovies, E. S. Alférez-Baquero, E. J. D. Andrés-Galiana, A. Cernea, Z. Fernández-Muñiz, and J. Martínez, "Detection of Breast Cancer Using Infrared Thermography and Deep Neural Networks," pp. 514–523, 2019, doi: 10.1007/978-3-030-17935-9_46.
- [8] M. Medjeded, S. Mahmoudi, A. Chenine, and M. Chikh, "A New Triplet Convolutional Neural Network for Classification of Lesions on Mammograms," *Rev Intell. Artif.*, vol. 33, pp. 213–217, 2019, doi: 10.18280/ria.330307.
- [9] R. Agarwal, O. Díaz, M. H. Yap, X. Lladó, and R. Martí, "Deep learning for mass detection in Full Field Digital Mammograms," *Comput. Biol. Med.*, vol. 121, p. 103774, 2020, doi: 10.1016/j.combiomed.2020.103774.
- [10] Y. J. Tan, K. Sim, and F. F. Ting, "Breast cancer detection using convolutional neural networks for mammogram imaging system," 2017 Int. Conf. Robot. Autom. Sci. ICORAS, pp. 1–5, 2017, doi: 10.1109/ICORAS.2017.8308076.
- [11] S. Alanazi et al., "Boosting Breast Cancer Detection Using Convolutional Neural Network," *J. Healthc. Eng.*, vol. 2021, 2021, doi: 10.1155/2021/5528622.
- [12] F. F. Ting, Y. J. Tan, and K. Sim, "Convolutional neural network improvement for breast cancer classification," *Expert Syst Appl*, vol. 120, pp. 103–115, 2019, doi: 10.1016/j.eswa.2018.11.008.
- [13] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "A Dataset for Breast Cancer Histopathological Image Classification," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 7, pp. 1455–1462, 2016, doi: 10.1109/TBME.2015.2496264.
- [14] Y. Benhammou, B. Achchab, F. Herrera, and S. Tabik, "BreakHis based breast cancer automatic diagnosis using deep learning: Taxonomy, survey and insights," *Neurocomputing*, vol. 375, pp. 9–24, 2020, doi: 10.1016/j.neucom.2019.09.044.
- [15] N. Brancati et al., "BRACS: A Dataset for BReAst Carcinoma Subtyping in H&E Histology Images," *Database J. Biol. Databases Curation*, vol. 2022, 2021, doi: 10.1093/database/baac093.
- [16] A. Aksaç, D. Demetrick, T. Ozyer, and R. Alhajj, "BreCaHAD: a dataset for breast cancer histopathological annotation and diagnosis," *BMC Res. Notes*, vol. 12, 2019, doi: 10.1186/s13104-019-4121-7.
- [17] M. Koziarski et al., "DiagSet: a dataset for prostate cancer histopathological image classification." 2021.
- [18] D. Sitnik et al., "A dataset and a methodology for intraoperative computer-aided diagnosis of a metastatic colon cancer in a liver," *Biomed. Signal Process. Control*, vol. 66, p. 102402, 2021, doi: https://doi.org/10.1016/j.bspc.2020.102402.
- [19] A. A. Borkowski, M. M. Bui, L. B. Thomas, C. P. Wilson, L. A. DeLand, and S. M. Mastorides, "Lung and Colon Cancer Histopathological Image Dataset (LC25000)." 2019.